

Computer-Assisted Volumetric Tumour Assessment for the Evaluation of Patient Response in Malignant Pleural Mesothelioma

Mitchell Chen
St. Anne's College



Supervisors: Professor Sir J. M. Brady
Dr. J. A. Schnabel

Institute of Biomedical Engineering
Department of Engineering Science
University of Oxford

A thesis submitted for the degree
Doctor of Philosophy
Trinity Term 2011

Computer-Assisted Volumetric Tumour Assessment for the Evaluation of Patient Response in Malignant Pleural Mesothelioma

Mitchell Chen
St. Anne's College

Doctor of Philosophy
Trinity Term 2011

Abstract

Malignant pleural mesothelioma (MPM) is a form of aggressive tumour that is almost always associated with prior exposure to asbestos. Currently responsible for over 47,000 deaths worldwide each year and rising, it poses a serious threat to global public health. Many clinical studies of MPM, including its diagnosis, prognostic planning, and the evaluation of a treatment, necessitate the accurate quantification of tumours based on medical image scans, primarily computed tomography (CT). Currently, clinical best practice requires application of the MPM-adapted Response Evaluation Criteria in Solid Tumours (MPM-RECIST) scheme, which provides a uni-dimensional measure of the tumour's size. However, the low CT contrast between the tumour and surrounding tissues, the extensive elongated growth pattern characteristic of MPM, and, as a consequence, the pronounced partial volume effect, collectively contribute to the significant intra- and inter-observer variations in MPM-RECIST values seen in clinical practice, which in turn greatly affect clinical judgement and outcome. In this thesis, we present a novel computer-assisted approach to evaluate MPM patient response to treatments, based on the volumetric segmentation of tumours (VTA) on CT.

We have developed a 3D segmentation routine based on the Random Walk (RW) segmentation framework by L. Grady, which is notable for its good performance in handling weak tissue boundaries and the ability to segment any arbitrary shapes with appropriately placed initialisation points. Results also show its benefit with regard to computation time, as compared to other candidate methods such as level sets. We have also added a boundary enhancement regulariser to RW, to improve its performance with smooth MPM boundaries. The regulariser is inspired by anisotropic diffusion. To reduce the required level of user supervision, we developed a registration-assisted segmentation option. Finally, we achieved effective and highly manoeuvrable partial volume correction by applying a reverse diffusion-based interpolation.

To assess its clinical utility, we applied our method to a set of 48 CT studies from a group of 15 MPM patients and compared the findings to the MPM-RECIST observations made by a clinical specialist. Correlations confirm the utility of our algorithm for assessing MPM treatment response. Furthermore, our 3D algorithm found applications in monitoring the patient quality of life and palliative care planning. For example, segmented aerated lungs demonstrated very good correlation with the VTA-derived patient responses, suggesting their use in assessing the pulmonary function impairment caused by the disease. Likewise, segmented fluids highlight sites of pleural effusion and may potentially assist in intra-pleural fluid drainage planning.

Throughout this thesis, to meet the demands of probabilistic analyses of data, we have used the Non-Parametric Windows (NPW) probability density estimator. NPW outperforms the histogram in terms of its smoothness and kernel density estimator in its parameter setting, and preserves signal properties such as the order of occurrence and band-limitedness of the sample, which are important for tissue reconstruction from discrete image data. We have also worked on extending this estimator to analysing vector-valued quantities; which are essential for multi-feature studies involving values such as image colour, texture, heterogeneity and entropy.

Acknowledgements

Recalling the day when I first walked into the lab as an anxious fourth year undergrad, I believe that my years as a graduate student have witnessed a steep learning curve, as well as the long awaited determination to commit a lifetime career to medical science. To this, I owe my heartfelt gratitude to my academic supervisor and mentor, Mike Brady. His omniscient intelligence, enthusiasm and passion about science, have always been a source of great inspiration and encouragement to me. In addition to guiding me through my study, Mike's supervision has opened up to me many absorbing facades of scientific research, not just limited to those of medical vision, and greatly broadened my prospects for the future.

I feel equally compelled to express my cordial thanks to Niranjan Joshi, who has led me through some of the most difficult times as a newcomer to medical vision. His dedication, masterly grasp of knowledge and eloquent teaching style were pivotal at all stages during my career as a research student.

I am also truly grateful to Dr. Emma Helm, for kindly providing the essential clinical support throughout my research. None of the key research findings would have been possible without her help. I would also like to thank Dr. Fergus Glesson of Churchill Hospital, Oxford, for generously providing the funding and clinical data for my project, and my academic co-supervisor, Julia Schnabel, for all her support during the course of my study.

There are also many others who I feel deeply indebted to in many ways. In particular, to my 'lab buddies' who have supported, in one way or another, my research works presented in this thesis: Gordon Stevenson, Mattias Heinrich, Habib Baluwala, Tünde Szilagyi, Lydia Tanner, and the many others with whom I have had excellent work discussions and would like to extend my thanks to.

Perhaps it is just as important to mention that the seven long years in Oxford could not go without my many friends. I was fortunate to be accompanied by you all through the many joys and tears all this time. It has always been pleasant to take part in the music rehearsals, photography walking tours, debates, wine-tasting sessions and the many other 'socials' we've had!

Special thanks to Mom and Dad, for always by my side and carried me through all my emotional ups and downs.

Finally, I would like to acknowledge and thank BRC/GE Healthcare, for funding my research works.

Contents

1	Introduction	1
1.1	The Clinical Problem: Origin and Unfolding of the Story	1
1.2	Morphology and Pathophysiology of MPM	7
1.3	Treatment and Medical Challenges	10
1.4	Design of the Clinical Trial	14
1.5	Quantifying MPM	16
1.5.1	RECIST	16
1.5.2	Computer-Assisted Volumetric Segmentation	18
1.6	Thesis Structure	22
2	Medical Imaging Techniques	25
2.1	Introduction	25
2.2	Computed Tomography	26
2.2.1	Principles	27
2.2.2	Advantages	32
2.2.3	Limitations	34
2.3	Image Artefacts of CT	34
2.3.1	Physics-based Artefacts	35
2.3.2	Patient-based Artefacts	36
2.3.3	Scanner-based Artefact	38
2.3.4	Helical and Multi-slice Artefacts	38
2.4	Positron Emission Tomography	39
2.4.1	Principles	39
2.4.2	Advantages	40
2.4.3	Limitations	41
2.4.4	Quantification	42
2.5	Experimental Data Acquisition and Chapter Summary	42
3	Probabilistic Estimation of Tissue Signals for Image Segmentation	44
3.1	Introduction	44
3.2	Non-Parametric Probability Density Estimation	46
3.2.1	Histograms	47
3.2.2	Kernel Density Estimator	49
3.3	Non-Parametric Windows Estimator	51
3.3.1	Theory	52
3.3.2	Implementation	57
3.3.3	Results and Discussion	65
3.4	Application to Medical Image Analysis	66
3.4.1	Applying NPW to MPM Data	66
3.4.2	Segmenting Mesothelioma	74
3.5	Extending NPW to Vector-valued Signals	75

3.5.1	Joint PDF for a Pair of 1D Signals	77
3.5.2	Results and Discussions	78
3.5.3	Joint PDF Estimation for 2D Signals	81
3.6	Summary and Future Works	82
4	Computational Methods for MPM Segmentation	84
4.1	Introduction	84
4.2	Related Work	86
4.3	Overview of Segmentation Methods in Medical Vision	89
4.4	NPW-based Level Sets	93
4.4.1	Theory	93
4.4.2	Results and Discussion	98
4.5	Summary	101
5	Random Walk-Based Method for MPM Segmentation	102
5.1	Introduction	102
5.2	Theory	103
5.3	Results and Discussion	108
5.4	Boundary Enhancement Regulariser	120
5.5	Volumetric Extension	124
5.6	Registration-Assisted Segmentation	136
5.7	Summary and Future Works	147
6	Partial Volume Effect Correction	150
6.1	Introduction	150
6.2	Partial Volume Estimators	152
6.2.1	Overview	152
6.2.2	Gaussian Mixture Model	153
6.2.3	Bayes' Theorem Mixture Model	156
6.2.4	Hidden Markov Random Fields	158
6.2.5	Interpolation by Reverse Diffusion	161
6.3	Results and Discussion	164
6.4	Extension to 3-D	168
6.5	Computation Time Consideration	171
6.6	Summary	174
7	Evaluation of Tumour Responses and Validation	175
7.1	Image Data	176
7.2	RECIST Analysis	177
7.2.1	MPM-RECIST	177
7.2.2	Observer Variabilities	178
7.3	Results from Volumetric Segmentation	187
7.4	Validation and Discussion	191
7.4.1	Statistical Validation with MPM-RECIST	194
7.4.2	VTA Response Rate to IV-Vinflunine	195
7.5	Quality of Life Assessment	197
7.5.1	Pulmonary Function Impairment	199
7.5.2	Pleural Effusion	205
7.6	Summary	206

8	Conclusions	212
8.1	Summary	212
8.1.1	Solution to the Clinical Problem	212
8.1.2	Contribution to Medical Image Analysis	217
8.2	Future Works	219
8.2.1	Clinical Future Works	219
8.2.2	PET-CT	220
8.2.3	Improving the Algorithm	222
	Bibliography	227
A	The International Mesothelioma Interest Group Staging System	233
B	Specifications of RECIST as Used in This Thesis	235
C	Relevant Publications	237

List of Tables

1.1	Asbestos: world production by country	2
1.2	Asbestos: world consumption by country	2
1.3	Comparison of the original and updated versions of RECIST	19
1.4	Comparison of potential computer-aided segmentation routine to MPM-RECIST	21
2.1	Categorising imaging modalities	26
2.2	Hounsfield unit ranges for various tissues in CT	33
3.1	Boundaries for the 2-tuple NPW	80
3.2	Accuracy and computational time of NPW in estimating joint PDF	81
5.1	Typical computation times of different algorithms	120
5.2	Computation time comparison of the four trials in the pilot study	131
5.3	Computation time comparison of the original and registration-assisted methods	143
6.1	Computation time comparison of the PVE-corrected versus the original method	174
7.1	Overall response rate (modified RECIST for MPM) on the available clinical data from IV-Vinflunine phase II trials	177
7.2	MPM-RECIST measures - clinical observer	183
7.3	MPM-RECIST measures - independent observer	183
7.4	Time point observations comparison of MPM-RECIST from two independent observers	184
7.5	Overall response rate comparison of MPM-RECIST from two independent observers	185
7.6	Segmented MPM volumes	193
7.7	MPM segmentation computation times	193
7.8	Provisional response criteria for VTA	195
7.9	Time point observations comparison of MPM-RECIST and VTA on the IV-Vinflunine data	196
7.10	Segmented aerated lung volumes	203
7.11	Aerated lung segmentation computation times	204
8.1	Differential diagnosis of MPM	223
8.2	A quick benchmark study of various sparse iterative solvers	226
A.1	Staging of MPM	234

List of Figures

1.1	Classification of asbestos	4
1.2	Potential sites of pleural mesothelioma growth	5
1.3	Number of pleural mesothelioma-related deaths in the UK 1968-2008	5
1.4	Typical patient journey for mesothelioma	6
1.5	Survival rate at different stages of MPM	8
1.6	Lead time bias for cancer diagnosis	8
1.7	Pathological pathway of MPM	9
1.8	Sample CT image slice and manual delineation of key tissues	10
1.9	Applying modified/MPM-RECIST to a CT scan of MPM	19
1.10	Probability density representations of the overall image and individual tissue intensities	22
2.1	X-ray production	28
2.2	Interaction mechanisms of X-ray with matter	29
2.3	Image from X-ray attenuation measurements	30
2.4	Flow of acquiring a CT scan	30
2.5	Gantry of a CT scanner	31
2.6	CT image reconstruction	32
2.7	Cupping effect of CT scans and correction through filtration	36
2.8	Partial volume shading effect	37
2.9	Illustration of the partial volume shading effect on MPM image data	37
2.10	Annihilation radiation	40
2.11	PET-CT scan of a MPM patient	41
3.1	Integration ranges for the NPW 2D bilinear implementation	55
3.2	Performance of 1D, 2D and 3D NPW for estimating the PDF of synthetic scalar signals, as compared to those of the histograms and KDE methods.	67
3.3	Non-parametric probability estimation for a T2-weighted brain MR section	68
3.4	Estimated overall PDF of a typical MPM scan slice and ‘guided’ PDF estimation for key tissue classes found in its intensity range	71
3.5	2-D PDF response assessment for key tissues in two sample MPM patients	72
3.6	2-D PDF response assessment for for key tissues in two further patients	73
3.7	Preliminary segmentation using PDF estimates	76
3.8	Histogram Estimation for 2-tuple Vector	79
3.9	Illustration of the NPW estimator for a 2-tuple vector	80
3.10	Applying the 2-tuple NPW to a lung CT scan	81
4.1	Probability density representations of the overall image and individual tissues	93
4.2	The workings of NPW-LS in segmenting MPM	97
4.3	More difficult cases of thoracic scans of MPM patients with key tissues manual delineations	99
4.4	Unsuccessful application of NPW-LS segmentation to more typical cases of MPM	100

5.1	Illustration of the random walk-based segmentation for a 5x5 image patch of unit weights	104
5.2	Flowchart outlining key steps in the RW-based method.	107
5.3	Synthetic data with varying image noise	110
5.4	Segmentation results for the low and critical image noise cases	111
5.5	Robustness to noise with varying thickness and contrast	112
5.6	‘Leakage’ study with synthetic image with varying gap width	113
5.7	Random walk-based segmentations for the two difficult cases of MPM from Fig. 4.3	115
5.8	DICE coefficients of the 2-D random walk testing trials	116
5.9	Specificity (blue) and sensitivity (red) plotted against threshold level for random walk testing trials	117
5.10	ROC curves for random walk testing trials	117
5.11	Control of the RW-based method performance by the choice of β	118
5.12	The effect of user-interaction on accuracy	119
5.13	Isotropic vs. anisotropic diffusion	121
5.14	Control of the boundary regularised method by the choice of β	124
5.15	Boundary enhanced-RW segmentation results	125
5.16	Isocontour probability map comparison of boundary enhanced versus original segmentations	126
5.17	PDF estimates of the MPM tumour BE-segmentations	127
5.18	DICE measures from the 2-D trials for the boundary enhanced method as compared the original method	128
5.19	Computation time comparison of the original and boundary enhanced RW method for the ten trials	129
5.20	A comparison of different connectivity cases for a 3-D random walk lattice	132
5.21	Volumetric segmentation of MPM with 3-D random walk	133
5.22	Further volumetric segmentation MPM with 3-D random walk	134
5.23	Accuracy (DICE) analysis of the four trials in the pilot study	135
5.24	Flowchart for the registration-assisted segmentation	138
5.25	Illustration of the registration-assisted algorithm	139
5.26	Evaluation of the computed deformation fields for the registration-assisted segmentation	141
5.27	Sample slices showing steps in the registration-assist algorithm	142
5.28	Registration-assisted segmentation for the scan at T_1	144
5.29	Registration-assisted segmentation for the scan at T_2	145
5.30	Accuracy (DICE) analysis of the registration-assisted algorithm	146
5.31	Comparison of the tumour response results found by four main ways of segmenting MPM	148
6.1	A sample PVE-affected tissue	151
6.2	PVE-affected boundary voxels in a MPM scan	152
6.3	Coronal scan of a thick-slice (5mm) scan showing the dominant axial PVE	153
6.4	Bayes’ theorem mixture model	156
6.5	Steps in the interpolation-based PVE reduction method for the unidimensional case	161
6.6	Depiction of the 8-connected neighbourhood of a subvoxel	163
6.7	MPM data interpolated with reverse diffusion for PVE reduction	165
6.8	Total flow and entropy over time during RD interpolation with s.f of 2	166
6.9	MPM data interpolated with reverse diffusion for PVE reduction	167
6.10	DICE coefficients comparison for the original and PVE-corrected data, as matched to the reference truth	168
6.11	The effect of interpolation scaling factor on segmentation speed and accuracy	169

6.12	3-D reverse diffusion interpolation in the axial direction (s.f = 4)	171
6.13	PVE-corrected volumetric segmentation results	172
6.14	Accuracy (DICE) analysis of the PVE-corrected result	173
7.1	Tumour measurements based on RECIST 1.1 and MPM-adapted RECIST . . .	179
7.2	Illustration of the intra-observer variability assessment	181
7.3	Scatter plots with least square linear fits showing the correlation between measurements made by the same observer at different sites	182
7.4	Scatter plots with least square linear fits showing correlation between measurements made by different observers	185
7.5	PDF of % variation from the clinical RECIST measurements	186
7.6	Volumetric tumour estimation for a patient undergoing 4 trial cycles	189
7.7	Volumetric tumour estimation for a second patient	190
7.8	Patient responses presented as volumetric shape changes	192
7.9	Scatter plots with linear fit showing correlation of VTA and RECIST responses .	194
7.10	Tumour responses based on VTA as compared to MPM-RECIST	196
7.11	The effect of varying response criteria thresholds on the overall patient response results	198
7.12	Diagram showing lung capacity measures in a spirometry output	199
7.13	Manually delineated scans showing changes in aerated lung volumes over the trial cycles	201
7.14	Aerated lung segmentation results	202
7.15	PDFs of the segmented aerated lung, combined lung region and the unsegmented portion	208
7.16	Correlation of aerated lung response with VTA and MPM-RECIST	209
7.17	Volumetric change in aerated lung over the treatment cycles as a measure of pulmonary function impairment	209
7.18	Combined segmented tissues	210
7.19	Combined tissue segmentation with its location in the rib cage	211
8.1	Segmented PET scan for MPM tumour	221
8.2	Iteration steps of the minimum residual method	225

Chapter 1

Introduction

1.1 The Clinical Problem: Origin and Unfolding of the Story

During his renowned journey to the Orient, whilst travelling in the Great Empire of Tartary, Marco Polo was presented with some miraculous garments impervious to fire, described as being made of ‘fire salamander’s wool’ by his local host. Although Polo himself was not entirely convinced of the tale of ‘a small lizard that’s immune to fire’, the false belief was not dispelled until some time later. The material, *linum asbesti*, or simply asbestos, was in fact known to the Greeks as early as 300BC and is remarkable for its extreme resistance to fire and chemical breakdown, tensile strength as well as the fibrous structure that allows great structural flexibility when woven as a textile. For centuries, it has been simply used for entertainment and socio-religious purposes (e.g wealthy nobility amusing their guests with ‘fire-proof’ cloth) across the world. It was not until the industrial revolution did the world saw widespread use of the material in numerous sectors, especially in the construction and shipping industries. However, with the increasing awareness of numerous health hazards associated with the use of asbestos, most countries have now imposed strict regulations on the use and mining of this material. By turn of the 21st century, the asbestos mining centres of the world have mostly relocated from the western hemisphere to developing countries; with Russia, China and Kazakhstan collectively

Country	1900	1940	1970	2000	2005
<i>Former Soviet Union</i>	NA	102 000	1 065 943	—	—
Russia	—	—	—	750 000	925 000
Khazakstan	—	—	—	271 300	355 000
China	—	20 015	172 365	370 000	520 000
United States	956	18 198	113 683	5 260	—
Canada	26 436	313 514	1 507 420	320 000	200 000
World Total	31 587	573 728	3 493 800	2 070 000	2 400 000

Table 1.1: Asbestos: world production by country (metric tonnes) [81]

Country	1920	1950	1975	1990	2000
Former Soviet Union	1 629	136 458	1 286 679	2 151 800	507 125
China	—	102	15 000	185 748	410 190
India	1 847	11 160	61 826	118 964	124 516
United States	153 000	660 000	803 000	41 000	14 600
United Kingdom	21 199	107 606	137 487	15 731	244
Japan	4 965	12 245	255 551	592 701	98 595

Table 1.2: Asbestos: world consumption by country (metric tonnes) [81]

producing over 75% of the global output in 2005 [80]. It should also be noted that production levels in these countries over the past decade had not only been largely unaffected by the world's trend in restricting asbestos usage but in some cases had even increased (Table 1.1). A similar observation can be made about the consumption of asbestos, with China and India emerging as major consumers of the material, most probably due to their rapidly growing heavy industry sectors (Table 1.2). Given the size of population in these countries, the effect of this long-term hazard may be particularly detrimental.

Perhaps the most notable and notorious health complication associated with asbestos is malignant pleural mesothelioma (MPM), an aggressive form of tumour found in the pleura¹, or the outer lining, of the lungs. The disease is known to have affected the lives of tens of thousands, perhaps more, over the past few decades [36]. Aetiological studies [41] have identified prior exposure to asbestos as the primary cause of the malignancy. In mesothelioma,

¹membrane of the thoracic, or chest cavity

malignant cells develop mostly in the pleura of the lungs and internal chest wall. They may also develop, although this is much less likely, in the peritoneum², the pericardium³ or tunica vaginalis/serosa⁴. These tissue membranes are collectively known as the mesothelia; hence the name given to the malignancy. For our research purposes, we are mainly concerned with pleural mesothelioma and will conduct our studies on medical images taken in the chest and upper abdominal region. The pleural form of mesothelioma generally starts on the parietal lung membrane (outer pleura) and then propagates to the neighbouring visceral mesothelium (inner pleura) before extending to soft tissues of the chest wall, other mesothelial locations or distant sites of metastasis such as the lymph nodes, liver, brain or bone. Fig. 1.2 shows the potential sites of pleural mesothelioma growth in the lungs. The nature of a tumour is considered malignant if metastases or anaplasia (dedifferentiation of cells) have occurred. The malignant form of pleural mesothelioma is studied in this thesis. Clinical features of MPM include: dyspnoea (shortage of breath), weight loss and chest wall pain (due to chest wall invasion) and the accumulation of fluid (effusion) in the pleural space. The pathophysiology of MPM will be described in greater detail in Section 1.2.

Epidemiology

The epidemiological burden of MPM has attracted a great deal of attention in recent years, as legal issues arise on the occupational cases of the disease. Responsible for over 47,000 deaths worldwide each year [88], MPM poses a serious threat to global public health. It is also the greatest single cause of work-related death in many countries [51]. In the UK alone, over 1,800 people were diagnosed with MPM in 2008. The incidence of MPM ranges from 7 to 40 per million in developed countries [62]. Given the gradual cessation of asbestos production in the 1980s and the disease's long latency period, typically between 30 to 40 years, the incidence of mesothelioma in the UK is expected to continue to increase and ultimately peak in 2020 [81]. The trend of mesothelioma-related deaths in the UK over the years is shown in

²The lining of the abdominal cavity

³A sac surrounding the heart

⁴membrane of male/female inner reproductive organ

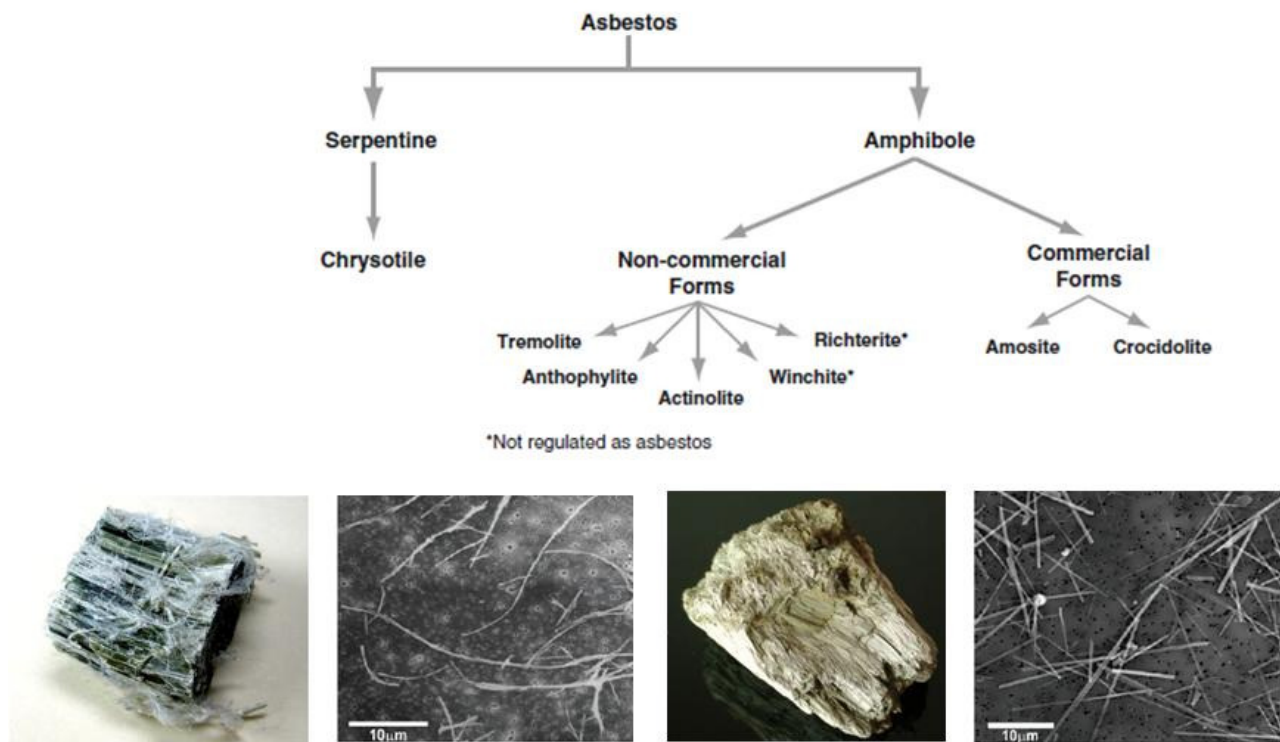


Figure 1.1: (*top*) Classification of asbestos, (*bottom left to right*) Chrysotile and amphibole asbestos and microscopic scans of their fibrous structures, respectively [72]. Previously it was claimed that only the amphibole forms of asbestos cause mesothelioma, due to their longer persistence in the lungs when inhaled. Recent studies [36], however, have suggested a link of chrysotile asbestos exposure to mesothelioma. Also given that asbestos rarely exist in pure chrysotile form - contamination with amphibole asbestos is common. It is advisable that all forms of asbestos be handled with care.

Fig. 1.3. Currently, the incidence is much lower in developing countries. In China for example, the current incidence is only 4 per million [85]. However, given China's large population and the fact that asbestos production in the country is still rising, MPM is likely to emerge as a more serious health concern in the years to come. Moreover, the average age of MPM onset is 60 years in developed countries, whereas in China it is only 45.2 years [85]. This may be attributed an earlier, more widespread, and frequent exposure to asbestos amongst the population in China and a future outbreak of the disease. Experts have also predicted a similar upcoming MPM endemic in India [75].

Occupational exposure to asbestos is not the sole cause of MPM. Bystander and environmental exposures to asbestos, even at low doses, are also shown to be aetiologically relevant [36].

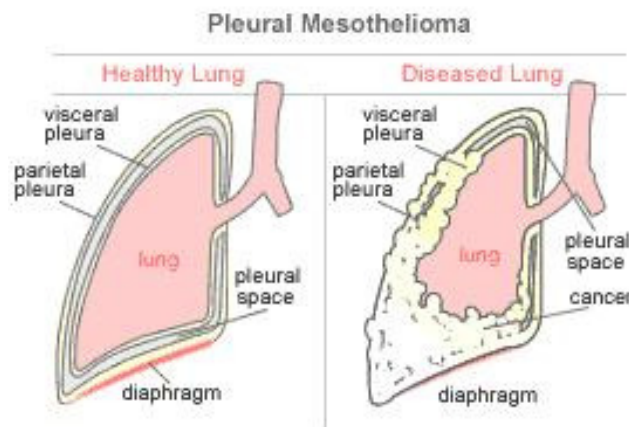


Figure 1.2: Potential sites of pleural mesothelioma growth [69]

In this regard, MPM presents a potential health hazard to a large percentage of the population born prior to the introduction of legal restrictions on asbestos usage. Currently, stringent control regulations [27] are in place in the UK for maintaining existing building structures with asbestos. However, regulations are considerably less strenuously enforced in the developing countries, where asbestos continues to be used with little or no protection, therefore exposing their populations to a greater risk to mesothelioma.

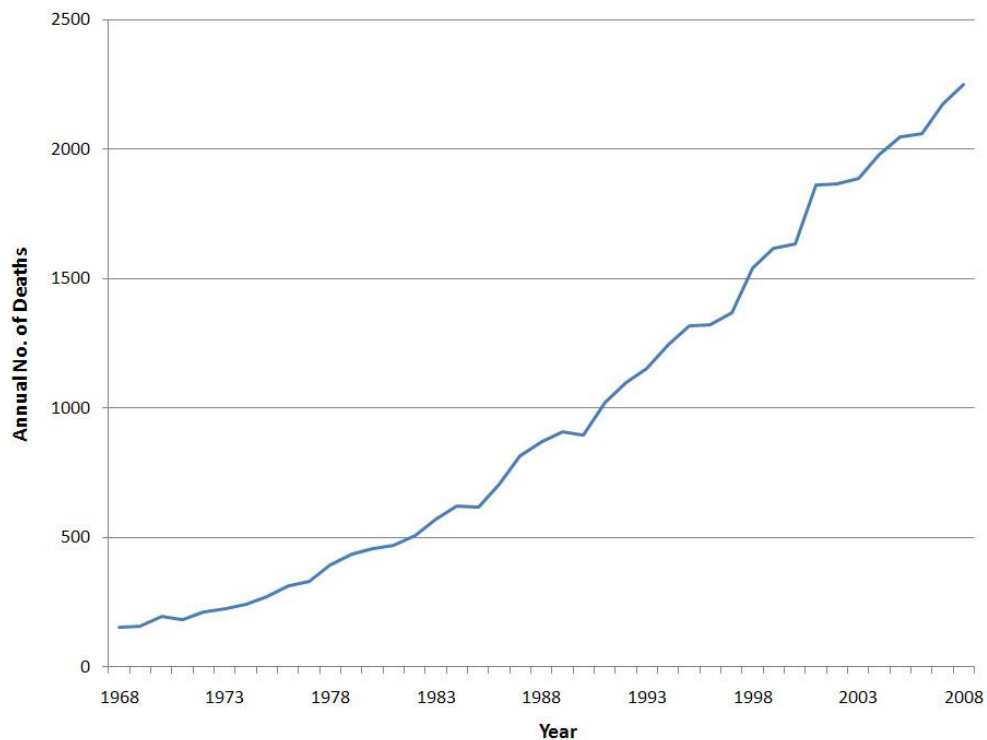


Figure 1.3: Annual number of pleural mesothelioma-related deaths in the UK 1968-2008 [27]

Diagnosis

Diagnosis of mesothelioma remains a key challenge in the field of medicine. A combination of invasive (spirometry, pleural fluid aspiration, open or closed biopsy) and imaging methods are currently employed for this purpose. A detailed account of the pathology of MPM is given in Section 1.2.

Prognosis

Prognosis of MPM is currently very poor; with an average survival time of only nine months from time of diagnosis [36]. Several treatment methods are available, including chemotherapy, radiotherapy, and surgery. A typical patient journey for mesothelioma diagnosis and treatment is illustrated in Fig. 1.4.

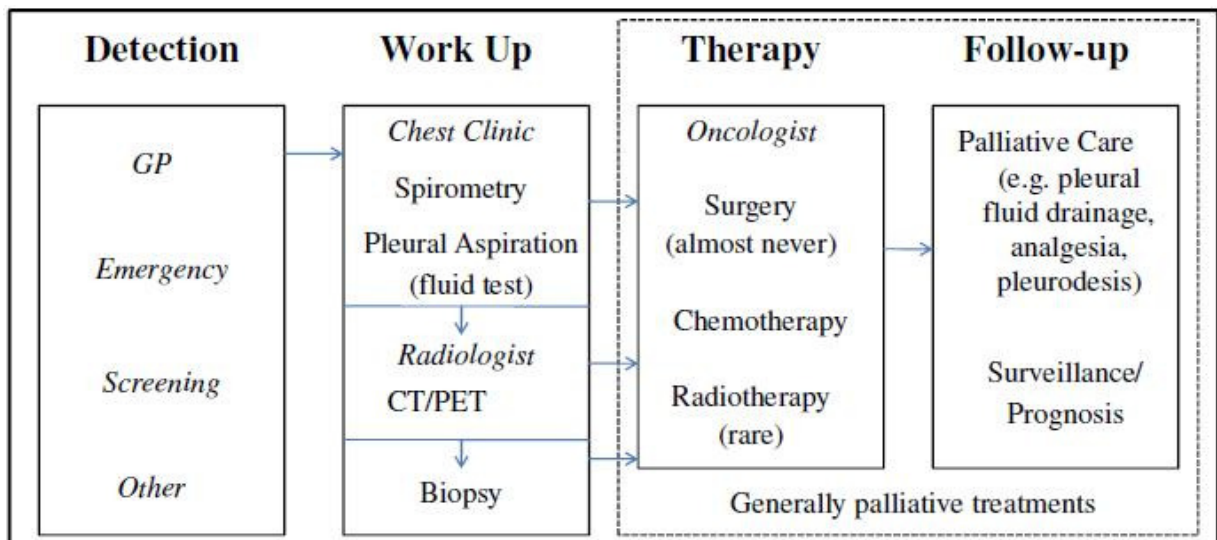


Figure 1.4: Typical patient journey for mesothelioma

There is currently no cure for the disease - most treatments only work as far as improving the quality of life of a patient, ie are palliative in nature. We describe these treatment approaches and prospects for a more curative answer to MPM in Sections 1.3 and 1.4.

The imaging assessment of disease progression in MPM focuses on changes in the volume of pleural fluid and tumour and an assessment of the underlying lung. An increase in tumour volume results in decreased chest wall compliance and reduced inflation of the underlying lung.

Pleural fluid also causes atelectasis (collapse) of the underlying lung. Currently, clinicians assess tumour progression with a set of guidelines known as Response Evaluation Criteria in Solid Tumours (RECIST). However, we will discuss the major shortcomings of this approach in Section 1.5, and the emerging need for a computer-aided volumetric segmentation routine. In this thesis, we develop a computer-aided 3-D segmentation method that measures changes in MPM, which offers a number of potential benefits and sheds light on the treatment and prognostic care of MPM. The steps of the development of the method are presented throughout this thesis, as outlined in Section 1.6.

1.2 Morphology and Pathophysiology of MPM

In MPM, the tumour typically originates from the parietal pleura of the lungs and subsequently extends to other thoracic membranes. The tumour also has a tendency to grow along the interlobular space, ultimately enclosing the lung. Advanced forms occur when the disease progresses beyond the mesothelial sites, such as to the lymph nodes or bones - leading to metastatic spread of the cancer. A standard staging system (Appendix A) by the International Mesothelioma Interest Group (IMIG) [71] classifies the morphological progression of MPM based on key tumour stages including: tumour spread to the visceral pleura, diaphragmatic muscle, soft tissues of the chest wall and other mesothelial sites such as pericardium and myocardium. Distant metastatic sites (liver, adrenal gland, brain, bone and kidney) and lymph nodes are also key indicators in determining stage of the disease. This staging system stratifies patients according to prognosis; with the survival probability at a specific time duration notably declining with tumour stage progression, as shown in Fig. 1.5.

Therefore, it is apparent that diagnosis of the disease in its early stage, like that of the many other forms of malignancy, should improve the prognosis and survival rate of patients. However, lead time bias (Fig. 1.6) should be taken into account in the survival rate analysis; it is the difference between the cancer diagnosis time through screening and that through symptoms [79].

In the early days of MPM study, patients with obvious MPM symptoms such as pleural

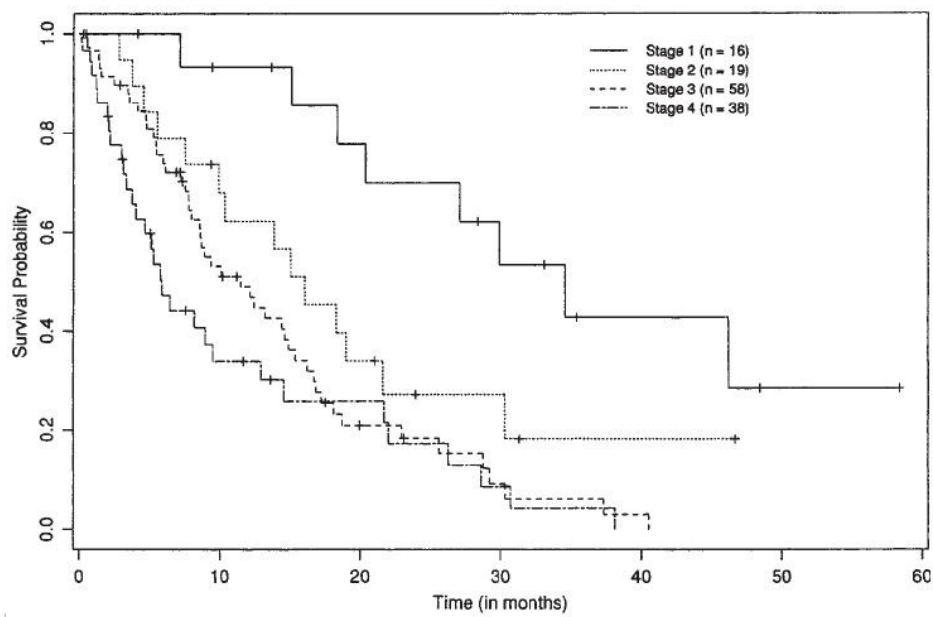


Figure 1.5: Survival rate at different stages of MPM [65]

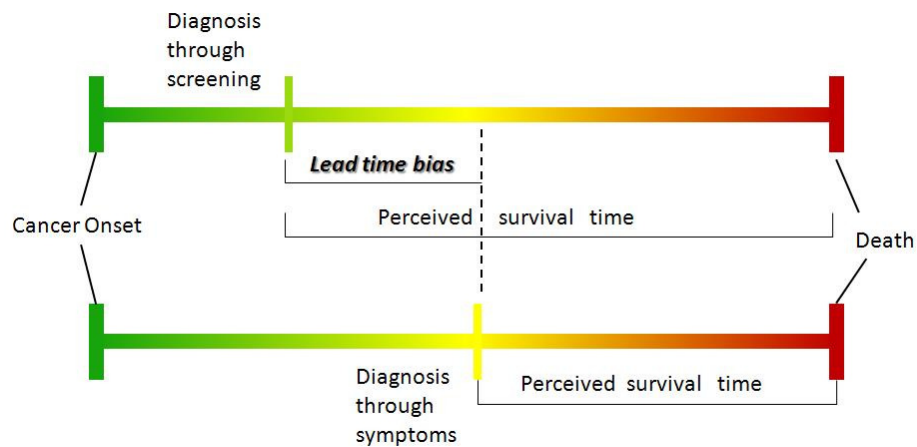


Figure 1.6: Lead time bias for cancer diagnosis

effusion and chest wall pain are usually suspected of having the disease and passed on for diagnosis [36]. This usually begins with a review of the patient's medical record and past exposure to asbestos. Spirometry can sometimes be employed to analyse symptoms of dyspnoea and measure the volumetric capacity of the thorax. Currently, the most reliable way of diagnosing MPM is biopsy; tissue samples can be obtained by either closed or open biopsy. Closed or percutaneous biopsy is usually performed with image guidance, typically using ultrasound or CT. Open biopsy on the other hand, can be performed via video assisted thoracoscopy (VATS) using a rigid endoscope or a flexible scope (medical thoracoscopy).

The lack of clinical evidence to support early diagnosis presents a major challenge to treating MPM. Existing diagnostic methods are only used on patients with symptoms. These symptoms often do not appear until a late stage, at which point metastasis may well have occurred. This leads to the aforementioned poor prognosis and low survival rate for MPM patients. A detailed pathological pathway is given and illustrated in Fig. 1.7. Today, effective treatment of the disease manifests its early diagnosis, possibly through medical imaging. However, medical images are currently less capable of showing the presence of early stage pleural tumours. This is due to a number of complications, not least the similarity in densities (attenuations), hence pixel intensities, of tissues in the abdominal region, as well as the complexity of the anatomical geometry. To illustrate these points, sample CT images are shown Fig. 1.8 along with a mask containing tissues delineated manually by an experienced clinician showing the tumour boundaries.

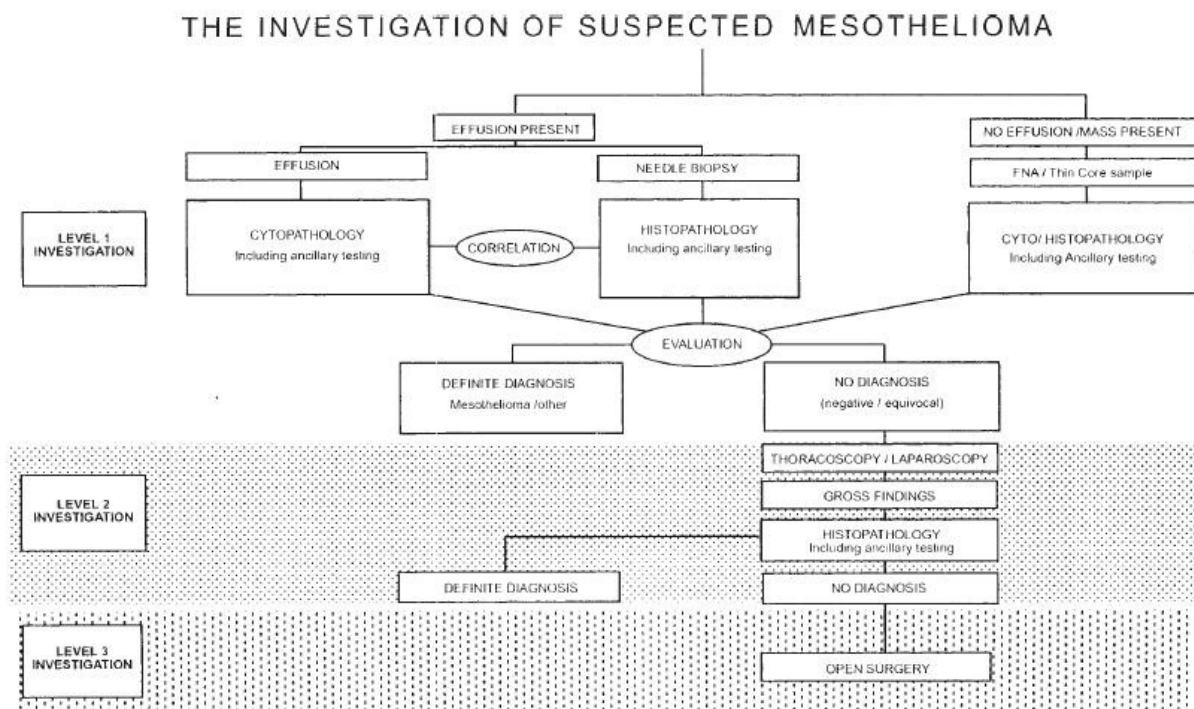


Figure 1.7: Pathological pathway of MPM [69]

In addition to its application in diagnosing MPM, medical imaging also sees use in treatment planning and prognosis. This is explained further in the next section.

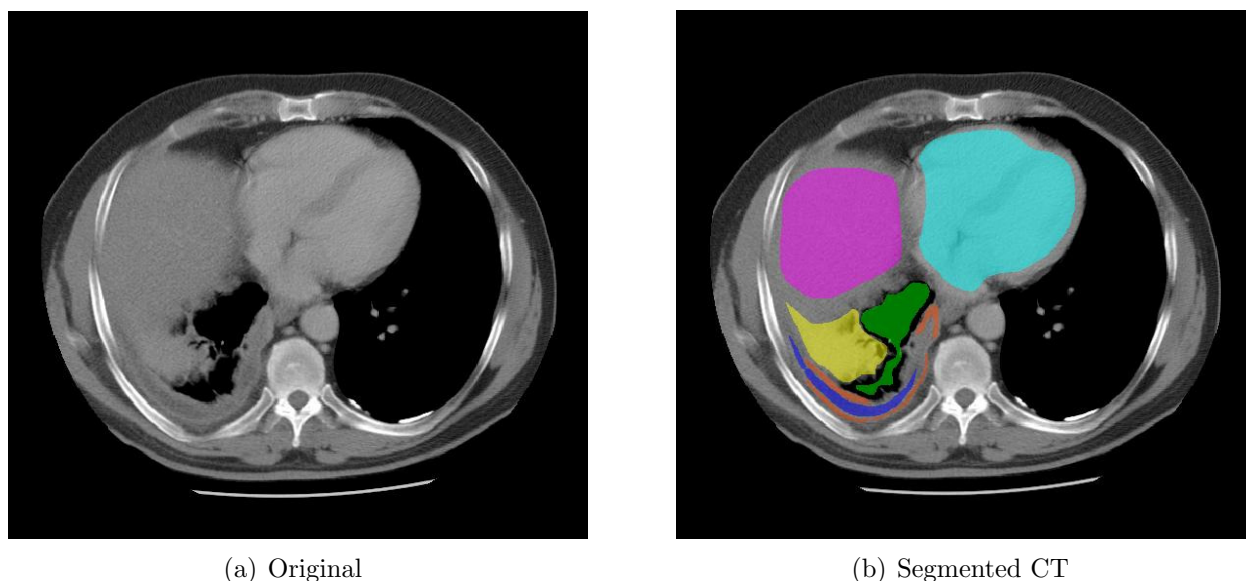


Figure 1.8: Sample CT image slice delineated manually with key tissues shown in different colours: tumour (*orange*), fluid (*blue*), aerated lung (*green*), collapsed lung (*yellow*), liver (*pink*) and mediastinum (*cyan*).

1.3 Treatment and Medical Challenges

As noted above, the prognosis in MPM is poor, with a median survival time of less than a year and a lower than 1% five-year survival rate [14]. Existing treatments are mostly palliative, rather than curative. However, in recent years, with the many clinical and scientific advancements which have greatly enhanced our understanding of the disease, significant improvement in MPM treatment or even a possible cure is no longer a distant vision. In this section, we review four main approaches to treating MPM; namely surgery, and radiation, gene and chemo- therapies.

Surgery

Due to the late diagnosis and ultra-thin structure of tumours along the lung membrane in many cases, surgery for MPM today mostly provide forms of palliation instead of cures. Control of the pleural effusion alleviates the shortness of breath symptom by clearing the lung spaces of serous fluid. For example, a chest tube might be placed for complete drainage of the effusion fluid. Pleurodesis, or the obliteration of the pleural space, can then be achieved by applying chemical sclerosing agents after full expansion of the lungs, which reduces the possibility of

recurrent effusion. However, this is only plausible in the earlier stages of MPM since the complete encirclement of the pleura by the tumour can trap the effusion and make it difficult for both fluid drainage and pleurodesis. To demonstrate this point, a late stage CT scan is presented in Fig. 1.8.

Although rarely performed, there are also surgical procedures aimed at curing early diagnosed cases of MPM. These generally involve the removal of the affected pleura, also known as pleurectomy or, much more aggressively, extrapleural pneumonectomy, which involves removal of parts of the lungs. Typically, after these operations, there would be residual microscopic tumour which would be addressed by adjuvant treatment such as chemo- or radiotherapy. More recently, photodynamic therapy (PDT) has been introduced as a novel adjuvant treatment for this purpose, where light sensitisers responding to light of a particular wavelength are administered. These chemicals, one example being porphyrin, a naturally occurring respiratory pigment present in plant and animal cells, are expected to produce toxic oxygen-free radicals which trigger cell death when activated by light. In PDT, the photosensitisers are designed to target only the tumour cells. However, despite various advancements in this approach, out of the reported surgical cases during the period 1976-1997, less than one third of patients receiving pleurectomy survived beyond two years; a similar number is observed for extrapleural pneumonectomy [69]. Today, most MPM cases are still handled by the other forms of cancer treatment, mostly notably chemotherapy.

Radiation-based Therapy

To date, radiotherapy has had only limited use in MPM treatment. A major obstacle is the anatomical complexity and extent of the disease, which typically involves the entire pleural surface of the lung. This impedes the application of a highly focused dose of radiation onto the tumour whilst sparing adjacent normal tissues. Various techniques have been proposed to handle this difficulty, including photon-electron beam treatments, customised blocks and CT scans for treatment planning. However, none of these innovations have yet shown benefit in patient survival. In fact, despite some showing significant palliation, most results from

radiotherapeutic studies show either no improvements or even worsening on the patient survival aspect of the disease [14], probably due to the absorption of harmful radiation by otherwise healthy tissues in the body.

Immunotherapy

Immunotherapy is a novel field of research and has shown promising results for treating conventionally difficult malignancies. The principle of the treatment lies in stimulating the anti-tumour cell immune response in the patient host. There are several proposed immunotherapeutic strategies for treating MPM, one of which is the introduction of a ‘suicide’ gene into the tumour cell. This transduces the genetic coding of the neoplastic cell with the one for the herpes simplex virus enzyme, thymidine kinase (HSVtk), with the aim of inhibiting further DNA replication of the cell through a series of chemical binding processes with the drug ganciclovir (GCV). In addition, the process can release toxic GCV metabolite to even non-transduced tumour cells via gap junctions or induce an anti-tumour immune response capable of killing the non-transduced tumour cells at a distance. Phase I clinical trials [47] have been conducted showing positive results for non-toxicity and general safety of the treatment. It was also concluded by the authors that the treatment might lead to some degree of tumour reduction. However, higher doses of the drug and gene inference are required to justify the clinical efficacy of the treatment in further phases of clinical trials.

A second gene therapy method [17] is based on augmenting the immune response to tumours in the body. Intrapleural administration of cytokines⁵ is known to partially overcome mesothelioma’s resistance to immune destruction but is also toxic, which limits its clinical use. However, intratumoural delivery of cytokine genes in a murine model has been found to be safe and has been successful in animal trials. Phase I human clinical trials using either a recombinant vaccine virus or a gene encoding for the bacterial heat shock protein (HSP-65) are underway to investigate the practicality of this treatment.

A third immunotherapeutic approach is combination gene therapy which, as its name sug-

⁵a protein molecule known for its transport signalling function in the nervous system

gests, applies both the toxic prodrug and genetic inhibition gene operations. Experimental results in some cases have confirmed the effectiveness of the treatment in killing tumour cells although further in-depth clinical studies are still needed to support this observation [38]. Despite the considerable light which immunotherapeutic approaches have shone on treating MPM, they remain largely experimental at the present, with most studies still in their infancy and far from clinical verification.

Chemotherapy

Chemotherapy is by far the most commonly adopted and mature treatment for MPM in clinical practice. It involves the application of a single or combination of several anti-neoplastic drugs to kill cells that divide rapidly, a characteristic property of cancer cells. Recent phase III trials of combination therapy with pemetrexed or raltitrexed and cisplatin have shown significant benefits in terms of progression free survival and alleviation of symptoms [82]. This evidence was the basis for the approval of pemetrexed and cisplatin as the standard of care for mesothelioma in England and Wales by NICE in January 2008⁶. In addition, there is evidence from a retrospective analysis of patients in this trial that second line therapy may also have survival benefit [2]. These encouraging studies have spurred interest in the development of new treatments and combinations of existing treatments. A recent study examined the efficacy of IV Vinflunine [74], a fluorinated Vinca alkaloid first discovered and presented in [20] and has reported some notable clinical results. This study forms the backbone of data acquisition for this thesis. This clinical project is detailed in Section 1.4.

The main difficulties of clinical trials in mesothelioma lie in its resistance to drugs, the complexity of tumour shape, as well as the short survival period following diagnosis. In addition, the scattered patient population, and the unwillingness of mostly advanced stage cancer patients to take part in experimental studies, has further hindered the execution and limited the scale of clinical trials on any proposed treatments.

⁶<http://www.nice.org.uk/TA135>

Treatment Evaluation

Assessment of a disease's response to therapy is essential for the evaluation of any clinical studies. Only then can clinicians analyse the efficacy of the proposed treatment and carry out the clinical trials of a new therapy effectively. In the case of mesothelioma, this necessitates the quantification of tumour size and establishment of a standard tumour response criteria based on such measurements. We introduce two such candidates in Section 1.5, namely a modified RECIST and computer-aided volumetric segmentation method, both used for CT scans. Other ways of MPM tumour assessment are possible, such as using the maximum standard uptake value (SUVmax), glycolytic volume or the metabolic heterogeneity of tracer uptake (FDGhetero), on PET images.

1.4 Design of the Clinical Trial

All the experimental data used in our research and reported in this thesis were collected in a recent Phase II study of the chemotherapeutic drug IV Vinflunine, that was carried out as a joint venture by pharmaceutical corporation Pierre-Fabre and nine research centres in France, Germany and the UK. The experiment, its design and findings, are presented in [74].

Patient Recruitment

A number of MPM patients (= 67) were recruited and their responses to the drug recorded. The primary objective of the study was to assess the efficacy of Vinflunine in terms of response rate in patients without prior chemotherapy. Other objectives included the assessment of progression-free survival and safety of the treatment. One of the patient selection criteria was that the patient had at least one lesion that satisfied one of the following measurability criteria:

- 1) for diffuse pleural thickening: diameter ≥ 5 mm assessed in 3 different areas; sum of the three measurements ≥ 20 mm.
- 2) for nodular lesion and any other lesion: diameter ≥ 20 mm with conventional techniques or ≥ 10 mm with spiral CT scan.

Written consent was obtained from each and every patient prior to participation in the trials.

Methodology

All patients were prevented from receiving concomitant treatment of chemotherapy during the course of the study; although other forms of treatments such as radiotherapy or most palliative measures were allowed. During the experiment, Vinflunine was administered intra-venously in doses of $320\text{mg}/\text{m}^2$ during 10-minute intervals, one per cycle. Computed Tomography (CT) was one of the ways by which the effect of the drug was studied. A baseline scan was first taken and then followed by one, two or three subsequent scans after 2, 4 or 6 cycles of Vinflunine treatment, respectively. This corresponds to a timeframe of 6 weeks per assessment. Further chemotherapeutic doses would stop if progressive disease was observed in the patients at any time or if they willingly withdrew from the clinical trial. Toxicity of the drug was recorded throughout the study period according to Common Toxicity Criteria (CTC) 2.0, which classifies the side effect of a drug on a scale from 1 to 5, with 1 being the mildest and 5 the most severe, leading to certain death. In this study, most patient reactions were graded as either 1 or 2 - Vinflunine was found to have good safety profile with respect to MPM patients. However, as with many other chemotherapeutic agents, a number of adverse effects such as anaemia, fatigue and nausea, were still observed.

To evaluate disease progression over time and hence the clinical efficacy of the drug, the change in volume is monitored as the primary outcome measure. The modified RECIST proposed by Byrne and Nowak [12] was used to quantify this, which is a set of guidelines that is currently widely regarded as the clinical standard for assessing imaged lesions in clinical practice. The disease response is determined from the uni-dimensional measurements of the lesions, known as the longest diameters (LD). We describe this system in greater detail in Section 1.5.1. The following experimental specifications are also in line with this protocol:

- 1) Tumour response: evaluation of the tumour based on the CT scans.
- 2) Response criteria when RECIST is used:

Complete response: disappearance of all target lesions;

Partial Response: at least 30% decrease in the sum of LD of target lesions (reference to the smallest baseline sum);

Progression of disease: at least 20% increase in the sum of LD of targets lesions (reference to the baseline);

Stable disease: $\leq 30\%$ decrease and $\leq 20\%$ increase in the sum of LD of target lesions (reference to the baseline);

Clinical Findings

Of the recruited patients, 70% had advanced disease (stages III to IV), as classified by the IMIG system (Appendix A). For late stage MPM patients, sites of metastasis included the lungs, liver, lymph nodes, soft tissues, and the skin. The median time from disease diagnosis to trial entry was 0.4 years, with a range between 0.04 to 7.20 years. Under the RECIST scheme, response rate to Vinflunine was found to be 14.5% in assessment patients with 95% confidence interval (CI), 6.9% to 25.8%. Regarding the palliative aspect of the therapy, the Karnofsky Performance status (a scale index system allowing patients to be classified according to their functional impairment) was improved in 13.8% of the patients, with 60% of the participants maintaining their baseline status. The median progression-free period was found to be 3.2 months. The overall survival time was 10.8 months with a 1-year survival rate of 36.9%.

In the following sections, we describe possible ways of applying our understanding of advanced medical image processing techniques to aid the assessment of this clinical study.

1.5 Quantifying MPM

1.5.1 RECIST

For many years, RECIST has been established for a range of tumours as the hallmark for assessing tumour sizes and response to treatments in clinical studies. In this section we introduce the main idea and working principles of this protocol for medical image evaluation.

In 1979, the World Health Organization (WHO) first recommended a bi-dimensional quantification system; but the WHO criteria [45] were designed for spherical shapes and hence poorly suited to evaluating long, thin, or other irregular shapes found in oncological studies. For this reason, in 2000, a set of standardised rules, RECIST guidelines, were proposed, suggesting the use of a uni-dimensional quantity; namely the sum of the longest diameters (LD) of all measurable lesions [77]. The system also included specification of a measurability threshold and the maximum number of lesions to follow in a prognostic study. There are four possible disease response states, namely, complete response (CR), partial response (PR), progressive disease (PD) and stable disease (SD). Updates to the RECIST measures were introduced in [19] for improving the clinical accuracy and applicability of these guidelines, hence giving rise to RECIST 1.1. In particular, the lesion threshold size, or measurability burden, is changed to 10mm for CT image scans. The maximum number of lesions that should be measured is limited to five (two per organ) instead of the previously specified ten. It was found that with these changes, the trade-off between the rate of misclassification of tumour response and measurements is optimised [19]. Moreover, a measurement of a lymph node's short axis, in the perpendicular axis to the longest measurement, is also taken into account. The lesion would be a target lesion if the short axis ≥ 15 mm, non-target if between 10 to 15mm and neglected otherwise. Additionally, to account for the situation where a malignancy with a small LD sum is labeled PD with a size increase as small as 2-3mm (which accounts for a $\geq 20\%$ growth), a second requirement is introduced such that the tumour must grow by at least 5mm over the previous lowest sum in order to be labeled as such. A further improvement is the requirement for response confirmation with a repeated assessment within 4 weeks of the initial measurement, which applies to response primary endpoints (e.g phase II single agent trial). Other imaging modalities such as FDG-PET can be used to complement CT in cases of uncertainty such as the appearance of new lesions; this ensures that the new lesion findings are unequivocal, that is, not merely attributable to the difference in the scanning technique used.

However, when applying RECIST 1.1 to MPM, selection of measurement sites poses a major challenge. The measurability criterion of 10mm in short axis thickness would rule out

most cases of pleural thickening. Additionally, criteria may be applied differently by different investigators, thereby creating unnecessary ambiguities in the measurement. This is largely because of the irregular boundary shape and uneven growth pattern of the tumour which do not lend itself to accurate assessment using a single unidimensional measurement. For this reason, a modified RECIST [12] has been suggested for use on MPM. In this method, tumour thickness perpendicular to the chest wall or mediastinum is measured in two positions at three separate levels on transverse cuts of the image scan. The sum of these six measurements is then defined as a single pleural uni-dimensional sum and used to represent the tumour at a specific point in time. One drawback of this method is that it only measures tumour size in 6 positions and is also known to be prone to inter- and intra-observer variabilities. In one literature [76], major disagreement was found to occur in 40% and minor disagreement in 10.5% of the cases. Reasons for these disagreements may have included errors in tumour measurement, errors in selection of measurable targets, and radiologic technical problems such as the partial volume effect which caused ambiguous tissue boundaries. Such discrepancies could potentially impair the assessment of treatment response in clinical trials. A study of this variability is given in [2]. We investigate the inter- and intra-observer variabilities of RECIST as part of this thesis, in the context of tumour response findings from MPM image scans.

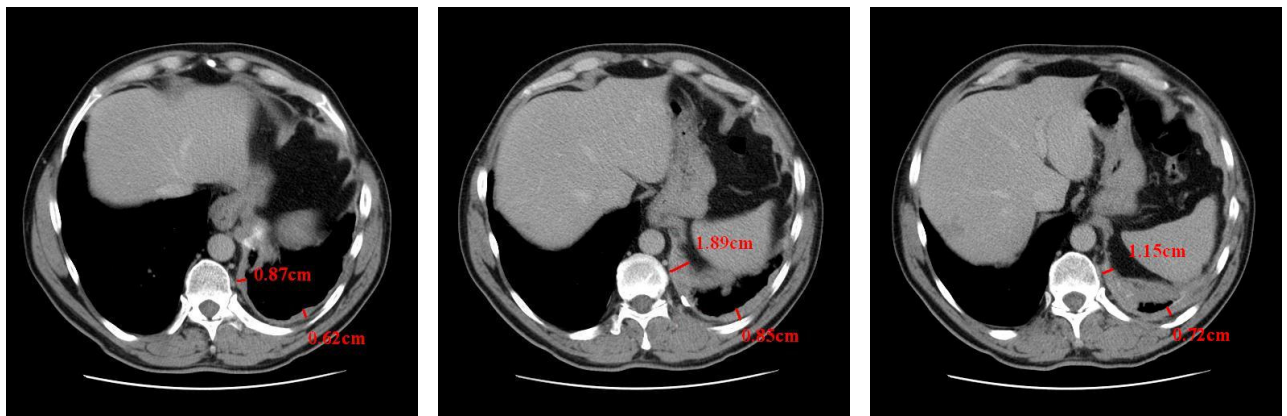
A chart detailing comparison of the three RECIST systems is given in Table 1.3. Fig. 1.9 illustrates the application of the modified criteria to a single baseline MPM CT scan. Note the longitudinal shape of the lesions, which renders the uni-dimensional measurements less representative of the actual disease progression. The full specifications of the RECIST adopted in our study are given in Appendix B.

1.5.2 Computer-Assisted Volumetric Segmentation

Despite recent efforts to improve and adapt RECIST to the measurement of MPM, the simple geometric uni-dimensional nature of the criteria makes them highly liable to a range of systematic errors and inaccuracies. Changes in tumour shape over time may significantly alter the measurement, hence making the assessment less effective for prognostic evaluation. Also,

Specification	RECIST 1.0	RECIST 1.1	Modified RECIST
Measurability Criteria	10mm spiral CT, 20mm otherwise	$\geq 10mm$ for CT and MRI, $\geq 15mm$ for lymph nodes and $\geq 20mm$ for chest X-ray	pleural thickening: $\geq 5mm$, nodular lesion $\geq 10mm$ for CT scans
Measurement Methodology	uni-dimensional. 1-10 targets, 5 per organ	uni-dimensional. 1-5 targets and 2 per organ (all longitudinal), short axis for lymph nodes	two thickness measures on three separate slices any two of which are at least 10mm apart
Lymph Node	Measure long axis as for other lesions. Silent on normal size	Measure short axis. Define normal size.	same as in 1.1
Disease Progression Definition	20% increase in sum	20% increase and at least 5mm absolute increase	or the appearance of one or more new lesions
New Lesions	Not required	Confirmation with FDG-PET required	same as in 1.1

Table 1.3: Comparison of the original and updated versions of RECIST



(a) MPM-RECIST (slice 1) (b) MPM-RECIST (slice 2 - 10mm) (c) MPM-RECIST (slice 3 - 20mm)

Figure 1.9: Applying modified/MPM-RECIST to a CT scan of MPM

the unequivocal progression in non-measurable disease still cannot be assessed by the latest RECIST standard. For these reasons, there is an urgent need for a better method of tumour assessment. This is the main problem that we tackle in this thesis.

Over the years, advanced tools in medical image analysis have been established as either alternatives, or improvements, to traditional manual methods adopted by clinicians and researchers. This is an inter-disciplinary field where the combined knowledge of computer vision, physics and medicine jointly contribute to the better handling of complex images for the diagnosis and management of a range of diseases. Medical images, in particular, are notorious for their poor signal-to-noise ratios (SNR). Moreover, applications in medicine require a high degree of certainty despite the limited availability of patient data, usually due to geographical and/or ethical difficulties in patient recruitment for clinical study purposes. To this end, anatomical, physiological information and an understanding of the physics of medical imaging are all applied to the field of computer vision in order to extend the applicability of imaging techniques to solving medical problems. One of the key areas in medical image analysis, image segmentation, is concerned with delineating structures of interest in an image. A number of segmentation algorithms have been proposed and subsequently applied to a range of clinical applications. The importance of such methods in addressing our clinical problem are highly significant. To summarise, the following aspects are most relevant to the measurement of MPM:

- 1) the segmentation methods are generally automatic or semi-automatic, thus allowing them to be more effective and to reduce the amount of user interaction required, compared with the slow process of applying manual measurements and deciding on the sites of measurements.
- 2) The computer-based nature of the algorithms render them both faster and more reproducible, hence greatly reducing or even eliminating the influence of intra- and inter-observer variations.
- 3) Computer algorithms are ideal for working with large data samples, thus making volumetric segmentations possible which in turn notably improves accuracy of the assessment.

Table 1.4 summarises the main advantages of a computer-aided segmentation routine over MPM-RECIST.

Aspect	MPM-RECIST	Computer-aided Volume Segmentation
Accuracy	Subject to a large number of complications including but not limited to drastic tumour shape changes over time, measurement errors and observer bias and variation.	Highly accurate method giving volumetric estimates based on all available image data; allows a direct assessment of the tumour volume changes
Computation Speed	Slow due to having to decide and choose the sites of measurements and make the actual manual measurements	Fast, especially with an efficient computational algorithm optimised for a particular clinical application such as MPM
Robustness	Poor performance with both intra- and inter-observer variations expected	Better reproducibility with consistent results
User Interaction	Requires extensive user interaction and manual measurement work	Depending on the computational nature of the algorithm, usually very little, mostly during the initialisation phase of the algorithm
Lesions	Only the primary lesion is considered	All lesions are considered and added to the overall tumour volume

Table 1.4: Comparison of potential computer-aided segmentation routine to MPM-RECIST

However despite the obvious benefits presented above, designing such a routine for segmenting MPM is not a simple task. Comparing to other forms of tumour, MPM presents a number of notable segmentation challenges, one of which is the similarity in the pixel intensity of MPM lesions and neighbouring tissues (mostly soft) in the thorax, as displayed on CT scans. This can be shown by the probability density plots of intensities of the overall image scan and key thoracic tissues in Fig. 1.10; note the overlap of these intensity distributions. The image noise, geometric complexity of the imaged thoracic region, and the long thin shape of pleural mesothelioma present additional obstacles to our segmentation task.

When evaluating a novel medical treatment, the improved quality of life is assessed alongside the patients survival benefits. In this regard, prognoses of MPM treatments may encompass the monitoring of forced vital capacity (FVC), that is, the maximum amount of air that can be expelled from the lung following a maximum inspiration. Anatomically this is directly linked to volume of the aerated lung, or the remnant functional portion of the tumour-affected

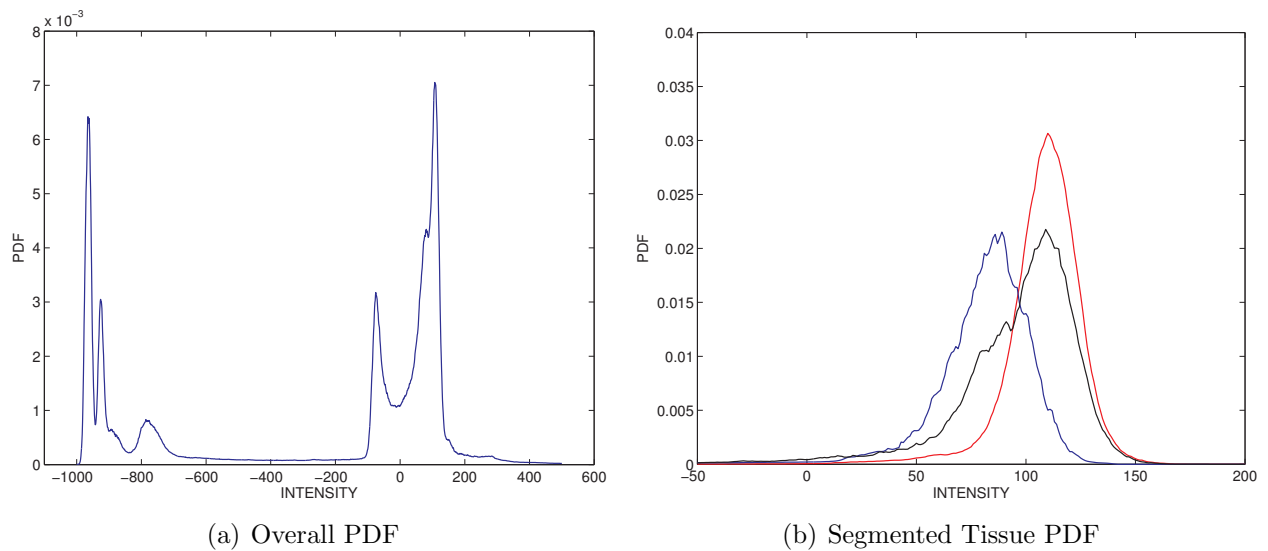


Figure 1.10: Probability density representations of the overall image and individual tissue intensities according to the manual delineation data. Shown here are tumour (*blue*), fluid (*red*) and mediastinum (*black*).

lung. By convention, FVC is measured clinically by spirometry. The availability of a tissue volume classifier would enable us to compute the lung volumes from the medical scans by segmenting the aerated lungs on the images. Therefore in addition to studying the tumour itself, we have also examined the aerated lung response in our experiment. It can be seen from Fig. 1.8 that determining the aerated lung boundaries closely depends on the accurate delineation of a tumour, a much more difficult imaging problem to tackle. Consequentially, during the developmental phase of our computer-aided segmentation algorithm, the emphasis remains on the segmentation of the tumour itself.

1.6 Thesis Structure

In this chapter, we have outlined the clinical problem of malignant pleural mesothelioma, including a brief description of its origin, epidemiology, means of diagnosis, and treatments. The established treatments of the disease in clinical practice include surgery, radiotherapy and most commonly, chemotherapy. Most current treatments, though, remain largely palliative. This is largely due to the late diagnosis of the disease in most cases. Over the years, however, clinicians

have gained considerable knowledge of the effects of numerous chemotherapeutic drugs, some of which have demonstrated good effects in treating or restraining the growth of MPM. To better tackle the issue of treating the disease, an effective way to assess tumour responses to experimental therapies is essential. Currently, the standard way for assessing tumour sizes from medical images is RECIST, a set of guidelines designed to quantify the tumour from a number of uni-dimensional measurements. However, this results in a great amount of uncertainty due to observer biases in a practical setting. To this end, we have designed and tested a computer-aided segmentation routine for MPM based on advanced computational tools in medical image processing, which may potentially be greatly superior to RECIST. This may in turn significantly improve the accuracy and reliability of our clinical studies of MPM therapies.

CT has been established as an effective imaging technique for the evaluation of MPM [2]. Many therapeutic trials have studied responses to treatment in mesothelioma by measuring changes in disease burden on CT. However, CT imaging has several limitations in terms of assessing tumour volume and growth. Therefore, the first priority of our research is to benchmark the available technologies for imaging the human body and justify our choice of using Computed Tomography (CT) for studying MPM. This discussion, along with possible ways of improving the imaging quality by incorporating information from other modalities, are given in Chapter 2. It should be noted that for MPM, the geometric complexity of the imaged region and tumour itself, similarity in CT intensity of neighbouring tissues and image noise usually rule out the use of any simple thresholding routine as an adequate mean of segmentation. Thus we are obliged to evaluate a number of advanced image segmentation techniques. However, before we proceed onto the topic of image segmentation, one key objective would be to ascertain that we are equipped with the necessary tools to devise such a segmentation scheme. Tissues in the thorax (e.g consolidated lung, pleural effusion) often have similar attenuations to the primary tumour; rendering segmentation of tumour from surrounding structures problematic. When clinicians look at scans of patients with mesothelioma, additional features such as texture, heterogeneity and knowledge of normal anatomy are frequently taken into account. For this reason, our initial experimental work was to investigate the possibility of incorporating these additional measures

into our segmentation method with the use of probability density functions in the vector space. In this regard, we have worked towards furthering our understanding of the non-parametric windows-based probability density estimator. More specifically, we worked on extending the theoretical basis of non-parametric windows to vector-valued quantities. This is further discussed in Chapter 3. In Chapter 4, we review a number of related works on MPM segmentation and present and compare a series of existing image segmentation methods. Chapter 5 presents the segmentation method of our choice, namely that based on random walk, and a series of improvements and extensions we have made towards optimising it for the case of MPM. In addition, for segmenting longitudinal structures like MPM, it is vital that we have a solution to correct for the pronounced partial volume effect (PVE) in CT. Works on this are presented in Chapter 6. Evaluation of the results, clinical deductions as well as a solution to our clinical problem are presented in Chapter 7. As our method sees functionality beyond MPM, also included in this chapter are results from other applications of clinical and scientific interest, demonstrating the versatility of our segmentation routine. Finally, we conclude this thesis in Chapter 8 by presenting a summary of our contributions to the clinical topic and general field of medical vision, and a discussion of potential future works.

Chapter 2

Medical Imaging Techniques

Medical imaging provides the first step to understanding tumour disease progression and response to therapy. To begin our work on addressing the clinical problem of Malignant Pleural Mesothelioma (MPM), we first review the underlying principles of a number of suitable medical imaging techniques and relate their utilities to the study of MPM and tumours in general.

2.1 Introduction

In clinical studies, medical imaging provides a fast and non-invasive way for obtaining patient information to support the diagnosis and prognosis of disease. Different physical phenomena, such as X-rays, radioactivity, ultrasound, and magnetic resonance have been harnessed for imaging uses. Imaging modalities can be categorised based on the type of energy source used: internal, external and hybrid. Details are given in Table 2.1. Tomography is a process by which the imaging is carried out in sections through the use of an electromagnetic wave. Most imaging modalities used today are forms of tomography. The specific choice of imaging modality is based on the required type of information (anatomical/metabolic), tissue contrast, spatial resolution, cost of scanning, as well as other relevant factors such as patient safety and scanning speed.

Thoracic scans generally involve the use of CT, which is less vulnerable to chest motion blurring and is readily available due to its relatively low cost. However, CT at the same time suffers from poor soft tissue contrast and is associated with a series of imaging artefacts. The

Type	Example	Physical Phenomenon
Internal	SPECT, PET	Radioactivity from intravenously administered imaging tracers
External	CT, Ultrasound	Monitoring the body's response to waves, which may or may not be ionizing
Hybrid	MRI, fMRI	External magnetic field and RF waves with excited nuclei acting as the internal source of energy

Table 2.1: Categorising imaging modalities

physical principle, advantages and limitations of CT are discussed in Section 2.2. The imaging artefacts of CT are covered in Section 2.3. A functional/metabolic imaging technique, PET, is reviewed in Section 2.4. We conclude this chapter in Section 2.5 with a discussion to justify the use of CT for collecting the experimental data in the IV-Vinflunine phase II trials, and the possibility of applying the latest multi-modal imaging techniques, such as PET-CT, to enhance our clinical study.

2.2 Computed Tomography

Since its discovery in 1895 by German physicist Wilhelm Röntgen, x-ray has been widely applied for imaging the human body for a range of medical uses. The method is traditionally referred to as x-ray radiography and is especially useful for viewing the skeletal systems and detecting calcifications such as gallstones and kidney stones [9]. The imaging aspect of x-rays derives from its different attenuations through body tissues of varying density, which information is collected and processed for a depiction of their presence and locations.

Computed tomography (CT) can be considered as a 3-D version of conventional x-ray radiography. It uses X-rays, an ionising radiation, and measures the corresponding attenuation coefficient to determine the densities of tissues in a body region. Better scans with higher resolution and contrast are obtained from regions consisting of tissues of higher atomic numbers

compared to those in the surrounding area. For this reason, CT has seen widespread applications in imaging the bones (detecting breakages, density, and calcifications), blood vessels, and the bowel region, and, more importantly for our purposes, the thorax (where the body tissues are surrounded by ribs and air in the lung). CT scans can be improved further with the use of contrast-enhancing agents, which are mostly iodine/barium-based, which have high attenuation coefficients.

2.2.1 Principles

X-rays are a type of electromagnetic radiation that interact with matter primarily through absorption and scattering, resulting in an attenuated beam, a property that gives rise to their use as an imaging source of data. X-rays are produced when a bombarding electron enters an atomic shell and strikes an electron, with sufficient energy so that it causes the electron to leave its shell. In response to this change, and to maintain atomic stability, an electron from an outer shell would attempt to fill the missing gap by moving to the inner shell, emitting an x-ray whose energy is the binding energy difference between the two shells involved, as illustrated in Fig. 2.1. In the design of an x-ray tube, tungsten is often used for colliding with the incoming electrons to produce x-ray emissions because of its high atomic shell energies. A vacuum is maintained in the tube to prevent unwanted ionisation of the air, which would cause a reduction in x-ray production efficiency.

The imaging functionality of x-rays is governed by the photoelectric effect, which occurs in much the same way as x-ray production. The x-ray enters an atomic shell causing the displacement of an electron from its shell, or its ionisation, as well as a ‘freed’ photoelectron of energy $E = E_0 - E_{BE}$ where E_0 and E_{BE} are the energies of the incoming x-ray and electron binding shell, respectively. The ejected photoelectron causes the ionisation of further atoms in the medium. It is also possible for scattering to occur where ionisation does not take place. In this case (Rayleigh scattering), the trajectory of the striking radiation is deflected as a result of the atomic collision. In Compton scattering, however, the x-ray both induces an ionisation of the atom and experiences a change in its trajectory. Finally, pair and triplet productions are

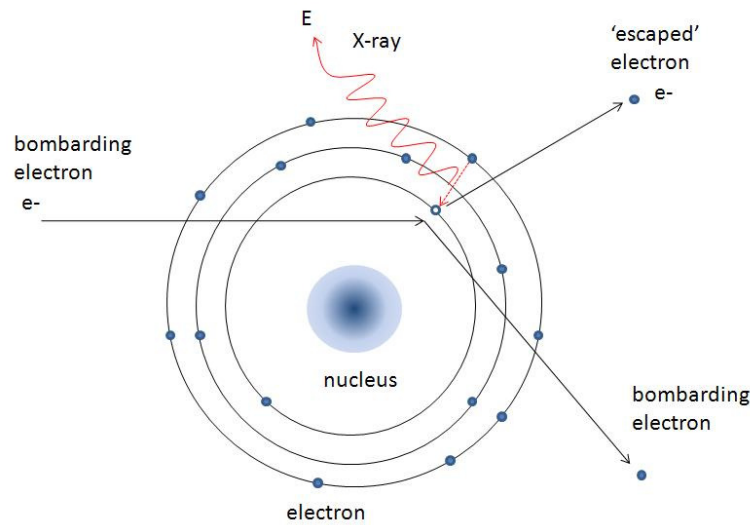


Figure 2.1: X-ray production begins with the bombardment of an atom with a striking electron. Ejection of the electron occurs when the energy of the striking electron is sufficiently large (greater than the shell binding energy). Then an electron from an outer shell would attempt to fill the vacancy, emitting characteristic x-rays, the energy of which is the binding energy difference of the two shells involved.

possible with incident x-rays of sufficiently high energy. When such an x-ray passes near the atomic nucleus, a pair of ions are ejected, positron e^+ and negatron e^- . Triplet production takes place when three particles are released, the aforementioned pair of ions plus a photoelectrically ejected electron, as a result of the powerful x-ray passing into the electron orbitals. In both cases, as soon as the positron has lost all of its initial kinetic energy, it quickly re-combines with any available electron; leading to annihilation radiation. These interaction mechanisms are illustrated in Fig. 2.2. In general, for a given atomic electron distribution, at low x-ray energy levels, the photoelectric effect is dominant and proportional to the atomic number; this is gradually overtaken by Compton scattering at high energy levels. The latter is the source of some key image artefacts (e.g beam hardening) in x-ray-based modalities, and will be explained in more detail later in this chapter.

Upon the absorption and/or scattering of the x-ray photons by the struck atoms, the original x-ray beam is attenuated. Attenuation is the process by which the beam loses its original intensity and is characterised by the attenuation coefficient μ and Lambert-Beers Law [9],

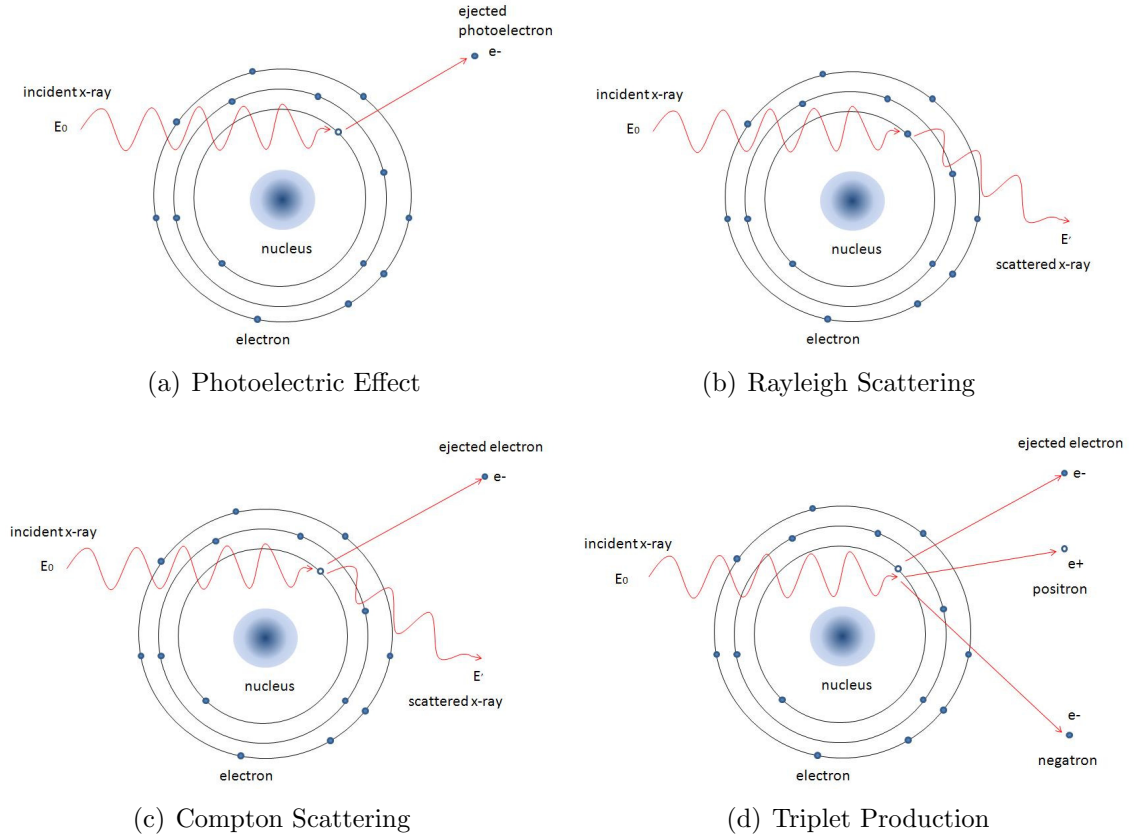


Figure 2.2: Interaction mechanisms of X-ray with matter. Ionisation occurs during the photoelectric effect whereas in Rayleigh scattering, the trajectory of the incident x-ray is affected.

where, for a beam of initial intensity I_0 , the attenuated beam intensity is given by:

$$I = I_0 e^{-\mu x} \quad (2.1)$$

The density, energy and atomic number of a material are all factors affecting the attenuation of an x-ray beam passing through it. Therefore by capturing the attenuated beams, it is possible to infer the composition of the material it passed through. Planar image information is acquired by projecting x-rays from different directions into the imaged body, as shown in Fig. 2.3.

Despite its early and wide usage, conventional x-ray radiography is unable to accurately image a scanned object due to the collapsing of 3D structures onto a single 2D image plane. For this reason, computed tomography (CT) was developed by Godfrey Hounsfield and Allan Cormack and first introduced to medical practice in 1971 [46]; it is capable of producing a 3D scan from a series of x-ray attenuation measurements. X-ray tubes with the tungsten vacuum

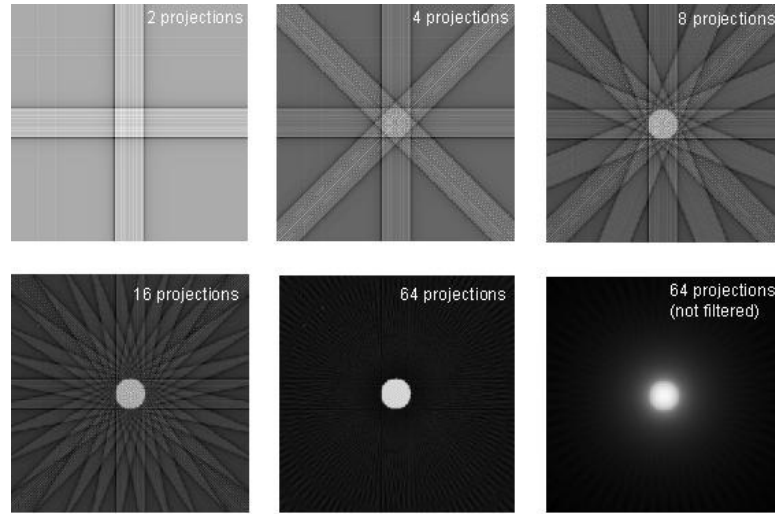


Figure 2.3: Image from X-ray attenuation measurements - radiation projections from different angles gives the 2-D image [59].

design explained earlier are used to produce the initial radiation. Collimation is applied before and after the x-rays passed through the patient to ensure a constant width of the x-ray beam. The detection of the attenuated signal is achieved using scintillators, which are made of crystal materials that respond to the absorption of ionising radiation by emitting light. The emitted light is captured by photodiodes which transform the light to an electrical analogue signal, or a sinogram, and passed on for image reconstruction. In modern digital CT scanners, the analogue reading is converted into a digital signal through an analogue-to-digital converter. The flow of a CT scan is outlined in Fig. 2.4.

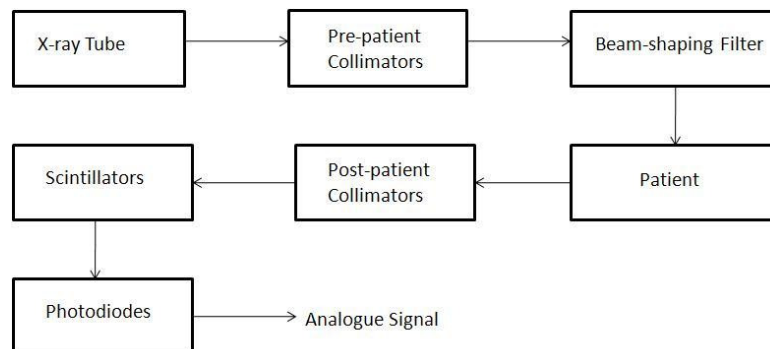


Figure 2.4: Flow of acquiring a CT scan

Modern scanners generally employ helical/spiral scanning, where the x-ray beam is cone-shaped and follows a helical trajectory relative to the patient. The attenuations are measured

by a moving set of scintillators and used to produce the image data. The practical advantages of the helical scanning is that it allows the simultaneous scanning of multiple image slices, thus greatly speeding up the scanning process as well as allowing for better spatial resolution, and the scanning of larger body volumes. Flexible image reconstruction at any position or interval is also made possible. The scanning speed can be further enhanced by using a multi-slice per gantry rotation detection design. Basic components of a CT scanner gantry are shown in Fig. 2.5.

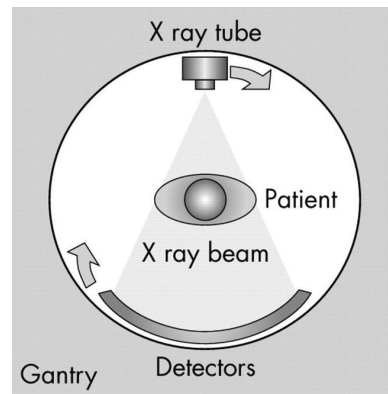


Figure 2.5: Gantry of a CT scanner [46]

A common image reconstruction procedure is ‘filtered backprojection’ (FBP), during which the linear attenuation data from each angular direction are convolved with an imaging filter and projected across a pixel field at the same angle to give the final image. Details of this procedure are provided in Fig. 2.6.

Finally, the CT intensity values are calculated and presented in Hounsfield units. To do this, the attenuation coefficient measurements μ_X are scaled to that of water according to Eq. 2.2. Hounsfield values for key body tissue components are given in Table 2.2.

$$CT \text{ number} = \frac{\mu_X - \mu_{water}}{\mu_{water}} \times 1000 \quad (2.2)$$

Contrast-enhancing(CE) agents are capable of boosting x-ray attenuation and can be administered intravenously to improve image contrast; one example being x-ray angiography where the blood vessels are highlighted using an iodine-based CE agent.

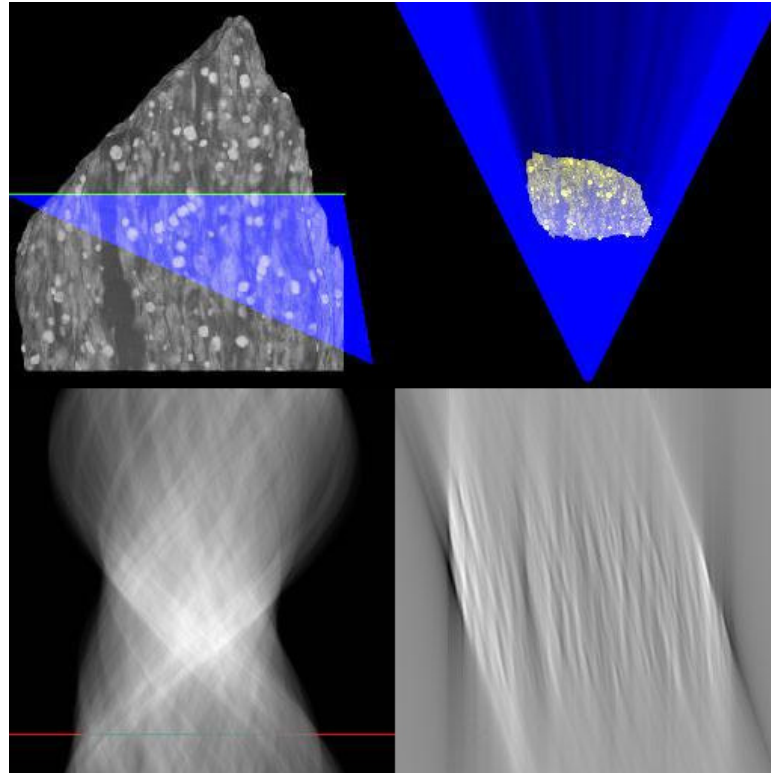


Figure 2.6: CT image reconstruction - the imaged object is scanned by rotating the scanner around it and its mid-section exposed to an angular x-ray beam. The attenuation of the beam is shown in the upper right, recorded and shown as the brightness in the bottom left figure, as a sinogram, where the horizontal axis is the detector channel and vertical axis the rotation angle. The sinogram is then convolved with a filter and projected on the pixel matrix (bottom right) on a row-by-row basis (indicated by the *red* scanline), for each and every angle interval. The image is complete once all angular readings are processed from the sinogram. [59]

2.2.2 Advantages

Compared to conventional 2D x-ray radiography, CT completely eliminates the overlapping of 3D structures from outside the area of interest. The good spatial resolution contributed by the helical scanning technique makes it possible to achieve a nearly isotropic scan, with a voxel size currently as small as $0.6 \times 0.6 \times 0.6 \text{ mm}$. This reduces, but does not completely eliminate, the influence of partial volume effect, which we will discuss more extensively in Chapter 6. In general, the signal-to-noise ratio is reasonably good for medical images, though this can be influenced by a series of scanning parameters such as exposure time, radiation dose used, collimation effectiveness, reconstruction algorithm and helical scanning speed. Moreover, CT has a high image contrast that can differentiate tissues of physical density difference as low

Tissue Type	Hounsfield Unit	Range
Bone (spinal cord, ribs)	+400 to +1000	
Soft tissues (tumour, liver and mediastinum)	+20 to +120	
Water (pleural effusion - part water)	0	
Fat	-60 to -200	
Air (air-filled lung spaces: aerated lung, unaffected lung)	-1000	

Table 2.2: Hounsfield unit ranges for various tissues in CT

as 1% [13]. Key tissue boundaries such as those of bone, air-filled cavities, and fat are clearly delineated on CT. A CT head scan can effectively detect infarctions, tumours, calcifications, haemorrhages, and bone trauma in and near the brain. Bone fractures and extremities in other parts of the body can also be quickly detected. Contrast-enhanced CT is useful for producing angiograms for diagnosing pulmonary embolism from an image of the pulmonary arteries.

Perhaps the most notable advantage of CT is its quick scanning speed, greatly facilitated by its helical scanning and multi-slice detector design. This is particularly useful in cases of significant motion such as that due to breathing. With a short scanning time of less than one breath hold [59](less than 5 minutes scanning time, as compared to 30 minutes for an MRI scan), CT is the predominant mode of imaging for the thoracic region. Comparing to MRI, it also resolves the issue of potential claustrophobia and avoids the need for anaesthesia for affected patients.

Lastly, as a low cost high resolution imaging modality, CT is readily available in most clinical settings and provides the quick access to meet urgent patient needs. For example, there are currently an average of 22.8 CT scanners per million people in OECD countries¹, comparing to 11 MR and 5.7 PET scanners for the same population [53].

¹Organisation for Economic Co-operation and Development, consisted of developed countries with generally high standards of living and healthcare

2.2.3 Limitations

The limitations of CT are equally obvious. Table 2.2 shows that although bones, water and air are clearly distinguishable from soft tissues, different soft tissues share much the same intensity range on the Hounsfield scale, mostly because of their similar physical densities and atomic compositions. This implies a low soft tissue contrast on CT, which makes the quantification of soft tissues on CT difficult.

Health and safety concerns are another disadvantage of CT compared to other modes of imaging. The adverse effect of radiation exposure during CT scanning remains a matter of controversy. Some studies [11] have claimed up to 0.35% additional risk of cancer mortality in children who have undergone CT head or abdominal scans (compared to the background risk of cancer mortality at 23%). The latest iodine-based CE agents used for contrast CT are generally safe although they may occasionally trigger allergies or induce kidney damage in patients with pre-existing renal or diabetic problems [11].

2.3 Image Artefacts of CT

There are a number of important artefacts associated with medical image scans. In particular, CT is prone to artefacts mostly because the scans are reconstructed from numerous independent x-ray detectors. The reconstruction algorithm assumes all such measurements are consistent and hence any errors present in the individual measurements are reflected in the reconstructed images. The artefacts can be categorised according to their causes. *Physics-based* artefacts result from the physical process during image acquisition. *Patient-based* artefacts are caused by patient movement or the presence of metallic materials in patients, for example pacemakers. *Scanner-based* artefacts derive from imperfections in the scanner function, and, finally, *Helical and multi-slice* artefacts are due to the reconstruction method employed. Most of these artefacts can be corrected by careful patient positioning, appropriate selection of scan parameter, as well as by post-scanning processing tools. It is the latter which we will predominantly investigate in our study.

2.3.1 Physics-based Artefacts

Physics-based artefacts may take the form of beam hardening [44]. As an x-ray beam is composed of photons of varying energies, when it passes through an object, the low-energy photons are absorbed much more rapidly than the high-energy ones, increasing the mean energy, that is, *hardening* the beam. In this regard, when the x-ray beam passes through the middle portion of a cylindrical object, the attenuation would be lower than those on the edges, giving rise to a ‘cup-shaped’ profile to the CT image along its cross-section. Moreover, for heterogeneous imaged cross-sections, the beam hardens less in some parts than in others, which might produce dark bands. This is particularly prevalent in bony regions of the body, such as the sternum and spine. Beam hardening may potentially misguide clinical decisions if left uncorrected. Solutions to hardening include filtration, which involves the use of a flat-piece filter that eliminates the lower energy photons prior to the beam passing through the patient body. The cupping effect can be compensated for by using a bowtie-shaped filter that ‘pre-hardens’ the edges of the beam [44], as illustrated in Fig. 2.7. Calibration provides another means to compensate for artefacts, for example the detectors may be adjusted and tailored to allow for beam hardening in the imaged part of the patient. An iterative correction algorithm [6] can be applied for the same purpose to reduce the effect of the dark band and also helps minimise blurring of the bone/soft tissue. Physics-based artefacts can also be caused by under-sampling which gives rise to aliasing where fine striations appear in the image [6]. However, this does not have much impact on the clinically relevant content of the image unless resolution of fine details is required. Furthermore, photon starvation occurs when the beam attempts to pass large areas with high attenuation such as the shoulder. In this case, insufficient photons may be captured by the detectors, leading to very noisy projections. When this noise is magnified during reconstruction, horizontal streaks are observed. This problem can be overcome with an increased x-ray dose, which however leads to the patient being exposed to an often unnecessary additional radiation dose. Alternative ways of minimising photon starvation include automatic tube current modulation where the scanner design allows the tube current to be varied automatically during the rotation, letting sufficient photons to pass through wider parts of the

body. Adaptive filtration is a software-based technique that targets places of low signal during the reconstruction stage.

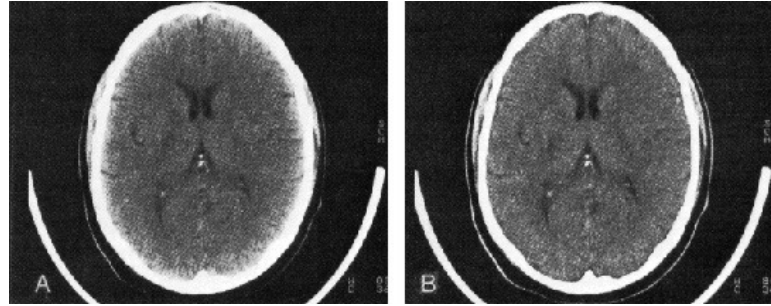


Figure 2.7: Cupping effect of CT scans and correction through filtration. Shown in the figure are the original (*left*) and corrected (*right*) phantom scans. Note the effect of the beam hardening correction. ‘Pre-hardening’ alleviated the ‘cup-shaped’ radiation profile in this scan [6].

The partial volume effect (PVE) is an intrinsic imaging phenomenon where the voxels on the boundary of an image region represent a combination of different tissue types. The overlap of different tissues creates problems in identifying the classes represented in the voxel, increasing inaccuracy on the overall pathology information inferred from the image. This can occur either in a planar form or most commonly along the axial (in/out-of-plane) direction. This tends to be exaggerated when long-thin structures, such as mesothelioma, are scanned and evaluated. The error due to PVE might in some cases be close to the estimated volume itself. Later in Chapter 6 of this thesis, we investigate solutions to reduce this problem. In addition, partial volume can also cause a shading artefact for dense objects, due to their off-centre alignments, as shown in Fig. 2.8. This can be ameliorated with a thinner slice scan, as illustrated in Fig. 2.9 with better defined boundaries. Note that in this case, the planar voxel size remains unchanged; the improvement in image accuracy derives entirely from the reduction of the partial volume shading effect.

2.3.2 Patient-based Artefacts

A number of patient-based artefacts are also problematic in CT image acquisition. As mentioned earlier, metallic objects such as pacemakers, dental fillings and surgical clips can affect the attenuation of the x-ray beam and hence the overall image accuracy. For this reason, avoid-

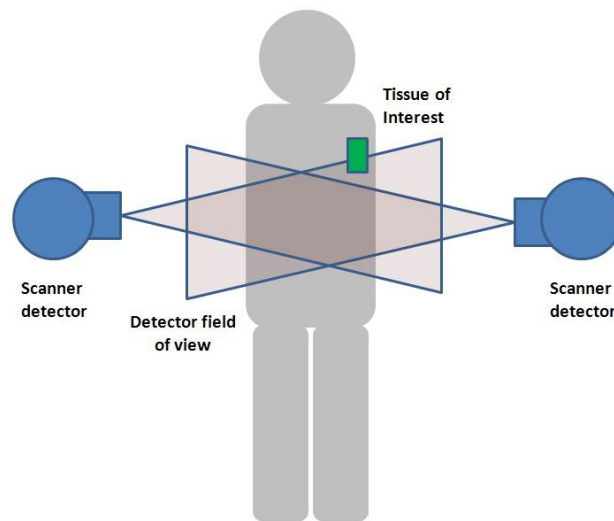
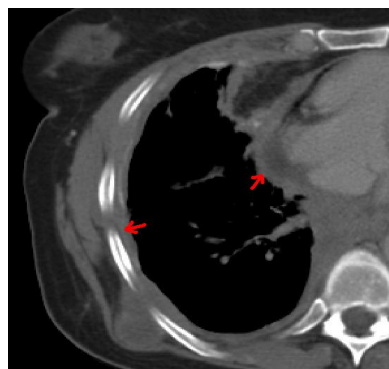
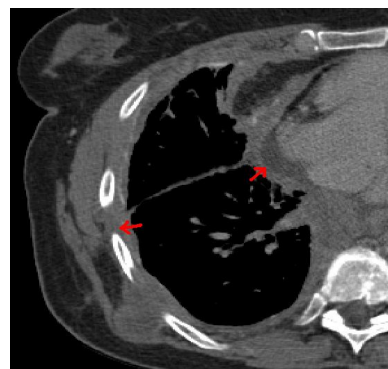


Figure 2.8: Partial volume shading effect



(a) Thick slice scan (5.0mm)



(b) Thin slice scan (0.625mm)

Figure 2.9: Illustration of the partial volume shading effect on MPM image data; pointed by arrows are the most severely affected regions.

ance of such objects is important during scan planning. Additional software corrections can be applied during reconstruction to compensate for situations where removal or avoidance of such materials is not possible.

Patient motion, as in most imaging modalities, poses a notable problem and has given rise to a fair amount of research works on motion correction. This is usually manifested on a scan in the form of shading or blurring. Avoidance of patient motion, such as the use of positioning aids or sedation, can be employed. Breath holding can be used to minimise the effect of respiratory motion during thoracic scans and often general instructions are given to all patients whose scan involves organs that move considerably, not least the lungs. The alignment of the scanning tube

in the motion direction can be used to accommodate for motion: for example, in the case of thoracic scans, dorsal and ventral to the patient. Some scanners can also have built-in features that scan for extra anatomical information beyond the 360° rotation. This information can be recombined later and reconstructed to compensate for the motion artefact.

A third type of patient-based artefacts is where patient dimensions exceed the scan field. This can be avoided by careful scan planning to ensure that no parts of the body lie outside the scan field or designing a scanner with greater scanning field of view [6].

2.3.3 Scanner-based Artefact

The ‘ring-artefact’ occurs on some CT scanners when one of the detectors is out of calibration. This presents as visible circular rings, which is the result of erroneous reading at each angular position. Re-calibration of the detectors resolves the problem.

2.3.4 Helical and Multi-slice Artefacts

For helical scanning, in regions with rapidly changing structures in the axial direction, the scanned shapes can appear distorted due to the weighting function used in helical interpolation and reconstruction. To solve this problem, steps can be taken to minimise the z-direction variation, by thinner slices, the use of a pitch of 1 and a 180° instead of a 360° helical interpolation [6].

With multi-slice scanning (defined in an earlier section in this chapter), the speed of the scan greatly reduces the effect of motion artefact. The use of narrower and overlapping slices also means that sharper edge resolution is possible. However, this creates an unwanted side effect, namely the zebra artefact, appearing as faint stripes and is caused by noise inhomogeneity along the z-axis from the helical interpolation. This is exacerbated in off-axis rotation cases, due to an increased noise inhomogeneity. Moreover, there might be trans-axial image distortion caused by the helical interpolator, which can be reduced with a better z-filter interpolator [6].

The cone beam effect occurs when the detectors rotate around the patient and the data acquired practically overlaps the cone-shaped field of view. In this case, there might be dis-

crepancies in the outer planes away from central scanning axis. This is usually resolved with the use of better reconstruction procedures that specifically adjusts for this artefact in thinner slices, where it is most pronounced.

2.4 Positron Emission Tomography

First developed in the late 1950s by David Kuhl and Roy Edwards [4], positron emission tomography (PET) is a 3D nuclear imaging technique that works by measuring the decay of a radioactive tracer.

2.4.1 Principles

Earlier in this chapter, we introduced the phenomenon of annihilation radiation in relation to certain x-ray interactions with matter. Before taking a PET scan, a radioactive tracer (typically fluorodeoxyglucose (^{18}F), or FDG), with a half-life of just under two hours is injected into the patient. When such a radioactive tracer undergoes positron emission decay during the scan, similar to the case of high energy x-ray triplet production, a pair of ions (positron and negatron) are released and eventually produce annihilation radiation upon the positron losing all of its initial kinetic energy and recombines with a free electron. Since the annihilation takes place almost at rest, and momentum has to be conserved, two annihilation photons are produced, moving in opposite directions, forming a straight line (shown in Fig. 2.10). Localisation of the annihilation event is made possible by tracking this straight line, referred to as the line of response, or LOR. As in the case of CT, a scintillator then detects these photons and responds by creating a burst of light. This is subsequently captured by a photodiode and transformed into an analogue signal. The raw signal is a sinogram that is analogous to that which is produced in CT, though with significantly fewer data counts.

Image reconstruction can be achieved in much the same way as in CT, using the FBP method, though the limited number of angular samples means that other methods have become more common, particularly Ordered Subsets Expectation Maximisation (OSEM). Because the

photons encounter tissues of varying thickness, they are attenuated to different extents in the decay process and must be corrected in attenuation in order to be of any practical use. In modern scanners, an x-ray CT scan is taken alongside the PET scan to build an attenuation map of density differences and used for the attenuation correction [4].

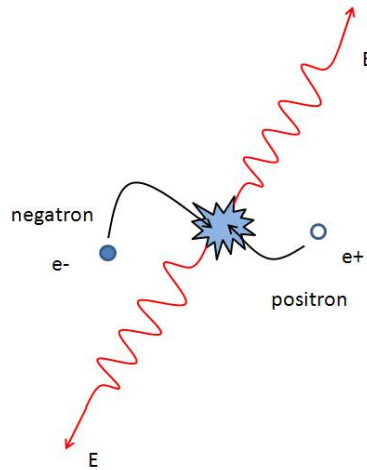


Figure 2.10: Annihilation radiation

2.4.2 Advantages

Unlike CT and MRI (excluding fMRI) where anatomical information is collected, PET measures metabolic activity. This is because the radioactive tracer travels via the circulatory system and accumulates in regions of high metabolism, marked by their glucose/radioactive marker intake. More signs of radioactivity are therefore expected to be measured from these areas. This can be particularly useful in cancer studies for detecting the presence of possible malignancy or lesions where metabolism is abnormally high.

In addition to showing good contrast for regions of high metabolic activity, PET offers a good spatial resolution for metabolic activity imaging [86].

Two types of PET scans are possible, namely, static and dynamic PET. The former produces a single image scan giving the anatomical distribution of metabolic activities. The latter consists of a sequence of contiguous PET scans taken at each time interval showing changes in metabolism due to outside interventions such as drug delivery.

As CT is required to correct for attenuation in PET in any case, multi-modal PET-CT scanners are becoming increasingly common. They allow for the simultaneous acquisition of both anatomical and metabolic information of the patient and their alignment for quick radiological interpretation. A sample PET-CT scan of a patient with MPM is shown in Fig. 2.11.

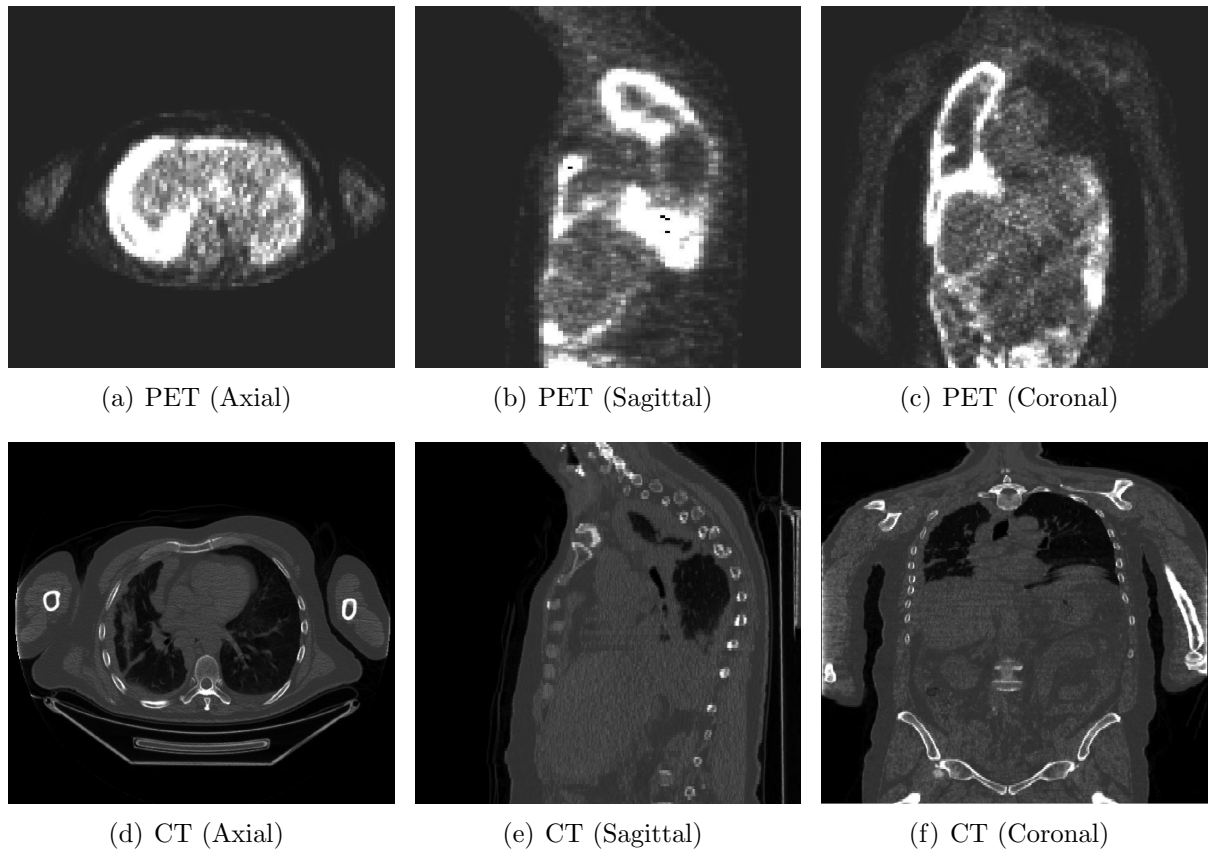


Figure 2.11: PET-CT scan of a MPM patient, with an axial scanning thickness of 2.5mm. Note acquisition of both anatomical and metabolic information.

2.4.3 Limitations

Like MRI, PET suffers from scanning motion in most cases. This is particularly prevalent in the case of thoracic scan. On the safety side, PET exposes the patient to substantial ionising radiation. This is exacerbated with the use PET-CT scanning where radiation doses as large as 23 – 26mSv (for a 70kg patient) [8] may be administered. By comparison, the annual background radiation in the UK is 2.2mSv, and thoracic CT scans typically use radiation doses of 6.5 – 8mSv [8].

The high costs of PET scanner and isotopes (£500k-1m) can be restrictive in some cases and prohibit their widespread usage.

2.4.4 Quantification

Quantification of an FDG-PET often uses the standardised uptake value (SUV), calculated as the average uptake of a radiotracer $C_{PET}(T)$, at a fixed time interval T after its initial injected, normalised by the injected dose amount over the patient body weight:

$$SUV = \frac{C_{PET}(T)}{\text{Injected dose/Patient's weight}} \quad (2.3)$$

SUV of the pixel with the highest image intensity is defined as SUV_{max} and provides a quantitative measure to assess the patient response to cancer treatment [8]. Total glycolytic volume (TGV) contains sites of glycolysis². It is a composite of tumour volume and glycolytic activity and has been shown to provide an alternative way of quantifying tumour on PET [49]. Finally, FDGhetero measures the metabolic heterogeneity of FDG uptake and is yet another way of quantifying tumours based on PET scans [35].

2.5 Experimental Data Acquisition and Chapter Summary

In this chapter, we have reviewed two imaging modalities including their advantages and limitations, namely CT and PET.

In our clinical study of MPM, helical CT is used as the primary modality to image the anatomy of the tumour and the surrounding structures, largely because of its short scanning time, which minimises chest motion artefacts during a scan. Some recent works [57][58] have suggested the use of MRI in thoracic tumour quantification, motivated by its avoidance of ionising radiations and better soft tissue contrast. For MPM scans, although soft tissues might

²breakdown of blood glucose

potentially be better delineated in MRI with motion correction, CT remains the gold standard for scanning the thoracic region in common practice [57]. Our experiments in this thesis are thus based on the quantification of MPM tumour from CT scans.

Clearly, the assessment of tumour response over the course of a treatment requires both anatomical and physiological information. Many clinicians in the past have based their decisions purely on the quantification of the tumour from anatomical scans. Novel multi-modal scanning techniques such as PET-CT have revolutionised the medical imaging field and cast new light on the topic of tumour quantification. A combination of anatomical and metabolic information helps in the understanding of the disease progression but is only useful when the images are well aligned. As a future work, PET-CT scans can be employed to provide additional means of evaluating tumour changes over time, from one or more of the three measures (SUVmax, TGV and FDGhetero) mentioned in the chapter. However, in order to incorporate this added knowledge to benefit anatomical CT quantification, cross-modal (PET to CT) image registration may be required. Finally, we note that motion correction may be necessary to render the PET scans useful for our study.

Chapter 3

Probabilistic Estimation of Tissue Signals for Image Segmentation

For CT scans of malignant pleural mesothelioma (MPM), tissues in the thorax (e.g consolidated lung, pleural effusion) often have similar attenuations to the primary tumour; rendering segmentation of the tumour from surrounding structures difficult. So it may be worthwhile to examine these tissues and the tumour from a probabilistic perspective. For this reason, the first component of our research is on the development of an effective tool for probabilistic density function (PDF) estimation.

3.1 Introduction

PDFs are central to many segmentation and registration techniques in medical vision. For instance, in image segmentation, each class of object (e.g tissue) represented in the image is associated with an observed likelihood PDF of the image intensities corresponding to that class. Such PDFs are then used, typically in a Bayesian framework, to segment the image into regions which correspond to the various tissue classes present. Another important aspect of the PDFs is that they provide a basis for computing other quantities, such as entropy and mutual information, which can be used to detect image features and to enable non-rigid alignment of images. Thirdly, PDFs provide a way to evaluate the similarity between images and hence

helps assessing the accuracy of experimental results against a certain ground truth.

The aim is to estimate the PDF of a continuous random variable X , given a randomly drawn sample $X_1, X_2, \dots, X_n$ of size n . In the cases we are interested in, $X_1, X_2, \dots, X_n$ correspond to a signal or image. All PDF estimates, denoted here by $f(x)$, satisfy the following properties:

$$f(x) \geq 0; \quad \int f(x)dx = 1 \quad (3.1)$$

Existing PDF estimators can be categorised broadly into parametric, non-parametric, and semi-parametric methods. Parametric estimators assume a particular parametric form (e.g. Gaussian, Poisson etc.), then determine the associated number of parameters (usually few). This can be solved using maximum likelihood in a Bayesian framework, given the prior and/or posterior information from the input data. However, for most medical applications, it is neither correct nor sufficient to assume a particular parametric form. For example, CT image noise typically do not assume a certain parametric form - though the quantum contribution to the noise, that caused by the number of incoming x-ray photons, follows a Poisson distribution [61]; anatomical structures are complex and variable; and as we have shown in the previous chapter, there are various imaging artefacts associated with CT. Semi-parametric estimators work by assuming a parametric form for component PDFs and find appropriate mixture proportions to approximate the overall PDF. A notable example is the Gaussian Mixture Model (GMM) estimator. Methods of this type are discussed more extensively in [31]. Like parametric estimators, they are often inadequate for solving difficult medical image problems. For these reasons, probabilistic analyses in this thesis assume a non-parametric nature, unless otherwise stated, where no assumptions are made on the form of the distribution.

We begin our discussion in Section 3.2 by introducing two widely-used PDF estimators, namely, histograms and the kernel density estimator (KDE). In Section 3.3, we describe the Non-Parametric Windows (NPW) estimator, a relatively new method for PDF estimation¹.

¹The theory and framework of the scalar version of this method were developed by current and former members of the research group, including Timor Kadir, Niranjana Joshi and Michael Brady. For this project, I have independently re-implemented this method for scalar quantities and extended it to the vector space.

We also present the steps we have taken to implement this method and apply it to analyse our clinical data, as well as the results of a quick pilot segmentation of MPM with a NPW-based segmentation method. However, the results suggest that for most MPM scans, due to the overlapping nature of tissue PDFs, image intensity alone might not be sufficient to give accurate and reliable segmentations. Thus, in Section 3.5, to investigate the possibility of incorporating other image features into our analysis, we develop an extension of the method for solving vector-valued quantities. Finally, we summarise this chapter in Section 3.6 and discuss potential future works.

3.2 Non-Parametric Probability Density Estimation

According to [30], non-parametric probability density estimations are most effective and advantageous in the following scenarios: 1) *exploratory* analysis, where the interest is on the descriptive features of the PDF such as multi-modality, tail-behavior and skewness. In this case NP methods are more flexible; 2) *confirmatory* studies such as classification, testing for mode and random variate testing. NP methods are more unbiased as they make no prior assumption about the PDF's parametric form; and 3) *presentational* purposes; because NP methods are intrinsically data-driven and can be easily communicated. To summarise, NP methods generally have the following favourable statistical properties: unbiasedness, consistency, and yielding *bona fide* estimates. The last of these properties relates to the results of a method which conforms to the two constraints for PDFs as specified in Eq. 3.1. Computational performance can sometimes be improved by relaxing one of these constraints. However, negativity of results may occur, as a result of data sparseness in certain regions. To remedy such issues, one may truncate the estimates for only the positive part followed by renormalisation. Alternatively, it is possible to apply a transformation to an estimate $f(x)$ such as $\log(f)$ or $f^{1/2}$ and inversely transform to re-construct the original results so that they are non-negative. In fact, it has been proposed [30] that any non-*bona fide* density estimate can be made to converge to a *bona fide* form.

In this section, we describe two popular non-parametric PDF estimators, namely, histograms

and the kernel density estimator (KDE).

3.2.1 Histograms

Histograms are by far the simplest and most widely used PDF estimator. A histogram is a mapping m_i , which is defined by the number of observations that fall into a series of disjoint intervals, or *bins*. More formally, given a total of n observations and k bins, the histogram m_i satisfies the following condition:

$$n = \sum_{i=1}^k m_i \quad (3.2)$$

However, despite its conceptual and computational simplicity, PDF estimation based on histograms is known to suffer from a number of complications including the binning problem. This refers to the lack of a scientifically informed method to estimate the optimal bin size, where to set the bin origin, or the relation between the bin and image sizes. A smaller than optimal bin size gives rise to an PDF estimate that is highly prone to signal noise, while one that is too large is overly smooth and incapable of presenting an accurate distribution representation [60]. A number of solutions have been proposed to address this problem. These include Stuges's formula which suggests a bin number of $k = \lceil \log_2 n + 1 \rceil$. This is known to perform poorly when $n < 30$. Additionally, a method proposed by Scott et al. [30] suggests a binwidth of $h = \frac{3.5\sigma}{n^{1/3}}$, where σ is the standard deviation of the input signal.

Consider a PDF f defined on an interval I , the actual bin is defined by binwidth h and a reference point $x_0 \in I$ [30]:

$$[x_0 + jh, x_0 + (j + 1)h]; \quad N_j/kh \quad (3.3)$$

The second expression gives the height of the histogram $H(x)$, so that the overall histogram is normalised. Formally, to assign x to bin B_i given an input signal X_i , the histogram is given by:

$$H(x) = \frac{1}{kh} \sum_{i=1}^k \sum_j J(X_i \in B_j) J(x \in B_j) \quad (3.4)$$

such that

$$J(X_i \in B_j) = \begin{cases} 1, & X_i \in B_j, \\ 0, & \text{otherwise.} \end{cases} \quad (3.5)$$

$$(3.5')$$

This means that the histogram in a certain bin B_j is only incremented when both the input signal X_i and sample x fall within that specific bin. It is now essential to assess the accuracy and convergence rate of a histogram estimator. The mean square error (MSE), defined below, is used to assess the accuracy of an estimation result.

$$\delta^2 = E \int_I [H(x) - f(x)]^2 dx \quad (3.6)$$

Given a number of mathematical assumptions, as listed in [22], we define the following variables:

$$\gamma = \int_I f'(x)^2 dx > 0; \quad \beta = 1/4 * 6^{2/3} * \gamma^{1/3}; \quad \alpha = 6^{1/3} \gamma^{-1/3} \quad (3.7)$$

With these notations, it is proposed in [22] that the optimal bin width, i.e to the one that minimises the MSE, is $\alpha k^{-1/3} + O(k^{-1/2})$ and the minimum MSE itself is $\delta^2 = \beta k^{-2/3} + O(k^{-1})$. In other words, we choose a cell width that is always twice the interquartile range of the data and divided by the cubic root of the sample size. Relaxing some of the mathematical assumptions, under weaker conditions, we found the optimal bin width and minimum MSE to be:

$$\alpha k^{-1/3} + O(k^{-1/3}) \quad (3.8)$$

$$\delta^2 = \beta k^{-2/3} + O(k^{-2/3}) \quad (3.9)$$

We note that Eq. 3.8 gives a more detailed version of Scott's formula. From Eq. 3.9, it can be noted that the convergence rate of the histograms estimator is of the order $k^{-2/3}$. This rate can be improved under asymptotic assumptions to $k^{-4/5}$, as proven in [22]. However, despite all these efforts, the limitations of the histograms estimator remain, as both Scott's and the Eq. 3.8 rely on prior knowledge about the signal's PDF (in the former it is the standard deviation, in the latter interquartile bounds), which are not generally available in applications such as ours. In addition, histograms are not differentiable, so the choice of origin becomes another performance-influencing factor. This can, however, be avoided using continuous histograms, or *histospline*; obtained by fitting a spline between values in a bin. Scott et al. [30] proposed another solution to this with the *averaged shifted histogram* or ASH that is constructed by averaging histograms of equal bin width and different bin locations.

3.2.2 Kernel Density Estimator

In comparison to the histograms, the kernel density estimator (KDE), also known as Parzen windows estimator, provides a good solution to the binning problem [26] and avoids the choice of origin. KDE is based on an idea that is similar to the histograms; but instead of defining the value of PDF $f_h(x)$ as $\frac{1}{kh} \times \text{observations that fall into a small interval containing } x$, it is defined as $\frac{1}{kh} \times \text{observations that fall into a small interval around } x$. With the notations as previously used in the histograms estimator definition, except for the binwidth h which now denotes the bandwidth, the general form of kernel density estimator is formally defined as [54],

$$f_h(x) = \frac{1}{k} \sum_{i=1}^k K_h(x - X_i) \quad (3.10)$$

$$\text{where } K_h(\bullet) = \frac{1}{h} K(\bullet/h).$$

Choice of the bandwidth has a major impact on the quality of the KDE estimation of the underlying PDF. Calculation of the optimal h forms the core of the KDE computation and implementation. The choice of the kernel function (typically a PDF in itself, most often a

Gaussian is chosen), on the other hand, is less important. We only discuss this briefly here.

A kernel function can be categorised into orders s . Univariate kernels that integrate to unity are defined as order 0, while kernels of order s are those whose $s - 1$ moments vanish, with their first finite terms being the s th moment. In this categorisation, second-order kernels have zero mean and finite variance. Kernels of the order $s \geq 0$ are only possible if they take negative values. These are very important for bias reduction and improving the mean integrated squared error convergence rate (MISE, defined in Eq. 3.11) of an estimate f_h to actual PDF f .

$$MISE\{f_h\} = E \int_{-\infty}^{\infty} [f_h(x) - f(x)]^2 dx = \int_{-\infty}^{\infty} MSE\{f_h(x)\} dx \quad (3.11)$$

It has been found that this rate is $O(n^{-2s/(2s+1)})$. For a fourth order kernel, for instance, this rate becomes $O(n^{-8/9})$, which is better than the histograms best convergence rate of $O(n^{-4/5})$.

We now discuss the computation of an optimal bandwidth for KDE. A popular approach to computing the optimal bandwidth is cross-validation. This involves removing a single value of the sample and computing the appropriate probability density estimate at this value, from the remaining sample values. Mathematically, this is expressed as:

$$f_{n,i}(X_i) = \frac{1}{(n-1)h} \sum_{j \neq i} K\left(\frac{X_i - X_j}{h}\right) \quad (3.12)$$

and h is chosen to optimise some given criterion involving values of $f_{n,i}(X_i)$. This criterion, for the case of likelihood cross-validation, is that h maximises the pseudo-likelihood given by $L(h) = \prod_{i=1}^n f_{n,i}(X_i)$. For the case of least squares cross-validation, h is chosen so as to minimise $LS(h) = R(f_h) - (2/n) \sum_{i=1}^n f_{n,i}(X_i)$. This is because this least square measure $LS(h)$ is exactly unbiased for $MISE - R(f)$. The performance of cross-validation is however quite mixed. The likelihood-based algorithm yields variable results. Much work has been done on evaluating its performance [30]. The criterion is found to be highly sensitive to outliers. It has also been found that the tails of the kernel K and estimate f have a complex influence on the bandwidth yielded

by the algorithm. The least squares-based method does not exhibit some of these variabilities but is also known to perform less well for large sample sizes. Computational burden is another issue when using these bandwidth optimisation schemes. Therefore, in general, for kernel density estimators, the goal is to achieve desirable results while being computationally efficient.

Other non-parametric methods are described in [30]. These include algorithms utilising locally adaptive smoothing to achieve good estimates. Built on the foundation of KDE, they aim to handle the clustering of sample values and take local peculiarities such as data regional clumping and sparseness into account. One such example is the adaptive kernel estimator, which is a two-step algorithm for computing a data-driven bandwidth. Orthogonal series estimators exploit the representation of the PDF as a series of orthonormals, which can be chosen from the list of candidates including Fourier, trigonometric and Legendre. Cross-validation-based algorithms are used to find the weighting of the individual components. In addition, the optimal number of terms will have to be determined and remains an open area of research.

3.3 Non-Parametric Windows Estimator

Four criteria are available to assess an estimation method: computational complexity, accuracy, sample requirement to attain a specific stability, and ease of setting up the parameters. According to these criteria, the histograms estimator are conceptually simple and computationally fast but require a large sample size to yield an accurate estimate. Moreover, they suffer from the binning and origin-defining problems. Kernel density estimators solve these issues and have a better convergence rate. However, determining the optimal bandwidth remains difficult: the currently used cross-validation-based algorithms are computationally demanding.

An alternative PDF estimator was developed independently by Kadir and Brady [34] and by Rajwade et al. [60]. It makes use of the concept of non-parametric windowing of the sample for probability density estimation, which we describe in this section.

3.3.1 Theory

Joshi, Kadir and Brady [33] noted that in most PDF estimators, the data samples are assumed to be independently and identically distributed (i.i.d). However, certain fundamental properties of a signal, such as the order of occurrence of the samples and band-limitedness, are ignored because most PDF estimators are based on statistical methods, which were originally developed for population data where these considerations do not apply. This assumption also helps in simplifying the analytical treatment of the estimation method. However, the omitted information can sometimes be useful for improving the estimation accuracy. For a continuous time signal, a pre-filter is first applied to bandlimit the signal, then it is sampled to create the output discrete signal. For exact reconstruction of the original signal, the Nyquist criterion requires that the signal be band-limited and that the sampling frequency be at least twice the signal bandwidth. In fact, Whittaker-Shannon theory states that three pieces of information are required to specify the original continuous signal: the samples, their order and pre-filter characteristics [34].

When we aim to smoothen a noisy signal in order to reconstruct the original continuous signal, for instance, by upsampling with linear interpolation, it is only possible to do so if we know the order of occurrence and the bandwidth from the pre-filter. However, a simple upsampling process can be impractical due to its computational requirements. Instead of upsampling by a finite factor, we may represent each piecewise polynomial section of the interpolated signal analytically. In this way we are able to bypass the need for the cumbersome numerical up-sampling process. This also yields a continuous estimate and requires no setting of parameters such as binwidth/bandwidth. This is the basic idea behind the non-parametric (NP) windows estimator, which only considers samples that are band-limited and critically sampled (BL-CS). NPW employs an interpolation model that builds a continuous signal from discrete signal first and then performs PDF estimation.

This method estimates the target PDF by fitting functions piecewise to sections of the signal. Note that the re-sampling procedure common in interpolation schemes is avoided for better accuracy and computational efficiency. In this case, the accuracy depends only on the

piecewise representation and not the number of samples or bins. Splines are chosen in [34] as the piecewise function to use. This choice was motivated by mathematical tractability and computational speed. This is because a zero-order spline would give a piecewise constant interpolation with the overall estimation equivalent to the estimation by histograms. An infinite order spline, on the other hand, corresponds to a Gaussian kernel. So the order of the spline controls the estimation procedure.

In this method, we follow the following procedures for PDF estimation [34]:

- 1) Calculate the polynomial coefficients for the data sample.
- 2) Calculate the PDF for each piecewise section.
- 3) Populate the appropriate bins for each piecewise section.

To introduce the method in more detail, consider a 1D signal $y(x)$. We follow [34] and start with the simplest example of all: for a uniform PDF of the domain variable x that is taken to be unity over its range: $f_x(x) = 1, 0 \leq x \leq 1$, the PDF estimate of the section is equivalent to mapping the PDF of the domain variable to that of the sample. This transformation, for a monotonic function, is given by:

$$f_y(y) = \frac{1}{|dy/dx|} f_x(x) \quad (3.13)$$

For the univariate case, two scenarios are possible, depending on the choice of the piecewise section. For the linear case, $y(x) = ax + b$ and the sectional PDF estimate is:

$$f_y(y) = \frac{1}{|a|} f_x\left(\frac{y-b}{a}\right) = \frac{1}{|a|}, \quad b \leq y \leq a+b \quad (3.14)$$

The overall PDF can then be found by superimposing all piecewise functions over their corresponding ranges, as specified by $[b, a+b]$. Similarly for the quadratic case, we use $y(x) = ax^2 + bx + c$. However, this case presents a challenge in that quadratics are often not monotonic over the range. This can be either solved by first detecting points of non-monotonicity

and adjusting the PDF estimate appropriately or by re-sampling the spline spans so that each section is monotonic. For design simplicity, we choose the former approach. Returning to the quadratic assumption, we know its inverse form is,

$$x(y) = \frac{-b \pm \sqrt{b^2 - 4a(c - y)}}{2a} \quad (3.15)$$

To deal with non-monotonicity, we observe that within the spline case, Eq. 3.15 may exhibit one of the following three characteristics: single valued, multiple valued, or a combination of both. Because the quadratics are symmetrical about their extrema, it is possible to simplify the calculation by considering only one root and doubling the PDF estimate for that section. This can be summarised as the following:

$$f_y(y) = \frac{1}{|2ax + b|} f_x(x) = \frac{1}{|\sqrt{b^2 - 4a(c - y)}|}, \quad c \leq y \leq a + b + c \quad (3.16)$$

Now we move onto the bivariate case where the intensity y_1 depends on two positional variables x_1 and x_2 . A number of implementations are possible but in [34], only the bilinear case is considered. Improved implementation ideas will be discussed in the next section. In this case, we introduce a dummy variable y_2 . The definitions of key variables are as follows:

$$y_1(x_1, x_2) = ax_1x_2 + bx_1 + cx_2 + d \quad y_2(x_1, x_2) = x_1 \quad (3.17)$$

$$x_2(y_1, y_2) = \frac{y_1 - by_2 - d}{ay_2 + c} \quad x_1(y_1, y_2) = y_2 \quad (3.18)$$

To find the transformation, we use the equivalence of a derivative in the multivariate space, that is, a Jacobian,

$$|J| = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \frac{1}{ay_2 + c} \quad (3.19)$$

The joint PDF of y_1 and y_2 is:

$$f_{y_1, y_2} = f_{x_1, x_2}(y_2, \frac{y_1 - by_2 - d}{ay_2 + c})|J| = \frac{1}{ay_2 + c}$$

$$0 \leq y_2 \leq 1, \quad by_2 + d \leq y_1 \leq y_2(a + b) + c + d \quad (3.20)$$

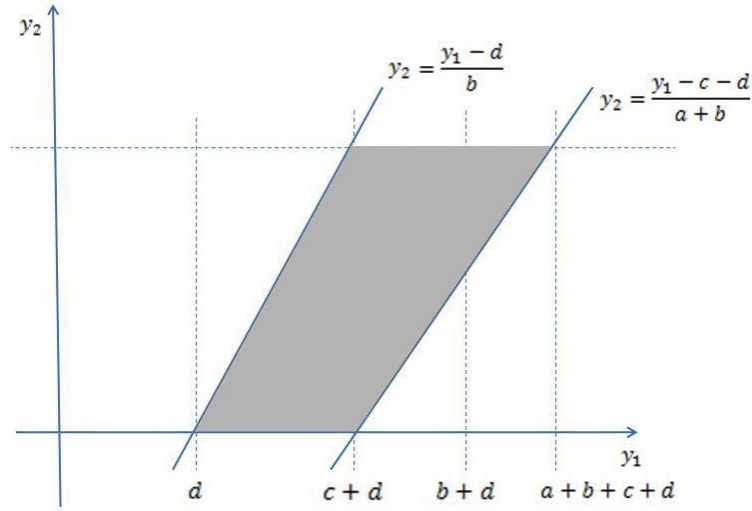


Figure 3.1: Integration ranges for the NPW 2D bilinear implementation, as shown in *grey*

Given Eq. 3.20, we can define the integration ranges graphically, as shown in Fig. 3.1. Therefore, the final equations can be derived as the following, based on Eq. 3.20 and Fig. 3.1:

$$\frac{1}{a} \ln\left(\frac{ay_1 - d + cb}{cb}\right) : d \leq y_1 \leq d + c \quad (3.21)$$

$$\frac{1}{a} \ln\left(\frac{a + b}{b}\right) : d + c \leq y_1 \leq b + d \quad (3.22)$$

$$\frac{1}{a} \ln\left(\frac{(a + c)(a + b)}{ay_1 - d + cb}\right) : d + b \leq y_1 \leq a + b + c + d \quad (3.23)$$

Finally, a joint PDF estimator can be developed in a similar manner. For simplicity of illustration, we use the 1D univariate linear case:

$$y_1(x_1, x_2) = ax_1x_2 + b \quad y_2(x_1, x_2) = cx_2 + d \quad (3.24)$$

$$x_1(y_1, y_2) = \frac{y_1 - b}{a} \quad x_2(y_1, y_2) = \frac{y_2 - d}{c} \quad (3.25)$$

Finding the Jacobian from Eq. 3.24 and 3.25, the joint PDF of y_1 and y_2 can be written as:

$$f_{y_1, y_2} = f_{x_1, x_2}\left(\frac{y_1 - b}{a}, \frac{y_2 - d}{c}\right) |J| = \frac{1}{ac}, b \leq y_1 \leq a + b; d \leq y_2 \leq c + d \quad (3.26)$$

NPW was benchmarked against the histograms and kernel density estimators [34]. It was found that this method is superior to the histograms estimator in its estimation smoothness and to the KDE in its parameter setting and computational requirement. It is worth noting that in [34], only the 1D linear and quadratic and 2D bilinear cases were considered and analysed. Results were assessed based on accuracy and stability measures and compared to the histogram and KDE estimators.

An independent development of the NPW estimator was carried out by Rajwade et al. [60]. In their work, singularities in the density estimate were considered and their method was applied to a mutual information (MI)-based image registration. More specifically, divergent behaviour of the regional joint density was analysed.

The NP windows estimator can be applied to compute entropy for feature detection [31]. The performance is generally superior to that using histograms. The computation time of the NPW estimator was studied and compared to other PDF estimators in [31]. It was found that NP windows estimator is of the order $O(10^3)$ faster than KDE for processing the same signal.

As noted earlier, a 1D windows estimator has convergence rate of order n^{-1} . 2D can be slower. However both are significantly faster than the histograms estimator for the same estimation accuracy and smoothness [31]. Compared to KDE, NP windows performs better for 1D signals and similar for 2D signals in terms of accuracy. However, the computation time performance of the NP windows estimator is much better.

3.3.2 Implementation

Theories and ideas for implementing the scalar NP windows estimator have been developed earlier by Joshi [31]. Though in my study, I have re-applied and coded these ideas, for learning and adaptation purposes.

NPW for 1D signals

Key steps in developing an NP window estimator can be summarised as follows: given a signal $y(x)$, where y is a function of the random variable x whose PDF is assumed to be known.

- 1) Find the dependency relationship between x and y . (for linear interpolation this is $y = ax + b$) Here y is the dependent variable and x is the independent variable.
- 2) Find the inverse relationship to that in step 1. i.e x : dependent variable, y : independent variable.
- 3) Assume a PDF distribution for $x : f_x(x)$. (i.e Gaussian, uniform) Uniform distribution is assumed for this case.
- 4) Find $f_y(y)$ from $f_x(x)$, for the 1D case, $f_y(y) = \frac{1}{|dy/dx|} f_x(x)$.
- 5) Then using the equation from step 4 and result from step 3, find the analytical form of the PDF approximation for y , in the 1D case, we have $f_Y(y) = \frac{1}{|a|}, b \leq y \leq a + b$. The boundaries are found by $y = ax + b$, for $x = 0, 1$.

NPW for 2D signals

For 2-D signals, we have two positional variables x_1, x_2 . We let y_1 be the intensity value. For analytical convenience and constructing a Jacobian, we also introduce a dummy variable y_2 . To find the PDF estimate of y_1 , four different ways of implementation methods have been proposed. Out of the four NP windows implementations, the box basis is the fastest for medium-sized signals but is slower than planar approximation for large images. Joshi [31] also showed that a log basis method is good for smaller quantity of data samples.

I. Bilinear interpolation-based approximation

We have already discussed this in the previous section. Here, we only summarise its steps for completeness. The implementation works by assuming that y_1 relates to positional variables through bilinear interpolation over an area made up of four mutually neighbouring pixels. Uniform distribution is assumed for positional variables x_1 and x_2 over this region. The estimator is then developed as follows:

- 1), 2) We use the definitions from Eq. 3.17 and 3.18.
- 3) Find the assumed form of $f_x : f_{y_1, y_2}(y_1, y_2) = f_{x_1, x_2}(y_2, (y_1 - by_2 - d)/(ay_2 + c))|J| = \left| \frac{1}{(ay_2 + c)} \right|$
- 4) $f_{y_1, y_2}(y_1, y_2) = \left| \frac{1}{ay_2 + c} \right|$ is subject to boundaries $0 \leq y_2 \leq 1$ and $0 \leq \frac{y_1 - by_2 - d}{ay_2 + c} \leq 1$
- 5) Marginal PDF $F_{y_1}(y_1)$ can be found by integrating out the dummy variable y_2 over the ranges given. i.e $f_{y_1}(y_1) = \int_0^1 f_{y_1, y_2}(y_1, y_2) dy_2$.

In this method, there are 4! coordinate frames. However, by a judicious choice of coordinate frame, we can reduce this number significantly if we always set the origin to be the minimum of the four pixels in the patch. Sorting all pixels as such, we are able to reduce the number of possible configurations to three. Individual PDF estimates obtained over each bilinear section can then be added and normalised to produce a PDF estimate of the full data.

II. Planar approximation

This implementation assumes that the intensity surface which is sampled by the image is piecewise planar. This linear relationship is defined by three mutually neighbouring pixels that form a triangle. Similar to the bilinear case, a uniform distribution is assumed for positional variables over this triangular region. The method is developed as follows:

- 1) $y_1(x_1, x_2) = ax_1 + bx_2 + c$, $y_2(x_1, x_2) = x_1$.
- 2) $x_2(y_1, y_2) = \frac{y_1 - ay_2 - c}{b}$, $x_1(y_1, y_2) = y_2$

- 3) $f(y_1, y_2)(y_1, y_2) = f_{x_1, x_2}(y_2, \frac{y_1 - ay_2 - c}{b})|J| = \frac{b}{2}$ since $|J| = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} a & 1 \\ b & 0 \end{vmatrix} = -b$ and $f_{x_1, x_2}(y_2, \frac{y_1 - ay_2 - c}{b}) = -\frac{1}{2}$.
- 4) The above expression is subject to $0 \leq y_2 \leq 1$; $0 \leq \frac{y_1 - ay_2 - c}{b} \leq 1$; $\frac{y_1 - ay_2 + by_2 - c}{b} \leq 1$

It is easier to choose an origin on the right angle of the triangle. The x_1 axis is chosen to be on the vertex that is the smaller of the remaining two vertices. With the three vertex pixels sorted v_{min} , v_{middle} and v_{max} , a new coordinate system is used, such that $c_{new} = v_{min}$, $a_{new} = v_{middle} - c_{new}$, and $b_{new} = v_{max} - c_{new}$.

III. Box-basis approximation

This implementation relies on transforming the PDF of a linear section to a box function with appropriate translation and scaling. The box function is defined by:

$$B(y_1) = \begin{cases} 1, & \text{if } -0.5 \leq y_1 \leq 0.5, \\ 0, & \text{otherwise.} \end{cases} \quad (3.27)$$

$$(3.27')$$

$$B_{\mu, \sigma}(y_1) = \frac{1}{\sigma} \times B\left(\frac{y_1 - \mu}{\sigma}\right) \quad (3.28)$$

Also given above is the normalised version of the box function, with μ being the mean and σ the width; The development of this implementation is outlined below:

- 1) $y_1 = (ax_1 + c)x_2 + (bx_1 + d)$.
- 2) $\mu(x_1) = \frac{ax_1 + c}{2} + bx_1 + d$, $\sigma(x_1) = ax_1 + c$.
- 3) Represented by box basis, the PDF over the pixel patch is: $f_{y_1}(y_1) = \int_0^1 B_{\mu(x_1)\sigma(x_1)}(y_1)dx_1$.
- 4) Typically the equation in step 3 is calculated numerically as: $f_{y_1}(y_1) = \sum_{n=1}^N B_{\mu_N \sigma_N}(y_1)$.

IV. Log basis approximation

This method is based on bilinear interpolation and enjoys a considerable computational advantage over the box basis method. In this case, consider the bilinear case with b and c being zero and $a \geq 0$. The component PDF can then be approximated by $f_{Y_1}(y_1) = -\frac{1}{a} \ln(\frac{y_1-d}{a})$ for $d \leq y_1 \leq a+d$. Note that when b and c are zero, the lines joining the corners $d, d+b$ and $c+d$ are zero slope lines. For these lines, linear relationship $y_1 = (ax_1 + c)x_2 + (bx_1 + d)$ gives the corresponding values for x_1 as $x_1 = -\frac{c}{a}$ and $x_2 = -\frac{b}{a}$. The constant intensity value is $y_1 = d - \frac{bc}{a}$. In summary, we note that we can subdivide the pixel into four subpixels, formed by the two zero slope lines. Out of the four corners, three have constant intensity values because they lie on the zero slope lines. The fourth takes an arbitrary value. Weighted by the areas of the subpixels w_i , we have the PDF of the whole pixel as $f_{y_1}(y_1) = \sum_{i=1}^4 -\frac{w_i}{|y_1(p_i) - y_1(p_5)|} \ln \frac{y_1 - y_1(p_5)}{y_1(p_i) - y_1(p_5)}$.

NPW for 3D signals

In most medical applications, the aim is to construct a 3D picture of the imaged volume. Extending the NP windows estimator to 3D follows the procedure outlined in [33]: Initially, start with a planar approximation implementation; instead of a 2D pixel path, consider a 3D volume: five different tetrahedra can be formed to complete the cube. Therefore in this implementation, it is essential to add all possible tetrahedral cases individually to find the PDF representation of the overall 3D signal. As before, we approach the problem by letting y_1 denote intensity and using positional variables x_1, x_2 and x_3 ; also introduce dummy variables y_2 and y_3 . Key steps are given below:

- 1) $y_1(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3 + d$; $y_2(x_1, x_2, x_3) = x_2$; $y_3(x_1, x_2, x_3) = x_3$;
- 2) Uniform distribution is assumed for the positional variables such that $0 \leq x_1, x_2, x_3 \leq 1$ and $x_1 + x_2 + x_3 \leq 1$. This means $f_{x_1, x_2, x_3}(x_1, x_2, x_3) = 6$.
- 3) The next step is to develop the approximation equation. First, the inverse equations are found: $x_1 = \frac{y_1 - by_2 - cy_3 - d}{a} = x_2 = y_2$ and $x_3 = y_3$, which are subjected to the following constraints:

$$0 \leq y_2 \leq 1 \text{ and } 0 \leq y_3 \leq 1$$

$$\frac{y_1 - cy_3 - (a+d)}{b} \leq y_2 \leq \frac{y_1 - cy_3 - d}{b} \text{ and,}$$

$$\frac{y_1 + (a-c)y_3 - (a+d)}{b-a} \geq y_2$$

Calculating the Jacobian we obtain $f_{y_1, y_2, y_3} = \frac{6}{a}$.

- 4) Then the dummy variables are eliminated for marginal PDF by integrating according to the constraints given above. Finally, the following equations are found:

$$f_{y_1}(y_1) = \frac{3(y_1 - I_0)^2}{abc} \text{ for } d \leq y_1 \leq c + d \quad (3.29)$$

$$f_{y_1}(y_1) = \frac{3}{a} \left[\left(\frac{y_2 - I_0}{b} \right) \left(\frac{y_1 - I_2}{c - b} \right) + \left(\frac{y_1 - I_1}{c - a} \right) \left(\frac{y_1 - I_3}{b - c} \right) \right] \text{ for } c + d \leq y_1 \leq b + d \quad (3.30)$$

$$f_{y_1}(y_1) = \frac{3(y_1 - I_1)^2}{a(b - a)(c - a)} \text{ for } b + d \leq y_1 \leq a + d \quad (3.31)$$

where $I_0 = d, I_1 = a + d, I_2 = b + d, I_3 = c + d$.

Considering different coordinate frames, a new set of coefficients is used in the implementation with vertices sorted by intensity values in increasing order: $d_{new} = v_1, c_{new} = v_2 - v_1, b_{new} = v_3 - v_1, a_{new} = v_4 - v_1$.

Coding

To better understand the method, we re-implemented NPW estimators in MATLAB, coded in .m files. For the reasons explained above, the histograms estimator is implemented and used as the ground truth for comparison and assessment purposes. Pseudo-codes for the implementations and corresponding results are presented below. In our 2D implementation we have used the planar approximation method. This choice is mainly motivated by its proven performance and conceptual simplicity.

First for univariate (1D) samples (Algorithm 3.1),

Algorithm 3.1: Pseudo-code for 1D Windows

```

generate a Gaussian signal of given length and parameters  $\mu$  and  $\sigma$ ;
find and save its max and min values; final PDF estimate = 0, whose length is found from
the above max and min values;
set binwidth, usually in powers of two;
for each element of the generated signal
    find parameters a and b according to linear relationship  $y = ax + b$ ;
    calculate PDF estimate for the current bin according to NP estimator equation;
    find the slot in the final estimate that the current estimate ...
    belongs to and add this to the existing final PDF estimate:
    if current bin estimate < next bin estimate
        while currentEstimate(index) < current bin value
            continue search and increment index
        end
    end
    else
        while currentEstimate(index) < next bin value
            continue search and increment index
        end
    end
end
plot the final PDF estimate and compare with the ground truth;

```

Similarly for the 2-D (Algorithm 3.2) and 3-D cases (Algorithm 3.3):

Algorithm 3.2: Pseudo-code for 2D Windows

```

generate a 1D Gaussian signal of given length and parameters  $\mu$  and  $\sigma$ ;
reshape the 1D signal to create a 2D signal of appropriate length;
find max and min values of the signal;
final PDF estimate = 0, whose length is found from the above max and min values,
with two arrays, the second one representing values;
initialise an array for storing the pixel triplet; set binwidth;
for each element of the signal array
    take the triplet pixels from signal array;
    assign  $a_{new}$ ,  $b_{new}$  and  $c_{new}$  from values of pixels in the triplet;

    if both  $a_{new}$  and  $b_{new}$  are not 0
        for range 1
            apply relevant equation from ( 3.3.2);
        for range 2
            apply relevant equation from ( 3.3.2);
        end
        ... similarly for all other cases; add bin PDF estimate to the overall PDF value;
    end

for bin of the signal array this time for the second half of the possible triplet combinations
    perform similar operation as in the previous loop;
end
add results from both loops and plot for visualisation;

```

Algorithm 3.3: Pseudo-code for 3D Windows

```

generate a 1D Gaussian signal of given length and parameters  $\mu$  and  $\sigma$ ;
reshape the 1D signal twice to create a 3D signal of appropriate length;
find max and min values of the signal;
final PDF estimate = 0, whose length is found from the above max and min values, with
two arrays, the second one representing values;
initialise an array for storing the pixel quadruple; set binwidth;
for each bin of the signal array
    take the quadruple from the signal array;
    sort the quadruple pixels in increasing order;
    assign  $a_{new}, b_{new}, c_{new}$  and  $d_{new}$  from the sorted pixels;
    assign  $I_0, I_1, I_2$  and  $I_3$ 
if  $a_{new}, b_{new}$  and  $c_{new}$  are all equal to zero
    for range 1
        apply formula to calculate PDF;
    for range 2
        apply formula to calculate PDF;
    end
    ...similarly for other ranges and boundary constraints
    add to existing final PDF estimate;
end
for four other tetrahedra
    add to existing final PDF estimate;
end
plot for visualization;

```

The implementations were first applied to synthetic Gaussian signals with noise ($\mu = 10, \sigma = 20$); the results are shown in Fig. 3.2. Note that the NPW estimator performs quite well, with the expected smoothing effect. The computation times for different implementations are given

in Table 3.2; these are based on using a PC with Pentium-D CPU 3.39GHz with 2GB of RAM. Note that the histograms estimator runs much faster than the NPW estimator. Histogram can produce smooth estimates with upsampling of the data samples. However, this presents several issues: First, the upsampling factor is generally not known and needs to be determined at the outset. We also need to interpolate all the samples. Both of these add to the computational requirement of the algorithm. Moreover, upsampling is only possible if the order of occurrence of the samples is given, which is usually not the case for histograms-based reconstruction, as discussed in Section 3.3.1. We assess the performance of NPW when applied to MPM CT scans in the next subsection.

3.3.3 Results and Discussion

Ground Truth

Before assessing the real-time performance of the NP windows estimator, the issue of ground truth must first be addressed. Since the data that is necessary to find the ground truth PDF is lost in the discrete sampling process, cubic spline interpolation can be applied to construct the original continuous signal; which is superior to linear interpolation because of its better accuracy. The histograms estimator is then used with a sufficiently large number of interpolated samples, which is known to converge to the true PDF. The term ‘ground truth’ will be used in this context for non-MPM-related works. For the MPM-related works reported in this thesis, we have used manual delineations of tumour and key anatomical features present on the thoracic CT scans. These are drawn by a clinical radiology specialist at the Churchill Hospital, Oxford, who is an expert in both CT and MPM and has provided clinical support and counselling throughout our study. Because these segmentations are highly prone to human error, for clarity, we refer to them as the reference truth for works presented in this chapter.

The 1D, 2D and 3D scalar implementations of NPW, re-implemented in our study based on ideas developed earlier by Joshi [31], are first validated with the reference truth in Fig. 3.2. The KDE optimal bandwidth is estimated using the cross-validation selector proposed in [25]. In terms of its computation time, NPW outperforms KDE by an order of 1000; this is mainly

due to the time required to complete the bandwidth estimation step. The ‘smoothing effect’ of NPW over the histograms estimator is also clearly observed in these results. Note that in this case, we are not evaluating the accuracy of NPW, since random signals do not satisfy the BL-CS assumption described earlier.

3.4 Application to Medical Image Analysis

PDFs have a wide range of applications in medical image analysis. For instance, in image segmentation, each class depicted in the image is associated with an observed likelihood PDF of image intensities. These PDFs are then typically used in a Bayesian framework to segment the image accordingly. Another important aspect of the PDFs is that they give knowledge about other quantities of the image, such as entropy and mutual information, which have proven utilities in medical image tasks such as image registration and partial volume correction. Given the accuracy and excellent real-time performance of the NPW estimator demonstrated in the previous section, we are now ready to apply this concept to the field of mesothelioma detection and quantification.

First, to illustrate an obvious application of NPW in medical image analysis, a T2-weighted brain MR scan is used. As given in Fig. 3.3, the estimated PDF shows distinct components which represent the three key areas in the region, namely, in the order from dark to light on the scan, cerebral spinal fluid (CSF), grey (GM) and white matters (WM). The overlapping of the component curves is minimal in this case therefore rendering the PDF highly useful and predicts good applicability of PDF-assisted thresholding for segmentation. However, as we soon will show in this thesis, this is hardly the case for our MPM data.

3.4.1 Applying NPW to MPM Data

Procedure

We aim to apply NPW to estimate the PDFs of intensity values of individual tissues defined by axial slices of the image volume, following partial manual segmentation by the clinician.

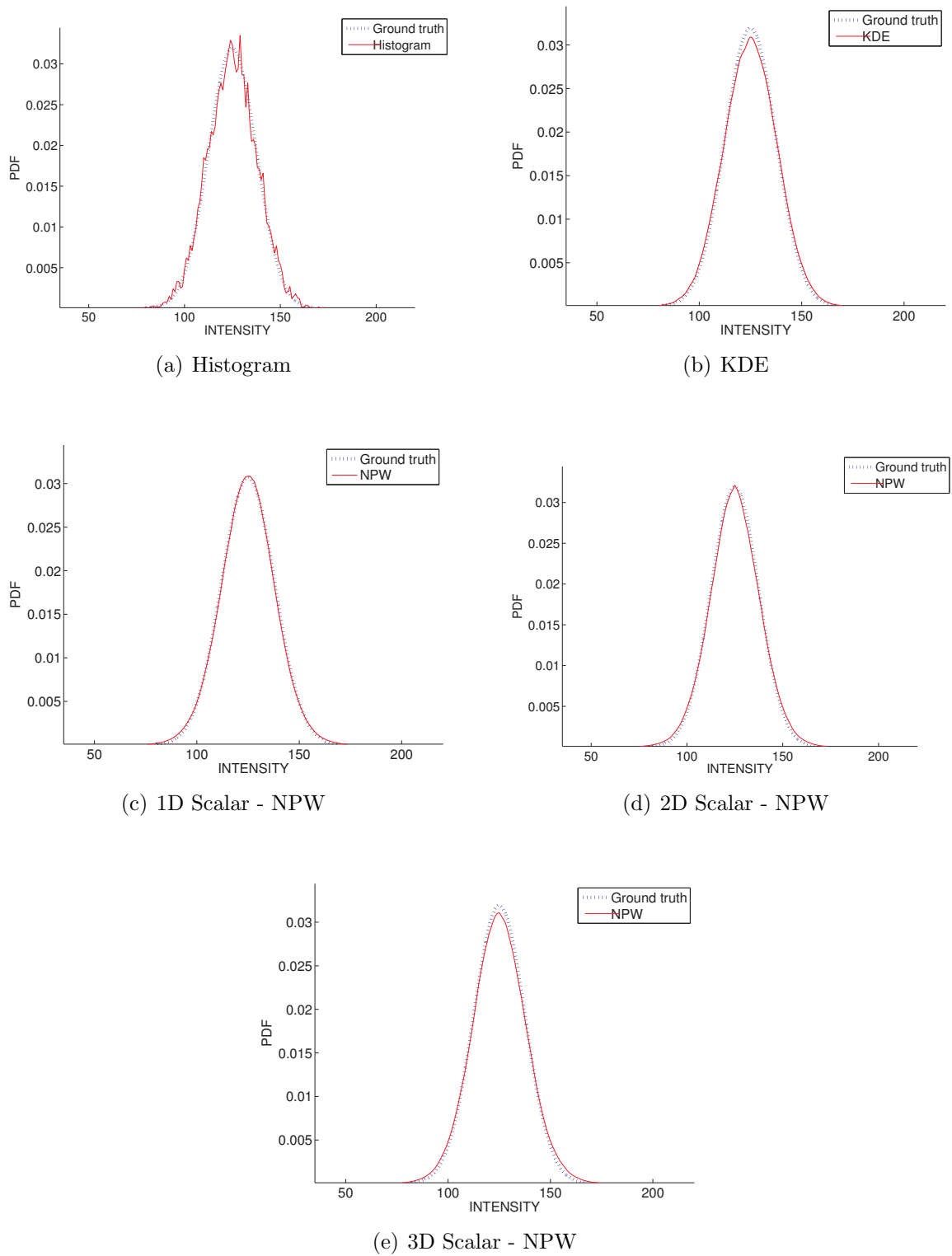


Figure 3.2: Performance of 1D, 2D and 3D NPW for estimating the PDF of synthetic scalar signals, as compared to those of the histograms and KDE methods. Normal distributions of 3000 samples are used in all cases; Ground truths are the actual synthetic distributions themselves, with specified means and standard deviations.

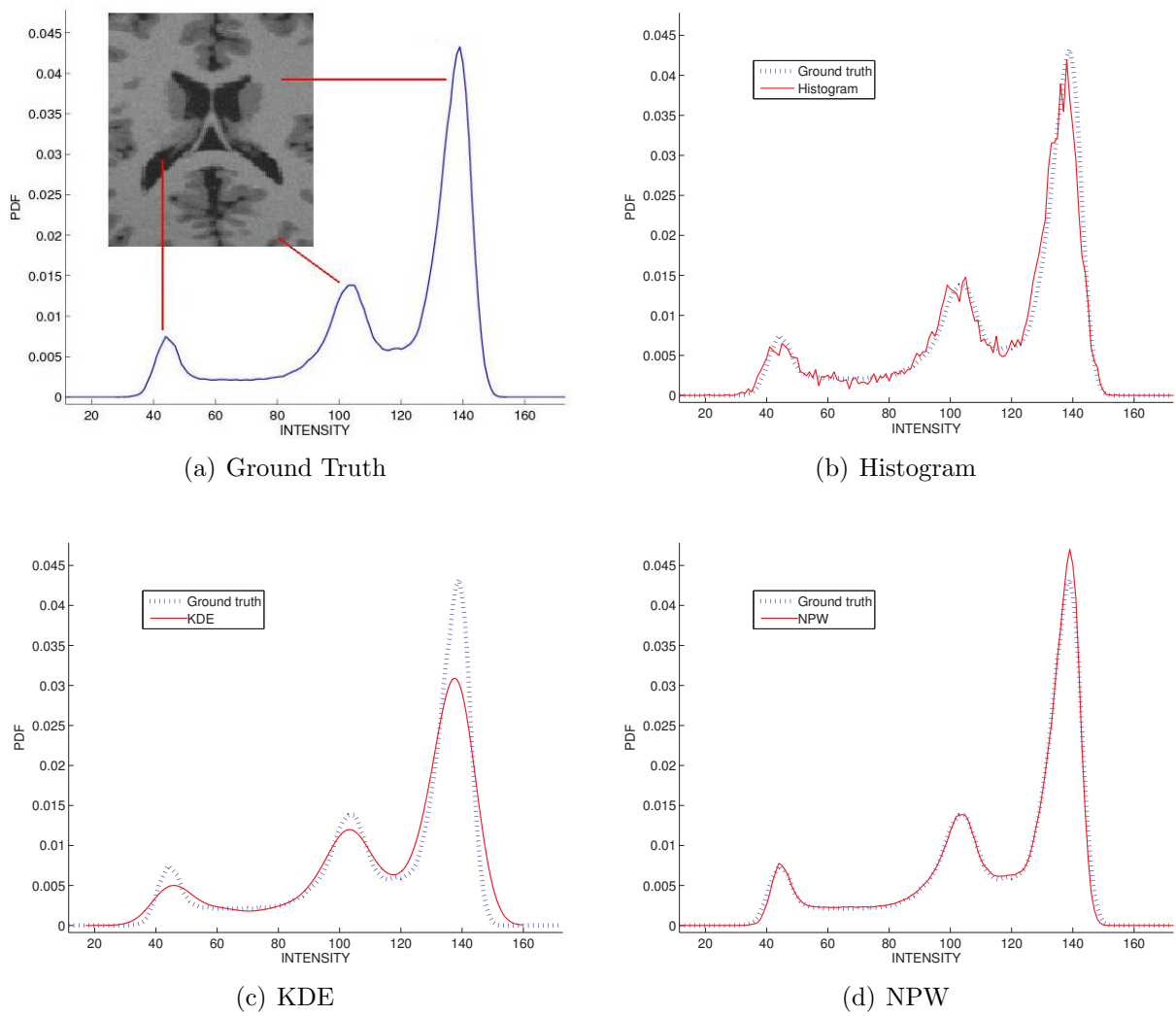


Figure 3.3: Non-parametric probability estimation for a T2-weighted brain MR section. Note the clarity of the separation of key tissue classes in this case, showing that it is a relatively easy segmentation problem, attainable by simple thresholding.

These estimates can then be used to formulate computer-aided segmentation algorithms. The key issue is the accuracy of the PDF estimation, since it directly affects the subsequent segmentation step. Evidently, there is a trade-off between having sufficient data to improve the accuracy of PDF estimation and minimising the amount of manual segmentation required. To achieve this, we perform manual segmentation on just three to five of the 60 to 90 image slices for the regions of interests (ROI); in this case, the classes of interests are: tumour; aerated lung; fluid; liver; and mediastinum.

A mathematical morphological erosion operation is first applied to the manual segmentation data to reduce the partial volume effect on the tissue boundaries. The NPW estimator is then applied to individual tissue classes in an eroded 2D image slice in each volume, as defined by the manual segmentation by the clinician, or prior information, to produce ‘guided’ estimates. The image slices are chosen so that equivalent thoracic sections are studied. This is a reasonable assumption since a 3D volume would include slices closer to the neck and diaphragm where the uneven tissue boundaries, coupled with image artefacts such as PVE, deteriorates the essential medical implications that can be derived from the PDF estimates. Estimation results are shown in Fig. 3.4. To evaluate the performance of NPW, we use the L_2 norm defined by $L_2 = \sqrt{\sum_i (u_{His}(i) - u_{NPW}(i))^2}$ where $u_{His}(i)$ and $u_{NPW}(i)$ are the histograms and NPW estimations, respectively. It has been found that the norm is of the order 10^{-4} over the range shown. Compared to an average data value of the order 10^{-3} , we conclude that the error norm between the histograms and NPW estimations is relatively small and falls within the tolerance range. This justifies the accuracy and use of NPW for PDF estimation in MPM studies.

Results

Comparison of NPW with histograms estimation for an arbitrarily chosen thoracic CT image slice is given in Fig. 3.4. Note that NPW estimator offers the advantage of producing smoother estimates. This improves the interpretability of the estimated PDF. Incorporating the manual segmentation data, we can estimate the PDFs for three notable tissue classes, also shown in Fig. 3.4. Note that all of these tissues correspond to a single ‘peak’ in the ‘unguided’ PDF estimation. This illustrates the difficulties in segmenting the tissues described earlier. This

issue is solved with a ‘guided’ estimation, that is, with the use of manual segmentation data; here each of these tissue classes is represented separately by an independent ‘peak’.

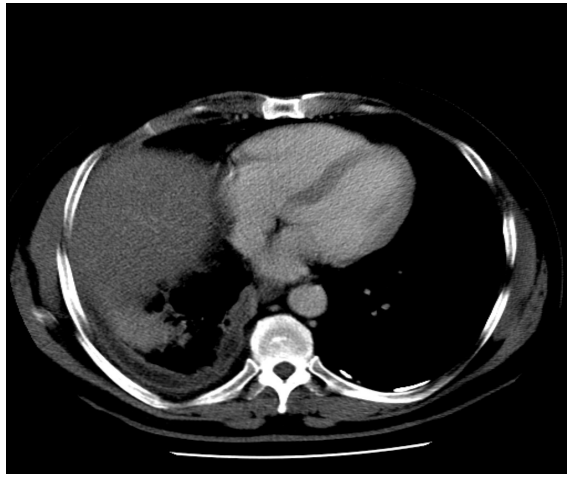
Discussion

Results from the CT images volumes from two mesothelioma-diagnosed patients are shown in Figs. 3.5 and 3.6. Assessment of treatment response is performed on each of these patients. The aim is to evaluate the feasibility of a computer-aided segmentation based on these PDF results. Four such studies are carried out on patient 1, one baseline scan plus three other scans taken after every two cycles of treatment. A comparison of the tumour PDF at different stages for patient 1 is given in Fig. 3.6(e).

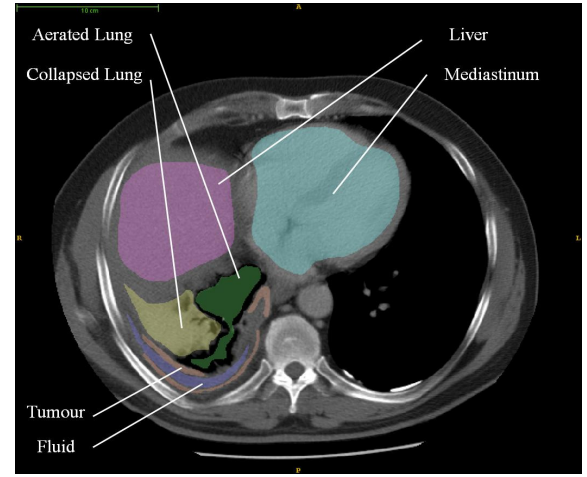
Before proceeding onto applying NPW to the actual segmentation problem, we make a few observations connecting the PDF findings to the medical understanding of MPM and show how these results help in verifying some of our clinical assumptions.

Anatomically, the thoracic tissues are well reflected by their PDFs. In particular, liver is always represented by a narrow smooth distribution largely separated from other body tissues. Therefore, we have omitted them from our analysis in Fig. 3.6. Mediastinum, with its myriad of cardiovascular structures, is represented by a much wider spread on the probabilistic spectrum. This sometimes overlaps with that of the tumour, though in most cases has distinct peaks well away from the tumour PDF. Therefore, mediastinum would not impact greatly on the tumour segmentation problem, because the overlap can be viewed as equivalent to a low level signal noise. On the other hand, distinguishing tumour from the pleural fluid is not a simple task. It should be noted that fluid is not always present in patients (such as its absence from Patient I). However, when they do appear the segmentation problem is made much harder because of a multitude of reasons including their location - usually engulfed or partially surrounded by the tumour and their similarity in radiation attenuation to the tumour. In terms of the PDF, a dominant overlap of the fluid with the tumour verifies this fact (Figs. 3.5 (e-f) and 3.6 (a-d)).

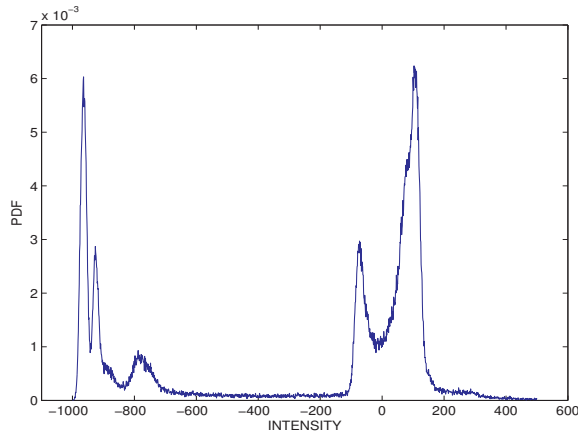
It has been observed that the measured thoracic organ tissues (liver and mediastinum) are largely unchanged in different scans. Similarly, the relative stability of the tumour may be accounted for by the fact that IV-Vinflunine treatments have been palliative and aimed at



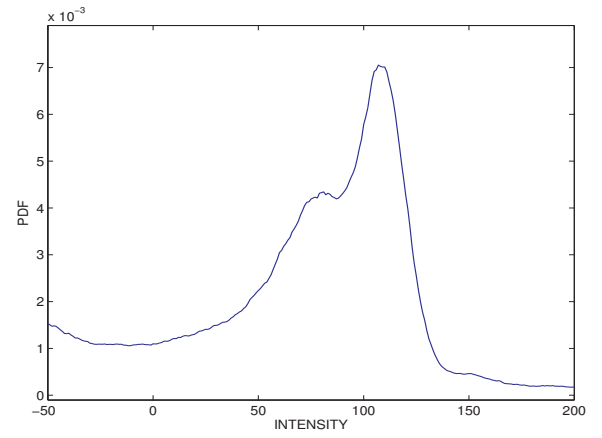
(a) Image data



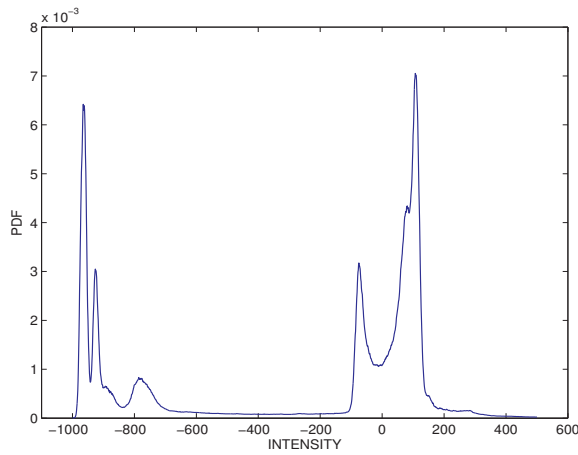
(b) Image Data with Clinical Segmentation



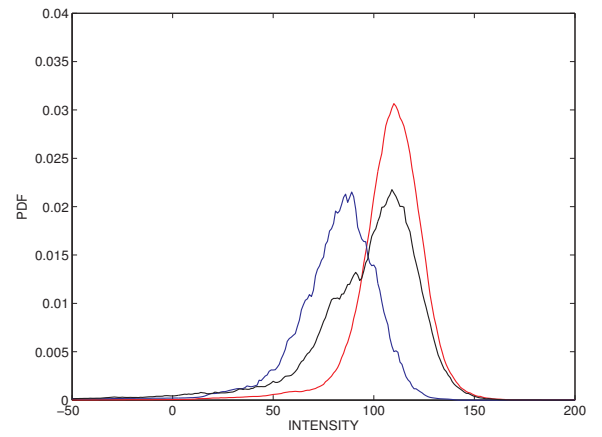
(c) Histogram



(d) Overall PDF (magnified)

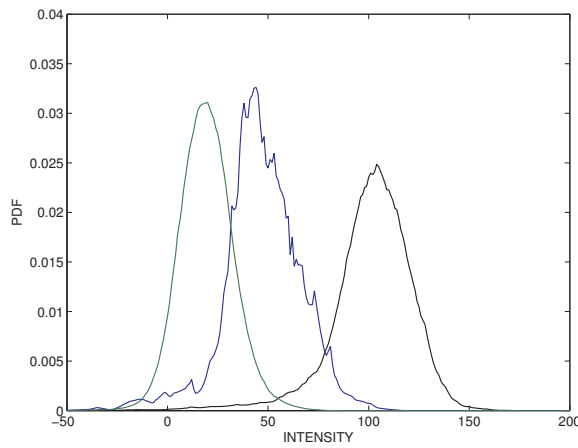


(e) NPW

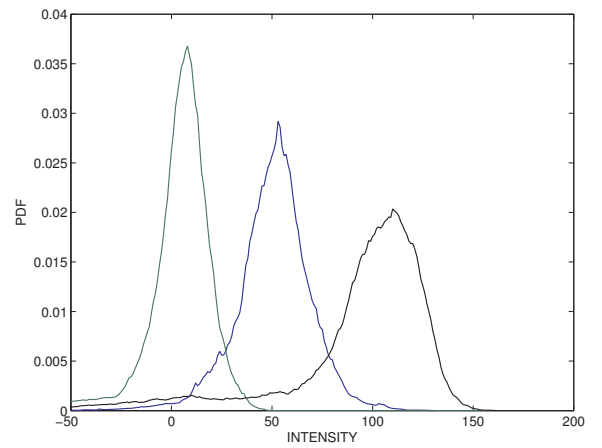


(f) Segmented Tissue PDF

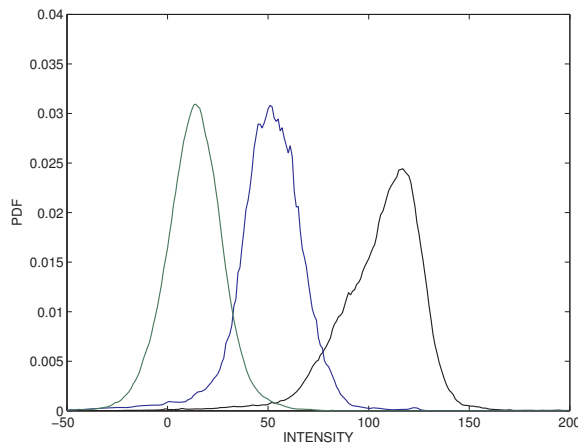
Figure 3.4: Estimated overall PDF of a typical MPM scan slice and ‘guided’ PDF estimation for key tissue classes found in its intensity range, namely, tumour (*blue, approx. $37 \times 37px$*), fluid (*red, approx. $62 \times 62px$*) and mediastinum (*black, approx. $137 \times 137px$*). Given the small number of data points available for each tissue class, NPW performs well with good smoothness and accuracy.



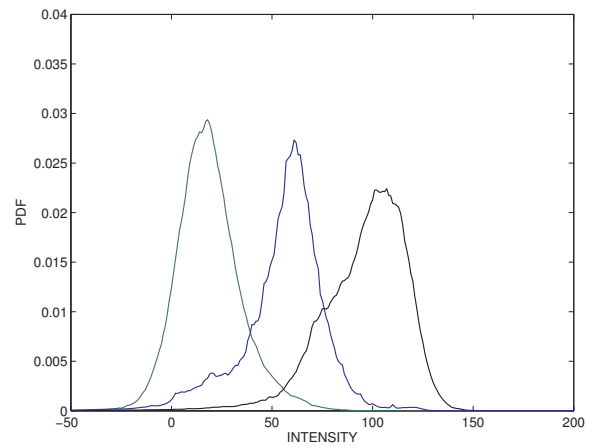
(a) Patient 1-I



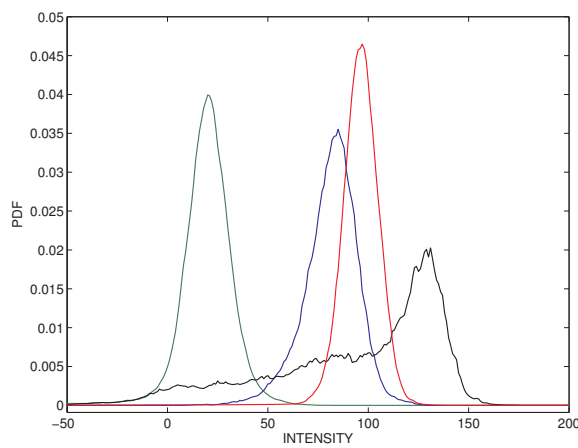
(b) Patient 1-II



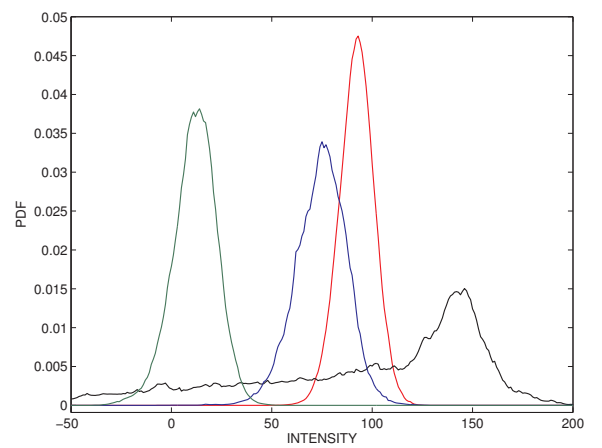
(c) Patient 1-III



(d) Patient 1-IV



(e) Patient 2-I



(f) Patient 2-II

Figure 3.5: 2-D PDF response assessment for MPM patients 1 and 2. Shown here are tumour (*blue*), fluid (*red*), liver (*green*) and mediastinum (*black*).

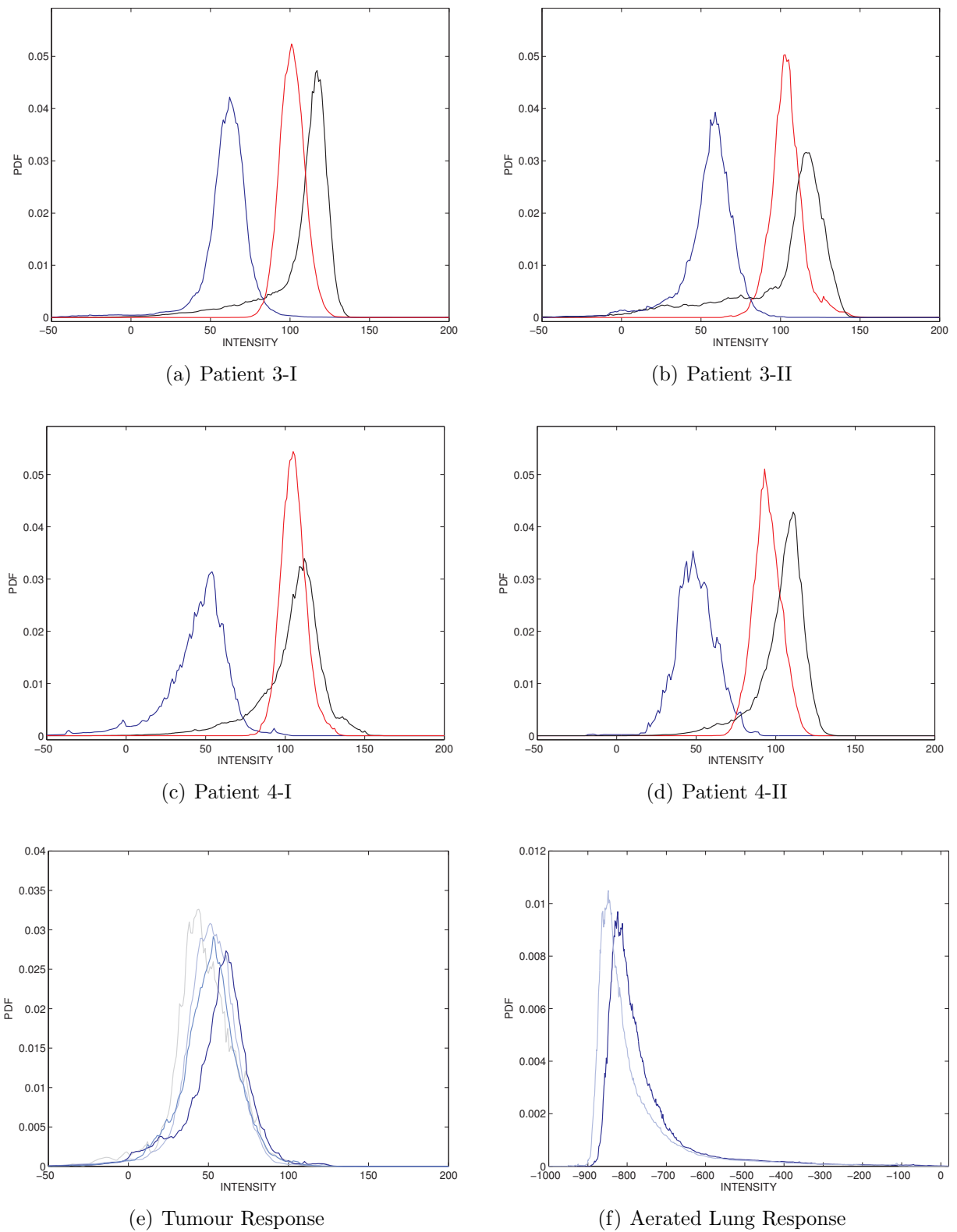


Figure 3.6: 2-D PDF response assessment for MPM patients 3 and 4 and tumour and aerated lung progression during treatment. Shown here are tumour (*blue*), fluid (*red*), and mediastinum (*black*). The tumour and aerated lung responses for Patient 1 are represented by (*light – dark blue*), in the order from early to later stages of the treatment.

slowing progression of the disease, rather than achieving significant regression. We, however, observe a trend of increase in tumour intensity over the trials, which might support a possible occurrence of lesion hardening. The response of the aerated lung to the drug trials is another indicator of the patient's quality of life. A reduced aerated lung volume might clinically imply worsening of dyspnoea. This idea is explored further in Chapter 7.

One final observation is that the uneven tissue PDF shapes reinforce our non-Gaussian assumption and further justify our use of non-parametric methods for probabilistic estimations in our study. In general, the overlap of tissue intensity distributions suggests that intensity thresholding alone would not produce good segmentation results for such MPM scans.

3.4.2 Segmenting Mesothelioma

We observe that in some cases, the PDFs show a good distinction between individual tissue classes. Therefore we are able to segment the images based on the image intensity values and manual segmentation data. In Fig. 3.7 we show some initial experiments for semi-automatic segmentation. Here we have used a Bhattacharya flow-driven level sets segmentation, which is described in greater detail in Chapter 4. From visual inspection it can be noted that the PDF estimates provide a good basis for the segmenting operation, as they clearly separated the tissues of interest in each of the two cases. To justify its accuracy, we need to assess the results numerically by registering our segmented classes to the reference truth, in case the clinically segmented image mask. We do this by comparing the PDFs of the two images, which are shown in Fig. 3.7(d) and (g).

As before, we quantify the similarity of the two PDFs using the L_2 norm. We find that in this case, it is in the order 10^{-1} . This shows the potential of using NPW to segment mesothelioma, although it may be necessary to incorporate additional information from the image. This is especially true if we aim to alleviate the need for manual segmentation and develop a more automated segmentation algorithm. In addition, we have only attempted to segment the simplest cases. For most MPM patients, the segmentation problem is generally not as straightforward, as we will show in Section 4.4. One obvious approach is to extend our prior

knowledge beyond scalar(intensity value)-based segmentation to segment these tissues more generically. Furthermore, this would equip us with a useful tool for estimating joint probabilistic distributions of two non-independent quantities in a number of imaging applications, such as calculating mutual information [18] and entropy.

3.5 Extending NPW to Vector-valued Signals

So far in this chapter, we have made observations on PDFs, and shown that PDF-based segmentation for MPM is feasible, as supported by the semi-automatic segmentation results. We however have concentrated on less complex scans with relatively small areas of consolidated lung. For most MPM scans, due to the overlapping nature of tissue PDFs, image intensity alone might not be sufficient to give accurate and reliable segmentations. This point is demonstrated in Section 4.4.

For this reason, we investigated further the application of the NPW estimator for automatic segmentation. A good starting point is to examine ways in which the manual segmentations are conventionally accomplished. Our study has shown that in addition to image pixel intensities, texture; tissue heterogeneity; and knowledge of normal anatomy may help in identifying a tissue's presence. To incorporate these quantities into our segmentation routine we need to develop a probability density estimation method for vector-valued image signals. Image texture is mostly image technique-dependent and is hard to accurately quantify. Tissue heterogeneity can, for example, be measured by information-theoretic measures such as entropy $H = -\sum_i P(i) \log P(i)$ where $P(i)$ is the probability at value i . Higher entropy values usually imply a more heterogeneous intensity distribution and vice versa. For assessing MPM response to treatment, monitoring the heterogeneity change is another area of good clinical interests. These additional measures might provide good insights into developing a better segmentation algorithm. The aforementioned points prompt our research on vector-valued NPW estimation for the joint PDF representation of more than one variable.

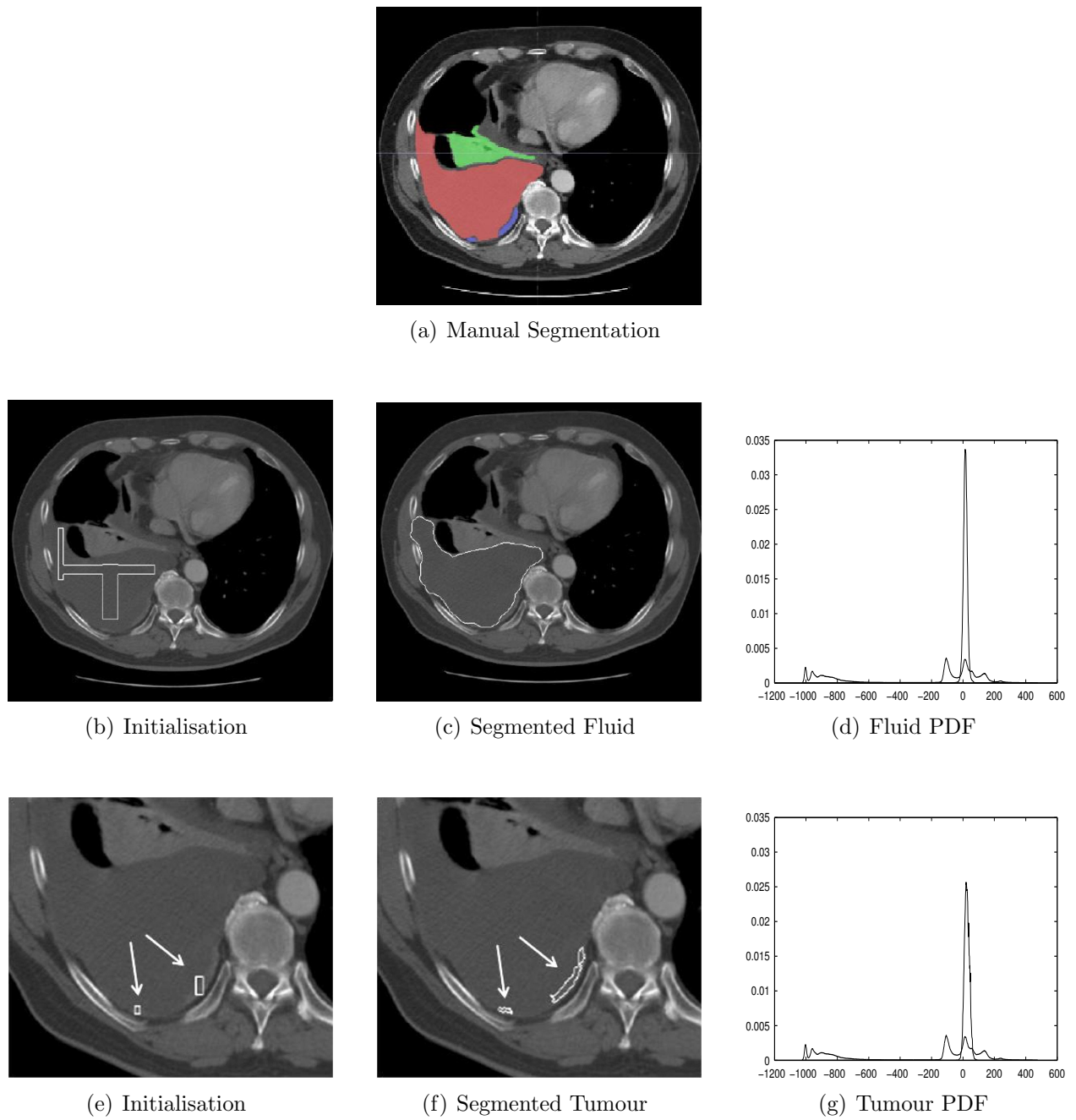


Figure 3.7: Preliminary segmentation using PDF estimates from Fig. 3.5d). Tumour, fluid and collapsed lung are denoted in *blue*, *red*, and *green*, respectively. The PDFs show the comparison of the tissue, P_{in} (*the peak*) against overall PDF, P_{out} (*the base curve*). Note this is a relatively simple case of MPM. A more thorough analysis is required for the general MPM cases.

3.5.1 Joint PDF for a Pair of 1D Signals

We begin with a 2-tuple vector $F_{y_1, y_2}(x)$, where for each x there are two associated quantities. This can be a combination of any two arbitrary pieces of information, y_1 and y_2 , given by an image sample. For instance, in a colourmap image, they can be the red and green components of the RGB triplet defining the colour of a pixel or the x- and y- components of a motion flow field. Alternatively, in our case, they can be the intensity and a texture measure in a greyscale CT scan. As before, we find the assumed form of these two distribution by fitting piecewise linear functions. For simplicity, we initially use the linear function $y = ax + b$. We have $y_1 = a_1x + b_1$ and $y_2 = a_2x + b_2$, giving two sets of parameters: (a_1, b_1) and (a_2, b_2) . In vector notation, this becomes:

$$\vec{y} = \vec{a}x + \vec{b} \quad (3.32)$$

$$\text{where } \vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \vec{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \text{ and } \vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \text{ for } 0 \leq x \leq 1.$$

Now, similar to the derivation of 1D scalar NPW, we assume a uniform distribution for $x : F_x(x)$.

$$x_i = \frac{\vec{y} - \vec{b}}{a_i} : F_x(x) = 1; \quad i = \{1, 2\}; \quad (3.33)$$

The joint distribution, $F_{y_1, y_2}(x)$ or $F_{\vec{y}}(\vec{y})$, is then given by:

$$F_{\vec{y}}(\vec{y}) = \frac{1}{|d\vec{y}/dx|} F_x(x) = \frac{1}{|d\vec{y}/dx|} F_x\left(\frac{\vec{y} - \vec{b}}{\vec{a}}\right) \quad (3.34)$$

$$\text{such that } \left[\frac{d\vec{y}}{dx}\right] = \begin{bmatrix} \frac{dy_1}{dx} \\ \frac{dy_2}{dx} \end{bmatrix} \text{ for } b_1 \leq y_1 \leq a_1 + b_1, b_2 \leq y_2 \leq a_2 + b_2.$$

In this case, the modulus is the diagonal length of a right triangle formed by a_1 and a_2 , so,

$$F_{\vec{y}}(\vec{y}) = \frac{1}{\sqrt{a_1^2 + a_2^2}} F_x\left(\frac{\vec{y} - \vec{b}}{\vec{a}}\right) = \frac{1}{\sqrt{a_1^2 + a_2^2}} \quad (3.35)$$

Therefore, the 1-D NPW estimate for a 2-tuple vector is found as:

$$F_{\vec{y}}(\vec{y}) = \begin{cases} \frac{1}{\sqrt{a_1^2 + a_2^2}} & \text{for region A and } \vec{a} \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad (3.36a)$$

$$(3.36b)$$

Note that we still need to determine the boundaries of region A. To this end, we first extend the histograms estimator to the 2-tuple vector space, to predict our PDF results, and to give insights for determining region A. Also note that each point (y_1, y_2) on the base plane formed by Y_1 and Y_2 has a joint probability, which gives the height to the PDF at that point; this is defined as:

$$P(Y_1 = y_1, Y_2 = y_2) = P(Y_1 = y_1 | Y_2 = y_2)P(Y_2 = y_2) = P(Y_2 = y_2 | Y_1 = y_1)P(Y_1 = y_1) \quad (3.37)$$

where $P(x|y)$ is the conditional probability of x , given the occurrence of y . The right hand side of the equation follows from Bayes' Theorem. With this derivation, we implement a histograms estimator for 2-tuple vector signals, whose pseudo-code is given below (Algorithm 3.4).

3.5.2 Results and Discussions

In this case, we make use of the Matlab® function for finding the histograms for scalar-valued samples, *hist*².

²documentation available from <http://www.mathworks.com>

Algorithm 3.4: Pseudo-code of Histogram Implementation for 2-tuple Vectors

Individually find bin centre coordinates of the y_1 and y_2 using command hist

for every bin in y_2

 assign [lower bound, upper bound] for each bin as

 [average of current last bin centre, average of current and next bin centre]

 for the two end cases, assign the bounds to be -inf and inf, respectively.

 find the overall histogram by

 hist($y_1((y_2 \geq \text{lower bound}) \ \&\& \ (y_2 < \text{upper bound}))$), y_1 bin center coordinates)

end

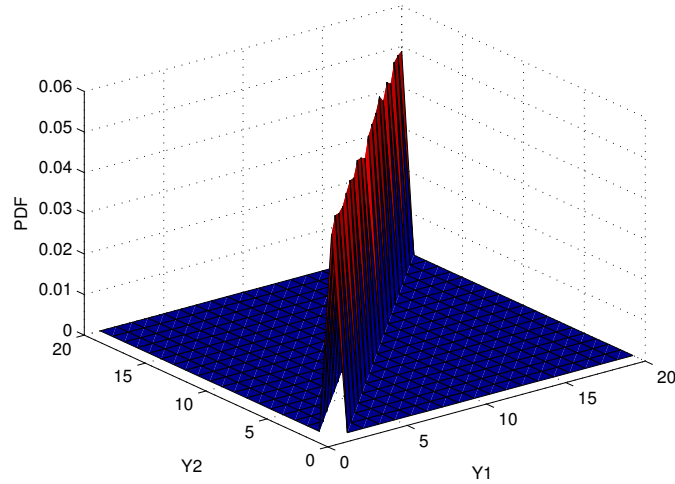


Figure 3.8: Histogram Estimation for 2-tuple Vector

Note that the volume shown in Fig. 3.8 is projected onto the base y_1, y_2 plane as a line. This suggests that A is actually the diagonal line crossing the highlighted region shown in Fig. 3.9. More specifically, NPW boundaries A can be written analytically, as given in Table 3.1.

To validate our implementation of the 1-D 2-tuple NPW, we estimate the averaged joint distributions of two-colour channels (red and green) in an angiographically-marked CT reconstruction of the lung (used for diagnosing emphysema, a lung disease characterised by abnormal enlargement of airspaces distal to terminal bronchioles, shown in Fig. 3.10). The purpose is to

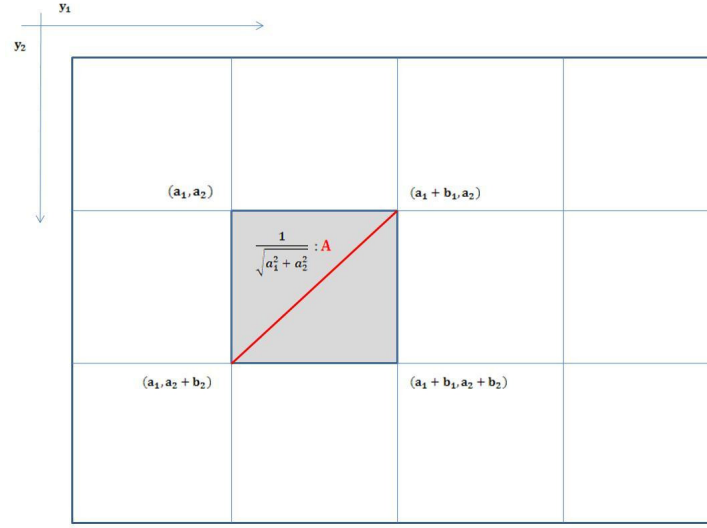


Figure 3.9: Illustration of the NPW estimator for a 2-tuple vector, range A is a diagonal crossing the region highlighted in grey. Shown here is one of the seven possible cases where $a_1, a_2 > 0$

Conditions	Case
$a_1, a_2 > 0$ or $a_1, a_2 < 0$	$\frac{a_2}{a_1}y_1 + b_2 - \frac{a_2}{a_1}b_1 = y_2$
$a_1 > 0, a_2 < 0$	$-\frac{a_2}{a_1}y_1 + a_2 + b_2 + \frac{b_1 a_2}{a_1} = y_2$
$a_1 < 0, a_2 > 0$	$-\frac{a_2}{a_1}y_1 + a_2 + b_2 + \frac{a_2 b_1}{a_1} = y_2$
$a_2 = 0, a_1 \neq 0$	$y_2 = b_2$
$a_1 = 0, a_2 \neq 0$	$y_1 = b_1$
$a_1, a_2 = 0$	a point at (b_1, b_2)

Table 3.1: Specifying boundaries for the 2-tuple NPW

assess the functionality of our extension by comparing the results to the ground truth, which in this case, is the 1-D 2-tuple histograms estimator.

In both experiments, 2-tuple NPW shows a consistent high level of accuracy and good computational efficiency as compared to the histograms estimator. The smoothing effect of NPW over the histograms can also be clearly observed. We observe the two peaks in Fig. 3.10, which correspond to the two dominant colours in the scan.

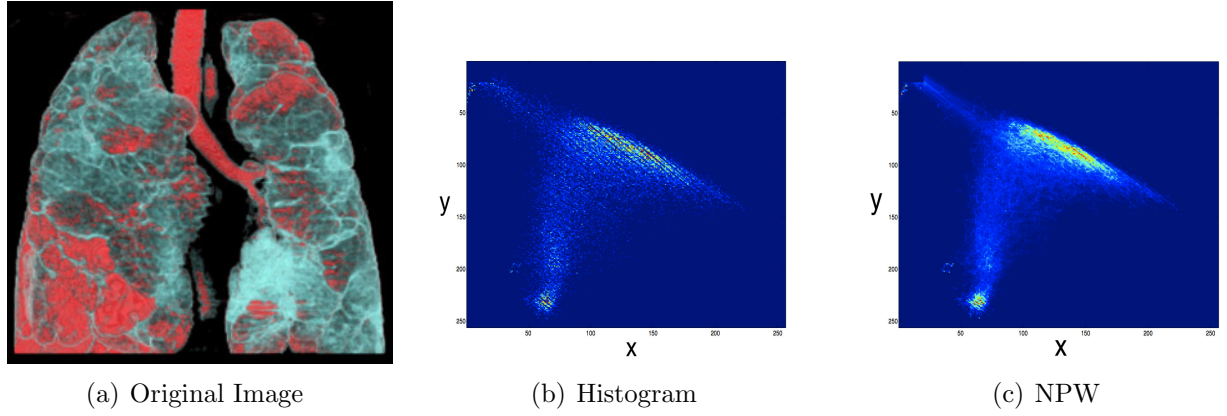


Figure 3.10: Lung CT for diagnosing emphysema, performed at the same time as coronary artery CT, giving values for the red and green channels of the artificially dyed image, with red showing lung damage; b) and c) show the peak compositions in these channels that make up the dominant colours in the scan.

Time-Hist(s)	Time-NPW (s)	L_2 norm
0.008395	0.008396	7.68e-3

Table 3.2: Accuracy and computational time of NPW in estimating joint PDF; accuracy is assessed by an L_2 norm, taking the histograms estimate as reference truth.

3.5.3 Joint PDF Estimation for 2D Signals

Having extended NPW to finding joint distributions of 1D vectors, we may now proceed with similar steps as given in Section 3.3.2 for the case of 2D vector signals. For simplicity, we work with the planar approximation assumption. Assigning equations to link the variables y and x : $y_1 = a_1x_1 + b_1x_2 + c_1$ and $y_2 = a_2x_1 + b_2x_2 + c_2$. In vector notation, this would become:

$$\vec{y} = \vec{A}\vec{x} + \vec{c} \quad (3.38)$$

$$\text{where } \vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \vec{A} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \text{ and } \vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ and } \vec{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \text{ for } 0 \leq x_1, x_2 \leq 1.$$

By substitution and rearranging the vector equation, the inverse functions are found to be:

$$x_1 = \frac{c_1y_2 - y_1c_2 - b_2y_1 + b_1y_2}{b_1a_2 - a_1b_2}, \quad x_2 = \frac{c_1a_2 - a_1c_2 - a_2y_1 + a_1y_2}{-b_1a_2 + a_1b_2}. \quad (3.39)$$

which yield a Jacobian, as follows,

$$|J| = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} \frac{-b_2}{a_1 b_2 - b_1 a_2} & \frac{b_1}{a_1 b_2 - b_1 a_2} \\ \frac{-a_2}{a_1 b_2 - b_1 a_2} & \frac{a_1}{a_1 b_2 - b_1 a_2} \end{vmatrix} = \frac{1}{|a_1 b_2 - b_1 a_2|} \quad (3.40)$$

And the overall joint probabilistic distribution can be calculated as:

$$F_{\vec{y}}(\vec{y}) = \begin{cases} \frac{1}{|a_1 b_2 - b_1 a_2|} & \text{for region A and } a_1 b_2 \neq b_1 a_1 \\ 0 & \text{otherwise} \end{cases} \quad (3.41a)$$

$$(3.41b)$$

The remainder of the derivation is left as future work, as it is not directly linked to our development of an algorithm for MPM segmentation, for reasons that are explained in the next section.

3.6 Summary and Future Works

In this chapter, we reviewed the relevant methods for probability density estimation. We also discussed them in relation to our research goal and worked on establishing an essential probabilistic tool for our project. To highlight key developments, we described and demonstrated the advantages offered by NPW for estimating PDFs over all existing non-parametric methods such as the histograms and KDE. We also implemented the NPW estimator for scalar-valued samples for the 1D, 2D and 3D cases, applied them to CT thoracic scans, with and without manual segmentation. The results showed the possibility to develop a fully automatic and accurate segmentation method using probability density estimations. However, due to the similarities in CT attenuation of soft tissues in CT scans, intensity values alone might not be sufficient for our segmentation task. In response to this we worked on furthering our understanding of PDF estimation by extending the idea of NPW to vector-valued data. This opened up the possibility to explore other features and information supplied by the CT scans, such as textures and entropy; or any other image features additional to the pixel intensity, that are commonly taken

into consideration during the clinical manual segmentation procedure.

To this end, we derived and implemented the theories of NPW estimation for 1D 2-tuple vector signals. The immediate next step is the extension and implementation of NPW to estimating joint distributions of 2D signals. This would enable us to apply the vector NPW method to a wider range of applications. This includes a good use of the theories in the field of multi-modal registration where both image intensity and entropy are involved. Additionally, it is possible as a future work to apply the method to estimate the joint distributions of image intensities with other key image quantities such as texture and entropy. Image texture is mostly image technique-dependent and is hard to accurately quantify. Tissue heterogeneity can, for example, be measured by information-theoretic entropy $H = -\sum_i P(i) \log P(i)$ where $P(i)$ is the probability at value i . Higher entropy values suggest a more heterogeneous intensity distribution and vice versa. Due to the fact that our initially proposed NPW-level-sets image segmentation had failed to present sufficient robustness to MPM scans, as we will show in Section 4.4, this idea will not be pursued further in this thesis. However, this extension can still be viewed as an important future work that is likely to benefit the field of medical vision in many ways. We discuss these possibilities in Chapter 8.

Having developed the probabilistic analytical tools for our application, we, in the next chapter, review some notable segmentation techniques and discuss potential candidate methods for the segmentation of MPM. As we will show later in this thesis, NPW as a probability estimation method sees application beyond those already presented in this chapter. For instance, in our development of a boundary regulariser for the random walk-based method in Chapter 5, it is used to compute the PDF priors which in turn provides a good way to achieve good boundary smoothness in the segmentation. It can also be applied to evaluate probabilistic distributions of pure and boundary voxels which enables partial volume correction. In sum, probability density estimation forms an integral part of many segmentation algorithms, in which cases NPW may prove to be highly useful.

Chapter 4

Computational Methods for MPM Segmentation

The first stage in assessing disease burden in most forms of cancer is the segmentation of the tumour from surrounding tissues. In this chapter, we introduce and briefly discuss a number of methods for medical image segmentation. Particular attention is paid to their applicability to our clinical problem, that is, the measurement of malignant pleural mesothelioma (MPM) from CT scans.

4.1 Introduction

Image segmentation is defined as the problem of partitioning an image into non-overlapping regions with respect to some features such as image intensity or texture. Mathematically, for a domain Ω_I representing the image I , a K -class segmentation can be specified as the problem of finding K regions $S_k \in \Omega_I$ such that [56]:

$$\Omega_I = \bigcup_{k=1}^K (S_k) \tag{4.1}$$

where $S_k \cap S_j = \emptyset$ for $k \neq j$. The segmentation task is to find these sets so that they distinctly and exclusively represent the tissues of interest.

For the purpose of our study, a two-class segmentation (i.e $K = 2$) is considered. The dimensionality of a segmentation method refers to whether it is 2D (planar) or 3D (volumetric). Ultimately, for our application, we desire a 3D volumetric method instead of a planar one that is applied sequentially to a series of image slices. This is because, given the intrinsic difficulties associated with MPM segmentation, as presented in the last chapter, a 3D segmentation is more likely to produce an accurate result given the additional information available from neighbouring image slices. However, this would make the implementation more complex and increases the computational burden; the latter in particular affects the long-term functionality of our algorithm.

Another consideration for all segmentation routines is the level of user supervision that they require. Manual intervention from the user can significantly improve the performance of a method, essentially by introducing prior information. For medical images, with their substantial and often non-parametric image noise and prevalence of image artefacts, the importance of user supervision is often key: user interaction almost always play a key role in ‘initialising’ the segmentation to produce clinically useful results. Because initialisations generally require a laborious process of pre-defining anatomical structures, ideally, a less onerous supervised scheme is preferred. Therefore, a trade-off between the amount of required user interaction and the segmentation reliability and accuracy must be sought. The general aim is to maximise the effectiveness of the segmentation method whilst ensuring its good clinical robustness and performance.

Over the years, a series of computer-assisted segmentation algorithms have been proposed for use in medical image analysis. They can be broadly categorised according to their mechanisms of operation: 1) Thresholding, 2) Feature space classifiers, 3) Clustering, 4) Markov Random Field (MRF) models, 5) Graph-based methods, 6) Level sets, and 7) Atlas-guided approaches.

In this chapter, we begin our discussion in Section 4.2 with a literature review of related works on MPM segmentation and demonstrate the need for an improved segmentation method. In Section 4.3, we overview a number of image segmentation methods and discuss their applicability to MPM. Particular interests are given to a level set-based approach, namely the Bhat-

tacharya distance-driven level sets method [32], in Section 4.4. We conclude our discussions in Section 4.5 with a summary of the chapter and motivations leading to our choice of random walk (RW) as the backbone of our main method.

4.2 Related Work

There have been a number of previous attempts to adapt and apply tools in medical image analysis to segment MPM on CT image scans.

A thresholding-based segmentation scheme is presented in Zhao et al. [91]. In this method, the volume containing the thorax is first masked from the CT scans using interpolation from the rib bones, which appear distinctively in the image. This allows a masking of the lung area where tumour lesions are to be found. The actual segmentation follows separating three key tissue classes: 1) the mediastinum, 2) tumour 3) lung parenchyma and blood vessels. Intensity histograms are studied for observable ‘peaks’ whose ranges are taken for calculating the controlling thresholds and used for tissue class segmentation. Despite showing some fair results, the method is a very rudimentary approach to the problem. It requires a highly refined mask where there are three distributions distinctly separated on the probability density function plot. In fact, this is a weakness common to many thresholding-based methods. Even the slightest peculiarity in the patient data, such as rib movement and tumour propagation, may affect the method’s performance.

A further tumour quantification method is described in Chaisaowong et al. [15], where the lung volume containing the tumour is first segmented by thresholding followed by the finding of contours using an adjacency tree of connected components [15]. The largest contour is taken to be the pleura. A healthy pleura is modelled as being convex shaped and used to outline ‘concave irregularities’ on the outer surface of the contour, which are treated as ‘candidate sites’ of pleural thickening. These sites are evaluated by a set of thresholding and geometric criteria to give the final tumour estimate. Note that in addition to being prone to the usual weakness of thresholding, this method makes numerous assumptions during contour selection, which greatly affect its robustness.

The works of Armato et al. [2] [3] presented an indirect approach to the problem. Instead of directly assessing the volume of the lung, it first masks the lung volume by thresholding. By giving a user-defined point on the pleura, tumour thickness is measured by and treated as a measure of tumour size. The method was tested on a group of 44 thoracic CT scans from 22 MPM diagnosed patients and reported an approval rate of 86% by clinical observers. Note, however, this method is closer to being a computer-assisted way of computing RECIST instead of presenting a new quantification method of its own. Therefore, it is subject to a fair amount of user bias and suffers from issues like reproducibility and robustness.

A later study, Ak et al. [1], presented a volumetric measurement method for assessing the tumour response to chemotherapy on a group of 55 MPM diagnosed patients. The method is based on the Cavalieri principle of stereology, which is a grid point counting system-based algorithm that computes the tumour volume from a large number of evenly distributed grid points. The system is governed by the formula:

$$V = t \times [SU/SL \times d]^2 \times \sum P \quad (4.2)$$

where V is the tumour volume, t the axial thickness of the scan, SU the measurement made on the scan, SL the scale of the scan to real object, d the distance between the grid points and $\sum P$ the total number of grid points that land in the tumour, as counted by the clinical user. The correlation coefficient of this method with the WHO [45] and the modified RECIST [19] criteria, in terms of their response assessment results, were found to be $CC = 0.64$ and $CC = 0.52$ respectively. Despite being only rough estimations of the actual tumour volume, results from this method demonstrated the possibility of establishing a volumetric method to quantify MPM for assessing its response to treatment. This paved way to our research presented in this thesis.

More recently, Liu et al. [42] have adopted a ‘sequential segmentation strategy’, in which a combination of chest-rib interpolation and a gradient vector flow snake technique was first used to separate the soft tissue region containing the heart, stomach, lung parenchyma and tumour from rest of the image. An estimate of the tumour is then found from this region

using a multiple thresholding method and edited by an experienced radiology specialist. The segmented results are reported to be strongly associated with patient survival, and shown to be more indicative of patient survival than their RECIST measure counterparts [42]. This method, with its final step of manual editing, is heavily affected by observer bias as well as suffering from the usual thresholding-related weakness.

Frauenfelder et al. [21] have suggested the use of volume segmentation for assessing therapy response for MPM. In their work, a commercial segmentation software, Myrian® , is used for finding the MPM volumes. Compared to the high level of inter-observer variation of RECIST derived in the paper, results from the volume segmentation were reported to show good reproducibility. However, details of the software are unknown and no assessments of the segmentation accuracy, such as how the results relate to the clinician-defined manual tumour delineations, nor its validation against clinical data, were given in the paper. Thus, it is difficult for us to evaluate the effectiveness of the software in segmenting MPM and determine whether it would be suited for use with our clinical data.

In summary, responding to the many weaknesses of RECIST in its application to MPM, most works on MPM segmentation to date support the use of segmented tumour volumes to establish an alternative method for assessing patient treatment responses. However, despite the numerous efforts, we are yet to see an efficient, automatic (subject to minimal required user interaction) 3D segmentation algorithm, that is designed specifically for use on MPM. In particular, many key issues such as the partial volume effect, geometric complexity of the tumour shape, and similar CT attenuation of the tumour and neighbouring soft tissues, all need to be addressed and accounted for in the design and optimisation of a potential segmentation algorithm.

For the above reasons, we have developed a segmentation algorithm addressing these MPM-specific difficulties. In the next few sections, we overview a number of notable image segmentation techniques in medical vision and discuss their potentials to be adapted for our need.

4.3 Overview of Segmentation Methods in Medical Vision

In this section, we overview some of the more prominent segmentation methods used in medical vision. Note that most of these methods are seldomly used alone. Established segmentation routines are often based on ideas drawn from two or more of the general segmentation methods given below.

Thresholding

Thresholding-based methods are by far the simplest conceptually. Image scans are partitioned according to a fixed set of criteria known as thresholds which are pre-defined ranges of the image intensity values. For mesothelioma, the geometric complexity of the imaged region and tumour, the intensity value similarity of neighbouring tissues, and image noise generally rule out the use of thresholding as an adequate mean of segmentation. We have already observed this from our probabilistic density findings in Chapter 3.

Feature Space Classifiers

Feature space classifiers partition the target image by operating on some feature space, which can be as simple as a plot of its overall probability density function (PDF) of the pixel intensity or in the 2D case, a plot of two independent scalar image properties such as the T1 and T2 values of MRI [56]. User-defined training data are required as references for segmenting new images. In the example of nearest-neighbour classifier (NNC), pixels are classified as belonging to the same class as the training datum with the closest image intensity [56]. A k-nearest-neighbour classifier (KNNC) provides a non-parametric generalisation to this by classifying a pixel to the same class as the majority of the k-closest training data [39]. The non-parametric property is especially important for our application, for reasons already given in Section 3.2. Classifiers are also relatively efficient, given their non-iterative nature. The main weakness of classifier methods is that they do not perform any spatial modelling, that is, they do not take

into account local image variations such as noise and inhomogeneities. Moreover, the use of a single training dataset for a large number of scans can cause user bias, because anatomical differences between image scans are not considered.

Clustering

The main advantage of clustering over feature space classifiers is that they do not require a training set [56]. However, they do require initialisation. One example is the K-means algorithm, which starts on a user-defined initialisation partition, and uses an iterative process in which a mean intensity is computed for each class and each pixel is classified as belonging to the class with the closest mean [84]. Like classifiers, clustering-based methods are lacking in spatial modelling and can be liable to inaccuracies caused by image noise and inhomogeneities.

Markov Random Field (MRF) Models

To address the need for spatial modelling, Markov Random Field (MRF) models can be used [90]. MRFs model the local neighbourhood at each pixel. They are used because a pixel is likely to belong to the same class as most of its neighbours. MRFs are particularly useful for modelling intensity inhomogeneities that are important for accurately defining tissue boundaries.

The main challenges of using MRF models include their high computational burden and setting the parameters controlling the spatial interactions, which, if set too high, would result in an overly smooth segmentation that is deficient in structural details. Strength of MRF lies in its local operation; global information such as that relating to the tumour shape, are very weak and implicit. But this is precisely what we need to capture in order to accommodate the low CT contrast between the tumour and surrounding tissues for MPM. This is a main problem preventing the adaptation of MRF to our clinical problem.

Graph-based Methods

In graph-based methods, the image is modelled as a graph consisting of nodes connected by edges, which represent the image pixels. The weights of the edges are assigned according to a method-specific weighting function. One such example is graph cuts [10], where a maximum flow minimum cut (with the least number of edges crossed for the maximum possible edge weights) is found from the source to the sink, segmenting the image as a result. It is shown that this produces a good segmentation in many cases [10]. However, graph cuts sometimes produce small cuts or localised segmentation fragments. We present a generalisation to graph cuts, and a solution to this particular problem, in Chapter 5.

Level Sets

In level sets-based methods, the tissue classes are segmented according to the interaction of ‘internal’ and ‘external forces’ working on some tissue initialisation contours [78]. The internal forces are computed from the pixels within the contour so as to maintain the smoothness of the segmentation. External forces, on the other hand, are derived from the outside pixels, with the aim to drive the segmentation towards the desired tissue boundaries as marked by some detectable image features. One example is active contour [78], where an energy function is defined for the contour, comprising of terms based on the external and internal forces. Iteration stops when this energy function is minimised.

Level set-based methods are notable for their ability to form closed segmentation shapes and to adapt to changes in topology (i.e when two regions split or merge), which improves robustness to noise and spurious edges [32]. The main limitations of level sets are their required initialisation and the setting of a series of parameters for controlling the acting forces.

Atlas-based Methods

Instead of directly segmenting an image, atlas-based methods [56] or registration-assisted segmentation, use image registration to map a pre-existing segmentation onto a new image. The

deformation field linking the two images is first computed and used to transform the segmentation of the first to the second. This can be very useful as it avoids user interaction altogether for the new image segmentation, but is also limited by performance of the image registration and the associated additional computation time. An application of this concept is found in Section 5.6.

Other Methods

Other segmentation methods include the model-fitting method where geometric shapes such as ellipse or parabola are used to fit locations of extracted image features [56]. The watershed algorithm first performs an edge detection, followed by mathematical morphology to divide the image into homogeneous regions. The fragments are then post-processed to merge them according to the tissue classes present [56]. Note that both of these methods are heavily relying on the accurate detection of image features.

Discussion

We now discuss the above segmentation methods in relation to our clinical problem.

First, given the non-parametric nature of the intensity distributions in our segmentation problem, many parametric methods such as Bayes' classifiers [56] and clustering-based Expectation-Maximisation (EM) [56] obviously do not offer a solution to our study because parametric assumptions do not lend well to medical image applications, as explains earlier in Chapter 3. Furthermore, methods such as model-fitting and the watershed require good feature detection, which can be a challenge on its own, given the geometric complexity and intensity range overlaps characteristic of MPM, making features hard to identify and isolate. The latter is also a major problem limiting the use of most global segmentation algorithms with little or no user supervision. To illustrate this, results from Chapter 3 are reproduced in Fig. 4.1, showing the failure of the overall plot to distinctly show individual tissue PDFs. This once again justifies the exclusion of thresholding as an adequate segmentation mechanism.

We conclude that effective segmentation needs to involve at least two considerations: (a)

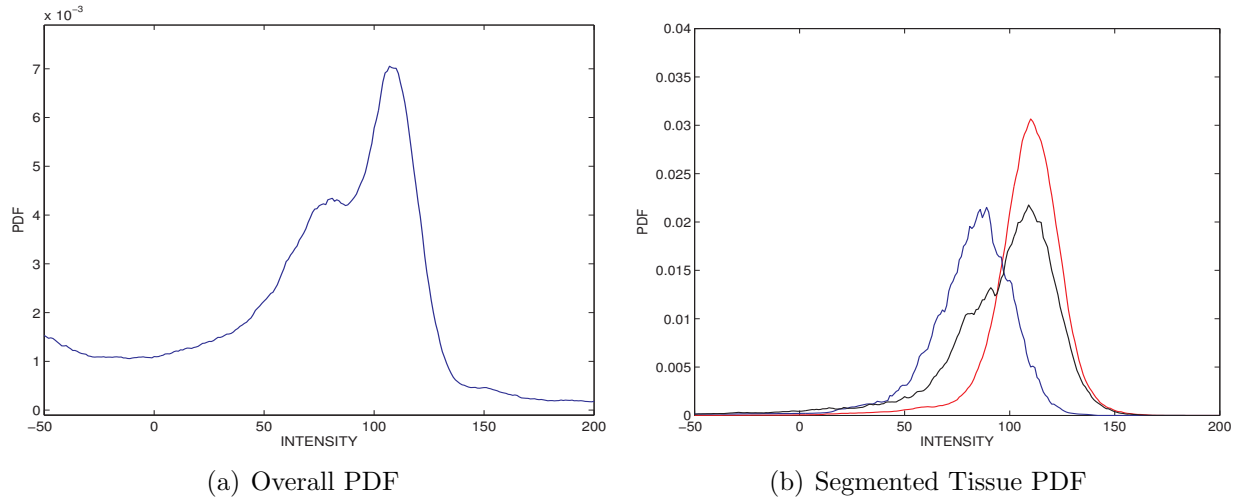


Figure 4.1: Probability density representations of the overall image and individual tissues according to the manual segmentation data. Shown here are the tumour (*blue, solid*), fluid (*red, dot-dash*), collapsed lung (*green, dash*) and mediastinum (*black, dot*).

the tumour forms a closed volume - suggesting the use of level sets; (b) characterisation of the regional properties of key tissues in the region. The second consideration mandates the estimation of a probability density function, so the Non-Parametric Windows estimator (NPW) that we introduced in Chapter 3 becomes important.

Other key considerations for finding a suitable segmentation method include computational requirement, the level of required supervision and robustness.

4.4 NPW-based Level Sets

In this section, we introduce a level set-based segmentation algorithm (NPW-LS) that utilises the idea of NPW for evolving level sets.

4.4.1 Theory

Level set-based methods involve the evolution of a level set function to the boundaries of the segmentation classes in an image. Fundamentally this is a curve evolution and, due to their implicit representation, they can handle changes in topology and do not require parameterisation of the curve [32]. Two approaches are available to evolve such curves; one being forming partial

differential equations according to an underlying physical principle such as temporal motion along the normal direction to the boundaries. Alternatively, curve evolution can be regarded as an energy minimisation problem, from which we solve for the Euler-Lagrange partial differential equation (PDE). Then the PDE is solved iteratively until completion. Statistical properties of the regions to be segmented often improve level set methods and bring robustness to deal with noise and other image properties. Regional statistics can be described using probability density functions. Therefore it is possible to develop a level set algorithm based on probability density estimations.

In a level sets method, a curve C is evolved along the normal direction with the following propagation equation:

$$\frac{\partial C}{\partial t} = F\mathcal{N} \quad (4.3)$$

with t being the time, F the speed function and $\mathcal{N}$ the unit normal vector. This curve is then embedded into a higher dimensional function ϕ , known as the level set function, defined as follows, with Ω being the image domain:

$$C = \{x \in \Omega | \phi(x) = 0\} \quad (4.4)$$

Note that $\phi(C(t), t) = 0$ and $\mathcal{N} = \nabla\phi/|\nabla\phi|$. The evolution of ϕ is defined by taking the total derivative of this, giving:

$$\frac{\partial\phi}{\partial t} = -\nabla\phi \times F \frac{\nabla\phi}{|\nabla\phi|} = -F|\nabla\phi| \quad (4.5)$$

This is the partial differential equation that was mentioned earlier in this section. An alternative evolution equation can be reached with the energy-minimisation approach as follows. First, we define the energy as:

$$E(\phi) = \int_{\Omega} g(x, \phi, \phi_x) dx; \quad \text{where } \phi_x = \frac{\partial \phi}{\partial x} \quad (4.6)$$

To minimise $E(\phi)$, we need to solve the corresponding Euler-Lagrange equation:

$$\frac{\partial E}{\partial \phi} = \frac{\partial g}{\partial \phi} - \frac{\partial}{\partial x} \frac{\partial g}{\partial \phi_x} = 0 \quad (4.7)$$

We can evolve the level set function as [32]:

$$\frac{\partial \phi}{\partial t} = -\frac{\partial E}{\partial \phi} = -\frac{\partial g}{\partial \phi} + \frac{\partial}{\partial x} \frac{\partial g}{\partial \phi_x} \quad (4.8)$$

As we noted above, the geometric complexity of tissues found in a thoracic scan presents a substantial challenge to image segmentation, since different tissues of similar densities usually form homogeneous structures shown on the images. For this reason, propagation-based segmentation algorithms, such as active contours, are less effective in achieving good segmentation. In response to this, we use a segmentation algorithm based on the Bhattacharyya flow.

Bhattacharyya distance is a measure of the dissimilarity between two PDFs. For PDFs P_{in} and P_{out} , denoting the regions inside and outside of a curve C , respectively, this measure is given as:

$$B = \int_{\mathcal{Y}} \sqrt{P_{in}(y)P_{out}(y)} dy \quad (4.9)$$

where $y \in \mathcal{Y}$ is the intensity variable. This measure approaches unity for two similar PDFs and zero when they are different. It was proposed [32] that Eq. 4.9 be minimised to evolve the level set function. For an image $I(x)$ with spatial variable x , we use a level set function $\phi(x)$, such that it is negative for the inside of the curve C and positive outside. Using the Heaviside step function, $H(-\phi(x))$ and $H(\phi(x))$ respectively denote the inside and outside of the curve.

Recall the kernel density estimator, in the case for P_{out} and P_{in} :

$$P_{out}(y) = \frac{\int_{\Omega} K_{\sigma}(y - I(x))H(\phi)dx}{\int_{\Omega} H(\phi)dx} \quad (4.10)$$

$$P_{in}(y) = \frac{\int_{\Omega} K_{\sigma}(y - I(x))H(-\phi)dx}{\int_{\Omega} H(-\phi)dx} \quad (4.11)$$

where we used a kernel K_{σ} with bandwidth σ . Then taking the derivative of P_{out} and P_{in} against the level set function and noting $A_{out} = \int_{\Omega} H(\phi)$ and $\frac{\partial}{\partial \phi} H(\phi) = \delta(\phi)$ which is the Dirac delta function,

$$\frac{\partial P_{out}}{\partial \phi} = \frac{\delta(\phi)}{A_{out}}(K_{\sigma}(y - I(x)) - P_{out}(y)) \quad (4.12)$$

$$\frac{\partial P_{in}}{\partial \phi} = \frac{\delta(\phi)}{A_{in}}(P_{in}(y) - K_{\sigma}(y - I(x))) \quad (4.13)$$

We now use Eqs. 4.12 - 4.13 to find the gradient of the Bhattacharyya measure given in Eq. 4.9:

$$\nabla_{\phi} B = \frac{B\delta(\phi)}{2} \left(\frac{1}{A_{out}} - \frac{1}{A_{in}} \right) + \frac{\delta(\phi)}{2} \times \int_{\Omega} K_{\sigma}(y - I(x)) \left(\frac{1}{A_{in}} \sqrt{\frac{P_{out}}{P_{in}}} - \frac{1}{A_{out}} \sqrt{\frac{P_{in}}{P_{out}}} \right) dy \quad (4.14)$$

and to apply this to a level sets evolution, we specify its energy function as:

$$E(\phi) = B(\phi) + \int_{\Omega} |H(\phi)| dx \quad (4.15)$$

and finally find the evolution function as:

$$\frac{\partial \phi}{\partial t} = \delta(\phi)(F + \alpha \kappa) \quad \text{where} \quad F\delta(\phi) = -\nabla_{\phi} B \quad (4.16)$$

Note that only KDE has been used in the above analysis. Applying NPW can produce an unsupervised level set segmentation algorithm [32]. First, we estimate P_{out} and P_{in} using the

NPW estimator instead of KDE. As explained previously, NPW requires knowledge of the three neighbouring pixels which are represented by x_N . Then the elementary distribution is denoted by $P(y; x, x_N, I(x), I(x_N))$. So, assuming a greyscale 8-bit image signal, where only a finite number of intensity values are possible:

$$\nabla_{\phi} B = \frac{B\delta(\phi)}{2} \left(\frac{1}{A_{out}} - \frac{1}{A_{in}} \right) + \frac{\delta(\phi)}{2} \times \sum_y P(y; x, x_N, I(x), I(x_N)) \left(\frac{1}{A_{in}} \sqrt{\frac{P_{out}}{P_{in}}} - \frac{1}{A_{out}} \sqrt{\frac{P_{in}}{P_{out}}} \right) dy \quad (4.17)$$

An illustration of the workings of NPW-LS is given in Fig. 4.2.

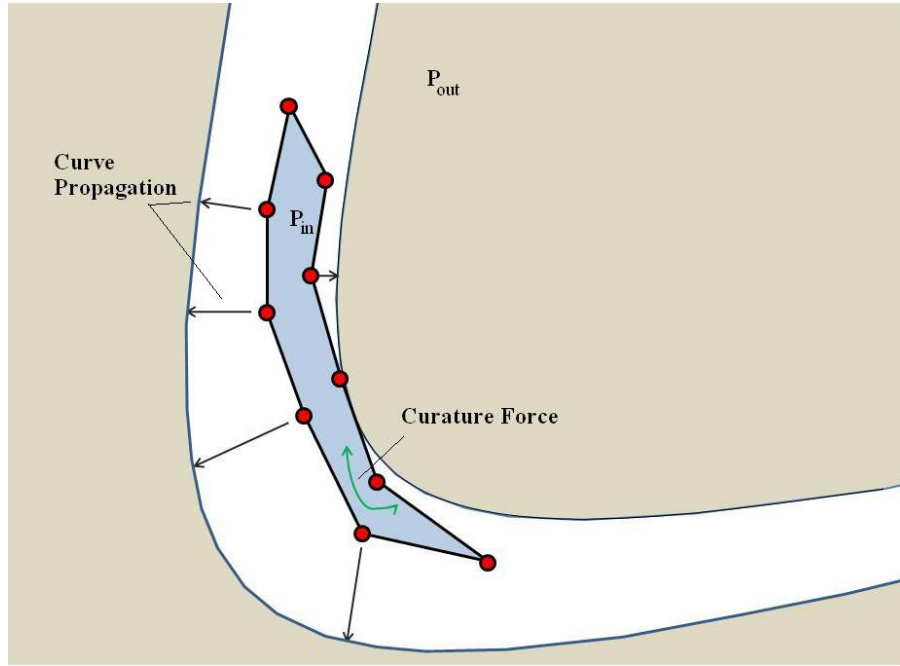


Figure 4.2: The workings of NPW-LS in segmenting MPM; A hypothetical situation of MPM is assumed where the tumour boundaries are outlined and bound by ‘greyed out’ regions. The curve propagates normal to the likely tumour edges. A mixture of curve expansion/contraction and curvature forces work together to control the evolution of the curve. At each iteration step, the forces depend on the probability density distributions of the inside and outside pixels, namely, P_{in} and P_{out} . Note that the longitudinal tumour shape requires highly distinct boundaries to prevent the curve propagating in those directions whilst allowing it to grow along the pleura; this significantly adds to the segmentation challenge.

4.4.2 Results and Discussion

We experimented with the NPW-LS algorithm that was implemented earlier by Joshi and Brady [32], the results were presented in Chapter 3 (Fig. 3.7). In this method, there are five user-defined parameters controlling the level sets propagation, defined as the following:

Propagation Scaling (PS): controls the response to region-based properties (eg. image intensity);

Ballooning Factor (BF): controls the rate of expansion of the contour;

No. of Iteration (NOI): number of iterations;

Advection Scaling (AS): controls the attraction of the detected edges;

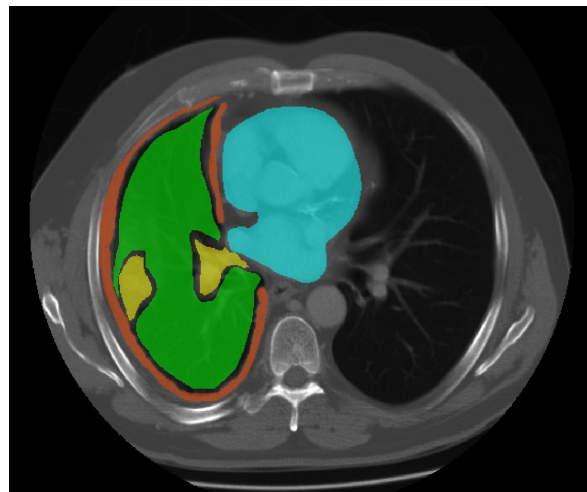
Curvature Scaling (CS): controls the response to boundary curvature changes.

These parameters are set based on user experience and a process of trial-and-error. An average of 2-5 trials were required to set all the parameters. Results in Fig. 3.7 have shown the good performance of this method in segmenting MPM in cases where the neighbour tissues show dissimilar PDFs. However, for most MPM patients, the segmentation problem is generally not so straightforward. Fig. 4.3 shows two general cases of MPM. In Case 1, we observe a complex circumferential shape for the tumour, while in Case 2 the difficulty lies in the interference coming from the surrounding tissues (esp. spleen) of similar pixel intensities. The failed segmentations of these image data are given in Fig. 4.4.

In addition to the poor robustness observed in the general cases of MPM, other limitations of the NPW-LS include difficulties in setting parameters for controlling the level sets evolution. The method had an average running time of 15min for a typical 2D slice of a MPM CT scan. The slow computation renders the method unsuitable for our application and would be further exacerbated in a 3D case. Also, the running time also depends on the user defined tissue initialisation, which can be a further burden to using the NPW-LS segmentation on MPM.



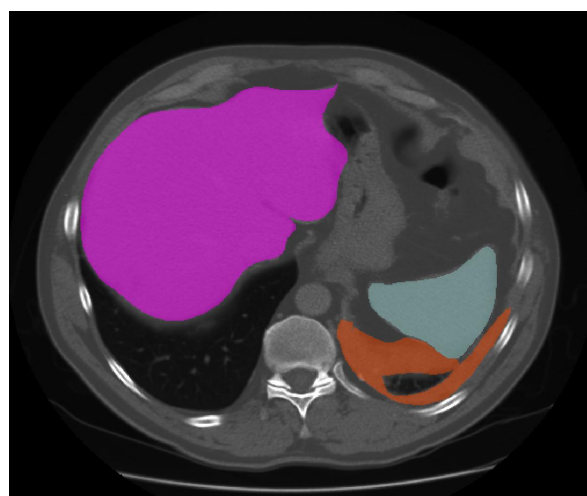
(a) Case 1 CT



(b) Case 1 with Manual Mask



(c) Case 2 CT



(d) Case 2 with Manual Mask

Figure 4.3: More difficult cases of thoracic scans of MPM patients; key tissues are shown in different colours: tumour (*orange*), spleen (*pale blue*), collapsed lung (*yellow*), aerated lung (*green*), liver (*pink*) and mediastinum (*cyan*).

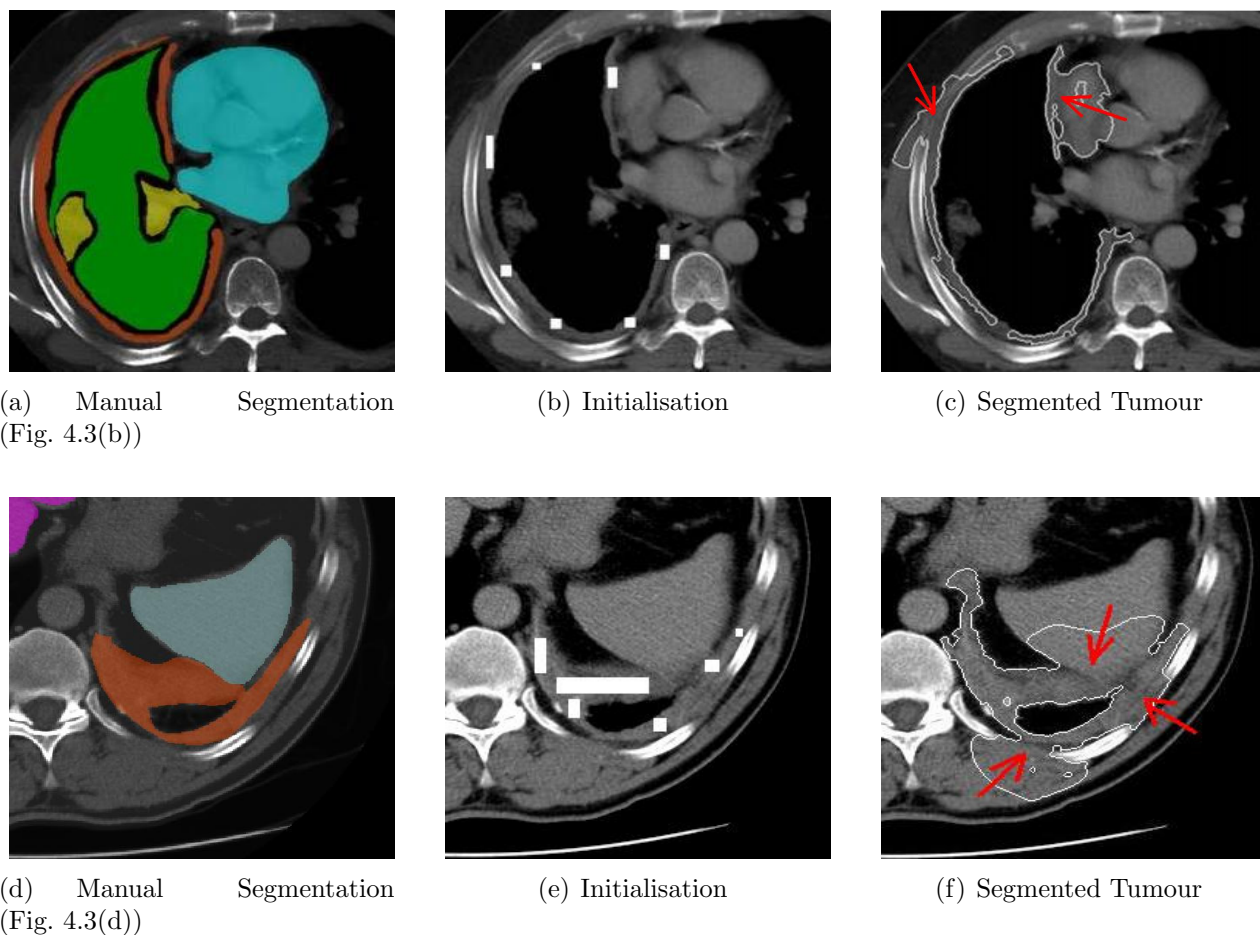


Figure 4.4: Unsuccessful application of NPW-LS segmentation to more typical cases of MPM. Note the longitudinal shape and weak boundaries of the tumour. (*tumour: orange, aerated lung: green, collapsed lung: yellow, mediastinum: cyan*). Parameters used: *top* PS - 1, BF - 22, NOI - 2000, AS - 10, CS - 10. *bottom* PS - 1, BF - 30, NOI - 1000, AS - 10, CS - 10. Note the major 'leakages' occurring at weak tumour edges, indicated by the *red* arrows.

4.5 Summary

In this chapter, we have reviewed related work on the segmentation of MPM for patient treatment response monitoring. Many works [1][2][21][42][57] have shown good evidence in justifying the use of segmented tumour volumes in assessing tumour response to treatments, as an alternative to modified RECIST. However, we also observed that the lack of the application of an MPM-adapted advanced segmentation algorithm has so far prevented an accurate and efficient computation of the tumour volumes. Therefore, there is a need to develop an improved method to tackle the clinical problem of MPM response assessment.

After a brief overview of the main segmentation methods in medical vision, we were motivated to adopt a NPW-based level sets method, for its ability to segment closed tumour shapes and robustness to image noises and spurious edges. However, despite showing promising results in the initial experiment phase, the method had failed to demonstrate similar performances when applied to the more general cases of MPM, producing unsuccessful segmentation results showing major segmentation ‘leakages’ at weak tumour boundaries. Additionally, we have found numerous other difficulties associated with using the method on MPM, including the issues of defining optimal control parameters and slow initialisation-dependent running times. We therefore are further convinced that conventional methods are difficult to apply to MPM. For this reason, in the next chapter, we present and investigate a novel graph-based segmentation method using the concept of random walk, which, as we will show, offers a good solution to our clinical problem.

Chapter 5

Random Walk-Based Method for MPM Segmentation

In our discussion of existing advanced segmentation methods, we noted a number of substantial difficulties in most of the apparently suitable candidate techniques for segmenting malignant pleural mesothelioma (MPM). To address these issues, we have implemented and applied the relatively recent random walk method to our data. We find that it offers a number of notable advantages and addresses many of the weaknesses of the other segmentation methods reviewed in Chapter 4.

5.1 Introduction

First established in [23] for image segmentation applications, the random walk (RW)-based method provides an effective interactive way of segmenting difficult pixel maps.

This chapter is organised as follows. We begin our discussion in Section 5.2 with a description of the theory of random walk and how the mathematical model is applied to the problem of image segmentation; this discussion is largely based on [23]. Segmentation results from the original random walk method are presented and analysed in Section 5.3. In Section 5.4 we introduce an enhancement to the method, namely a boundary regulariser with an updated weighting function that utilises the PDF estimation method expertise we developed earlier in

this thesis. Then as we will show, the problem of uneven tissue boundary edges, as opposed to the clinically expected smooth lesions, is addressed in Section 5.4. Subsequently, a volumetric extension of the method is presented along with results in Section 5.5. One final adaptation of the method to MPM data is made in Section 5.6, demonstrating the effectiveness of applying non-rigid registration to assist in our segmentation procedure.

5.2 Theory

Random walk (RW) is a mathematical formalisation of the trajectory traced by an object undergoing Brownian motion; that is, when it takes successive random steps. This is a naturally occurring phenomenon that is most commonly described by the molecular motion in liquids and gases. Random walks are typically assumed to be Markov chains where the future steps only depend on the current position but not the past. It is also possible to have random walks that exclusively ‘walk’ on pre-defined structures such as planes, groups or graphs. It is the latter case that the segmentation method proposed by [23] uses. It is possible to find the likelihoods of occurrence of different RW outcomes, which correspond to their probabilities.

The principle of the RW-based segmentation lies in the construction of an undirected graph $G = (V, E)$; with nodes $v \in V$ corresponding to the image pixels and connecting edges $e \in E \subseteq V \times V$. A graphical illustration of the method is given in Fig. 5.1 for a 5×5 image of unit weights. The initialisation consists of user-supplied seeds, which are used to solve for the probabilities computed for the unseeded nodes. The label of the seed giving the highest probability to the node is assigned.

In practice, Grady [23] suggests the following weighting function for computing the weight w_{ij} for edge e_{ij} that connects nodes v_i and v_j :

$$w_{ij} = \exp(-\beta(g_i - g_j)^2) \quad (5.1)$$

where β is a free parameter and g_i the image intensity at pixel(px) i . This weight is equivalent to the likelihood that the random walker will cross that particular edge. The formulation of the

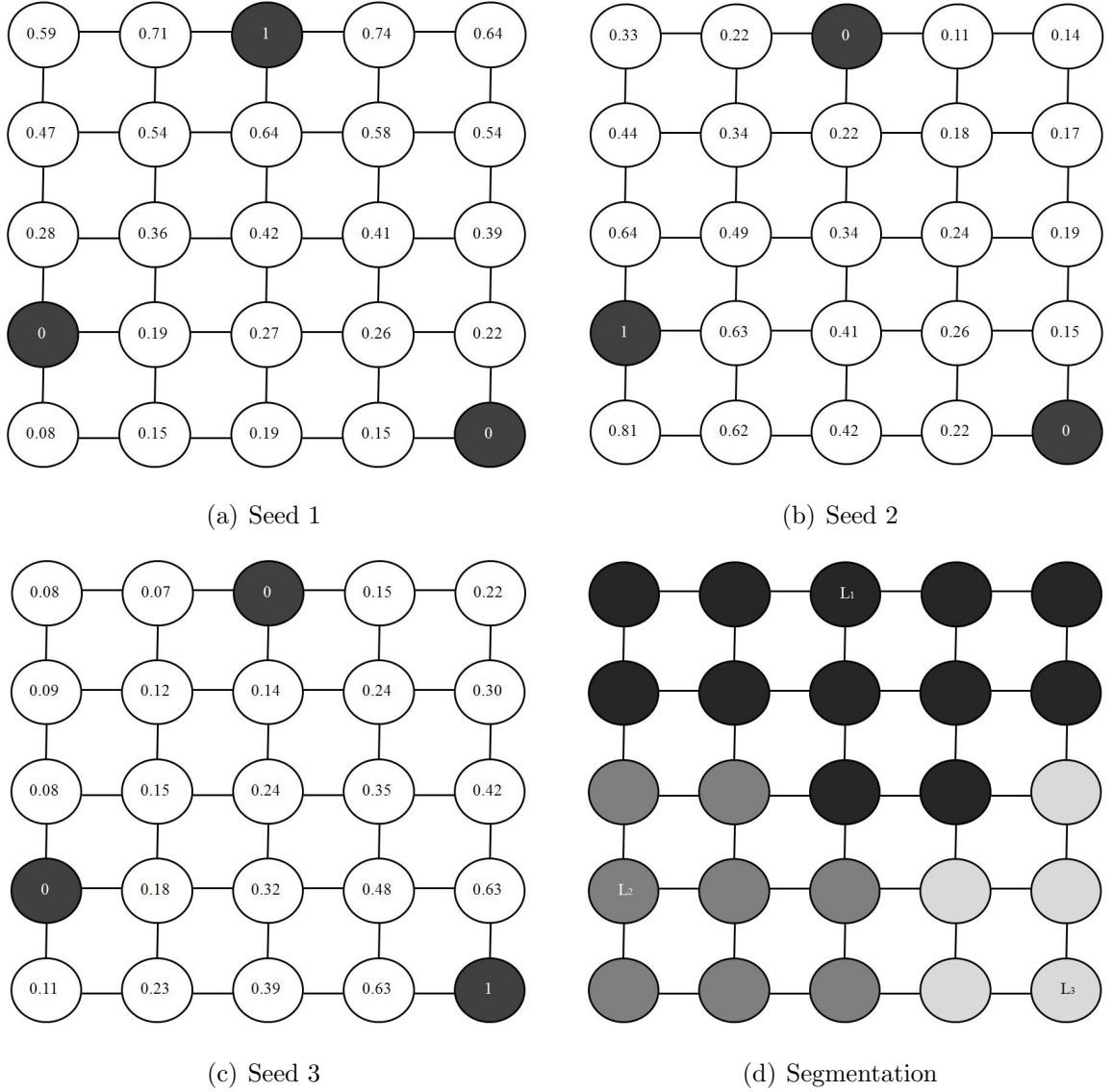


Figure 5.1: Illustration of the random walk-based segmentation for a 5x5 image of unit weights. Individual seeds are sequentially set to unity. The probability of a Brownian motion walker starting at each of the unseeded nodes reaching the unity seed first instead of the other seeds is then determined and given in the circles. Note that because of this the three probabilities (from the three seeds) for each node always sum to one. The nodes are classified based on which seed label yields the highest probability, as shown in (d)

image on a graph allows for the use of combinatorial operators and eliminates the ambiguities associated with discretisation. It is shown that finding such probabilities is identical to solving the Dirichlet problem [23]. The Dirichlet integral is defined as:

$$D[u] = \frac{1}{2} \int_{\Omega} |\nabla u|^2 d\Omega \quad (5.2)$$

for a field u over region Ω . The Dirichlet problem is to find u which also satisfies the Laplace equation: $\nabla^2 u = 0$. In light of solving for u , we define the combinatorial Laplacian matrix as:

$$L_{ij} = \begin{cases} d_i & \text{if } i = j, \\ -w_{ij} & \text{if } v_i \text{ and } v_j \text{ are adjacent nodes,} \\ 0 & \text{otherwise} \end{cases} \quad (5.3)$$

and the incidence matrix as

$$A_{e_{ij}v_k} = \begin{cases} +1 & \text{if } i = k, \\ -1 & \text{if } j = k, \\ 0 & \text{otherwise} \end{cases} \quad (5.4)$$

Defining the constitutive matrix C as a diagonal matrix with the edge weights assigned to its diagonal and all its other entries set to zero, Eq. 5.2 can be re-stated as:

$$D[x] = \frac{1}{2} (Ax)^T C (Ax) = \frac{1}{2} x^T L x = \frac{1}{2} \sum_{e_{ij} \in E} w_{ij} (x_i - x_j)^2 \quad (5.5)$$

such that x is a combinatorial harmonic that minimises Eq. 5.5. Partition the vertices into V_M and V_U for seeded and unseeded nodes, respectively. Decomposing the above equation gives:

$$\begin{aligned} D[x_U] &= \frac{1}{2} [x_M^T x_U^T] \begin{bmatrix} L_M & B \\ B^T & L_U \end{bmatrix} \begin{bmatrix} x_M \\ x_U \end{bmatrix} \\ &= \frac{1}{2} (x_M^T L_M x_M + 2x_U^T B^T x_M + x_U^T L_U x_U) \end{aligned} \quad (5.6)$$

where B is derived from the seeds. The critical point is found by differentiating $D[x_U]$ with respect to x_U . Let V_M be the space of marked nodes. Defining the overall seed vector for each label, s , at node $v_j \in V_M$ as

$$m_j^s = \begin{cases} 1 & \text{if } Q(v_j) = s, \\ 0 & \text{if } Q(v_j) \neq s \end{cases} \quad (5.7)$$

where $Q(v_j) = s, \forall v_j \in V_M$ such that $s \in \mathbb{Z}, 0 < s \leq K$, K being the total number of labels, and let m^s and x^s be columns of M and X , respectively, we have:

$$L_U x_U = -B^T x_M \Leftrightarrow L_U X = -B^T M \quad (5.8)$$

which will be used to solve for the probabilities. For every pre-defined seed, the probability of a random walker starting from each node reaching that seed node first is calculated. These are then applied to Eq. 5.8 to solve for X . Finally for each node v_i , the label corresponding to $\max_s(x_i^s)$ is assigned as its overall probability; with x_i^s denoting the probability of node v_i being label s .

A flowchart is shown in Fig. 5.2 to summarise the main steps in this method. First, we create a graph lattice based on the image and assign weights to the lattice edges based on the user-defined seeds, as in Eq. 5.1. Our algorithm then finds the laplacian and boundary matrix based on Eqs. 5.2 and 5.6, respectively, which also depend on the defined seeds. Next, Eq. 5.8 is applied to solve for the segmentation label probabilities. Finally, the nodes are classified according to the label giving the maximum probability, as illustrated in Fig. 5.1. The whole process can then be repeated with updated seed points until user satisfaction.

Numerical Practicalities

The most computationally intensive part of the algorithm is the solving of the lattice for nodal probabilities. This is especially true for medical image applications where large amounts of data points are present. In our current implementation, we have applied the concept of sparse

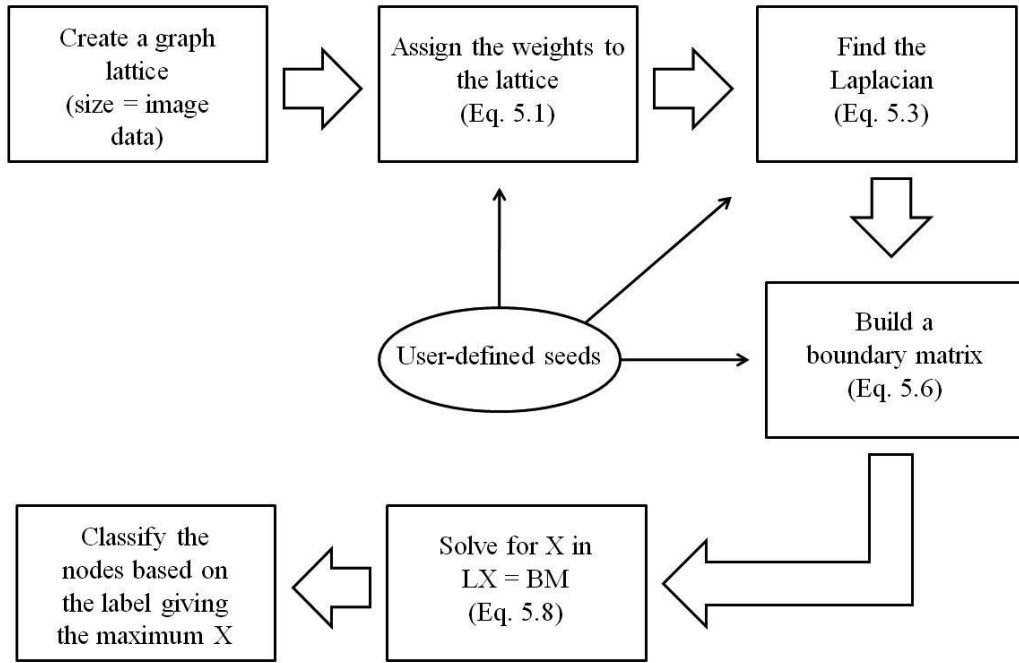


Figure 5.2: Flowchart outlining key steps in the RW-based method. Bulk of the computation lies in the solving of the matrices, or the Dirichlet equation (Eq. 5.8).

matrices for efficient storage of the data entries. For simplicity, at the initial planar 2-D phase, the built-in least squares-based direct solver provided by Matlab® is used. Although this solver can feasibly solve 2-D images, its ability to tackle larger data sizes is very limited. In fact, with the workstation used for our experiments (Intel Pentium D 3.39GHz CPU with 2GB of RAM), we have found the largest image size that can be handled under the Matlab® programming environment, is around $1000px \times 1000px$, roughly equal to 4 slices of the MPM CT data. Therefore, in order to extend the RW-based method to volumetric applications, it is essential that we find an alternative way of implementing the method.

It was briefly mentioned in the original paper [23] that iterative solvers offer the advantage of small memory requirement as well as the ability to represent the matrix-vector multiplication as a function and hence are highly efficient for solving sparse matrices. From our knowledge, we also understand iterative solvers offer us a control over the iterations. Hence it may be possible to guide the iteration in a certain anisotropic way so as to assist the convergence to an accurate segmentation. This may additionally help in achieving the boundary smoothness that we need.

For memory efficiency considerations, the volumetric extension of the RW-method is carried out in an Visual C++ environment, with the aid of the ITK and VTK toolkits. This is discussed further in Section 5.5.

The edge weights are normalised to remove any potential image intensity bias. To avoid the situation where the weight becomes zero due to numerical precision or the choice of β , a small positive constant is attached to all weights, as suggested by [23].

5.3 Results and Discussion

For our studies, we implemented RW-based method according to the following framework (Algorithm 5.1):

Algorithm 5.1: Pseudo-code for Implementing the RW-based Segmentation

```

read in the image data;
read in user-defined seeds, establish their location and label correspondences;
assign a 4-connected graph lattice of the image data size;
set the weights of the lattice, as defined by Eq. 5.1;
find a laplacian based weights and graph lattice, from Eq. 5.2.;
rearrange the seeds according to their labels - seeds of the same label are grouped;
for each label
    assign an element of the boundary matrix, as give in Eq. 5.6;
end
solve for the probability using the laplacian, seed map and boundary matrices, according
to Eq. 5.8 ; a direct equation solver is employed;
for each node
    find the label that has produced the maximum;
    assign the label to the node;
end
output the segmented nodes as the result.

```

To appreciate its effectiveness in segmenting elongated thin structures, we evaluated the RW method on a series of synthetic images, which are modelled after typical lesions found in our MPM data. In this case, we investigated the performance of the method in response to changes in a number of variables, namely, thickness of the segmentation target (pleural lesion), contrast of the lesion tissue against its surrounding structures, and noise level. In the synthetic images, the background is set to represent the aerated lung, spleen and fluid in the MPM data. These varying features are presented along with the synthetic image itself in Fig. 5.13. The Weber definition is used for the image contrast, that is, $C = \frac{I-I_b}{I_b}$, where I and I_b are the tumour and background (inner ellipse) intensities, respectively. We define the critical noise as the point where the method begins to fail to accurately segment the image, i.e with a DICE measure of lower than 0.9. β is the only parameter in the method and is taken as 100 at this stage of the experiment, justification on this will follow.

These responses were then assessed against their robustness to noise. We have applied to the synthetic data a white Gaussian noise of varying variance and with a fixed mean at zero. The Gaussian form is exceptionally used in this case for the ease of controlling the synthetic data. The noise levels are measured in SNR (dB) such that $SNR = \mu_I^2/\sigma^2$, where μ_I and σ^2 are the mean intensity and noise variance respectively. The standard deviation or variance is gradually increased to monitor how the performance of the method deteriorates and eventually falls drastically, indicating the failure of the method. For consistency, the same seed map is used for all our experiments on the synthetic data. Seeds will be taken as individual points for the purpose of evaluating and understanding random walk in 2-D. However, later, for our studies in 3-D, the initialisation of 2-D seed regions will be used for greater accuracy and computational speeds. The segmentation and response studies results are given in Figs. 5.4 and 5.5, respectively. The low critical noise level of the method for the thin cases of the data with low contrast can be as low as 1% of the dynamic range. This can be accounted for by the intrinsic difficulties in segmenting MPM, and/or any thin circumferential structures of low image contrast.

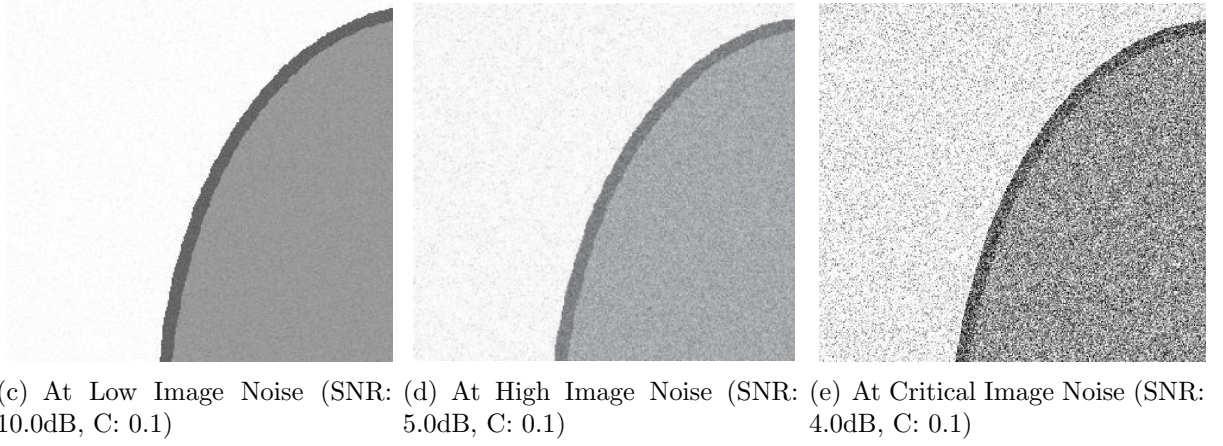
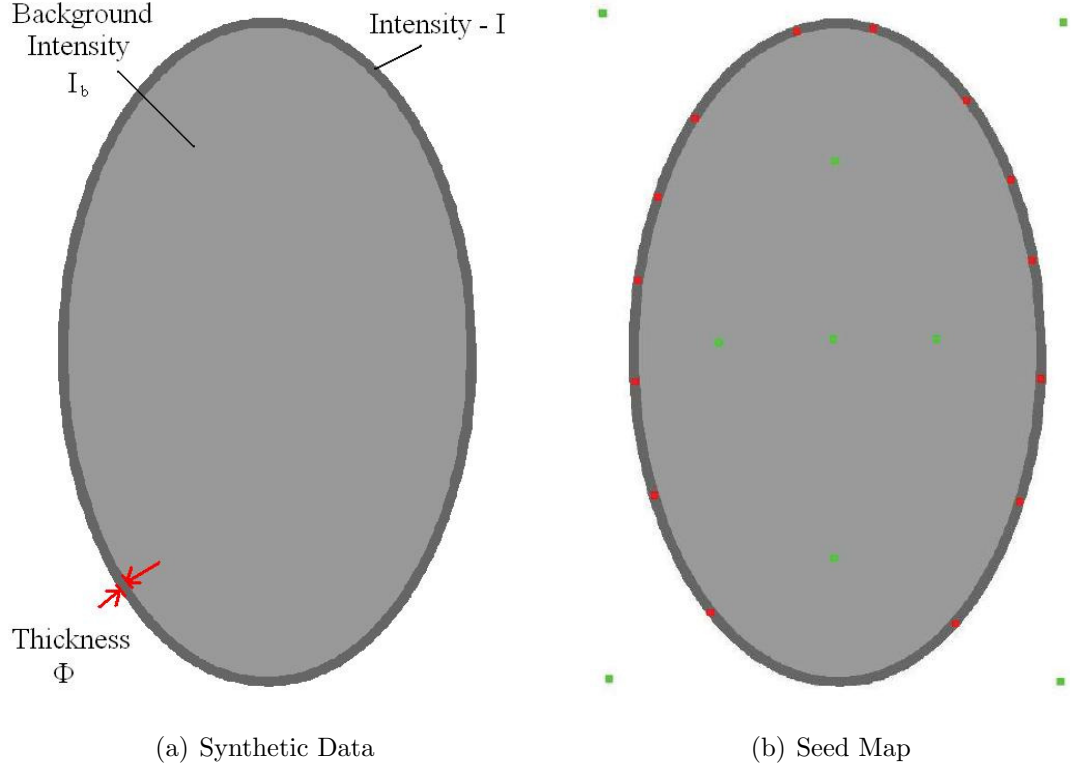


Figure 5.3: Synthetic data with varying noise, critical noise is defined at where the method fails. The synthetic patch measures $250px \times 500px$ and is chosen so as to match the size of a primary MPM pleural lesion. NB: the image contrast is deliberately enhanced for better printing (much lower contrasts were used in our study, to simulate the attenuation similarity of MPM and neighbouring tissues such as fluid and aerated lungs).

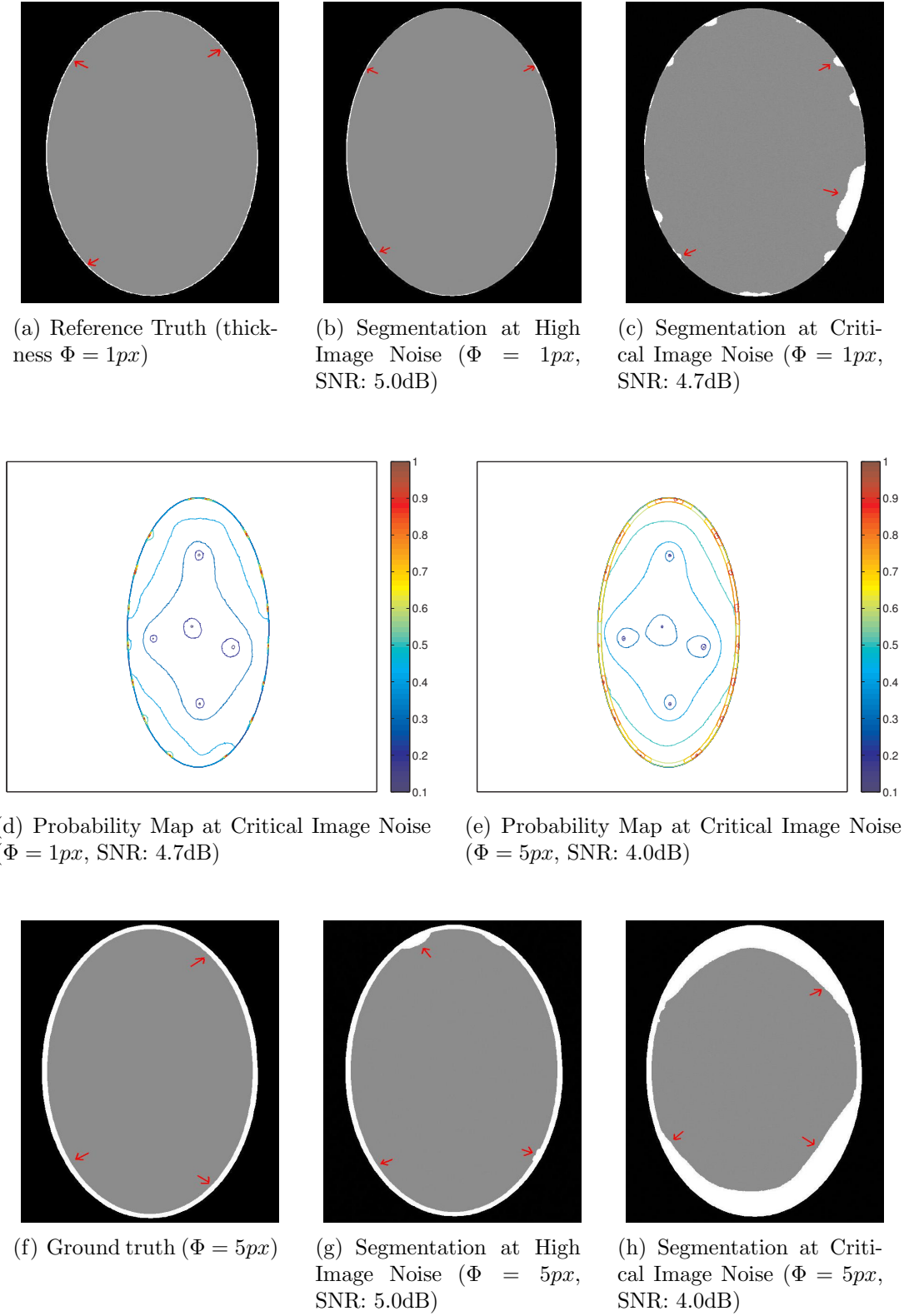
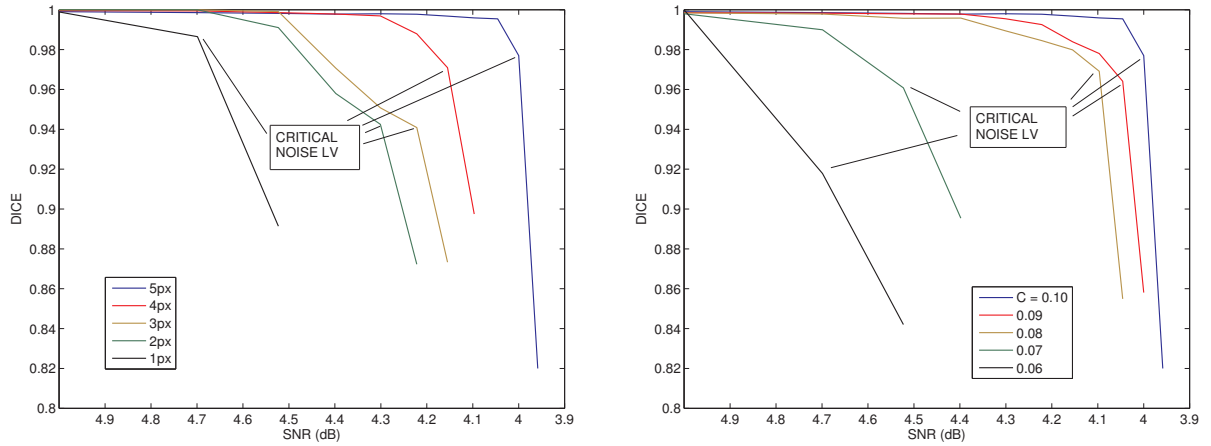


Figure 5.4: Segmentation results for the low and critical noise cases for thin $1px$ and thick $5px$ synthetic data. Probability maps at critical noise levels are also given. Note the propagation and control of the random walk segmentation errors with increasing noise.



(a) Robustness to Image Noise (varying thickness Φ) (b) Robustness to Image Noise (varying contrast C)

Figure 5.5: Robustness to noise with varying thickness and contrast. Note that the critical noise level decreases with a lower thickness or contrast. The contrast and thickness ranges are chosen to match the sizes of potential regions of segmentation difficulty in our clinical data.

It can be observed from Fig. 5.4 that one major reason for a failed RW-based segmentation is the occurrence of a ‘leakage’, that is, when the segmentation perfuses through the gaps in between the seeds and/or intensity gradient boundaries when weak boundaries are present. We therefore examine ways in which the method solves for gaps and look for its limitations at handling ‘boundary gaps’. The results on the synthetic ‘gap’ image are presented in Fig. 5.6. Critical leakage is defined to be when erroneous segmentation through the gap becomes significant. For smaller gap cases, a brief comparison with the result from NPW-LS is made. Here the level sets coefficients and number of iterations are chosen so as to enable the contour to just cover the whole area to the right of the boundary.

Having made the above observations on synthetic data results, we note a number of key advantages offered by the random walk-based segmentation. Most notably, it gives a superior performance to NPW-LS for smaller gap cases though significant (critical) leaks occur for wider gaps. In addition to its good accuracy, the tedious task of selecting level sets parameters is completely avoided. Despite these advantages, however, the RW-based method is still susceptible to the weak boundary problem where the gap width exceeds the critical leakage limit. In addition, image noise continues to cause segmentation inaccuracies and potentially segmen-

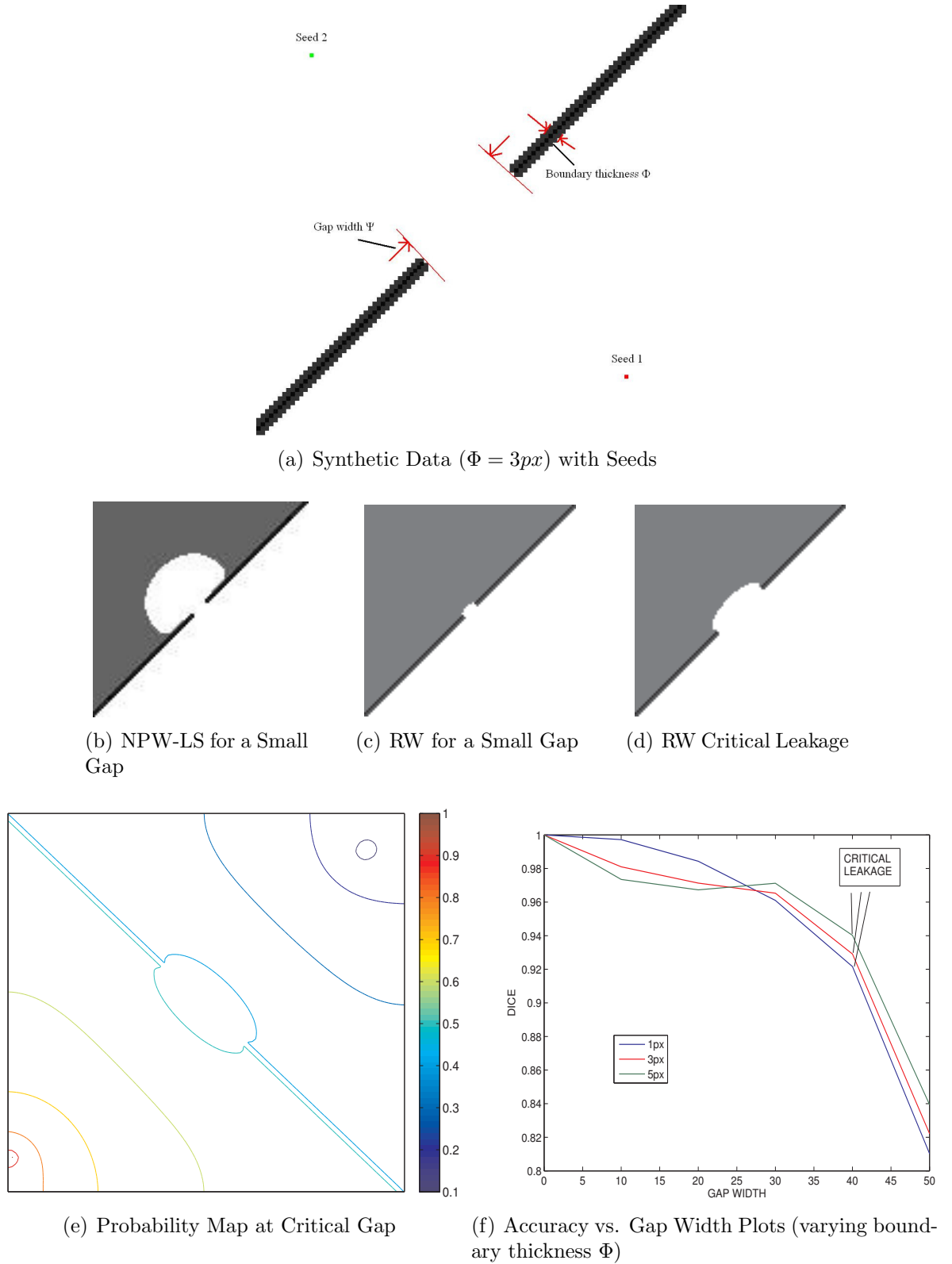


Figure 5.6: ‘Leakage’ study with synthetic image with varying gap width. The RW-based method has superior performance to NPW-LS for smaller gap cases though significant (critical) leak occurs for wider gaps. The parameters of NPW-LS are optimally chosen through a process of trial-and-error. The plot of accuracy vs. gap width shows the critical gap width for a series of boundary thicknesses. Strong correlation between the boundary thickness and critical gap width is not found.

tation failures when the critical noise level is exceeded. These are the two major points that require our attention when applying the method to clinical studies.

Fig. 5.7 presents the segmentation results for the difficult tumour cases from Fig. 4.3. Random walk-based segmentation is applied to eight further image slices with notably challenging image features. The accuracy of the segmentation results is evaluated by calculating the DICE coefficients according to Eq. 5.9. The DICE measures are plotted and given in Fig. 5.8. To address the partial volume effect, morphological erosion is applied to both the reference truth and computed segmentations, the resulting segmentations are also assessed in their accuracy using Eq. 5.9. To show the influence of the choice of threshold on algorithmic accuracy, specificity (true negative rate) and sensitivity (true positive rate) are plotted against threshold level in Fig. 5.9. The intersections of the two plots validate our choice of probability threshold level in this study. As an additional accuracy evaluation, receiver operating characteristics (ROC) curves are presented in Fig. 5.10.

$$DICE = \frac{2|X \cap Y|}{|X| + |Y|} \quad (5.9)$$

for vectors X and Y .

Good performance of the random walk method is clearly demonstrated; with all ten trials yielding accurate segmentations (DICE coefficients exceeding 0.8). This reflects good handling of the weak boundaries by random walk and its ability to segment any arbitrary shapes with appropriately placed seeds. As a side note, the accuracy improvement offered by the use of morphological erosion on the segmented result is clear. This demonstrates the critical role played by the partial volume effect in the RW segmentation process (and in fact in image segmentation in general). Extensive discussions on this topic will follow in Chapter 6.

Another observation made on the basis of these initial segmentation results is that they lack the boundary smoothness expected for most MPM lesions. This may be the root cause of the inaccuracies in our results (where DICE falls below 0.90). We will investigate this further in Section 5.4 and show our steps to developing a regulariser for achieving the required smoothness in the boundary of the segmented tumour. Finally, we justify our choice of β by a

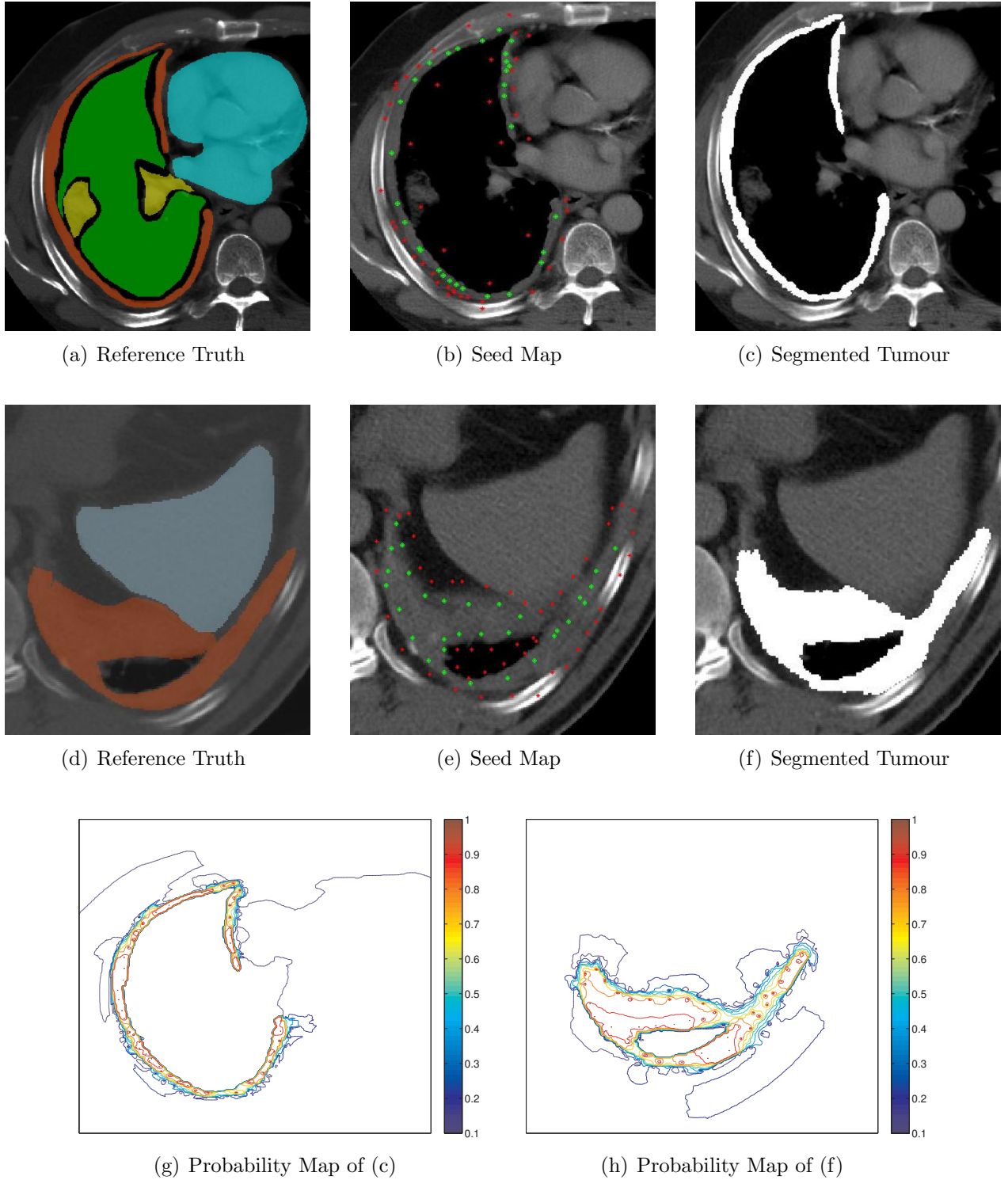


Figure 5.7: Random walk-based segmentations for the two difficult cases of MPM from Fig. 4.3. (*tumour: orange, aerated lung: green, collapsed lung: yellow, mediastinum: cyan, spleen: grey*) Green seeds are those that are inside the tumour and red seeds are those outside. Also shown here are the probability maps for the segmentations, which show likely regions of leakage. Good accuracy may be observed in both cases.

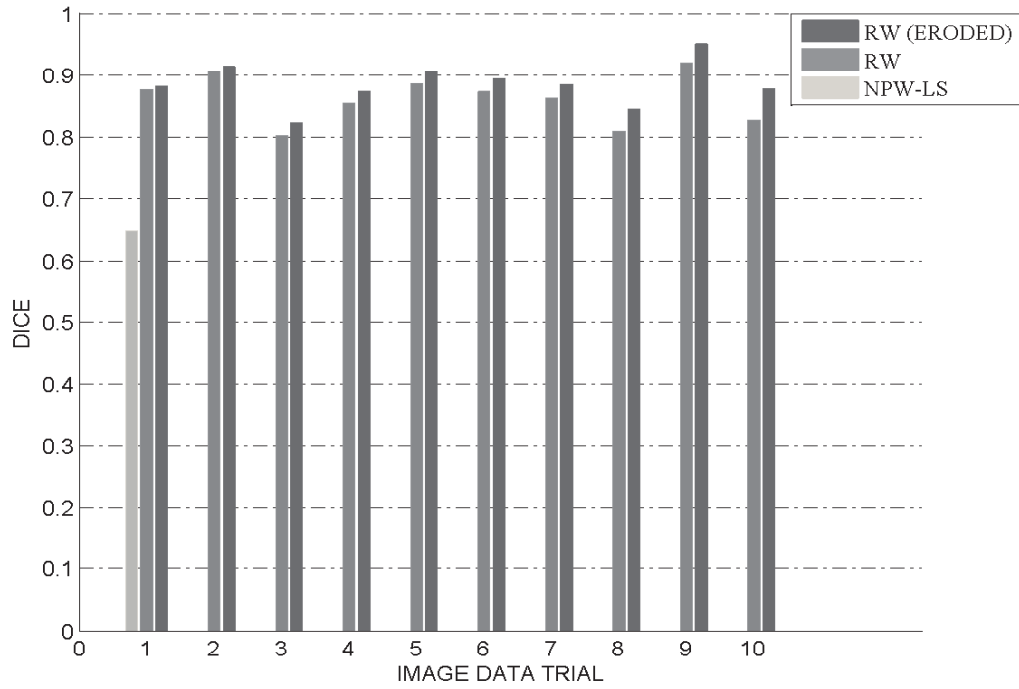
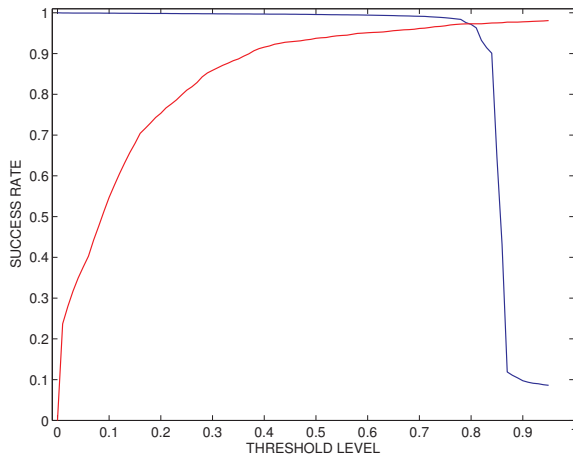


Figure 5.8: DICE coefficients of the 2-D random walk testing trials, as matched to the reference truth; trials 1 & 2 are already given in Fig. 5.7 (a) and (d), respectively. NPW-LS trials are deemed to have ‘failed’ when large areas of ‘leakage’ exceeding the area of the tumour itself are present, which is the case for trials 2 through 10.

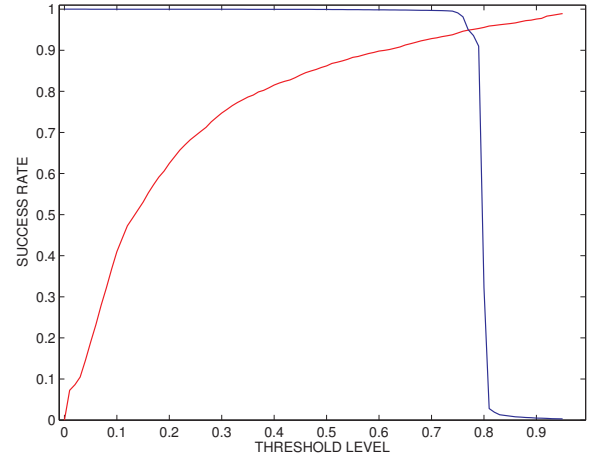
plot of accuracy against the free parameter in Fig. 5.11 where the best performance is attained when $\beta = 100$. Mean accuracies from the ten testing trials are given in this plot.

Comparison to Other Segmentation Methods

The random walk-based segmentation method is shown to be robust to weak boundaries, due to its Brownian motion characteristic, and requires only one free parameter, β from Eq. 5.1, to be set. In addition, the algorithm allows for fast editing of the segmentation by assigning additional seeds after a segmentation is completed and the results are largely free of the ‘small cuts’, or localised segmentation fragments, found in other graph-based segmentation routines, such as graph cuts [23]. Other benefits of the method include: fast computation, robustness to noise, ease of implementation, as well as the ability to produce any arbitrary segmentation from the user-defined seeds [23]. The last property is especially useful, as we have observed in our initial results, since it guarantees a successful segmentation with an acceptable degree of

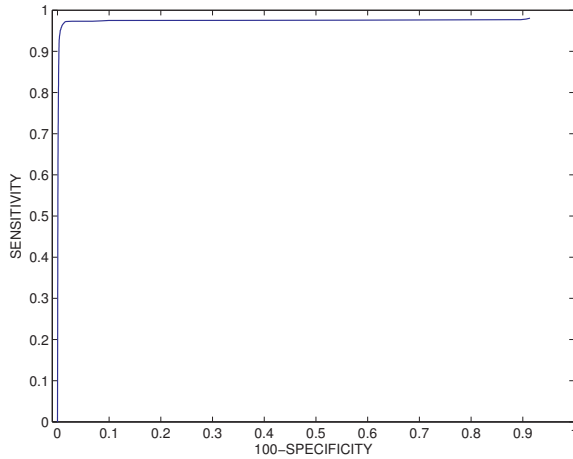


(a) For Trials from Fig. 5.7(a)

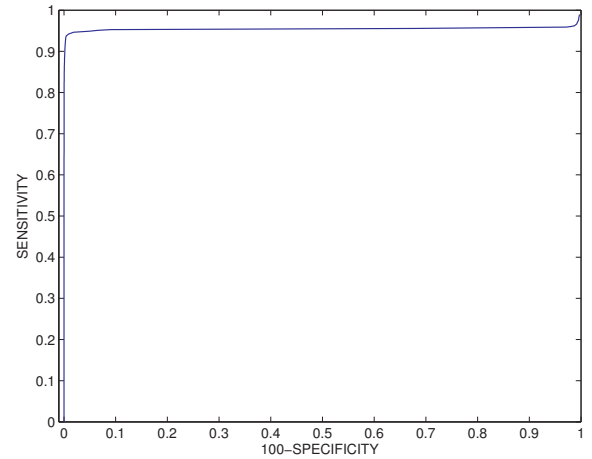


(b) For Trials from Fig. 5.7(d)

Figure 5.9: Specificity (blue) and sensitivity (red) plotted against threshold level for trials from Fig. 5.7(a) and (d). The plots show the good accuracy of RW classifications where both specificity and sensitivity curves are well placed, as expected.



(a) For Trials from Fig. 5.7(a)



(b) For Trials from Fig. 5.7(d)

Figure 5.10: ROC curves for trials from Fig. 5.7(a) and (d). Sensitivity: true positive rate, 100-specificity: false positive rate. The plots show the good accuracy of RW classifications where both specificity and sensitivity measures are optimally correlated, as expected.

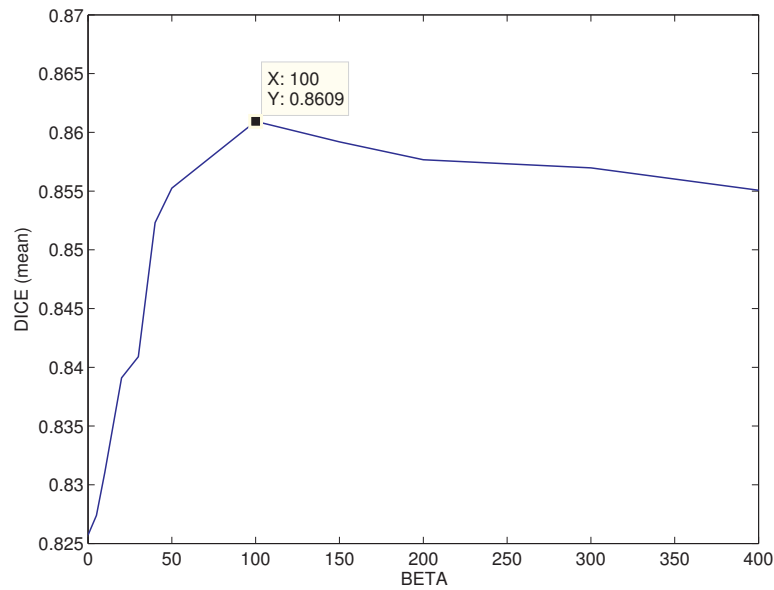


Figure 5.11: Control of the RW-based method performance by the choice of β . Note peak accuracy is attained when $\beta = 100$.

user interaction.

The amount of user interaction is well balanced in this method and depends solely on the image data itself. For most of our difficult cases, the algorithm performs extremely well with 50-150 seed points. The process of seed definition is also very straightforward, taking an average of less than one minute to complete. This is a major improvement over the manual segmentation routine where scrupulous selection of a large number of accurate boundary points from a highly trained clinician on a tablet PC was required, for each and every image slice. Potential regions of difficulty such as weak tissue boundaries or elongated shapes have been easily overcome with the specification of a few additional seeds after the preliminary segmentation. This shows that random walk provides a very generic routine for segmenting our image data. Next, we note from Fig 5.12 that the correlation between the number of seeds used and the resulting accuracy is not obvious, so we suggest that the required seed number mostly depends on the image data itself and adding more seeds beyond the required level does not impact greatly on the accuracy measure.

In terms of noise robustness, our experiments with the synthetic data with increasing noise clearly showed good performance of the random walk-based method in handling image noise.

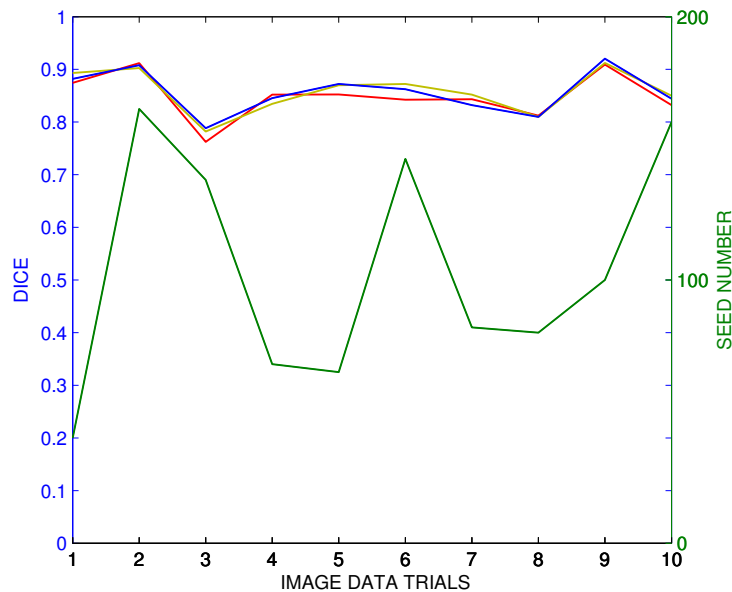


Figure 5.12: The effect of user-interaction on accuracy. The amount of interaction required to achieve a successful segmentation mostly only depends on the image data itself and does not have a direct effect on segmentation accuracy. Repeated trials with different seed allocations but the same seed numbers are shown in different colours (*red, blue, yellow*). We note the good repeatability and robustness of the method.

However, we should also note that this still depends on other regional image properties such as contrast and the thickness of the interesting tissue. In general, regions that cause difficulties can always be compensated for with additional clinical guidance in the form of initialisation seeds.

Consider the trial cases whose DICE measures are shown in Fig. 5.8. As we reported earlier in the thesis, NPW-LS was able to segment successfully only the first of these ten cases, and even then yielded a less accurate result, with a DICE coefficient of just 0.625, compared to 0.885 for random walk with morphological erosion. The accuracy of random walk-based segmentation is again shown by the ROC curves in Fig. 5.10, in which rapid ascents to the optimal sensitivity level are observed. A comparison of the computation speeds of random walk (RW) and NPW-LS methods, given in Table 5.1, clearly illustrates the fast computation property of random walk. All of the experiments were carried out on a workstation with 3.39GHz CPU with 2GB of RAM.

Experiment	LS-NPW	2-D RW
Time (min)	15	0.05

Table 5.1: Typical computation times of different algorithms (in minutes). The computation time varies according to the actual image and the figure given is only an indicator of the nearest order of magnitude.

We have thus seen the superior performance of random walk for segmenting mesothelioma. In the later sections of this chapter, we will present steps we have taken to improve the existing method to better adapt it to our clinical needs in the MPM study.

5.4 Boundary Enhancement Regulariser

From our clinical study of MPM, we understand that during the earlier to mid- stages of MPM (before the occurrence of large-scale metastases), its growth pattern retains a smooth surface. This requires us to address the issue of boundary smoothness in our segmentation. With the initial version of the algorithm, reported above, the results generally lack the expected smoothness. The ‘rough’ segmentation boundaries may be explained by the presence of excessive image noise in our data. Earlier we have noted, from our working with the synthetic data, that random walk, despite its good robustness to noise, is still liable to substantial inaccuracies when the noise exceeds a critical threshold. Based on this observation, we consider it to be important to develop a boundary enhancement regulariser for the existing RW-based method. The main goal of developing the regulariser is to increase the clinical reliability of the RW method and to make it better adapted to our MPM experiments.

Perona and Malik [55] first introduced the use of anisotropic diffusion for image smoothing. By convolving the image with a non-linear space-variant filter, it allows for noise reduction without compromising the image content and features. Parameters of the filter applied vary according to the local information of every image voxel. In this manner, the resulting filtered image is somewhat ‘guided’ by its own information. By contrast, an isotropic diffusion is where the filter parameters remain constant regardless of the presence of image features in the

proximity. Difference between these two ideas is illustrated in Fig. 5.13:

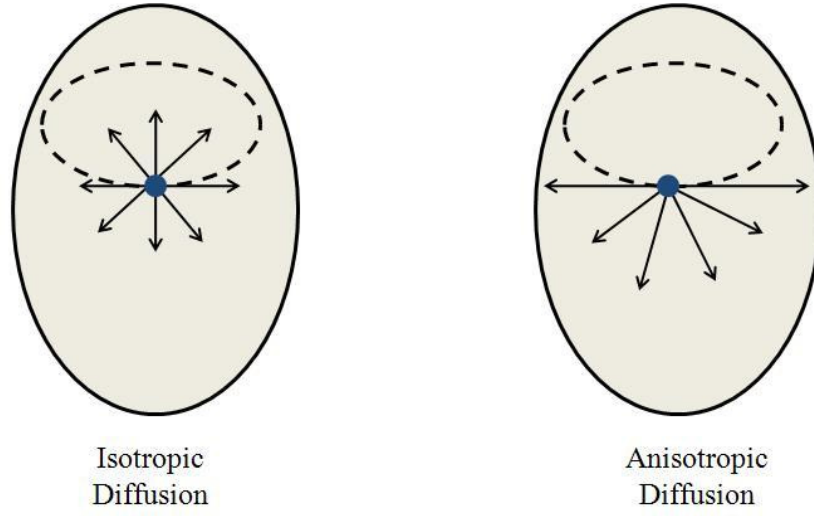


Figure 5.13: Isotropic vs. anisotropic diffusion. The elliptical shape represents the lung cavity; the dotted shape inside models the segmentation contour of the tumour that is set to grow in the directions pointed by the arrows, for the blue dot. The diffusion propagates equally in all directions in the isotropic case and avoids direction of the nearby edge in its anisotropic counterpart.

The governing equation of anisotropic diffusion for a greyscale image I is defined as:

$$\frac{\partial I}{\partial t} = \text{div}(c(x, y, t) \nabla I) = \nabla c \cdot \nabla I + c(x, y, t) \Delta I \quad (5.10)$$

where ∇ , Δ and $\text{div}(\dots)$ are Laplacian, gradient and divergence operators respectively. $c(x, y, t)$ controls the rate of diffusion and is a function of the image gradient and chosen to best preserve key image features such as lines and edges. Perona and Malik [55] proposed two forms for this function:

$$C(\|\nabla I\|) = \exp(-\|\nabla I\|/K)^2 \quad (5.11)$$

$$C(\|\nabla I\|) = \frac{1}{1 + \frac{\|\nabla I\|^2}{K}} \quad (5.12)$$

where K is a edge sensitivity measure, which can be found either empirically or as a function of the image noise. We note that the weighting function used in the current implementation (Eq. 5.1) already assumed Eq. 5.11, reproduced below for the ease of reference:

$$w_{ij} = \exp(-\beta(g_i - g_j)^2) \quad (5.13)$$

where $\beta = 1/K$. This equation assigns edge weights that guide the solving of a graph in an anisotropic way based on walking along (instead of crossing over) the high gradient boundaries of image intensities. This is useful in that there is only one free parameter and facilitates calibration of the algorithm. However, the above function does not make full use of the available image and seed information. To improve the boundary smoothness performance, we could utilise the probabilistic information that can be easily deduced from the initialisation seeds. Moreover, the existing method relies on the Gaussian assumption, which is generally not applicable to MPM images. This motivates the adoption of a non-parametric formulation.

Based on our knowledge of segmented thoracic tissues from Chapter 3, we may assume that the intensities in the foreground (seeded) voxels are drawn from a single intensity distribution. By estimating this distribution with a non-parametric method (one of those presented in Chapter 3), it would be possible to update the above weighting equation to incorporate some essential probabilistic prior information drawn from the pre-defined seeds. The significance of this enhancement is that it makes the edge weights of the RW graph more meaningful and better reflects the image processing situation. The new lattice should be expected to reflect the presence of foreground seeds, which are set to resist diffusions in certain boundary directions and carry higher weightings in the Dirichlet matrices. The anisotropic walk is also expected to converge faster in this manner. Consequentially, we propose the following updated weighting function:

$$w_{ij} = \exp(-\beta(F(g_i) - F(g_j))^2) \quad (5.14)$$

where $F(g)$ is the probability of pixel intensity g from the probability density function (PDF) F , which can be computed using one of the PDF estimators discussed in Chapter 3.

Grady and Jolly [24] have described a similar weighting function in their paper for the

purpose of benchmarking a number of weighting functions and connectivity cases for the graph cuts segmentation method. In their paper, they used the kernel density estimator (KDE) for computing the probabilistic distributions. However, as we noted earlier in Chapter 3, KDE is highly inefficient and orders of magnitude slower than the non-parametric windows-based method (NPW), which is used throughout our study. It should be noted that the histogram estimator would generally not work in this case due to the small data sample size (limited by the number of seeds). The advantages of using the NPW estimator include its ability to produce smooth PDFs given small sample sizes as well as its better accuracy and fast computation relative to KDE.

The results, given in Fig. 5.15, clearly show good boundary segmentation that supersedes the existing algorithm and justifies the effectiveness of the new boundary regulariser. The probability maps in Fig. 5.16 demonstrate an improved more seed-oriented probability distribution pattern to that of the original method, which is expected given that now probabilistic prior information is taken into account during the construction of RW lattice. The choice of the free parameter β and its effect on image accuracy are shown for both the original and boundary enhanced (BE) methods in Fig. 5.14. The accuracy performance of the BE-method at the optimal β is found to be superior to that of the original method, at a new optimal parameter setting of $\beta = 150$. We observe from this the effect of β on optimising the RW method performance.

To assess the expected accuracy improvement, we perform segmentation with boundary enhancement on the same ten images as in the previous section; the results are given in Fig 5.18. We note that despite the boundary smoothness which meets clinical expectation, the boundary regulariser does not necessarily improve the DICE accuracy measure of the segmentation. We believe this to be caused by the uncertainties in the reference truth itself. In particular, observer variabilities and partial volume effect are known factors that affect the clinical validity of our reference truth. As we have explained earlier in Chapter 3, because the manual segmentations carry much uncertainty in themselves, we view them only as referencing guides in our studies. A lower DICE measure, in other words, does not necessarily imply a worse segmentation. All segmentations and numerical findings presented in the result section of this thesis, Chapter 7,

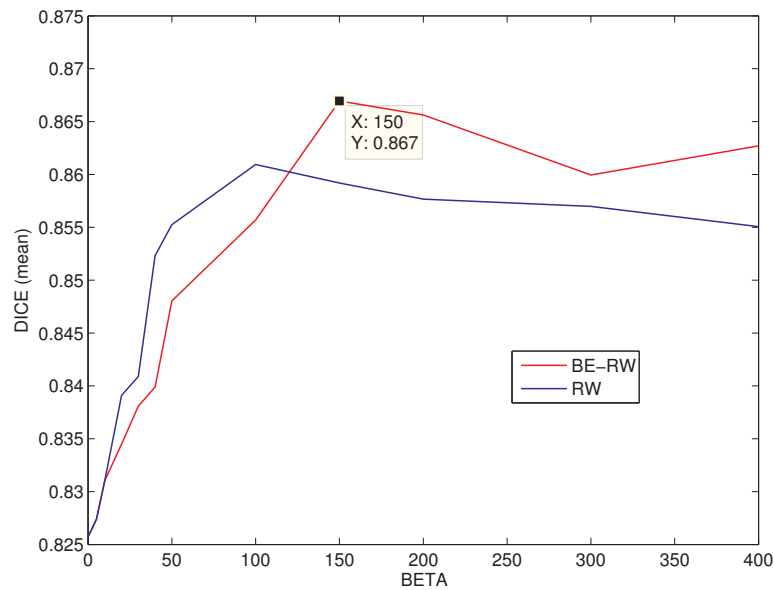


Figure 5.14: Control of the boundary regularised method by the choice of β , as compared to the original unenhanced method. Optimal performance is attained at $\beta = 150$. Mean DICE measures taken from all ten trials are shown.

have been examined and verified by the MPM radiology specialist.

The PDFs of these two image trials, their foreground seedmaps as well as the segmented tumour and reference truth, are given in Fig. 5.17. These results confirm the role played by the seedmap PDFs in anisotropically assigning edge weights and subsequently isolate and segment the tumours from their surroundings.

Finally, the additional computation time of the method with the boundary regulariser is relatively insignificant (Fig. 5.19). NPW, being a fast and highly efficient PDF estimator, is able to compute the foreground seeds distribution within fractions of a second. For most cases, the increase is less than 10% of the overall computation time of the original which is negligible given that the computation time fluctuates up to 100% depending on the image data content.

5.5 Volumetric Extension

A volumetric extension of random walk follows and will form the basis of the development of a computer-aided algorithm equivalent to, or even superior to, the modified RECIST criteria used at the present. With knowledge of the tumour for only a number of images, estimation

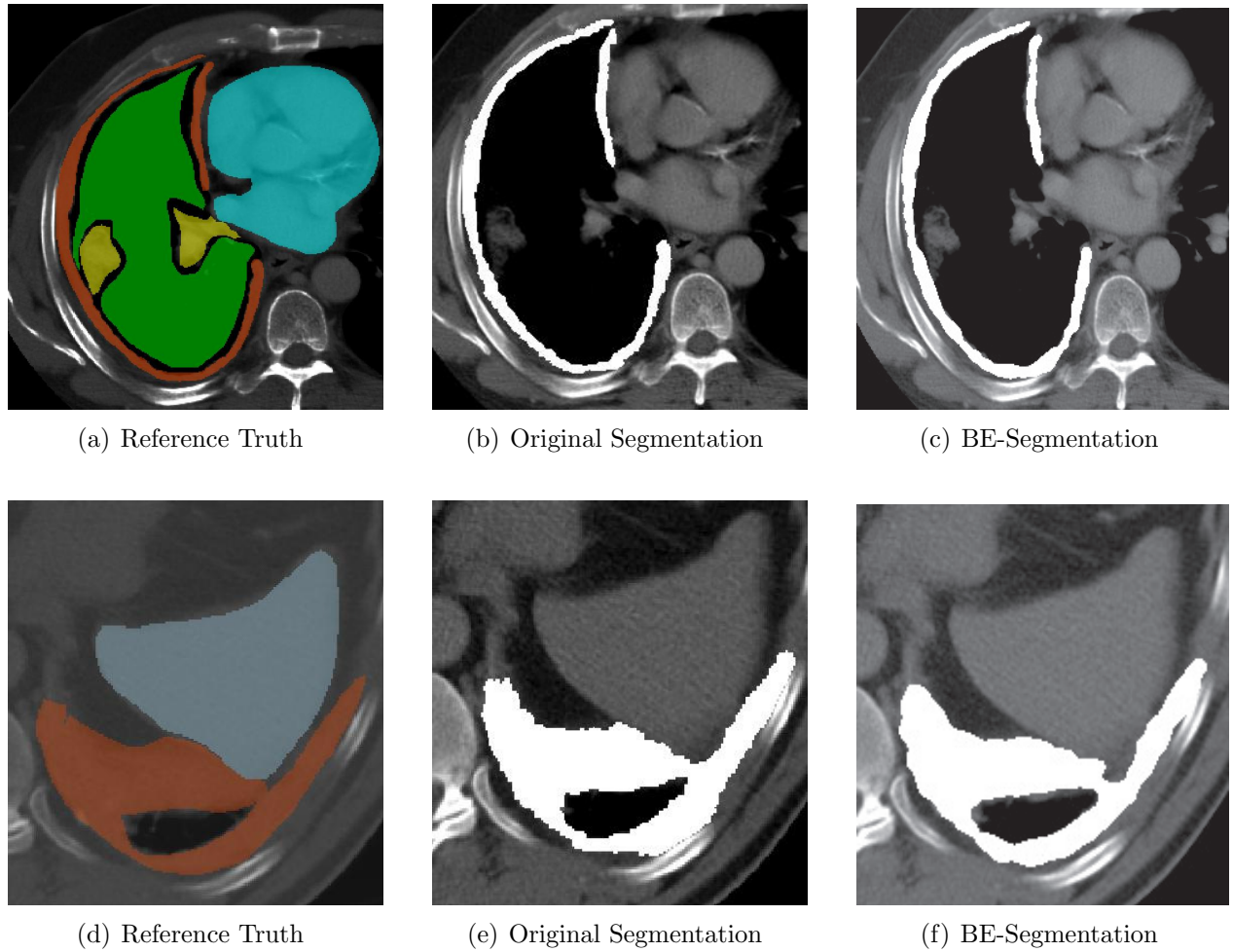


Figure 5.15: BE-RW segmentation for the same MPM data segmented in Fig 5.7. (*tumour: orange, aerated lung: green, collapsed lung: yellow, mediastinum: cyan, spleen: grey*). Boundary smoothness enhancement is clearly observed in both cases.

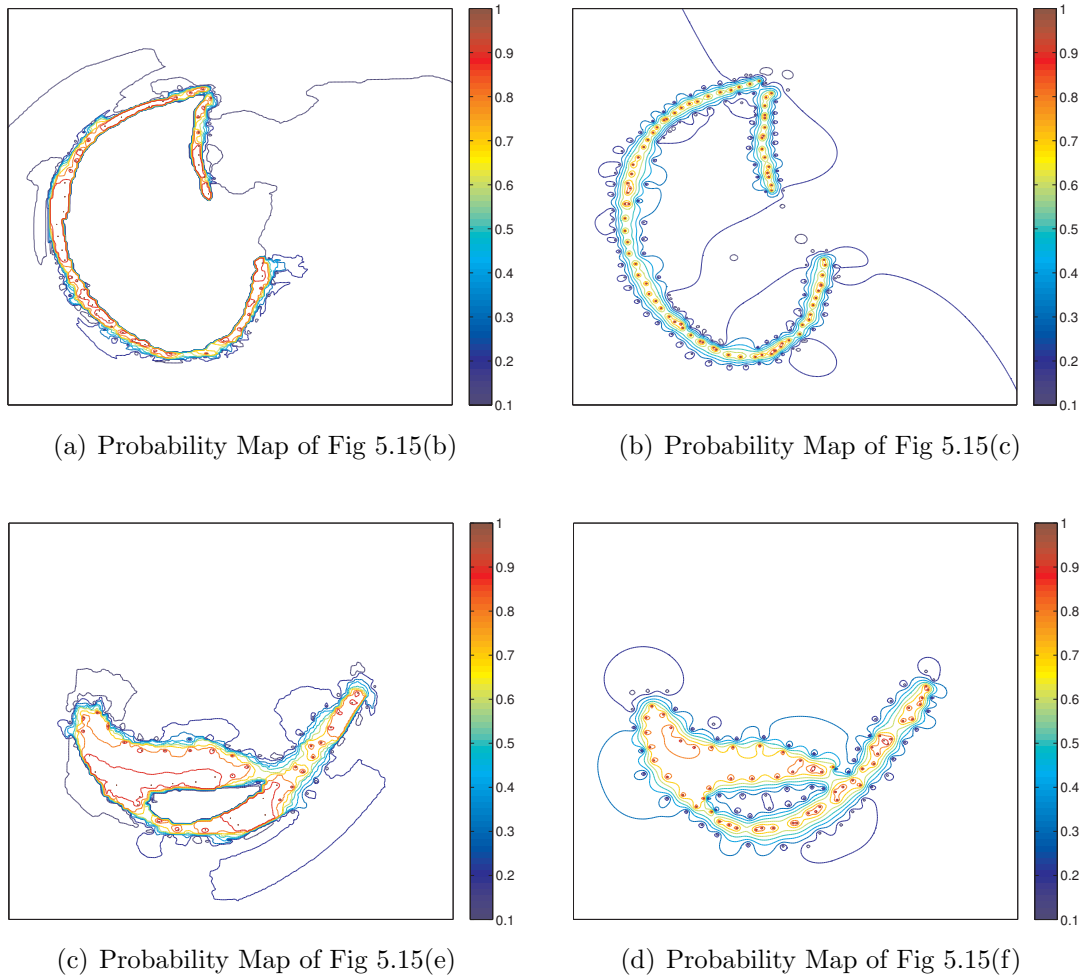
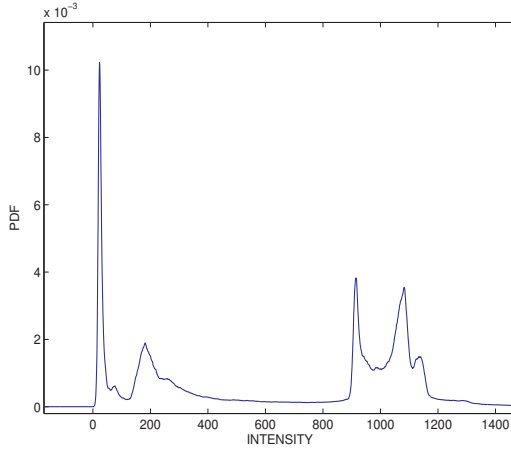
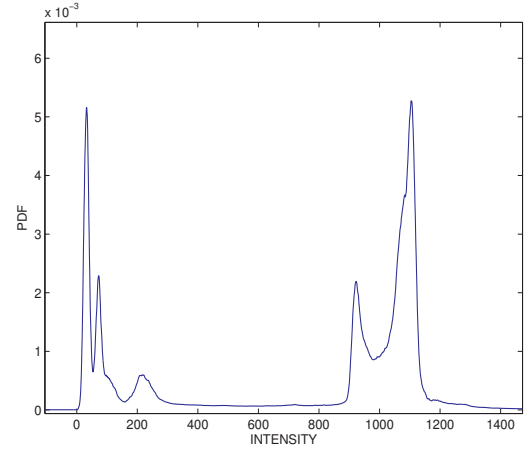


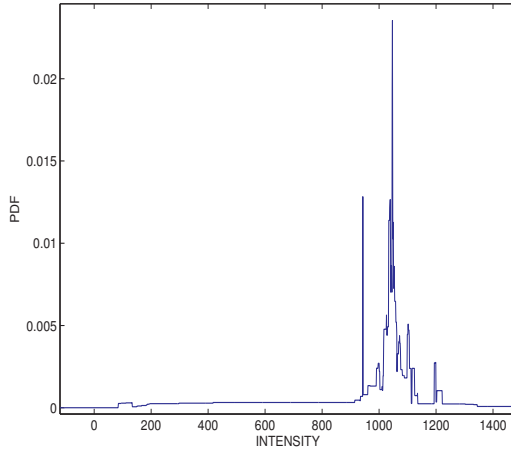
Figure 5.16: Isocontour probability map comparison of boundary enhanced versus original segmentations. The probability maps indicate an improved probability pattern attributable to the new boundary regulariser.



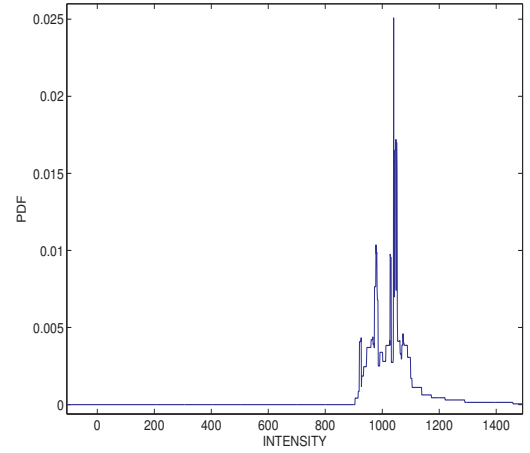
(a) Overall PDF of Fig 5.15(a)



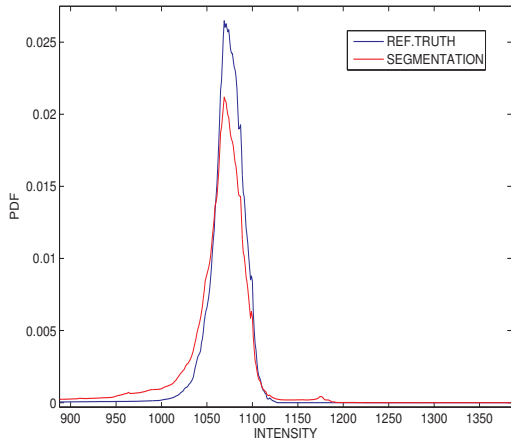
(b) Overall PDF of Fig 5.15(d)



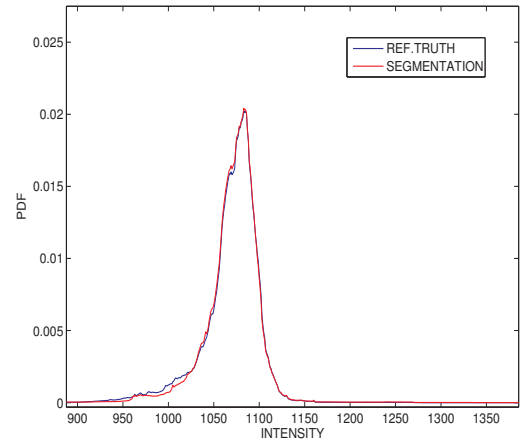
(c) Tumour Seed PDF



(d) Tumour Seed PDF



(e) BE-Segmented Tumour PDF vs. Reference Truth



(f) BE-Segmented Tumour PDF vs. Reference Truth

Figure 5.17: PDF estimates of the MPM tumour segmented in Fig 5.15, compared to the reference truth and foreground seed map PDF. It can be seen that the foreground seed distributions play a major role in allocating the region of interest. This information is used as prior in our boundary smoothness regulariser.

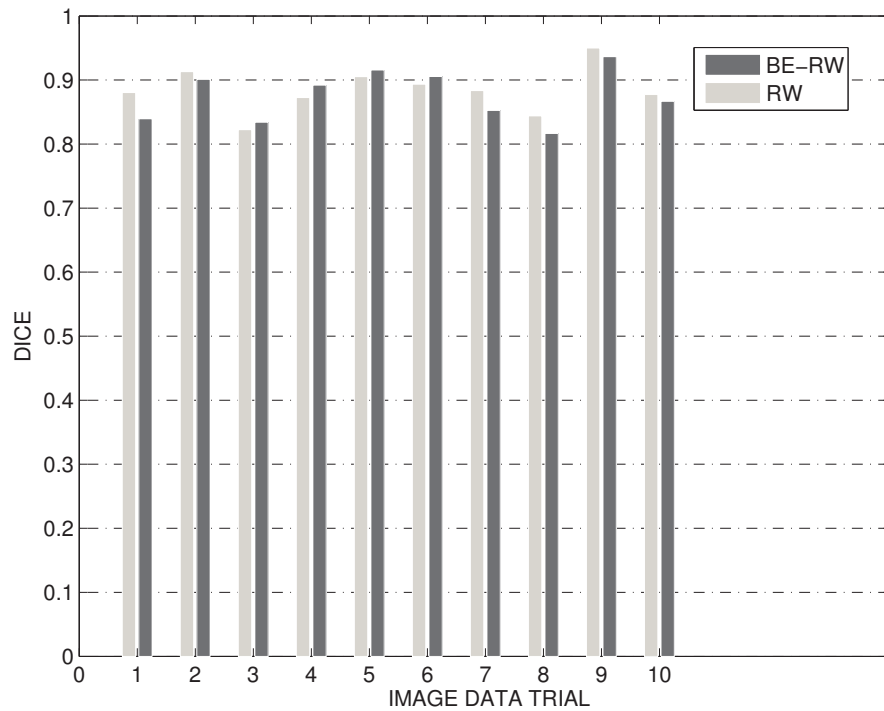


Figure 5.18: DICE measures from the 2-D trials for the boundary enhanced (BE) method as compared to the original method. Results from both methods underwent erosion prior to being assessed in accuracy. Note that despite the boundary smoothness, which conforms to clinical expectations, the boundary regulariser does not always improve the DICE accuracy measure of the segmentation. We believe this to be caused by the uncertainties in the reference truth itself.

of the full volume would still be possible, where information can be gathered from the nearest slices containing initialisation seeds. However, with the significant increase in data sizes for the full 3-D data (typically 512x512x128 pixels), it would be computationally impractical to run the code under the Matlab® environment used for the planar implementation.

The Insight Segmentation and Registration Toolkit (ITK)¹ was developed by the National Library of Medicine for image processing in medical vision and provides a good solution to our problem. The main advantages of ITK are the following:

- 1) Provides resources for many image processing tasks
- 2) Compatibility with a wide range of platforms
- 3) Open-source, object-oriented nature

¹The ITK code is available as a set of software packages from www.itk.org.

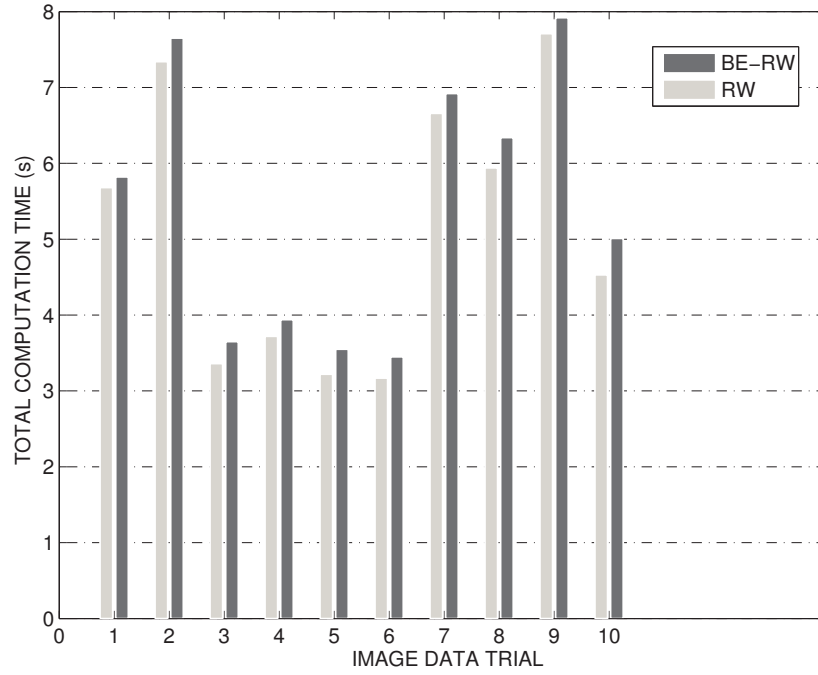


Figure 5.19: Computation time comparison of the original and boundary enhanced RW method for the ten trials. Additional computational time required for PDF estimation is relatively insignificant compared to the level of overall computation time fluctuation between image trials.

- 4) Constantly updated and a sophisticated user distribution list

For these reasons, we have used it to carry out the volumetric segmentation. All codes in the extension are written in C++ in the MS Visual Studio 2005® environment, with essential tools from ITK. This allows for a more efficient computation and handling of larger image data, and has proven to be sufficient for our clinical studies.

Perhaps the most important choice to make for the random walk volumetric extension is the 3-D topology. We considered a number of possible connectivity cases, their pros and cons, and finally chose one that best suits our needs in terms of accuracy, computational efficiency and equally importantly, ease of implementation.

The simplest connectivity of a 3-D lattice is the 6-connected case (Fig. 5.20(a)), where the edge set is defined as:

$$E = \{i, j \mid \|C(v_i) - C(v_j)\| \leq 1\} \quad (5.15)$$

where $C(v)$ maps node v to its 3-D coordinates. $C(v_i) - C(v_j)$ is the distance between connecting

nodes. It is also possible as a second scenario, for the centre node to be connected to the planar adjacent corner nodes; such lattices are known as 10-connected (Fig. 5.20(b)). The formal definition is given below:

$$E = \{i, j \mid \|C(v_i) - C(v_j)\| \leq 1\} \cup \{i, j \mid \|C(v_i) - C(v_j)\| \leq \sqrt{2}, \forall C_z(v_i) = C_z(v_j)\} \quad (5.16)$$

where $C_z(v)$ is $C(v)$ on plane z . A Euclidean distance adjustment term would be needed in the weighting function for cases with non-unity distance adjacent nodes, adding to the numerical complexity. A third connectivity case is where all corner adjacent nodes are considered, regardless of the plane they are in, gives rise to a 26-connected local lattice (Fig. 5.20(c)). The edge set is defined as:

$$E = \{i, j \mid \|C(v_i) - C(v_j)\| \leq \sqrt{3}\} \quad (5.17)$$

Conventionally we understand that stronger connectivity gives better results [24], as more information are considered at the local level. Noting its computational and implementation complexities, we dismiss the 26-connected topology from our consideration and will limit our choice to between the 6- and 10-connected cases. The 10-connected case, however, is liable to negative bias because the addition of planar nodes reduces the influence of inter-planar interaction and subsequently the significance of the volumetric extension. Therefore, the 6-connected case is selected for our work, for its theoretical simplicity, ease of implementation and the best overall expected performance out of the three topological scenarios. Previously for our study of random walk in 2-D, we have considered a 4-connected case, so this is merely a straightforward extension from our early works².

²Our current 3-D implementation is built on a base framework coded by a lab collaborator, Gordon Stevenson. In his work, Gordon has applied the VNL matrix solver for dealing with the sparse matrices in the Dirichlet problem (Eq. 5.8). The VNL solver is an iterative solver based on the conjugate gradient method. This is much easier to control (by adjusting the tolerance) and faster than the direct solver used in 2-D. However, I have compared a number of iterative solvers for solving the Dirichlet problem and found that the generalised minimum residual method gives the best computation speed at any given tolerance level. Therefore, the computation speed may still potentially be improved using a more efficient iterative solver, such as minimum residual. This is left as future work; the results of our works to date in this direction are given in Section 8.2.3.

Image Data	RW	BE-RW
Fig. 5.21	16m27s	19m45s
Fig. 5.22	20m37s	26m29s

Table 5.2: Computation time comparison of the four trials presented in this section (3.39GHz, 2GB RAM). The times are significantly shorter than those of other segmentation methods such as NPW-LS (20 minutes/slice). Also note that the use of boundary regulariser necessitates additional time for computing the PDF. However, this extra time is small compared to the variation of the overall computation time due to that of the captured image content. This confirms our 2-D findings.

The 3-D method is applied to the image from Fig. 5.7(a), whose results are given in Fig. 5.21 and 5.22. As explained earlier in this chapter, we now use seed regions for initialising our volumetric segmentation; whilst requiring little additional user interaction, a reduced number of unmarked voxels decreases the amount of computation required and contributes much more to the prior knowledge in the algorithm. All initialisations at this stage are based on inputs from the clinical expert, that is, the reference truth on a planar basis. The overall image data is masked for the region of interest (ROI) based on the positions of the relevant seeds. Only the masked regions are passed onto the matrix solver for a faster computation. To validate the accuracy of our volumetric method, we have used two fully segmented volumes and take one out of (on average) every 3-4 segmented slices as initialisation. The unmarked slices are expected to receive implicit seed information from their adjacent slices.

From the segmentation results presented, the superior performance of random walk with the boundary regulariser is clearly demonstrated, with smooth boundaries matching closely to those of the reference truth. This is in line with our findings from the 2-D experiments earlier in the chapter. To validate the volumetric method, the results are assessed against the full volume reference truths and represented as DICE measures in Fig. 5.23. The computation times are given for all four trials in Table 5.2.

In this section, we have studied only two image datasets. Much of our hypotheses and observations on the volumetric extension of random walk can only be validated by assessing the method against a much larger sample size. However, this is not currently possible in our

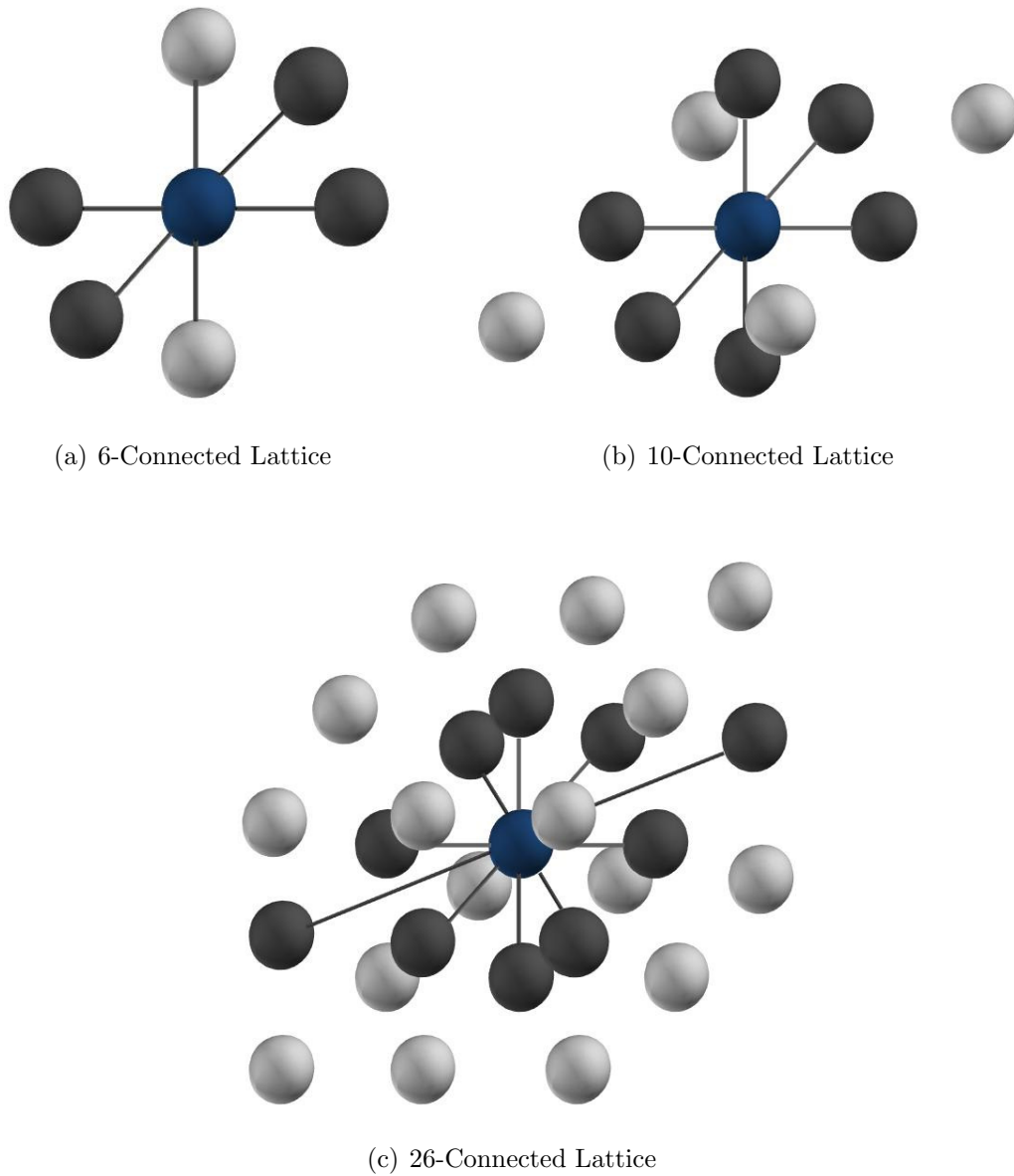


Figure 5.20: A comparison of different connectivity cases for a 3-D random walk lattice. The centre voxel *blue* is surrounded by existing *black* and newly connected nodes *grey* in each case. Note the asymmetrical structure of the 10-connected lattice, which introduces a negative bias.

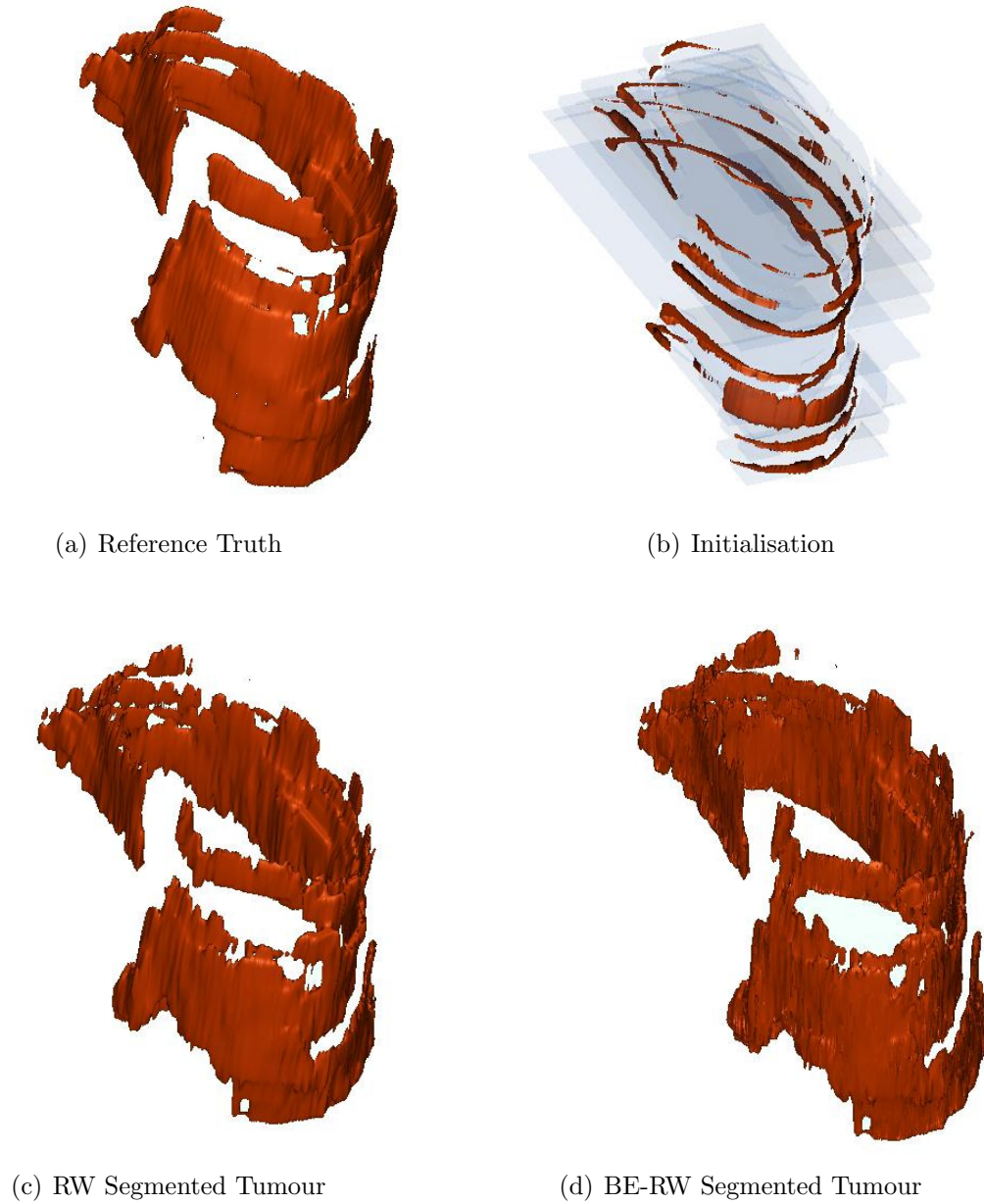


Figure 5.21: Volumetric segmentation of MPM with 3-D random walk (data from Fig. 5.7(a)). The superior performance of random walk with the boundary regulariser is clearly observed, with smooth boundaries resembling closely those of the reference truth. The initialisation are based on the manual clinical segmentation; As given in (b), the ‘in seeds’ are displayed in *orange* and ‘out seeds’ in *blue*.



(a) Reference Truth



(b) Initialisation

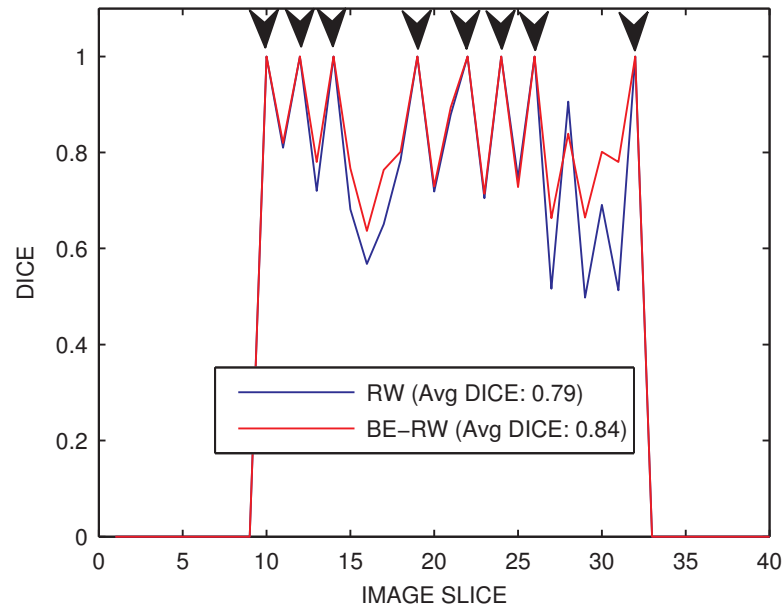


(c) RW Segmented Tumour

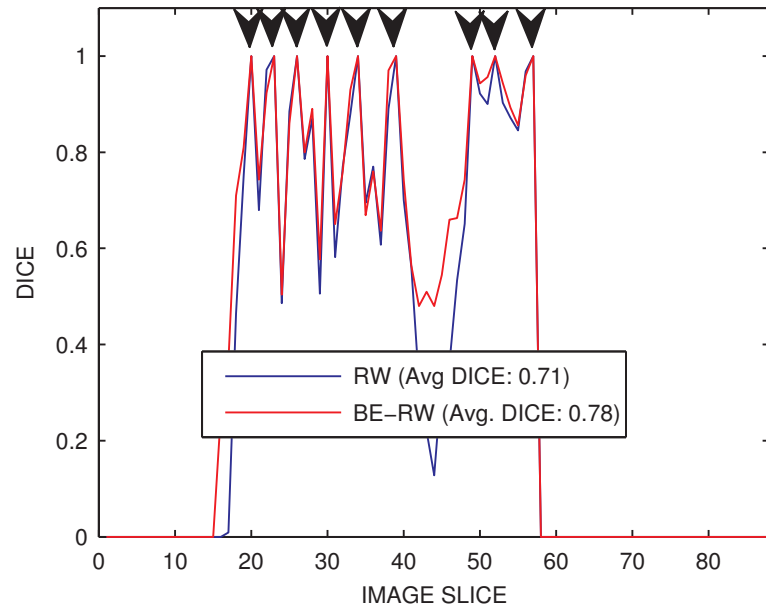


(d) BE-RW Segmented Tumour

Figure 5.22: Volumetric segmentation MPM with 3-D random walk (data from Fig. 5.7(d)). Note the good accuracy achieved towards the top of the image where the initialisation slices are more densely concentrated. The opposite is also true in this case, for the middle section where no initialisation slices are found. This shows the influence of seed regions to nearby image slices in a 3-D context. The initialisation ‘in’ seeds are displayed in *orange* and ‘out’ seeds in *blue*, as shown in (b).



(a) Accuracy of the Segmentation from Fig. 5.21



(b) Accuracy of the Segmentation from Fig. 5.22

Figure 5.23: Accuracy (DICE) analysis of the four trials presented in this section. Results are shown for the full volume data, where accuracy is zero outside the range of tumour as defined by the reference truth. Reasonably good accuracies are observed across the four trials. We note that unlike in the 2-D case, both results this time show better performance of the BE-method over the unmodified random walk across the range of image slice. This may be due to more certainty within the reference truth when it is available for larger data and averages out the impact of intra-observer variability and regional partial volume effect. However, a formal justification of this observation would need more thorough experimental study for which additional complete clinical segmentations would be required. Moreover, with BE-RW, the minimum DICE has improved in both cases, meaning BE results are more robust because the DICE results are guaranteed not to fall below the average DICE by a large margin.

project because of the scarcity of clinically verifiable manual segmentations. In general, only thick slice (5mm, 40-65 slices/volume) scans can be fully validated in this manner, where full-volume manual segmentations are feasible. Because much of our data are thin-slice scans (up to 512 slices/volume), we would need to rely on grounds established from our experimental findings of the 2-D random walk. As we have seen so far, much of our knowledge from the first sections of the chapter apply well to the volumetric extension of random walk.

A notable advantage of the 3-D random walk segmentation over a stack of 2-D segmentations is that it requires initialisation only on a small number of image slices. Specifying pre-defined seeds on a single slice is sufficient to produce a full-volume segmentation, though the accuracy significantly improves with the addition of more seeded regions. In the next section, we introduce a novel way to efficiently formulate additional segmentation information without increasing the user interaction. This is particularly useful in our clinical application, where only limited prior knowledge of the MPM tumour is available for initialisation use.

5.6 Registration-Assisted Segmentation

From the volume results presented in the last section, we have seen that having sufficient knowledge of the tumour’s approximate location is the key to an accurate segmentation of MPM. However, the topic of effectively collecting such data is in itself a major challenge. First, manually defining lesion contours is a very lengthy process even with a high level of specialist expertise in the field. On average, 2-5 minutes would be needed for each image slice. Clearly, this is not viable for clinical studies with large number of thin-slice scans, such as ours. Secondly, the presence of the partial volume effect and image noise amplifies observer variabilities and uncertainty in the manual segmentation to match the actual tumour growth. In this case, we believe that a segmentation scheme more reliant on a good smooth computed results based on ‘seed flows’ from the adjacent slices may resemble more closely to the actual ground truth. This suggests that we apply more spread out and evenly distributed seed regions along slices in the volume axial direction. Despite this, the intrinsic difficulty in obtaining accurate seed information prevents a straightforward application of the current segmentation method for most

of our medical data.

Aside from image segmentation, another main branch of research in medical image analysis is image registration. This refers to the process of geometrically transforming different sets of medical data into a unified coordinates system. Driven by the idea of harnessing ‘seeding knowledge’ from a different time-point scans of the same patient through such a transformation, we are motivated to apply available techniques in image registration to help us tackle the problem of ‘seed point scarcity’ in our MPM data. In this section, we show steps we have taken to achieve this and the performance advantage offered by this addition to our existing algorithm, as well as its computation time consideration.

We begin with an outline of the key steps in the registration-assisted segmentation that we have developed. First, given two scans of a patient at times T_1 and T_2 , a deformation field can be computed by registering the data at T_1 (source, denoted as $I(T_1)$) to that at T_2 (target, $I(T_2)$). Then, initialise $I(T_1)$ with seed regions towards one end of the scan and pass it through the random walk algorithm in order to find the segmented tumour mask at T_1 , namely, $I_{ts}(T_1)$, with emphasis on the seeds end. This then is applied to the deformation field to find the registered tumour mask at T_2 : $I_{tR}(T_2)$. Next we seed $I(T_2)$ with seed regions near the other end of the scan and segment it accordingly, giving $I_{ts}(T_2)$. Apply the two segmented volumes to the following equation for finding the combined segmentation $I_t(T_2)$:

$$I_t(T_2) = I_{ts}(T_2) \cup I_{ts}(T_1) \quad (5.18)$$

The later steps are repeated for finding $I_t(T_1)$ from the same segmented volumes $I_{ts}(T_1)$ and $I_{ts}(T_2)$. Finally, the end results, $I_t(T_1)$ and $I_t(T_2)$, are compared to assess the tumour response to treatment over the time interval $[T_1, T_2]$. The overall flow of the proposed algorithm is illustrated in Figs. 5.24 and 5.25. Note the obvious advantage of applying the additional registration step in the algorithm as compared to direct segmentation followed by volumetric comparison: it avoids the tedious process of defining a large number of seed regions as required previously.

Having proposed the use of image registration to assist in our algorithm, we are now in a good

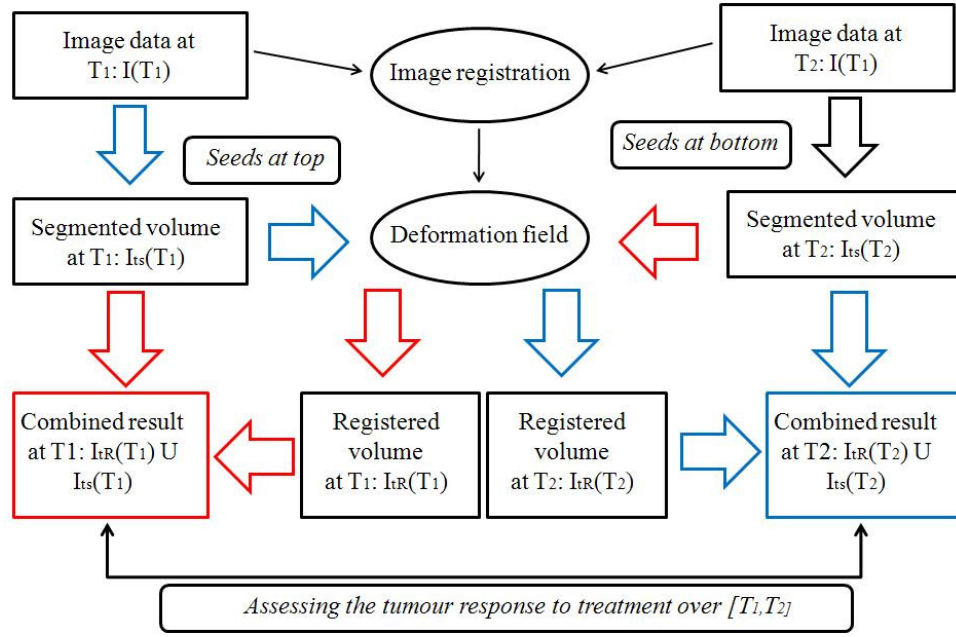


Figure 5.24: Flowchart for the registration-assisted segmentation. The bold arrows indicate the main flows of the data. The thin arrows show the building of the deformation field. Construction steps of enhanced volume at T_1 , $I_t(T_1)$ are given in *red*. Steps for $I_t(T_2)$ are in *blue*. The end results, $I_t(T_1)$ and $I_t(T_2)$, are compared to assess the tumour response to treatment over $[T_1, T_2]$.

position to discuss our choice of registration algorithm. For image registration, the registered datasets can be either different scans of the same patients or different patients and of a single or multiple imaging modalities. In this thesis, we consider only the registration of the same patient scans using a single imaging modality (CT). Linear and deformable transformation models are available. Since our data are not generally well-aligned, we need to apply a deformable model. Intensity-based methods are useful for cases where intensity provides a good measure for establishing links between the datasets at each voxel. Feature-based methods, on the other hand, build the correspondence based on image features such as points or boundaries. Given that we need to be able to apply seeds at any arbitrary location on the image, we would require an intensity-based deformation field. Lastly, since we propose to apply the registration to raw data without any markers, an automated method is necessary.

To fulfil the above requirements, we have applied a non-rigid mutual information (MI)-based registration framework by Heinrich et al. [28], which works in the spatial domain and more importantly known for its robust performance at the time of our study ³. Another

³A fully working coded implementation was kindly provided by lab co-worker, Mattias Heinrich.

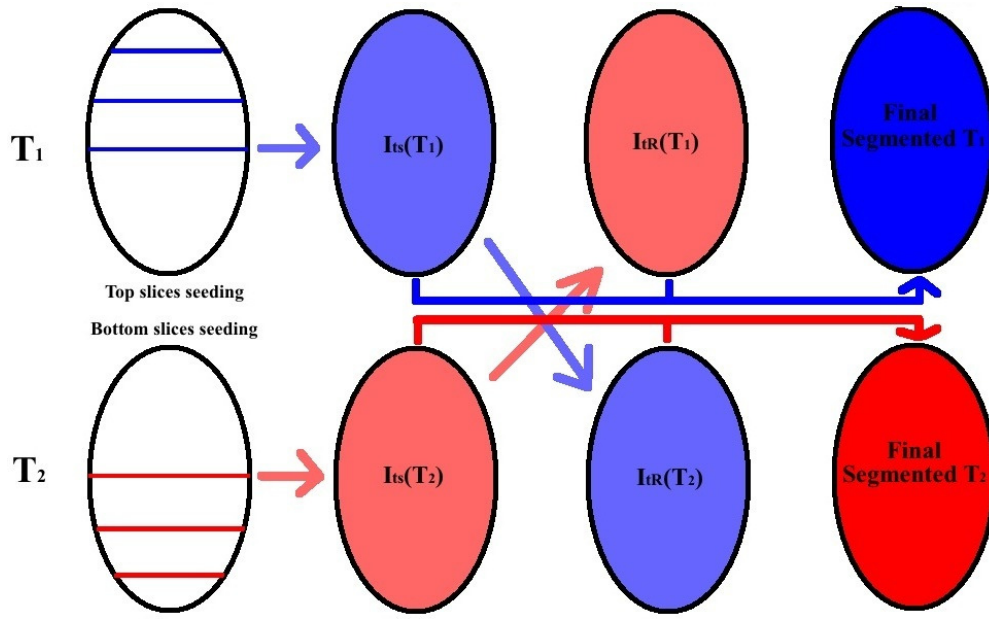


Figure 5.25: Illustration of the registration-assisted algorithm. Note the complete segmentation is not achieved until the final step, which requires the union of the registered and segmented image masks.

motivating factor for our choice was the works presented in Sensakovic et al. [70], where a mutual information-based optimisation scheme is used to allocate the equivalent of a baseline CT section in follow-up scans and demonstrated a clinically validated matching rate of 81.8%.

Starting with a cost function with respect to the deformation field $\vec{u} = (u, v, w)^T$:

$$\operatorname{argmin}_{\vec{u}} = \int_{\Omega} S(I_1(\vec{x}), I_2(\vec{x} + \vec{u}))^2 + \alpha \operatorname{tr}(\nabla \vec{u}(\vec{x})^T \nabla \vec{u}(\vec{x}))^2 d\vec{x} \quad (5.19)$$

where $S(I_1, I_2)$ is the similarity measure of image data I_1 and I_2 . The second term in Eq. 5.19 is a diffusion regularisation term dependent on the deformation field $\vec{u}$. To minimise the cost function, a Gauss-Newton optimisation method is used. More specifically, letting $\nabla S = (\frac{\delta S}{\delta u}, \frac{\delta S}{\delta v}, \frac{\delta S}{\delta w})^T$ and $\Delta \vec{u} = \nabla(\nabla(\vec{u}(\vec{x})))$, the iteration step can be written as:

$$(\nabla S^T \nabla S + \alpha \Delta) \vec{u}_{new} = -(\nabla S^T S + \alpha \Delta \vec{u}_{previous}) \quad (5.20)$$

The free parameter α is set as unity to balance the similarity term with the regulariser. In this method, local normalised mutual information (LNMI) is used as the similarity measure. The

local term refers to the evaluation of the similarity function at each location. It is applied for variational optimisation reasons explained in Rogelj et al. [64]. For an intensity pair $\vec{i} = (i_1, i_2)^T$, the joint probability $p_{12}(\vec{i})$ and marginal probabilities $p_1(i_1)$ and $p_2(i_2)$, LNMI is defined for a location x as:

$$LNMI(\vec{x}) = \log\left(\frac{p_{12}(I_1(\vec{x}), I_2(\vec{x}))}{p_1(I_1(\vec{x})) \cdot p_2(I_2(\vec{x}))}\right) \frac{1}{\int_{\vec{x}} p_1(I_1(\vec{x})) \log(p_1(I_1(\vec{x}))) d\vec{x}} \quad (5.21)$$

where global entropy of I_1 is used for normalisation. The joint and marginal PDFs were previously calculated at each iteration using the kernel density estimator. We, for the consideration of improving its computation time, adopt the 1-D NPW method described in Chapter 3.

The forward ($I(T_1)$ to $I(T_2)$, ‘source’ to ‘target’) and reverse ($I(T_2)$ to $I(T_1)$, ‘target’ to ‘source’) deformation fields are first computed. The cost functions are reduced over the course of iterations according to the optimisation condition (Eq. 5.19). The resulting deformation fields are evaluated using image difference maps in Fig. 5.26, where the pre-registration maps show notable misalignments that are effectively corrected by the registrations, as shown by the registered maps.

With the computed deformation fields, the segmented volumes are mapped onto their corresponding target images. The registered images have a number of interesting characteristics. First, they provide a good source of knowledge to the segmented volume of the target image for regions where initialisation seeds are not found. The confidence of the registered data, however, remains an area of considerable interest. It can be shown here that the clinical accuracy of our reference truth is somewhat controversial, as we have expected. We have stated earlier that the reference truth used in our studies are delineations of the tumour by a well trained clinical expert. However, Moltz et al. [48] suggested that for CT applications, the optimal number of independent observers is three, to ensure the best overall balance between unbiased observer opinion and redundant inter-observer variance. Therefore, Fig. 5.27 shows a part of the neoplasm that has been omitted during the manual segmentation process and thus not included in the reference truth. By registering the corresponding slice from the scan at T_1 to the the one at T_2 , we can see that although the tumour has greatly reduced in size, it still maintains a very

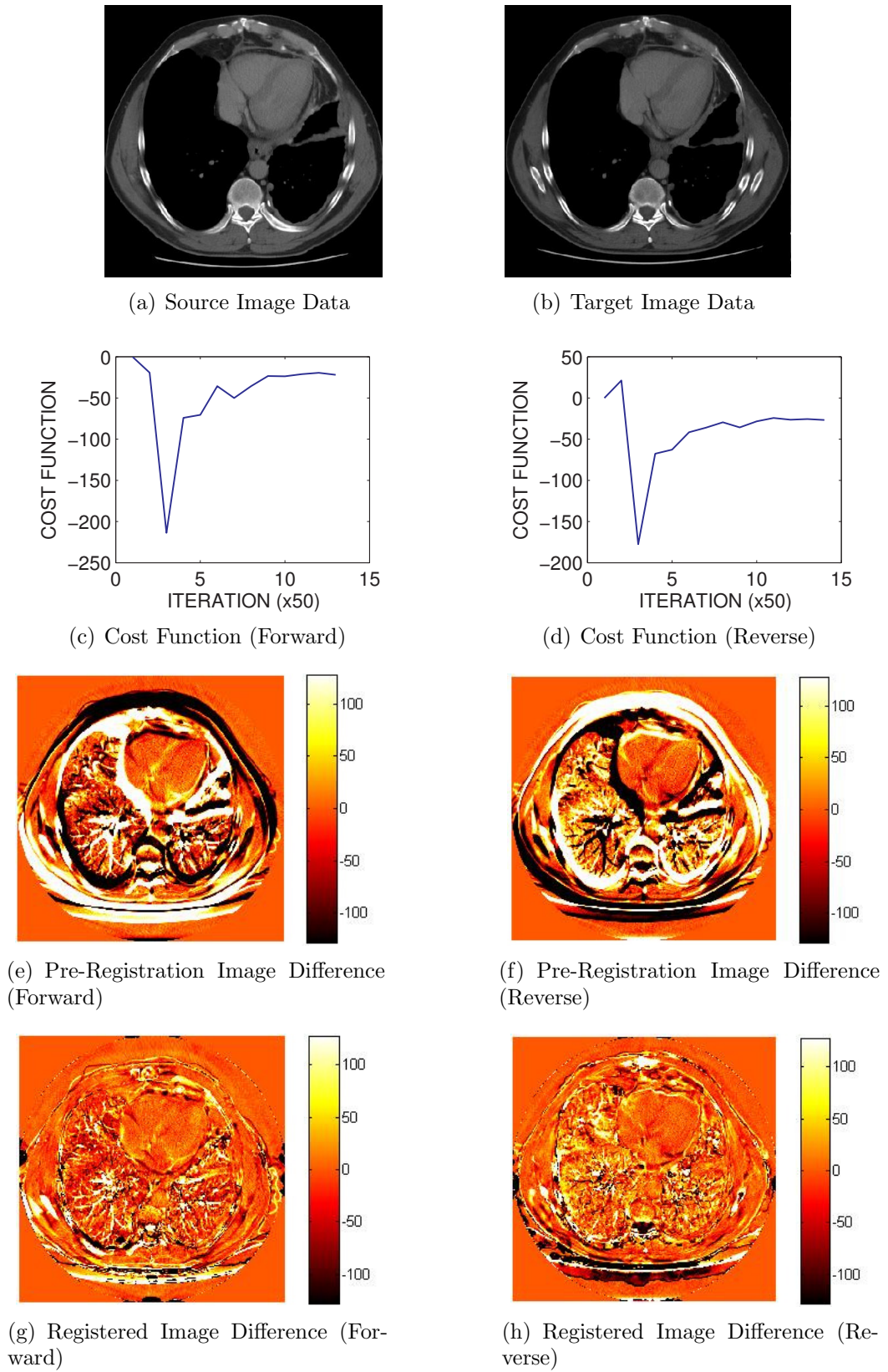


Figure 5.26: Evaluation of the computed deformation fields. Slice samples from the full image data volumes are shown. Note the minimisation of the absolute cost functions over the iterations according to the optimisation condition (Eq. 5.19). The pre-registration difference maps show notable misalignments that are effectively corrected by the registrations.

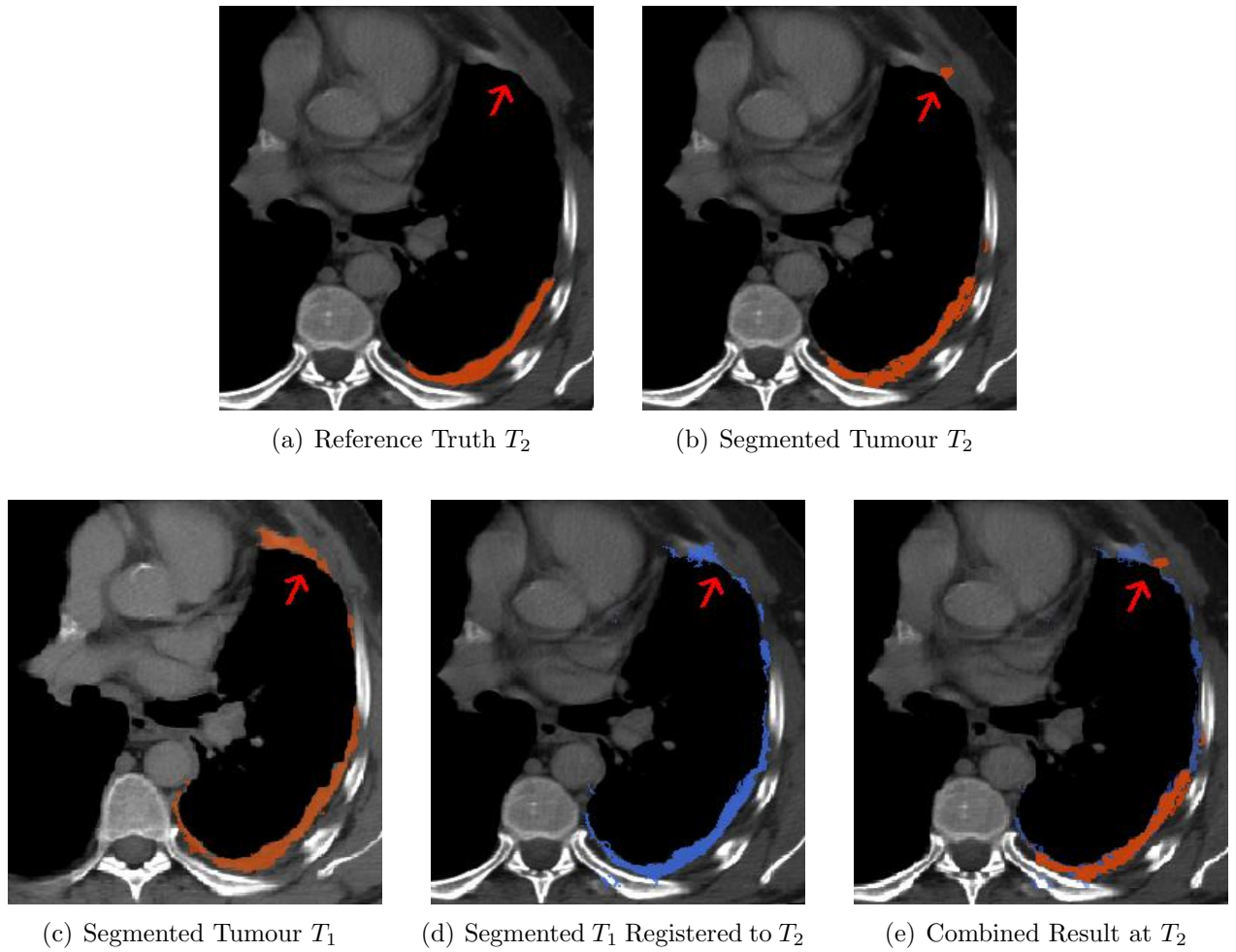


Figure 5.27: Sample slices showing steps in the registration-assist algorithm. Pointed by arrows are places where the reference truth might have been fallible to observer errors. The validity of the thin tumour presence is verified by both the registered T_1 segmentation and volumetric influence from adjacent image slices.

thin presence along the chest walls.

After finding the mapping of segmented volume at T_1 onto the T_2 scan, we can make use of that information to assist in our segmentation. As the accuracy of the mapping is predetermined by seeding of the source, only half of the mapped data is used and superimposed to correct the corresponding section on the segmented volume at T_2 . This process is repeated using the reverse deformation field for correcting segmented T_1 . In both operations, we note that the initialisations used to segmented each of the images are reduced by half. This is the key advantage of our registration-assisted algorithm. Results are presented in Figs. 5.28 and 5.29.

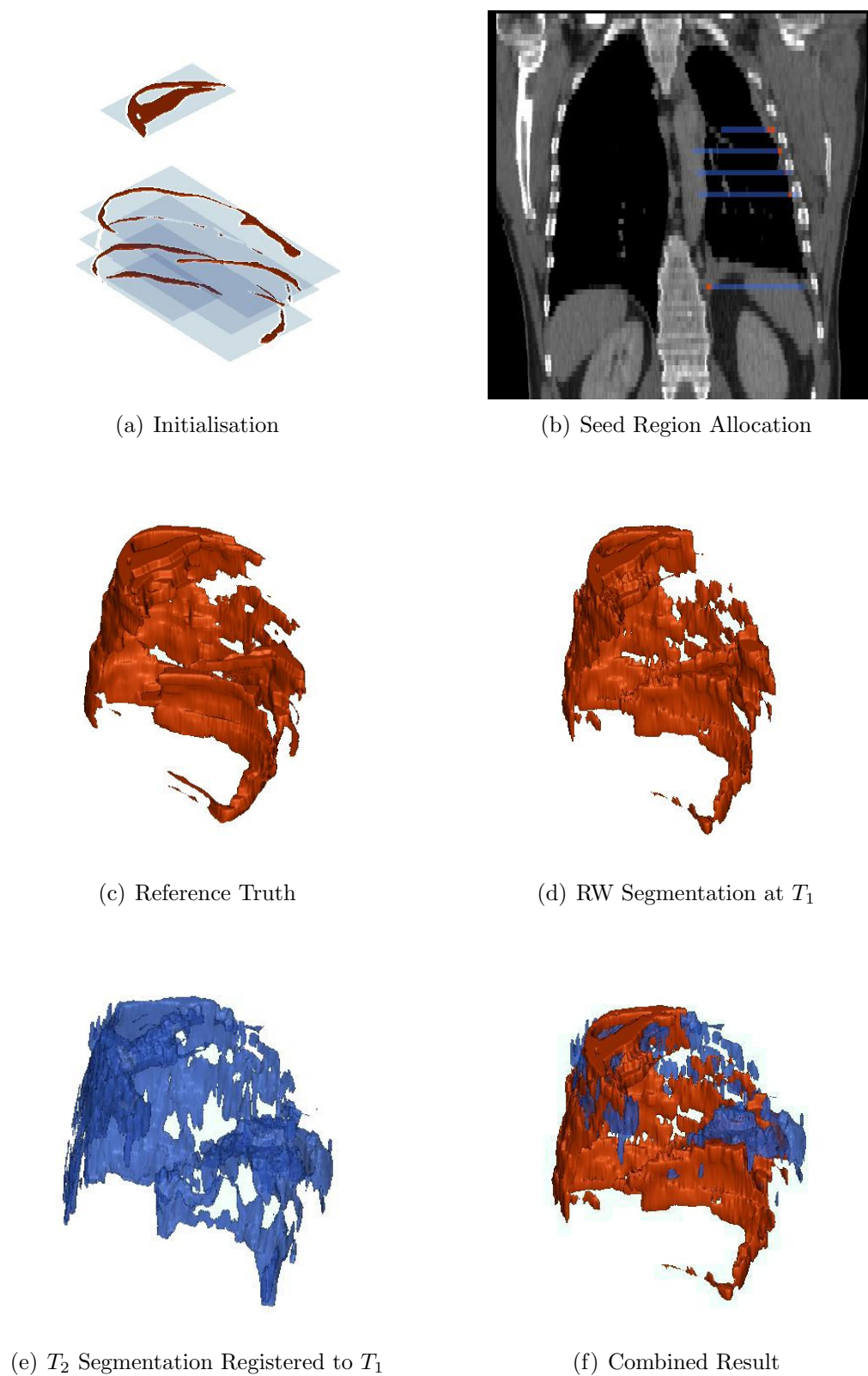


Figure 5.28: Registration-assisted segmentation for the scan at T_1 . Initialisations are only given for the top slices. The mapped image in (e) is used to correct the upper half of the segmented volume at T_1 . Comparing the combined result to the reference truth shows the functionality of our proposed algorithm.

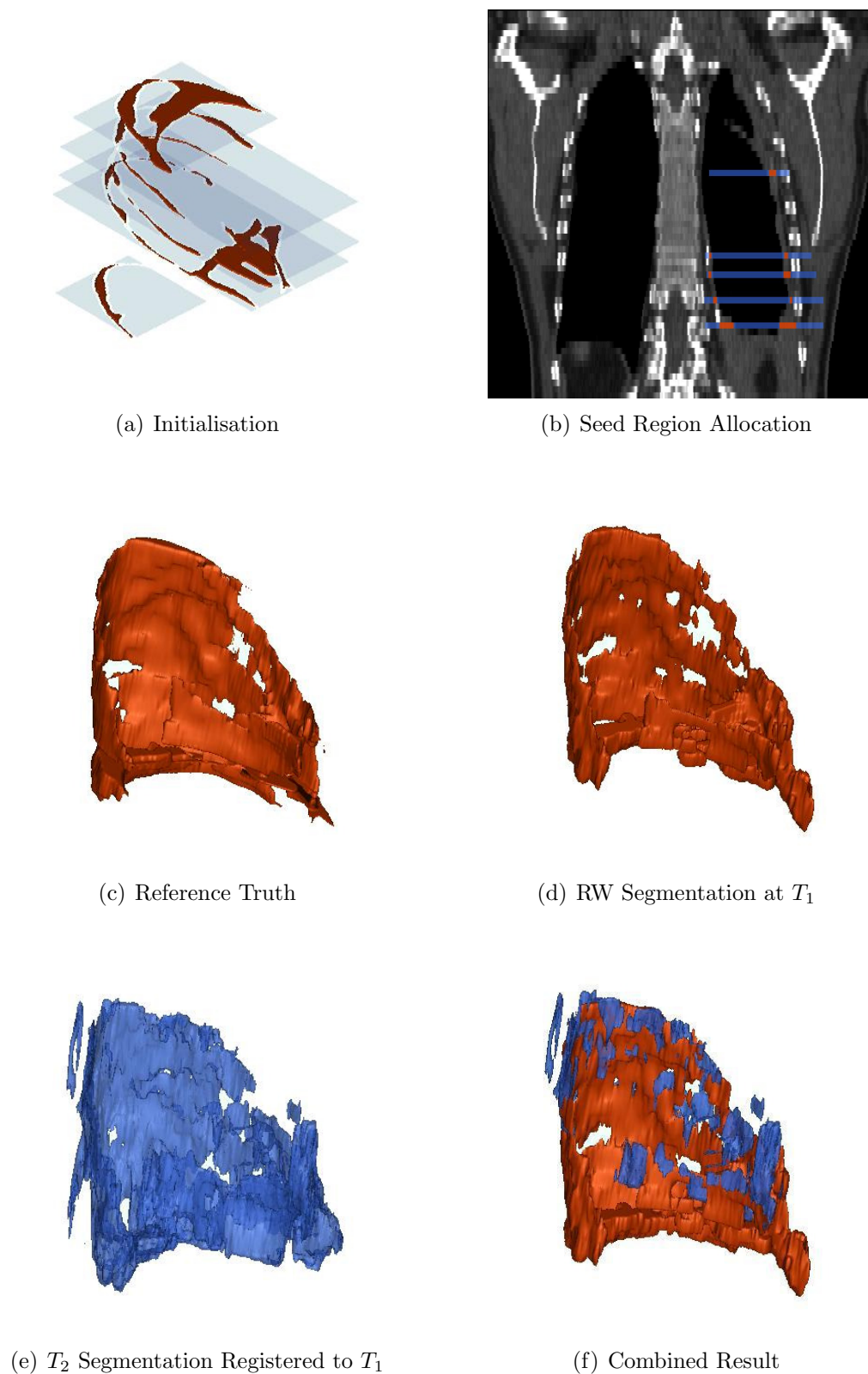


Figure 5.29: Registration-assisted segmentation for the scan at T_2 . Initialisations are only given for the bottom slices. The mapped image in (e) is used to correct the upper half of the segmented volume at T_2 .

Image Data	Segmentation Only	With Deformation Field Computation
T_1	21m29s	54m13s
T_2	18m53s	56m19s

Table 5.3: Computation time comparison of the original and registration-assisted methods (3.39GHz, 2GB RAM). The additional time for deformation field computation can be easily accounted for by the time saved for delineating the manual segmentations on part of the clinician.

The accuracy of our registration-assisted algorithm is evaluated by its DICE measure to reference truth in Fig. 5.30. Keeping in mind the validity issue of the reference truth, we observe a good overall performance of the algorithm on the two datasets. The effect of the registration is obvious in both cases where the overall accuracy improves when the additional information offered by the mapped image data is applied. The computation times of the original and newly proposed methods are benchmarked in Table 5.3. Despite the extra time required for computing the deformation fields, the required level of initialisation to achieve a good overall segmentation is greatly reduced. Moreover, with the inevitable uncertainties for any manual segmented MPM shapes, it might be beneficial to compute the registered volume regardless.

As a final point in this chapter, we present registration as an important tool for visualisation. In order to visualise the responses of the tumour over time in terms of volume shape changes, the segmented volumes must be registered for precise alignments. We have used the same non-rigid registration method to first build a deformation field linking the segmented volumes of a pair of chemotherapy monitoring scans (the one already studied in great detail in this section) and then overlay the aligned post-treatment volume to the pre-treatment one. This allows us to directly view the change of tumour volume over time in geometric terms. All four methods described so far in this section are applied, with their results given in Fig. 5.31, for a direct comparison. These results are not evaluated for accuracy, given our discussions on the reference truth validity. However, the similarity of the four results can be easily seen. Also, by visual inspection, the 8-slice BE-segmentation gives the closest match to the reference truth. However, with less clinical knowledge of the tumour being available, the registration-assisted

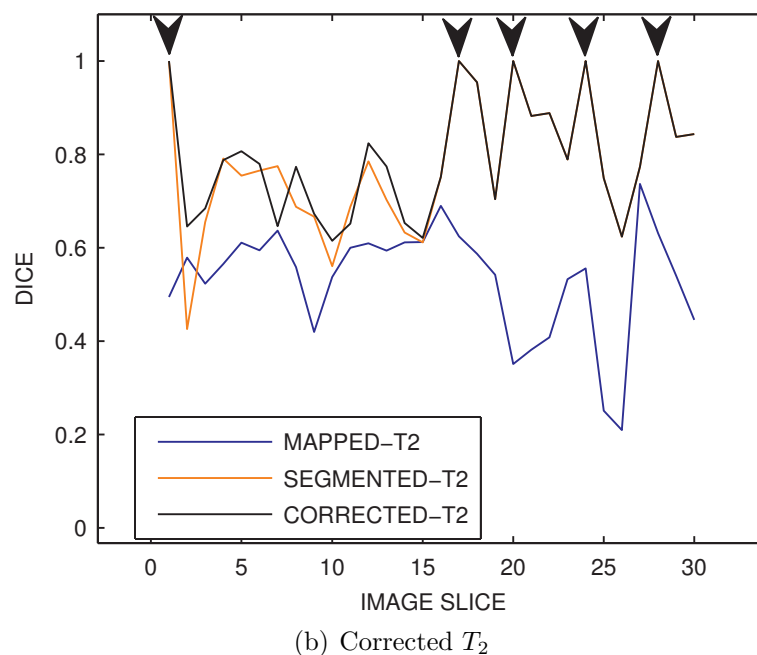
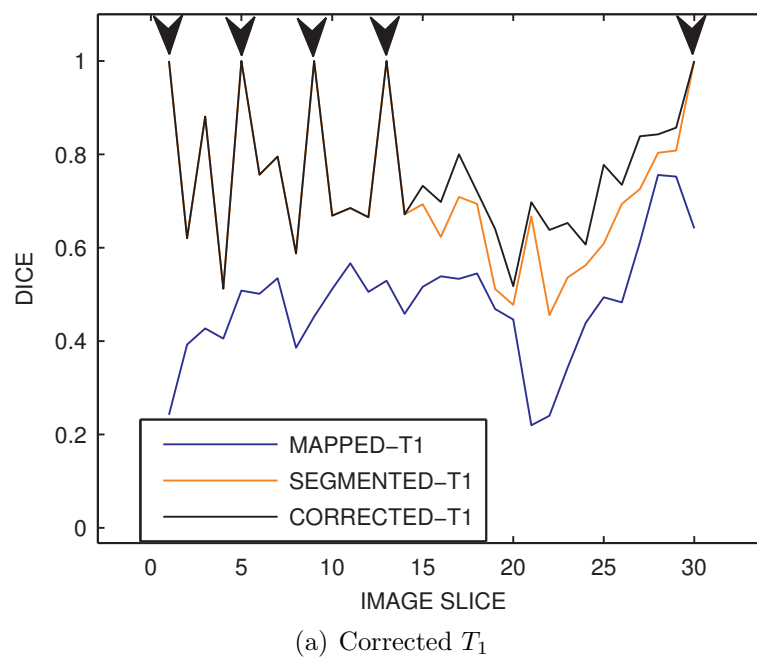


Figure 5.30: Accuracy (DICE) analysis of the registration-assisted algorithm. Note the effect of registration ('Corrected $T_1/2$ ') on enhancing the accuracy of the 'under-initialised' region ('Mapped- $T_1/2$ ') of the scan in both cases, and as compared to the BE-method used on its own ('Segmented $T_1/2$ ').

segmentation yields a better estimate than the 4-slice BE-segmentation.

In summary, for our clinical studies, we used BE-volume segmentation with sufficient initialisation where possible and in cases of uncertainty such as difficult thin tumour boundaries, image registration is applied to assist in the segmentation process.

5.7 Summary and Future Works

In this chapter, we have introduced the random walk-based segmentation method. After initial experiments using the method on a collection of synthetic and medical data, we have discovered the many advantages it offers to our application. First, compared to the methods we have reviewed in Chapter 4, it handles the segmentation of complex geometric shapes very well, even under considerable image noise. In terms of noise robustness, our experiments with the synthetic data of increasing noise clearly show the good performance of random walk in handling image noise of varying degree. The method is also highly automatic, requiring little user control, with only one free parameter as preset. Lastly but not the least, the random walk method is able to segment any arbitrary shapes given sufficient user inputs, a characteristic highly useful given the complexities of MPM segmentation.

Our findings have shown that the original RW-method lacked an adequate level of boundary smoothness for MPM tumour shapes. Therefore, in the spirit of anisotropic diffusion and adding a non-parametric nature to the method, we have implemented a boundary regulariser by updating the weighting function to incorporate NPW estimates of the probabilistic prior estimation. The results of this enhancement are significant and successfully removed the problem of ‘rough’ tissue boundaries.

We then moved onto extending the method into the 3-D domain. This is an essential step in making the method applicable to our clinical application, where we need to assess the volumetric response of the mesothelioma tumour to chemotherapy. The results showed a good performance and an accuracy dependency on initialisation seeds for which we have developed a registration-assisted routine. With this amendment we were able to generate equally accurate segmentation with less initialisation seed regions. This is particularly useful given our discussions on the

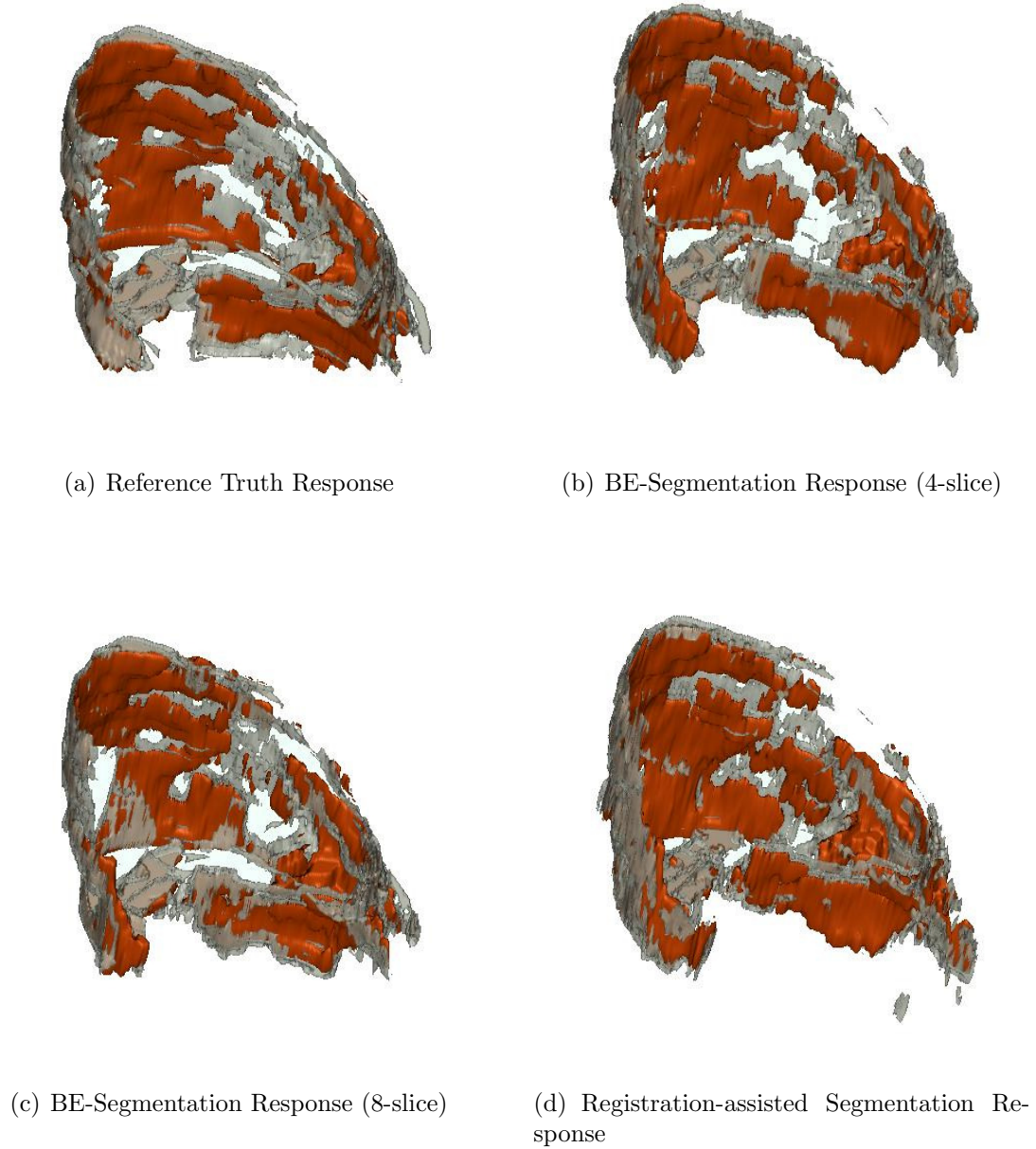


Figure 5.31: Tumour response comparison of the results by four main ways of segmenting MPM. Orange is the tumour at T_2 , white shows the tumour shrinkage since T_1 ; therefore suggesting a possible response to IV-Vinflunine. By visual inspection, the 8-slice BE-segmentation gives the closest match to the reference truth. However, with less manually defined initialisation available to us, the registration-assisted segmentation yields a better estimate than its 4-slice BE-segmentation equivalent.

clinical validity of manual segmentation. In this regard, it is in our interests to minimise the use of excessive amount of manual seeds whilst maintaining a good level of segmentation accuracy.

As future work, a possible improvement to the boundary regulariser can be studied. Better control of the anisotropic diffusion might be achieved by adopting a more efficient iterative solver. This would allow us to examine the matrix solving in greater detail and hence control the iterations in an anisotropic manner with optic flows [87]. Another possibility is the use of subjective contours, a concept proposed by Rutkowski [66], to fill the contour gaps and overcome the issue of random walk ‘leakage’. This idea is discussed further in Chapter 8.

With regards to the volumetric extension, the 26-connected case potentially offers a better controlled flow between the planar image slices. This might improve the segmentation accuracy in between initialisation slices, especially those further away from the seeded regions. Therefore, the implementation of the 26-connected topology is considered as an important future point of interests. Additionally, as we have witnessed in this chapter, image registration is an important component of medical image analysis and our application. It sees wide usage both in providing good segmentation knowledge in the form of mapped volumes and visualisation of tumour responses over time. Currently, we have applied a state-of-the-art deformable framework that best suits our study needs. Though as with any imaging methods, this method still has its limitations; for example in terms of accuracy, where mapped points lands outside of the rib cage of the target image scan. Therefore, it may be worthwhile to investigate further ways of registration and select one that better answers to our clinical needs.

Chapter 6

Partial Volume Effect Correction

So far in this thesis, we have attempted to finesse the issue of partial volume effect (PVE) with the use of morphological erosion. However, it is essential that we have a way to tackle this problem more systematically if we are to achieve segmentations with better accuracy. In this chapter, in an attempt to reduce PVE on our data to facilitate better segmentation results, we review a number of techniques that address PVE in segmentation. Our discussions are primarily focused on the method based on entropy boundaries, the one applied in our segmentation algorithm for mesothelioma (MPM). Based on these discussions, we choose a robust algorithm that best suits our computational needs; adapt and then implement it for our problem.

6.1 Introduction

Due to the discrete sampling nature of medical imaging, voxels may represent either a single tissue type or a combination of such types. The latter situation gives rise to the PVE. When a discretely sampled image is acquired with an anti-aliasing filter, signals at frequencies higher than the Nyquist threshold (half of the sampling frequency) are permanently lost. The absence of high frequency signals causes blurring of the tissue boundaries and lowers the image contrast.

For image analysis applications in which a high degree of accuracy of tissue boundary definition is required, which is certainly true for our application, voxel-level segmentations such

as those introduced in Chapters 4 and 5 are insufficient for clinical purposes. In our study of MPM, the impact of PVE is particularly severe, since the tissue of interest, the MPM tumours often have an elongated shape where extended boundaries with neighbouring tissues are present.

Given any shape or region of interest (ROI) (Fig. 6.1), an error measure can be established based on the ratio of the affected shape boundary (B) and PVE in place, as given in Eq. 6.1.

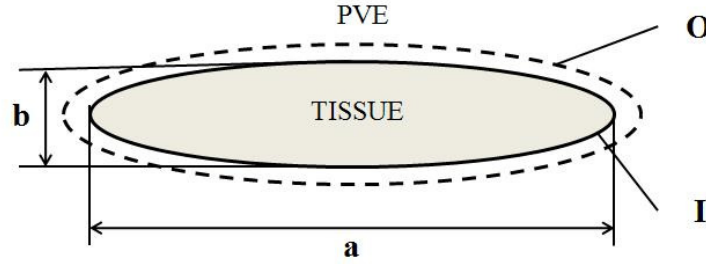


Figure 6.1: A sample PVE-affected tissue, elongated to match the case where the PVE becomes more pronounced, as found in MPM data. Boundary/PVE ratio decreases with an increase in ratio a/b , a and b being the two dimensions of the longitudinal figure.

$$\frac{B}{PVE} = \frac{O - I}{1/2(I + O)}\% \quad (6.1)$$

The shape with the lowest such ratio are those which are elongated, as given in Fig. 6.1. In fact, the RECIST criteria explicitly report a much reduced functionality when applied to longitudinal malignancies such as MPM [19]. For the segmentation of MPM, this problem becomes highly pronounced as the B/PVE ratio increases to beyond 40%. One such example is given in Fig. 6.2. It is thus essential that we find an effective way to estimate voxels affected by PVE, so as to render our segmented tumour data clinically useful.

Most of our image data voxels are not isotropic. This means the PVE is more dominant in the axial direction than planar-wise on transverse scans. To illustrate this point, the coronal view of a thick-slice (5mm) scan is given in Fig. 6.3. It is therefore important that we take the axial PVE into consideration and perform PVE correction in all three dimensional directions.

In this chapter, an overview of some available methods to partial volume estimation is given in Section 6.2. The choice of the estimator will be based on a number of criteria such as

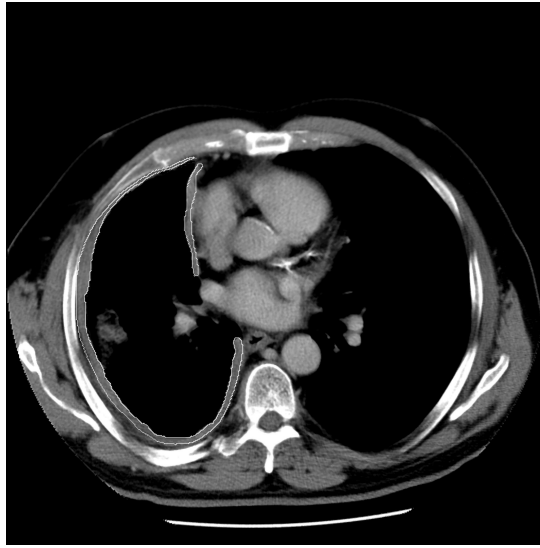


Figure 6.2: PVE-affected boundary voxels in a MPM scan. Note the longitudinal shape of the primary lesion hence high level of uncertainty (uncertainty-to-area-ratio to the segmentation due to PVE).

accuracy, computational requirement, as well as the ease of implementation. These factors will be weighed against each other in our decision-making process and the method that best suits our need will be selected for use in our application. Experimental results from the estimator of choice are then given and discussed, with respect to its links to our clinical targets, in Section 6.3. In response to axial PVE, the method is extended to the 3-D space in Section 6.4. The extra running times incurred from the PVE-correction scheme are analysed in Section 6.5, with discussions on possible ways of optimising the overall computational performance of a PVE-corrected algorithm. We conclude this chapter in Section 6.6 along with a description of possible future works on this topic.

6.2 Partial Volume Estimators

6.2.1 Overview

To begin our discussion on partial volume estimation, we first suggest the requirements of an estimator that would suit our needs. In MPM, there are a number of key tissues present, so clearly we need a multi-class model, that is, voxels can either represent ‘pure tissues’ or a

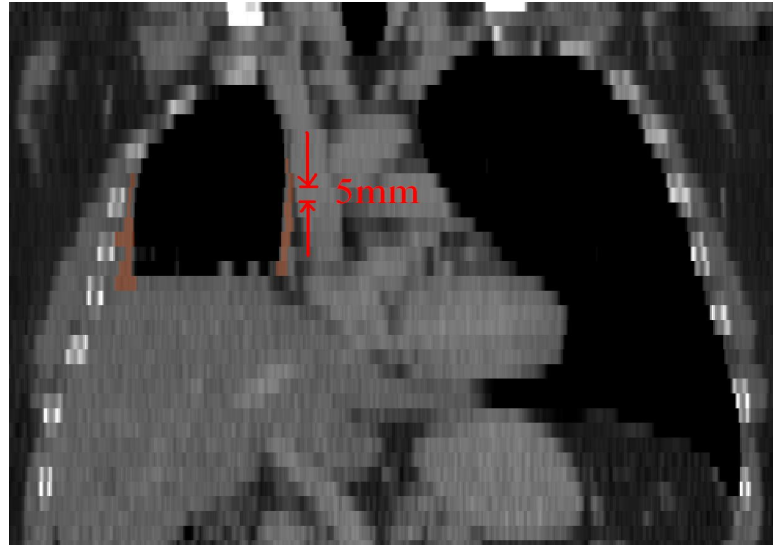


Figure 6.3: Coronal scan of a thick-slice (5mm) scan showing the dominant axial PVE. Note the overwhelming PVE in the axial direction, which severely ‘blurs’ the major tumour boundaries and affects the accuracy of the manual segmentation. Tumour (reference truth) is highlighted in *orange*.

combination of two or more tissues. Unlike in the cases of MR, for CT, we will not consider K-map blurring which leads to the spread of PVE to non-boundary voxels. Additionally, given the complexity and probabilistic overlapping nature of MPM tumour and surrounding tissues, the method must be robust to difficult image problems.

It should be noted that most of the literature on PVE classification is centred on MR data. This is in part due to the lower spatial resolution of MR with respect to CT; but primarily because, unlike CT, MRI is mainly used to image regions of high soft tissue heterogeneity, whose boundaries are susceptible to the influence of PVE and tend to require the use of PVE estimators. CT, on the other hand, is generally applied to regions of distinct bone to soft tissue contrast. But for reasons explained earlier, CT scans of MPM are a notable exception to this. Therefore it is important for the method to have good applicability to CT image data.

6.2.2 Gaussian Mixture Model

A PVE estimator has been proposed in [83] for brain CT data, that makes use of the Gaussian Mixture Model (GMM). The probability density is assumed to take the form of GMM, with x being the intensity of individual pixels:

$$f(x) = \sum_{i=1}^N \pi_i f_i(x) \quad (6.2)$$

where $f_i(x)$ are the GMM components, π_i are the weights such that $\sum_{i=1}^N \pi_i = 1, 0 \leq \pi_i \leq 1$. The number of components is chosen to be equal to that of pure tissue classes present. The distributions of the f are assumed to be $f_i(x) = N(x|\mu_i, \sigma_i)$ where N is the Gaussian PDF with mean μ and standard deviation σ :

$$N(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (6.3)$$

An Expectation-Maximisation (EM) algorithm is employed to estimate the parameters for individual GMM components, which are subsequently used to find the distributions of the boundary, or PVE pixels. For example, given a voxel for which α denotes the proportion that is tissue 1 and $1 - \alpha$ the proportion of other tissues, the PVE distribution would then be a convolution of the PDF of tissue 1 and other tissues, denoted as f_1 and f_2 respectively:

$$g(z) = g((1 - \alpha)x + \alpha y) = \int_x f_1(z - (1 - \alpha)x) f_2(\alpha x) dx \quad (6.4)$$

where $z = (1 - \alpha)x + \alpha y$ such that x and y are independent random variables with Gaussian distributions. Combining Eqs. 6.2-6.4, the parameters, μ_g and σ_g for the PVE distribution, which is itself a Gaussian, we would have:

$$\mu_g = (1 - \alpha)\mu_1 + \alpha\mu_2 \quad (6.5)$$

$$\sigma_g = \sqrt{(1 - \alpha)^2\sigma_1^2 + \alpha^2\sigma_2^2} \quad (6.6)$$

Given these equations, we may calculate α using the estimated μ_g , μ_1 and μ_2 , which can be found based on the manually segmented reference truth for tissues 1, 2 and boundary voxels. A threshold value T_{PVE} is then defined, equal to the most probable value of PVE voxels, μ_g . A mask is created containing an one-contour layer of pixels immediately inside and outside of the boundary voxels. Amongst those which are masked, voxels with intensities below the threshold are classified as tissue 1. The other voxels will be passed onto an additional classification scheme based on values of its neighbours. For a masked voxel q of intensity x , a function is defined as follows,

$$e(x) = \begin{cases} 0, & \text{if the pixel } q \text{ is outside the mask,} \\ \max(\frac{\alpha N(x|\mu_1, \sigma_1)}{N(x|\mu_1, \sigma_1)}, \xi), & \text{otherwise.} \end{cases} \quad (6.7a)$$

$$(6.7b)$$

This describes the likelihood of the voxel being a PVE voxel. ξ specifically defines a possible share of tissue 1 in the voxel, and is calculated in a fashion similar to α . Given three neighbouring voxels p, q and r with intensities $g(p), g(q)$ and $g(r)$: If p is already marked as tissue 1 whereas q and r are not, ξ is found from $g(q) = (1 - \xi)g(r) + \xi g(p)$. Otherwise, $\xi = 0$. When $\xi \geq \alpha$, $g(q) \leq (1 - \alpha)g(r) + \alpha g(p)$ is satisfied, so q is assigned to tissue 1.

Finally a snake algorithm is applied to smoothen the contour.

When applied to brain CT data, the results [83] showed segmentation results with smooth boundaries representing CSF, WM and GM of the brain. There are a few things that we should note in this case. First, brain CT data is a very simple segmentation case with distinct separation of different tissue distributions. The overlapping and ambiguity present in the MPM data, as shown in Chapter 3, may prevent the method from performing well in our application. Additionally, small regions of PVE ($\sim \leq 1\%$ of the total data) may sometimes be disguised by image noise which typically means that the independent PDF for the PVE is effectively unattainable. Moreover, the intrinsic reliance on reference truth by the method means it is highly susceptible to uncertainties caused by inter- and intra-observer variabilities. The problem of finding an optimal solution therefore becomes a matter of user-judgement and introduces a great deal of subjectivity.

Despite these concerns, this method works for many scenarios in medical imaging and forms the backbone of PVE classifiers found in generic image processing toolkits such as SPM99. However, it is based on an Gaussian assumption, which does not hold for our application, as we have explained in Chapter 3. For these reasons, we introduce a second PVE estimator in the next section, which is void of this assumption and offers better performance in many regards.

6.2.3 Bayes' Theorem Mixture Model

Ballester et al. [5] proposed a method that relies on the Bayes' theorem:

$$p(\alpha^i | I^i) = \frac{p(I^i | \alpha^i) p(\alpha^i)}{p(I^i)} \quad (6.8)$$

where I^i is the intensity at pixel i , otherwise notations are the same as in the previous section, such that $I^i = \alpha^i I_1^i + (1 - \alpha^i) I_2^i$. A mixture model is constructed for yielding the overall distribution of the image for a given α and distributions of individual tissues. For the two-tissue case, this idea is illustrated in 6.4.

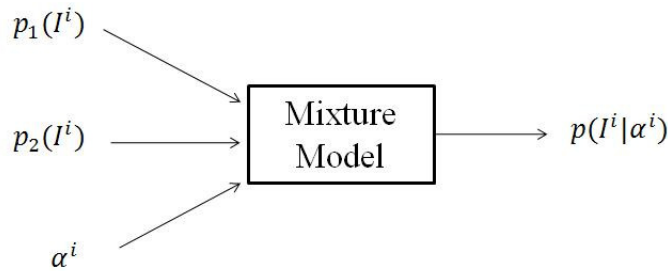


Figure 6.4: Bayes' theorem mixture model. The mixture model outputs the conditional probability of the full mixed voxel intensity based on the probabilistic distributions of individual component tissues and proportion of each tissue in the mixed voxel.

An expression for the model can be derived analytically as follows. In Eq. 6.8, the normalising constant is defined as $p(I^i) = \int p(I^i | \alpha^i) p(\alpha^i) d\alpha$. $p(\alpha^i)$ is simply the prior on α where, for a masked region of only boundary voxels, a uniform distribution, $p(\alpha^i) = 1, \forall i$, is assumed.

Furthermore, $p(I^i|\alpha^i)$ is taken to be independent across all voxels; that is, the distribution of intensity depends only on that of p_1 , p_2 and α , without any influence from neighbouring pixels. Therefore, given the distributions of individual tissues and the analytically computed $p(I^i|\alpha^i)$, it is possible to find $p(\alpha^i|I^i)$ from Bayes' Theorem and the above assumptions. The most probable value of α is in this case the maximum of the PDF, denoted by α_{mode} . $p(\alpha^i|I^i)$ then give the upper (α_{lower}) and lower bounds (α_{upper}) at a particular confidence level, which can be used to create a mask of the boundary PVE voxels.

More specifically, this can be achieved by first choosing a confidence interval (CI) and then applying:

$$\int_{\alpha_{lower}^i}^{\alpha_{mode}^i} p(\alpha^i|I^i)(\alpha)d\alpha = \int_{\alpha_{mode}^i}^{\alpha_{upper}^i} p(\alpha^i|I^i)(\alpha)d\alpha = 1/2CI \quad (6.9)$$

The final PVE volume is estimated as:

$$V = v_{voxel}(n_{pure} + \sum_{i \in PVE} \alpha^i) \quad (6.10)$$

where v_{voxel} is the volume of a voxel and n_{pure} the number of pure voxels. Upper and lower volume estimates can be found by replacing the V and α terms in Eq. 6.10 with their upper/lower bound equivalents.

Alternatively it is possible to generate the distribution of V using the Monte Carlo method [5], which however can be computationally expensive.

The author [5] applied the method to a range of applications, including synthetic and brain MR data and has shown the benefit of applying his method to segmentation problems; in the synthetic data case offering an improvement of two orders of magnitude in the spatial resolution of the data. The main challenge to adapting the method for our use, however, is its implementation, which requires a great number of complex geometric visualisations and consideration of numerous scenarios. Since the code was not available and would require substantial time to

implement correctly, we have decided against its use.

6.2.4 Hidden Markov Random Fields

Given that we are not considering the effect of K-map blurring, it is sensible to make the Markov assumption; that is, only the neighbouring pixels have an effect on any arbitrary pixel in the image.

Zhang presented in his thesis [89] an HMRF-based algorithm to address the PVE. Following the steps given in [89], we define two random vectors $\vec{Y}$ and $\vec{Z}$, based on the index set $S = \{1, 2, \dots, N\}$ and which are respectively m - and k -dimensional. $\vec{Z}$ satisfies the condition $\sum_{\ell} \vec{Z}_{\ell} = 1, \ell \in L$ where $L = \{1, \dots, k\}$ is the set of labels. Specific configurations of $\vec{Y}$ and $\vec{Z}$ are defined as $\vec{y}$ and $\vec{z}$, respectively.

It is assumed that given $\vec{z}$, then $\vec{y}$ follows a Gaussian distribution with mean $\vec{E}_i$ and covariance Ψ_i such that,

$$\Psi_i = \sum_{l \in \ell} z_{il}^2 \psi_l \quad (6.11)$$

where the covariance matrix for class l is ψ_l . Finally the probability distribution of y_i can be written as,

$$p(\vec{y}|\vec{z}) = g(y_i; \Sigma_i, \Psi_i) = (2\pi)^{-m/2} |\Psi_i|^{-1/2} \exp\left(-\frac{1}{2}(\vec{y}_i - \Sigma_i)^T \Psi_i^{-1} (\vec{y}_i - \Sigma_i)\right) \quad (6.12)$$

Here $\vec{y}$ is the image and $\vec{z}$ the PVE classification. A random field Z can then be modelled as an MRF with probability distribution,

$$P(\vec{z}) = \frac{1}{M} \exp\{-U(\vec{z})\} \quad (6.13)$$

with M and ζ being the normalising factor and neighbourhood parameter. We then subsequently define the following:

$$U(\vec{z}) = \zeta \sum_{c \in C} V_c(\vec{z}) \quad (6.14)$$

where $V_c(\vec{z}_i, \vec{z}_j) = ||\vec{z}_i - \vec{z}_j||^2$. Conditional independence is assumed in that case such that for any $\vec{z}$, the random variable $\vec{Y}_i$ satisfies the condition:

$$P(\vec{y}|\vec{z}) = \prod_{i \in S} P(\vec{y}_i|\vec{z}_i) \quad (6.15)$$

The Gaussian Hidden Markov Random Field (GHMRF) is then found from the marginal probability distribution of $\vec{Y}_i$:

$$p(\vec{y}_i|\vec{z}_i, \theta_i) = \int_{\vec{z}} p(\vec{y}_i, \vec{z}_i|\vec{z}, \theta_i) d\vec{z} = \int_{\vec{z}} g(\vec{y}_i; \theta_i) p(\vec{z}_i|\vec{z}) d\vec{z} \quad (6.16)$$

Each voxel is subsequently assigned to a mixel $\vec{z}_i$ such that $\vec{z}_i = [z_i^1, \dots, z_i^k]^T$ so that it contains a component for each of the k classes represented by the image. The sum of all components adds up to unity. $\vec{y}_i$ on the other hand defines the greyscale intensity of voxel i .

A Gaussian distribution is assumed in this case for all the tissue classes as well as individual voxels. Additionally, the assumption of piecewise smoothness is made. Finally an MRF-MAP algorithm is used to compute the result.

$$\hat{\vec{z}} = \arg \max_{\vec{z}} \{P(\vec{y}|\vec{z})P(\vec{z})\} = \arg \min_{\vec{z}} \{U(\vec{y}|\vec{z})U(\vec{z})\} \quad (6.17)$$

Incorporating Eq. 6.14 and definitions of mean $\vec{E}_i$ and covariance Ψ_i , the final function for computing $\vec{z}$ is found to be:

$$\hat{\vec{z}} = \arg \min_{\vec{z}} \sum_{i \in S} \left[\frac{1}{2} (\vec{y}_i - \vec{E}_i)^T \Psi_i^{-1} (\vec{y}_i - \vec{E}_i) - \log(\sqrt{|\Psi_i|}) + \zeta \sum_{j \in N_i} \|\vec{z}_i - \vec{z}_j\|^2 \right] \quad (6.18)$$

This minimisation problem is then solved by adopting a sampling method that only allows changes towards a lower energy state. This ensures that each iteration either reduces the energy, or leaves it constant. An Expectation-Maximisation (EM) algorithm is also used alongside such that an initial estimation of $\theta^{(0)}$ is first made, which is used to compute the log-likelihood in the **E-Step** such that,

$$Q(\theta|\theta^{(t)}) = \epsilon [\log P(\vec{y}, \vec{z}|\theta|\vec{y}, \theta^{(t)})] = \int_{\vec{z}} P(\vec{z}|\vec{y}, \theta^{(t)}) \log p(\vec{y}, \vec{z}|\theta) d\vec{z} \quad (6.19)$$

Then in the **M-Step**, $Q(\theta|\theta^{(t)})$ is maximised such that $\theta^{(t+1)} = \arg \max_{\theta} Q(\theta|\theta^{(t)})$. The E- and M- steps continue until a satisfactory result is achieved. The author recommended that a discrete segmentation be first used to produce a likelihood probability for each voxel, which will then be used to start the PVE classification.

To assess the performance of the algorithm, the author [89] compared the method to the GMM-based algorithm in terms of accuracy and robustness measures. They found that at different noise levels, their method always outperforms GMM, as indicated by a smaller mean square error (Eq. 3.6). Robustness was examined by testing the method on different images of the same brain volume. A smaller difference between segmentations would imply a better robustness. In this case, the average difference is found to be 50% that of the GMM-based method.

When applied to CT data, the vector $\vec{z}$ can be reduced to a scalar with the absence of the multi-spectral nature in CT. However, as the algorithm is primarily designed for MRI and that our MPM data are much harder to effectively segment, we will examine a fourth PVE estimator in the following section.

6.2.5 Interpolation by Reverse Diffusion

A more straightforward and logical approach to solving for PVE is by using interpolation. Usually linear and cubic methods are employed but their results are somewhat lacking in both accuracy and robustness. Salvado et al. [67] have presented a novel interpolation method by reverse diffusion (RD). This refers to a flow direction of low to high greyscale intensity gradient. Compared to traditional ways of interpolation such as bilinear or bicubic spline, reverse diffusion relieves the band-limitedness requirement on the interpolated signal. This is particularly true for PVE voxels where step edges with an infinite bandwidth are typically observed. Overall, this method solves for PVE by recovering the high frequency information lost in the anti-aliasing process.

In this method, voxels are first interpolated into subvoxels, or *subbins*. Subvoxels are then subject to *diffusive flows* whose magnitude and direction are determined by neighbouring subvoxels, in the form of anisotropic diffusion. As a result, boundary voxels are expected to produce better edge resolution with a reduced effect of PVE blurring. During the diffusion process, flow is conserved within voxels; that is, no flows would take place between full voxels. For the uni-dimensional case, the working of the method is illustrated in Fig. 6.5.

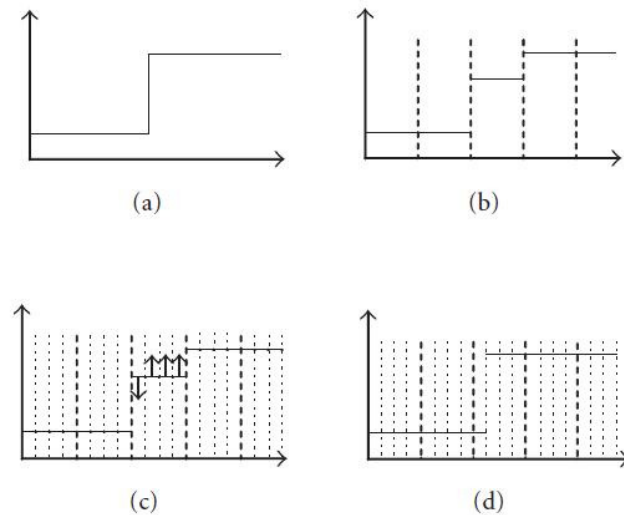


Figure 6.5: Steps in the interpolation-based PVE reduction method for the unidimensional case. (a) shows the original image signal, which is interpolated and upsampled in (b), then flow directions are found in (c) and the final PVE-corrected image is given in (d) [67]

In this case, the observed signal is denoted $\tilde{y}_l$ with l being the index of the bin. The interpolated signal is represented by y_i where i is the subbin index. An iterative step is defined such that:

$$y_i^{t+1} = y_i^t - a[(y_{i+1}^t - y_i^t) + (y_{i-1}^t - y_i^t)] \quad (6.20)$$

where the constant $a > 0$ adjusts for the speed of the flow. Rearranging Eq. 6.20 gives:

$$\frac{dy_i}{dt} = a' \frac{d^2 y_i}{di^2} \quad (6.21)$$

which is the diffusion partial differential equation when $a' > 0$. In this case, however, reverse diffusion is needed so a positive a' is applied. This usually leads to unstable behaviour. Therefore to ensure numerical stability, the flow must be constrained with an upper and lower bounds. Specifically for N_i neighbour subbins, the constraints are found as:

$$Q_i^{\max} = \frac{\max_{i' \in N_i}(y_{i'}) - y_i}{2}; \quad Q_i^{\min} = \frac{y_i - \min_{i' \in N_i}(y_{i'})}{2} \quad (6.22)$$

Here the flows are bi-directional and so they are divided by two. The overall flow computation and iterative steps can be written as:

$$Q_{i,i+1} = \max[-Q_i^{\max}, -Q_{i+1}^{\min}, \min[(Q_{i+1}^{\max}, Q_i^{\min}, y_{i+1}^t - y_i^t)]] \quad (6.23)$$

$$y_i^{t+1} = y_i^t - Q_{i,i+1}; \quad y_{i+1}^{t+1} = y_{i+1}^t - Q_{i,i+1} \quad (6.24)$$

Note that Eq. 6.23 restricts inward flow when $y_i > y_{i+1}$, in accordance to the reverse diffusion criterion.

For 2D images, 8-connected voxels are considered. The same notations are employed with the addition of R such that every voxel is divided into R^2 subvoxels. Naturally, out of the 9

members in the connected ‘patch’, the lowest and highest components would be taken as the lower and upper constraints. But in order to consider at least 3-connected voxels as edges, they are taken as the 4th and 6th highest members instead; this is shown in Fig. 6.6.

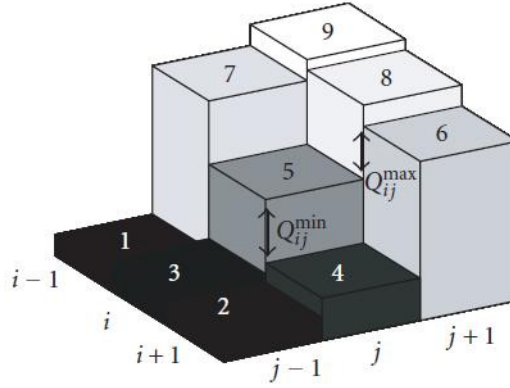


Figure 6.6: Depiction of the 8-connected neighbourhood of a subvoxel [67]

The bounds are thus given by:

$$Q_{ij}^{\max} = \frac{\text{ord}_{\{i',j'\} \in N_{ij}}(6, y_{i'j'}) - y_{ij}}{4}; \quad Q_{ij}^{\min} = \frac{y_{ij} - \text{ord}_{\{i',j'\} \in N_{ij}}(4, y_{i'j'})}{4} \quad (6.25)$$

It has been suggested [67] that a Gaussian kernel be applied to the interpolated data to remedy for the problem of noise-induced aggregation of regional intensity clusters: $\hat{Y} = Y \otimes G(0, R/2)$ where the standard deviation is chosen to be half of the interpolation scale. The horizontal and vertical flow equations are simple extensions of Eq. 6.22 into the 2D space.

$$Q_i = \max[-Q_{ij}^{\max}, -Q_{i+1j}^{\min}, \min[(Q_{i+1j}^{\max}, Q_{ij}^{\min}, \hat{y}_{i+1j}^t - \hat{y}_{ij}^t)]] \quad (6.26)$$

$$Q_j = \max[-Q_{ij}^{\max}, -Q_{ij+1}^{\min}, \min[(Q_{ij+1}^{\max}, Q_{ij}^{\min}, \hat{y}_{ij+1}^t - \hat{y}_{ij}^t)]] \quad (6.27)$$

Finally, to prevent ‘overflow’ and satisfy the voxel intensity conservation assumption, a term B_{ij} is introduced into the iterative equations which is 0 when $i \bmod R = 0$ or $j \bmod R = 0$

and 1 otherwise.

$$y_{ij}^{t+1} = y_{ij}^t - B_{ij}(Q_i + Q_j); \quad y_{i+1j}^{t+1} = y_{i+1j}^t + B_{ij}Q_i; \quad y_{ij+1}^{t+1} = y_{ij+1}^t + B_{ij}Q_j \quad (6.28)$$

The method iterates through Eqs. 6.25 to 6.28 for all voxels until the total flow converges to zero.

Notable features of this method include its independence from PDF estimation and its unsupervised nature, without the need to specify the number of iterations. In their paper [67], Salvado et al. applied the method to a series of synthetic and medical images and analysed the results by comparing them to those from other interpolation-based PVE optimisers such as cubic splines. The method was found to be insensitive to noise and superior performance was observed across all experimental studies. The authors have further suggested the method as a generic method for PVE reduction, applicable to all medical imaging modalities including x-ray and CT.

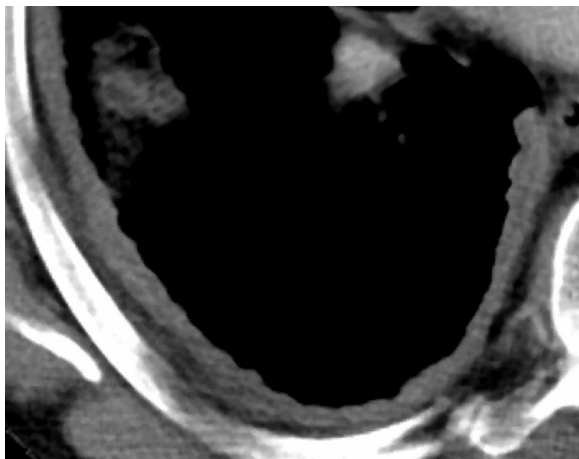
For this very reason of generality, we have applied this method¹ to our MPM data and present the results in the next section.

6.3 Results and Discussion

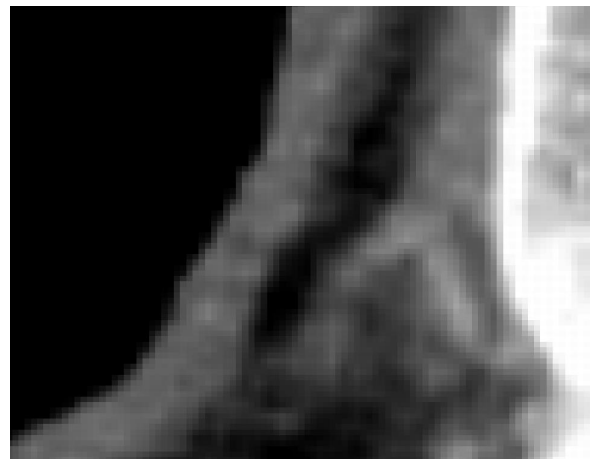
Aside from the convergence tolerance level, the only free parameter in the RD interpolation method is the scaling factor. We investigated the effect of varying its value by applying the method to MPM with a series of scaling factors. Interpolated MPM data (from Fig. 6.2) with scaling factors (s.f) of 2 and 4 are given in Fig. 6.7.

The total flow measure is used to control the iteration. The reverse diffusion process is stopped when the total flow reaches zero. Entropy is shown as an additional measure to indicate the gradual decrease of random uncertainty over the iterations. The probabilistic terms in the entropy calculation are estimated using NPW. Both plots, for the case of interpolation with s.f of 2, are given in Fig. 6.8.

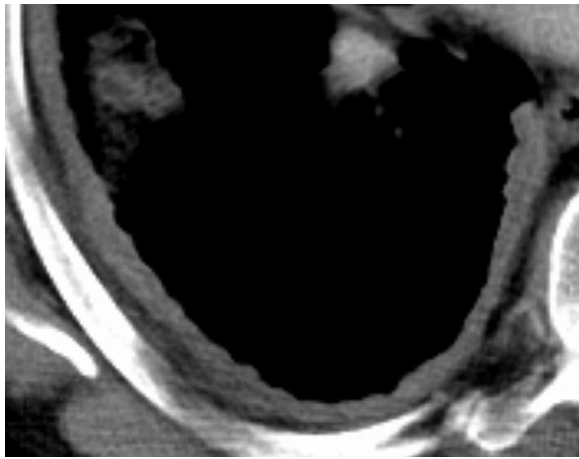
¹An implementation of the method was kindly provided by the one of the original authors, Olivier Salvado



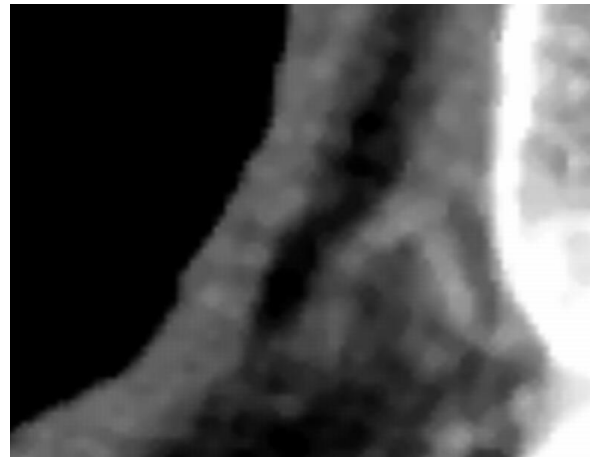
(a) Original Image (x5)



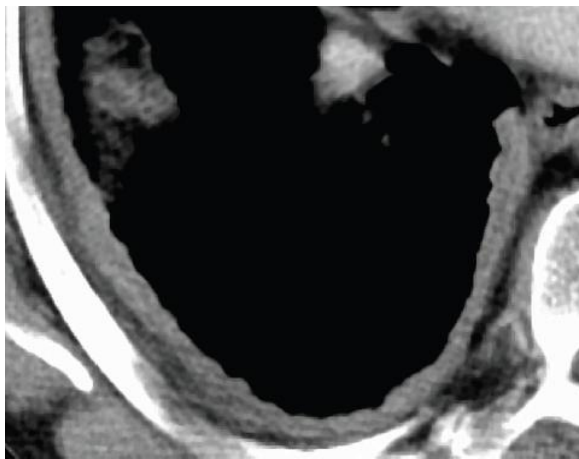
(b) Original Image (x20)



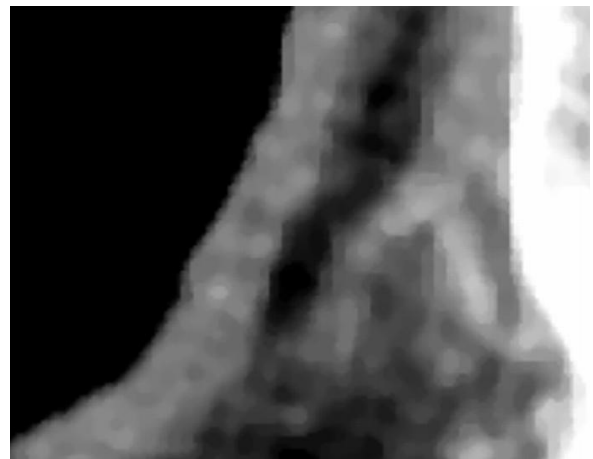
(c) Interpolated s.f:2 (x5)



(d) Interpolated s.f:2 (x20)



(e) Interpolated s.f:4 (x5)



(f) Interpolated s.f:4 (x20)

Figure 6.7: MPM data interpolated with reverse diffusion for PVE reduction. The interpolation results are shown magnified 5 and 20 times. Also note the ‘smoothing’ effect with the interpolated data and that at high scaling factor (i.e: 4), some image features may be ‘blurred’ out. Therefore, we must optimise the choice of s.f in the experiments, both for accuracy and computational time considerations.

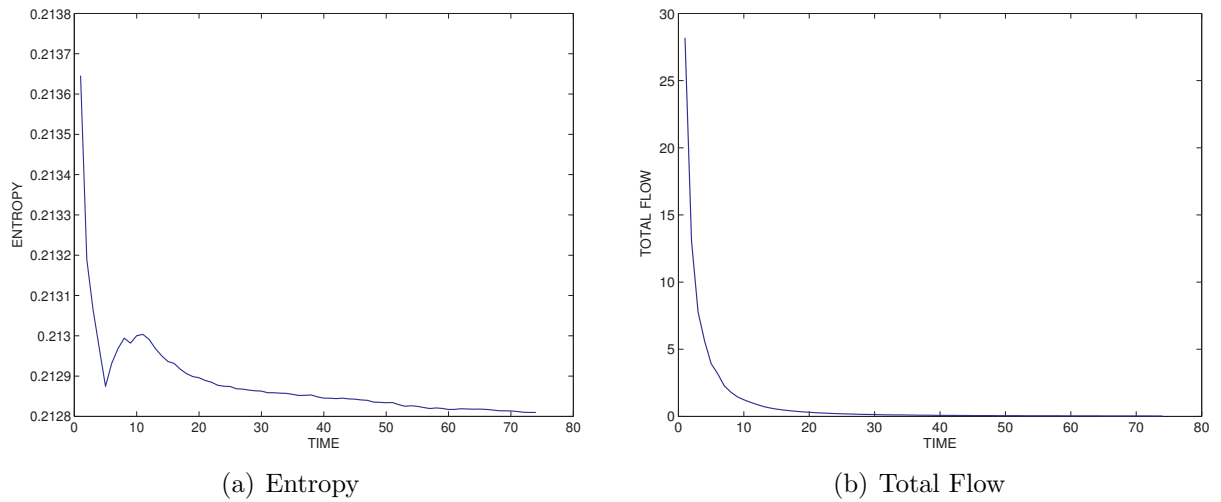


Figure 6.8: Total flow and entropy over time during RD interpolation with s.f of 2. The entropy of the image is at first reduced with the incorporation of initial diffusion flows. Due to overflow, the entropy rises again at end of the initial flow phase before the reverse flows are instigated to counter this, bringing the final flow to its minimal level. Note the gradual decrease and minimisation of both entropy and total flow over the time/iterations.

The RW segmentation result with the interpolated data, as compared to that using the original data, is given in Fig. 6.9. Note that we have deliberately used a relatively small seed group, to highlight the benefits of the PVE estimator in producing accurate segmentations. The same seed group is applied in both segmentations.

The improvement in accuracy is significant: compared to a DICE coefficient of 0.80 in the original case, the PVE-enhanced gave a DICE result of 0.88 with the same seed map. In addition to producing notably better defined tissue boundaries, the effect of interpolation in preventing leakage in the RW segmentation can also be easily observed. However, there are still locations of thin tumour growth where our method failed to correctly classify. This can be solved by adding additional seeds to such locations. A more comprehensive accuracy study has been carried for the ten pilot image data slices from Chapter 5. This time, fewer seeds have been applied as initialisation to monitor the effect of PVE correction on the DICE measure with the reference truth. A comparison of the original and PVE-corrected results are given in Fig. 6.10. Because of the successful prevention of leakage in most cases, an improvement in accuracy is observed in all ten trials.

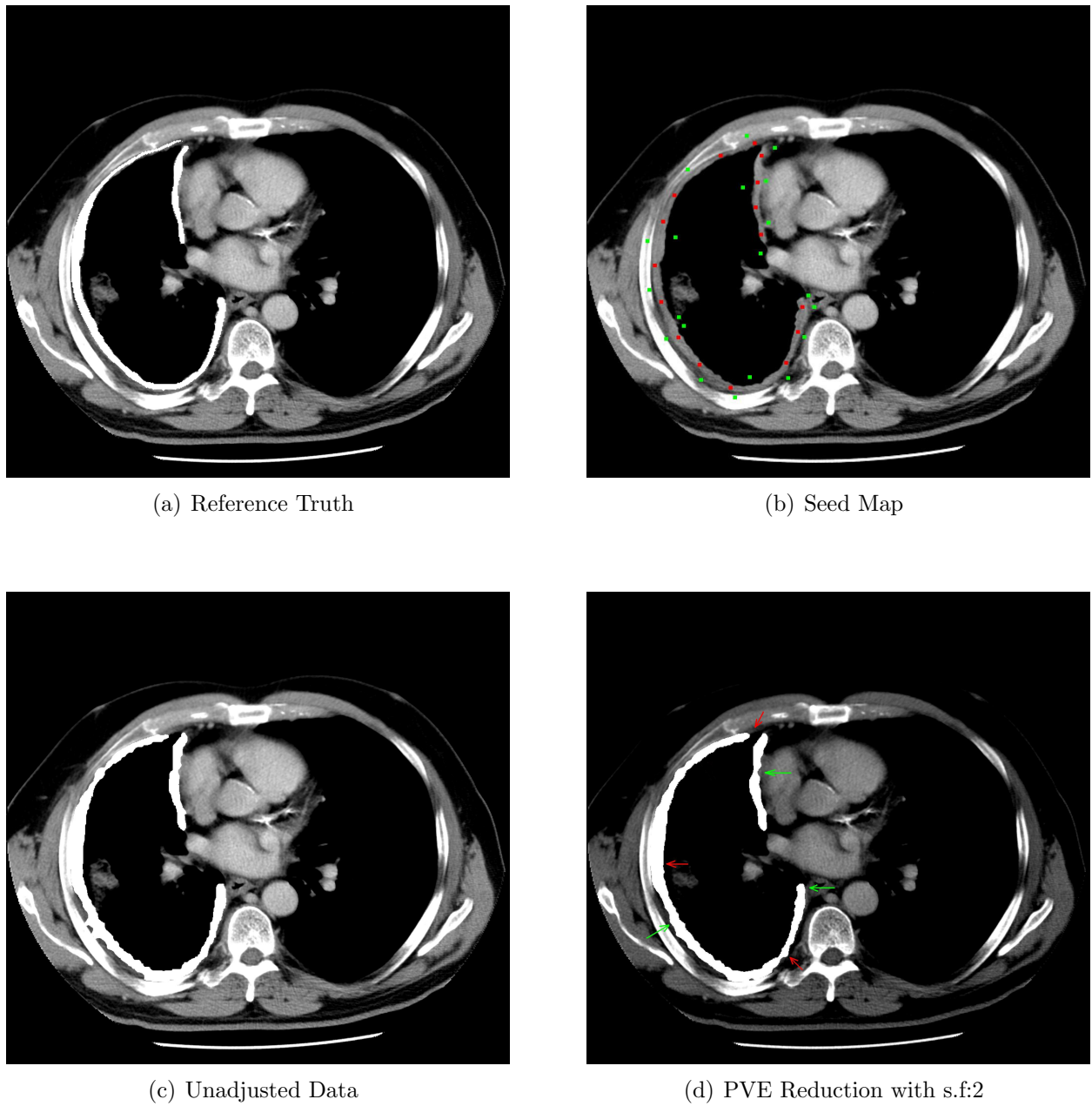


Figure 6.9: MPM data interpolated with reverse diffusion for PVE reduction. Improved performance of the random walk method is clearly observed, with the same seed map applied as initialisation. Note that most boundary leaks are effectively prevented in the PVE-corrected segmentation (indicated by *green arrows*). However, there are still locations of thin tumour growth where our method failed to correctly classify (indicated by *red arrows*).

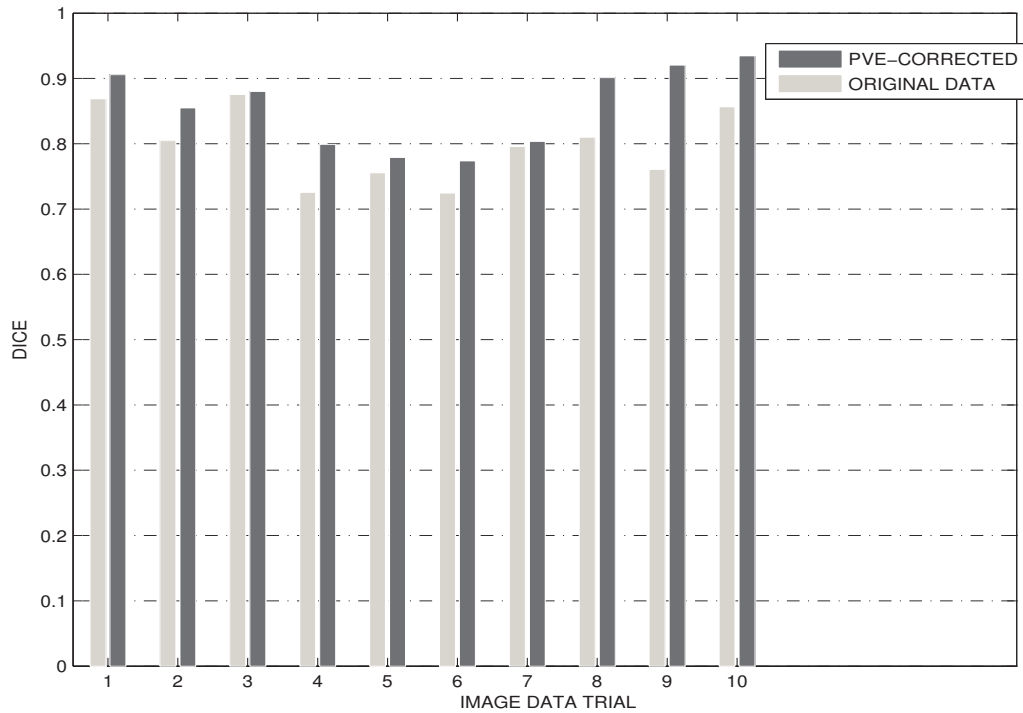


Figure 6.10: DICE coefficient comparison for the original and PVE-corrected data, as matched to the reference truth. Because of the successful prevention of leakage in most cases, an improvement in accuracy is observed in all ten trials. Trials 3 and 7 do not show significant improvements, possibly because the reference truth did not take the enhanced PVE-boundary voxels fully into account.

As mentioned earlier, we studied the effect of applying different scaling factors in the interpolation. Accuracy and computation speed changes are monitored against changes in scaling factor, as shown in Fig. 6.11.

We note that the computational burden increases drastically as the scaling factor increases, though this does not result in a corresponding enhancement in segmentation accuracy. Therefore we chose to use a consistent scaling factor of 2 throughout our study.

6.4 Extension to 3-D

Most of our image data are not represented by isotropic voxels. In fact, as explained earlier in this chapter, PVE is most dominant in the axial direction. In general, to correct PVE in a full image volume, it would not be sufficient to simply perform a series of planar PVE estimations. Rather, we need to extend the RD-PVE estimator to 3-D to take all three dimensions into

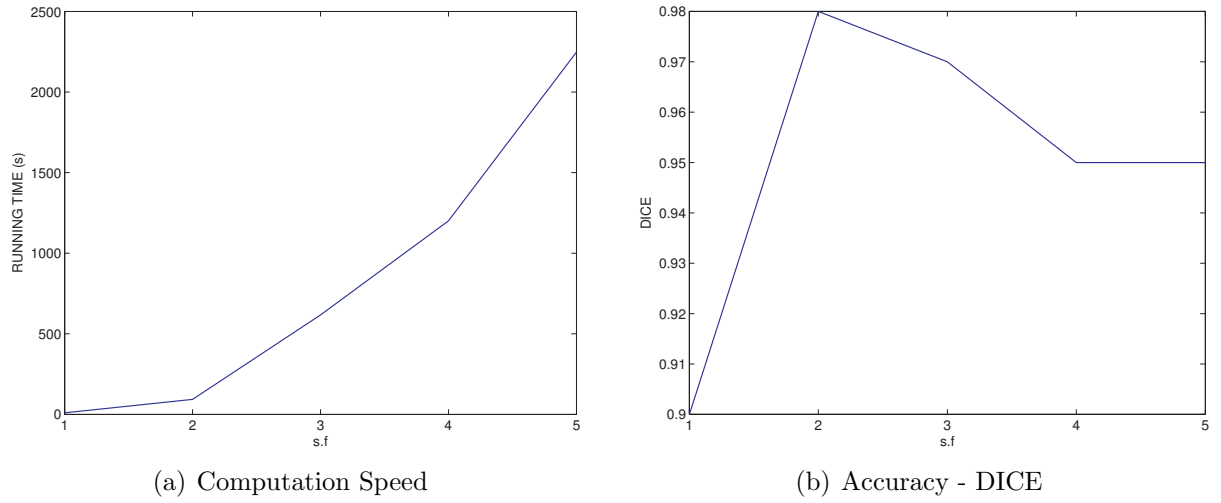


Figure 6.11: The effect of the interpolation scaling factor on segmentation speed and accuracy. Scaling factor (s.f) is best chosen as 2; because a larger s.f does not improve accuracy whilst it increases the computational burden drastically.

account for a wholly effective PVE correction.

In theory, the 3-D extension is simply a combination of a series of 2-D planar PVE corrections and 1-D corrections in the axial direction. However, as the 1-D axial correction would need to be performed for each and every image pixel, the overall computational requirement is expected to be very intensive. For our project, building onto the planar RD implementation by O. Salvado, we have implemented an extension of the method to 3-D, the pseudo-code of which is given below (Algorithm 6.1):

Algorithm 6.1: Pseudo-code for 3-D Reverse Diffusion Interpolation

define a new volume of size: Original Volume $\times (s.f_x \times s.f_y \times s.f_z)$, where $s.f_x$, $s.f_y$ and $s.f_z$ are scaling factors in the planar x,y and axial z directions, respectively.

for each voxel in the image data

set the 3-D voxel neighbourhood lattice;

compute the intensity gradients in all three directions;

find the conduction coefficients based on Eq. 6.20;

find the flow constraints based on Eq. 6.22 ;

compute the final diffusion flow for the local lattice from the flow constraints
 and conduction coefficients;

Algorithm 6.1 cont'd

update relevant voxels in the interpolated volume according to Eq. 6.22;

end

output the interpolated volume as the result.

Unlike in the planar case where the scaling factor is consistent in both the x- and y- directions, a higher scaling factor would generally be required in the axial direction for the 3-D case. This is because PVE is larger for most non-isotropic image scans and to ascertain effective PVE correction, higher order axial interpolations would be necessary. For our experiment in this section, we have used a s.f of 4 in the axial direction and 2 on the transverse planes. A coronal view of the interpolated data is shown and compared to the original in Fig. 6.12.

Following interpolation, the boundary-enhanced segmentation method is applied to the image, with interpolated identical seed maps as previously used on the original data segmentation in Chapter 5. The result is given in Fig. 6.13. Downsampling the segmented volume to match to the reference truth, a lower average overlap (DICE) is found and given in Fig. 6.14. It should be noted that the reference truth was acquired without the PVE adjustment and deliberately excluded uncertain tissue boundaries, hence the additional boundary information offered by

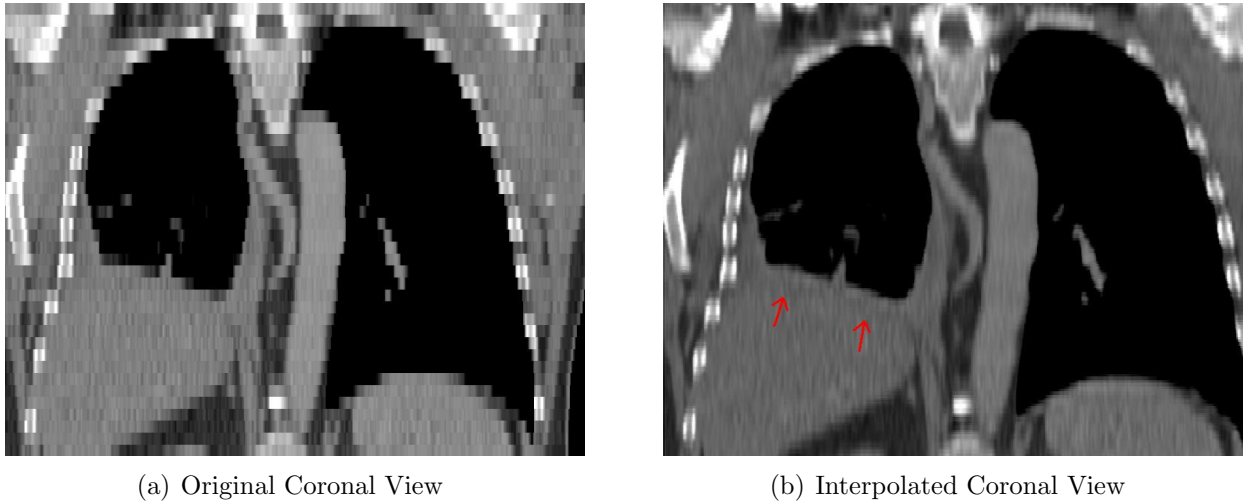


Figure 6.12: 3-D reverse diffusion interpolation in the axial direction ($s.f = 4$). Note the drastic overall improvement in axial definition and appearance of an identifiable tumour boundary in the coronal view (indicated by arrows).

the interpolation was not considered. Therefore it is difficult to assess accuracy of the PVE-corrected result by reference truth alone. Additional clinical consultation on an individual basis is necessary. Moreover, as each image slice in the original data corresponds to four slices in the interpolated data, this accuracy comparison is inconclusive because it fails to recognise the additional inter-slice knowledge in the interpolated data.

6.5 Computation Time Consideration

The final point of interest in this chapter is the computation time performance of the PVE-corrected method. As shown in Table 6.1, the extra computation time in the PVE-corrected case is mainly attributable to handling the additional interpolated data. For the 3-D segmentation this is particularly significant where a ten-fold increase in the computation time is observed. Ways of improving the computation time performance include a more discrete choice of scaling factor (smaller $s.f$ possible for the axial direction), and using a faster workstation (the one used for our study is largely obsolete by 2011 standards). We should note that in this pilot study, to illustrate the effect of PVE, we have intentionally applied the method to a very thick image scan (5mm) with coarse planar resolution (0.75mm planar voxel spacing). For most of

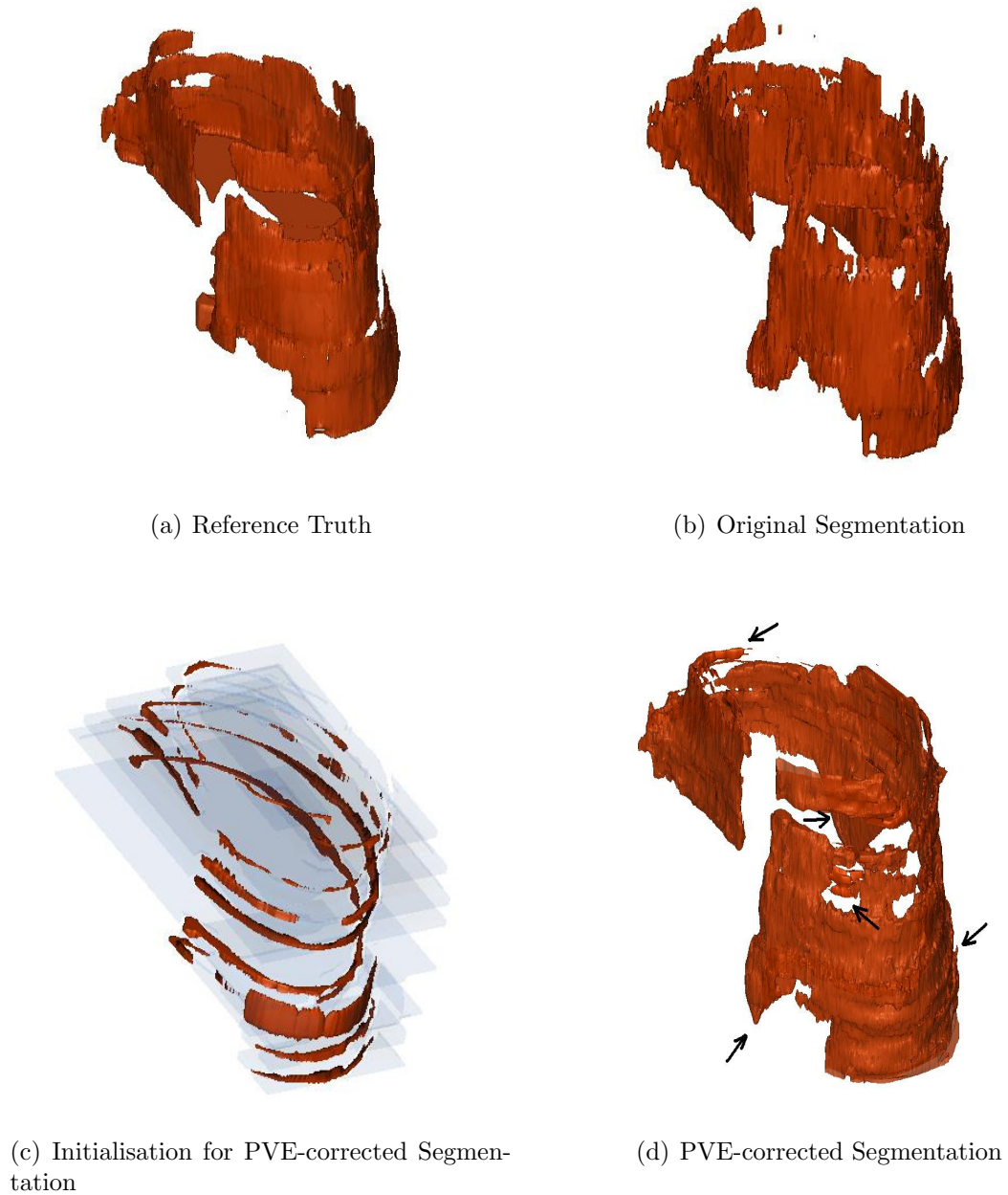


Figure 6.13: PVE-corrected volumetric segmentation results (same data as in Fig. 5.21). The initialisation ‘in’ seeds are displayed in *orange* and ‘out’ seeds in *blue*. Pointed in arrows are regions where discrepancies are found. Note the reference truth was acquired without PVE correction therefore it is difficult to assess accuracy of the PVE-corrected result by reference truth alone. Additional clinical consultation on an individual basis is necessary.

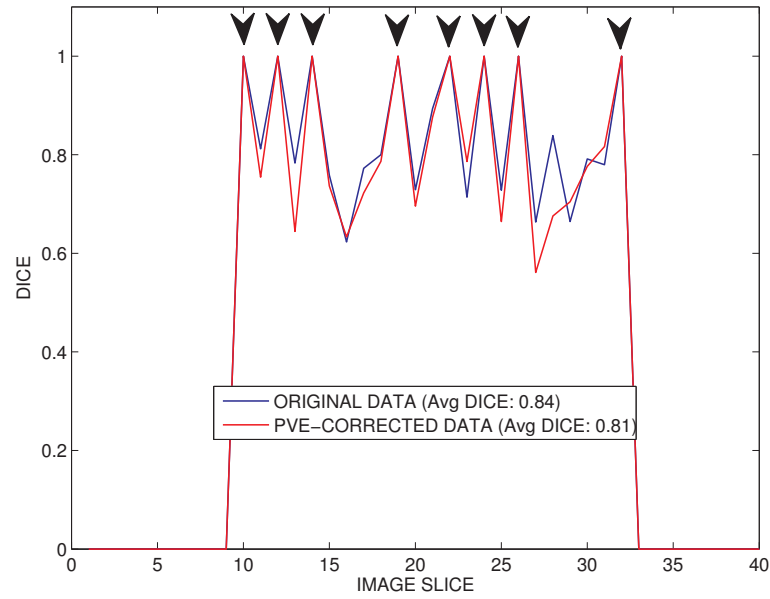


Figure 6.14: Accuracy (DICE) analysis of the PVE-corrected result, downsampled to match the reference truth. Each image slice in the downsampled data corresponds to four slices in the interpolated data therefore the comparison is inconclusive.

our clinical data and by convention of today's radiological practice, axial spacings of 2.5mm to 0.625mm are used for CT thoracic scans. Better planar resolutions are also possible. In this regard, for the most part in our main clinical experiment, we will use the lowest scaling factor of 2 for all directions for better efficiency. Repeating the 3-D experiment on a faster workstation with quad-core 3GHz CPU and 8GB of RAM gives an improved overall computation time of 52min37s for the 3-D PVE corrected case. In sum, at this stage, the impact of computation time on the general workflow of our algorithm is significant and the weighting of PVE correction must be carefully assessed against its benefits. We would suggest the use of this part of our algorithm only when necessary, such as for low resolution or thicker slice scans where PVE is more pronounced. This would remain the case until a better computation time can be achieved, rendering the PVE correction component more attractive from a computational efficiency perspective.

Experiment	2-D Overall	3-D Interpolation	3-D Overall
Unadjusted	5-15s	N/A	16m27s
PVE-corrected	10-40s	20m13s	164m20s

Table 6.1: Computation time comparison (3.39GHz, 2GB RAM) of the PVE-corrected versus the original method. Both 3-D experiments are for only a masked region of interest. The extra computation time in the PVE-corrected case is mainly attributable to handling the additional interpolated data.

6.6 Summary

In this chapter, we have reviewed a number of notable PVE classifiers. Our choice of the RD interpolation method is motivated by its simplicity, superior performance, as well as robustness to different imaging data. The extra computation time incurred by this addition to our algorithm can be ameliorated by interpolating only a masked region containing the ROI for segmentation. We have also implemented the method in 3-D, mainly to correct for the dominant PVE in the axial direction. The volume results are compared to the unadjusted case in terms of its accuracy and computation time and discussed.

We now briefly discuss the reference truth validity. During manual segmentation, it is known that there is a great deal of subjectivity while classifying boundary voxels. It is sometimes entirely arbitrary when it comes to difficult voxels of great uncertainty. This gives rise to a large amount of inter-observer and even intra-observer variability in the segmentation. Having a good way of computationally dealing with PVE is thus of high clinical value for MPM study. When used as a reference guide, computationally-derived boundary classification may offer good help to clinicians for segmenting boundaries of uncertainty. Therefore the advantage of an accurate PVE-voxel classifier extends beyond the improvement of automatically segmented results, which we have studied in this chapter.

Chapter 7

Evaluation of Tumour Responses and Validation

Having developed the tools necessary to evaluate the response of mesothelioma to chemotherapy over time, we are now in a position to address some of the clinical issues stated for the IV-Vinflunine phase II trials in Chapter 1.

We begin this chapter in Section 7.1 by describing the collection of the imaging data that is to be studied. Next, we analyse the universally used Response Evaluation Criteria in Solid Tumours (RECIST) system in terms of its observer variability, performance and limitations (Section 7.2). In Section 7.3 we perform volume segmentations on the complete clinical data library and establish patient responses to chemotherapy from a new perspective, taking into account the partial volume effect correction presented in Chapter 6. In Section 7.4, we study the correlation of established RECIST-based tumour responses and those from volumetric segmentation; specifically, we evaluate the performance of the volume segmentation method for assessing tumour progression in MPM, as compared to RECIST. Then, by updating the clinical response evaluation of relevant trials, we recommend a computer-assisted algorithm for future assessments of MPM response and discuss the applicability of extending this algorithm to other clinical situations. Finally in Section 7.5, changes in aerated lung and pleural effusion with time are determined as a way of evaluating the potential symptomatic effects of a

treatment to patients.

The focus of our study is on the RECIST measurements with the goal of improving the existing measurement protocols. The notion of reference truth is not considered in this chapter. Instead, we base our discussions on the current clinical and anatomical understanding of MPM growth and, without bias, benchmark the computer-aided method to the RECIST technique for quantifying MPM. Our analysis highlights areas of performance which are essential to image-based quantification including clinical accuracy, robustness, accessibility and time efficiency.

7.1 Image Data

A total of 65 patients were initially recruited for the IV-Vinflunine experiment. Of the recruited patients, 70% had advanced disease (stages III to IV), as classified by the IMIG system (Appendix A). The median time from disease diagnosis to trial entry was 0.4 years, with a range between 0.04 to 7.20 years.

Due to unforeseeable circumstances such as voluntary withdrawal and the fact that only the UK centre data was available for our study, we had access to 48 full thoracic CT scans, taken from 15 different patients. All patients had histologically and cytologically confirmed cases of mesothelioma and at least one measurable lesion that satisfied the measurability criteria specified in Section 1.4. Each participant received one baseline, and between one to three follow-up scans, for monitoring the effect of the trial. Scans were planned at baseline (1) and after 2 cycles (2), 4 cycles (3), and 6 cycles (4) of chemotherapy. Each cycle lasts 3 weeks (one single administration of Vinflunine). Therefore, all scans are valid for assessing tumour responses under RECIST (at least four weeks between eligible scans). Patients were treated until either disease progression or unacceptable toxicity.

The CT examinations were performed on a LightSpeed Ultra CT scanner (GE Medical Systems). Each CT scan volume consists of multiple axial slices of 512 x 512 pixels. An assortment of thick (5mm), thin (2.5mm) and quasi-isotropic (0.625mm) scans are available. Planar voxel spacings are either 0.68mm or 0.74mm.

Of the 15 patients, RECIST showed no confirmed responders. Instead, stable disease was

Response	No	%
Complete response (CR)	0	0
Partial response (PR)	0	0
Stable disease (SD)	12	80
Progressive disease (PD)	3	20
Total no. of patients	15	100

Table 7.1: Overall response rate (modified RECIST for MPM) on the available clinical data from IV-Vinflunine phase II trials.

found in 80% of the cases and progression was seen in the remaining 20% of the patients. This is shown in Table 7.1.

7.2 RECIST Analysis

As we noted in Chapter 1, the RECIST system is a quantitative tumour response assessment protocol that was developed in an effort to revise the WHO criteria [45], which were originally designed to quantify near-spherical shapes and are no longer appropriate for the growing demand to study various forms of tumour growths in modern clinical oncology. The RECIST protocol has since been updated several times to accommodate various additional clinical scenarios. The main designations of RECIST include the original version (1.0), the modified version (1.1), and RECIST adapted for MPM (MPM-RECIST), whose pros and cons have been discussed extensively and compared in Section 1.5.1.

7.2.1 MPM-RECIST

To commence our analysis, we applied RECIST 1.1 and MPM-RECIST to our clinical data. In RECIST 1.1, the individual tumour's longest axis is measured uni-dimensionally, as shown in Fig 7.1. For lymph nodes, the short axis perpendicular to the longitudinal measurement, is also taken into account, yielding a bi-dimensional measurement. However, the main problem with RECIST 1.1 is that it is designed for near-spherical shapes, and so the longitudinal structures characteristic of MPM tend to render it less accurate. MPM-RECIST, in an attempt to solve

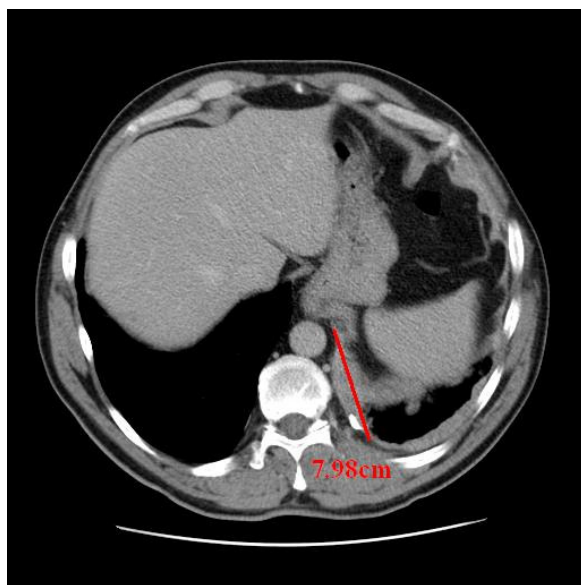
this problem, measures the tumour thickness perpendicular to fixed structures such as the chest wall and/or vertebral column, and specifies measurements in two positions at three separate levels on transverse cuts of the CT scan. The sum of these six measurements is then defined as a single pleural uni-dimensional sum and used to represent the tumour at a specific point in time. MPM-RECIST does not materially alter the response rates of RECIST 1.1 but merely resolves the ambiguity between uni-dimensional and bi-dimensional lesions [12]. For the remainder of this chapter, we adhere to the recommendations given in [12] and concentrate our discussions on MPM-RECIST. Full specification of MPM-RECIST is given in Appendix B.

7.2.2 Observer Variabilities

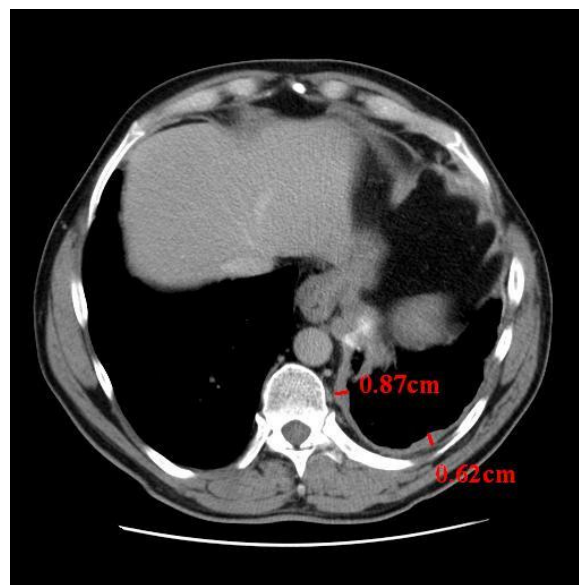
We next investigate the intra- and inter-observer variabilities of MPM-RECIST in order to assess the reliability of the method for assessing the tumour response in the IV-Vinflunine trials. Intra-observer variability can be examined by performing a series of RECIST measurements at different locations relative to fixed thoracic structures and checking for repeatability. Initially, we used a 6'-9' combination (6 and 9'o clock locations, with the image centre being the 'clock origin', as illustrated in Fig. 7.2) as the measurement sites for right lung lesions (3'-6' for the left). Subsequently, the measurements were taken at 3'-9' and 3'-6' (or 6'-9' and 3'-9' for left lung lesions). Scatter plots are given in Fig. 7.3, showing the correlation between observations made by the same observer at different measurement sites. To quantify this correlation, Pearson's coefficient of correlation (CC) was used, defined as follows, for two data samples, $X = \{x_i\}$ and $Y = \{y_i\}$ of length n whose means and standard deviations are $\bar{x}, \bar{y}$ and s_x, s_y respectively:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}} \quad (7.1)$$

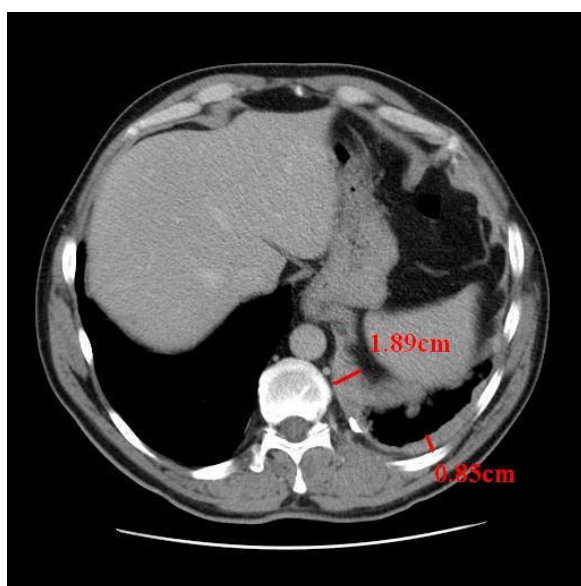
The p-value was also computed for each pair of datasets according to student's t-test. The presence of a correlation between the two datasets is assumed when $p < 0.05$. Measurements taken at individual measurement sites were found to show correlation except for the case of 6' and 9' where the p-value is found to be greater than 0.05 (0.0533). This suggests that a



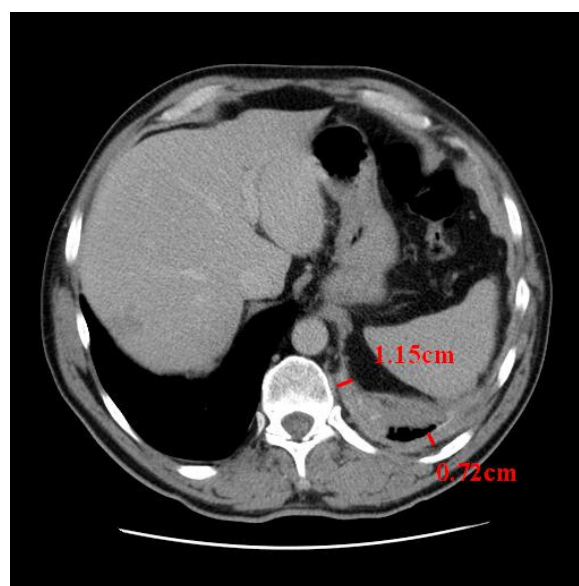
(a) RECIST 1.1



(b) MPM-RECIST (slice 1)



(c) MPM-RECIST (slice 2 - 10mm)



(d) MPM-RECIST (slice 3 - 20mm)

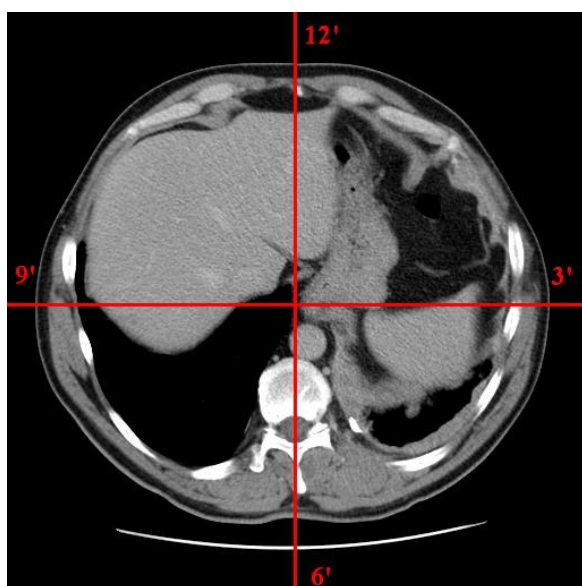
Figure 7.1: Tumour measurements based on RECIST 1.1 and MPM-adapted RECIST. The tumour is first deemed as ‘measurable’ by both systems. The latter is performed on three transverse cuts of the scan which are 10mm apart. The individual measurements are added up to give an overall sum of 6.5cm. In contrast, RECIST 1.1 represents the tumour as a single measurement of 7.98cm. The tumour measurements are subsequently compared to those taken from other time point scans for an assessment of the tumour response according to the specifications given in Appendix B.

substantial degree of variability may be found depending on the measurement sites selected. However, as MPM-RECIST specifies measurements taken at two sites, this variation is reduced, as given in Fig. 7.3(d) - (f). Strong correlations are found for the sums of two measurements. In fact, we suggest a linear correlation for the RECIST sums. The least mean squares method is applied to perform such regression on the data, where the least squares sum: $R^2 = \sum_i [y_i - f(x_i)]^2$ for data points f_i and linear fit $f(x) = ax + b$ is minimised such that $\frac{\partial(R^2)}{\partial a_i} = 0$. The goodness of fit is evaluated by the coefficient of determination R^2 , defined for a dataset of observed values y_i and modelled values f_i whose algebraic mean is $\bar{y} = \frac{1}{n} \sum_i y_i$:

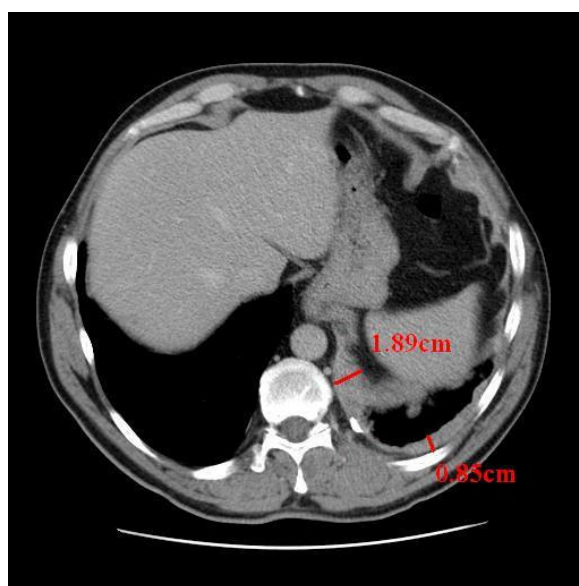
$$R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (7.2)$$

The intra-observer results show a considerable degree of variability (0.5-0.7 CC between observations, with a mean variation of between 20-30% and standard deviation of 70%, of the total measurements). The linear correlation does not capture well the extensive scattering pattern for individual site measurements. However, a strong correlation is found amongst the RECIST sums. This explains the benefit of using multiple measurement sites in MPM-RECIST. To summarise, it is found that the MPM-RECIST protocol is prone to substantial intra-user variation and can thus be less reliable for situations where high degrees of accuracy and reproducibility are required. However, despite this variability, we note that, in our patients, the choice of measurement location does not play a decisive role in the overall response classification. In our study, we have obtained the same patient response results with the tumours assessed at different measurement sites.

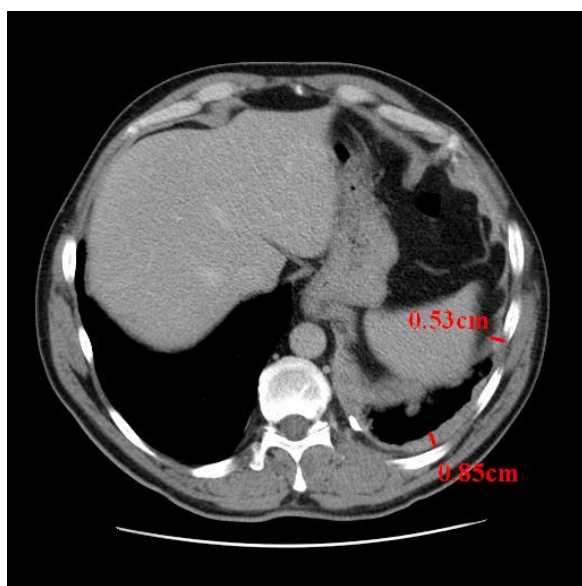
We assessed the inter-observer variability of RECIST by repeating the same measurements with an independent(non-specialist, non-clinical) observer and compared them to the findings made by the clinical investigator. Note the involvement of a non-specialist observer tended to exaggerate our inter-observer variation finding. This is primarily due to the lack of expertise in areas such as thoracic anatomy and oncological progression, leading to delineations different from those by the clinical observer. However, for the purpose of this study, the non-specialist



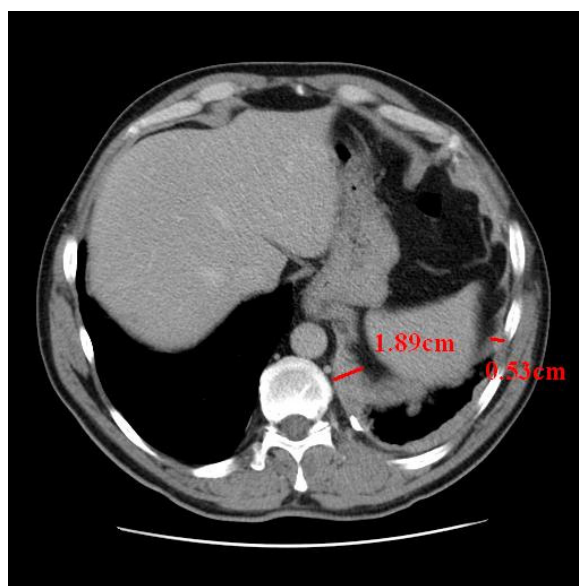
(a) Measurement Sites Reference Coordinates System



(b) 6'-9'

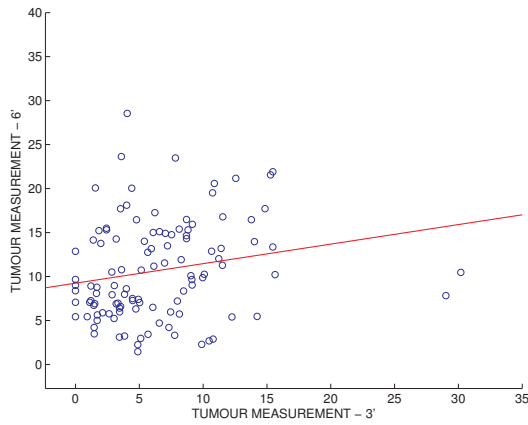


(c) 3'-6'

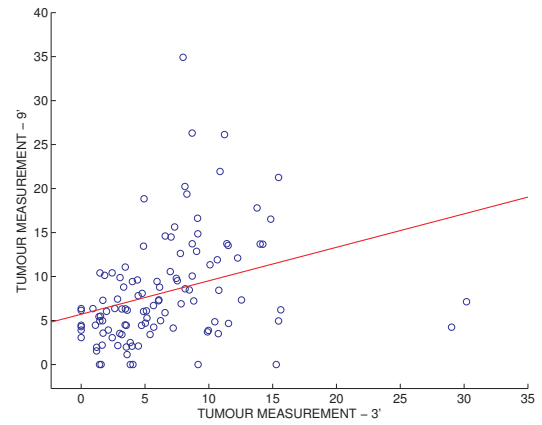


(d) 3'-9'

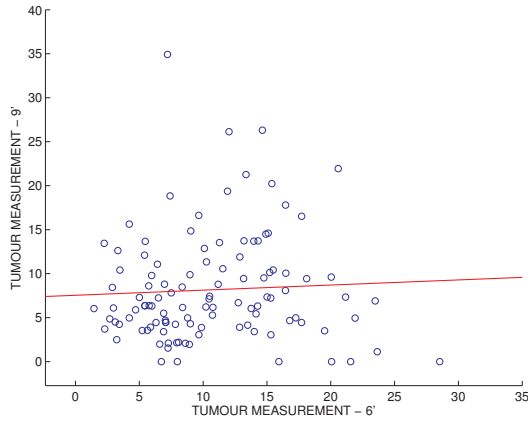
Figure 7.2: Illustration of the intra-observer variability assessment, by comparing measurements taken at different sites along the pleural fixture. The 3', 6' and 9' locations are chosen because MPM is most likely to grow near these sites. The coordinate system is used only as a guidance. Specific sites are selected based on both tumour measurability and image clarity.



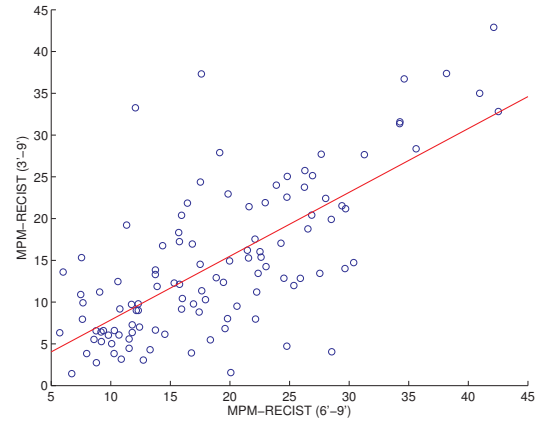
(a) 3' vs. 6', CC: 0.2049, p-value: 0.0310, R^2 : 0.0420



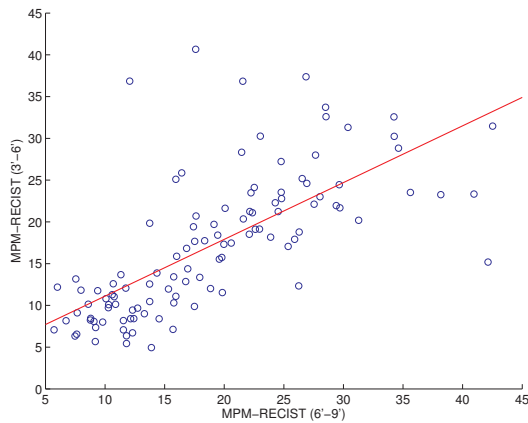
(b) 3' vs. 9', CC: 0.3211, p-value: 0.00059, R^2 : 0.1031



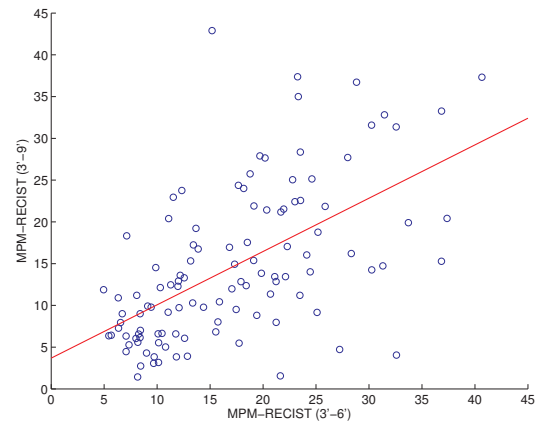
(c) 6' vs. 9', CC: 0.5787, p-value: 0.0533, R^2 : 0.0028



(d) 6'-9' vs. 3'-9', CC: 0.7078, p-value: 3.78e-18, R^2 : 0.5009



(e) 6'-9' vs. 3'-6', CC: 0.6920, p-value: 4.06e-17, R^2 : 0.4789



(f) 3'-6' vs. 3'-9', CC: 0.5202, p-value: 2.42e-11, R^2 : 0.3370

Figure 7.3: Scatter plots with least square linear fits showing the correlation between observations by the same observer. The linear correlation does not well explain the scattering pattern for individual site measurements. However a strong correlation is observed amongst the RECIST sums. This explains the benefit of using multiple measurement sites in MPM-RECIST. In sum, a notable degree of intra-observer variability is found, quite significantly affecting the reproducibility of RECIST measures.

Patient No	Scans				Response
	Baseline	2 Cycles	4 Cycles	6 Cycles	
130101	35	32	34	36	SD
130102	30	28	29	35	SD
130103	30	29	33		SD
130104	75	42			SD
130105	42	33			SD
130106	100	75	120		PD
130107	12	12			SD
130108	38	36	43	36	SD
130109	39	50			PD
130110	94	96	102		SD
130111	31	32	34	36	SD
130112 (A)	76	57	79		SD
130113	35	34	30		SD
130114	34	49	42		PD
130115 (B)	70	59	60	68	SD

Table 7.2: MPM-RECIST measures - clinical observer (mm)

Patient No	Scans				Response
	Baseline	2 Cycles	4 Cycles	6 Cycles	
130101	50	57	81	112	PD
130102	30	29	25	34	SD
130103	52	74	33		SD
130104	77	35			SD
130105	30	34			SD
130106	43	32	40		SD
130107	57	73			PD
130108	46	48	47	59	SD
130109	59	66			SD
130110	84	69	89		SD
130111	52	63	68	63	PD
130112 (A)	80	54	78		SD
130113	38	38	43		SD
130114	66	71	65		PD
130115 (B)	59	54	61	79	SD

Table 7.3: MPM-RECIST measures - independent observer (non-specialist) (mm).

Clinical Observer	Independent Observer				
	CR	PR	SD	PD	Total
CR	0	0	0	0	0
PR	0	1	0	0	1
SD	0	2	14	9	25
PD	0	0	2	3	5
Total	0	3	16	12	31

Table 7.4: Time point observations comparison of MPM-RECIST from two independent observers. The classifications of tumour response vary greatly between the two observers, in accordance to the large inter-observer variability, as expected.

observer has been trained thoroughly by the clinical observer for MPM delineation. In addition, even the outlining of tumour lesions by the clinical observer was subjected to notable intra-observer variability, as shown earlier. Therefore, from an image processing point of view, we expect the bias imposed by the non-specialist observations to be within the acceptable threshold. To further eliminate observer bias, a blinded experiment was conducted where one observer made his/her measurements without knowledge about those made by the other. Complete experimental findings on these RECIST observations are given in Tables 7.2 and 7.3. For the results presented in Fig. 7.4, Tables 7.4 and 7.5, the mean absolute difference ($(B-A)/A \times 100\%$) between the two sets of measures (denoted as A and B) is found to be 31.41% with a standard deviation of 42.01%, and a relatively weak correlation of 0.3808. The correlation itself, however, is supported by an affirmative p-value of 0.0201. The linear regression once again fails to effectively predict the relationship between the two data ($R^2 = 0.1450$). Finally, treating the clinical RECIST measurements as the ground truth, we compute the variation of the repeated measurement as a distance to these measurements. The probability density function (PDF) is shown in Fig. 7.5. Distinct modes are found at 6.43% and 58.57%. The above results show that measurements made by two independent observers have low correlation, indicating a high degree of inter-observer variability of the RECIST system.

It should be noted that comprehensive comparison of different RECIST protocols requires

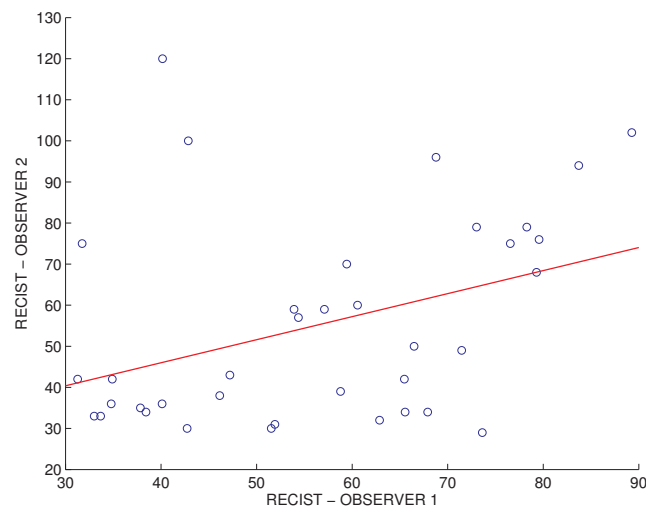


Figure 7.4: Scatter plots with least square linear fits showing correlation between observations made by different observers. CC: 0.3808, p-value: 0.0201, R^2 : 0.1450. This shows that measurements made by two different users are low in correlation and in turn indicates a very high degree of inter-observer variability in the RECIST system.

Clinical Observer	Independent Observer				
	CR	PR	SD	PD	Total
CR	0	0	0	0	0
PR	0	0	0	0	0
SD	0	0	9	3	12
PD	0	0	3	0	3
Total	0	0	12	3	15

Table 7.5: Overall response rate comparison of MPM-RECIST from two independent observers. Classification discrepancies, as in the time point case, are overwhelmingly common.

their extensive application to very large sets of data, a condition that our data library simply does not meet. Data from over 6500 patients were originally used to support RECIST 1.1 [19]. Moreover, in this section, we have only conducted a very brief study of both variabilities, so our findings are far from being conclusive. It would be necessary, as a future work, to invite additional observers to take part in the inter-observer variability study for a more thorough review of the RECIST reliability; nevertheless we would expect it to confirm the initial results presented here. Furthermore, the assessment of inter-observer variability typically requires comparing measurements from more independent observers to reduce sampling bias. In fact,

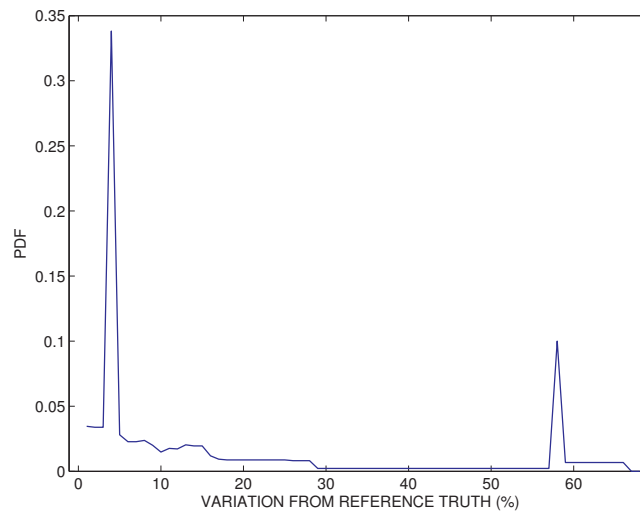


Figure 7.5: PDF of % variation from the clinical RECIST measurements. Statistical modes are found to be at 6.43% and 58.57%. The principal mode is found $< 10\%$, therefore suggesting a notable link between the two observations.

three specialist observers is normally used for studies of this kind, and is supported by an investigation into the trade-off between experimental accuracy and viability [48].

We conclude this section with a discussion of RECIST based on our findings thus far. As stated in Chapter 1, a major drawback of the MPM-RECIST protocol is that it only measures tumour size in 6 positions and is known to be prone to inter and intra-observer variabilities. Major disagreements occurred in 40% and minor disagreements in 10.5% of all cases [76]. This is supported by our own findings of intra- and inter-observer variabilities accounting for average variations of 20-30% and 31% of the total measurements, respectively. Reasons for these disparities include errors in tumour measurements; errors in selection of measurable targets, and radiologic technical problems such as the persisting partial volume effect and other image artefacts. Such disagreements may distort the assessment of treatment response in clinical trials. In addition, inconsistencies in the scanning thickness and/or axial slice alignment may produce additional problems in selecting equivalent slices on subsequent scans. This may explain the large inter-observer variability, since for measurements repeated by the same observer, equivalent image slices were used, which is not the case for measurements made by different users. Lastly, the authors of MPM-RECIST [12] suggested that tumours may also be non-evaluable; this manifests a multitude of restrictions, such as the use of intravenous

contrast enhancing agent, proper timing, scan delay and breathing etc, all of which must be well coordinated in order to make the scan assessable under RECIST.

To address the above issues, we propose an image segmentation scheme that assesses tumour responses based on volumetric changes, namely volumetric tumour assessment (VTA). In the next few sections, we compare the performance of this scheme to that of MPM-RECIST and illustrate the key advantages offered by our new way of tackling MPM response monitoring.

7.3 Results from Volumetric Segmentation

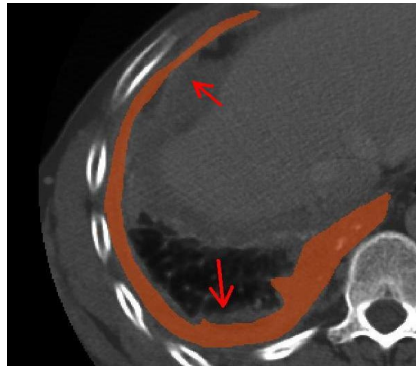
Using the volumetric segmentation developed in Chapter 5, we segmented the complete collection of MPM image data for each tumour. Sample results from two patients (A & B) are presented in Figs. 7.6 and 7.7. For our experiment, we specified seed initialisations so as to result in an effective segmentation without, in most cases, necessitating the use of computationally expensive processes such as high order RD-based PVE correction or the image registration sub-routine developed in Section 5.6. Note that the ‘efficacy’ of a segmentation is entirely subjective: in this case at the author’s discretion. It is determined based on immediate visual inspection and cross-validation with the clinician.

Excessive initialisation is likewise avoided for two reasons: 1) Our aim in this thesis is to demonstrate the functionality of our computer-aided segmentation method. 2) Initialisation, or the manual sketching of 2-D tumour boundaries, is in itself a tedious procedure and can be particularly problematic for a non-specialist in the field. In this regard, we have delineated seed regions based on our given knowledge about MPM, as the initialisations, and defined them on between 6 to 10 axial slices for each scan (with 60-110 raw image slices depending on the scanning voxel depth), in other words, roughly one initialisation used for every 10 raw image slices. An average of 40-60 seconds was spent on delineating seed regions per initialised slice. The initialisation slices are not evenly distributed across the volume scan; they tend to concentrate around regions of difficulties, marked by the presence of weak tumour boundary, fluid and/or collapsed lung. Additionally, to facilitate an effective response evaluation, the same axial locations are selected for initialisation for the scans of the same patient. This is

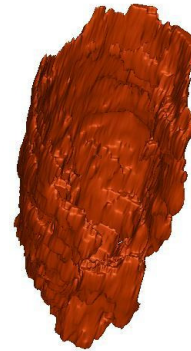
comparable to MPM-RECIST, where measurements are taken from the same transverse cuts for different scans of the same patient. However, due to the lack of a standard scanning protocol during the clinical trials, which resulted in a series of scanning thicknesses being employed even for scans of the same patient, the exact slice association between different scans was not always possible. The RD-interpolation from Section 6.4, along with the registration method introduced in Section 5.6, are applied to upsample/downsample the slices and compute the slice correspondence for difficult cases.

For Patient A (Fig. 7.6), at baseline - (a), (b), the shape of tumour resembles that typical of MPM, surrounding the pleura with near even distribution of thickening. After two cycles of IV-Vinflunine treatment, a reduction in the pleural thickening may be observed, and most notably a part (upper lung) of the primary lesion appears to becoming detached - (d) (pointed by *green* arrows). Fluid secretion is also observed in the planar view - (c) (pointed by *red* arrows). Overall, these observations may explain the significant reduction in the RECIST measure (over 20%, classified as PR). However, a further scan (taken after 4 cycles) rebuts this hypothesis and shows that the tumour response occurs only over the short term. Volumetrically - (f) this is demonstrated by the significant growth in the primary lesion and reconnection occurs between the formerly ‘detached’ upper section and the main lesion (pointed by *red* arrows). From the planar perspective - (e), This is once again supported by the RECIST measurement of 79mm, indicating strong evidence of relapse. The tumour is classified in RECIST as SD after 4 cycles of treatment, which is also supported by our computed tumour volume change ($9.398e6\text{mm}^3 - 6.305e6\text{mm}^3 - 8.733e6\text{mm}^3$).

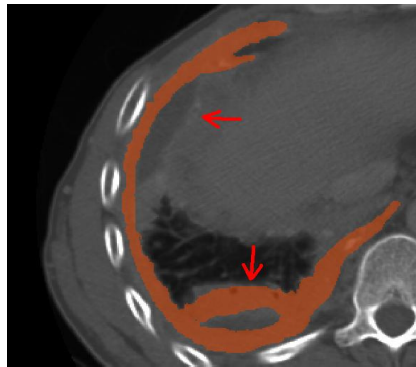
For Patient B (Fig. 7.7), MPM pleural thickening is observed at baseline - (a), (b). After two cycles - (c), (d), although a reduction in thickness is observed in some areas of the lesions (*green* arrows), a new lesion has initiated in the lower thorax (*red* arrows). This growth pattern continues after 4 cycles - (e),(f). Note that despite its good performance in (a) and (e), our segmentation routine fails to segment the calcification in (c) (*yellow* arrow). This may be caused by the lack of ‘seed guidance’ from the immediate surrounding slices, coupled with the strong ‘tissue boundary’ suggested by the calcification intensity, similar to that of the ribs. In other



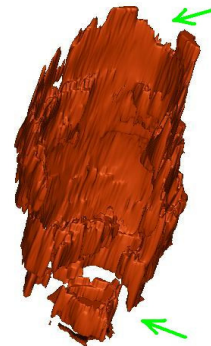
(a) Patient A - Scan 1 (Baseline, 2003/12/29) - Segmented Slice



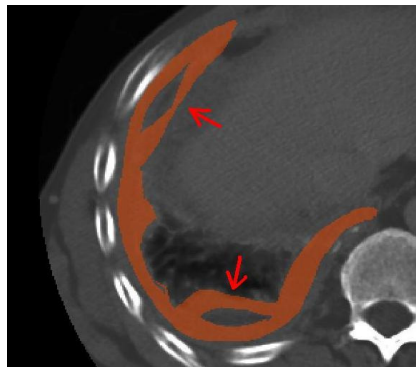
(b) Patient A - Scan 1 (2003/12/29) - Segmented Volume: $9.398e6\text{mm}^3$, RECIST: 76mm



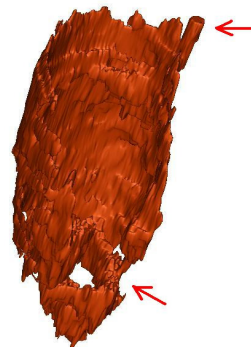
(c) Patient A - Scan 2 (2 Cycles, 2004/03/01) - Segmented Slice



(d) Patient A - Scan 2 (2004/03/01) - Segmented Volume: $6.305e6\text{mm}^3$, RECIST: 57mm



(e) Patient A - Scan 3 (4 Cycles, 2004/04/29) - Segmented Slice

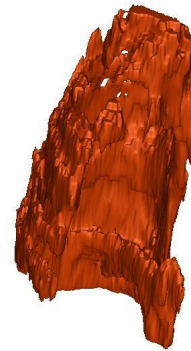


(f) Patient A - Scan 3 (2004/04/29) - Segmented Volume: $8.733e6\text{mm}^3$, RECIST: 79mm

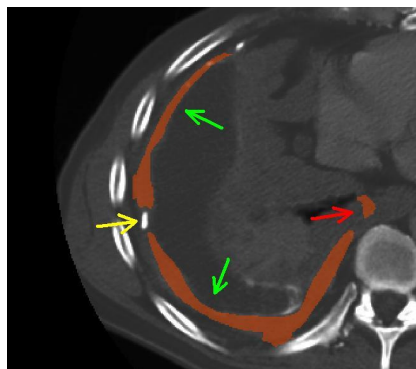
Figure 7.6: Volumetric tumour estimations for Patient A. At baseline - (a), (b), the shape of the tumour resembles that typical of MPM, surrounding the pleura with nearly evenly distributed thickening. After two cycles of treatment - (c), (d), a reduction in the pleural thickening may be observed, most notably a part (upper lung) of the primary lesion appears to becoming detached - (d) (*green* arrows). However, a further scan - (e), (f) (taken after 4 cycles) rebuts this hypothesis and shows that the tumour response occurs only over the short term. Volumetrically, this is demonstrated by the significant growth in the primary lesion. Reconnection occurs between the formerly 'semi-detached' upper section and the main lesion (*red* arrows).



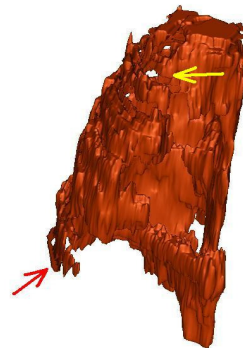
(a) Patient B - Scan 1 (Baseline, 2004/07/12) - Segmented Slice



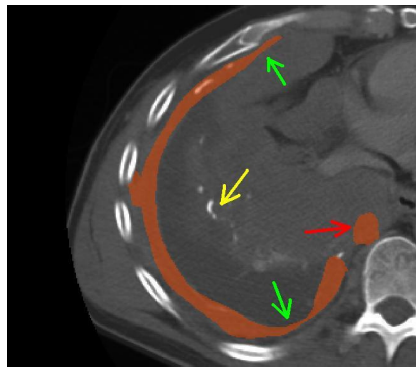
(b) Patient B - Scan 1 (2004/07/12) - Segmented Volume: $4.682e6\text{mm}^3$, RECIST: 59mm



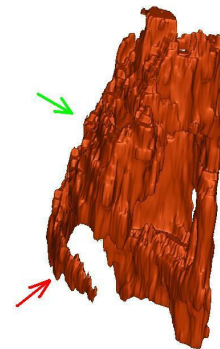
(c) Patient B - Scan 2 (2 Cycles, 2004/08/18) - Segmented Slice



(d) Patient B - Scan 2 (2004/08/18) - Segmented Volume: $5.050e6\text{mm}^3$, RECIST: 60mm



(e) Patient B - Scan 3 (4 Cycles, 2004/11/01) - Segmented Slice



(f) Patient B - Scan 3 (2004/11/01) - Segmented Volume: $6.345e6\text{mm}^3$, RECIST: 68mm

Figure 7.7: Volumetric tumour estimations for Patient B. As in Patient A, MPM pleural thickening is observed at baseline - (a), (b). After two cycles - (c), (d), although a reduction in thickness is observed in some areas of the lesions (*green* arrows), a new lesion longitudinal growth has initiated in the lower thorax (*red* arrows). This growth pattern continues after 4 cycles - (e), (f). Note that despite its good performance in (a) and (e), our segmentation routine fails to segment the calcification in (c) (*yellow* arrow), this may be due to the fact that our segmentation was designed to work within a low contrast soft tissue framework and rejects any strong boundaries not specifically explained by the initialisation seed map.

words, our segmentation was designed to work within a low contrast soft tissue framework and rejects any strong boundaries not specifically explained by the initialisation seed map. The computed tumour volumes exhibit a strong correlation with RECIST measurements. We investigate this correlation more closely in Section 7.4.

To visually illustrate the above discussion points, patient responses as regional changes in volume are given in Fig. 7.8. This is accomplished by first registering the segmented tumour volumes using the method introduced in Section 5.6, then applying a search algorithm to detect and map local differences. For Patient A, the computed reduction in tumour is found to be a general reduction in the pleural lesion thickness (*green* arrows). This is naturally accompanied by a similar change in RECIST, which measures the tumour based on its lesion thicknesses. However, these reductions are negated by tumour growth (*red* arrows) after subsequent trials, indicating a disease relapse - (b). For Patient B, the gradual development/elongation of the lower thoracic lesion is clearly observed in (c) and (d). Note that unlike for Patient A, the drug is largely ineffective at halting the disease progression for Patient B, even in the initial phase (i.e after 2 trial cycles). The angle of viewing is chosen so as to best illustrate the key observations made. A number of 3-D visualisation tools are available to better show tumour volume changes. This may prove to be vital to any potential clinical implementations of our method.

7.4 Validation and Discussion

The initial volumetric findings presented in the previous section suggest a strong correlation between RECIST and volumetric changes. We investigate this relationship more closely in the following analysis. Complete experimental findings on volumetric tumour segmentation are presented in Tables 7.6 and 7.7, alongside the clinically verified RECIST measures for comparison. This supports our volumetric findings from a clinical perspective.

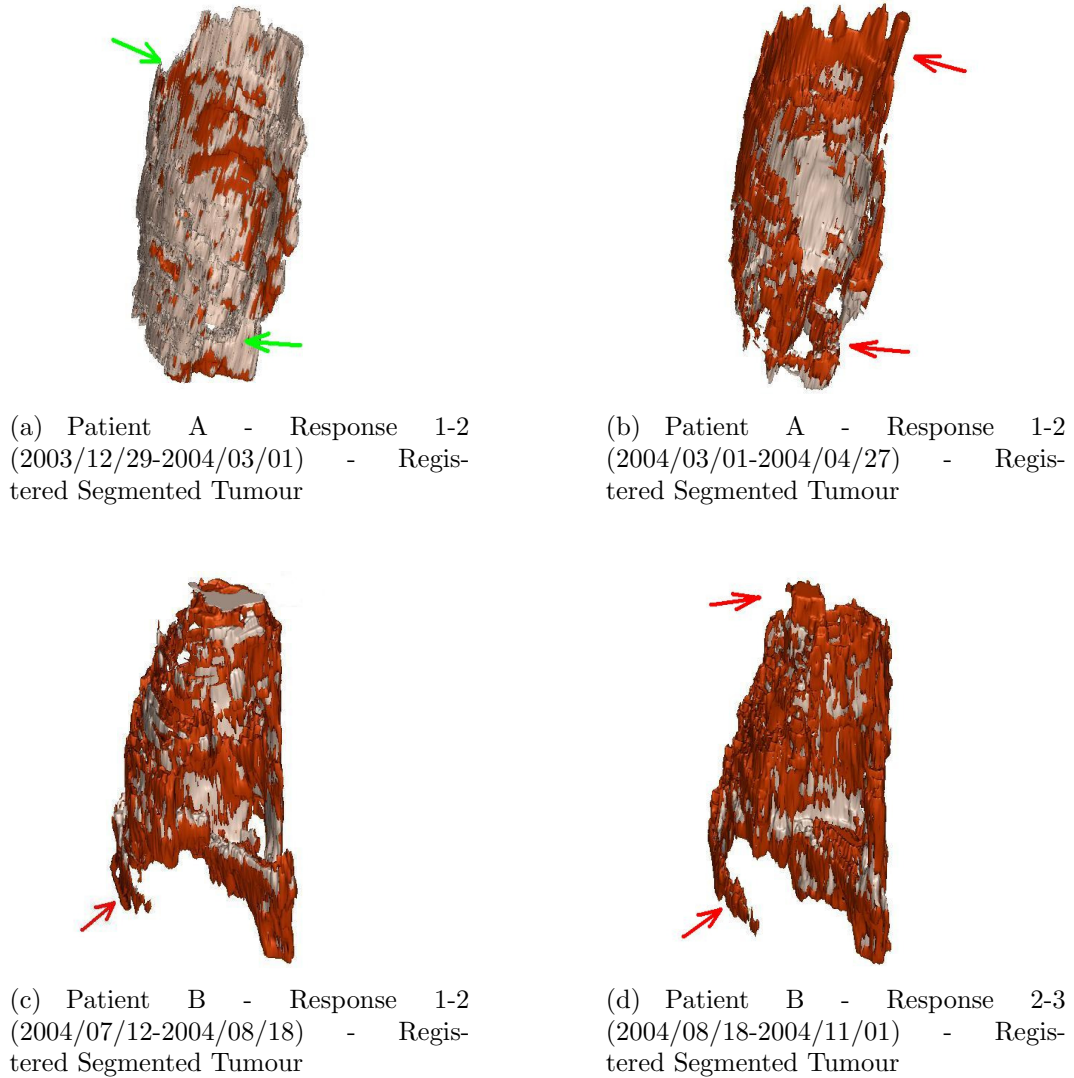


Figure 7.8: Volumetric shape changes for patients A & B. Tumour growth and reduction are shown in *orange*, and *white*, respectively. In (a), the computed reduction in tumour is found to be a general reduction in the pleural lesion thickness (*green* arrows). This gives rise to a similar change in RECIST, which measures the tumour based on its lesion thicknesses. These reductions are negated by tumour growth (*red* arrows) after subsequent trials, indicating a disease relapse - (b). For Patient B, the gradual development/elongation of the lower thorax lesion is clearly observed in (c) and (d). The angle of viewing is chosen so as to best illustrate the key observations described above.

Patient No.	Scans				%Change
	Baseline	2 Cycles	4 Cycles	6 Cycles	
130101	229031	274025	165920	185376	-19.1
130102	333650	289575	188461	232001	-30.5
130103	231266	274968	285437	0	23.4
130104	523928	338702	0	0	-35.4
130105	216690	165982	0	0	-23.4
130106	492666	180036	389923	0	-20.9
130107	348011	526396	0	0	51.3
130108	315052	340443	330652	291806	-7.4
130109	813416	870872	0	0	7.1
130110	106348	132937	182732	0	71.8
130111	523828	348642	397824	333911	-36.3
130112 (A)	939842	630513	873334	0	-7.1
130113	211220	209103	152571	0	-27.8
130114 (B)	163116	208185	238774	0	46.4
130115	476943	505000	461913	634501	33.0

Table 7.6: Segmented MPM volumes (mm³)

Patient No.	Scans				Mean
	Baseline	2 Cycles	4 Cycles	6 Cycles	
130101	814	655	684	458	653
130102	935	1011	455	815	804
130103	2557	780	900		1413
130104	1400	1346			1373
130105	1474	1147			1311
130106	1609	1940	985		1511
130107	1466	1314			1390
130108	2067	1461	1656	1832	1754
130109	1846	1754			1800
130110	1063	1233	1543		1280
130111	1288	823	579	930	905
130112 (A)	2538	2491	1624		2218
130113	1471	1161	1145		1259
130114	1572	1301	1281		1385
130115 (B)	1623	1721	1476	1716	1634
Mean					1346

Table 7.7: MPM segmentation computation times (s)

7.4.1 Statistical Validation with MPM-RECIST

The correlation between the segmented tumour volumetric changes and RECIST measure changes can be examined on a scatter plot, given in Fig. 7.9. As the RECIST measures are uni-dimensional, the cubic powers of the RECIST measures must be taken to compare them to changes in volume. The RECIST responses are found according to the following equation, for RECIST measures R_1 and R_2 from two consecutive time point scans, as a percentage:

$$(\Delta RECIST)^3 = \left[\left(\frac{R_2}{R_1} \right)^3 - 1 \right] \times 100\% \quad (7.3)$$

The tumour volume responses between subsequent scans are found in a similar manner and represented as a percentage, for segmented volumes V_1 and V_2 from the same scans:

$$\Delta Vol = \left[\frac{V_2}{V_1} - 1 \right] \times 100\% \quad (7.4)$$

where volumes are computed as a product of transverse slice segmented area and the PVE-corrected slice thickness.

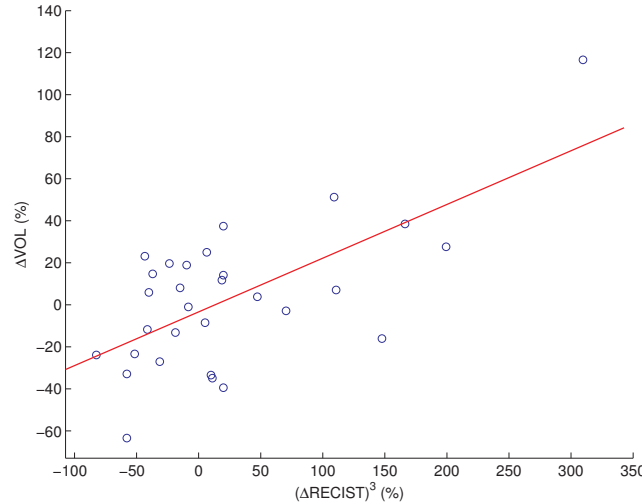


Figure 7.9: Scatter plots showing correlation of VTA and RECIST responses. CC: 0.6593, p-value: 7.4e-5, R^2 : 0.4347. This indicates a strong link between the two measures and tentatively confirms the proposed linear correlation.

The Pearson's correlation coefficient (Eq. 7.1) is found to be 0.6593, with a p-value much lower than 0.05 (7.4e-5). A linear fit is assumed and calculated minimising the least square

Tumour Response	Criterion
Complete Response	disappearance of all target lesion volumes; confirmed patient response requires non-relapsing observations on two occasions, 4 weeks apart.
Partial Response	at least 30% decrease in the total tumour volume; confirmed patient response required as in CR.
Progressive Disease	at least 20% increase in the total tumour volume over the course of the trial monitoring.
Stable Disease	$\leq 30\%$ decrease and $\leq 20\%$ increase in the total tumour volume

Table 7.8: Provisional response criteria for VTA, based on those for MPM-RECIST

differences, with a relatively satisfactory level of confidence (goodness of fit: $R^2 = 0.4347$). This shows that tumour responses found using the segmented volumes are well correlated with those predicted by RECIST.

7.4.2 VTA Response Rate to IV-Vinflunine

Having analysed the utility of VTA in the context of MPM-RECIST, we now compute an updated tumour response to IV-Vinflunine. To do so, we first develop a set of response criteria for VTA. The primary mechanism by which MPM-RECIST assesses tumour response is by examining the change in perpendicular axis of the lesion to thoracic fixtures, or simply the change in pleural thickening. Thus, given the MPM-RECIST response criteria (Appendix B), we provisionally translate them to VTA. Note that at this stage, we are uncertain of the actual volumetric implications of changes in the uni-dimensional MPM-RECIST measure. Thus to maintain a conservative approach at this point, we have retained the same numerical % thresholds for classifying volumetric changes as in MPM-RECIST. The proposed response criteria are listed in Table 7.8. The time observation tumour responses for our patient cohort have been updated accordingly and are given in Fig. 7.10. The patient response rates are shown as a percentage of responders (CR/PR), out of the MPM population ($= 15$) used in our study.

We note that the provisional VTA response criteria presented in Table 7.8 produces different responses to MPM-RECIST, for many time point and patient observations. This argues for

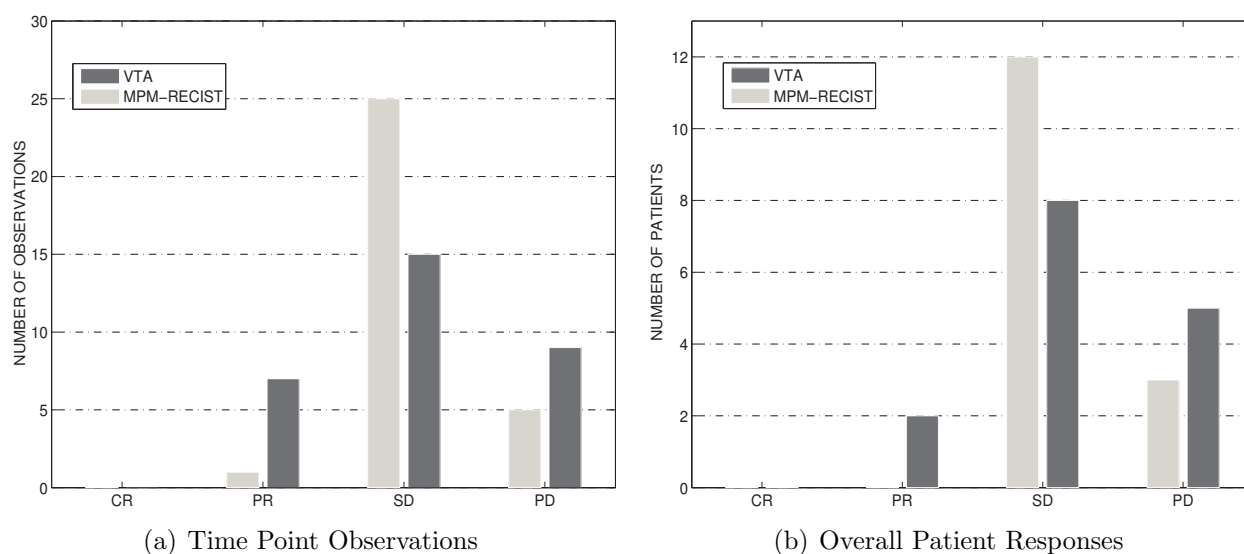


Figure 7.10: Tumour responses based on VTA as compared to MPM-RECIST. Note that with the same thresholds applied, VTA is less conservative than MPM-RECIST for classifying responses. The response rate to the IV-Vinflunine treatment is increased to 14% (VTA), compared to the original 0% (RECIST). Additional clinical information (e.g individual patient survival time) is required to assert the accuracy of either result.

MPM-RECIST	Volumetric Change				
	CR	PR	SD	PD	Total
CR	0	0	0	0	0
PR	0	0	0	0	0
SD	0	6	14	6	26
PD	0	0	2	3	5
Total	0	6	16	9	31

Table 7.9: Time point observations comparison of MPM-RECIST and VTA on the IV-Vinflunine data, assuming the provisional response criteria in Table 7.8. Large discrepancies are observed, prompting more extensive studies on the topic of response criteria for VTA.

more extensive studies on the topic of response criteria for volumetric tumour assessment. In Fig. 7.11, we show the effect of varying the PR and PD criteria thresholds on patient response classification. However, we do not at present have sufficient clinical data to suggest any meaningful criteria thresholds for VTA. To validate any proposed response criteria for VTA, clinical study and verification of a much larger data sample size is required.

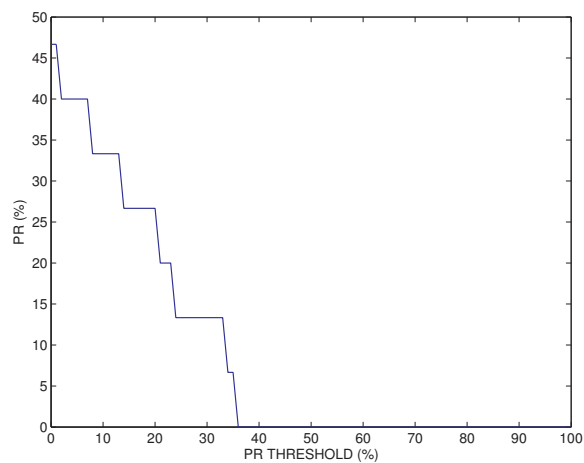
It should be noted that although MPM-RECIST is currently used as the standard clinical protocol for assessing MPM, due to a myriad of unresolved issues such as high intra- and inter-observer variabilities, measurement ambiguity and the lack of comprehensive clinical verification¹, it can not be considered as a reference truth for our design of VTA criteria.

In this section, we have presented steps to validating our proposed VTA method based on the MPM-RECIST scheme. A strong linear correlation between changes in segmented tumour volume and RECIST measures was statistically established. A set of provisional response criteria is compiled based on the RECIST criteria. However, a great degree of discrepancy is found between the two response assessment methods. This motivates the need to acquire and analyse a much larger data size for proposing more meaningful volumetric response criteria. In the following section, we suggest a number of additional prognostic benefits offered by the volume segmentation, in addition to its use in VTA for tumour response assessment.

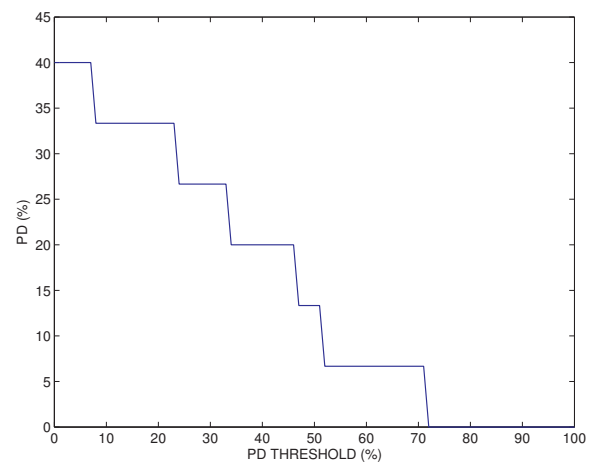
7.5 Quality of Life Assessment

The quality of life (QOL) is a clinical measure defining the well-being of the patient and forms an integral part of disease prognoses. The change in QOL assesses the palliative (symptom-alleviating) effect of a medical treatment and is usually as important as patient survival and pathological (i.e tumour response) benefits for evaluating the clinical potentials of a treatment. This is especially true for many chemotherapeutic solutions to MPM, whose main aim remains of a palliative nature. For MPM patients, pulmonary function impairment and pleural effusion are two important factors affecting patient QOL. By convention, both are measured using

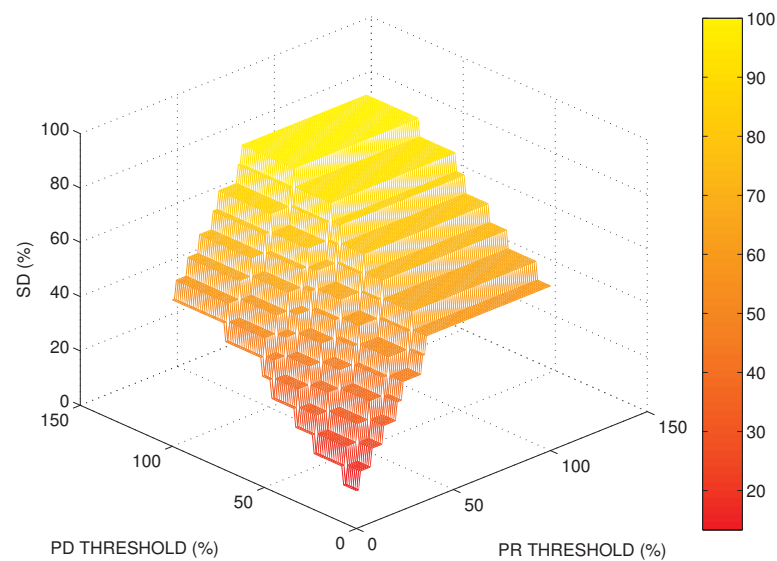
¹Although the original and 1.1 RECIST response criteria were devised as a result of large number of patient data studies, it is not true for the MPM-RECIST, where only 73 patients were employed to support the proposed thresholds [12].



(a) Partial Response/Overall Patient Response Rate



(b) Progressive Disease



(c) Stable Disease

Figure 7.11: The effect of varying response criteria thresholds on the overall patient response results. To validate any proposed response criteria for VTA, clinical study and verification of a much larger data sample size is required.

separate clinical examinations. In this section, we demonstrate ways in which CT scans, along with the segmentation method we have developed in this thesis, are able to assist in measuring these two factors. It should be noted that RECIST, as an one-dimensional measurement system developed specifically for tumour assessment, can not be applied to quantifying other tissues in the thoracic region without relevant large volume patient studies and validation.

7.5.1 Pulmonary Function Impairment

Late stage MPM is usually characterised by severe dyspnoea. Pulmonary function tests (PFT) quantify this symptom and provide an indicator of the extent of respiratory impairment caused by the tumour. The extent of the symptom can be directly assessed via the functional vital capacity (FVC), or the maximum amount of air a person can expel from the lungs after a full inspiration. Under regular clinical settings, FVC is measured by a spirometer. The relationship of FVC to other pulmonary function measures, as given in a spirometer reading, is illustrated in Fig. 7.12. By monitoring FVC over time, a measure of respiratory impairment be derived and forms part of treatment assessment and planning. For MPM, studies [12][57] have shown a strong correlation between spirometric measurements of FVC and the tumour responses based on RECIST.

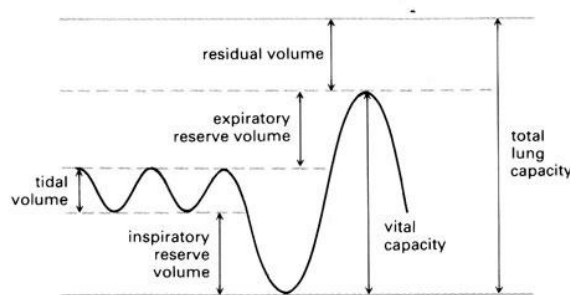


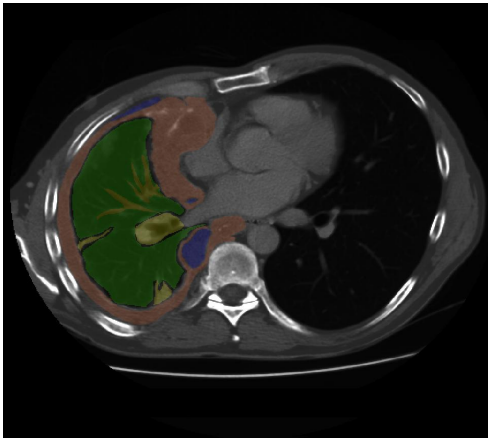
Figure 7.12: Diagram showing lung capacity measures in a spirometer output, which are useful in analysing the pulmonary function at any point in time.

In MPM, pleural lesions affect pulmonary function by inflicting damage on the alveoli, causing their collapse, which reduces the gas exchange at the alveolar sites and subsequently capacity of the affected lung; a condition known as chronic atelectasis. Regions of the lung

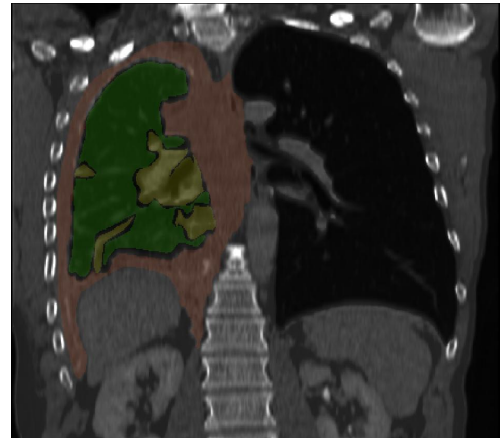
where alveolar collapse has taken place are referred to as collapsed lung and usually appear on CT image with an intensity that is very similar to that of the tumour (as shown in Chapter 3). On the other hand, the part of the lung that remains filled with air is named aerated lung and is easily distinguishable from nearby tissues on CT scans. Delineations of both tissues for scans of Patient A from Section 7.3 are given in Fig. 7.13. Clinically, we understand that the change in lung capacity as a result of the disease is closely related to the change in volume of the aerated lung. Thus, we may derive a measure for the disease severity by segmenting for the aerated lung and computing its changes in volume over time. This may provide a measure of the pulmonary function impairment.

As observed in Fig. 7.13, the deleterious effect of the tumour is directly linked to the amount of atelectasis and shrinkage of the aerated lung. The tumour responses derived from both MPM-RECIST and VTA, ie partial response after two trial cycles, is reflected by the slight decrease in atelectatic lung. Relapse occurs after 4 cycles and is shown by the gradual reappearance of the atelectatic lung section. Note that the 2-D image slices are not perfectly aligned and give only a rough estimate of the 2-D pathological dynamics. These images are subsequently segmented using BE-RW and presented in Fig. 7.14. Here the volume changes are in accordance with our response expectation and suggest an inverse linear relationship between the volumes of tumour and aerated lung. As we observe from the manual segmentations in Fig. 7.13, for Patient A, it is obvious that the aerated lung volume is directly linked to that of the surrounding tumour and atelectatic lung. Therefore, an increase in either measure would be likely reflected in a decrease in the aerated lung volume and vice versa.

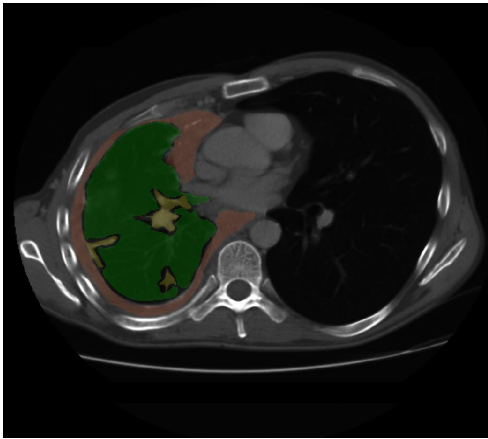
Discrepancies are found between the aerated lung segmentation and manual delineations. In this case, because we are no longer taking the manual segmentations as our reference truth, we analyse this situation further without simply dismissing it as an inaccuracy on the part of the segmentation method. The original CT scans given in Fig 7.13 typically show the air-filled aerated lung space as low attenuation. However, by visual inspection, there are thin strands of medium attenuation outside the regions labeled as collapsed/atelectatic lung. It is likely that these represent sites of more minor atelectasis. However, due to PVE in both the planar and



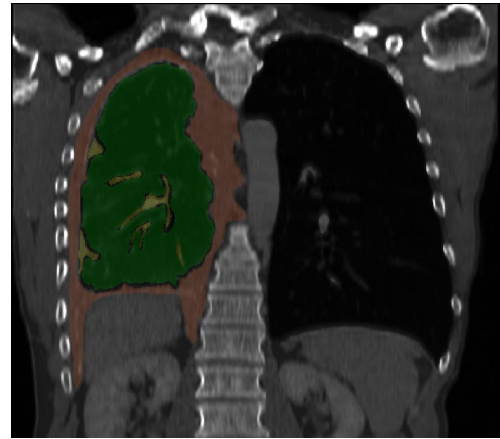
(a) Patient A Baseline Scan - Axial View



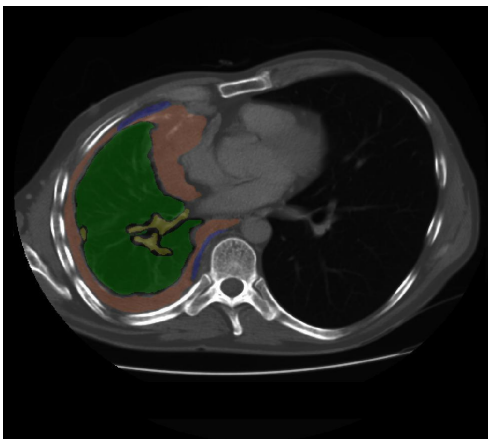
(b) Patient A Baseline Scan - Coronal View



(c) Patient A After 2 Cycles - Axial View



(d) Patient A After 2 Cycles - Coronal View

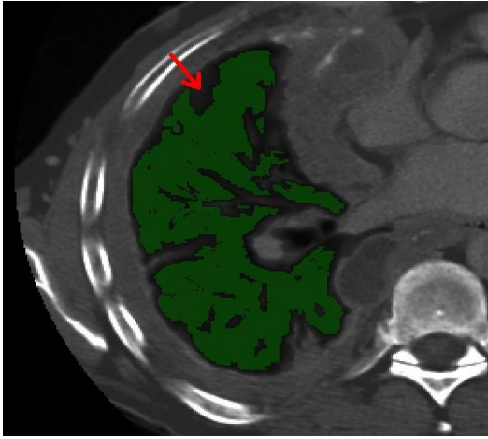


(e) Patient A After 4 Cycles - Axial View

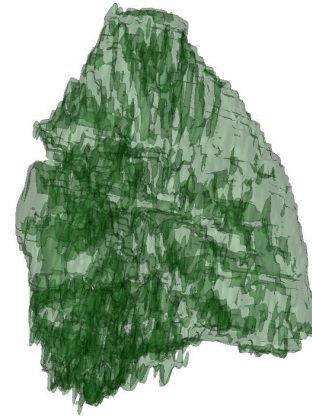


(f) Patient A After 4 Cycles - Coronal View

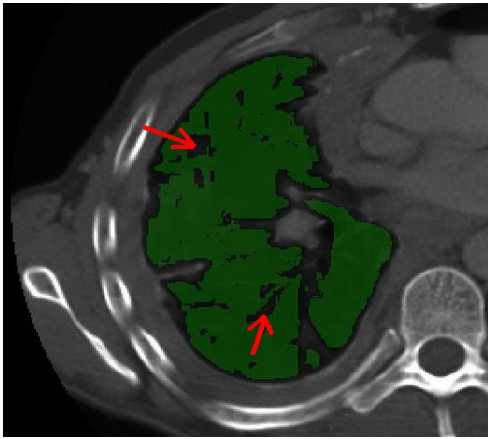
Figure 7.13: Manually delineated scans showing changes in aerated lung volumes over the trial cycles for Patient A. All images shown are PVE-corrected ($s.f_{x,y} = 2, s.f_z = 4$). Aerated lung, atelectatic lung, tumour and effusion fluid are denoted by colours *green*, *yellow*, *orange* and *blue*, respectively. The atelectatic effect of the tumour is directly linked to the amount of atelectasis and shrinkage of the aerated lung. Understanding from the tumour responses derived earlier from MPM-RECIST and VTA, partial response after two trial cycles is reflected by a slight decrease in atelectatic lung. Relapses occur after 4 cycles and is shown by the gradual reappearance of the atelectatic lung section. These images are not perfectly aligned and give only a rough estimate of the 2-D pathological dynamics.



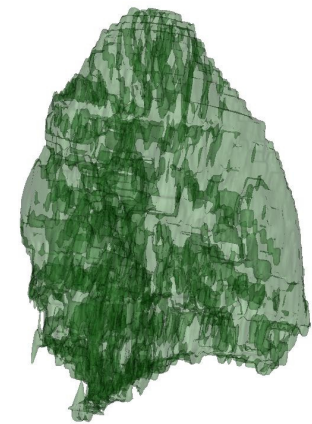
(a) Patient A Baseline Scan - Segmented Aerated Lung Slice



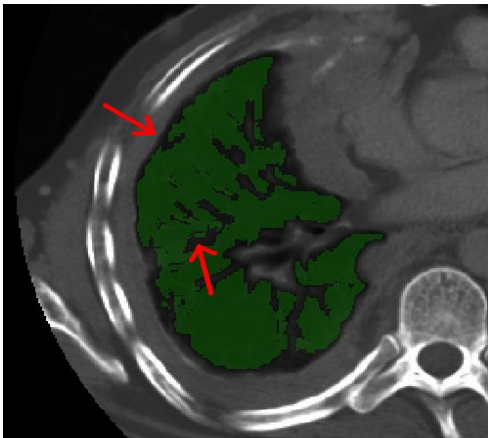
(b) Patient A Baseline Scan - Segmented Aerated Lung Volume: 785390mm^3



(c) Patient A After 2 Cycles - Segmented Aerated Lung Slice



(d) Patient A After 2 Cycles - Segmented Aerated Lung Volume: 924382mm^3



(e) Patient A After 4 Cycles - Segmented Aerated Lung Slice



(f) Patient A After 4 Cycles - Segmented Aerated Lung Volume: 781669mm^3

Figure 7.14: Aerated lung segmentation results for Patient A. Segmented aerated lung is shown in *green*. The quantitative volume changes are in accordance with our response expectation and suggest an inverse linear relationship between the volumes of tumour and aerated lung. Note the discrepancies of the segmentation with the manual delineations, pointed by the *red* arrows. These are likely due to the failure of the manual delineations in accurately outlining all the relevant sites of atelectasis. We investigate this further in Fig. 7.15.

Patient No	Scans				%Change
	Baseline	2 Cycles	4 Cycles	6 Cycles	
130101	551415	362815	330961	264291	-52.1
130102	438238	531818	768073	642913	46.7
130103	425879	440948	524215		23.1
130104	542953	642182			18.3
130105	752749	785421			4.3
130106	899005	973485	712345		-20.8
130107	370962	207853			-44.0
130108	953375	943286	999846	986213	3.4
130109	320288	258563			-19.3
130110	189490	201543	164587		-13.1
130111	328705	328333	284305	293594	-10.7
130112 (A)	785390	924382	781669		-0.5
130113	610785	574885	573810		-6.1
130114	680308	731926	711215		4.5
130115	538730	564389	464048	351071	-34.8

Table 7.10: Segmented aerated lung volumes (mm³)

axial direction, manually delineating all the atelectatic sites is very difficult, if not impossible, as well as being overly tedious. Our method, on the other hand, is corrected for PVE in all three directions and automatically considers all sites of potential atelectasis. Therefore, the observed discrepancies between the computed and manual segmentations are likely to be caused by the failure of the manual segmentations in accurately defining the relevant voxels. We verified this by applying NPW to find the PDF of the segmented aerated lung, the overall lung region involved (as specified by the tumour inner boundary in this case) and the unsegmented portion of this region. The PDFs are given and analysed in Fig. 7.15. Complete experimental findings on aerated lung segmentation are given in Tables 7.10 and 7.11.

We investigate the correlation between tumour and aerated lung volumes. This relationship is examined statistically in Fig. 7.15 with a scatter plot and linear regression. A linear correlation is found with fair confidence ($R^2 = 0.3733, p < 0.0005$) between aerated lung response and VTA, though a much weaker correlation appears to exist between VTA and MPM-RECIST.

The aforementioned linear correlation with segmented tumour volumes suggests that we

Patient No	Scans				Mean
	Baseline	2 Cycles	4 Cycles	6 Cycles	
130101	1103	1143	1087	936	1067
130102	691	613	691	663	664
130103	420	456	509		462
130104	993	841			917
130105	871	832			851
130106	694	820	557		690
130107	1169	1366			1268
130108	962	1083	1017	714	944
130109	1182	1456			1319
130110	1074	1038	978		1030
130111	867	914	631	586	749
130112 (A)	802	799	627		742
130113	1179	1128	1003		1103
130114	978	846	923		915
130115	736	666	788	673	716
Mean					872

Table 7.11: Aerated lung segmentation computation times (s)

should use VTA for analysing the aerated lung results. We proceed onto verifying our proposed way of using aerated lung volume as a measurement of response to treatment by plotting its change for all patients over the course of treatment cycles, categorised by their responses to IV-Vinflunine, as per VTA with the provisional response criteria, in Fig. 7.17. The aerated lung change shows an increasing trend for the one responding (PR) patient whereas it decreases in non-responding patients (SD and PD). More specifically, after 4 cycles of treatment and beyond, patients with SD showed a somewhat less sharp decrease in aerated lung volume than those with PD. This is in line with clinical expectations. Patients responding to the treatment had reduced tumour bulk, hence a similar reduction in atelectasis. This increased alveolar recruitment and allowed for better aeration with a corresponding increase in the volume of lung observed on the image. Amongst the non-responding patients, those with SD experienced little tumour change over time thus showing a less reduction in aerated/usable lung space. However, it should be noted that even in patients with highly stable disease with absolutely no changes over time, factors such as tumour shape change and fluid secretion may still influence

atelectatic activity and alter the aerated lung volume. We investigate the latter factor in the next subsection. In order to support these observations from a wider clinical point of view, we would need more patient data for a better statistical analysis, especially with additional responding patients in our analysis.

From an imaging perspective, we found that it is possible to apply our segmentation method to establish the aerated lung volume from the image scan, and use it as a measure of pulmonary function. The air-filled nature of the aerated lung rendered our computer-aided segmentation algorithm highly accurate, especially when coupled with the PVE-correction scheme introduced in Chapter 6. This is somewhat analogous to using vital capacity in PFT, though the clinical justification of this would require validation with the actual spirometer readings. We also note that in order to attach any clinical utility to the volume, it would require the application of a standard protocol regulating patient breathing and other relevant factors during image acquisition. Currently, the use of image scans to assess pulmonary function is under investigation and is already applied to dynamic CT [43]. In addition to evaluating patient QOL, past works [40] have shown the prognostic value of segmented lung volumes for MPM patients, such that patients with a shrinking aerated lung are found to have a shorter average survival than those with an expanding lung aeration.

7.5.2 Pleural Effusion

In MPM, the symptom of dyspnoea can also be caused by the build-up of excessive fluid between pleural layers of the affected lung, commonly referred to as pleural effusion. The fluid is secreted by the MPM tumour and contained in the interstitial space between the tumour and pleura. Associated pleural inflammation also induces chest pain in many patients. Thus the level of pleural effusion is a strong influence on patient QOL and is an important factor in treatment planning. Chest drainage followed by pleurodesis is typically used as a palliative measure to alleviate the symptoms of pleural effusion in MPM patients.

As we have already seen in images presented in this thesis including the one in Fig. 7.13(a), the segmentation of the fluid is challenging because it is almost always surrounded by tumours

in small and scattered quantities and thus is highly vulnerable to PVE. Correcting for this, we applied our BE-RW segmentation method to the images. The segmentation results are given with respect to the other segmented tissues (MPM tumour and aerated lung) in Fig. 7.18. The position of the segmented tissues in the rib cage is shown in Fig. 7.19, where the bones are effectively segmented out by simple thresholding.

Unlike tumour and aerated lung volumes, pleural effusion is not a validated measure of disease progression and may relate to various other factors. For this reason, effusion volumes presented in Fig. 7.18 do not appear to have a correlation with the time observations of either MPM-RECIST or VTA. Between the scans, palliative measures such as fluid drainage may have been employed to reduce effusion volume and is beyond our control in this experiment. However, the computed fluid volumes from our segmentation can be clinically useful for planning and monitoring palliative treatments.

In summary, we have reviewed the use of image segmentation for assessing the two important QOL factors in MPM, namely pulmonary function impairment and pleural effusion. In addition to its clinical benefits, addressing the issue of QOL is also important from an economic perspective. The quality-adjusted life year (QALY) takes the QOL factors into account for any added survival benefit offered by a treatment. As a reference, the National Institute of Clinical Excellence (NICE) recommends a target of £20,000-£30,000 per QALY for funding new treatment schemes through the National Health Service (NHS) [50].

7.6 Summary

In this chapter, we first analysed the accuracy and robustness of MPM-RECIST, the tumour assessment method used in the IV-Vinflunine phase II trials. In addition to recognising the time-consuming nature of applying the currently used RECIST scheme to image data, we examined its intra- and inter-observer variabilities and concluded that it is vulnerable to both and is therefore not robust for quantifying MPM. Reasons for observer disagreements in RECIST include errors in tumour measurements; errors in selection of measurable targets, and radiologic

technical problems such as the persisting partial volume effect and other image artefacts. Such disagreements may distort the assessment of treatment response in clinical trials. For this reason, we applied the numerous advanced imaging tools we have developed throughout the thesis to our complete medical image dataset, with an aim to seek a way to substitute or improve on the RECIST system. The BE-RW segmentation method was in most cases able to accurately define the MPM tumour volumes in the images. Validating this against the set of RECIST measures taken by a clinical radiology specialist, we found the segmented tumour volumes to be statistically significant and showed an affirmative linear correlation to MPM-RECIST. A fully automated computation time of between 20 to 30 minutes also makes our method very applicable in a practical setting and avoids the tedious procedure of measuring tumour dimensions, as required in MPM-RECIST.

In essence, RECIST was designed as a way to reflect the tumour volumetric changes, before an automated algorithm could be established to accurately and reliably segment such changes. Therefore, so long as we establish a good way to achieve such targets in a volumetric tumour segmentation, MPM-RECIST should be easily replaced. However, having investigated the issue of response criteria, we note that proposing any thresholds to classify patient tumour response would require the study of a much larger dataset, given that more than 6,000 patient data were originally studied for the stipulation of RECIST.

We also introduced in this chapter possible ways of applying our segmentation method to study two important QOL measures in MPM, namely pulmonary function impairment and pleural effusion. They both have important clinical implications, so addressing these problems may potentially lead to better treatment planning and evaluation.

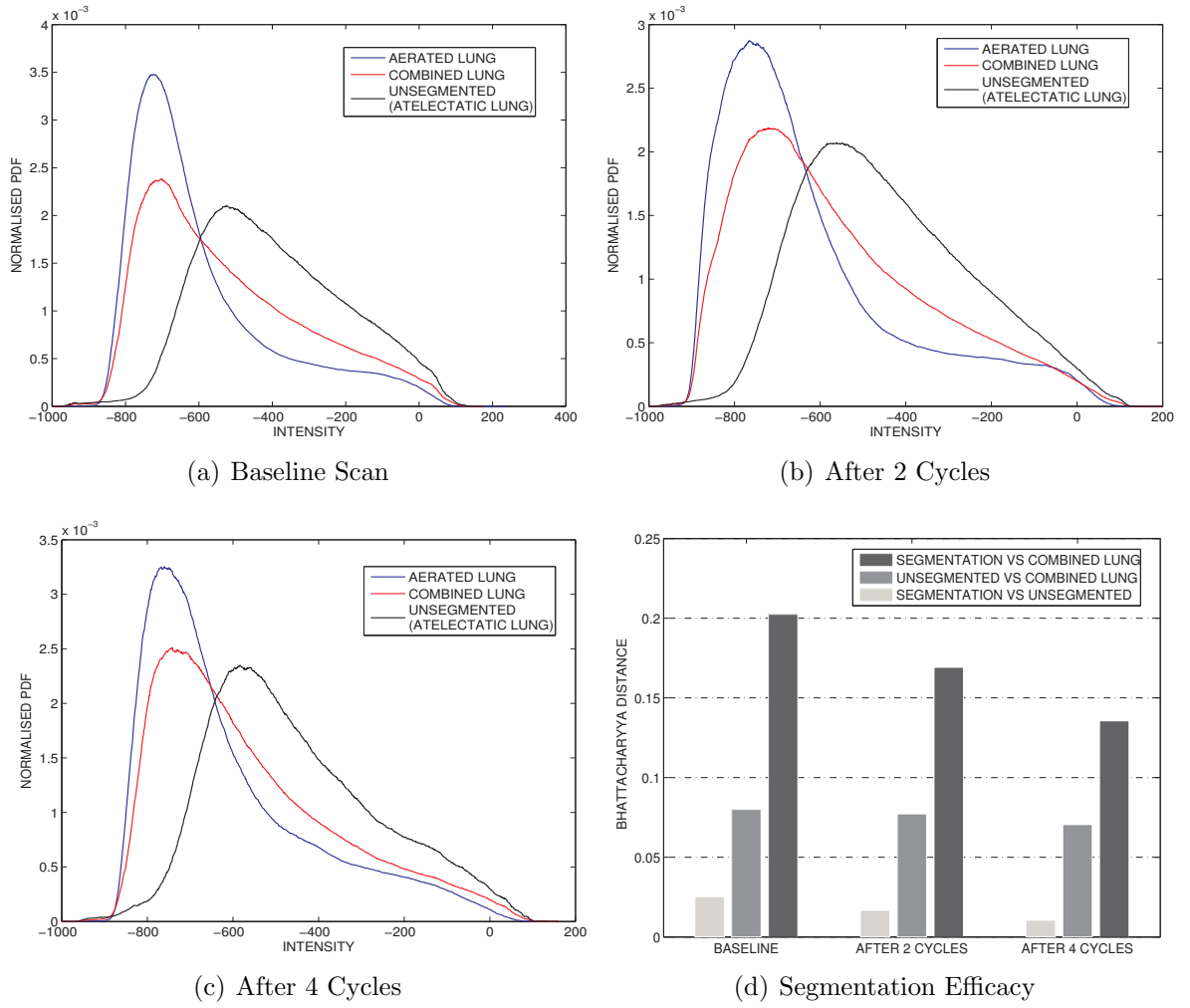


Figure 7.15: PDFs of the segmented aerated lung, combined lung region and the unsegmented portion. The combined lung region is decomposed by the segmentation into two distributions of distinct modes, respectively representing the aerated lung and the unsegmented region which would correspond to atelectatic lung. This is quantitatively assessed using Bhattacharyya distance (Eq. 4.9) measures, which shows our segmentation is successful in dividing the combined lung into two distinct classes and thus separating the aerated lung from sites of atelectasis.

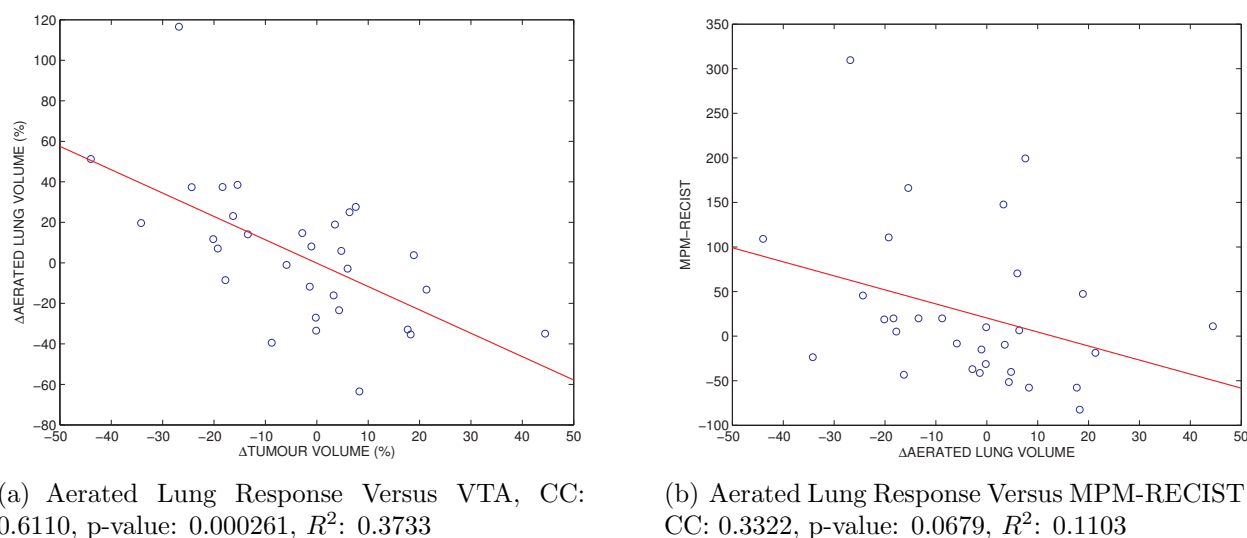


Figure 7.16: Correlation of aerated lung response with VTA and MPM-RECIST. As expected, a notable inverse linear correlation is found between aerated lung and VTA responses. A similar correlation, though much weaker, is found with MPM-RECIST.

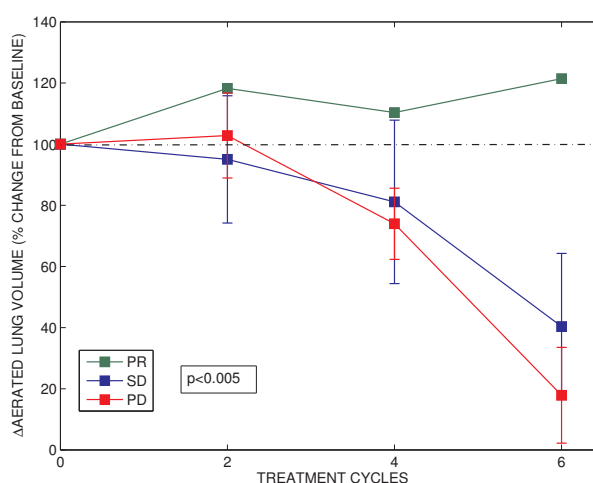
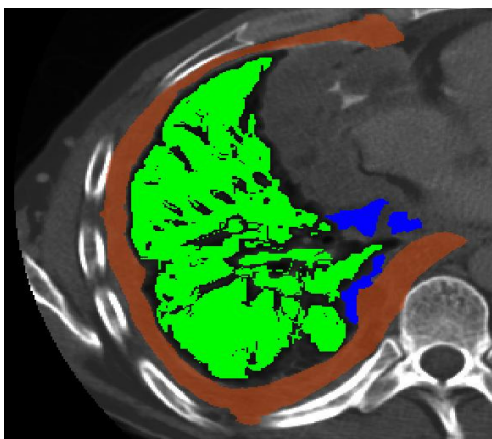
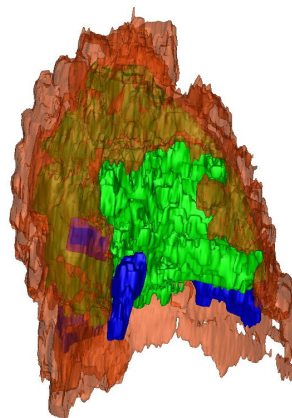


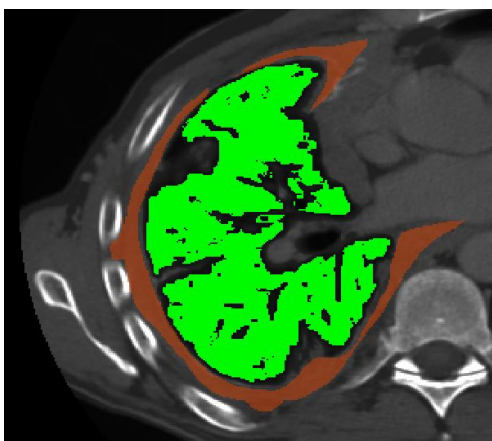
Figure 7.17: Volumetric change in aerated lung over the treatment cycles as a measure of pulmonary function impairment, for all patients, categorised by patient responses to IV-Vinflunine, as per MPM-RECIST and averaged. Note the difference in aerated lung change direction between responding (PR) and non-responding patients (SD and PD); p-value: 0.0064.



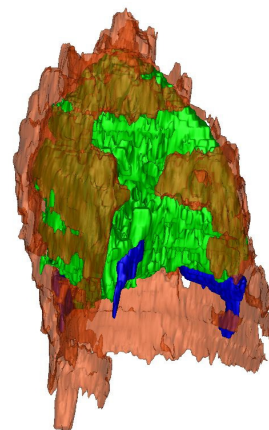
(a) Patient A Baseline Scan - Combined Segmented Slice



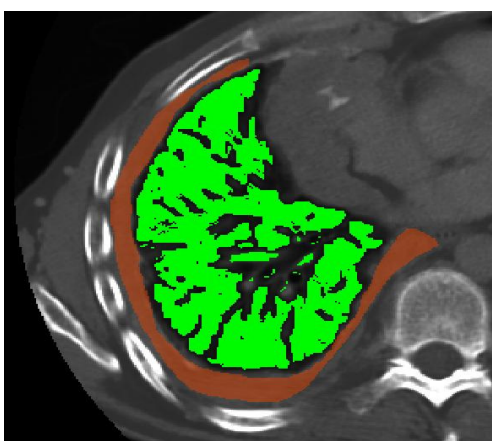
(b) Patient A Baseline Scan - Combined Segmented Volumes, Effusion Volume: 447323mm³



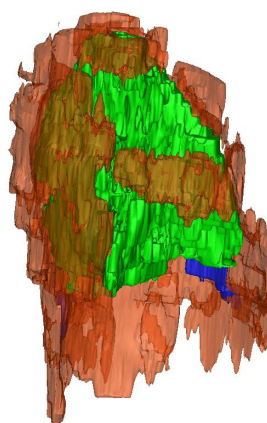
(c) Patient A After 2 Cycles - Combined Segmented Slice



(d) Patient A After 2 Cycles - Combined Segmented Volumes, Effusion Volume: 26323mm³



(e) Patient A After 4 Cycles - Combined Segmented Slice



(f) Patient A After 4 Cycles - Combined Segmented Volumes, Effusion Volume: 17352mm³

Figure 7.18: Combined tissue segmentation results for Patient A. Segmented aerated lung, atelectatic lung and tumour are denoted respectively by the colours *green*, *yellow* and *orange*. Effusion volumes vary between scans and do not appear to correlate with the time observations of either MPM-RECIST or VTA. Palliative measures such as fluid drainage may have helped reducing the effusion volume and is beyond our control in this experiment.

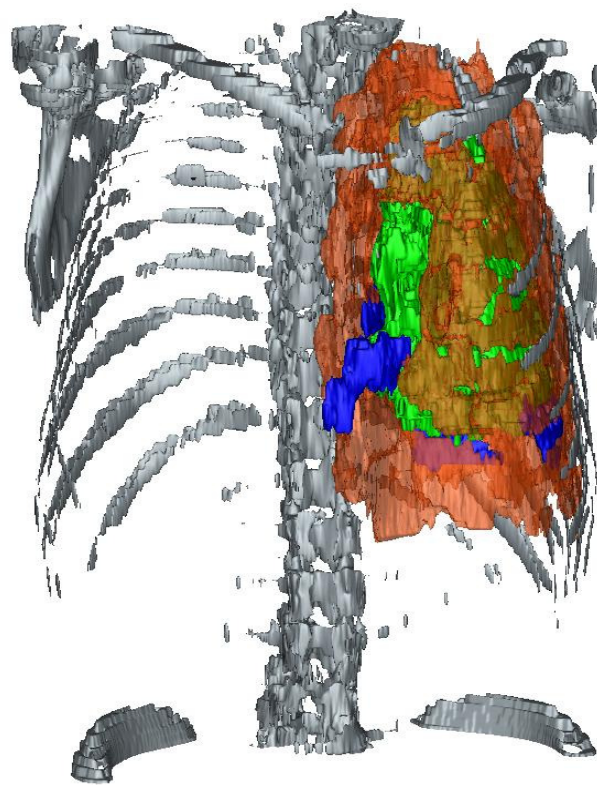


Figure 7.19: Combined tissue segmentation with its location in the rib cage. The bones are effectively segmented by a simple procedure of intensity thresholding.

Chapter 8

Conclusions

In this chapter, we summarise the works presented in the thesis and highlight our main clinical and engineering contributions. We then discuss prospective future works, in relation to our clinical problem and to the field of medical vision.

8.1 Summary

This thesis intertwines two main themes. The first is to address a clinical problem, namely the establishment of a novel volume-based assessment of malignant pleural mesothelioma (MPM) patient response to chemotherapy, as an alternative to the currently used MPM-adapted Response Evaluation Criteria in Solid Tumours (MPM-RECIST) scheme. This required the development of a novel image segmentation algorithm adapted to measuring MPM from CT scans. The second theme evolves around the general contributions offered by our works on developing such a segmentation algorithm to advancing the understanding of medical vision. We re-visit our works on each of these themes in the following sections.

8.1.1 Solution to the Clinical Problem

In MPM, quantification of a tumour is the key to evaluating the efficacy of chemotherapy drugs such as IV-Vinflunine. In the design of the original experiment [74], MPM-RECIST was applied to provide a uni-dimensional measure of tumour's size. It stipulates that the measurement of

a tumour be made at two different sites perpendicular to thoracic fixtures, on three separate transverse levels which are at least 10mm apart.

We have investigated the reproducibility of MPM-RECIST by examining its intra- and inter-observer variation and concluded that it is vulnerable to both. Given the weak linear correlation between individual site measurements (Pearson's correlation coefficient $CC = 0.2049 \sim 0.5787$; coefficient of determination R^2 ranging between 0.0028 - 0.1031), we conclude that the MPM-RECIST protocol is prone to substantial intra-user variation due to the subjective choice of measuring sites by the user. Similarly, the measurements made by two independent observers have shown a low linear correlation of $R^2 = 0.1450$ and $CC = 0.3808$, indicating a high level of inter-observer variability for the system.

Our findings further showed that the intra- and inter-observer differences account for average variations of 20-30% and 31% of the overall measurement, respectively. The fundamental reasons for these disparities included errors in tumour measurements; errors in the selection of measurable targets, and radiologic technical problems such as the partial volume effect and other image artefacts. Such disagreements may distort the assessment of treatment response in clinical trials. In addition, inconsistencies in the scanning thickness of slice alignment cause additional problems of finding equivalent slices in future scans of the same patient. Lastly, the authors of MPM-RECIST [12] confirm that tumours are likely to be non-evaluable; this imposes a multitude of restrictions from a radiological perspective, such as the use of intravenous contrast enhancing agent, proper timing and scanning delay/breathing etc, all of which must be well coordinated in order to make the scan assessable under RECIST. Consequentially, our study suggests that MPM-RECIST may be less suited to situations where high degrees of robustness and reproducibility are required. This prompts for volumetric changes of the tumour to be directly monitored, and necessitates the need for a computer-assisted volume segmentation that demonstrates sufficient accuracy and robustness to support the clinical needs in MPM treatment analysis.

Previous works on MPM [1][2][21][42][57] have provided strong evidence to justify the use of segmented tumour volumes for assessing tumour responses to treatments, as an alternative

to RECIST/MPM-RECIST. However, we noted that the lack of an MPM-adapted advanced segmentation algorithm has prevented the accurate and efficient computation of the tumour volumes in these referenced works. In response to this, we proposed a computerised image segmentation scheme that assesses tumour responses based on volumetric changes, namely volumetric tumour assessment (VTA).

The low CT contrast between the tumour and its anatomical neighbours, the extensive elongated growth pattern characteristic of MPM, and, as a consequence, the pronounced partial volume effect, collectively prevent the direct application of many established medical vision methods to our segmentation problem. To tackle these difficulties, after initial pilot experiments confirming its applicability to simulated and 2D MPM data, we implemented and extended the recently developed [23] random walk-based segmentation to the 3D space. The method provides good handling of weak tissue boundaries and is able to segment any arbitrary shapes with appropriately placed seeds.

This volumetric extension is attractive in that additional image information from adjacent axial slices is considered in classifying the image voxels. Also, the 3D segmentation does not require exhaustive initialisation on all image axial slices. Specifying pre-defined seeds on a single axial slice is sufficient to produce a full-volume segmentation, though the accuracy drastically improves with the application of more seeded regions. Despite this benefit, defining seed regions on a large number of image slices can be very laborious and substantially increases the level of user interaction.

To address the above issue, we developed a registration-assisted segmentation routine using a non-rigid regularisation method from a recent literature [28]. In addition to reducing the required user interaction by up to 50%, this extra component is useful in cases where ambiguous tumour boundaries are encountered, where the uncertain manual delineations can be avoided with ‘mapped’ tumour knowledge from other scans of the same patient. The computed deformation fields from the image registration also facilitate the 3D visualisation of tumour responses over time, by allowing the mapping of segmented tumours from different points in time.

Unlike most forms of malignant tumours, MPM, due to its growth pattern in the form of pleural thickening, most often exhibits smooth boundaries. A boundary-enhancement regulariser based on the idea of anisotropic diffusion, has been developed and successfully applied to improve the boundary edge performance of the segmentation. Effect of the regularisation is visible in both 2D and 3D cases.

With the aforementioned adaptation and enhancements, our segmentation algorithm offers good performance in segmenting MPM, as judged by its accuracy, robustness, running time and the level of user interaction. Validating against a set of clinician-defined reference truth, results from the non-registration-assisted method demonstrated a DICE overlap of 80% to 90% in most cases, based on initialisation seed contours applied to one in every 5 axial slices. A mean running time of 22 minutes was observed. The registration-assisted method, on the other hand, produced results with a DICE overlap of 55% to 90%. A total (registration and segmentation) running time of approximately 60 minutes was observed.

We addressed the partial volume effect using a method based on reverse diffusion interpolation introduced in [67]. For our study, we have implemented the method in 3D and applied it to cases where highly pronounced PVEs are found, either due to the scanning resolution used or difficult tumour boundaries present.

Finally, based on the tools we have studied and developed, we segmented MPM on a total number of 48 CT thoracic scans, taken from 15 patients drawn from a pool of 65 patients recruited for the IV-Vinflunine phase II trial. A strong linear correlation between changes in segmented tumour volume and RECIST measures was statistically established: correlation of the volumetric changes findings to those of clinically observed MPM-RECIST cubed is found to be $CC = 0.6593$, with a good degree of certainty ($p \ll 0.05$); A linear correlation with least squares fit goodness of $R^2 = 0.4347$ was also seen. This showed that tumour responses found using the segmented volumes are well related to those predicted by RECIST and supported our use of volumetric segmentation from a clinical perspective. We then compiled a set of provisional response criteria based on the RECIST criteria. However, a considerable degree of discrepancy is found between the two response assessment methods. This motivates the need to

acquire and analyse a much larger data size for proposing more meaningful volumetric response criteria.

We have also presented experiments showing the prognostic benefit offered by our volume segmentation algorithm in assessing two key quality of life measures for MPM, namely, pulmonary function impairment and pleural effusion. In particular, we found that the segmented aerated lung volumes are inversely correlated to the tumour changes ($CC = 0.6110$, p-value: 0.000261 , $R^2 = 0.3733$) and exhibit a good link to patient responses to treatment. We also noted that the air-filled nature of the aerated lung renders our computer-aided segmentation algorithm very accurate, especially when coupled with PVE correction. This is comparable to the spirometric examination of vital capacity, though clinical verification would require validation with the actual spirometer readings. Finally, we have segmented the pleural fluid and suggested the use of segmentation as a clinical tool for planning palliative treatments such as pleural fluid drainage and pleurodesis.

In conclusion, based on our experimental findings and their validation to the clinically-defined MPM-RECIST measures, we believe there is good evidence that our segmentation algorithm performs satisfactorily with good robustness in its finding of the tumour volumes in MPM. To suggest it as a general alternative to RECIST, however, necessitates the formulation of a set of VTA-specific response criteria, which would require the analysis of a much larger data size (~ 6500 patients as in the case of original RECIST [19]). In addition, VTA presented many other notable advantages to the conventionally used MPM-RECIST such that compared to RECIST, it is largely void of observer variations and provides the anatomical information needed to visualise the tumour shape. Our segmentation allows a fast semi-automated way to efficiently acquire this information.

Overall, the ideas and results presented in this thesis support strongly the contribution our segmentation algorithm provides in assessing the efficacy of a treatment of MPM and for treatment planning. Note that treatment in this context is not limited to chemotherapy. Our findings have already shown prospects in the use of segmented fluid to planning and evaluating fluid drainage. It is hoped that upon completing the future works discussed later

in this chapter, our suggested algorithm would better justify its clinical utilities, although as with any applications of medical vision, it is very unlikely such computer-aided method would ever fully replace the role of a clinician. We believe, however, the use of advanced tools from medical vision would undoubtedly see more application in solving the ongoing MPM burden in many developed countries and predicted outbreak of MPM cases in developing countries where asbestos continues to be harvested and/or used.

8.1.2 Contribution to Medical Image Analysis

The second central theme of this thesis relates to the contributions of our work to the general advancement and understanding of the field of medical vision.

Initially, to better understand the segmentation problem from a probabilistic perspective, we studied the available tools for probability density function (PDF) estimation. We first explained the non-parametric nature of our problem and reviewed the relevant methods for non-parametric PDF estimation, namely the histograms and kernel density estimators (KDE). Having discussed them in relation to our research goals, we studied the method of Non-Parametric Windows (NPW). Our findings show that NPW outperforms the histograms in its smoothness and KDE in its parameter setting. Because NPW does not require the computationally intensive process of finding an optimal bandwidth in KDE, it presents significant advantage over KDE in terms of its computation time, by many folds. In addition, unlike the histograms and KDE, NPW takes into account important signal properties such as the order of occurrence and band-limitedness of the sample. Both properties are important for tissue reconstruction from discrete image data and are therefore desirable in the analysis of medical images.

We have implemented the NPW method in 1D, 2D and 3D, and conducted a pilot study assessing the burden of segmenting a MPM image purely based on thresholding and its pixel intensity values. The results from NPW showed good accuracy performance and fast computation time, in segmenting the overall image scan and individual thoracic tissues, based on their manual delineations by a clinical expert and offered good evidence of the computational superiority of NPW.

Applying NPW results to image segmentation, we studied the suitability of a Bhattacharyya flow-driven level sets method to our image data. We note that for difficult cases of malignancies such as MPM, scalar intensity values might not be sufficient to support the effective application of image analysis techniques. To this end, we have furthered our understanding of PDF estimation by extending the concept of NPW to vector-valued data. This opens up the possibility to explore other features and information supplied by the medical image scans, such as image texture, heterogeneity and entropy; or any other relevant image features that are commonly considered by clinicians in making their decisions. To this end, we have derived and implemented the theory of NPW estimation for 1D 2-tuple vector signals. Due to the fact that our initially proposed NPW-level-sets segmentation had failed to present sufficient robustness to MPM scans, this idea was not pursued further in this thesis.

The 3D extension of the random walk-based method presented in this thesis extends the application of our segmentation method to many medical vision problems, especially where a good segmentation robustness is needed. In comparison to segmenting a stack of 2D slices, in the 3D case, image data from the adjacent slices are considered, giving rise to a more accurate segmentation less affected by local image noise and artefacts. The extra components we have developed for the random walk-based method may improve the utility of the method in many situations beyond that of MPM. Most notably, smooth segmentation boundaries can now be attained under high noise and PVE using the boundary enhancement regulariser. Moreover, the required level of user interaction is reduced under the registration-assisted routine. Both of these modifications are possible without considerably adding to the computation time.

Lastly, we validated the functionality of a reverse diffusion-based partial volume correction technique to some difficult medical image data such as MPM; this potentially offers a solution for correcting any planar or axial partial volume effect present in the image scans. This technique is notable for its good user manoeuvrability, with many parameters such as scaling factors controlled by the user and can be applied to many situations depending on the accuracy and computation time needs.

In the aforementioned aspects, the many unresolved difficult segmentation problems in

medical imaging are likely to benefit from the advantages offered by our algorithm.

8.2 Future Works

Future works are possible in a number of directions, as explained in the following sections.

8.2.1 Clinical Future Works

In our assessment of inter-observer variability, the second set of RECIST observations was acquired by a non-specialist in the field. Such assessments, however, typically necessitate the comparison of measurements from two or more independent specialist observers to reduce sampling bias. In fact, three specialist observers should be optimally recruited for studies of this kind [48]. Involvement of two other clinical observers is needed in order to better justify our inter-observer findings.

MPM-RECIST was proposed based on a set of 73 patient scans [12]. The majority of past works on MPM volumetric segmentation [1][2][21][42][57] have used patient samples of similar sizes. To better establish the clinical basis of VTA, it is necessary to perform validation with a larger database of patient studies. This is also necessary to support the finding of response categorisation criteria for VTA. Recruitment of additional MPM patients is therefore an important clinical future work. In our study, we have used the volumetric changes to develop VTA as a RECIST alternative. However, as absolute volumetric changes do not always reflect those in the tumour shape and anatomical growth, it is important that the clinician visualise the tumour segmentations directly and study the tumour shape changes over time, in the manner depicted in Fig. 7.8.

In order to better assess the accuracy of our volume segmentation, it would be interesting as a future work to acquire autopsy data to compare to the segmented tumour volumes on the scans taken just before the patient death.

Assessing the efficacy of a treatment generally involves examining the correlation of tumour response to survival rate benefits. Acquiring these data for the patients from the IV-Vinflunine

trials is thus necessary to better demonstrate the utility of VTA as an alternative to RECIST. It may therefore be worthwhile investigating the correlation of our predicted patient responses to visible prognostic measures such as patient survival.

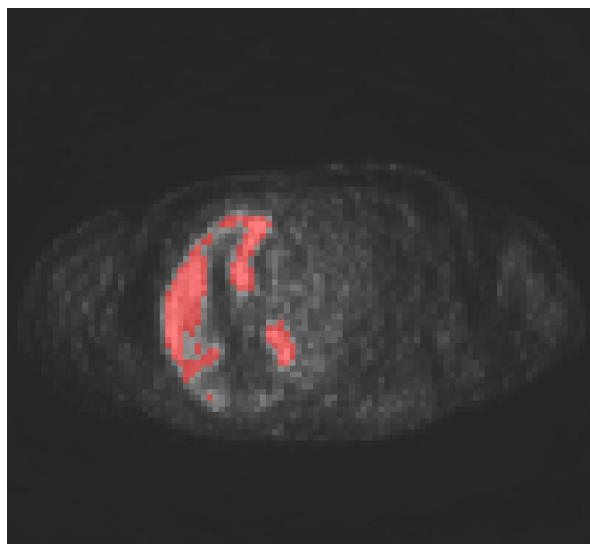
To better justify the use of our segmentation method in monitoring pulmonary impairment, we propose the adoption of a standard clinical protocol regulating patient breathing and other relevant factors during image acquisition at different points in time.

8.2.2 PET-CT

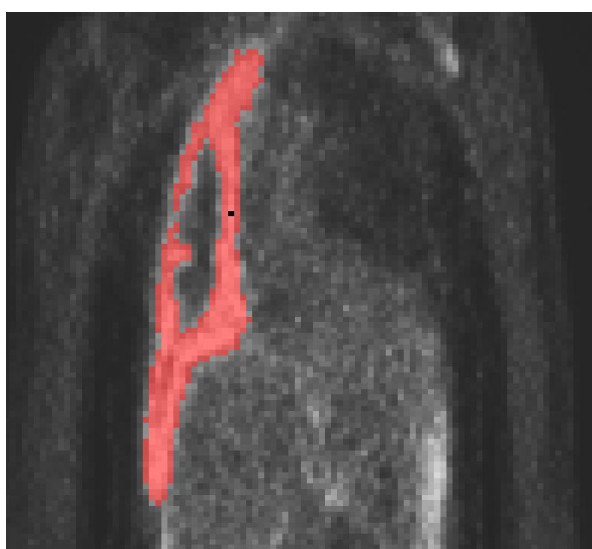
In modern clinical oncology, PET-CT is increasingly used for its ability to acquire both anatomical and metabolic information. The availability of metabolic information from PET scans would assist in our CT segmentation by providing a secondary source of information, which is useful both in validating our VTA results and ascertaining regions of uncertainty (e.g weak soft tissue boundaries) in the segmentation.

One or more of the three quantification measures (SUVmax, glycolytic volume, FDGHetero) introduced earlier in this thesis can be employed for the purpose of VTA validation. To support CT in handling regions of uncertainty, we would first require the segmentation of tumours on PET scans. Note that although regions of high metabolic activity display distinct radioactivity (SUV) levels (and hence image intensity) different from that of the background, not all such regions are tumours because other body tissues such the brain, kidney and heart are also known sites of notable metabolism. Therefore, identifying the tracer-marked regions representative of the tumour is the key to a successful segmentation. This, however, can be easily achieved with the simple application of our random walk-based method with one single-seed initialisation for each class, to allocate the malignancy sites based on the clinical understanding of the disease. Partial volume correction should be applied where necessary. Segmentation result from a MPM PET-CT scan is given in Fig. 8.1.

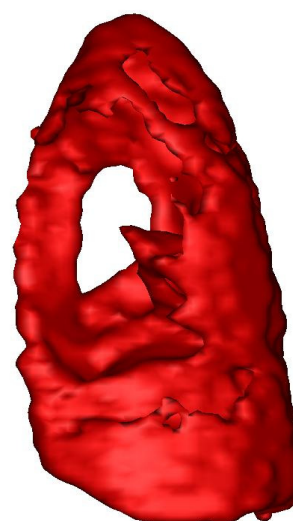
Applying segmented PET data to our CT segmentation requires a point-based cross-modal registration. This allows for the mapping of the segmented PET voxels onto the CT segmentation. The CT segmentation may incorporate the additional information by applying a 2D



(a) PET Axial Slice Segmentation



(b) PET Coronal Slice Segmentation



(c) Segmented PET MPM Lesion

Figure 8.1: Segmented PET scan for MPM tumour. The fast and accurate segmentation of PET images brings light to the future of MPM segmentation, using CT-PET scans of the patient.

feature space classifier-like concept. In PET-CT, because the patient is fixed during the taking of both the PET and CT scans, a rigid registration might be sufficient. PET, due to its longer scanning time as compared to CT, may require breath motion correction.

To study patient response to a treatment, we may also wish to monitor small pulmonary lung nodules. This can be difficult because for example, a small nodule with a given PET standard uptake value (SUV) would be reported as FDG avid; but the same nodule near the diaphragm would likely be reported as much less likely to be FDG avid. In addition to their detection and characterisation, monitoring the presence of small lungs nodules help in planning the appropriate radiotherapy and avoids damage to the mediastinum.

We have also described in this thesis the challenges in diagnosing MPM. It may be possible to apply our segmentation algorithm in combination with PET to assist in MPM diagnosis. In the early stages of the disease, the unique growth pattern of MPM makes its CT boundaries highly vulnerable to PVE. Coupled with a myriad of image noise and scanning artefacts, this could prevent a successful image-based diagnosis of the disease. The differential diagnosis of MPM requires the preclusion of other likely causes of symptoms and CT appearance similar to those of MPM. A summary of these causes is given in Table 8.1. An effective segmentation of MPM on PET scans in combination with CT scan would facilitate an easier clinical interpretation of a patient scan amidst these shortcomings. In this regard, the chances of detecting MPM from routine screening at an early stage could be drastically increased. This may in turn help improving the prognosis of the disease. Also note that most of the differential diagnostic works require the use of CT guided biopsy, which in itself can be aided by better defined MPM lesions on the image scan, achieved through an accurate tumour segmentation.

Currently, only one set of PET-CT data was available to us. It is expected that additional scans would be acquired in the near future as part of an ongoing clinical project on the treatment of MPM. The analysis and incorporation of these imaging data form important future work tasks.

Condition	Description	Differentiating techniques on CT
Benign reactive mesothelial hyperplasia	Caused by lymphatic immune reaction to infections and present as lymph tissue enlargement that leads to mesothelial thickening	Pleural biopsy is usually necessary to distinguish benign mesothelial hyperplasia from MPM; on CT, invasion of underlying tissues favours mesothelioma.
Benign asbestos-related pleural reactions	A benign reaction caused by exposure to asbestos that may lead to pleural effusion, plaques, fibrosis and rounded atelectasis.	Plaques on CT suggest a likelihood for MPM, though this depends on asbestos exposure. Circumferential or nodular pleural thickening, as well as the involvement of mediastinal pleural or enlarged regional lymph nodes, are suggestive of MPM. Biopsy to confirm diagnosis.
Non-small cell lung cancer	The most common form of lung cancer with symptoms similar to that of MPM	CT present the characteristics of the primary tumour, which is evaluated for hilar and/or mediastinal lymphadenopathy and distant metastases. Bronchoscopy and biopsy required for affirmative diagnosis.
Small cell lung cancer	More aggressive form of lung cancer that is likely accompanied by distant metastases (two thirds of the cases)	Chest CT show notable lymphadenopathy and direct mediastinal invasion. Bronchoscopy and biopsy required for affirmative diagnosis.
Metastatic cancer	Late stage spread of various forms of cancer to the pleura	Chest CT may identify primary tumours and show multiple nodes from diffuse micro-nodular shadows to well-defined masses.

Table 8.1: Differential diagnosis of MPM [7]. Note that most of the differential diagnostic works require the use of CT guided biopsy, which in itself can be aided by better defined MPM lesions on the image scan, achieved through an accurate tumour segmentation.

8.2.3 Improving the Algorithm

The immediate next step is the extension and implementation of NPW to estimating joint distributions of 2D signals. This would enable us to apply the vector-valued NPW method to a wider range of applications. It includes a good use of the theories in the field of cross-modal registration where both image intensity and entropy are involved. Additionally, it is possible as a future work to apply the method to estimate the joint distributions of image intensities with other key image quantities such as texture and entropy. This work may present some good insights into enhancing our existing segmentation through the use of multiple intensity-derived image measures, possibly in a form of feature space classifier.

The random walk method itself may be further improved by replacing the conjugate gradients method with a better sparse iterative solver. This would allow a faster computation time and perhaps better control of the anisotropic diffusion through interventions at certain iteration steps with optic flows [87]. Another possibility is the use of subjective contours, a concept proposed by Rutkowski [66], to fill the contour gaps and overcome the issue of random walk ‘leakage’. Using a simple 2-seed elliptical phantom, we conducted a quick benchmark study of the various built-in sparse matrix solvers provided by Matlab® (release 2008b). We have found that the minimum residual method offers the fastest convergence, as shown by Table 8.2. The iterative steps of this method is further analysed and presented in Fig. 8.2. Note the gradual convergence to eventual segmentation of the ring-shaped MPM phantom, after 864 iterative steps.

With regard to the volumetric extension, as a future work we should consider the implementation of 26-voxel connectivity, which may potentially provide a better controlled flow between the planar image slices. This would in turn improve the segmentation accuracy in between initialisation slices, especially those further away from the seeded regions. However, this may increase the method’s computational burden. A parallelised solving of the sparse matrix on the GPU has been previously suggested [23] to counter such effects.

Partial volume correction is another topic requiring further investigation. We have so far only applied one of the four PVE estimators reviewed in this thesis. The reverse diffusion

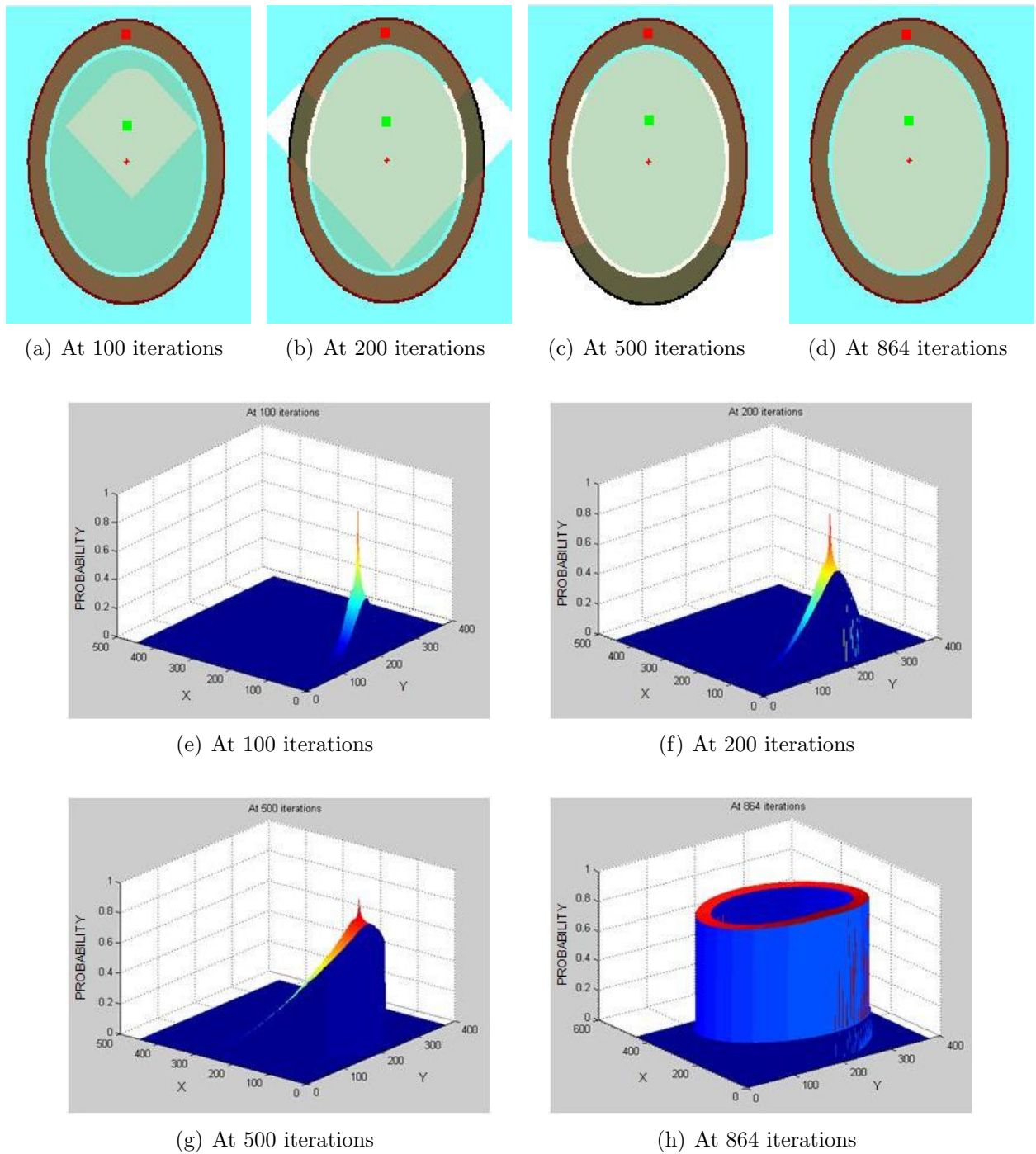


Figure 8.2: Iteration steps of the minimum residual method; (a)-(d) show the 2D contour plots: starting from the initialisation seeds, propagation of the segmented area (shown as *white* region superimposed on the background) over the iterations, and (e)-(h) show the equivalent 3D changes in probability. Note the gradual convergence to segmentation of the ring shaped MPM phantom, after 864 steps. Initialisation seeds: *red* - in, *green* - out.

Method	Matlab Abbr.	Running Time (s)	Observations
Biconjuate gradients	bicg	93.99	normal result, two seed case worked.
Biconjuate gradients stabilized method	bicgstab	111.31	everything worked fine as before, faster convergence at 850 steps
Conjugate gradients squared method	cgs	89.78	not working even at 2000 iterations
Generalised minimum residual method	gmres	N.A	Initial guess issue, failed
Minimum residual method	minres	72.65	worked and faster than bicg, converged
Preconditioned conjugate gradients method	pcg	77.14	failed at 1000 iterations
Quasi-minimal residual method	qmr	111.12	without convergence but worked
Symmetric LQ method	symmlq	71.39	worked fine, without convergence
Least squares method	lsqr	78.16	non-convergence

Table 8.2: A quick benchmark study of various sparse iterative solvers provide by Matlab (release 2008b)

method was chosen primarily because of its generality, user manoeuvrability and the ease of implementation. It may be worthwhile investigating the performance of other key non-parametric PVE estimators [5][89] described in this thesis.

Lastly, having witnessed image registration as an important component of medical image analysis and to our application, we expect its wider usage both in assisting segmentation and tumour response visualisation. In our research so far, we have applied a state-of-the-art deformable framework that best suited to our study needs. Though as with any imaging methods, this method still has its limitations; for example in terms of accuracy, where mapped points land outside of the rib cage of the target image scan. Therefore, it may be worthwhile to examine further ways of registration and select one that better answers to our clinical needs.

Bibliography

- [1] G. Ak, M. Metintas, S. Metintas et al. “Three-dimensional evaluation of chemotherapy response in malignant pleural mesothelioma.” In *EJR*, **4221**, 2009.
- [2] S. G. Armato, J. L. Ogarek, A. Starkey et al. “Variability in mesothelioma tumor response classification.” In *AJR*, **5**, pp.76, 2005.
- [3] S. G. Armato, G. R. Oxnard, M. Kocherginsky et al. “Evaluation of semiautomated measurements of mesothelioma tumor thickness on CT scans.” In *Acad Radiol*, **12**, pp.1301–1309, 2005.
- [4] D. L. Bailey, D. W. Townsend, P. E. Valk et al. “Positron emission tomography: basic science.” *Springer*, 2005.
- [5] M. Ballester, A. P. Zisserman & M. Brady. “Estimation of the partial volume effect in MRI.” In *Medical Image Analysis*, **6**, pp.389–405, 2002.
- [6] J. Barrett & N. Keat. “Artefacts in CT: recognition and avoidance” *RSNA*, 2003.
- [7] BMJ Best Practice. “Mesothelioma: differential diagnosis.” *BMJ Best Practice Website*, URL: <http://bestpractice.bmj.com/best-practice/monograph/1184/diagnosis/differential.html>, Retrieved in March 2012.
- [8] R. Boellaard. “Methodological aspects of multicenter studies with quantitative PET.” In *Positron Emission Tomography*, *Springer*, pp.335–349, 2011.
- [9] J. M. Boone. “X-ray production, interaction, and detection in diagnostic imaging.” In *Handbook of Medical Imaging*, *SPIE Press*, pp.1–77, 2000.
- [10] Y. Boykov, O. Veksler & R. Zabih. “Fast approximate energy minimization via graph cuts.” In *IEEE TPAMI*, **22**, pp.1222–1239, 2001.
- [11] D. Brenner, C. Elliston, E. Hall et al. “Estimated risks of radiation-induced fatal cancer from pediatric CT.” In *AJR Am J Roentgenol*, **176(2)**, pp.289–296, 2001.
- [12] M. J. Byrne & A. K. Nowak. “Modified RECIST criteria for assessment of response in pleural mesothelioma.” In *Annals of Oncology*, **15**, pp.257–260, 2000.
- [13] L. Y. Carreon, S. D. Glassman, J. D. Schwender et al. “Reliability and accuracy of fine-cut computed tomography scans to determine the status of anterior interbody fusions with metallic cages” In *The Spine Journal*, **8(6)**, pp.998–1002, 2008.
- [14] A. P. Chahinian. “Treatment of malignant mesothelioma: radiotherapy and chemotherapy.” In *Mesothelioma*, *Taylor & Francis Group*, pp.209–222, 2005.

- [15] K. Chaisaowong, P. Jager, S. Vogel et al. "Computer-assisted diagnosis for early stage pleural mesothelioma: towards automated detection and quantitative assessment of pleural thickenings from thoracic CT images" In *Methods of Information for Medicine*, **46(3)**, pp.324-331, 2007.
- [16] A. Churg, V. Roggli, F. Gleteau-Salle et al. "World health organization classification of tumours. pathology and genetics of tumours of the lung, pleura, tymus and heart." *IARC Press Lyon*, 2004.
- [17] J. A. Davidson, A. W. Musk, R. Wood et al. "Intralesional cytokine therapy in cancer; a pilot study of GM-SCF infusion in mesothelioma." In *J Immunother*, **21**, pp.389-398, 1998.
- [18] N. Dowson, T. Kadir & R. Bowden. "Estimating the joint statistics of images using nonparametric windows with application to registration using mutual information." In *IEEE TPAMI*, **30(10)**, pp.1841-1857, 2008.
- [19] E. A. Eisenhauer, P. Therasse, J. Bogaerts et al. "New response evaluation criteria in solid tumours; Revised RECIST guidelines (version 1.1)." In *European Journal of Cancer*, **45**, pp.228-247, 2009.
- [20] J. Fahy, A. Duflos, J. Ribet et al. "Vinca alkaloids in superacidic media: A method for creating a new family of antitumor derivatives" In *J Am Chem Soc*, **119(36)**, pp.8576-8577, 1997.
- [21] T. Fraunfelder, M. Tutic, W. Weder et al. "Volumetry: an alternative to assess therapy response for malignant pleural mesothelioma?" In *European Respiratory Journal*, **38**, pp.162-168, 2011.
- [22] D. Freedman & P. Diaconis. "On the histogram as a density estimator: L2 theory." In *Probability Theory and Related Fields*, *Springer Berlin/Heidelberg*, **57(4)**, pp.453-476, 1981.
- [23] L. Grady. "Random walks for image segmentation." In *IEEE TPAMI*, **28(11)**, pp.1768-1783, 2006.
- [24] L. Grady & M. Jolly. "Weights and topology: a study of the effects of graph construction on 3D image segmentation." In *MICCAI*, **I**, pp.153-161, 2008.
- [25] P. Hall, J. S. Marron & B. U. Park. "Probability theory and related fields: smoothed cross-validation." *Springer*, **92**, pp.1-20, 1992.
- [26] W. Hardle, M. Muller, S. Sperlich et al. "Nonparametric and semiparametric models." *Springer*, 2004.
- [27] Health Safety Executive. "Mesothelioma: Overall scale of disease including trends" *HSE Website*, URL: <http://www.hse.gov.uk/statistics/causdis/mesothelioma/scale.htm>, Retrieved in May 2011.
- [28] M. Heinrich, M. Jenkinson, M. Bhushan et al. "Non-local shape descriptor: A new similarity metric for deformable multimodal registration." In *MICCAI*, **6892**, pp.541-548, 2011.
- [29] G. T. Herman, J. Zheng, C. A. Bucholts et al. "Shape-based interpolation." In *IEEE Comp Graph Appl*, **12**, pp.69-79, 1992.

- [30] A. K. Izenman. “Recent developments in non-parametric density estimation.” In *Journal of the American Statistical Association*, **86(413)**, pp.205–224, 1991.
- [31] N. Joshi. “Non-parametric probability density function estimation for medical images”. *D.Phil Thesis*, University of Oxford, UK, 2007.
- [32] N. Joshi & M. Brady. “Unsupervised level set evolution using NP windows density estimation and Bhattacharya distance.” *Technical Report, OUEL*, **2311/09**, 2009.
- [33] N. Joshi, T. Kadir & M. Brady. “Simplified computation for nonparametric windows method of probability density function estimation.” In *IEEE TPAMI*, **33(8)**, pp.1673–1680, 2011.
- [34] T. Kadir & M. Brady. “Non-parametric estimation of probability distributions from sampled signals.” *Technical Report, OUEL*, **2283/05**, 2005.
- [35] E. Kidd. “Intra-tumoral metabolic heterogeneity of cervical cancer” *RSNA*, 2008.
- [36] N. H. de Klerk & A. W. Musk. “Epidemiology of mesothelioma.” In *Mesothelioma, Taylor & Francis Group*, pp.339–349, 2005.
- [37] J. S. Kole & F. J. Beekman. “Evaluation of accelerated iterative x-ray CT image reconstruction using floating point graphics hardware” *Physics in medicine and biology*, **51**, pp.875 2006.
- [38] J. Kolls, S. Freeman, R. Ramesh et al. “The treatment of malignant pleural mesothelioma with gene modified cancer cells: a Phase I study.” In *Am J Respir Crit Care Med*, **157**, 1998.
- [39] L. I. Kuncheva & L. C. Jain. “Nearest neighbor classifier: simultaneous editing and feature selection.” In *Pattern Recognition Letters*, **20(11-13)**, pp1149–1156, 1999.
- [40] Z. Labby, W. Sensakovic, H. Kindler et al. “Prognostic value of automatically segmented lung volumes during chemotherapy for patients with mesothelioma.” In *Intl J of Med Phys Rsch and Practice*, **38**, pp.3464, 2011.
- [41] Y. C. Lee, R. I. Thompson, A. Dean et al. “Clinical and palliative care aspects of malignant mesothelioma.” In *Mesothelioma, Taylor & Francis Group*, pp.111–142, 2005.
- [42] F. Liu, B. Zhao, K. Lee et al. “Assessment of therapy responses and prediction of survival in malignant pleural mesothelioma through computer-aided volumetric measurement on computed tomography scans.” In *Journal of Thoracic Oncology*, **5(6)**, pp.879–884, 2010.
- [43] K. Markstaller, B. Eberle, H. U. Kauczor et al. “Temporal dynamics of lung aeration determined by dynamic CT in a porcine model of ARDS.” In *Br. J. Anaesth.*, **87(3)**, pp.459–468, 2001.
- [44] F. Meng, N. Zhang & W. Wang. “Virtual experimentation of beam hardening effect in x-ray CT measurement of multiphase flow” In *Powder Technology*, **194**, pp.153–157, 2009.
- [45] A. B. Miller, B. Hoogstraten, M. Staquet et al. “Reporting results of cancer treatment.” In *Cancer*, **47(1)**, pp207–214, 1981.
- [46] N. R. Mollet, F. Cademartiri & P. J. DeFeyter. “Non-invasive multislice CT coronary imaging” In *Heart*, **91(3)**, pp.401, 2005.

- [47] K. L. Molnar-Kimber, D. H. Sterman, M. Chang et al. "Humoral and cellular immune responses induced by adenoviral-based gene therapy for localized malignancy: results of a phase 1 clinical trial for malignant mesothelioma" In *Hum Gene Ther*, **9**, pp.2121-2133, 1998.
- [48] J. H. Moltz, S. Braunewell, J. Ruhaak et al. "Analysis of variability in manual liver tumor delineation in CT scans." In *ISBI*, pp1974-1977, 2011.
- [49] S. Nagamachi, H. Wakamatsu, S. Kiyohara et al. "Which FDG PET/CT quantitative indices (SUVmax, metabolic volume, total lesion glycolysis) correlate well with serum tumor markers in NSCLC, colon cancer and pancreas cancer?" In *The Journal of Nuclear Medicine*, **50(2)**, pp.1722, 2009.
- [50] National Institute of Health and Clinical Excellence. "Incorporating health economics in guidelines and assessing resource impact." *National Institute for Health and Clinical Excellence*, pp.54, 2007.
- [51] National Statistics. "Mesothelioma mortality in great britain: estimating the future burden." *London: Health and Safety Executive*, 2003.
- [52] A. K. Nowak, R. J. Francis, M. J. Millward et al. "A novel prognostic model for malignant mesothelioma incorporating quatitative FDG-PET imaging with clinical parameters" In *Clin Cancer Res.*, **16(8)**, pp.2409-2417, 2010.
- [53] Organisation for Economic Co-operation and Development (OECD). "Medical technologies (supply and use)" In *Health at a glance 2009: OECD Indicators*, pp42, 2009.
- [54] A. Papoulis & S. U Pillai. "Probability, random vaiables and stochastic processes" *McGraw-Hill*, 2002.
- [55] P. Perona & J. Malik. "Scale-space and edge detection using anisotropic diffusion." In *IEEE TPAMI*, **12(7)**, pp.629-639, 1990.
- [56] D. L. Pham, C. Xu & J. L. Prince. "Current methods in medical image segmentation." In *Annu. Re. Viomed. Eng. 2000.*, **2**, pp.315-337, 2000.
- [57] C. Plathow, M. Klopp, C. Thieke et al. "Therapy response in malignant pleural mesothelioma-role of MRI using RECIST, modified RECIST and volumetric approaches in comparison with CT." In *Eur Radiol*, **18**, pp.1635-1643, 2008.
- [58] C. Plathow, M. Schoebinger, M. Fink et al. "Quantification of lung tumor volume and rotation at 3D dynamic parallel MR imaging with view sharing: preliminary results." In *Radiology*, **240**, pp.537-545, 2006.
- [59] D. Platten. "Multi-slice helical CT physics and technology." *Impact scan* URL: <http://www.impactscan.org/slides/eanm2002/>, Retrieved in Sep 2011.
- [60] A. Rajwade, A. Banerjee & A. Rangarajan. "A new method of probability density estimation with application to mutual information based image registration." In *Proceedings of Computer vision and Pattern Recognition (CVPR) 2006*, pp.1769-1776, 2006.
- [61] S. J. Riederer, N. J. Pelc & D. A. Chesler. "The noise power spectrum in computed x-ray tomography." In *Physics in Medicine and Biology* **23(3)**, pp.446, 1978.
- [62] B. W. Robinson & R. A. Lake. "Advances in malignant mesothelioma." In *The New England Journal of Medicine* **15**, pp. 353, 2005.

- [63] B. W. Robinson, S. A. Mukherjee, S. A. Davidson et al. "Cytokine gene therapy or infusion as treatment for solid human cancer" In *J Immunother*, **21(3)**. pp.211–217, 1998.
- [64] P. Rogelj, S. Kovacic & J. C Gee. "Point similarity measures for non-rigid registration of multi-modal data" In *Computer Vision and Image Understanding*, **92(1)**. pp.112–140, 2003.
- [65] V. W. Rusch & E. Venkatraman. "The importance of surgical staging in the treatment of malignant pleural mesothelioma." In *Journal of Thoracic Cardiovascular Surgery*, **111**, pp.815–826, 1996.
- [66] W. S. Rutkowski. "Shape completion" In *Computer graphics and image processing*, **9**. pp.89–101, 1979.
- [67] O. Salvado, C. Hillenbrand & D. Wilson. "Partial volume reduction by interpolation with reverse diffusion" In *Int J of Biomed Imaging*, pp.1–13, 2006.
- [68] D. J. Sargent, L. Rubinstein, L. Schwartz et al. "Validation of novel imaging methodologies for use as cancer clinical trial end-points." In *EJR*, **45**, pp.290–299, 2009.
- [69] A. Segal, D. Whitaker, D. Henderson et al. "Pathology of mesothelioma" In *Mesothelioma*, Taylor & Francis Group, pp.143–184, 2005.
- [70] W. F. Sensakovic, S. G. Armato III, A. Starkey et al. "Automated matching of temporally sequential CT sections" In *Medical physics*, pp.3417–3424, 2004.
- [71] J. B. Shrager, D. Serman & L. Kaiser. "Surgery and staging of malignant mesothelioma." In *Mesothelioma*, Taylor & Francis Group, pp.185–199, 2005.
- [72] T. A. Sporn. "Mineralogy of asbestos" In *Malignant Mesothelioma*, Springer-Verlag Berlin, pp.1–13, 2011.
- [73] D. H. Serman, S. M. Albelda, S. Mukherjee et al. "Gene Therapy for malignant mesothelioma - 'suicide gene' and immunological approaches." In *Mesothelioma*, Taylor & Francis Group, pp.223–235, 2005.
- [74] D. Talbot, J. Margery, G. Dabouis et al. "Phase ii study of vinflunine in malignant pleural mesothelioma" In *J Clinical Oncology*, **25**, pp.4751–4756, 2007.
- [75] B. Talha. "Health experts concerned over Indias asbestos industry" In *The Lancet*, **375**, pp.626–627, 2010.
- [76] P. Thiesse, L. Ollivier D. DiStephan-Louineau et al. "Response rate accuracy in oncology trials: reasons for interobserver variability." In *Journal of Clinical Oncology*, **15**, pp.3507–3514, 1997.
- [77] P. Therass, S. G. Arbuck, E. A. Eisenhauer et al. "New guidelines to evaluate the response to treatment in solid tumors." In *J Natl Cancer Inst*, **92**, pp.205–215, 2000.
- [78] P. Torr. "Image Segmentation with Level Sets" *Oxford Brookes University: presentation slides*, 2005.
- [79] R. D. Tunnell, B. W. Millar & G. B. Smith. "The effect of lead time bias on severity of illness scoring, mortality prediction and standardised mortality ratio in intensive care - a pilot study." In *Anaesthesia*, **53(11)**, pp.1045–1053, 1998.

- [80] U.S. Department of the Interior. “Asbestos: world production, by country.” *United States Geological Survey Mineral Resources Program*, URL: <http://minerals.usgs.gov/>, Retrieved in May 2011.
- [81] L. Vogel. “Special report: asbestos in the world.” In *HESA Newsletter*, **27**, 2005.
- [82] N. J. Vogelzang, J. J. Rusthoven, J. Symanowski et al. “Phase iii study of pemetrexed in combination with cisplatin versus cisplatin alone in patients with malignant pleural mesothelioma.” In *J Clin Oncol*, **21**, 2003.
- [83] I. Volkau, F. Puspitasari & W. L. Nowinski. “Ventricle boundary in CT: partial volume effect and local threshold.” In *Int J Biomed Imaging Online*, 2010.
- [84] K. Wagstaff, C. Cardie, S. Rogers et al. “Constrained K-means clustering with background knowledge.” In *Proc. Intl Conf on Machine Learning*, pp.577–584, 2001.
- [85] H. Y. Wang & X. R. Zhang. “Advance in clinical diagnosis of mesothelioma of pleura.” In *Journal of Medical Research, China*, **33**, 2004.
- [86] M. Watanabe, K. Shimizu, T. Omura et al. “A new high-resolution PET scanner dedicated to brain research.” In *IEEE Trans Nuclear Science*, **49(2)**, 2002.
- [87] J. Weickert & C. Schnorr. “Variational optic flow computation with a spatio-temporal smoothness constraint.” In *Journal of Mathematical Imaging and Vision*, **14**, pp.245–255, 2001.
- [88] World Health Organization. “WHO calls for prevention of cancer through healthy workplaces” *World Health Organization Website*, URL: <http://www.who.int/mediacentre/news/notes/2007/np19/en/index.html>, Retrieved in May 2011.
- [89] Y. Zhang. “Partial volume classification with sub-voxel accuracy”. *D.Phil Thesis*, pp.116–129, University of Oxford, UK, 2001.
- [90] Y. Zhang, S. Smith & M. Brady. “Hidden markov random field model and segmentation of brain MR images” In *IEEE Transaction on Medical Imaging*, **20(1)**, 2001.
- [91] B. Zhao, L. Schwartz & R. Flores et al. “Automated segmentation of mesothelioma volume on CT scan.” In *Proceedings of Medical Imaging: Image Processing*, pp.866–875. SPIE, 2005.

Appendix A

The International Mesothelioma Interest Group Staging System

Table A.1: Staging of MPM [69]

Stage	Tumour Progression
Stage Ia	T1aN0M0
Stage Ib	T1bN0M0
Stage II	T2N0M0
Stage III	Any T3M0, Any N1M0, Any N2M0
Stage IV	Any T4, Any N3, Any M1
Primary Tumour Index	
T1a	Tumour limited to the parietal pleura, No involvement of visceral pleura
T1b	Tumour involving both the parietal and visceral pleura
T2	Further involvement of diaphragmatic muscle or extension of tumour into the underlying pulmonary parenchyma
T3	Locally advanced but potentially resectable tumour and extension into the mediastinal fat, endothoracic fascia, soft tissues of the chest wall, or pericardium
T4	Locally advanced but unresectable tumour, and diffuse extension of tumour in the chest wall, possibly with associated rib destruction, extension into the peritoneum, contralateral pleura, one or more mediastinal organs, spine, internal surface of the pericardium, or myocardium.
N-Lymph nodes	
NX	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastases
N1	Metastases in the ipsilateral bronchopulmonary or hilar lymph nodes
N2	Metastases in the subcarinal or the ipsilateral mediastinal lymph nodes, including the ipsilateral internal mammary nodes
N3	Metastases in the contralateral mediastinal, contralateral internal mammary, ipsilateral, or contralateral supraclavicular lymph nodes
M-Metastases	
MX	Presence of distant metastasis cannot be assessed
M0	No distant metastasis
M1	Distant metastasis present

Appendix B

Specifications of RECIST as Used in This Thesis

Lesion Measurability Criterion

- 1) for diffuse pleural thickening: diameter $\geq 5\text{mm}$ assessed in 3 different areas; sum of the three measurements $\geq 20\text{ mm}$.
- 2) for nodular lesion and any other lesion: diameter $\geq 20\text{ mm}$ with conventional techniques or $\geq 10\text{ mm}$ with spiral CT scan.

Measurement Methodology

Short axis measurement at two sites of greatest widths, perpendicular to the longest diameter, for three slices of the image scan, any two of which are at least 10mm apart.

Response Criteria

- 1) Tumour response: evaluation of the tumour based on the CT scans.
- 2) Response criteria when RECIST is used

Complete Response: disappearance of all target lesions

Partial Response: at least 30% decrease in the sum of measurements taken for the targets lesions

Progressive Disease: at least 20% increase in the sum of measurements taken for the targets lesions

Stable Disease: $\leq 30\%$ decrease and $\leq 20\%$ increase in the sum of measurements taken for the targets lesions

Responders are patients who demonstrate complete or partial response to the treatment. A confirmed response requires a repeated observation on two occasions at least 4 weeks apart, i.e no relapses. Non-responders are those with either stable or progressive diseases. Progressive disease can be identified in the patient if a net measure increase of 20% is observed during the trial.

Appendix C

Relevant Publications

Mitchell Chen, Emma Helm, Niranjana Joshi and Michael Brady. Random walk-based automated segmentation for the prognosis of malignant pleural mesothelioma, *IEEE International Symposium on Biomedical Imaging*, pp.1978-1981, 2011.

Mitchell Chen, Emma Helm, Niranjana Joshi and Michael Brady. Non-parametric windows-based estimation of probability density function in vector space, *Medical Image Understanding and Analysis*, pp.167-173, 2010.

Mitchell Chen, Emma Helm, Niranjana Joshi and Michael Brady. Automated tumour segmentation for assessment of treatment response in malignant pleural mesothelioma., *UK Radiological Congress*, 2010.

Mitchell Chen, Emma Helm, Niranjana Joshi, Michael Brady and Fergus Gleeson. Feasibility evaluation of thoracic ct automatic segmentation for malignant pleural mesothelioma treatment., *Medical Image Understanding and Analysis*, pp.214-218, 2009.