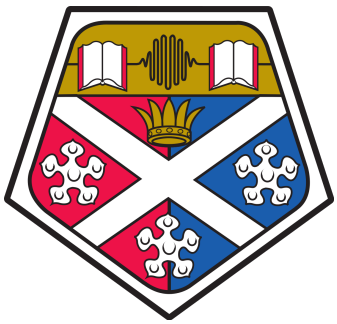
Registration of Pre-Operative Lung Cancer PET/CT Scans with Post-Operative Histopathology Images

Gabriel Reines March

A thesis presented for the degree of
Doctor of Engineering



Department of Electronic and Electrical Engineering
University of Strathclyde
2020

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed: 

Date: May 11, 2020

Abstract

Non-invasive imaging modalities used in the diagnosis of lung cancer, such as Positron Emission Tomography (PET) or Computed Tomography (CT), currently provide insufficient information about the cellular make-up of the lesion microenvironment, unless they are compared against the gold standard of histopathology. The aim of this retrospective study was to build a robust imaging framework for registering *in vivo* and post-operative scans from lung cancer patients, in order to have a global, pathology-validated multimodality map of the tumour and its surroundings.

Initial experiments were performed on tissue-mimicking phantoms, to test different shape reconstruction methods. The choice of interpolator and slice thickness were found to affect the algorithm's output, in terms of overall volume and local feature recovery. In the second phase of the study, nine lung cancer patients referred for radical lobectomy were recruited. Resected specimens were inflated with agar, sliced at 5 mm intervals, and each cross-section was photographed. The tumour area was delineated on the block-face pathology images and on the pre-operative PET/CT scans. Airway segments were also added to the reconstructed models, to act as anatomical fiducials. Binary shapes were pre-registered by aligning their minimal bounding box axes, and subsequently transformed using rigid registration. In addition, histopathology slides were matched to the block-face photographs using moving least squares algorithm.

A two-step validation process was used to evaluate the performance of the proposed method against manual registration carried out by experienced consultants. In two out of three cases, experts rated the results generated by the algorithm as the best output, suggesting that the developed framework outperforms the current standard practice.

Acknowledgements

This multidisciplinary project would have not been possible without the help of many wonderful people. First and foremost, I would like to thank my great supervisory team: Stephen Marshall, for his advice, support and guidance throughout the last four years; Xiangyang Ju, always available to answer any questions, suggest new experiments or engage in technical discussions; Stephen Harrow and Craig Dick, the clinical project leads who conceived this study and from whom I have learnt so much about oncology, pathology and some other finer things in life; and Stuart Parks, who facilitated this EngD position and gave me access to a wealth of clinical engineering expertise.

I would also like to say thanks to my fabulous colleagues and friends at the Medical Devices Unit, for their continuous support and help with a variety of scientific topics (software, informatics, mechanics, electronics), and not-so-scientific (climbing, running, Scottish patter, and many others!). A special mention goes to the Department of Mechanical Engineering (Robin Sayer, Tam Murdoch, John Hodgins, Alan McIntyre), who conceived, designed and built the soft tissue slicing rig, and helped me develop the 3D-printed phantom mould.

I would like to extend my gratitude to the staff at the Beatson West of Scotland Cancer Centre: Ronan Valentine, for his help with getting the PET/CT imaging data; Clare McKeown, who provided us with the CT calibration phantom; Eithne MacPherson and Suzanne Smith, for giving us access to the CT machines; and all the treatment radiographers, who operated the scanners. From the Golden Jubilee National Hospital, special thanks to Alan Kirk, thoracic surgeon and local project coordinator; and to Julie Buckley, Elaine Matthews and Elizabeth Boyd, research nurses who approached and consented all patients enrolled in this study. I would also like to acknowledge John Whitters, Lee Renfrew and Denise Cole, from the Glasgow Dental Hospital and School, for their help with laser measurements performed on some of the phantoms.

I am also indebted to Olivia Lala and Jennifer MacDougall for proof-reading every single page of this document, suggesting many improvements to make it look professional. Finally, I dedicate this thesis to my parents, Sebastià and Catalina, and siblings, Pere and Xisca. I could feel your support and encouragement every hour of every day, even from one thousand miles away.

G. Reines March

Glasgow, May 2020

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Lung cancer is the third most common cancer in the UK, with approximately 130 new cases being diagnosed every day [1]. Treatments available for this type of cancer include: surgical resection of the tumour; radiotherapy, which consists of delivering ionising radiation to the cancerous site in order to kill the malignant cells; and chemotherapy, where a combination of drugs is delivered to the tumour. The choice of treatment is determined by a multidisciplinary team of experts, and depends on several factors, such as cancer type, size, location, stage and the patient's overall fitness [2].

In the majority of cases, treatment decisions are based upon the patient's clinical history and the visual information extracted from medical scans. However, non-invasive imaging modalities, such as Positron Emission Tomography (PET) or Computed Tomography (CT), currently provide insufficient information about the cellular make-up of the lesion microenvironment. Identification and classification of different tumour substructures, such as regions with hypoxia (poor oxygen perfusion), necrosis (dead tissue) or inflammation, is of great importance during therapy planning. This is because treatment can be tailored depending on the intratumoural characteristics.

In most cases, tumour biopsies are performed in order to gain a better insight of the tumour type and pathology. This procedure, however, only offers extremely localised and discrete information. Multimodality imaging techniques, such as

combined PET/CT, partially overcome this issue, by providing complementary knowledge about the cancerous region, like its metabolic fingerprint and anatomy.

In addition to these limitations, information extracted from these modalities can be ambiguous. For example, a high-absorption area in a CT scan could correspond to a solid tumour mass, but also to calcified tissue. Similarly, a PET image showing a zone with a prominent glucose uptake could be indicative of a highly metabolic carcinoma, or instead a region with acute inflammation. The only way to confirm the nature of the imaging phenomenon would be to compare it against the gold standard of histopathology.

For all the reasons exposed above, having an enhanced imaging protocol that provides a global, pathology-validated map of the cancer environment may influence clinicians to make better-informed decisions regarding diagnosis, prognosis and the choice of treatment for a given patient.

1.2 Project Scope

This retrospective study aims to spatially align PET/CT scans, obtained from lung cancer patients, with their corresponding histopathology data. As histopathological processing can only take place *ex vivo*, our cohort is limited to patients referred for radical lobectomy. In these cases, the entirety of the diseased tissue is available post-surgery, and a pathologist can carry out a detailed analysis of the whole tumour.

The project's main objective is to build an imaging framework to enable the registration of pre- and post-operative scans of lung cancer patients. Our comprehensive approach not only focuses on image processing algorithms, but also on data acquisition and clinical validation of the results. The study's workflow is detailed below:

1. **Patient screening and selection**: thoracic surgery lists are vetted to identify candidates for this study. Suitable patients are approached and, if they wish to participate, their written consent is recorded.

2. **Image data acquisition**: immediately following surgery, the resected lobe is fixed, processed and photographed by a consultant pathologist. In addition, patient's pre-operative PET/CT scans are retrieved from the imaging servers.

3. **Tumour 3D reconstruction**: virtual, three-dimensional models of the tumour are generated for all modalities. This is done by interpolating their cross-sectional binary masks.

4. **Image registration**: tumour volumes are rigidly registered. Moreover, microscopic histology data is matched to their corresponding pathology photographs.

5. **Validation**: the algorithm's performance is tested using a two-step validation method, which qualitatively evaluates its clinical significance against expert and non-expert manual registrations.

The method proposed establishes a direct spatial correspondence between non-invasive, *in vivo* medical imaging modalities and the pathology ground truth. Its potential applications are manifold. First, the mapping could be used to retrospectively validate the image-based diagnosis of a patient, and to ascertain whether the estimated pathological features of the tumour, inferred from the pre-operative imaging data, were accurate. Moreover, pathology-matched PET/CT datasets would make it possible for researchers to explore the relationship between radiomic image features and a particular pathology outcome [3]. Another possible application would be to support different pharmaceutical validation studies, by providing a registration framework that maps the emission profile of a radiotracer to its gold standard. For instance, one of the current research questions within NHS Greater Glasgow and Clyde is to ascertain whether hypoxia-specific PET tracers ^{18}F-FAZA and ^{18}F-FMISO reliably identify areas of pathology-validated hypoxia in lung cancer.

1.3 Novel Contributions

In addressing the scope of the project, the following major contributions have been identified:

- **Pathology processing protocol**: a novel tissue dissection protocol was developed to overcome the difficulties posed by the mechanical properties of lung specimens, and ensure repeatable, precise pathological processing across all patients. The main contribution in this area was the design and construction of a soft tissue slicing rig. This mechanical device allowed pathologists to fix and immobilise the specimen, adjust the slice thickness,

perform flat, parallel cuts with uniform thickness, and photograph the exposed cross-section, while maintaining the integrity and clinical value of the tissue.

- **Phantom study**: an exhaustive study of several factors that influence the reconstruction of a 3D shape from sparse 2D cross-sectional images was performed. Different choices of slice spacing, imaging modality, interpolation and extrapolation functions, offset, and area segmentation methods were tested, to evaluate their effect on overall volume recovery and shape reconstruction. A bespoke tumour-mimicking phantom was used in these experiments. To the best of our knowledge, this comprehensive analysis has never been reported in the literature before.

- **Tumour reconstruction**: results obtained from the phantom experiments provided an invaluable insight into the process of lung tumour reconstruction from pathology photographs and PET/CT scans. Any discrepancies found between different modalities, in terms of shape and volume, were studied. This allowed us to perform a detailed analysis of the different sources of error involved in tumour segmentation and reconstruction, and to establish acceptable tolerance margins for each modality.

- **Validation of results**: the lack of a ground truth against which to evaluate our method rendered many conventional image registration performance metrics unsuitable. This required us to develop a novel, two-step validation process. Expert clinicians were first asked to perform a manual registration of both datasets, to establish a gold standard result. They were then asked to blindly rate the output from the proposed method against several manual registrations, including their own.

1.4 Thesis Structure

This thesis is structured as follows: Chapter 2 contains a literature review of different medical image registration methods. The first section is devoted to analysing the three main building blocks of any image registration problem. The second focuses on the particular question of *in vivo* to pathology registration. Chapter 3 provides a theoretical background on all concepts, methods and algorithms discussed in this study. Its contents are split into three sections: anatomy and medicine; medical imaging; and image processing.

Prior to starting with human lung tissue processing, several simulations were run on various synthetic, custom-made, tissue-mimicking phantoms. These are discussed in Chapter 4. Different imaging protocols and shape reconstruction algorithms were tested on these shapes, and their performance was assessed against the model's ground truth. Chapter 5 gives a detailed outline of the study protocol in terms of patient recruitment, image acquisition and multimodality lung tumour registration. Results for each of the registration sub-tasks are presented here. The main issues identified during the realisation of the experiments, such as movement artefacts affecting the PET signal, or the need for external registration markers, are also discussed. In Chapter 6, the output of the proposed method is assessed against the current standard practice using a randomised test. Finally, Chapter 7 presents the thesis conclusions and gives some suggestions for future work.

1.5 Publications Arising from this Work

The work presented in this thesis has generated the following research outputs:

- G. Reines March, X. Ju and S. Marshall, "Correlation of pre-operative cancer imaging techniques with post-operative gross and microscopic pathology images," in *Proceedings of the Irish Machine Vision and Image Processing Conference 2017*, J. McDonald *et al.*, Eds. 2017, pp. 257-260.
- G. Reines March and S. Harrow, "Correlation of pre-operative cancer imaging techniques with post-operative gross and microscopic pathology images," presented at the 6$^{\text{th}}$ Annual Scientific Meeting of the Scottish Radiotherapy Research Forum, Stirling, UK, Nov. 2, 2017.
- G. Reines March, X. Ju, S. Marshall, S. Harrow and C. Dick, "Correlation of pre-operative cancer imaging techniques with post-operative macro and microscopic lung pathology images," presented at USCAP 2018 Annual Meeting, Vancouver, Canada, Mar. 17-23, 2018.
- G. Reines March, S. Harrow, C. Dick, X. Ju and S. Marshall, "Lung Tumor 3D Reconstruction from Planar CT Scans and Registration to Micro and Macroscopic Pathology Imaging," presented at the 40$^{\text{th}}$ IEEE EMBS International Conference, Honolulu, HI, USA, Jul. 17-21, 2018.

- S. Harrow, G. Reines March, C. Dick, X. Ju and S. Marshall, "Correlation of pre-operative cancer imaging techniques with post-operative macro and microscopic lung pathology images," presented at IASLC World Conference in Lung Cancer 2018, Toronto, Canada, Sep. 23-26, 2018.
- G. Reines March, S. Marshall, X. Ju, C. Dick and S. Harrow, "Registration of pre-operative lung cancer PET/CT scans with post-operative histopathology maps," presented at IASLC World Conference in Lung Cancer 2019, Barcelona, Spain, Sep. 7-10, 2019.

Chapter 2

Literature Review

This chapter provides the reader with a comprehensive survey of published works related to medical image registration. In the first section, the registration problem is broken down into three building blocks: transformation model, similarity measure and optimisation. A review of different methods available to approach each task is detailed. The second section focuses on studies about registration of pathology data with *in vivo* imaging modalities.

2.1 Medical Image Registration Algorithms

Medical image registration is a term that encompasses many different approaches and techniques depending on its specific application. For instance, monomodal bone image registration may be tackled using linear transformations (also known as rigid registration), because these structures are hard tissues and suffer little deformation. In contrast, soft organs such as the lungs can experience local and global deformations, caused by cardiac movement or the patient's breathing. In this case, a non-linear or deformable method is used, regularised by the mechanical properties of the organ. Registration methods therefore depend on several parameters and choices. Due to this high variability, it can prove difficult to classify all the algorithms using a single criterion. Maintz and Viergever proposed a detailed scheme for classifying medical image registration techniques, based on nine criteria [4]. These are: 1) dimensionality of the problem; 2) nature of the registration basis; 3) nature of transformation; 4) domain of transformation; 5)

interactivity; 6) optimisation strategy; 7) modalities involved; 8) subject; and 9) organ registered.

Regardless of the methodology used, the aim of any registration algorithm is to find a transformation function $\hat{W}$ that optimises the following functional [5]:

$$\hat{W} = \arg\max_{W} \mathcal{S}(F, M \circ W) - \mathcal{R}(W) \tag{2.1}$$

where F is the fixed image, which remains static throughout the registration process, serving as a target; M is the moving image, which is deformed in order to match F; W is the transformation applied to M; $\mathcal{S}$ measures the similarity between the moving and fixed images; and $\mathcal{R}$ is a regularisation term, which enforces some constraints on the transformation model.

Equation 2.1 illustrates that all registration algorithms are composed of three main blocks: a transformation model, a similarity measure and an optimisation method. The number of transformation parameters to be estimated differs greatly depending on the model chosen [5]: in the case of global rigid deformations, only six parameters need to be estimated, corresponding to the translation and rotation of the object. Local free-form dense transformations, however, may deal with hundreds of thousands of parameters, which in many cases outnumber the model constraints. Therefore, many registration problems are ill-posed, and the solutions need to be computed using iterative approaches. Moreover, many cost functions used are non-convex, and the optimisation algorithm can get trapped in a local minimum, leading to a suboptimal result. This is one of the reasons for which a pre-registration step, that is, performing a coarse alignment of the images and use it as an initial solution, can be crucial in determining the quality of the final registration [6].

2.1.1 Transformation Model

The transformation model defines the deformation nature of the images to be registered. The more parameters it has, the richer its description is; however, computational complexity scales rapidly with increasing number of degrees of freedom, eventually leading to intractability. Another important characteristic of the transformation model is its symmetry. In a diffeomorphic model, the transformation is invertible and differentiable, meaning that the topology of the

images is preserved, and the interchangeability of the moving and fixed images does not influence the final result [5].

The protocol proposed in this thesis includes two different registration scenarios: firstly, histopathology slides of the tumour are matched to their corresponding block-face photographs. Secondly, the reconstructed 3D pathology tumour model is aligned to its PET/CT counterpart. In the first case, the two images belong to different modalities, but represent the same reality: the histological composition of the tumour. A unique, one-to-one mapping can be found that maximises their level of alignment. Due to the deformations suffered during its processing, histopathological samples present non-linear distortions with respect to the block-face images. Therefore, a non-rigid transformation model is appropriate for this scenario.

However, in the second case, pathology and PET/CT contain slightly different information about the tumour. Hence, it is not correct to assume that a total overlap between both modalities, achieved by deforming one in order to match the other, is the best solution. Our assumption throughout the thesis is that pathology and pre-operative tumour models have very similar volumes and shape, but not necessarily the same. For this reason, a rigid registration is best suited to this purpose, with the main objective to maximise alignment between both shapes while preserving their structural integrity.

Rigid transformations only include rotations and/or translations. Because the Euclidean distance between any two points in the image is preserved, scale and shape are not affected. Conversely, non-rigid deformation models have a higher number of degrees of freedom, and are used to model local image distortions. These can be classified in two main groups [6]: transformations guided by physical models, and free-form deformations based on interpolation theory.

Transformations guided by physical models

Several approaches have used physical models to guide transformations. Elastic body models were first proposed by [7] to solve a problem of identifying planar objects in 3D computed tomography images. The cost function proposed confronts deformation of the moving image versus its similarity with the target image, so that an external force deforms the image in order to achieve maximum similarity whilst an internal force, which models the elastic properties of the solid, tries to oppose this deformation. The optimum solution is found when these

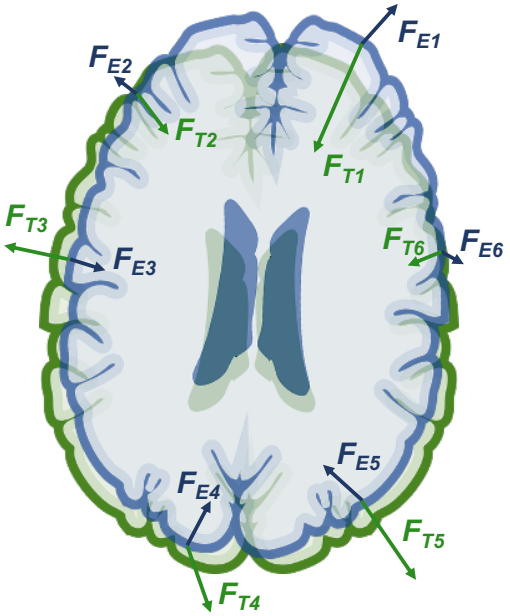


Figure 2.1: Diagram illustrating elastic body transformation. Moving image (in blue) is deformed by forces F_T to maximise similarity with template image (in green). These deformation fields are opposed by the elastic forces of the body, F_E

two forces reach equilibrium. This concept is illustrated in Figure 2.1. Many other authors have used this approach: Davatzikos [8] applied it to register brain cortical images, modelling them as inhomogeneous elastic bodies. Alexander *et al.* [9] proposed a multi-resolution elastic matching for aligning brain volumes, based on the work carried out by [10]. In [11], a combination of rigid and elastic transformations was used to register pre-treatment and intra-treatment prostate magnetic resonance (MRI) images.

Fluid-based models are another approach which are a relaxation of the elastic body models mentioned above, thus allowing higher degrees of deformation. This kind of transformation is ruled by the Navier-Stokes equations [5]. In Christensen *et al.* [12], a hierarchical approach was presented: firstly, an elastic body model was used to perform a coarse low-dimensional transformation, and the registration obtained was subsequently refined using a high-dimensional fluid model. D'Agostino *et al.* [13] also used a viscous fluid model deformed by the gradient of the mutual information criterion between multimodal MRI images.

Another kind of physical model transformation is based on the diffusion equation. Thirion [14], inspired by Maxwell's Demons paradox, created the so-called demons algorithm, which would later give foundation to many other diffusion-based al-

gorithms. This approach defines the object boundaries as semi-permeable membranes, containing perpendicular gradient vectors at each boundary point. The demons are randomly situated along the target object membrane, and they act by selectively pushing the inside points (i.e. the points enclosed inside the membrane) of the moving image into the model contour. It is an iterative approach, where at each repetition all the demon forces calculated are applied to deform the moving image [14]. Vercauteren *et al.* [15] proposed a faster variation of the demons algorithm which had diffeomorphic and symmetric properties. Cahill *et al.* [16] formulated a generalisation of the demons algorithm which accounts for locally adaptive regularisation. Gooya *et al.* [17] used Thirion's approach combined with an Expectation-Maximisation routine to register brain cancer patient scans with a glioblastoma atlas.

Additionally, biomechanical models of the organs can also be used as a source of prior information. Such methods are closely related to the physics-inspired models above, but are more realistic and consistent as they convey anatomical and physiological properties. Breast and prostate are two biomechanical models widely used in image registration tasks. The former is useful for regularising large deformations of the breast tissue, as in [18, 19]. Prostate models are used for MRI/ultrasound image registration in image-aided surgical interventions [20]. Tumour growth models have also been studied in [21, 22].

The majority of methods reviewed in this section are applied to images of the same modality, and the anatomical structures they represent have relatively low inter-patient variability, such as brain and prostate. In our case, due to the non-uniform distortions present in histopathology slides, a global physical model will not suffice to precisely define local tissue deformations. Moreover, due to their extremely thin cross-sections, histology slides do not share the same biomechanical properties as *bulk* lung tissue. Fortunately, the presence of common anatomical structures on both datasets, such as airways and blood vessels, can be used as landmarks to guide the registration process. For this reason, an interpolation-based deformation model is better suited to the task.

Interpolation-based deformations

In many free-form transformation models, the displacements of the moving image are only known in a discrete subset of points (i.e. the control points), meaning that the motion vectors need to be interpolated for the rest of the image, as shown

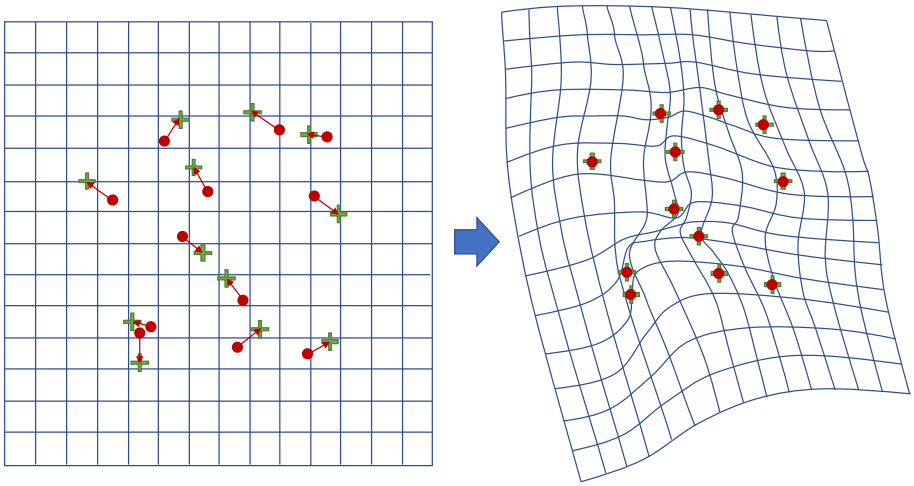


Figure 2.2: Diagram showing an example of interpolation-based deformation. Control points in the moving image (red circles) are matched to their corresponding features on the target image (green crosses). Displacement fields are estimated for the whole image by interpolating the known motion vectors

in Figure 2.2. One of the most widely used interpolators for this kind of transformation is the cubic B-spline. Rueckert *et al.* [23] used the free-form deformation model combined with cubic B-spline interpolation in order to describe the local motion of the breast for MRI registration. Rohlfing *et al.* [24] proposed a free-form transformation model based on B-splines with a volume-preservation constraint to register pre-contrast and post-contrast breast MRI images. Sdika [25] applied a symmetric non-rigid monomodal registration method based on the cubic B-spline. The advantages of this interpolator are its simplicity and efficiency to produce smooth deformations. However, the preservation of topology is not guaranteed, which may result in unrealistic transformations [5].

Another family of interpolators widely used in free-form image registration is thin-plate splines (TPS). Bookstein [26, 27] used this interpolator to construct anatomical atlases and study inter-patient variability. Meyer *et al.* [28] evaluated the performance of an automated multimodal image fusion algorithm using both linear and TPS registration approaches. Similar to cubic B-splines, TPS also produces smooth deformations, and it has a closed-form solution. However, the transformation is not diffeomorphic, and can produce non-uniform scaling and shearing [5, 29], both undesirable for our model. To overcome this problem, Schaefer *et al.* [29] proposed a method based on linear moving least squares. To limit the amount of local scaling and shearing, only a restricted set of transformation functions, including rotation, translation and uniform scaling, are permitted.

The resulting deformation is also smooth. Like thin-plate splines, this method also has a closed-form solution, which removes the need to use an optimisation algorithm. These properties make moving least squares an ideal candidate for our histology to block-face pathology registration task.

2.1.2 Similarity Measure

The similarity measure, also known as cost function or matching criterion, is used to quantitatively evaluate how aligned or *similar* the images to be registered are. Depending on which attributes of the image are used to evaluate their matching, it is possible to clearly differentiate two methods: intensity-based and feature-based.

Intensity-based methods

The family of intensity-based similarity measures rely on the intensity value of the voxels. This approach is usually more computationally demanding, as all image elements are taken into account for driving the registration process. The presence of noise can also undermine its performance. However, it is better suited for dense deformation patterns [5].

A relatively simple and straightforward intensity-based similarity measure is the sum of squared element-wise differences. This method is useful only in the case of monomodal image registration, where the images have been taken using the same imaging technique and the dynamic range of both are similar. This approach has been used in [30, 31] for brain MRI image registration and normalisation. In some cases, multimodal algorithms can be simplified to monomodal registration. This can be achieved by simulating one modality from the other, using prior knowledge about the physical features of the imaging technique. In [32], the authors translated a magnetic resonance image into ultrasound (US), using the MRI intensity and gradient information.

Perhaps the most commonly used methods for multimodal intensity-based registration are the so-called information theoretic approaches. These are based on the concept of mutual information (MI), defined as [6]:

$$I(X, Y) = H(X) + H(Y) - H(X, Y) \tag{2.2}$$

where $H(X) = -\sum_{i=1}^{N} p(x_i) \log p(x_i)$ represents the marginal Shannon's entropy of random variable X, and $H(X,Y) = -\sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j) \log p(x_i, y_j)$ is the expression for the joint Shannon's entropy. X and Y are random variables with possible outcomes $x_1, \ldots, x_N$ and $y_1, \ldots, y_M$, which represent the intensity values of the fixed and moving images, respectively. Mutual information measures the interdependence between them.

This approach was first used simultaneously by Collignon *et al.* [33] and Wells *et al.* [34]. Since then, it has received much attention from the research community. Pluim *et al.* [35] surveyed the field of image registration based on the mutual information similarity measure. Different entropy measures (Shannon, Jumarie, Rényi), pre-processing steps for noise and artefacts removal or interpolation strategies were reviewed, among other features. One of the drawbacks of these methods is that the entropy is measured single pairs of image elements, therefore neglecting any spatial information. In order to overcome this problem, Rueckert *et al.* [36] proposed an extension of the MI approach using second-order co-occurrence estimators. Studholme *et al.* [37] suggested the use of regional mutual information (RMI), in which each image region was seen as an independent source of information, contributing to the global relationship between the two images to be registered.

Although PET/CT and block-face pathology photographs are completely different modalities, our proposed 3D reconstruction method generates binary volumes that contain shape information only. After this conversion, both datasets share the same intensity domain. The rationale behind this simplification is that these two modalities show different aspects of the cancer. For this reason, there is not necessarily a value correspondence or statistical correlation between pixels representing the same spatial location.

The main advantage of handling binary volumes is that all element-wise differences can be computed using optimised, matrix-based binary logical operators. Computational time and memory requirements are drastically reduced, compared to equivalent greyscale problems. Speed becomes especially important when dealing with dense volumetric datasets. Moreover, in the binary case, sum of squared differences and mutual information metrics are interchangeable, as the joint probability term in Shannon's joint entropy formula is equivalent to the averaged sum of element-wise XNOR (exclusive negated disjunction) binary operations. Therefore, either can be used in our rigid registration task.

Feature-based methods

These methods are based on minimising the distance between sets of feature points or landmarks. This approach is robust for large deformations and is more efficient computationally, as only a subset of image elements is used (i.e. the feature points). However, one of the main challenges of this family of algorithms is to reliably extract the landmarks from the images. A naive solution would be to manually define the landmarks, based on the expertise of the clinical staff. Another approach is to place fiducial markers on the patient, such as metallic screws or skin marks. These constitute a set of salient and well-known feature points which are useful for guiding the registration process.

However, automatic landmark detection is often desirable. One method is based on the extraction of contours and spot-like regions. The multi-scale Laplacian of Gaussian (LoG) is one of the most used blob detectors in computer vision. The image is first smoothed with a Gaussian kernel, and then the Laplacian is applied. Lindeberg [38] proposed a feature detection algorithm based on LoG with automatic scale selection. In [39], a multiscale algorithm was used for extracting salient parts of an image based on the regional complexity of the image.

Scale-invariant feature transform (SIFT) is another method used for automatic extraction of keypoints, based on a multi-scale image analysis similar to Laplacian of Gaussian. First proposed by Lowe [40], this algorithm is composed of four main steps: 1) detection of local scale-space extreme points; 2) keypoint localisation, removing low contrast, candidates situated on edges and unstable points; 3) orientation assignment based on the local image gradient; and 4) labeling of each keypoint with a descriptor based on the local neighbourhood features.

SIFT provides a powerful method for identifying and matching features on images depicting different viewpoints of the same object or scene. This makes the algorithm suitable for tasks such as image stitching, object recognition or monomodal image fusion [40]. However, one of our registration problems concerns two dissimilar modalities (histology and block-face pathology), that contain different pixel values and statistical distributions. As a consequence, there is no guarantee that the keypoints identified by SIFT on both datasets will correspond to the same image features, or that these will be correctly matched. For this particular scenario, manual selection of landmarks is the most robust method. Our other registration task involves two binary 3D volumes, which by definition are featureless, meaning that a SIFT-based approach is also unsuitable.

Once the landmarks are extracted, the correspondences between them need to be calculated. One approach is to match those landmarks with similar descriptors. However, with this method the spatial location of the landmarks on the image plane is ignored. By imposing geometrical constraints, the correspondence problem becomes more robust. This is shown in [41], where the correspondences are modelled using graph theory. Another widely used algorithm for feature matching is the so-called iterative closest point (ICP), proposed in [42], where the correspondences between the moving and fixed images are recursively updated based on geometrical proximity. When descriptors are not available for each landmark, like in the proposed histology to block-face pathology registration problem, manual matching, based on prior knowledge about the scene, can be used.

2.1.3 Optimisation Algorithms

Image registration methods usually do not have a closed-form solution for inferring the deformation parameters, due to the non-convexity of the cost functions used. Therefore, iterative optimisation algorithms are required in order to find the best match between the two images. As already mentioned above, non-rigid registration problems are usually ill-posed, meaning that the final result will depend to a great extent on the optimisation algorithm used. According to Sotiras *et al.* [5], optimisation methods can be classified into two main families: continuous and discrete. Continuous methods have real-valued parameters and their cost function is continuously differentiable, whereas discrete optimisers deal with non-continuous functions and discrete variables. This review focuses on the former, due to their suitability for our particular problem.

One of the most popular continuous optimisation approaches is the gradient descent algorithm. It is based on iteratively updating the estimated parameters using the following rule:

$$\boldsymbol{x}_{t+1} = \boldsymbol{x}_t - \gamma \nabla(\mathcal{S}(\boldsymbol{x}_t)) \tag{2.3}$$

where $\boldsymbol{x}$ is the vector of deformation parameters at a given iteration, t is the iteration number, γ is the step size of the algorithm, and $\nabla(\mathcal{S}(\boldsymbol{x}_t))$ represents the derivative of the cost function. The estimated parameters move successively in the direction opposing the gradient, so that eventually a local minimum is reached, as depicted in Figure 2.3. Some variants of this method also reduce

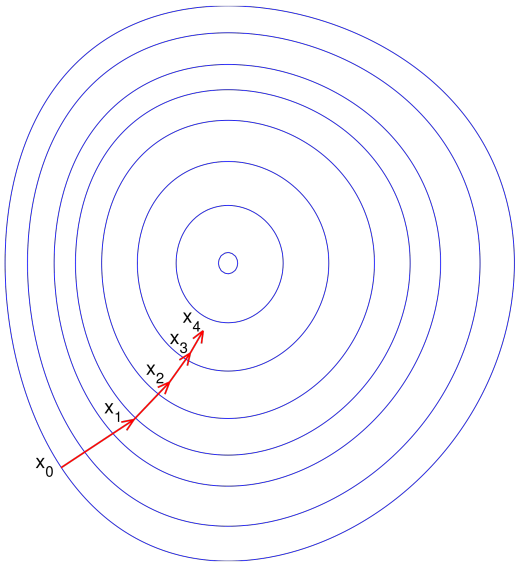


Figure 2.3: Illustration depicting the gradient descent method. Concentric lines are level sets representing the cost function $\mathcal{S}$, and the parameter vector $\boldsymbol{x}$ moves in the direction opposing the gradient of $\mathcal{S}$ until a local minimum is reached (source: `https://commons.wikimedia.org`)

the step size as the cost function approaches a local minimum, which leads to a more stable solution. Pennec *et al.* [43] showed that the demons registration algorithm [14] can be understood as a second-order gradient descent method. Two variants of this algorithm were used in [44], with different update strategies for the step size parameter.

This approach, however, can be computationally expensive, based on the fact that the parameters to be estimated have high dimensionality. This issue can be overcome by using stochastic gradient descent methods. This way, only a random subset of parameters is used to compute the minimisation direction. Although the estimations can be suboptimal, due to the fact that only a fraction of the parameters space is used, the algorithm eventually converges.

Another broadly used optimiser is the Gauss-Newton algorithm. It is useful for solving unconstrained least-squares problems, and does not require the calculation of second order derivatives, thus reducing its computational complexity. Ashburner and Friston [45] showed that the use of the Gauss-Newton optimisation method improved the efficiency of the registration algorithm in terms of memory requirements and convergence time for large deformation diffeomorphic maps. Salas-Gonzalez *et al.* [46] proposed a variant of the Gauss-Newton al-

gorithm which proved to converge faster than the standard Gauss-Newton or Levenberg-Marquardt algorithms.

As our rigid registration problem only involves binary image data and a very limited number of transformation parameters (rotation and translation), both optimisation algorithms are valid alternatives. In the non-linear registration case, moving least squares provides a closed-form solution, and therefore an optimiser is not needed.

2.2 Registration of Pathology Data with *in vivo* Imaging Modalities

Positron emission tomography (PET) combined with computed tomography (CT) is currently the main imaging tool for detection, treatment and assessment of most cancers. This type of non-invasive, combined scan provides the clinicians with a wealth of information regarding the anatomical layout of the tumour, as well as its metabolic fingerprint. However, the accuracy of these imaging modalities is low, due to their poor spatial resolution (especially in PET), the presence of artefacts generated during image acquisition, and the lack of a clinical ground truth. The only way of resolving these uncertainties is by validating PET/CT against tumour pathology [47].

Performing an accurate registration between pathology data and *in vivo* imaging modalities is a challenging task, due mainly to the following issues:

- **Resolution**: the in-plane resolution disparity between histopathology data and PET/CT scans can be of several orders of magnitude. The former, being a microscopy image, can achieve sub-micron pixel sizes, whereas the latter falls within the millimetre range. In contrast, interslice spacing is much greater in pathology imaging, due to the fact that the specimen needs to be physically sliced. Therefore, the mechanical properties of the tissue will limit the minimum thickness of the sections.
- **Deformation**: certain organs tend to collapse after surgical resection, especially those composed mainly of soft tissue, such as the lungs or the colon. This non-uniform deformation needs to be addressed in order to obtain an accurate result.

- **Modality**: pathology, PET and CT are different modalities, meaning that there may not be a direct correspondence between image features, landmarks and/or pixel values. This can potentially hinder registration accuracy.

To address this problem, some authors propose the use of an additional intermediate scan, namely a computed tomography (CT) or magnetic resonance (MRI) of the resected specimen [48–56]. This provides a way to bridge the gap between *in vivo* images and their corresponding histopatology: the intermediate dataset shares the same modality as the pre-operative scan, but shows the anatomy of the excised organ. Figure 2.4 contains a block diagram showing this relationship.

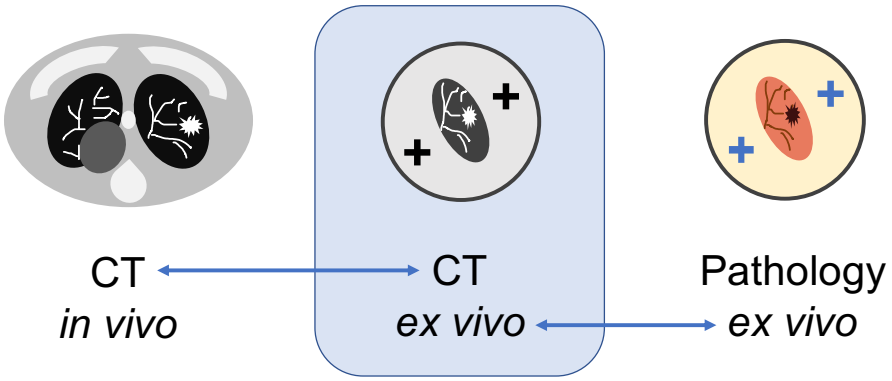


Figure 2.4: Block diagram showing the usage of an *ex vivo* scan as a bridging modality between the patient's *in vivo* tomography and the corresponding pathology of the specimen. In this example, the *ex vivo* CT scan shows the deformations that the organ suffered post-resection. The introduction of fiducial markers (+) visible in both *ex vivo* CT and pathology images facilitates their posterior registration

Validation of radiopharmaceutical studies is one application of this approach. Puri *et al.* [48] used this intermediate step to evaluate the accuracy of registration between PET and histopathology slices in laryngeal cancer patients. Six subjects underwent a PET/CT scan prior to radical laringectomy. Sea urchin spines were inserted into the specimen as fiducials, and after a 24-hour fixation period in formalin, it was CT scanned. The tissue was subsequently sliced and microtome sectioned. Through a series of rigid registrations, using the *ex vivo* CT as reference, the authors claimed an average registration error of 3 mm. Garcia-Parra *et al.* [49] performed PET to histology registration to evaluate the clinical usefulness of radiotracer ^{18}F-FAZA for identifying hypoxic regions in prostate cancer. A MRI scan of the prostatectomy specimen acts as bridging modality.

Another use of this technique is to compare tumour segmentation on *in vivo* imaging modalities to the ground truth of histopathology. Park *et al.* [50, 51] designed a system to validate tumour delineations made during radiotherapy planning in prostate cancer. Alignment between MRI/PET and histology sections was aided by an *ex vivo* MRI scan of the excised prostate. Their proposed registration method relies on maximising mutual information after deforming the images with thin plate splines, and consists of the following five steps:

1. Register each histology slide to its corresponding block-face photograph (macroscopic image of the whole specimen cross-section) using 6 control points situated on common landmarks, such as the prostate boundary.

2. Perform rigid registration between consecutive block-face images and stack them to create a volume. Pad the interslice space with 0's to preserve dimension ratio along z-axis.

3. Register reconstructed 3D volume from block-face photos with *ex vivo* MRI of the resected specimen, using 18 common landmarks as control points.

4. *ex vivo* and *in vivo* MRI scans of the prostate are registered. Deformations of the organ due to surgical removal are accounted for in this step.

5. Finally, by joining steps 1 to 4 above, histology is registered to its corresponding pre-operative MRI scan.

Zhan *et al.* [52] built a collection of different MRI appearances of prostate cancer, validated against their histology sections. Correlation between datasets were inferred from the prostate boundaries and salient structures visible on both modalities. Axente *et al.* [53] proposed an alternative method for testing the accuracy of their registration algorithm against the ground truth: they generated synthetic PET (sPET) volumes from autoradiography images of head and neck cancer specimens. sPET data was registered to its corresponding histopathological maps, and used to validate tumour segmentation for radiotherapy guidance. Meyer *et al.* [54] discussed different approaches for registering *in vivo* imaging with histopathology prostate data through the use of *ex vivo* specimen scans. In their review, they also listed some common false assumptions that oversimplify the deformation model and compromise registration accuracy.

Ohnishi *et al.* [55] proposed a two-step registration method to investigate the relationship between pathology findings and magnetic resonance signals in brain cancer specimens. Firstly, histology sections are aligned to the corresponding gross

photographs of the resected organ. This is performed using thin-plate splines, constrained by a set of common feature points. Next, each gross photograph is matched to its equivalent MRI slice using an affine transform. Using the parameters from both registrations, spatial correspondences can be drawn between histology maps and magnetic resonance images.

To overcome resolution limitations of clinical CT, Roth *et al.* [56] performed an *ex vivo* x-ray microtomography (μCT) of the resected lung tissue and aligned it to the pre-operative CT scan. Volumes were pre-registered by matching common landmarks, and then deformed using a multi-resolution approach. The authors claimed that their method could act as a bridging modality in the registration of clinical CT and histopatology images.

2.2.1 Registration of PET/CT and Pathology Lung Images

Lung tissue tends to collapse once extirpated from the patient, due to its mechanical properties. This makes the registration task even more difficult, as large non-linear deformations need to be accounted for. To mitigate this issue, in many cases the specimen is carefully inflated with different materials prior to pathological processing, so it resembles the *in vivo* shape and volume of the organ. A selection of studies addressing this particular registration problem are reviewed below.

Stroom *et al.* [57] validated the extent of the radiotherapy clinical target volume delineated on the planning PET/CT scan against histopathology. Five patients with non-small-cell lung cancer (NSCLC) due for lobectomy were recruited. The resected specimens were inflated with formalin to attain their pre-operative volumes, and subsequently sliced free-hand at 5 to 10 mm intervals. Orientation of the lobe was estimated using anatomical landmarks, such as the trachea and main bronchus. Registration was performed in two steps: CT axial planes were initially matched to the pathology cross-sections; and each PET/CT slice was deformed to match its corresponding block-face photographs. Similarly, van Loon *et al.* [58] investigated whether microscopic cancerous disease is included in the delineation of clinical target volumes for radiotherapy planning. 34 patients with NSCLC were recruited and PET/CT scanned prior to surgery. Resected lobes were inflated with 10% formaline through the main bronchus. Slicing was performed

free-hand along the longest axis of the lobe, with thicknesses ranging between 5 and 10 mm. CT scan slices were assumed co-planar with pathology cutting planes, and deformations were modeled by an ellipse. Gross tumour volumes were calculated for both datasets and compared.

Dahele *et al.* [59] developed a 3D tumour reconstruction algorithm to compare the information obtained from radiotherapy planning scans and histopathology. 12 cases of NSCLC were processed: excised tissue was inflated with agar, in an attempt to mimic original shape and volume, and dissected at 3 to 10 mm intervals. Each cross-section was photographed and sectioned. Tumour boundaries were traced on pathology and PET/CT images. Reconstructed tumour volumes were manually registered using a rigid transformation. Wu *et al.* [60] also aimed to compare radiotherapy target definition on PET/CT against pathology maps. 31 patients were recruited for this study. Lobectomy specimens were inflated with formalin and sliced every 3 to 5 mm. The tumour's largest diameter was measured on pathology block-face photographs, PET and CT datasets.

Yu *et al.* [61] attempted to empirically determine the cut-off standardised uptake value (SUV) in combined FDG-PET/CT that better matches the pathological tumour boundary. All 15 resected lobes were bisected along the longest axis immediately after surgery. Glass-mounted histopathology slides were obtained at 4 mm intervals, and an experienced pathologist delineated the tumour boundary in each of them. The pathological gross tumour volume obtained was compared with PET data, to determine the SUV that yields the better volume match. Meng *et al.* [62] also analysed whether cancerous microscopic extensions correlate with maximum SUV in FDG-PET images of non-small-cell lung cancer. 39 patients referred for lobectomy were recruited for this study. Specimens were sliced every 4 mm and stained using hematoxylin and eosin dyes. Extension of the tumour was delineated under the microscope by a senior pathologist. The resulting volume was correlated to the FDG-PET uptake data.

Wanet *et al.* [63] used a similar registration protocol to validate a gradient-based segmentation on FDG-PET against pathological findings. Ten NSCLC patients due for lobectomy were enrolled in the trial. Freshly excised lobes were inflated with 15% gelatin until uniformly filled, and left to solidify. These were subsequently placed inside a polystyrene box and fiducials were inserted (wooden sticks). The box was filled with gelatin to keep the tissue in place, frozen, and CT scanned. Afterwards, the container was sliced at 4 mm intervals and each

cross-section photographed. Contours were delineated on CT, FDG-PET and pathology datasets by experienced clinicians. All images were rigidly aligned manually until visual agreement was obtained. Schaefer *et al.* [64] also set out to assess the performance of a thresholding algorithm for PET tumour delineation compared to pathology. 15 cases were evaluated. Resected specimens were sliced every 4 to 5 mm. Each cross-section was photographed, and tumour contour was manually drawn. Pathology volumes were obtained by interpolating the tumour boundary along consecutive slices, and compared against PET and CT segmentation results.

Fu *et al.* [65] studied the correlation between lung tumour consolidation on CT imaging and pathology findings in 440 patients. Maximal tumour diameter was measured on two different CT window settings, namely lung and mediastinal. Lung window setting allows the user to observe areas of ground glass opacity in the lung, linked to inflammatory processes or alveolar thickening. Mediastinal window setting only displays solid sections of the carcinoma. Maximal tumour diameter was also measured on pathology specimens, and compared to the CT findings. A clear correlation was observed between the type of lung malignancy and its opacity. Park *et al.* [66] also compared tumour diameter in 135 resected lung specimens and their pre-operative computed tomography scans. Pathology samples were sliced along the lesion's largest cross-section and measured. CT volumes were re-sliced to display the greatest diameter of the tumour. Results showed that, in 27% of the cases, the cancer stage assigned at diagnosis, based on largest CT size, was overestimated compared to the post-operative pathology findings.

Comparison between published literature and proposed method

Registration of PET/CT and lung pathology images is a complex problem for many reasons: firstly, they belong to different imaging modalities, and consequently their pixel values have different physical meanings; secondly, datasets present heterogeneous resolutions, with pixel sizes spanning across three orders of magnitude; thirdly, the lungs are highly deformable organs and tend to collapse after surgical removal, meaning that the resected specimen hardly resembles its *in vivo* shape; and finally, the mechanical properties of the tissue make its precise dissection difficult.

To facilitate this process, it is often broken down into several sub-tasks, such as specimen dissection, lesion segmentation and shape alignment. The robustness of the overall method relies on the performance of each of these building blocks: for example, a very good image registration algorithm, based on the wrong assumption that specimen cross-sections are parallel, will probably produce sub-optimal results. Our approach aims to address and correct the shortcomings identified in the literature, in terms of pathological processing of the tissue, tumour reconstruction and image registration:

- **Pathology tissue processing**: the majority of the surveyed works dissected the specimen free-hand, which can easily lead to non-uniform, non-parallel slices. Wanet *et al.* [63] partially addressed the issue by freezing the specimen in a polystyrene box and subsequently cutting it with a saw. Dahele *et al.* [59] embedded it in agar gelatine and sliced with a rotary knife. Nevertheless, in neither case was the blade mechanically guided or constrained, and therefore it was still susceptible to error. In our method, the tissue is embedded in agar and sliced using a purpose-built soft tissue slicing rig (described in Section 4.2.3), which ensures parallel slices of uniform thickness.

  To facilitate the dissection process, some authors injected the specimen with a fixation agent, either formaldehide [57,58,60] or agar gelatine [59,63]. Our approach uses the latter to inflate the lung sample and achieve, approximately, the same volume and shape it had *in vivo*. Additionally, the stiffer texture achieved with the injection of agar facilitates the slicing of the tissue.

- **Pathology tumour model reconstruction**: building a 3D model of the tumour from the gross pathology photographs enables us to perform a volume-to-volume registration between pathology and PET/CT images. Out of ten studies analysed, only Wanet *et al.* [63] and Dahele *et al.* [59] followed this approach. The rest only used information contained within the planar tissue slice. Three studies [57, 58, 61] calculated the pathological gross tumour volume by multiplying the area of the lesion by the slice thickness for each cross-section, and adding them together. Schaefer *et al.* [64] performed linear interpolation between slices to calculate the overall pathlogical tumour volume. The remaining four works [60, 62, 65, 66] only carried out linear measurements on the images, to compare tumour extension in pathology and PET/CT.

- **Image registration**: the aim of this project is to align pre-operative PET/CT and post-operative histopathology images of lung tumours, enabling clinicians to draw spatial correspondences between both datasets. The proposed method uses a 3D reconstruction of the pathological lesion, similar to [59, 63]. Stroom *et al.* [57] performed a 2D-to-3D registration, where the CT volume was re-sliced to match the gross pathology cross-sections.

  Regarding registration optimisation, in [57,59,63], pathology and PET/CT volumes were rigidly aligned using manual rotations and translations, until visual agreement was reached. In [58, 61], PET/CT slices and pathology cutting planes were assumed co-planar. Our approach uses mathematical optimisation criteria to guide the registration problem, which results in a process that is semi-automatic (potentially reducing the need for supervision), consistent between executions, and free from human bias.

Table 2.1 summarises these findings. In conclusion, having identified and analysed the deficiencies found in similar studies, our protocol aims to deliver a robust and comprehensive solution to this multimodality image registration problem.

Author	Pathology tissue processing	Tumour reconstruction	Image registration
Stroom [57]	Specimen inflated with formalin, sliced free-hand (5-10 mm thick)	Pathology volume calculated by area projection (slice cross-section × thickness)	1) CT manually matched to pathology planes. 2) Manual warping of pathology images
van Loon [58]	Lobe inflated with formalin, sliced free-hand (5-10 mm) perpendicular to longest axis	Pathology volume calculated by area projection	No direct registration, only tumour equivalent diameters are compared
Dahele [59]	Formalin fixation, lobe embedded in agar and sliced free-hand (3-10 mm)	Stack cross-sectional areas, separated by equivalent slice thickness	Manual rotation of reconstructed tumour models
Wu [60]	Specimen inflated with formalin, sliced free-hand (3-5 mm)	No pathology tumour reconstruction, only maximal diameter is calculated	No registration, only maximal tumour diameters are compared
Yu [61]	Lobe inflated with formalin, sliced free-hand (4 mm)	Pathology volume calculated by area projection	No direct registration, only tumour volumes are compared
Meng [62]	Lobe inflated with formalin, sliced free-hand (4 mm)	No pathology tumour reconstruction, only microscopic disease extension (ME) is calculated	No registration, only correlation of maximum SUV and PET volume to pathology ME
Wanet [63]	Specimen inflated with agar, frozen inside PS box filled with agar, and sliced free-hand (4 mm)	Pathology volume calculated by area projection	Manual rotation and translation of segmented tumours
Schaefer [64]	Fresh specimen sliced free-hand (4-5 mm)	Linear interpolation of pathology cross-sections	No registration, only tumour volume and diameter are compared
Proposed method	Specimen inflated with agar, embedded and sliced using slicing rig (5 mm)	Phantom-validated spline interpolation method, to approximate tumour's true shape	Automatic rigid registration between PET/CT and pathology model, airways used as fiducials

Table 2.1: Comparison between several PET/CT to histology lung tumour registration protocols published in the literature and our proposed method

Chapter 3

Background Theory

The aim of this chapter is to provide a foundation on all methods, models and algorithms described in this thesis. The contents are divided into three main sections, according to their nature and discipline: anatomy and medicine; medical imaging; and image processing.

3.1 Anatomy and Medicine

3.1.1 The Human Lung

The lungs are a pair of stretchable organs belonging to the respiratory system. Their main function is to perform gas exchange between the body and the atmosphere, namely to expel carbon dioxide, produced by the body's cells, and to take in oxygen. The pulmonary artery transports de-oxygenated blood from the heart to the lungs, and it ramifies into small capillaries. Gas exchange occurs at the pulmonary alveoli. Oxygenated blood is then returned to the heart via the pulmonary vein. The lungs are connected to the outside via the trachea, allowing air to get in and out. This tubular, cartilaginous structure branches into the two primary bronchi, and once in the lungs it subdivides further into small bronchioles [67].

The lungs are surrounded by a mucous membrane called the pleura. Its function is two-fold: to lubricate the surfaces between the lungs and ribcage, to reduce friction during breathing; and also to enclose the organs, preventing spread of infection. The mediastinum is another membranous cavity, situated between the

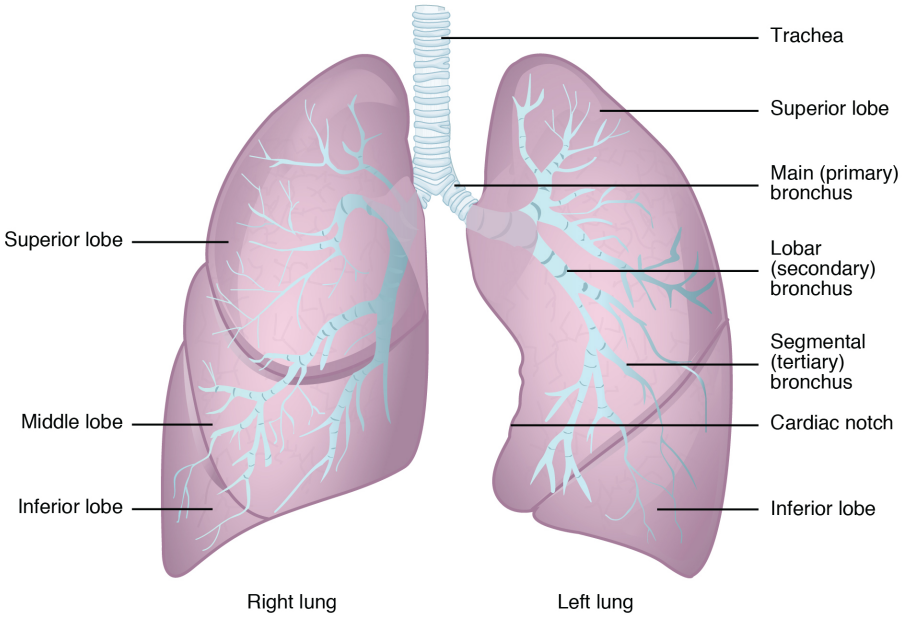


Figure 3.1: Diagram showing the gross anatomy of the lungs [67]

lungs, that contains the heart, its vessels and the lymph nodes of the chest. Each lung is divided into smaller units called lobes. The right lung has three lobes (superior, middle and inferior). The left lung is slightly smaller, as it surrounds the heart, and is composed of two lobes only (superior and inferior) [67]. Superior and inferior are also commonly called upper and lower, respectively. The abbreviated nomenclature used for referring to a particular lobe consists of three letters: R/L for right or left, U/M/L for upper, middle or lower, and L for lobe. For instance, RUL refers to Right Upper Lobe. Figure 3.1 shows a diagram with the gross anatomy of the lungs.

3.1.2 Lungs and Cancer

Cancer refers to a large family of health conditions that have a common trait: uncontrolled cell growth and reproduction. These rogue cells are capable of invading healthy tissue and organs, and build their own network of nutrient and oxygen supplies. Normal cells can become cancerous due to aberrations in their genetic material, acquired through gene inheritance or by spontaneous random DNA mutations. These can happen naturally, or due to exposure to certain factors, such as radiation, some chemicals, pollutants, cigarette smoke or substances present in food [1].

Lung cancer is one of the most common types of cancer in the UK, affecting nearly 47,000 people every year [2]. Current 1-year survival rates are at 32.1% for adults, with less than 1 in 20 surviving beyond 10 years following diagnosis [1]. If the disease originates in the lungs, it is called primary lung cancer. Depending on the type of cell affected, lung cancer can be divided into two main sub-types, the most prevalent being non-small-cell lung cancer (as opposed to small-cell). As cancer progresses, it starts spreading into surrounding tissue, eventually affecting other parts of the body, often situated far away from the origin (called cancer metastasis). Depending on the extent of the spread, lung cancer can be staged into four different categories [1]:

- **Stage 1**: the cancer is small and has not spread to the lymph nodes or distant organs. Generally concentrated in one single tumour no larger than 4 centimetres.
- **Stage 2**: if the tumour is smaller than 5 centimetres, it is classified as stage 2A. Stage 2B means that the tumour is between 5 and 7 centimetres or it has spread into other parts of the lung, lung lymph nodes, chest wall or mediastinum.
- **Stage 3**: it is divided into 3 sub-stages. 3A indicates that the tumour is up to 5 centimetres but it has invaded the lymph nodes in the centre of the chest, or that the cancer is between 5 and 7 centimetres and there is more than one tumour. It also means that the tumour has invaded nearby organs or structures, such as the heart or windpipe. In 3B, the tumour is large (5 to 7 centimetres) and has invaded multiple lymph nodes and/or proximal organs or structures. 3C refers to the cases where a large tumour has spread to the lymph nodes and also heart, nearby nerves or organs, or has progressed into several masses in different regions of the lung.
- **Stage 4**: there are tumours present on both lungs and/or membranes wrapping the lungs (pleura) or the heart, or it has spread to a distant organ (4A). If it has spread to several areas in one or multiple distant organs, then it is classified as 4B.

Choice of treatment for lung cancer depends on various factors, including the position of the tumour in the lung, its stage, and the patient's level of fitness [1]. Surgery is the most common treatment for cancers in stages 1 and 2. For more advanced cancers, chemotherapy, radiotherapy or a combination of both is often

recommended. Due to the nature of this study, the majority of patients recruited for our clinical trial fall within categories 1 to 2B.

3.1.3 Surgery as Lung Cancer Treatment

If doctors determine that cancer has not spread from the lungs to other parts of the body, surgery is generally the treatment of choice. The goal of this type of intervention is to remove the cancerous tissue completely. Depending on how much tissue is resected, the procedure can be classified into four classes [1]:

- **Wedge/segment**: if cancer is detected early and is concentrated in a small area, a wedge resection cuts only the affected area of the lung. A segmentectomy also removes the blood vessels and airways belonging to the area.
- **Lobe**: a lobectomy involves removing the lobe (or lobes) where the cancer is present. This is the most common type of surgery.
- **Lung**: if the cancerous tissue is spread over two or more lobes, the surgeon resects the whole affected lung, known as pneumonectomy.
- **Lymph nodes**: sometimes the surgeon decides to remove the nearby lymph nodes if there are suspicions that they also contain cancer cells. This is called lymphadenectomy. The procedure is normally combined with one of the above.

Open surgery is normally performed in lung cancer cases. The surgeon makes a cut or incision in the patient's chest and the affected part of the lung is removed. If the tissue to be resected is not very big, a keyhole surgery procedure called video-assisted thoracoscopic surgery (VATS) is chosen. The approach is less invasive, allowing quicker recovery and causing less pain to the patient. It is performed by making two or three small cuts in the subject's chest: one for a fibre optic camera, which allows the surgeon to see inside the body, and one or two for the surgical tools (see Figure 3.2). Once cut, the resected tissue is sucked out of the chest cavity through one of the small incisions.

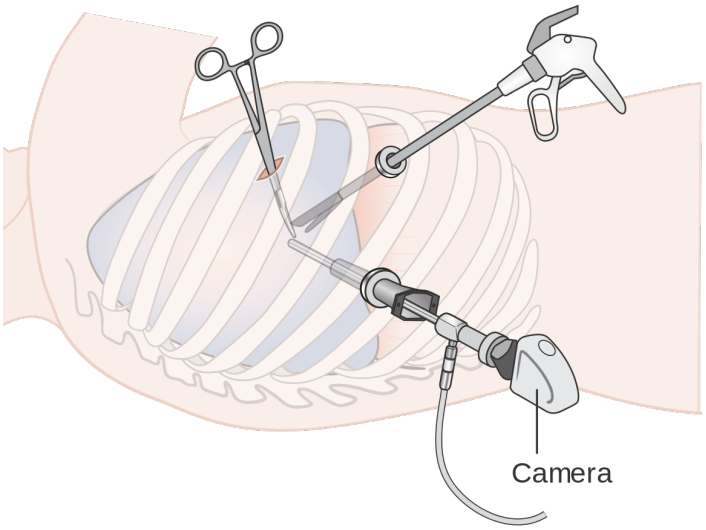


Figure 3.2: Depiction of keyhole lung surgery [1]

3.1.4 Pathological Processing of Cancer Specimens

Cancerous tissue resected by surgical intervention is sent to the pathology department for analysis. The specimen is immediately put into a chemical bath, composed mainly of formaldehyde, to prevent decomposition and remove any infectious agents. After 48 hours, the tissue is rinsed with water and prepared for dissection. A pathologist slices it several times along the tumour, for cross-sectional analysis of the cancer. Some healthy tissue is also collected to act as a baseline. Using a bench-mounted digital camera, they take photographs of the different sections, called block-face or gross pathology photographs.

Next, the slices of interest are cut into small square segments and placed in plastic cassettes. The tissue is fixed in paraffin wax, cut into extremely thin sections (approximately 10 microns thick) using a microtome, and placed on glass slides. These histology sections are then stained using hematoxylin and eosin. The dyes highlight sub-cellular structures in two different colours, namely purple for the cell nuclei and pink for the cytoplasm (see Figure 3.3). Consequently, with the help of a microscope, the pathologist is able to determine the shape, distribution and type of cells present in the tumour microenvironment. This analysis is crucial to determine the cancer stage and grade.

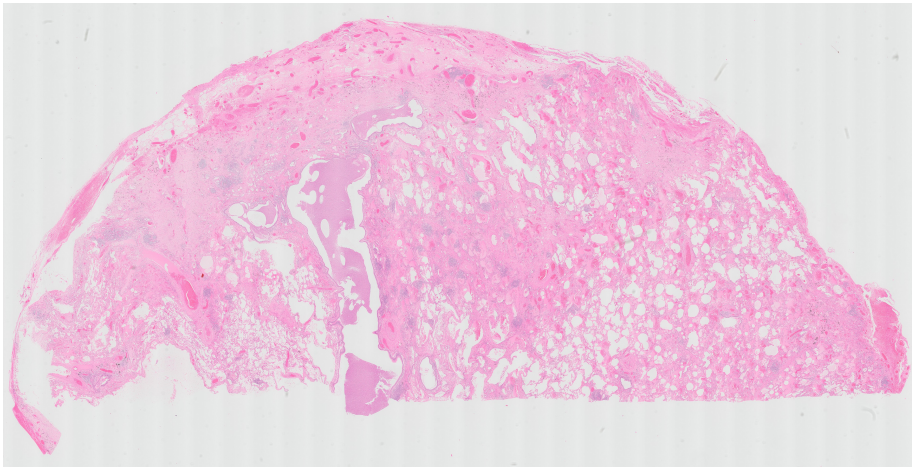


Figure 3.3: Histology section stained with hematoxylin and eosin

With the advent of digital pathology, laboratories are starting to scan and digitise all glass slides. This expedites diagnosis, enables faster collaborations and removes the need for physical retrieval of glass slides. It also paves the way for the introduction of digital image processing techniques into the workflow. Examples include automatic image analysis and machine learning applications, which would ultimately help pathologists deliver a faster and more precise diagnosis.

3.2 Medical Imaging

Medical imaging is a broad term that refers to the acquisition, compression, processing, visualisation and storage of images for clinical purposes. This section focuses on the technology and physics involved in their acquisition. Image processing is covered in Section 3.3.

A digital medical image is a discretised, numerical representation of some visual characteristics of a real-world object, usually a human body (or part of it). This information is distributed onto a n-dimensional grid, where n is 2 for planar images, 3 for volumes or 4 for spatio-temporal datasets. The smallest, undivisible units of this grid are called *pixels* (in the 2D case) or *voxels* (3D). The number of grid elements dictates the *resolution* of the image, and relates to the level of detail an image holds. For example, a standard High-Definition Video (HDV) image, which has 1920 columns by 1080 rows of pixels, will be able to unequivocally resolve finer image features compared to its older DVD counterpart (720×576

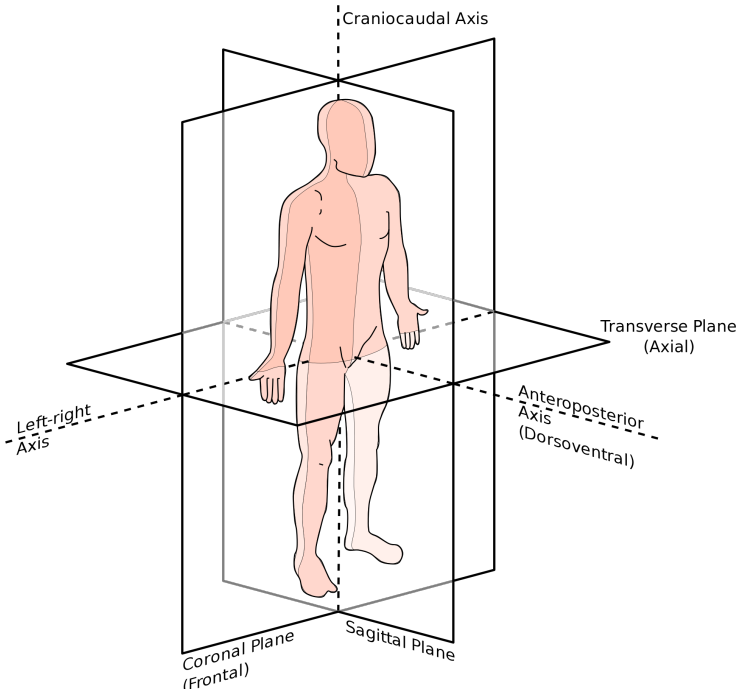


Figure 3.4: Anatomical axes and planes (source: `https://commons.wikimedia.org`)

pixels). Element size is normally expressed in real-world spatial units (e.g. millimetres) or in terms of density (e.g. pixels per inch, or PPI).

Volumetric medical images are normally acquired as a tomography, that is, stacks of parallel, 2D cross-sections of the body. These can be taken along any of the three anatomical axes: craniocaudal (head to toe); left-right; and anteroposterior (back to front). The anatomical planes perpendicular to said axes are called transverse or axial, sagittal, and coronal or frontal, respectively, and are illustrated in Figure 3.4. In most cases, the images acquired in tomographic modalities are axial cross-sections of the body, and run along the craniocaudal axis.

3.2.1 Computed Tomography

Computed Tomography (CT) is an imaging modality that uses multiple x-ray projections to produce cross-sectional images of the body [68]. These are acquired using a *CT scanner* machine. The device is formed of two main parts: the gantry and the sliding couch. The gantry is a ring-like structure with a central aperture of 50 to 70 centimetres in diameter, containing the x-ray source and detectors. The couch is where the patient lies. Its purpose is to move the subject through

Figure 3.5: Photography of a CT scanner showing the gantry with the circular aperture and the patient couch (source: `https://commons.wikimedia.org`)

the gantry aperture to obtain axial sections of their body. Figure 3.5 shows a picture of this type of scanner.

CT image reconstruction

As a convention, x and y axes lie on the transverse plane, and z is parallel to the craniocaudal axis. The imaging plane can be represented as a $m \times n$ matrix, where $n = m = 512$ in a standard setting. Every element of this matrix represents a linear attenuation coefficient $\mu(x, y)$. The x-ray source rotates 360 degrees inside the gantry, obtaining a set of projections of the object from different angles, as shown in Figure 3.6. Diametrically opposed to the source lies the detector array. Therefore, for a fixed acquisition angle θ, the relationship between the incident ($I_{\theta,i}$) and transmitted ($T_{\theta,i}$) x-ray intensities, for each array element i, is given by the Beer-Lambert law of absorption [68]:

$$P_{\theta,i} = \ln\left(\frac{I_{\theta,i}}{T_{\theta,i}}\right) \tag{3.1}$$

for $i = 1, \ldots, n$, where $P_{\theta,i}$ is the projection of all attenuation coefficients along the line i and angle θ, that is:

$$P_{\theta,i} = \mu(1, i) + \mu(2, i) + \cdots + \mu(m, i) \tag{3.2}$$

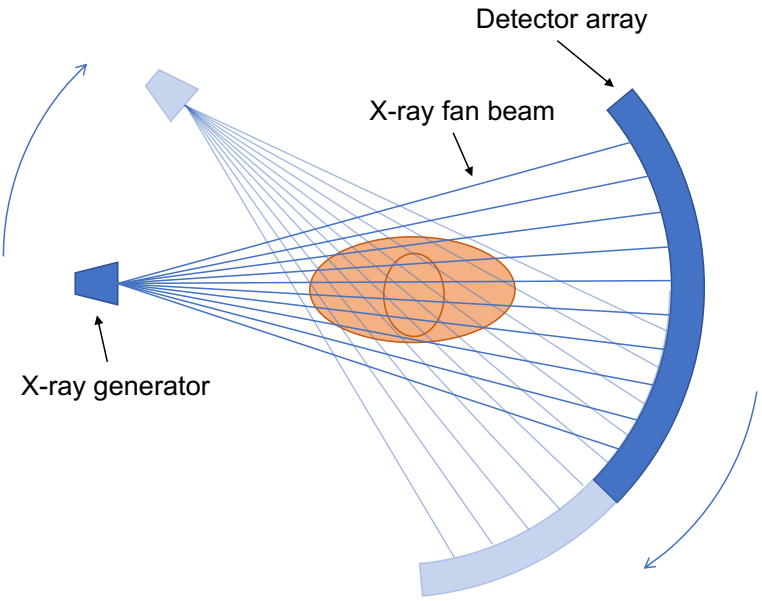


Figure 3.6: By rotating the x-ray source and detector array around the patient, different projections of the body can be obtained (adapted from [68])

A sufficiently large collection of 1D projections P_θ can be used to reconstruct a 2D sectional image of the object of interest. Many of the algorithms used for this purpose are based on either Filtered Back Projection or iterative methods [68]. The first involves projecting the 1D profiles back through the image, so that areas of constructive interference correspond to a common element in the image. A high-pass filter is applied to the data before the back-projection, to suppress flat areas generated by the projective nature of the acquisition method. The reconstructed image is thus sharpened.

Iterative methods make use of a set of mathematical functions to obtain an approximate solution to a given problem. The interim result is fed back recursively to the algorithm until an optimal solution is found. These methods generally start with an educated guess of the final solution, based on some knowledge extracted from the data. After an iteration, the difference between the model prediction and the actual data is calculated. The next cycle takes into account this error to adjust the model and predict new data, and so on.

Once the CT image has been reconstructed using either method, the attenuation coefficients calculated for each pixel $\mu(x, y)$ are normalised with respect to the attenuation coefficient of water, μ_{water}, using the following formula [68]:

$$H(x, y) = 1000 \cdot \frac{\mu(x, y) - \mu_{water}}{\mu_{water}} \tag{3.3}$$

Tissue	Absorption (HU)
Air	-1000
Lung	-200 to -500
Fat	-50 to -200
Water	0
Muscle	25 to 40
Bone	200 to 1000

Table 3.1: Typical CT absorption coefficients for different biological tissues expressed in the Hounsfield scale (adapted from [68])

where $H(x, y)$ is the attenuation in Hounsfield units (HU). Some typical attenuation values for different biological tissues are shown in Table 3.1. Due to the high sensitivity of the detector array, contrast differences of 0.2% can be discerned. This translates into approximately 500 different grey levels. However, the human eye only distinguishes up to about 30 shadows of grey, meaning that subtle changes in HU will not be appreciated. To overcome this issue, CT scans are normally examined on interactive displays, where the user can adjust the visualisation window in terms of width and level [68]. *Window level* refers to the mean CT attenuation level where the window is centred, proportional to the brightness of the image. *Window width* relates to the range of levels displayed, and therefore affects the contrast.

Image enhancement with contrast agents

Depending on the type of exam, radiocontrast agents can be used to highlight specific structures or organs. These are substances with different, known densities, which act by either offering more attenuation (positive contrast) or less attenuation (negative contrast) than surrounding regions. Examples of positive radiocontrast agents include iodine (for angiograms) and barium (for digestive tract studies). Carbon dioxide gas is often used as a negative agent for gastrointestinal examinations [68].

3.2.2 Positron Emission Tomography

Positron Emission Tomography (PET) is a functional imaging modality that makes use of radioactive agents to monitor metabolic processes of the body [68]. When these radionuclides interact with matter, they emit a pair of high-energy

photons, which are detected by a *PET scanner* machine. While the device looks very similar to a CT scanner, its working principle is completely different.

PET imaging principles

A range of different radionuclides is used for clinical imaging, each one suited to highlight a particular condition. The choice of radioactive molecule depends on its decay mode, positron energy, and most importantly, its half-life. A long half-life will expose the patient to radioactivity for a longer period, whereas a short-lived radionuclide may deplete even before the scan takes place. One of the most routinely used agents is a fluorine-labelled glucose molecule called fluorodeoxiglucose (^{18}F-FDG). It is utilised to detect areas of high metabolical activity, such as in most tumours. FDG has a half-life of just under two hours, therefore the molecule needs to be synthetised and injected into the patient within this time frame. For this to be logistically viable, most radionuclides are produced on-site, using a cyclotron.

Positron decay is a process in nuclear physics where an unstable atomic nucleus restores its charge balance by transforming a proton into a neutron. As a byproduct of this reaction, it also releases a positron (positive electron, e^+). The positron travels away from the nucleus until it interacts with a surrounding electron (e^-). When they annihilate, two identical γ photons are produced and emitted in opposite directions [68]. Two opposing detectors in the scanner can capture these photons simultaneously. If the two events are detected within a given time window τ (typically set between 3 and 9 nanoseconds) and both incident photons have similar energies, then a coincident event is recorded. The imaginary line joining the two detectors involved, along which the annihilation event happened, is known as line of response (LOR). Figure 3.7 illustrates this concept.

A fraction of the coincident events captured during a scan do not match the ideal scenario described above. In some cases, photons originating from the same annihilation event change their direction of travel due to Compton scattering. Although detected within the coincidence timing and energy windows, the associated LOR does not represent the actual event location. Similarly, two scattered photons, belonging to different annihilation processes, may reach the detector within coincidence windows. The system will regard them as belonging to the same decay process and therefore generate a false positive. Scattering events reduce image contrast and lower the accuracy of quantitative measurements [68].

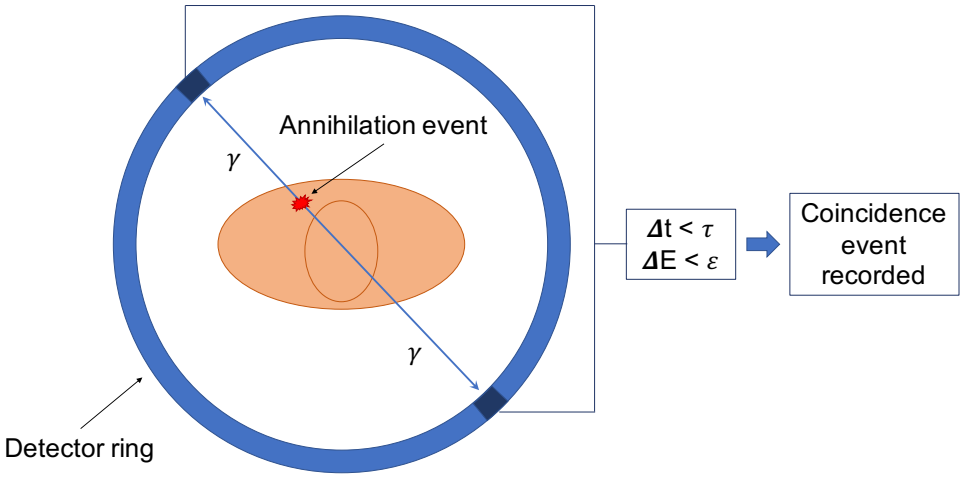


Figure 3.7: The two γ photons generated in the annihilation process are detected by opposing detectors. If detection falls within the set time and energy windows, it is recorded as a coincident event

After the acquisition phase, all projections from parallel LORs are arranged in a sinogram. Filtered Back Projection and iterative methods are the most common algorithms used to reconstruct the PET image, as described in Section 3.2.1.

PET/CT multimodality imaging

Sometimes, different image modalities are combined to obtain more information about a condition or organ of interest. These can be obtained concurrently on the same physical machine, or recorded in separate scanners or at separate times and fused posteriorly. PET and CT images are usually acquired on the same device, and therefore data is inherently aligned, both in space and time. This removes the need to use complex registration methods. However, alignment errors can still occur due to physiological motion, such as breathing. This is especially problematic for PET data, which is acquired in intervals of 2 to 5 minutes of duration [68]. To account for this movement, time-gated modalities have recently been introduced. Also known as 4D imaging protocols, they have the capacity to track physical motion during the scanning procedure. The acquired data is divided into a set number of bins, depending on time of capture in the breathing cycle. The technique reduces movement-induced blurring and allows for a more accurate analysis.

The main advantage of a combined PET/CT scan is that it presents the clinician with simultaneous anatomic and functional information, which is important for correct image interpretation. Moreover, CT data is often used to perform attenu-

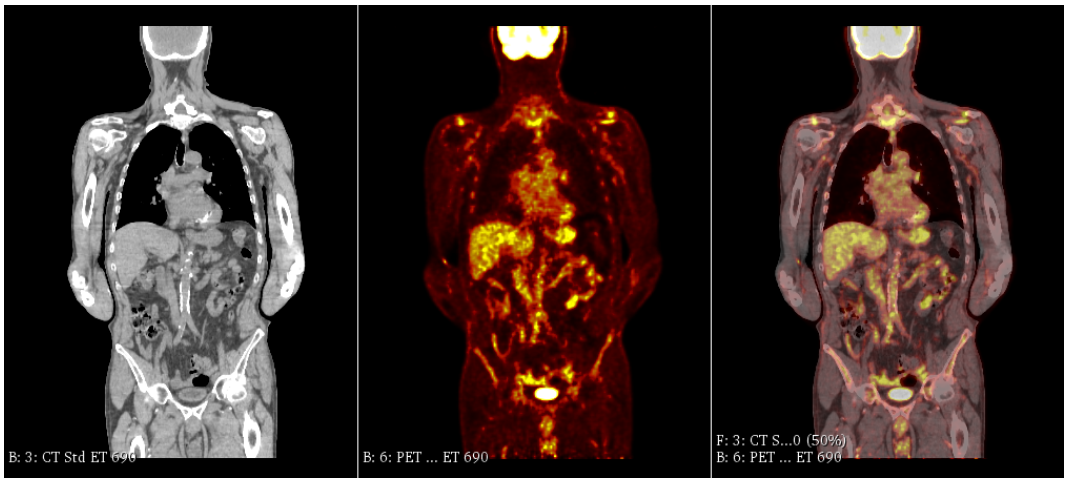


Figure 3.8: CT (left), PET (centre) and fused (right) coronal slices of a patient

ation correction on the PET image. Figure 3.8 shows CT, PET and fused coronal scans of a patient.

3.3 Image Processing

The term *digital image processing* refers to a set of computer routines and algorithms, aimed at enhancing certain features of an image, for the purpose of facilitating its interpretation by a human or a computer. Fuelled by the current availability of cheap and reliable computing power, digital image processing techniques have become ubiquitous. A non-exhaustive list of examples include image rendering and compression, texture analysis, feature extraction, segmentation and pattern recognition. Some of the methods used in this project are described next.

3.3.1 Interpolation

Interpolation is a mathematical method for inferring intermediate values of a function from a set of neighbouring, known data points. The well-known Nyquist-Shannon theorem states that a band-limited, continuous signal can be unequivocally recovered from its discrete version as long as the sampling rate is at least twice the highest frequency of that signal [69]. Some applications of interpolation include: discrete-to-continuous signal conversion; restoration of noise-corrupted samples; up-sampling of data for enhanced graphical visualisation; and recon-

struction of sparsely sampled signals [69]. In this work, different interpolation methods are investigated for the reconstruction of 3D volumes from sparse 2D cross-sections.

Depending on the choice of interpolant function, the resulting signal will have different properties in terms of computational cost, memory, continuity and smoothness. Similarly, the accuracy of the reconstruction depends on the entropy and non-stationary characteristics of the original signal, as well as the length of consecutive missing data samples [69]. Different interpolation methods used in this thesis are described below:

Nearest-neighbour interpolation

Nearest-neighbour is one of the simplest methods of interpolation. A given arbitrary point is assigned the value of its closest known neighbour, ignoring any other known values in the vicinity. The result is a discontinuous, piecewise constant function, as seen in Figure 3.9. This method has very low computational power and memory requirements.

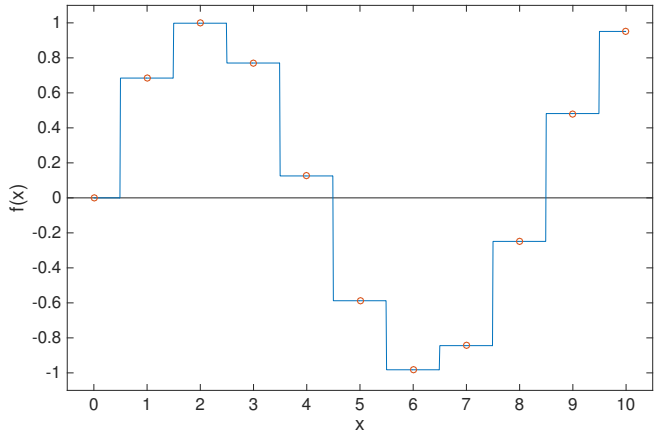


Figure 3.9: Discrete sinusoidal signal (circular markers) reconstructed using a nearest-neighbour interpolator

Polynomial interpolation

Interpolation can also be performed by fitting a polynomial function through the known data points. Given $N+1$ samples, the polynomial of order N that passes

through all of them is unique, and will have the form:

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_N x^N \tag{3.4}$$

where $a_0, \ldots, a_N$ are the polynomial coefficients that can be found by solving the following $N + 1$ system of equations:

$$\begin{array}{rcl} f(x_0) & = & a_0 + a_1 x_0 + a_2 x_0^2 + \cdots + a_N x_0^N \\ f(x_1) & = & a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_N x_1^N \\ \vdots & & \vdots \\ f(x_N) & = & a_0 + a_1 x_N + a_2 x_N^2 + \cdots + a_N x_N^N \end{array} \tag{3.5}$$

where $f(x_0), \ldots, f(x_N)$ are the known values of the function at points $x_0, \ldots, x_N$. For large signals, the system increases in complexity and becomes ill-conditioned, leading to unstable results [69]. Moreover, high-order polynomials normally present oscillatory responses, which may produce unrealistic results, as most natural signals may fit a smoother function.

A way of simplifying this problem is by using *piecewise polynomials* [70]. Let $[a, b]$, where $a = x_0$ and $b = x_N$, be an interval subdivided into $[x_i, x_{i+1}]$ subintervals, for $i = 0, \ldots, N - 1$. A piecewise polynomial is a function $p(x)$ defined in the interval $[a, b]$ as:

$$p(x) = \begin{cases} p_0(x) & \text{if } x_0 \leq x \leq x_1 \\ p_1(x) & \text{if } x_1 \leq x \leq x_2 \\ \vdots & \vdots \\ p_{N-1}(x) & \text{if } x_{N-1} \leq x \leq x_N \end{cases} \tag{3.6}$$

where each function $p_i(x)$ is a polynomial of at most M-th order and defined in the interval $[x_i, x_{i+1}]$, where $M \leq N$.

Depending on the order M and any other constraints imposed on $p_i(x)$, different outcomes are obtained. For instance, $M = 1$ results in the well-known **linear interpolator**. Let us define $\hat{x}$ as a point located between known values x_0 and x_1, that is, $x_0 \leq \hat{x} \leq x_1$. The corresponding linear interpolator will have the form

$$f(\hat{x}) = a_0 + a_1 \hat{x} \tag{3.7}$$

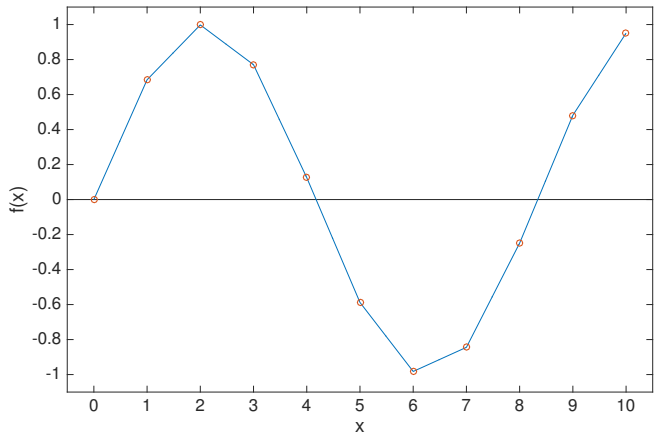


Figure 3.10: Discrete sinusoidal signal (circular markers) reconstructed using a linear interpolator

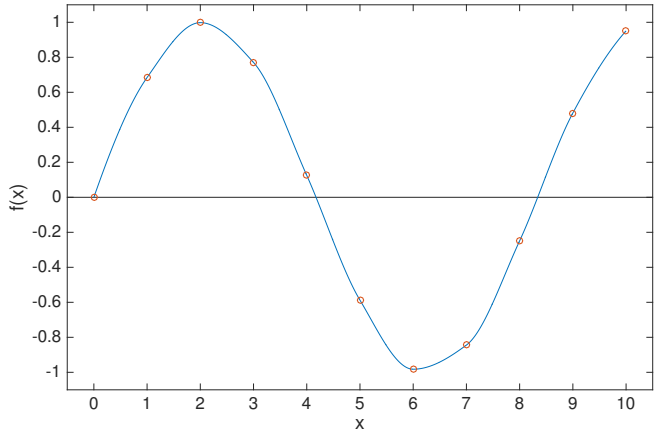


Figure 3.11: Discrete sinusoidal signal (circular markers) reconstructed using a cubic Hermite interpolator

As $f(x_0) = y_0$ and $f(x_1) = y_1$ are known points, coefficients a_0 and a_1 can be found by solving the following linear system:

$$\begin{aligned} f(x_0) &= a_0 + a_1 x_0 \\ f(x_1) &= a_0 + a_1 x_1 \end{aligned} \tag{3.8}$$

which yields $a_0 = y_0 - \dfrac{y_0 - y_1}{x_0 - x_1} x_0$ and $a_1 = \dfrac{y_0 - y_1}{x_0 - x_1}$. This corresponds to the equation of a line passing through points (x_0, y_0) and (x_1, y_1). Solving (3.6) for each interval results in a continuous, piecewise linear function, as shown in Figure 3.10.

Higher-order polynomials, along with a set of continuity constraints on their derivatives, can be used to obtain smoother interpolation functions. For example, cubic Hermite interpolators are a family of third order polynomials ($M = 3$) with specified first derivative values at their known points. They belong to differentiability class C^1, meaning that their first derivative is continuous [69]. An example is shown in Figure 3.11.

Another higher-order polynomial method is **spline interpolation**. Splines are piecewise, M-th order polynomials that have $M - 1$ continuous derivatives (class C^{M-1}). Cubic spline interpolation is one of the most common, where a third-order function is fitted between two known samples (see Figure 3.12). A cubic polynomial has four coefficients (a_0, a_1, a_2, a_3). Four constraints are therefore needed to ensure uniqueness of the solution. Two conditions are set by the known endpoints of the interval (also called nodes). The other two are imposed by ensuring that the first and second derivatives of the function are continuous at these points. The general expression for a piecewise cubic spline has the form:

$$s(x) = s_i(x) = a_{0,i} + a_{1,i}(x - x_i) + a_{2,i}(x - x_i)^2 + a_{3,i}(x - x_i)^3 \tag{3.9}$$

for $i = 0, 1, \ldots, N - 1$ and $x_i \leq x \leq x_{i+1}$. That is, a different cubic polynomial is fitted at each subinterval $[x_i, x_{i+1}]$. With this expression and the additional constraints, a system of equations can be built to calculate the coefficient values [69], as explained below.

Firstly, one of the conditions is that $s(x)$ has to fit the known data points. Therefore, if $s_i(x)$ is evaluated at $x = x_i$ (i.e. the beginning of the subinterval) we obtain:

$$s_i(x_i) = a_{0,i} = y_i \tag{3.10}$$

where $y_i = f(x_i)$ is the value of the function at x_i. To simplify the notation, let us now define $X_i = x_{i+1} - x_i$, as the spacing between known data points. Due to the continuity constraint, the function $s_i(x)$ at the end of the subinterval ($x = x_{i+1}$) needs to be equal to $s_{i+1}(x_{i+1})$. Therefore:

$$s_i(x_{i+1}) = a_{0,i} + a_{1,i}X_i + a_{2,i}X_i^2 + a_{3,i}X_i^3 = a_{0,i+1} = s_{i+1}(x_{i+1}) \tag{3.11}$$

Another condition is that the second derivative of the function needs to be continuous, meaning that, at the end of the subinterval $[x_i, x_{i+1}]$, $s_i''(x_{i+1}) = s_{i+1}''(x_{i+1})$.

Therefore:

$$s_i''(x_{i+1}) = 2a_{2,i} + 6a_{3,i}X_i = 2a_{2,i+1} = s_{i+1}''(x_{i+1}) \tag{3.12}$$

From Equation 3.12 it can be seen that at the beginning of the subinterval the following relationship will also hold:

$$s_i''(x_i) = 2a_{2,i} \implies a_{2,i} = \frac{s_i''(x_i)}{2} \tag{3.13}$$

Substituting $a_{2,i}$ into Equation 3.12, $a_{3,i}$ is obtained:

$$a_{3,i} = \frac{s_i''(x_{i+1}) - s_i''(x_i)}{6X_i} \tag{3.14}$$

Finally, $a_{1,i}$ can be calculated by inserting $a_{0,i}$, $a_{0,i+1}$, $a_{2,i}$ and $a_{3,i}$ into Equation 3.11:

$$a_{1,i} = \frac{y_{i+1} - y_i}{X_i} - \frac{s_i''(x_{i+1}) + 2s_i''(x_i)}{6}X_i \tag{3.15}$$

If all coefficients are substituted into the general spline expression (Equation 3.9):

$$\begin{aligned} s_i(x) = y_i + \left(\frac{y_{i+1} - y_i}{X_i} - \frac{s_i''(x_{i+1}) + 2s_i''(x_i)}{6}X_i\right)(x - x_i) + \frac{s_i''(x_i)}{2}(x - x_i)^2 \\ + \frac{s_i''(x_{i+1}) - s_i''(x_i)}{6X_i}(x - x_i)^3 \end{aligned} \tag{3.16}$$

it can be seen that they depend on second derivatives $s_i''(x_{i+1})$ and $s_i''(x_i)$. These can be calculated by applying the condition of continuity to the first derivatives, that is, $s_i'(x)$ must equal $s_{i-1}'(x)$ when evaluated at $x = x_i$:

$$s_i'(x)\Big|_{x=x_i} = -\frac{X_i}{6}(s_i''(x_{i+1}) + 2s_i''(x_i)) + \frac{1}{X_i}(y_{i+1} - y_i) \tag{3.17}$$

$$s_{i-1}'(x)\Big|_{x=x_i} = \frac{X_{i-1}}{6}(2s_{i-1}''(x_i) + s_{i-1}''(x_{i-1})) + \frac{1}{X_{i-1}}(y_i - y_{i-1}) \tag{3.18}$$

Taking into account that $s_i''(x_i) = s_{i-1}''(x_i)$, and that $2a_{2,i} = s_i''(x_i)$ (Equation 3.13), equating the right-hand sides of both Equation 3.17 and 3.18 gives:

$$\begin{aligned} X_{i-1}a_{2,i-1} + 2(X_{i-1} + X_i)a_{2,i} + X_i a_{2,i+1} \\ = 3\left(\frac{1}{X_{i-1}}y_{i-1} - \left(\frac{1}{X_{i-1}} + \frac{1}{X_i}\right)y_i + \frac{1}{X_i}y_{i+1}\right) \end{aligned} \tag{3.19}$$

for $i = 1, 2, \ldots, N-1$. That results in a system of $N-1$ linear equations with $N+1$ unknowns, $a_{2,0}, \ldots, a_{2,N}$. To obtain a unique solution, the behaviour of

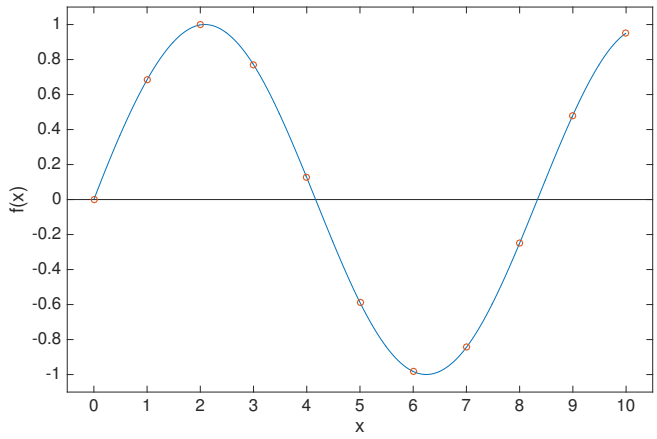


Figure 3.12: Discrete sinusoidal signal (circular markers) reconstructed using a cubic spline interpolator

the derivatives at the extreme points, x_0 and x_N, must be defined. These can be set to zero (free boundary conditions), or calculated over the mirrored signal (clamped boundary conditions) [69,70]. Assuming the former, the resulting linear system $A\mathbf{x} = \mathbf{b}$ is defined by:

$$A = \begin{pmatrix} 1 & 0 & 0 & \dots & \dots & 0 \\ X_0 & 2(X_0 + X_1) & X_1 & \ddots & & \vdots \\ 0 & X_1 & 2(X_1 + X_2) & X_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & X_{N-2} & 2(X_{N-2} + X_{N-1}) & X_{N-1} \\ 0 & \dots & \dots & 0 & 0 & 1 \end{pmatrix},$$

$$\mathbf{x} = \begin{pmatrix} a_{2,0} \\ a_{2,1} \\ \vdots \\ a_{2,N-1} \\ a_{2,N} \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} 0 \\ \frac{3}{X_1}(y_2 - y_1) - \frac{3}{X_0}(y_1 - y_0) \\ \vdots \\ \frac{3}{X_{N-1}}(y_N - y_{N-1}) - \frac{3}{X_{N-2}}(y_{N-1} - y_{N-2}) \\ 0 \end{pmatrix} \tag{3.20}$$

Once all a_2 coefficients have been computed by solving the above system, a_0, a_1 and a_3 can be obtained from Equations 3.10, 3.15 and 3.14, respectively. Finally, the spline interpolation function is calculated by substituting these sets of coef-

ficients into Equation 3.9. Matlab's `interpn()` standard function* was used to implement the interpolation methods described in this section.

Shape-based interpolation

Tomographic medical imaging modalities allow the user to reconstruct 3D models of organs or other anatomical structures of interest, by acquiring a series of 2D, parallel slices. These datasets are usually anisotropic, that is, the slice separation is higher than the in-plane pixel size. In order to obtain an accurate representation of such data, it is necessary to interpolate the missing information between known planes. One straightforward method is to interpolate the grey values between slices, followed by segmentation of the object of interest. However, according to Herman *et al.* [71], this approach poses two main problems: a) it puts more burden on the user performing segmentation, as the number of slices increases, and b) segmentation by thresholding is likely to produce artefacts due to sudden changes in boundary location.

An alternative approach is to perform shape-based interpolation. This is based on the propagation of the organ's boundary rather than its actual pixel values. The method works as follows [71]:

1. For each slice of size $N \times M$, the organ of interest is segmented.
2. Each segmented slice is converted to greyscale, where pixel values represent the distance of the pixel to the organ boundary. Any geometrical distance can be used (Manhattan, Euclidean...). A sign convention is assigned to the distances. For instance, pixels within the boundary can be negative, while outsiders are positive. This can be reversed without affecting the result. An example is shown in Figure 3.13.
3. Intermediate slices are estimated by interpolating each pixel in the $N \times M$ slice along the z axis. Any of the interpolation methods described above can be used, such as nearest-neighbour, linear or spline. The reconstructed shape is obtained by thresholding the resulting stack at zero.

Shape-based interpolation produces a 3D, isometric, binary model of the organ. Compared to the classical method of grey-value interpolation, this approach gives more accurate quantitative measures in volume and shape, as well as enhanced visualisation, particularly when inter-slice spacing is large [72].

*Matlab R2017b, The MathWorks Inc., Natick (US)

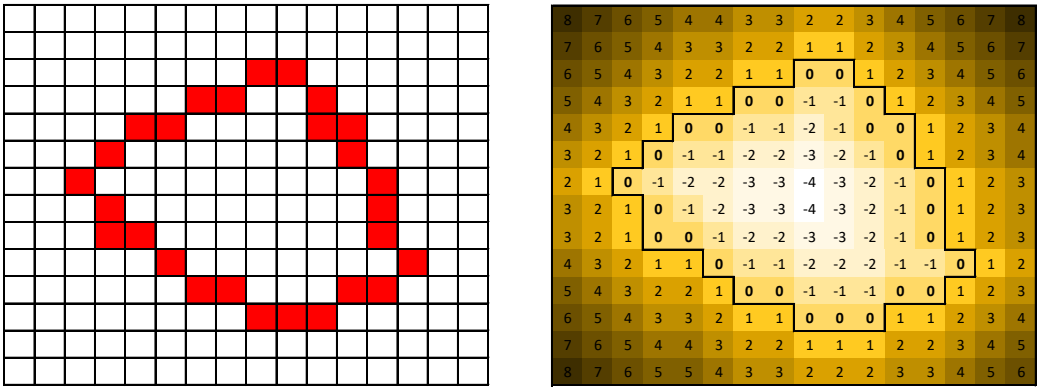


Figure 3.13: Example of a segmented slice (left) and its corresponding distance grey-value image (right). Pixels inside the boundary have negative values, and positive outside

One of the main drawbacks of this method is in reference to the reconstruction of the shape's extrema. With classical greyscale interpolation, slices immediately adjacent to the top and bottom of the object can provide some information about its shape at the endings. However, with shape-based interpolation, due to its binary nature, any more information beyond the slide where the object was last identified (and, equally, before the slide where it was first detected) cannot be inferred [71]. Therefore, a suitable extrapolation algorithm must be used in these extrema regions to produce natural-looking, smooth endings.

3.3.2 Image Registration

Image registration is the process by which a set of images is projected onto a common coordinate system, so that they are aligned and spatial correspondences can be established between them. It is a widely used technique in medical imaging, with many diagnostic and therapy planning tools relying on visual information obtained from different image modalities, such as CT, PET, magnetic resonance (MRI) or ultrasound (US). In many cases, different modalities provide complementary information about the organ of interest. For instance, the fusion of PET and CT images is commonly used in cancer diagnosis. The former shows the metabolic fingerprint of the tumour microenvironment, whereas the latter provides anatomical information. Similarly, the registration of cardiac MRI, CT and US data has been proven to enhance the quality of dynamic images of the heart. This translates into more accurate diagnoses and surgical interventions [73]. Another common application of medical image registration is the construction of

anatomical atlases, where a given organ or tissue is imaged in different subjects. The fusion of these images allows the study of inter-patient anatomical variability of a structure or organ.

In any registration task, one image will always serve as reference, remaining unchanged throughout the process, while a moving image is transformed to match the reference. If the images are regarded as rigid entities, they can be registered by simply applying rotation and translation. This is known as *rigid registration*. If the images are treated as deformable bodies, an affine transformation or a non-linear function is needed to match both datasets, and is called *non-rigid registration* [74]. The work presented in this thesis tackles two separate image registration problems: firstly, pathology and PET/CT tumour volumes are aligned using a global rigid registration. Secondly, histology cross-sections of the tumour are elastically deformed to match their corresponding block-face photographs. The methods used for these two different tasks are reviewed below.

Rigid registration

A rigid registration is performed by rotating and translating an image towards the reference. For 2D images, it can be written as:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{3.21}$$

where (x', y') are the transformed coordinates of point $P = (x, y)$, $T = (t_x, t_y)$ is the translation vector and θ represents the rotation of the moving image in counter-clockwise direction [74]. These parameters can be estimated using various methods.

One of such methods is **Iterative Closest Point** [42]. This is a recursive algorithm for registering two point clouds based on least squares minimisation of their element-wise distances. It alternates between two main processes: point matching and transformation.

In the *point matching* step, two point clouds $M, R \in \mathbb{R}^d$ are defined, with elements m_i and r_j, respectively, $i = 1, \ldots, p$ and $j = 1, \ldots, q$. M is to be moved to match the reference cloud R. We want to find a function $\psi : M \to R$ which finds one-to-one point correspondences that minimise the root mean squared distance

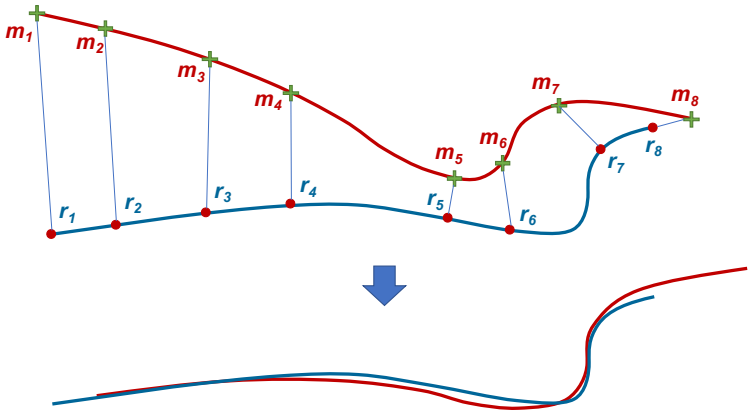


Figure 3.14: Diagram showing point matching and transformation steps in ICP algorithm. Once one-to-one point correspondences are established between the moving (red) and target (blue) datasets, a rigid transformation is applied, such that the sum of squared Euclidean distances between point pairs (m_i, r_i) is minimised

between sets M and R, that is:

$$\hat{\psi} = \arg\min_{\psi} \sqrt{\frac{1}{p} \sum_{m \in M} \|m - \psi(m)\|^2} \tag{3.22}$$

In other words, $\hat{\psi}$ finds the nearest neighbour in set R for each point $m_i \in M$. This can be calculated with Voronoi diagrams or k-d trees [75].

In the *transformation* step, once all one-to-one point matches have been optimised, $\hat{\psi}(m_i) = r_i$ can be defined. That is, point $m_i \in M$ is matched to $r_i \in R$, for $i = 1, \ldots, p$. Next, as expressed in Equation 3.23, the algorithm finds an optimum rigid transformation (orthogonal rotation matrix A and a translation vector b) that minimise the squared Euclidean norm between both point clouds. This concept is illustrated in Figure 3.14.

$$F(A, b) = \arg\min_{A,b} \sum_{i=1}^{n} \|Am_i - b - r_i\|^2 \tag{3.23}$$

Arun *et al.* show that the minimisation problem can be simplified by decoupling rotation and translation [76]. If $\mu(M)$ and $\mu(R)$ are the centroids of sets M and R, respectively, then the translational offset of both sets can be removed by subtracting the centroid from each point:

$$\tilde{m}_i = m_i - \mu(M) \quad \text{and} \quad \tilde{r}_i = r_i - \mu(R) \tag{3.24}$$

Taking into account that translation vector b can be expressed as:

$$b = A\mu(M) - \mu(R) \tag{3.25}$$

if we insert Equation 3.24 into Equation 3.23, we obtain the following simplified minimisation:

$$F(A) = \arg\min_{A} \sum_{i=1}^{n} \|A\tilde{m}_i - \tilde{r}_i\|^2 \tag{3.26}$$

Expanding the right-hand term in Equation 3.26 we obtain:

$$\begin{aligned} F(A) &= \arg\min_{A} \sum_{i=1}^{n} \left(A\tilde{m}_i - \tilde{r}_i\right)^T \left(A\tilde{m}_i - \tilde{r}_i\right) \\ &= \arg\min_{A} \sum_{i=1}^{n} \left(\tilde{m}_i^T A^T A\tilde{m}_i + \tilde{r}_i^T\tilde{r}_i - \tilde{m}_i^T A^T\tilde{r}_i - \tilde{r}_i^T A\tilde{m}_i\right) \\ &= \arg\min_{A} \sum_{i=1}^{n} \left(\tilde{m}_i^T\tilde{m}_i + \tilde{r}_i^T\tilde{r}_i - 2\tilde{r}_i^T A\tilde{m}_i\right) \end{aligned} \tag{3.27}$$

which is equivalent to maximising the last term of the sum:

$$\begin{aligned} F'(A) &= \arg\max_{A} \sum_{i=1}^{n} \left(\tilde{r}_i^T A\tilde{m}_i\right) \\ &= \arg\max_{A} \left[\mathrm{tr}\left(\sum_{i=1}^{n} A\tilde{m}_i\tilde{r}_i^T\right)\right] \\ &= \arg\max_{A} \left[\mathrm{tr}\left(AH\right)\right] \end{aligned} \tag{3.28}$$

where $\mathrm{tr}(\cdot)$ is the matrix trace operation and H is defined as $\sum_{i=1}^{n} \tilde{m}_i\tilde{r}_i^T$. If we perform Singular Value Decomposition (SVD) on H:

$$H = U\Sigma V^T \tag{3.29}$$

where U and V are orthogonal matrices and Σ is a diagonal matrix, it can be proved that the rotation matrix $\hat{A}$ that maximises Equation 3.28 is given by (see [76] for full derivation):

$$\hat{A} = VU^T \tag{3.30}$$

Finally, translation vector b can be calculated by substituting $\hat{A}$ into Equation 3.25.

The point matching and transformation steps are repeated until the algorithm converges, that is, when the change in root mean squared distance between consecutive iterations falls below a given threshold. In this work, iterative closest point provides a fast method for registering volume reconstructions of the same real-world object, taken with different imaging modalities. This applies to the phantom study, where the shape was reconstructed from CT and pathology cross-sectional images and aligned to the virtual CAD model.

Other rigid registration methods depend on higher hierarchical image features, related to the statistical and geometrical properties of the data. Principal component analysis and minimal bounding box are some examples of methods that fall within this category.

Principal Component Analysis (PCA) is a dimensionality reduction statistical procedure. It works by projecting a dataset onto a new, uncorrelated, orthogonal basis, in such a way that few of the basis vectors (known as *principal components*) represent most of the overall data variance [77]. One of its many uses is the registration of segmented shapes by aligning their principal components.

In the three-dimensional case, imagine we have M data points $\mathbf{x}^i = (x_1^i, x_2^i, x_3^i) \in \mathbb{R}^3$, where $i = (1, \ldots, M)$, that represent the pixels belonging to shape Ω. Our goal is to find a 3×3 matrix A that can project $\mathbf{x}$ onto a new basis so that the projection $\mathbf{y} = A^T\mathbf{x}$ is quasi-sparse, that is, all but a few of its elements are close (or equal) to zero. Assuming that data points are centred around the origin (zero mean), then the expected value E of $\mathbf{y}\mathbf{y}^T$ can be obtained from:

$$\begin{aligned} E(\mathbf{y}\mathbf{y}^T) &= E(A^T\mathbf{x}\mathbf{x}^T A) \\ &= A^T E(\mathbf{x}\mathbf{x}^T) A \\ &= A^T \Sigma_x A \end{aligned} \tag{3.31}$$

where Σ_x is the covariance matrix of $\mathbf{x}$, defined as:

$$\Sigma_x = \begin{pmatrix} E(x_1, x_1) & E(x_1, x_2) & E(x_1, x_3) \\ E(x_2, x_1) & E(x_2, x_2) & E(x_2, x_3) \\ E(x_3, x_1) & E(x_3, x_2) & E(x_3, x_3) \end{pmatrix} \tag{3.32}$$

where each element is computed from:

$$E(x_i, x_j) = \frac{1}{M}\sum_{k=0}^{M}(x_i^k x_j^k) \tag{3.33}$$

If we populate the columns of matrix A with the eigenvectors of Σ_x , then $A^T\Sigma_x A$ will become a diagonal matrix whose values $\boldsymbol{\lambda}$ correspond to the eigenvalues of Σ_x. If these are ordered so that $\lambda_1 \geq \lambda_2 \geq \lambda_3$, then the eigenvectors associated with the largest eigenvalues will be the main principal components of Ω [74]. If we repeat the same procedure for another shape Φ representing the same scene, then we can register them by aligning their principal components, as described next.

Let us define $u, v \in \mathbb{R}^3$ as the principal axes of shapes Ω and Φ, respectively. Let $\hat{u} = u/\|u\|$ and $\hat{v} = v/\|v\|$ be the unit vectors defining the principal axes. The rotation from $\hat{u}$ to $\hat{v}$ can be defined as a 2D rotation on the plane Π with normal equal to $\hat{u} \times \hat{v}$, where $\times$ represents the cross product operator. As shown in Equation 3.21, a standard rotation matrix can be expressed as:

$$R = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \hat{u}\cdot\hat{v} & -\|\hat{u}\times\hat{v}\| & 0 \\ \|\hat{u}\times\hat{v}\| & \hat{u}\cdot\hat{v} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.34}$$

as $\cos(\theta) = \hat{u} \cdot \hat{v}$ and $\sin(\theta) = \|\hat{u} \times \hat{v}\|$. However, this rotation is expressed in the basis B containing plane Π. This basis $B = (i, j, k)$ is defined by the following three vectors:

$$\begin{aligned} i &= \hat{u} && \text{(normalised projection of } \hat{v} \text{ onto } \hat{u}) \\ j &= \frac{\hat{v} - (\hat{u}\cdot\hat{v})\hat{u}}{\|(\hat{u}\cdot\hat{v})\hat{u}\|} && \text{(normalised rejection of } \hat{v} \text{ onto } \hat{u}) \\ k &= \hat{u} \times \hat{v} && \text{(cross product between } \hat{u} \text{ and } \hat{v}) \end{aligned} \tag{3.35}$$

Therefore, the rotation matrix R' defined in the canonical basis can be expressed as:

$$R' = B^{-1}RB \tag{3.36}$$

The translation vector can be easily computed by subtracting the centroids of both shapes after applying rotation.

Minimal Bounding Box (MBB) registration is another algorithm that relies on the geometry of the shape to determine its orientation. The minimal bounding box of a volume is the smallest parallelepiped that fully encloses it. Dura and Domingo [78] proposed a method for pose normalisation of anatomical structures based on the alignment of their minimal bounding boxes. Its performance was compared against the PCA-based approach previously described. The authors claim that MBB outperforms PCA when aligning liver shapes segmented from MRI scans, and also that MBB is more robust for registration of quasi-symmetrical volumes. The work presented in this thesis is based on the registration of lung tumour 3D models. These shapes are typically rounded, without protruding features or large asymmetries. Therefore, based on the findings reported in [78], minimal bounding box registration was regarded as the best fit for our particular application. A Matlab toolbox written by Korsawe [79] was used to compute the MBB for our tumour shapes.

Non-rigid registration

Deformable registration methods generate a non-linear mapping between the moving image and the reference. This type of transformation is applicable when the shapes involved correspond to elastic entities in the real world, as is the case for the majority of human organs and anatomical structures. Most non-rigid registration problems are ill-posed, meaning that they do not have a unique solution and the solution space is unfeasibly large. One way to overcome this is by introducing a regularisation term in the registration function, that is, we assume that the object has some physical properties that limit its deformation. There is a plethora of functions that can be used to perform these transformations, such as thin-plate splines, multiquadric interpolation, piecewise polynomials or moving least squares [74].

Moving Least Squares [29] is an interactive registration method where the user selects a number of control points on the image of interest, and when these are moved, the image deforms accordingly. Let p_i be a set of control points for $i = 1, \ldots, N$, and q_i the positions of p_i after applying the deformation. Given a point v in the image, we need to find the transformation function $l_v(x)$ that minimises the following expression [74]:

$$\sum_{i=1}^{N} w_i \|l_v(p_i) - q_i\|^2 \tag{3.37}$$

where

$$w_i = \frac{1}{\|p_i - v\|^\alpha} \tag{3.38}$$

is a non-negative monotonically decreasing weight function which ensures that a data point v will be most influenced by its closest control point p_i. The free parameter α controls the behaviour of the inverse distance relationship. Note that, in Equation 3.37, function $l_v(x)$ is particular to point v, which means that for each point in the image, the parameters of the function need to be re-estimated. Therefore, for the algorithm to be fast, $l_v(x)$ has to be simple. A popular choice is using an affine transformation model, defined by a linear map M and a translation vector t [29]:

$$l_v^A(x) = xM + t \tag{3.39}$$

As already shown in Equation 3.25, the problem can be simplified by ignoring the translation term and calculating it a posteriori as $t = \tilde{q} - \tilde{p}M$, where $\tilde{p}$ and $\tilde{q}$ are the weighted centroids of point sets p and q:

$$\tilde{p} = \frac{\sum_{i=1}^N w_i p_i}{\sum_{i=1}^N w_i}, \quad \tilde{q} = \frac{\sum_{i=1}^N w_i q_i}{\sum_{i=1}^N w_i} \tag{3.40}$$

If we substitute t in Equation 3.39, we obtain:

$$l_v^A(x) = (x - \tilde{p})M + \tilde{q} \tag{3.41}$$

and the least squares minimisation in Equation 3.37 becomes:

$$\sum_{i=1}^N w_i \|\dot{p}_i M - \dot{q}_i\|^2 \tag{3.42}$$

where $\dot{p}_i = p_i - \tilde{p}$ and $\dot{q}_i = q_i - \tilde{q}$ are the zero-mean point sets. There is a closed-form solution for M in Equation 3.42, based on the classical normal equation, which in the case of affine models takes the following form [29]:

$$M = \left(\sum_{i=1}^N \dot{p}_i{}^T w_i \dot{p}_i \right)^{-1} \sum_{j=1}^N w_j \dot{p}_j{}^T \dot{q}_j \tag{3.43}$$

Therefore, the deformation function for affine moving least squares $l_v^A(x)$ at a given point v becomes:

$$l_v^A(v) = (v - \tilde{p}) \left(\sum_{i=1}^{N} \dot{p_i}^T w_i \dot{p_i} \right)^{-1} \sum_{j=1}^{N} w_j \dot{p_j}^T \dot{q_j} + \tilde{q} \tag{3.44}$$

This deformation function indicates the new position of point v in the deformed image. When v approaches p_i, the weight function w_i tends to infinity and $l_a(v)$ interpolates, that is, $l_a(p_i) = q_i$.

Schaefer *et al.* [29] also suggest different deformation functions, that can be obtained by enforcing M to represent other linear transforms. An interesting case is the use of a rigid transformation matrix, as it avoids non-uniform scaling and shear. These effects are undesirable for our particular problem, as lung histology sections do not obey such deformation rules. Again, there is a known closed-form solution for this problem. The resulting deformation function for rigid moving least squares $l_v^R(x)$ is:

$$l_v^R(x) = \sum_{i=1}^{N} \dot{q_i} \frac{A_i}{\mu_r} + \tilde{q} \tag{3.45}$$

where

$$A_i = w_i \begin{pmatrix} \dot{p_i} \cdot (x - \tilde{p}) & -\dot{p_i} \cdot (x - \tilde{p})^\perp \\ -\dot{p_i}^\perp \cdot (x - \tilde{p}) & \dot{p_i}^\perp \cdot (x - \tilde{p})^\perp \end{pmatrix} \tag{3.46}$$

and

$$\mu_r = \sqrt{\left(\sum_{i=1}^{N} w_i \dot{q_i} \dot{p_i}^T \right)^2 + \left(\sum_{i=1}^{N} w_i \dot{q_i} \dot{p_i}^{\perp T} \right)^2} \tag{3.47}$$

and the 2D operator $a^\perp$ returns a perpendicular vector to a, that is, if a has coordinates (α, β), $a^\perp$ corresponds to $(-\beta, \alpha)$.

This registration algorithm was implemented using the ImageJ plugin *Moving Least Squares registration*, published by Schindelin [80].

Performance evaluation of registration tasks

Registration methods aim to achieve a meaningful overlap between two or more images. Their performance can be evaluated in several ways, depending on the metrics considered.

The most obvious and naive evaluation is **qualitative visual inspection** by an expert. A trained clinician is likely to identify common structures on both datasets and assess the quality of the registration. Although fast and simple, this method is not reproducible, as it is highly dependent on the individual carrying it out. Also, it does not provide a numerical outcome. However, in some cases, there is no objective ground truth to compare a registration algorithm to, due to the complexity of the task or the availability of partial information only. In these cases, the subjective opinion of an expert may be the only way to evaluate the clinical meaningfulness of the registration. Multiple evaluations performed independently by different individuals may remove potential sources of bias. A scoring system or rubric must be established beforehand, either binary (right/wrong) or graded. This enables examiners to use the same criteria, ultimately leading to more consistent results.

A quantitative assessment of how well aligned a set of feature points are is given by **Target Registration Error** [81]. This measures the root mean square distance between corresponding landmarks on the reference and moving images that have not been used for calculating the deformation map. It is defined as:

$$\mathbf{TRE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}|m_i - r_i|^2} \tag{3.48}$$

where m_i and r_i are the positions of the i-th feature point on the moving and reference images, respectively, for $i = 1, \ldots, N$.

Another approach is to measure the overlap between registered structures. The **Dice Similarity Coefficient** (DSC) calculates the ratio between the intersection of two sets and their individual areas [82]:

$$\mathbf{DSC}(A, B) = \frac{2 \cdot |A \cap B|}{|A| + |B|} \tag{3.49}$$

where $\cap$ represents the intersection operator between two sample collections and $|\cdot|$ indicates the cardinality of the set, that is, the number of elements in it. A DSC of 1 indicates complete overlap, and 0 no overlap at all. Figure 3.15 shows a graphical representation of the concept. DSC is widely used for evaluating medical image registration and segmentation tasks, where the output of an automatic algorithm is compared against the ground truth produced by an expert.

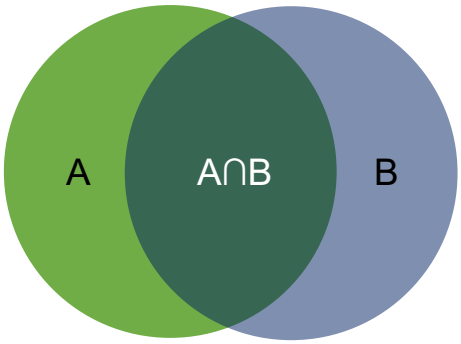


Figure 3.15: Dice Similarity Coefficient calculates the ratio between the intersection of sets A and B and their individual areas

The distance between the boundaries of two registered objects can also be useful to assess the performance of a registration algorithm. The **Hausdorff distance** measures the maximum distance of a point set to the closest point in the other set. Its mathematical formulation is defined as [83]:

$$d_H(A,B) = \max\{\sup_{a\in A}\inf_{b\in B} d(a,b), \sup_{b\in B}\inf_{a\in A} d(a,b)\} \tag{3.50}$$

where *sup* and *inf* are the supremum and infimum operators, respectively. In other words, for every point in set A, find the closest point in set B. Once every point in A has been evaluated, pick the largest of the shortest distances. As this measure is asymmetric, the same calculation is repeated from set B to set A. The largest of these two is defined as the Hausdorff distance between sets A and B.

Chapter 4

Phantom Study

The design of several tissue-mimicking phantoms, used to test our volume reconstruction algorithms, are discussed in this chapter. First, the reasons for using synthetic phantoms in our experiments are explored. In the second section, the materials and methods used to build and process each sample are explained. Results are presented and discussed in the last section.

4.1 Motivation

One of the main tasks in the project is to reconstruct an anatomical 3D model of the tumour from sequences of cross-sectional images, namely CT and block-face pathology photographs. However, in most tomographic imaging modalities, the researcher only has access to partial information. A volumetric scan is a stack of parallel slices separated by a fixed distance. The image values situated between these planes are either projected onto the closest slice, averaged, or lost. For this reason, evaluating the performance and accuracy of volume reconstruction algorithms becomes challenging.

The use of a synthetic phantom offers several advantages in this regard, some of which are listed below:

- **Ground truth**: the exact physical measurements of the phantom are known prior to imaging. Therefore, we can compare the reconstructed volume against the ground truth model to assess different characteristics of the reconstruction algorithm, such as overall volume recovery, or local shape reproduction.

- **Customisation**: a phantom can be purposefully designed to present certain shape or texture traits. These features can mimic real-world characteristics which are normally present in a tumour, or impose challenging conditions to test the performance of different reconstruction algorithms.
- **Reproducible**: a phantom can be easily duplicated, keeping the same physical attributes. Reproducibility is especially useful when comparing reconstructions obtained with the same method, but under different conditions.
- **Ethics**: unlike human tissue, synthetic phantoms are not subject to ethical or legislative regulations. Similarly, it avoids the need to obtain patient consent, which is required for the collection and use of human samples.

For these reasons, the initial experiments were performed on several phantoms. Different materials and configurations were tested to achieve similar conditions to those of lung tumours. These are reviewed in the following section.

4.2 Materials and Methods

4.2.1 Phantom Shape

The shape of the phantom has a direct impact on the performance of the volume reconstruction algorithms. Depending on the spacing between consecutive cuts, some details may not be recovered—especially finer or fast-changing features. Therefore, a phantom can be designed to test specific aspects of the reconstruction methods, for instance how well a particular shape is recovered.

Pyramidal

A square pyramidal shape was initially used, due to its simplicity and the availability of off-the-shelf moulds. The prototype used was 12 cm in height, with a base of 8 cm squared. Another advantage of this shape is that it has a constant, known slope. Therefore, the distance between any two slices can be verified from the geometry of their cross-sections.

Anatomical

A more complex shape was needed to approximate the actual geometry of a lung tumour, to test our reconstruction algorithms on a more realistic model. Some of the desired new features were: presence of both deep and shallow grooves along its slicing direction, and inclusion of high-frequency details on its surface. It also had to be easily reproducible and measurable.

The initial approach consisted of customising a cylindrical mould (12 cm high, 5 cm diameter) using plumber's putty, a pliable material made from clay and oil. The cylinder was cut in half lengthways, and several features made with putty were attached to its walls. This method allowed us to produce different prototypes in a short space of time. However, it also had several disadvantages:

1. The level of detail was limited by the mechanical properties of the putty. Only large, smooth features could be produced without them breaking or deforming.
2. The two halves of the mould have to be joined and sealed before injecting the phantom material. However, this task proved difficult due to the narrow cross-section of the cylinder (1 mm wall thickness).
3. Some of the features made with putty stuck to the phantom material and detached from the cylinder walls, preventing reproducibility.
4. The exact volume of the phantom could not be precisely calculated, due to the nature of the features.

To overcome these challenges, it was agreed that the most efficient way to proceed was to generate a virtual shape using computer-aided design (CAD) software, and use it to 3D-print a negative mould. The only requirement was that the mould had to be reusable, and so should not be destroyed when the cast was released. For this reason, the part had to be designed in such a way that it could be produced with a separable, two-part mould. The location of the *parting surface*, i.e. the plane along which the mould splits in two, greatly influences the design: the resulting halves should not present any undercuts (indentations or features that would prevent the shape from being released).

A bespoke phantom model was designed using SolidWorks*, shown in Figure 4.1. Grooves of different depth can be observed both on the upper and lower sides, as

*SolidWorks 2013, Dassault Systèmes, Villacoublay (France)

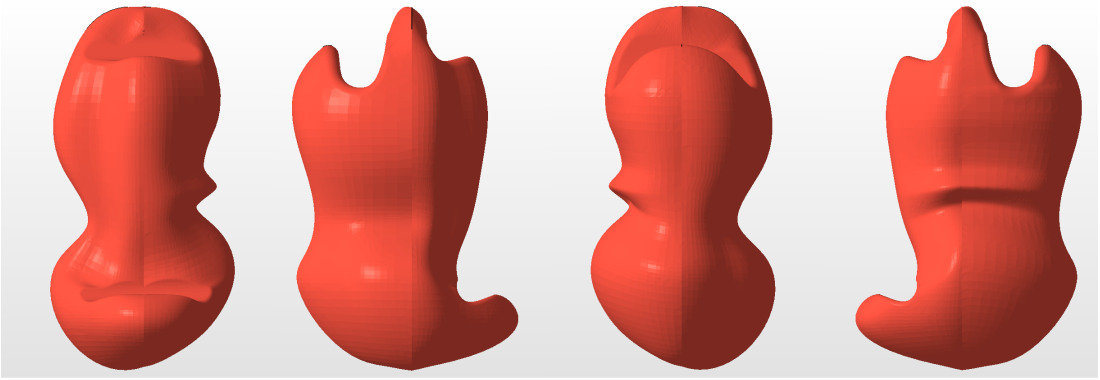


Figure 4.1: Different views of the anatomical phantom CAD model, showing its front, right, back and left sides, respectively

well as a fast-changing detail on the left side. The model was subsequently used to generate a separable injection mould, with the parting surface running along the phantom's saggital plane.

After completing and verifying the design, the mould was 3D-printed in two halves. The material used was acrylonitrile butadiene styrene (ABS), a thermoplastic polymer widely used in fused deposition printing processes. It is a tough, resilient material, with a high impact resistance and is easy to machine [84]. Once printed, an injection inlet was milled in one of the sides, and two holes were drilled on the opposite half to allow air to escape. Cardboard gaskets were laser cut and used to seal the joint between the parts, avoiding material spillage. Two protruding features were also printed on the mould corners to facilitate alignment when coupling both halves together. Figure 4.2 shows the mould with a cast.

To produce a sample of the phantom, a gasket was placed around the joint and the two halves were clamped together. A nozzle was screwed to the injection inlet. Using a syringe, the phantom material was slowly pumped into the cavity, shaking it regularly to avoid the formation of air bubbles. When the substance appeared at the air outlets, the pumping was stopped, and the phantom was left to harden. Once cured, the shape was dislodged, and any excess material was trimmed with a scalpel.

4.2.2 Phantom Material

The choice of a suitable phantom material was fundamental. Besides evaluating different reconstruction methods, phantoms were also used to simulate actual

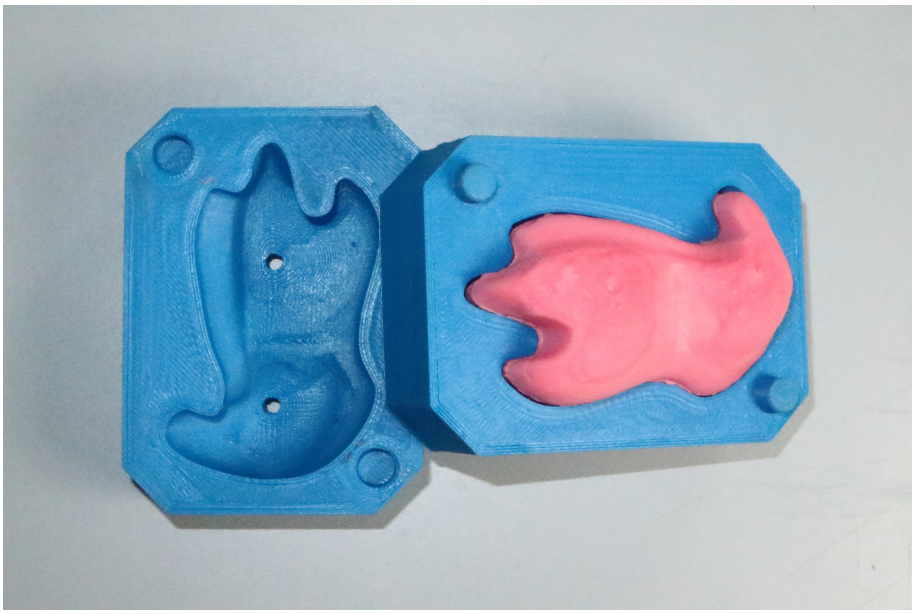


Figure 4.2: 3D-printed phantom mould. The negative alignment features and the two air outlets can be seen on the left-hand part. The right-hand half contains a cast of the phantom made of polyvinyl alcohol, along with the alignment protusions at the corners. Each half is 64 mm wide by 88 mm high

lung tumour tissue, in order to test the study protocol in realistic conditions. Therefore, the substance needed to present comparable characteristics: visible in x-ray and photography modalities, and similar mechanical properties.

Polyvinyl alcohol

Polyvinyl alcohol (PVA) is a non-toxic, water-soluble polymer. PVA hydrogel was prepared by mixing 7% by weight PVA granules* with water at room temperature. The solution was stirred and slowly heated to 95°C, until the granules were completely dissolved. The viscous mixture was allowed to cool down to room temperature, at which point a contrast agent could be added. The solution was then injected or poured into a mould and subjected to various freeze-thaw cycles. This was performed using an in-house, automatic thermocycler that froze the sample at -20°C for 12 hours. This was followed by a 12-hour thawing period at 20°C. The 24-hour cycle was repeated between one and three times, where more repetitions resulted in a stiffer hydrogel [85].

*Poly(vinyl alcohol) 99+% hydrolysed (prod. 341584), Sigma-Aldrich UK Ltd, Dorset (UK)

Agar

Agar is a jelly-like substance of vegetal origin. Its main component is agarose, a naturally-occurring polymer. Agar hydrogel was prepared by mixing 4% by weight agar powder* with water at room temperature. Once dissolved, the mixture was heated to 98°C while continuously stirring, until it became a thick fluid. At this point, a CT contrast agent could be mixed in. The solution had to be transferred to the mould while still hot, before it solidified at approximately 40°C.

Agar gel was also used as an embedding medium, to hold the specimens in position during pathology processing, and was prepared following the same procedure.

Addition silicone

Vinyl polysiloxane (VPS) is an addition-reaction, two-element silicone†. Its two components, base and catalyst, are stored separately. When mixed together, the substance quickly becomes a rubber-like, solid material. It is widely used in dentistry to produce impression moulds.

Due to the high viscosity of the material, it was unfeasible to inject the mixture through the inlet nozzle. Therefore, the substance was applied uniformly to both halves of the mould, until the cavities were filled, and the two parts were subsequently coupled together. The curing time for this material is 6 minutes, after which the cast is ejected from the mould.

Note that contrast agents could not be added to this material, again due to its high viscosity and short curing time.

Contrast agents

One of the phantom's purposes was to mimic visual and scattering attributes of lung tumours. For this reason, two different CT contrast agents were added to the preparations, namely aluminium oxide and graphite. Red food colouring was also used as a dye in some samples. Aluminium oxide‡ is a chemical compound with the formula Al_2O_3. It is an inert, white substance widely used in industry as plastic filler, electrical insulator and paint pigment. It is non-toxic

*Agar-Agar (E406) powder, Special Ingredients Ltd, Chesterfield (UK)
†Extrude^{TM}VPS Impression Material (prod. 300916), Kerr Corp., Orange (US)
‡Aluminium oxide powder (3 μm particle size), Logitech Ltd, Glasgow (UK)

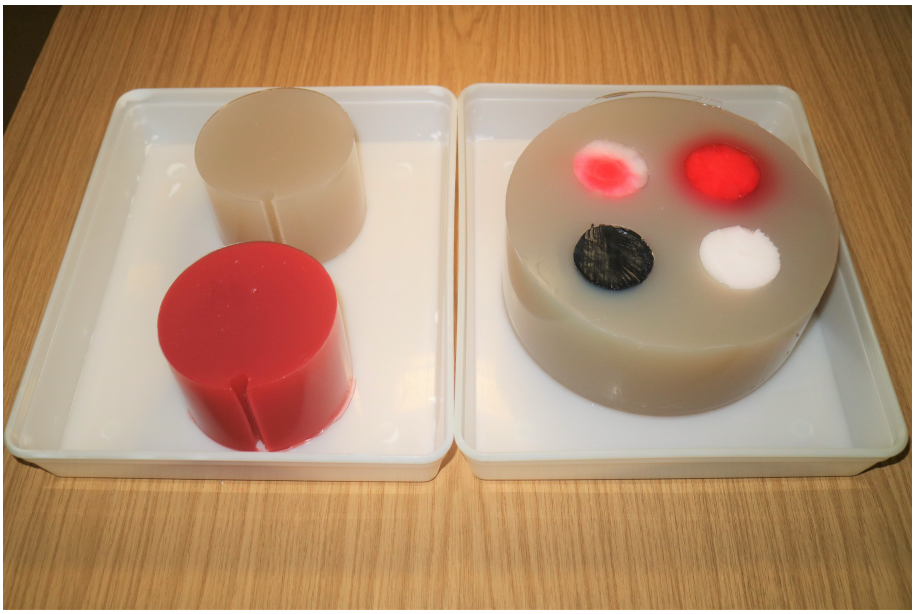


Figure 4.3: On the left, agar samples (110 mm in diameter) with 4% (red) and 2% (uncoloured) concentrations. On the right, four PVA samples (45 mm in diameter) with different scattering agents: high-density PVA rod (red and white), PVA control (red), PVA with aluminium oxide (white), and PVA with graphite (black). The four samples are embedded in 3% agar gel

and non-flammable, and therefore considered a low-risk substance. Graphite* is a crystalline form of carbon, commonly used in steelmaking and refractory industries. Its high electrical conductivity makes it an ideal material for electrical motor components and batteries. It is black in colour, and chemically stable under normal operating conditions.

To test their x-ray absorption properties, four different PVA samples were built: a) a control sample made with 7% PVA and food colouring, b) a high density (15%) PVA rod embedded in 7% PVA, c) a 7% PVA mixed with 1% graphite powder, and d) a 7% PVA mixed with 1% aluminium oxide. All four samples were embedded in 3% agar gel, as per protocol guidelines. Different agar concentrations (2% and 4%) were also tested. Figure 4.3 shows the two agar samples on the left tray, and the four embedded PVA samples with different contrast agents on the right.

All samples were scanned in a CT scanner† at the Beatson West of Scotland Cancer Centre (Glasgow, UK). Images were retrieved from the imaging servers and visualised using 3D Slicer [86]. A Gaussian blur filter with $\sigma = 1.5$ was

*Graphite powder (44 μm particle size), Easy Composites Ltd, Stoke on Trent (UK)
†GE Discovery CT590, GE Healthcare, Chicago (US)

Sample composition	Avg. absorption (HU)
PVA 7%	19
PVA 15%	28
PVA 7% with Al_2O_3 1%	25
PVA 7% with graphite 1%	23
Agar 2%	10
Agar 3%	14
Agar 4%	19
Agar 4% with Al_2O_3 1%	28
Agar 4% with Al_2O_3 3%	41
Agar 4% with Al_2O_3 5%	53

Table 4.1: Average absorption coefficients for different phantom samples

applied to remove noise and average neighbouring values. High-density PVA sample showed the highest absorption, between 27 and 29 Hounsfield Units (HU). With regard to the contrast agents used, aluminium oxide showed slightly higher attenuation compared to graphite. It was proven that a higher agar concentration translates to a higher absorption.

Different concentrations of aluminium oxide (1%, 3% and 5%) were also mixed with agar and CT-scanned. As expected, the sample with the highest Al_2O_3 concentration presented the highest contrast. Average absorption of 4% agar with 5% aluminium oxide was 54 HU, compared to 13 HU for the embedding medium (agar at 3% concentration). Results are summarised in Table 4.1.

4.2.3 Slicing Rig

Pathological dissection of a surgical specimen is usually performed free-hand by a pathologist. The tissue is manually sliced at several arbitrary points along the tumour, and a representative number of the samples are selected for histopathology analysis. Although fast, the method presents two main issues: a) it is imprecise in terms of slice thickness uniformity; and b) lung tissue is likely to tear and deform when sliced, especially if sections are thin, due to its mechanical properties.

Volume reconstruction would be less difficult when pathology slices are parallel and uniformly spaced. For this reason, a bespoke soft tissue slicing rig was designed and built in-house by the Department of Mechanical Engineering (Medical Devices Unit, NHS Greater Glasgow and Clyde, UK). It consists of a cylindri-

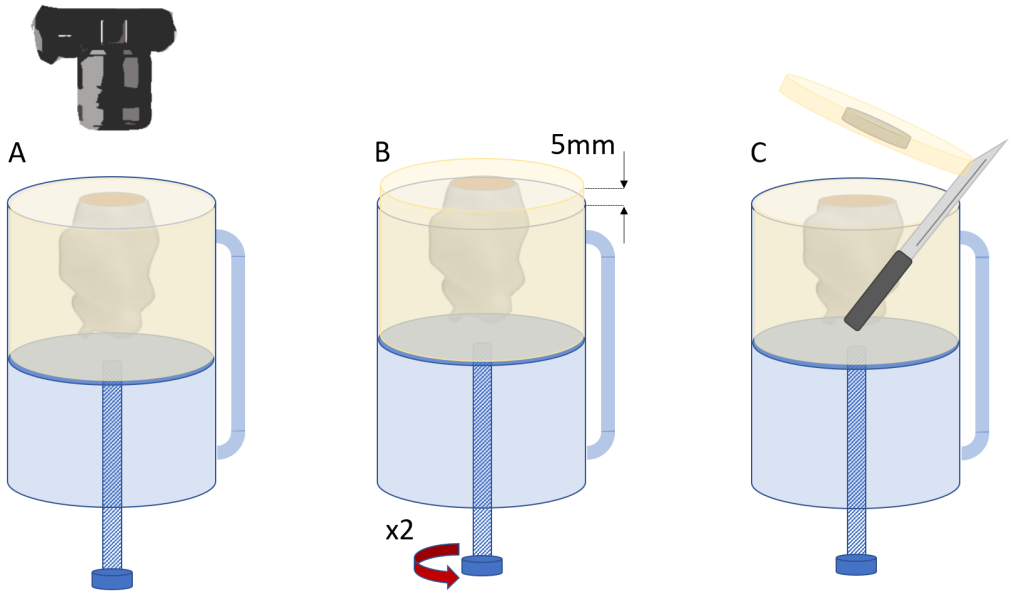


Figure 4.4: Diagram showing the iterative dissection process using the slicing rig: (A) photograph the exposed cross-section, (B) lift the embedded tissue 5 mm by performing two full turns of the threaded plunger, and (C) slice the specimen

cal tissue container that slides into an acrylic frame. This structure features a threaded plunger at the bottom, that pushes the tissue upwards, and a surgical blade at the top. Using the plunger, the tissue (embedded in agar) is lifted until the first slice of interest reaches the cutting plane. The knife, constrained by two horizontal slits, is then used to perform a clean, flat cut. The plunger thread has a known pitch of 2.5 mm, allowing the slice thickness to be adjusted with the number of turns. The exposed cross-section is photographed and the slice is stored away for further processing. The sample is then lifted 5 mm by performing two full turns, and subsequently sliced. The new exposed plane is photographed, and so on. This iterative process, shown in Figure 4.4, is repeated until the whole tissue has been imaged.

The first prototype of the slicing rig consisted of a small cylindrical container (110 mm inner diameter, 220 mm high) and a frame made of 5 mm-thick acrylic sheet. A 24-megapixel digital camera* was suspended horizontally over the slicing rig using a 90° rotating camera tripod†. External illumination sources were placed around the rig, to enhance the quality of the photographs. This set-up was

*Canon EOS M3 digital camera, Canon Inc., Tokyo (Japan)

†Manfrotto 190XPRO Aluminium 3-Section camera tripod, Vitec Imaging Solutions Spa, Ashby-de-la-Zouch (UK)

Figure 4.5: Second prototype of the slicing rig, mounted on its trolley

not robust in terms of preventing relative movement between the rig and the camera, which would hinder the posterior volume reconstruction methods. It was also difficult to operate due to the protrusion of the plunger thread when fully retracted, for which the rig had to be suspended above the workbench. Moreover, its size would not fit larger specimens.

For these reasons, after collecting feedback from all users, a second version was built. 8 mm acrylic sheet was used for the main frame, which would house a container 210 mm in diameter. A reinforced 4 cm base, with a built-in brass nut, provided stability to the threaded plunger. A bespoke trolley was constructed to hold the frame, with a central aperture through which the thread could retract. Stands for the light sources and the camera were attached to the trolley, therefore minimising the risk of movement between the camera and the specimen. Due to the increased size of the container, the 30 cm knife was more prone to bending at its middle section, yielding slices of non-uniform thickness. To solve this problem, the blade was adapted to fit a hacksaw frame, which was used to apply tension. The second prototype of the slicing rig, mounted on its trolley, is shown in Figure 4.5.

4.2.4 CT Calibration Phantom

A standardised PET/CT calibration phantom* was used to assess the precision of volume delineation tasks on a CT scan. It consisted of a phantom body made of 3 mm acrylic sheet, a lung insert and six fillable spheres with outer diameters of 12 mm, 15 mm, 19 mm, 24 mm, 30 mm and 39 mm. When filled with water, the spheres presented higher attenuation than the surrounding air, similar to that of a real lung tumour scenario. The design complied with the International Electrotechnical Commission (IEC) Standard 61675-1.

4.2.5 Image Processing Methods

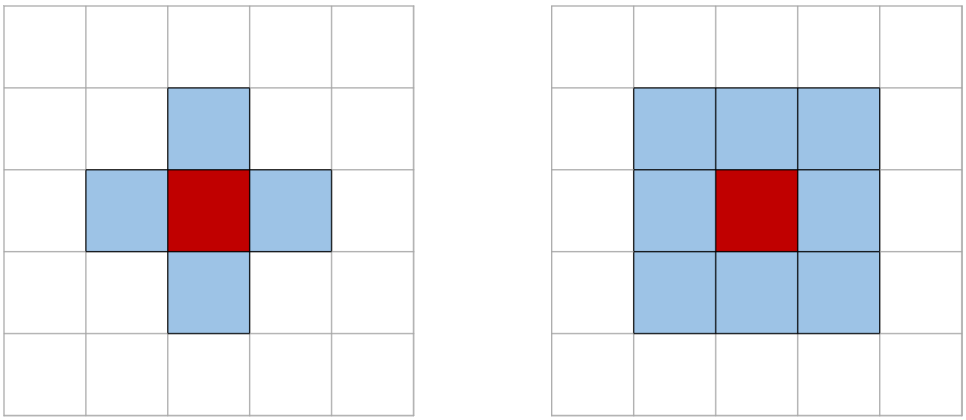
Region growing segmentation

Region growing is an interactive segmentation algorithm [87]. Unlike thresholding approaches, this method tolerates smooth colour variations within the target area, making it more robust in situations where the scene presents non-uniform illumination.

The user selects an arbitrary pixel (seed) within the area to be segmented. Next, the method performs one-to-one comparisons between the seed and its neighbouring pixels: if their colour difference (i.e. image gradient) is lower than a user-selected threshold, that neighbour is included in the region; if higher, the pixel is rejected. In the next iteration, the included pixels evaluate their corresponding neighbours in the same fashion, being added to the region if the condition is met. The algorithm recursively evaluates neighbouring boundary pixels until it converges and the region cannot grow anymore. After convergence, a morphological closing is performed on the resulting mask, to remove any small holes produced by noise.

A pixel's neighbours are defined by the pixel's *connectivity*. The neighbours of a 4-connected pixel include every pixel that touches one of its edges. The neighbours of an 8-connected pixel are all pixels in contact with either one of its edges or corners. This is illustrated in Figure 4.6. The user can select one of the two connectivity options in the algorithm described above.

*NEMA IEC Body Phantom Set™, model PET/IEC-BODY/P, Meditest SAS, Buc (France)



(a) **(b)**

Figure 4.6: Pixel (red) with its corresponding neighbours (blue), in (a) 4-connected and (b) 8-connected settings

Extrapolation of extrema regions

Shape interpolation methods often need to be complemented with an extrapolation algorithm, to produce smooth, rounded edges around the shape's end regions. An *end region* $\mathcal{R}$ is defined as a closed area of a cross-sectional image, present in slice i, but absent in the subsequent plane $i+1$. Extrema are particular cases of end regions. However, the behaviour of these algorithms needs to be modulated in the presence of nearby regions, co-habiting in either the same or contiguous planes. Two approaches are investigated: linear and curvature-based extrapolation.

With the **linear** method, an end region $\mathcal{R}$ is projected onto slice $i+1$, $\hat{\Pi}_{i+1}$, and distance transform $\Delta(\cdot)$ is applied. Next, a region $A = \mathcal{R} \oplus D$ is defined on $\hat{\Pi}_{i+1}$, where $\oplus$ indicates the dilation operation and D is a disk-shaped structural element of radius r. That is, A includes $\mathcal{R}$ and its adjacent neighbours within $r-1$ pixels. Then, for each pixel $p_n \in A$, with value $\Delta(\hat{\Pi}_{i+1}(p_n)) = \delta_n$, the adjusted distance $\dot{\delta}_n$ is calculated as:

$$\dot{\delta}_n = \min\Big(\max\left(1, \delta_n + h\right), \Delta(\Pi_{i+1}(p_n))\Big) \tag{4.1}$$

where h is the slice thickness in pixels, and $\Delta\left(\Pi_{i+1}(p_n)\right)$ is the distance transform of the original $i+1$-th slice at point p_n, i.e. without the projection of $\mathcal{R}$. In other words, the distance to the neighbouring slice is considered in and around the area of the end region $\mathcal{R}$. If, for any pixel within this area, the resulting corrected distance is lower than or equal to one, then it is clipped to 1. That is because the

end region should not appear on slice $i+1$, as enforced by the data. Moreover, if the distance d between a point $p_n \in A$ and a region $\mathcal{S} \in \Pi_{i+1}$ is smaller than the inter-slice distance, then $\dot{\delta}_n = d$. Note that, if the other sign convention is used for the distance transform (positive values inside a region, and negative outside), then Equation 4.1 still holds if $\dot{\delta}_n$ and δ_n are substituted by $-\dot{\delta}_n$ and $-\delta_n$, respectively.

The **curvature-based** method uses second-order differences, from the two preceding slices, to extrapolate the end region, following the same curvature trend. This approach is also applied locally around the area A of the end region $\mathcal{R}$. Let us assume that the last cross-section of region $\mathcal{R}$ is on slice i, and that the region exists on slices $i-1$ and $i-2$. For each pixel $p_n \in A$, the first-order differences between the distance transforms of slices i and $i-1$ are defined as:

$$\nabla'_i(p_n) = \Delta\left(\Pi_{i-1}(p_n)\right) - \Delta\left(\Pi_i(p_n)\right) \tag{4.2}$$

$$\nabla'_{i-1}(p_n) = \Delta\left(\Pi_{i-2}(p_n)\right) - \Delta\left(\Pi_{i-1}(p_n)\right) \tag{4.3}$$

Values obtained in Equations 4.2 and 4.3 are used to calculate the second-order differences:

$$\nabla''_i(p_n) = \nabla'_{i-1}(p_n) - \nabla'_i(p_n) \tag{4.4}$$

If we enforce that $\nabla''_i = \nabla''_{i+1}$, that is, second-order differences in slice i have to be the same for the extrapolated region in slice $i+1$, then the first-order differences for that slice can be computed as:

$$\nabla'_{i+1}(p_n) = \nabla'_i(p_n) - \nabla''_i(p_n) \tag{4.5}$$

Finally, the extrapolated distance values δ_n on slice $i+1$ can be calculated as:

$$\delta_n = \Delta\left(\Pi_{i+1}(p_n)\right) = \Delta\left(\Pi_i(p_n)\right) - \nabla'_{i+1}(p_n) \tag{4.6}$$

The resulting distance image also needs to be adjusted, as in the linear case, to enforce the same data constraints. Adjusted values $\dot{\delta}_n$ can be obtained using Equation 4.1.

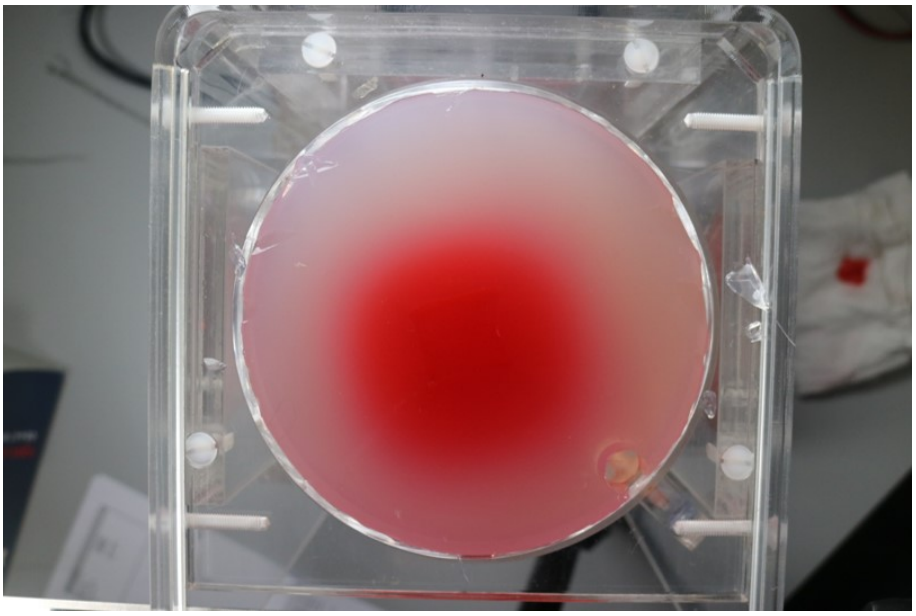


Figure 4.7: The food colouring used to dye the phantom permeated into the surrounding agar, thus blurring the edges of the shape

4.3 Results and Discussion

4.3.1 Phantom Material Testing

The first material tested was 4% agar mixed with red food colouring. The reasons for this initial choice were ease of preparation, and the substances involved being categorised as having minimal health and safety risk [88]. The pyramidal mould was used for this test. Once solid, the phantom was embedded in 2% agar, and left overnight to set. When the sample was sliced the next day in the rig, we discovered that the phantom's red dye had permeated into the surrounding agar, as shown in Figure 4.7. As a consequence, the boundary of the shape was very diffuse, thus hindering its segmentation. Dyed agar was ruled out as a phantom material.

The next substance tested was polyvinyl alcohol (PVA). Different concentrations and contrast agents were used to assess its physical and mechanical properties, as summarised in Table 4.1. The sample that presented suitable x-ray absorption properties and the easiest to slice was PVA at 7% with 1% aluminium oxide, cured for 72 hours (3 freeze-thaw cycles). A trace of red food colouring was also added to the mixture to enhance its contrast in the visible spectrum. Several phantoms were produced with this material, initially using the pyramid shape, and later

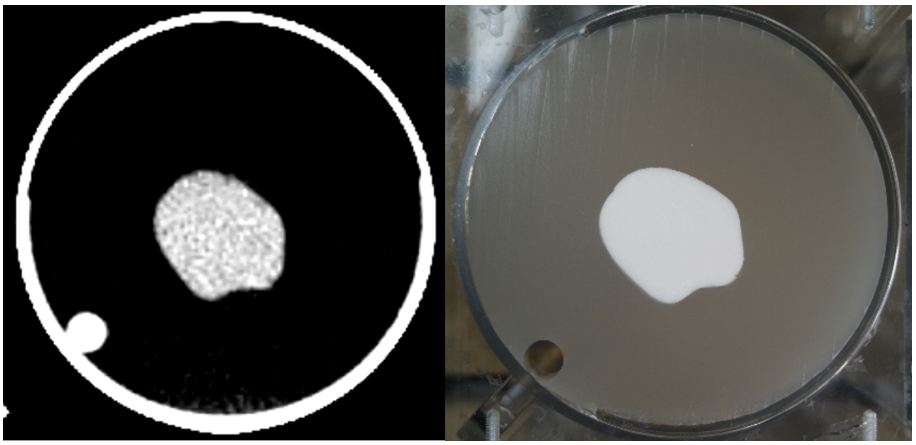


Figure 4.8: Cross-section of a phantom sample (4% agar with 5% aluminium oxide) in the x-ray (left) and visible (right) spectra

with the bespoke 3D-printed mould. In this case, there was no diffusion of dye into the medium, and therefore the cross-sectional contours were well preserved. However, due to the rubbery texture of PVA and its high friction coefficient, the blade was prone to getting trapped in the material when cutting the phantom on the slicing rig. This caused the phantom to move sideways with the movement of the knife, instead of being cut through. Eventually, it would cause the shape to become dislodged from the embedding medium, especially towards the last slices where agar support was weaker.

Agar gel proved easier to slice than PVA. However, some kind of additive was needed to make the phantom distinguishable from its surrounding agar, both in a photograph and in x-ray. The problem with food colouring is that it is a soluble substance, and it transfers to water-based materials easily. In contrast, aluminium oxide is non-soluble, and also gives the mixture a white colour. With this in mind, we created a sample with 4% agar and different Al_2O_3 concentrations, namely 1%, 3% and 5%. As expected, the resulting phantoms were white, and the contrast agent did not permeate into the neighbouring medium. The colour intensity was proportional to the aluminium oxide concentration: the 1% sample presented a vitreous white tone, slightly transparent, whereas at 5% it showed a solid, opaque white colour. Their x-ray absorption properties are listed in Table 4.1. Figure 4.8 shows an arbitrary cross-section of a phantom sample (4% agar with 5% aluminium oxide) in the x-ray and visible spectra. Its contrast is high in both modalities, facilitating segmentation. It is also observed that the cut is uniform across the slice, making agar an ideal candidate material for phantom preparation.

4.3.2 Volume Reconstruction

This section describes various experiments designed to test the performance of different 3D shape reconstruction methods, in terms of volume recovery. Different modalities, methods and set-ups are reviewed, performed on a variety of phantom shapes.

Physical phantom slicing

The first experiment consisted of slicing and photographing an anatomical agar phantom (4% agar with 5% Al_2O_3) at 5 mm intervals, using the slicing rig. A picture of a fiducial marker of known size was also taken to establish the relationship between pixel size and real world units. To reduce the computational complexity and running time of the algorithm, images were downsampled by a factor of 4 in each direction, that is, the original 6000×4000-pixel photographs were shrunk to 1500×1000 pixels. The resulting images had a real-world measure correspondence of 0.12 mm per pixel.

Phantom cross-sections were segmented using a region growing colour thresholding algorithm. Once the binary masks were obtained, the 3D shape was reconstructed using shape-based interpolation with a cubic spline interpolant function. After runing the experiment three times, the average volume recovered was 45,732 mm^3 (standard deviation: 804 mm^3). This value corresponds to 92.9% of the volume of the virtual CAD model (ground truth), which was 48,864 mm^3. After a qualitative inspection of the overall recovered shape, it was observed that none of the shallow valleys at the top or bottom of the model were recovered. This was due to the low axial sampling frequency. Moreover, the high-frequency wrinkle on the side was underestimated and only partially recovered. Figure 4.9 shows orthogonal and perspective views of the reconstructed phantom.

To investigate whether slice thickness had an effect on volume and shape reconstruction, a similar experiment was performed. We used the same set-up, doubling the axial sampling frequency (2.5 mm slices). The same region growing segmentation routine was utilised. On this occasion, the overall shape reconstruction was better from a qualitative point of view. As shown in Figure 4.10, the shallow trough at the top of the object was recovered, although the bottom one was still missed. The high-frequency feature on the left side was successfully

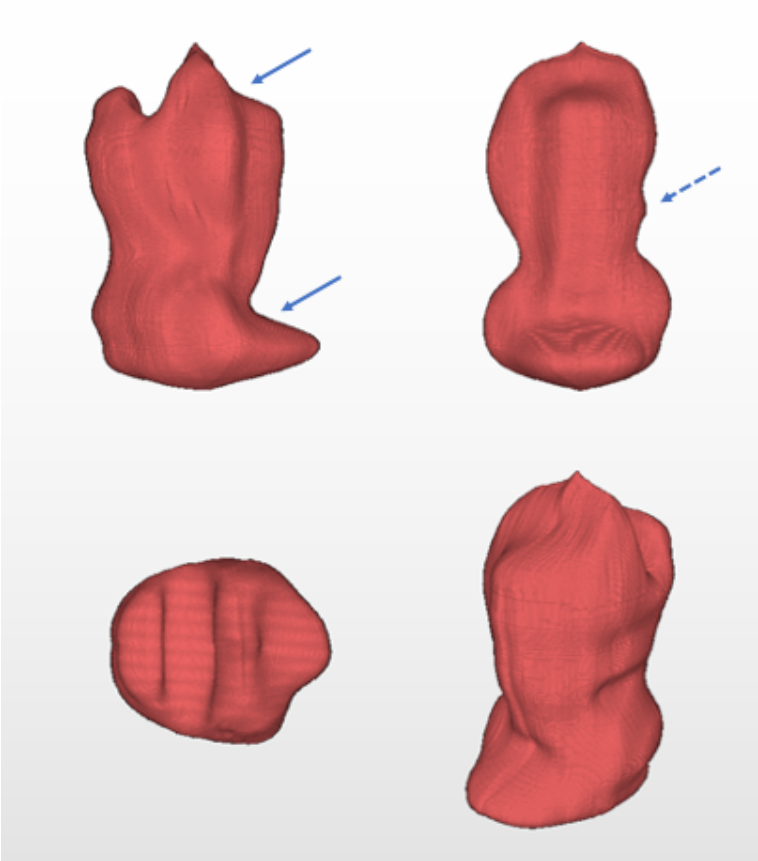


Figure 4.9: Orthogonal and perspective views of a phantom reconstructed from cross-sectional images taken every 5 mm. It can be seen that none of the shallow valleys were recovered (solid arrows), and the high-frquency feature on the left was underestimated (dashed arrow)

reconstructed. The average recovered volume after three runs was 44,950 mm^3 (standard deviation: 698 mm^3), that is, 92.0% of the ground truth volume.

The question was now whether this volume underestimation was intrinsic to the interpolation method, due to imprecisions in the segmentation routine, or a combination of both. To isolate the interpolation problem, a *virtual slicing* and reconstruction of the original CAD model was proposed. Using this method, the phantom cross-sectional areas are explicitly defined by the model, thus avoiding the need to segment the images. The procedure is described in the following section.

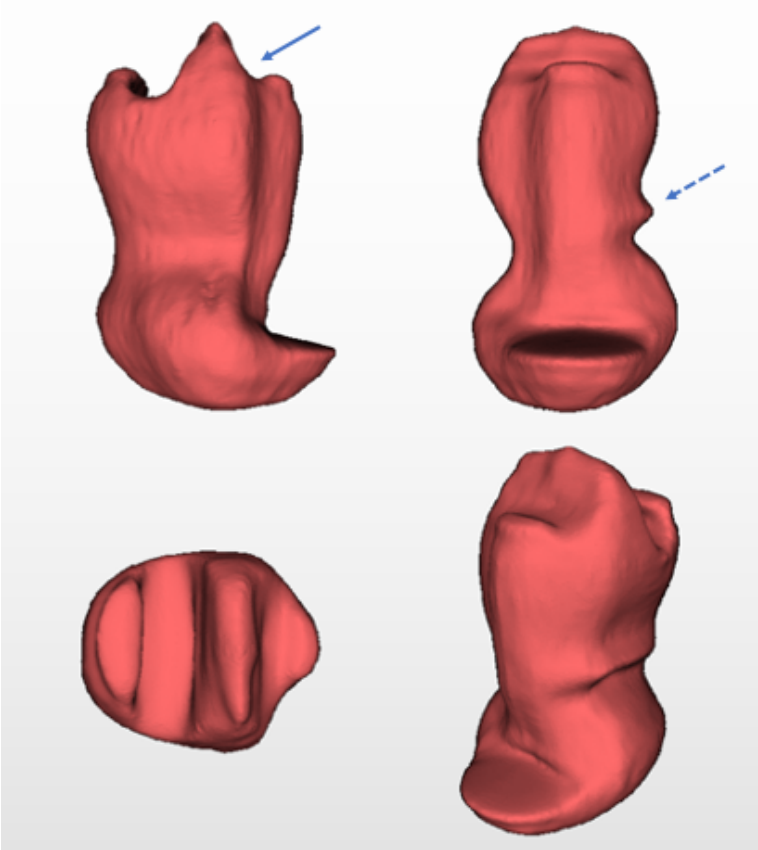


Figure 4.10: Orthogonal and perspective views of a phantom reconstructed from cross-sectional images taken every 2.5 mm. The top shallow valley was successfully recovered (solid arrow), and the high-frquency feature on the left side was correctly reproduced (dashed arrow)

Virtual phantom slicing

Virtual slicing is performed by selecting a subset of planar, parallel slices of a computer-generated 3D shape, in order to obtain its cross-sectional profiles. The advantages of this method are numerous:

- There is no need to segment the phantom area from the cross-sections, as it is explicitly defined by the virtual shape model.
- Simulations can be quickly set up and reproduced multiple times, without the need to produce a physical phantom sample every time. Multiple repetitions of an experiment will improve the statistical power of the results.
- Smaller slice thicknesses can be simulated without the mechanical constraints of a physical phantom.

During the design phase of the experiment, a new issue had to be taken into consideration: where to position the first slice in the virtual model. The default behaviour of the algorithm was to place it in line with the uppermost voxel of the shape. However, with a physical phantom, it is unfeasible to perform a cut that corresponds to the cross-section of the very first distinguishable feature situated at the phantom's apex. In addition, in a real tumour scenario, this is impossible, as the tumour is embedded within the lung tissue, which is opaque. Therefore, the randomness associated with the positioning of the first slice had to be considered, in order to perform a truthful simulation of the physical slicing process. The assumption was that the first slice could be situated anywhere between 0 (the apex of the shape) and S (one slice thickness), with uniform probability.

The simulation was set to run a parameter sweep over the slice thickness and interpolation functions, to find the optimum conditions for volume reconstruction. Slice thicknesses were set from 0.5 mm to 10 mm in 0.5 mm steps. Shape-based interpolation was performed using nearest neighbour, linear and cubic spline interpolants. The experiment was run 200 times for each combination of slice thickness and interpolation function, introducing a random offset to the first cut at each iteration.

Results are displayed in a box and whisker plot, shown in Figure 4.11. Regardless of slice thickness, nearest neighbour shows the most accurate results, with mean reconstructed volume consistent with the ground truth value. Similarly, recovered volume using cubic spline is relatively constant, although it shows a constant underestimation of 0.7% irrespective of slice thickness. In contrast, linear interpolation function shows a clear negative trend, where volume underestimation is proportional to slice frequency. Another observation is that data uncertainty increases with thickness. Interestingly, this trend is reversed for 9, 9.5 and 10 mm slices. One hypothesis to explain this behaviour is that at certain low slicing frequencies, most of the fast-changing features will be lost. Therefore, the randomness in the position of the first slice will not influence the recovered volume as much, because all fine details will be averaged regardless.

If this conjecture is true, it means that the performance of the algorithm is influenced by the particular shape of the phantom. To prove this assertion, we qualitatively analysed several of the phantom virtual cross-sections, obtained from the simulations with different slicing frequencies and random offsets. For thicknesses below 1.5 mm, all detailed features were recovered, including both

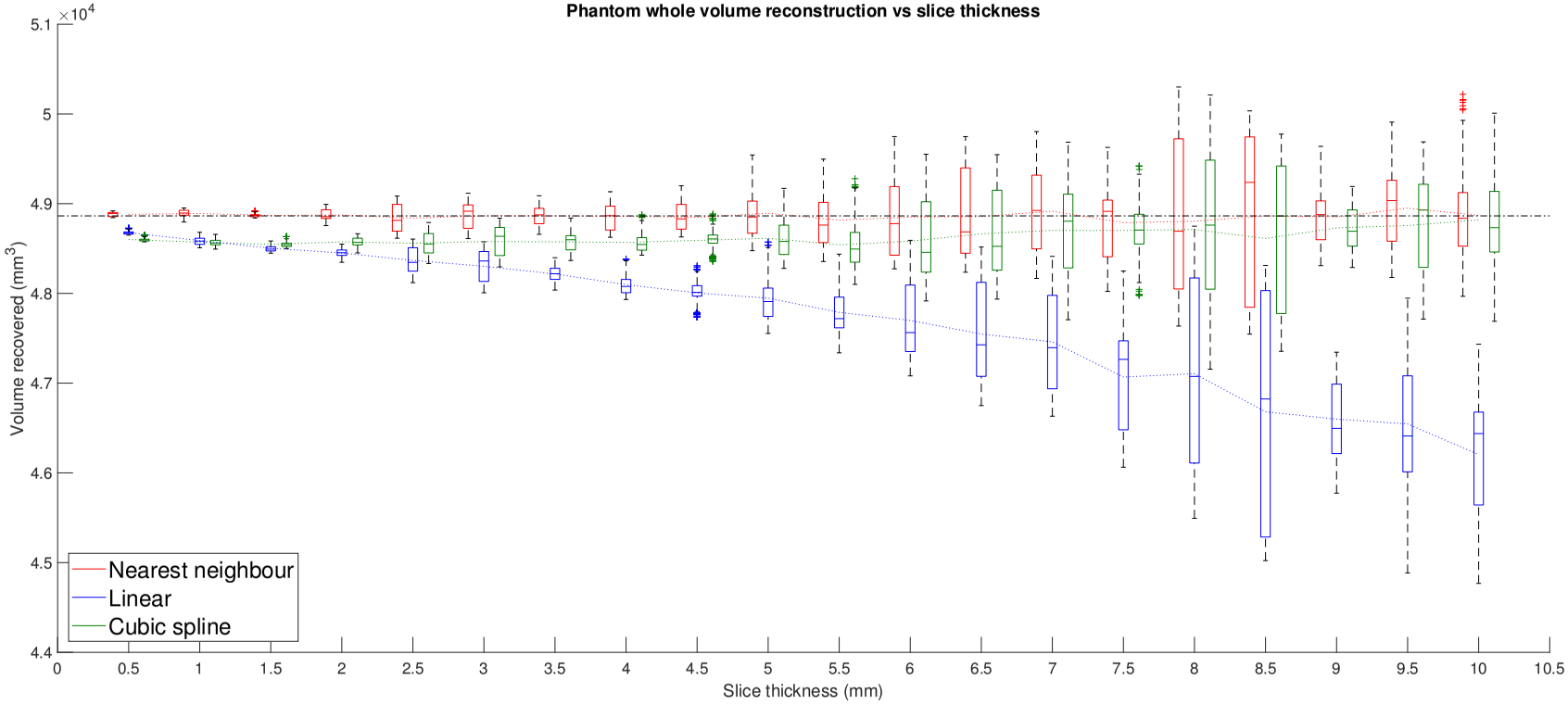


Figure 4.11: Box and whisker plot showing volume reconstruction results from virtual phantom slicing experiments. Slice thickness was set from 0.5 mm to 10 mm in 0.5 mm steps, represented on the x axis. Interpolants used were nearest neighbour (red), linear (blue) and cubic spline (green). The experiment was run 200 times for each combination of slice thickness and interpolation function, introducing a random offset to the position of the first cut. The dash-dotted horizontal black line represents the shape's ground truth volume (48,864 mm^3). Box extrema represent the 25^{th} and 75^{th} data percentiles, and the band inside the box corresponds to the median. Whiskers show the lowest and highest data points within 1.5 times the interquartile range of quartiles 1 and 3, respectively. Outliers are marked with +. Coloured dotted lines represent the mean volume obtained for each experiment

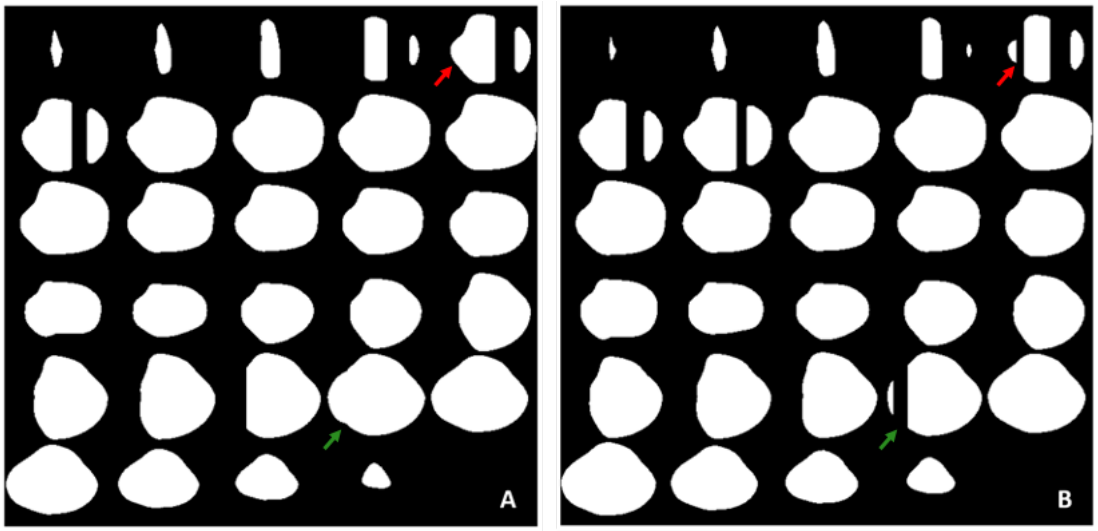


Figure 4.12: Two sets of virtual cross-sections of a phantom, sliced at 2.5 mm intervals, with different offsets for the position of the first slice: 1.9 mm (A) and 0.7 mm (B). Red and green arrows indicate the critical slices that influence the recovery of selected shape features

shallow valleys. However, for slices between 2.5 and 3.5 mm thick, the position of the first cut greatly influenced the shape recovered. This is illustrated in Figure 4.12: on set A, an offset of 1.9 mm was introduced to the first slice, and an offset of 0.7 mm was applied to set B. The transition between the fourth and fifth cross-sections in the first row of set A (red arrow) indicates that the shallow valley, situated at the top of the phantom, is missed. However, observing the same slices on set B, it can be seen that the feature is successfully recovered, as the spacing between the left and central regions is captured. Similarly, the fourth cross-section in row four shows how the bottom shallow valley is missed in A (green arrow), but reconstructed in B. Finally, for slices over 7.5 mm, the profile of the cuts was very similar at each iteration, only capturing the deep valley situated at the top of the phantom. All other detailed features present along the slicing direction were smoothed out.

These results suggested that the size of the features influence volume recovery. Also, the constant underestimation shown in spline reconstruction and the negative trend proportional to slice thickness observed with linear interpolation needed further study. To obtain more proof, the experiment was repeated on different object shapes.

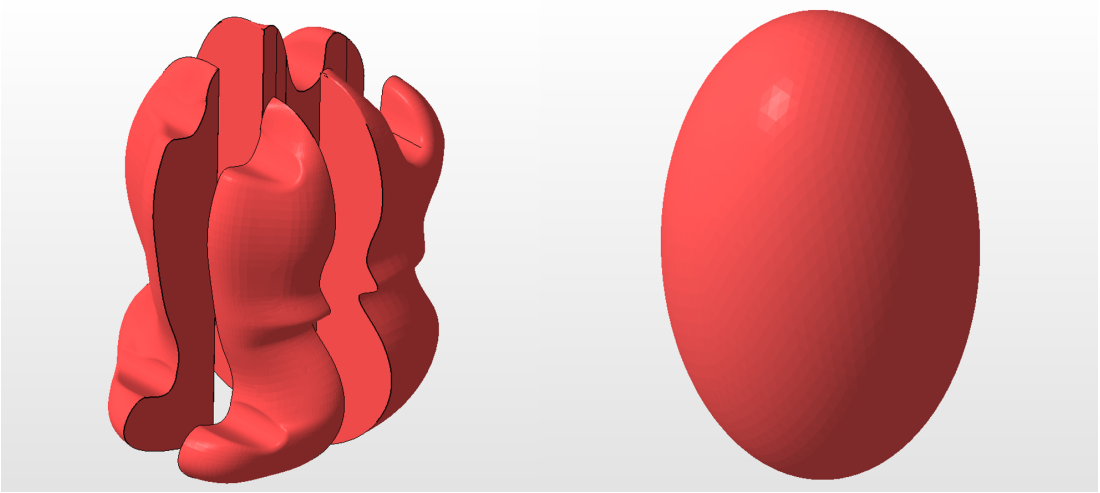


Figure 4.13: Anatomical phantom split into its four quadrants (left) and ovoid shape (right), used to validate the results obtained in the initial virtual slicing experiments

Virtual slicing on different shapes

Ideally, a number of different shapes would be designed in CAD software, virtually sliced and reconstructed. However, building a virtual object with a set of particular features is a lengthy process. A compromise was reached by splitting the original, anatomical shape into its four quadrants, and using each segment as an independent object. A simple ovoid shape was also used to evaluate the algorithm performance on smooth, regular shapes. These are displayed in Figure 4.13. Following the same protocol used in the previous experiments, all five new shapes were virtually sliced, with thicknesses ranging from 0.5 mm to 10 mm. Nearest neighbour, linear and cubic spline were used as interpolation functions, and a random offset was introduced to the position of the first cut. The script was run 200 times for each configuration, on all five objects. Phantom quadrants are labelled RR (rear right), RL (rear left), FR (front right) and FL (front left). Results are displayed in box and whisker plots, shown in Figures 4.14 (RR), 4.15 (RL), 4.16 (FR), 4.17 (FL) and 4.18 (ovoid).

All five shapes presented some common trends. Nearest neighbour appeared to be the most accurate method, with average recovered volume closest to ground truth values for any slice size. Linear interpolator showed a negative linear trend, with volume underestimation proportional to slice thickness. Cubic spline produced a constant volume undercalculation of 1% to 2% with respect to the ground truth volume, irrespective of slicing frequency. Moreover, all models presented very little data spread up to 2.5 mm thickness. In general, none of the five phantoms

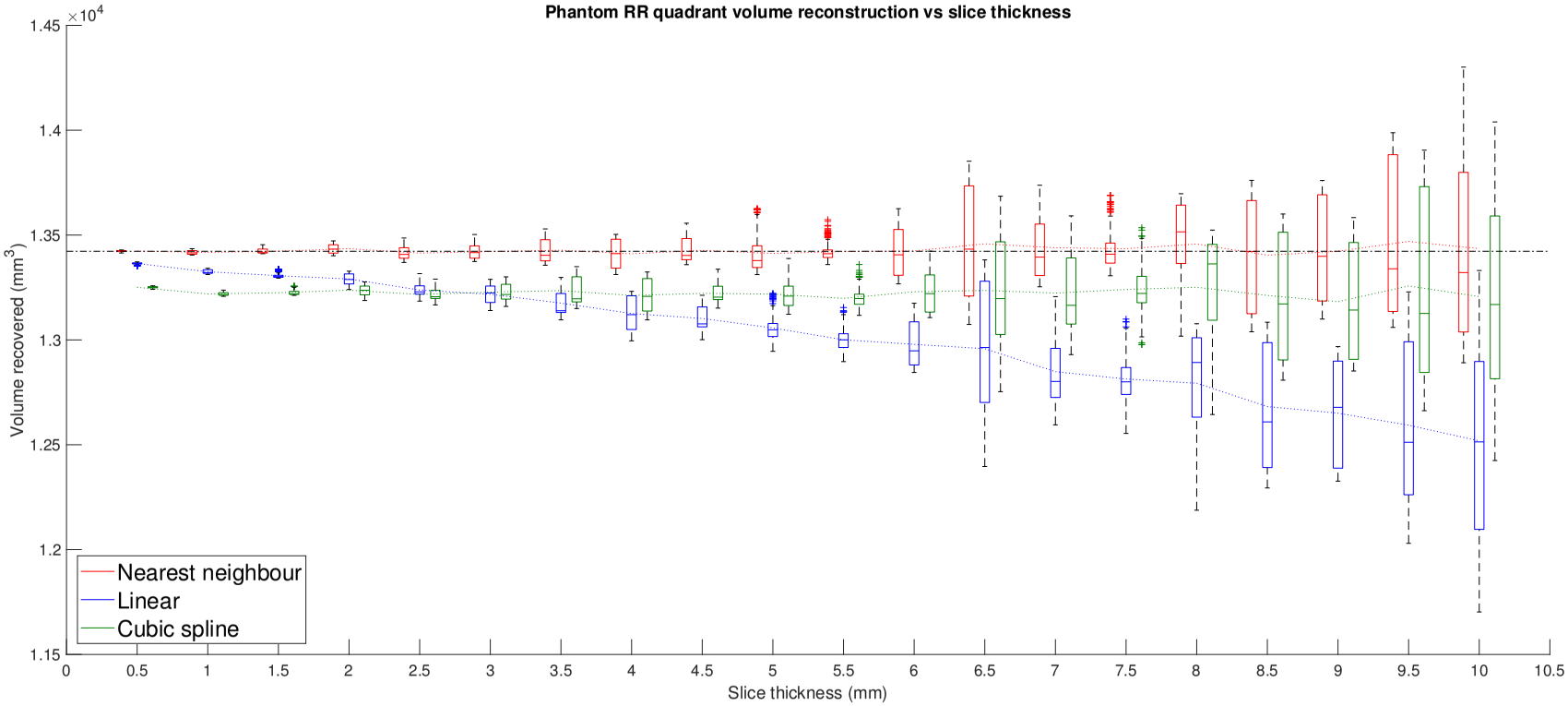


Figure 4.14: Box and whisker plot showing volume reconstruction results from virtual RR quadrant slicing experiments. Slice thickness was set from 0.5 mm to 10 mm in 0.5 mm steps, represented on the x axis. Interpolants used were nearest neighbour (red), linear (blue) and cubic spline (green). The experiment was run 200 times for each combination of slice thickness and interpolation function, introducing a random offset to the position of the first cut. The dash-dotted horizontal black line represents the shape's ground truth volume (13,423 mm^3). Box extrema represent the 25^{th} and 75^{th} data percentiles, and the band inside the box corresponds to the median. Whiskers show the lowest and highest data points within 1.5 times the interquartile range of quartiles 1 and 3, respectively. Outliers are marked with +. Coloured dotted lines represent the mean volume obtained for each experiment

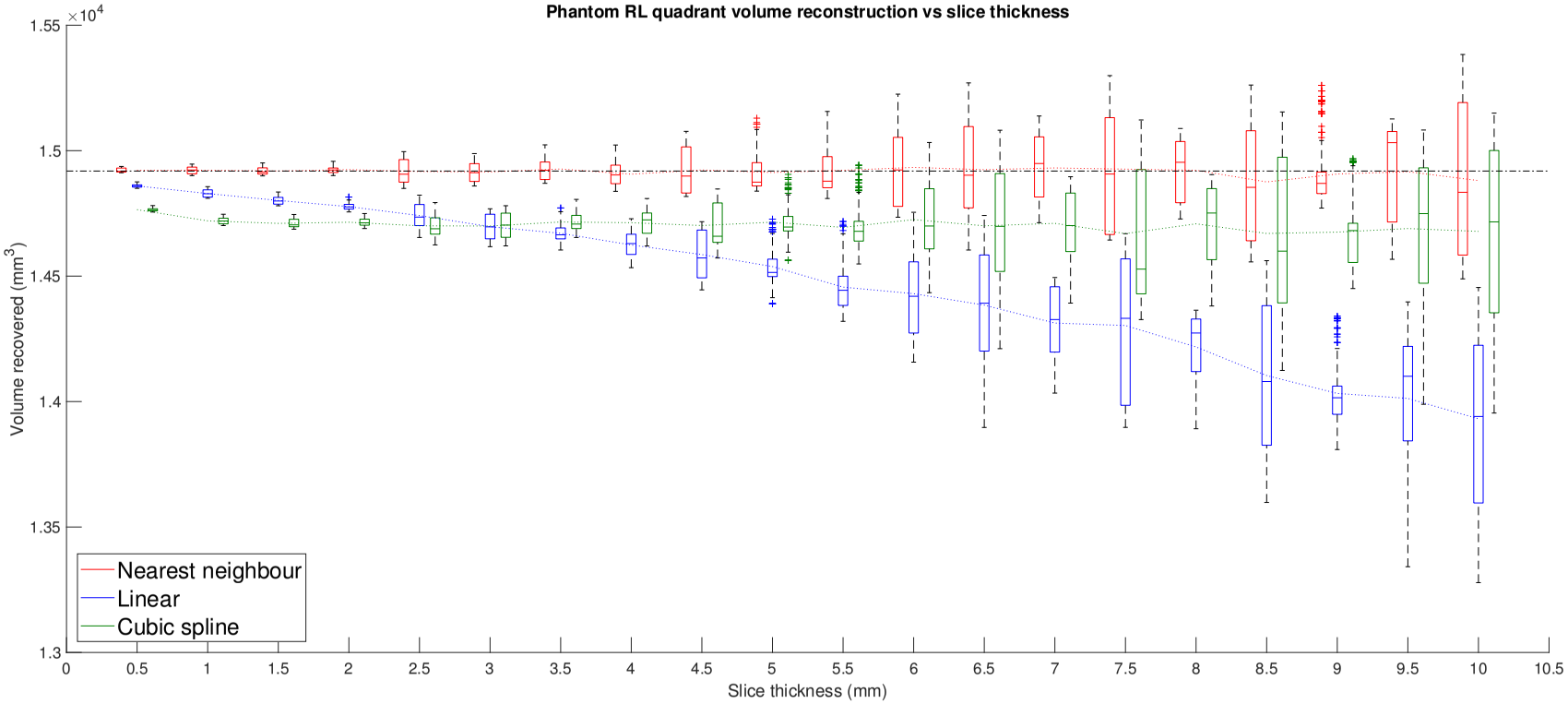


Figure 4.15: Box and whisker plot showing volume reconstruction results from virtual RL quadrant slicing experiments. Slice thickness was set from 0.5 mm to 10 mm in 0.5 mm steps, represented on the x axis. Interpolants used were nearest neighbour (red), linear (blue) and cubic spline (green). The experiment was run 200 times for each combination of slice thickness and interpolation function, introducing a random offset to the position of the first cut. The dash-dotted horizontal black line represents the shape's ground truth volume (14,919 mm^3). Box extrema represent the 25^{th} and 75^{th} data percentiles, and the band inside the box corresponds to the median. Whiskers show the lowest and highest data points within 1.5 times the interquartile range of quartiles 1 and 3, respectively. Outliers are marked with +. Coloured dotted lines represent the mean volume obtained for each experiment

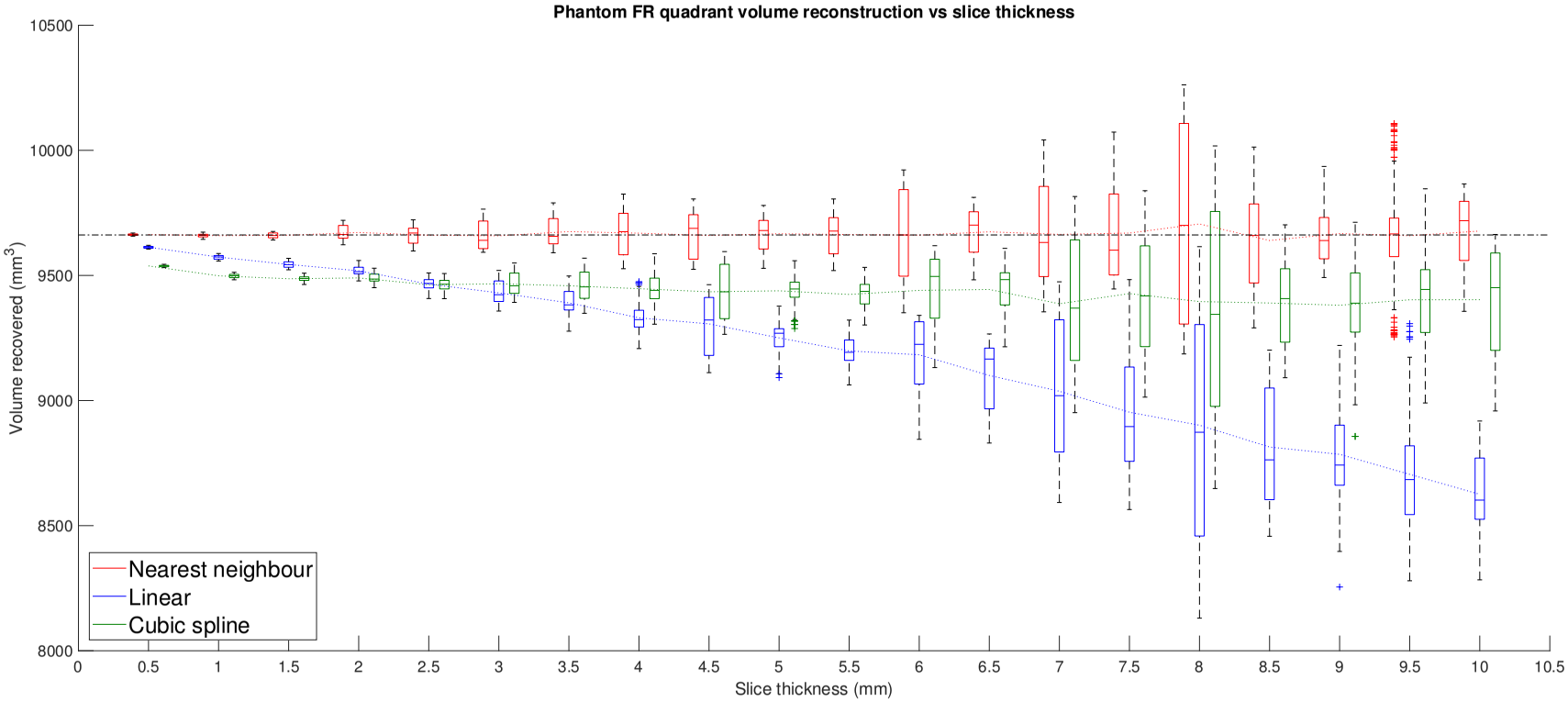


Figure 4.16: Box and whisker plot showing volume reconstruction results from virtual FR quadrant slicing experiments. Slice thickness was set from 0.5 mm to 10 mm in 0.5 mm steps, represented on the x axis. Interpolants used were nearest neighbour (red), linear (blue) and cubic spline (green). The experiment was run 200 times for each combination of slice thickness and interpolation function, introducing a random offset to the position of the first cut. The dash-dotted horizontal black line represents the shape's ground truth volume (9,662 mm^3). Box extrema represent the 25^{th} and 75^{th} data percentiles, and the band inside the box corresponds to the median. Whiskers show the lowest and highest data points within 1.5 times the interquartile range of quartiles 1 and 3, respectively. Outliers are marked with +. Coloured dotted lines represent the mean volume obtained for each experiment

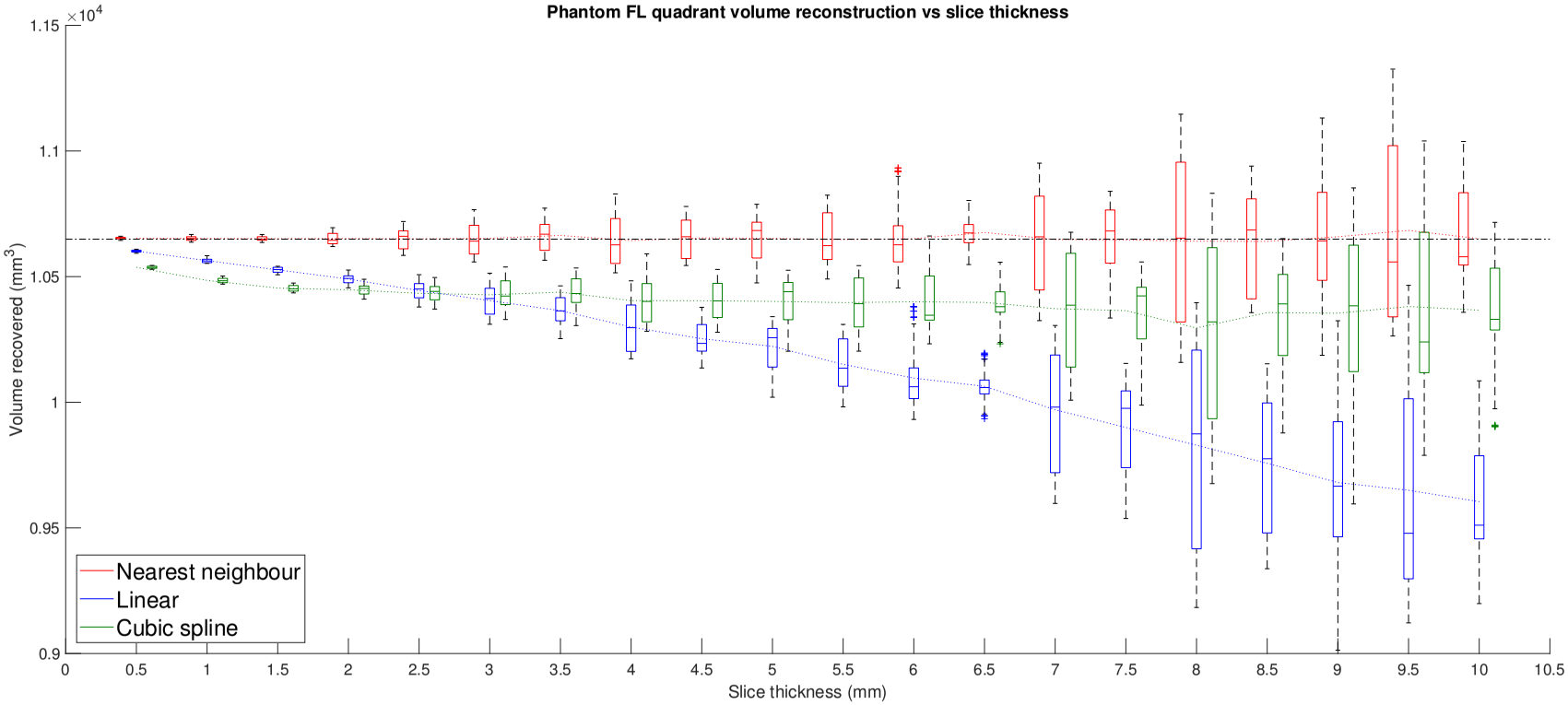


Figure 4.17: Box and whisker plot showing volume reconstruction results from virtual FL quadrant slicing experiments. Slice thickness was set from 0.5 mm to 10 mm in 0.5 mm steps, represented on the x axis. Interpolants used were nearest neighbour (red), linear (blue) and cubic spline (green). The experiment was run 200 times for each combination of slice thickness and interpolation function, introducing a random offset to the position of the first cut. The dash-dotted horizontal black line represents the shape's ground truth volume (10,649 mm^3). Box extrema represent the 25^{th} and 75^{th} data percentiles, and the band inside the box corresponds to the median. Whiskers show the lowest and highest data points within 1.5 times the interquartile range of quartiles 1 and 3, respectively. Outliers are marked with +. Coloured dotted lines represent the mean volume obtained for each experiment

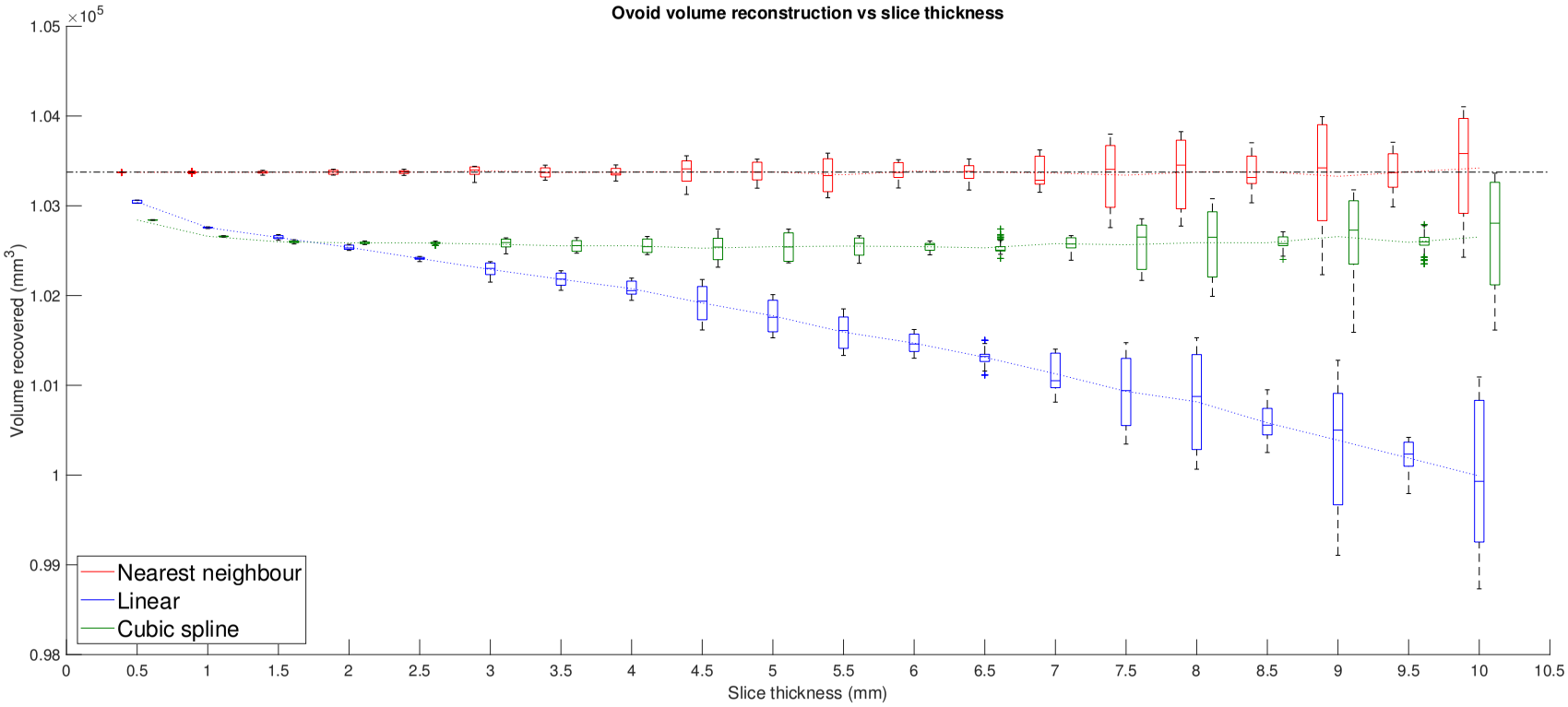


Figure 4.18: Box and whisker plot showing volume reconstruction results from virtual ovoid slicing experiments. Slice thickness was set from 0.5 mm to 10 mm in 0.5 mm steps, represented on the x axis. Interpolants used were nearest neighbour (red), linear (blue) and cubic spline (green). The experiment was run 200 times for each combination of slice thickness and interpolation function, introducing a random offset to the position of the first cut. The dash-dotted horizontal black line represents the shape's ground truth volume (103,375 mm^3). Box extrema represent the 25^{th} and 75^{th} data percentiles, and the band inside the box corresponds to the median. Whiskers show the lowest and highest data points within 1.5 times the interquartile range of quartiles 1 and 3, respectively. Outliers are marked with +. Coloured dotted lines represent the mean volume obtained for each experiment

exhibited a clear linear relationship between data spread and slice thickness. For instance, quadrant FL (Figure 4.17) showed very precise measurements for 6.5 mm thickness. This fact contrasts with the comparatively larger spread of 5.5 mm, 7 mm or 8 mm slices. Also, the ovoid shape (Figure 4.18) produced very consistent estimates for 8.5 mm and 9.5 mm slices, compared to their adjacent thickness values.

In addition to their statistical properties, we looked at the individual values obtained for each realisation of the experiment, to analyse the actual distribution of data points. One interesting observation was that, for certain shapes and slice frequency, the 200 points showed a tendency to group in clusters, rather than being uniformly distributed. In other words, for a range of offset values, the reconstructed volume was very similar. To further investigate, we plotted the volume figures against the offset applied to the first cut. The trends suggested that there is a strong correlation between slice thickness, offset and phantom feature size.

To better illustrate these findings and understand the behaviour of the experiments, a real example is analysed. In this case, we used quadrant FL, sliced at 9 mm intervals and reconstructed using linear interpolation. The first cut was randomly placed between 0 and 9 mm (one slice thickness), and the experiment was repeated 200 times. Figure 4.19a contains a graph that plots the offset applied to the first cut against reconstructed volume. It shows three critical points, marked with arrows, which are related to the position and size of the phantom's smaller details. Initially, the volume increased linearly, mainly due to the omission of the top shallow dip and consequent shape overestimation. At a 4 mm offset, two main events produced the sudden decrease in volume. Firstly, the side wrinkle is completely missed, as cuts 4 and 5 fall on either side of the protuberance. Secondly, the bottom protrusion is heavily underestimated due to the particular positioning of cutting planes 7 and 8, as shown in Figure 4.19b. At a 5.6 mm offset, the top dip and side wrinkle are roughly averaged out. However, as shown in Figure 4.19c, cutting planes 6 and 7 just miss the shallow feature at the bottom, causing major overestimation. As the offset increases, the side wrinkle is similarly exaggerated, which causes the overall estimated volume to increase. Finally, at a 7.3 mm offset, the bottom protrusion stabilises and the top feature starts to shrink, as cutting plane 1 encounters the top depression.

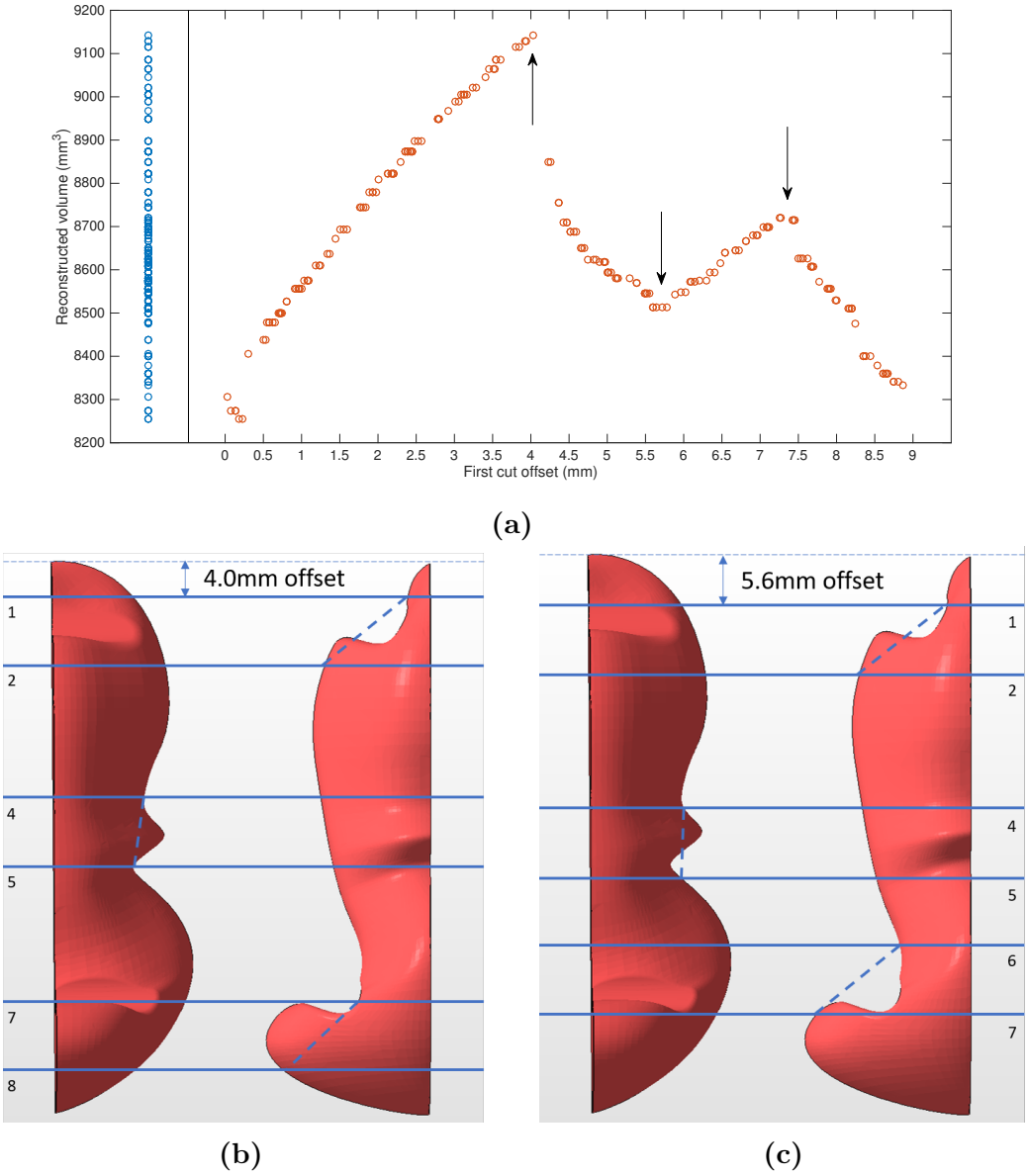


Figure 4.19: Virtual slicing example showing the influence of first cut offset over reconstructed volume. Shape used is quadrant FL, sliced at 9 mm intervals and reconstructed using linear interpolation. The experiment was run 200 times, introducing a random offset between 0 and 9 mm at every iteration. (a) Stem plot showing the volume recovered at each run (y axis) against the offset applied (x axis). Arrows indicate inflection points at 4.0 mm, 5.6 mm and 7.3 mm offset values. (b) Frontal and lateral views of the shape. Continuous lines indicate the position of the cutting planes at 4 mm offset. Dashed lines represent linear interpolation between planes. (c) Frontal and lateral views of the shape. Continuous lines indicate the position of the cutting planes at 5.6 mm offset. Dashed lines represent linear interpolation between planes

These experiments demonstrate that slicing frequency and offset greatly influence the final reconstruction, both in terms of shape and volume. Ideally, slice thickness should be half the size of the smallest feature, as described in Section 3.3.1. Moreover, the choice of interpolation function influences the overall volume reconstruction, with nearest neighbour being the most accurate method regardless of slice size. The effect of the interpolant on shape recovery is discussed in the following section.

4.3.3 Shape Analysis

So far, experimental analysis has focused solely on the evaluation of overall volume recovery. However, shape and volume are closely related: any inaccuracies in one will affect the other. Robust shape reconstruction is also of paramount importance, especially as our tumour registration approach relies on shape features to guide the alignment process.

We performed a quantitative evaluation of the shapes obtained by aligning the reconstructed phantom to the original CAD model, following these steps:

1. The outermost layer of voxels (shell) of the reconstructed shape and the model are extracted. Both datasets are converted to point clouds, by transforming each voxel into a point coordinate.
2. Reconstructed shape is rigidly registered to the model using Iterative Closest Point algorithm.
3. For each point in the registered shape, the distance to its nearest neighbour in the model is calculated.
4. These distances are colour-coded and represented on a 3D heatmap, which helps the user visualise the areas with most distortion.

This routine was repeated for different slice thicknesses and interpolation functions. As expected, nearest neighbour produced the most deformed shape, in terms of absolute sum of distances, due to the stepped profile of the reconstructions. The dips and phantom's extrema were the features that produced the most error, regardless of the interpolator. Figure 4.20 displays a shape distortion heatmap of a reconstructed anatomical phantom, sliced at 2.5 mm intervals, and interpolated using spline (left) and nearest-neighbour (right) functions. Nearest neighbour shows a persistent, small error on every slice, consistent with the

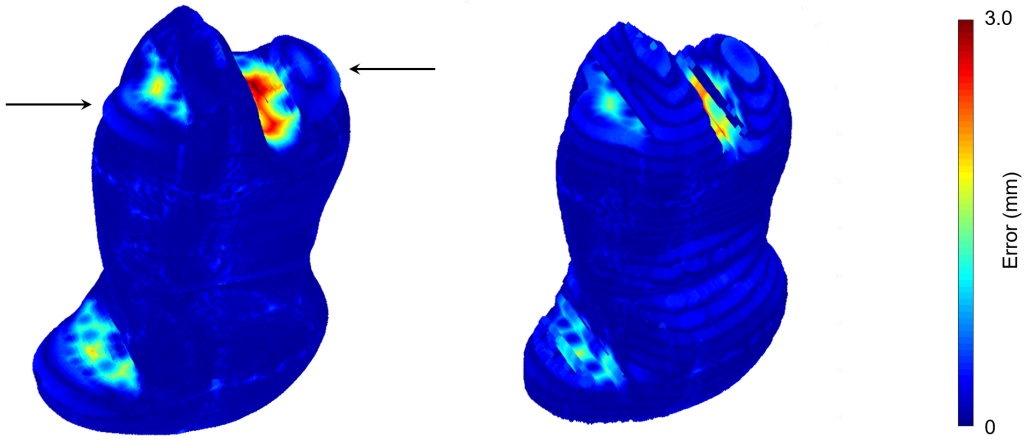


Figure 4.20: Heatmap showing shape distortion on reconstructed anatomical phantoms, sliced at 2.5 mm intervals, and interpolated using spline (left) and nearest-neighbour (right) functions. Arrows indicate "swelling" artefacts produced by spline interpolation

step-like nature of the reconstruction. On the other hand, spline interpolation produces *swelling* artefacts around some of the shape's extrema (marked with arrows). This can also be observed in Figure 4.10.

This phenomenon, characteristic of shape-based interpolation methods, is linked to some specific changes in topology between consecutive slices. When cross-sectional images are converted into distance functions, the transform is applied on each plane independently. That is, distance transform of slice i does not depend on neighbouring slices $i-1$ or $i+1$. In the particular case where a topological region $\mathcal{R}$, present in slice i, disappears in slice $i+1$, the distance transform function of slice $i+1$ does not have any prior knowledge of the existence of $\mathcal{R}$ in previous slices. This is illustrated in Figure 4.21: on the left, slice i has two regions, $\mathcal{R}$ and $\mathcal{S}$. Point $p_i = (a, b, i)$ (marked with a red +) belongs to region $\mathcal{R}$. When the distance transform is applied to the slice, point p_i will be assigned a relatively small negative number $-q$ (following the convention proposed in Section 3.3.1), because it is an inner point situated close to the region's boundary. On the right, slice $i+1$ has only one region $\mathcal{S}$. In this case, point $p_{i+1} = (a, b, i+1)$ is an outer point, placed a long distance from the closest boundary point of $\mathcal{S}$, and therefore will be assigned a large positive number d. Even though p_{i+1} may be closer to $\mathcal{R}$ (situated in the previous slice) than $\mathcal{S}$, it will be assigned a distance equal to d, because inter-slice distances are not considered. As $-q \ll d$, the interpolator will encounter a sudden change in the values of the function around

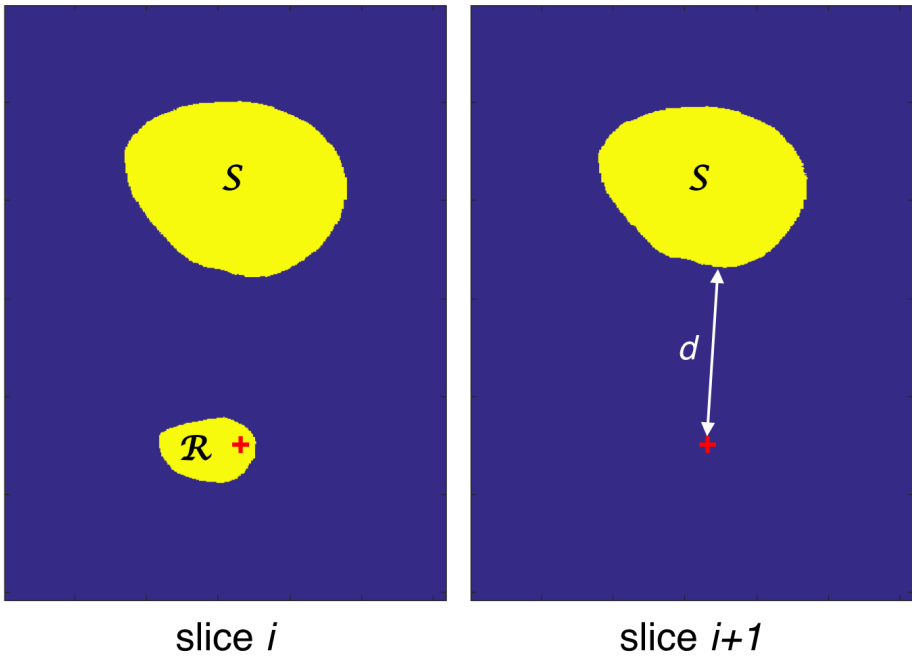


Figure 4.21: Example of topology change between consecutive slices. Region $\mathcal{R}$, present in slice i, contains point p_i (marked with a red +). When the region disappears in slice $i+1$, p_{i+1} becomes an outside point, away from region $\mathcal{S}$. White arrow indicates the minimum distance d between p_{i+1} and $\mathcal{S}$

the region that disappeared, which may affect its performance. In the case of cubic spline interpolation, the function produces ringing artefacts on each side of a discontinuity to preserve its class $\mathcal{C}^2$ differentiability condition.

To mitigate this effect, neighbouring slices shall be taken into account when computing distance transform. We investigated two approaches based on the extrapolation of end regions, to produce natural-looking, rounded extrema: linear and curvature-based (see Section 4.2.5). The performance of these two methods was evaluated on the anatomical phantom, which was virtually sliced at 2.5 mm intervals. The same initial offset was applied on both methods and cubic spline interpolator was used. Figure 4.22 shows the results obtained, with the extrapolated features marked with arrows. It can be observed that the linear extrapolation approach, on the left, produces a pointed end, which is especially salient in cases where the surrounding curvature is smooth. In contrast, the curvature-based method, on the right, displays a rounder finishing, more in line with the true shape. We obtained similar responses when the two extrapolation algorithms were used on different shapes. Therefore, the curvature-based approach was selected as the method of choice. To round off the top and bottom

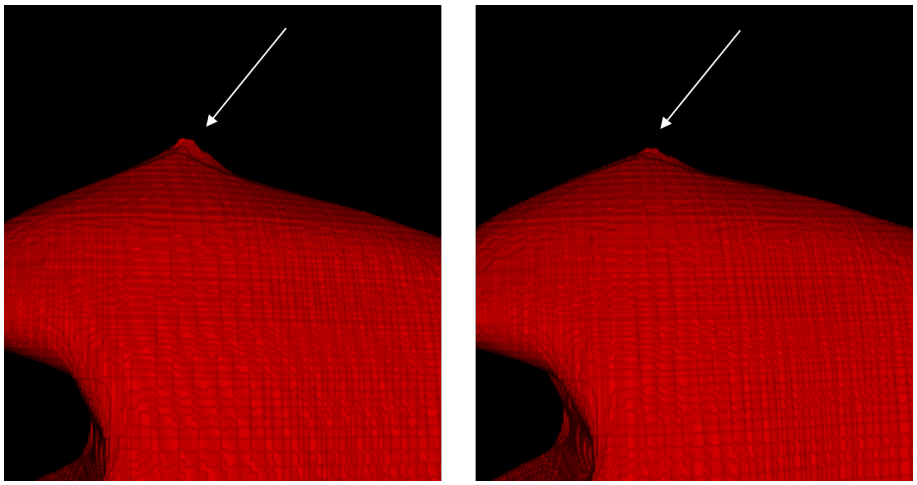


Figure 4.22: Comparison between linear (left) and curvature-based (right) methods for region extrapolation. Arrows indicate the extrapolated features

ends of the shape, this extrapolation routine was applied to the first and last slices of the phantom.

4.3.4 Phantom Reconstruction from CT Data

The next set of experiments was aimed at evaluating the shape reconstruction methods used on computed tomography (CT) data. For this purpose, an anatomical agar phantom (4% agar with 5% Al_2O_3) was produced and embedded in a cylindrical container using agar at 3% concentration. The sample was scanned in a CT scanner* at the Beatson West of Scotland Cancer Centre (Glasgow, UK). CT slice thickness was set at 2.5 mm. Next, for each slice, a consultant oncologist (S.H.) delineated the contour of the phantom on a commercial radiotherapy planning software tool†. Figure 4.23 shows the transverse, saggital and coronal planes of the specimen's CT scan, as well as a 3D representation of the reconstructed volume. The purple areas indicate the segmentation performed by the consultant.

The resulting shape, automatically reconstructed by the treatment planning software tool, presented an overall volume of 55,590 mm^3. That is a 13.7% overestimation with respect to the phantom's ground truth volume of 48,864 mm^3. This result was unexpected, especially after obtaining an 8% volume underestimation from the physical slicing experiments performed earlier (see Section 4.3.2). To

*GE Discovery CT590, GE Healthcare, Chicago (US)
†Eclipse^{TM}Treatment Planning System, Varian Medical Systems, Palo Alto (US)

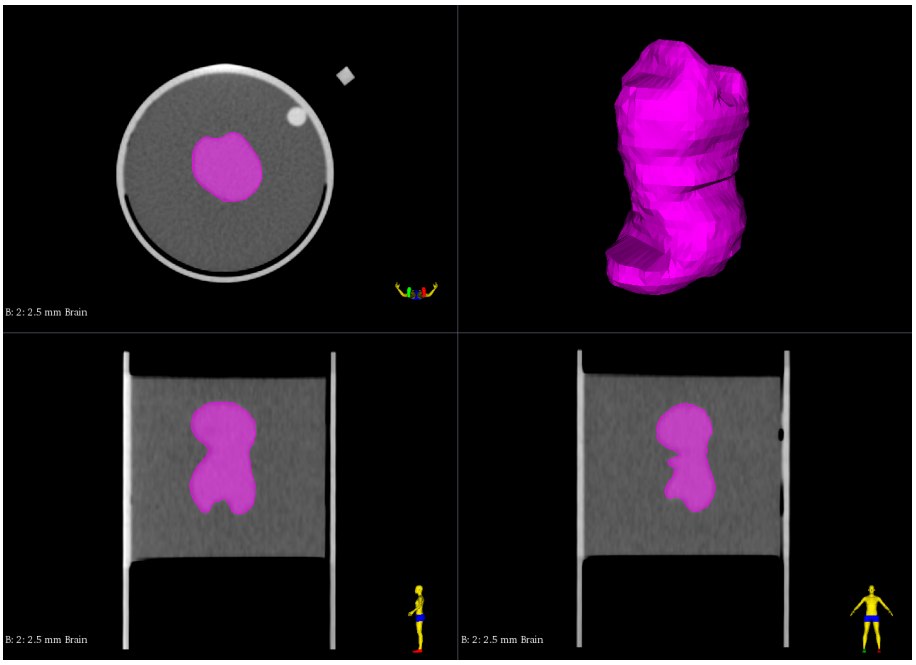


Figure 4.23: Transverse, saggital and coronal views of the CT-scanned phantom. Purple areas show the segmentation performed by the consultant. On the top-right corner, a 3D representation of the reconstructed volume is displayed

understand the origin of these discrepancies, two possible sources of error were identified and analysed: low precision of the delineated contours, and inconsistency between theoretical and actual phantom volume.

Contour precision

To evaluate the first source of error, the consultant oncologist (S.H.) was asked to review the segmentation several days after. This time, the visualisation window of the scan was modified to highlight the phantom structure and reduce background noise. The delineation was also checked independently by a consultant pathologist (C.D.). They both ascertained that the contours were as close to the phantom body as they would be in a real cancer patient situation, following standard practice. Consequently, it was agreed that shrinking the volume would be non-realistic.

Their delineated transverse cross-sections were compared against the output of our region-growing segmentation algorithm, P. A qualitative inspection revealed that the consultants' contours included, in general, an extra 1-voxel-thick layer around P. To quantify the effect of this slight increase in diameter, we calculated

how much volume an additional shell of voxels H would add to region P:

$$H = \Big|(P \oplus \mathbf{1}_3) - P\Big| * V \tag{4.7}$$

where $|\cdot|$ is the notation for cardinality, $\oplus$ indicates the morphological dilation operation, $\mathbf{1}_3$ is a 3-by-3 matrix of ones, and V is the voxel size. That is, the original phantom P is dilated by one pixel all around. Next, P is subtracted from its dilated version. The voxels of the resulting shell are counted and multiplied by the voxel volume. For a voxel size of $2.5 \times 0.98 \times 0.98\text{mm}^3$, the resulting shell H was 9,047 mm^3. Therefore, the 6,726 mm^3 overestimation in the consultant's phantom delineation was within tolerance.

Two conclusions can be drawn from these results: firstly, contours traced by clinical experts will tend to slightly overestimate the actual specimen area on CT scans. Secondly, due to the relatively large voxel size in diagnostic CT scanners, small delineation errors (i.e. inclusion versus exclusion of a voxel) have the potential to greatly influence the final overall volume. Therefore, manual contouring is one of the highest sources of error. A word of caution: the phrase *manual contouring as a major source of error* needs to be interpreted in the context of phantom segmentation for the purpose of this study. By no means does it imply that manual tumour segmentation, performed by trained medical staff in a real clinical scenario, is overestimated or indeed erroneous.

To further study the relationship between specimen size and overestimation error percentage, a CT calibration phantom (described in Section 4.2.4) was scanned, and three of its water-filled spheres (12, 24 and 39 mm in diameter) were segmented by the consultant oncologist using the same radiotherapy planning software. Figure 4.24 displays a transverse slice with the segmented cross-sections in different colours. The cyan region corresponds to the 12 mm sphere segmentation, red to 24 mm, and green to 39 mm. A quick visual inspection revealed that the drawn contours extended slightly beyond the solid area of the phantom, in line with our previous findings. Results obtained from these manual segmentations are shown in Table 4.2.

As expected, the overestimation percentage is larger for smaller volumes, and shows a tendency to stabilise with increasing sample size. Compared to the anatomical phantom shape, spheres present a lower overestimation ratio for similar volumes. For example, the 31,059 mm^3 sphere displays a 6.2% magnification, whereas the anatomical phantom, at 48,864 mm^3, shows a 13.7% enlargement.

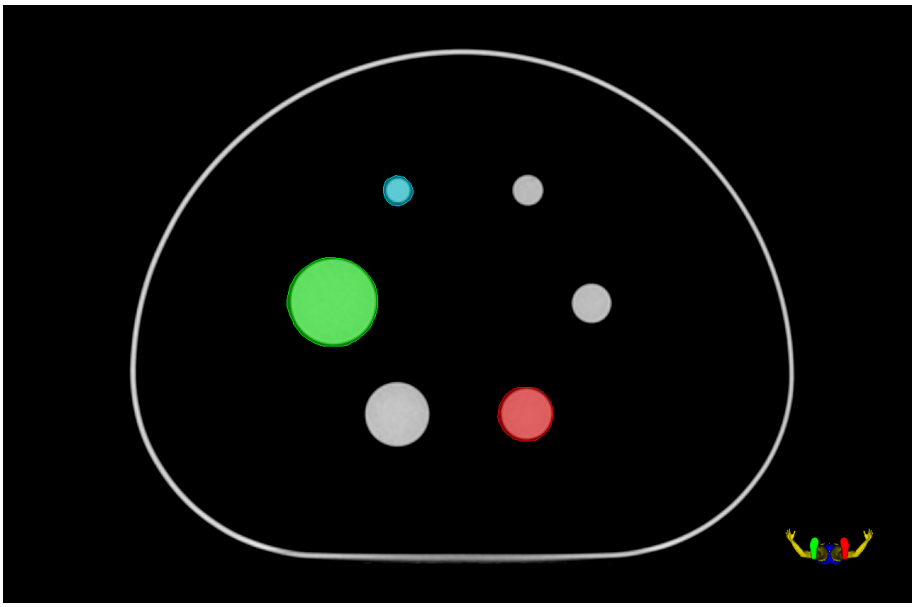


Figure 4.24: Transverse slice showing the segmentation performed by the consultant oncologist on a CT calibration phantom. Cyan, red and green regions correspond to the segmentation of the 12 mm, 24 mm and 39 mm-diametre spheres, respectively

Diameter	Manual segmentation	Ground truth	Discrepancy
12 mm	1199.0 mm^3	904.8 mm^3	32.5%
24 mm	8095.3 mm^3	7238.2 mm^3	11.8%
39 mm	32975.5 mm^3	31059.4 mm^3	6.2%

Table 4.2: Results from manual segmentations on CT calibration phantom, compared against the ground truth. Discrepancy between segmented and ground truth values are expressed as a percentage. Ground truth volumes were calculated using the sphere volume formula: $V_{sphere} = 4\pi r^3/3$

One probable cause for this behaviour is that spheres have the lowest surface-area-to-volume ratio of all the shapes. To illustrate this, suppose we have a sphere S and an irregular shape P, both with equal volume v. Both shapes are delineated including a 1-voxel-thick overestimation error around their cross-sections, resulting in volumes v_S and v_P, respectively. Due to the lower area-to-volume ratio of S, the quotient v_S/v will be smaller than v_P/v, thus explaining the results obtained in our experiments.

Virtual and actual phantom volume comparison

Lastly, we designed a set of experiments to ascertain whether the theoretical CAD and actual phantom volumes were consistent. The initial approach was

to use water displacement to measure the volume of the physical phantom. A one-litre glass beaker was filled with 750ml of water, and its level marked. Next, the phantom was immersed, and the new water level recorded. The volume of the liquid displaced was calculated using the cylinder volume formula ($V_{cyl} = \pi r^2 \Delta h$), with the beaker's inner diameter ($2r$) and the increase in water level (Δh). However, the uncertainty in the different measurements, namely water level and distance between water level markings, meant that we could only capture volume changes to a 1,000 mm^3 (1 ml) precision (0.1 mm change in water level). Weighing the amount of displaced water was also explored, but it was not possible to reach the desired level of precision either.

The next strategy adopted was to use a laser scanner*, to render a precise 3D model of an actual phantom sample. The first trial was performed on a phantom made from 4% agar and 5% aluminium oxide. However, due to the glossy, uniform surface of the sample, the scanner could not capture and track any features, and the measurement procedure was unsuccessful. In the next scan, we used silicone as the material of choice, which was correctly picked up by the scanning device. To get a 360° coverage of the shape, the procedure had to be performed twice, scanning one half at a time. A special stage holder was built to facilitate image acquisition. The resulting two meshes were subsequently stitched at post-processing using VRMesh†, producing a watertight model. However, the volume measurement obtained was 70,977 mm^3, 45% higher than the theoretical model. After a visual inspection of the reconstructed shape, several scanning artefacts could be observed on its surface. The raw mesh data was stitched again using a different software package‡, and the resulting volume was 53,076 mm^3. This large discrepancy between both measurements meant that reconstructions are highly dependent on the software used, and that there is not a unique way to stitch the raw data. For this reason, added to the impracticalities of the scanning procedure, this strategy was also abandoned.

As a last resort, we took several physical measurements on the phantom mould and sample using a digital caliper, and compared them against the CAD dimensions. Measurements were repeated 3 times and averaged. Results showed that all physical distances differed by less than 0.2 mm from the virtual model. As the potential error introduced by these discrepancies was an order of magnitude

*7Series Optical 3D Scanner (v13.0.6), Dental Wings Inc., Montreal (Canada)
†VRMesh 10.2 Studio, VirtualGrid, Bellevue (US)
‡Autodesk Netfabb Standard 2017.3, Autodesk Inc., San Rafael (US)

smaller than the error introduced by any manual delineation inaccuracies, this source of uncertainty was deemed irrelevant.

4.4 Conclusions

The different experiments described in this chapter provided a wealth of knowledge about the behaviour of our reconstruction algorithms. We proved that results are highly dependent on several factors: image modality; slice thickness; interpolation method; and segmentation. Virtual slicing showed that higher separation between slices produces larger uncertainty in terms of volume recovered, also affecting the reconstruction of small features. In addition, the position of the first cutting plane was found to influence the end result.

The three interpolation functions evaluated generated different outcomes. Nearest neighbour produced the most accurate volume estimation, regardless of slice thickness or phantom shape. However, it also was the method that produced the highest shape distortion, the best being cubic spline.

It was also observed that CT data segmented by clinical experts consistently shows an overestimation margin of 1-2 voxels. This has a significant effect on the overall volume reconstructed, both in relative and absolute terms. Resolution of commonly-used diagnostic CT scanners may limit the precision of the delineations.

Chapter 5

Lung Tumour Image Registration

This chapter focuses on the registration of multimodality lung tumour images. The first section describes the rationale behind the approach followed. Materials and methods are introduced in Section 5.2, built upon the experiments performed on phantoms (see Chapter 4). Finally, results are presented and discussed in the last section.

5.1 Introduction

The registration of two images belonging to completely different modalities, namely PET/CT and miscroscopy, is not a trivial task. On the one hand, PET/CT is a volumetric dataset, of relatively low resolution, that shows the anatomical and functional characteristics of the cancerous tissue and its surroundings. On the other hand, microscopic histopathology slides are inherently 2D, with very high in-plane resolution. These images contain detailed information about the cellular make-up of the tumour. Moreover, histopathology samples undergo some deformations during pathological processing.

Due to the heterogeneity, deformities and non-coplanarity of the images, it is difficult to draw spatial correspondences directly between both datasets. Therefore, in order to provide a robust solution, the overall registration algorithm is broken down into smaller sub-tasks, as illustrated in Figure 5.1:

1. Registration of microscopic histology slides to the corresponding gross photographs of the specimen.

2. Reconstruction of a macroscopic 3D model of the tumour from the gross pathology photographs.
3. Segmentation of the tumour volume on the pre-operative PET/CT scans.
4. Registration of pathology and pre-operative virtual tumour models.
5. Application of the transformation parameters to PET/CT data, in order to match the pre-operative scans to the pathology cutting planes.

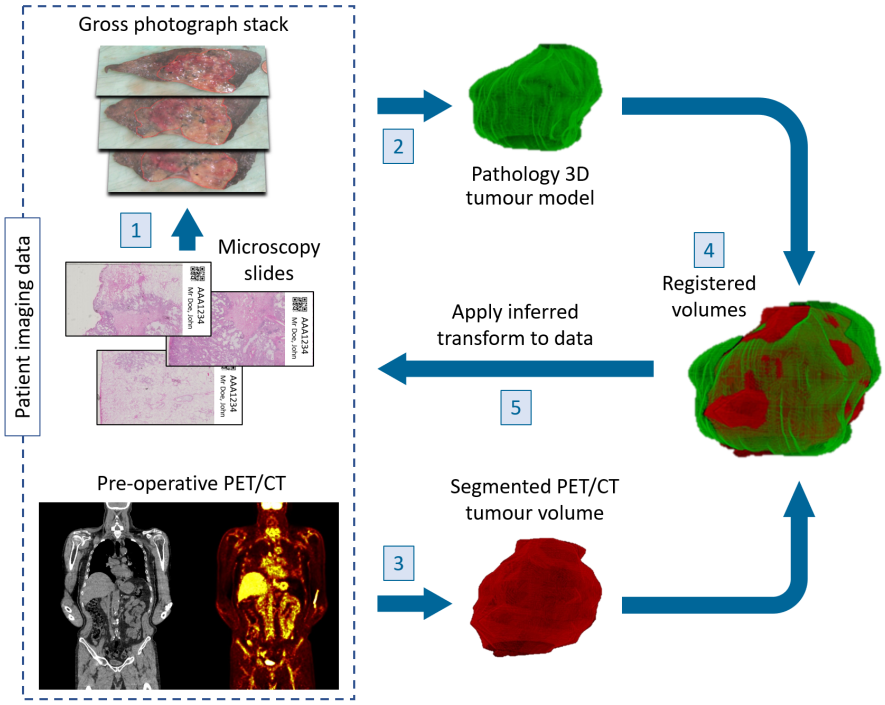


Figure 5.1: Flowchart depicting the different steps involved in the proposed lung tissue registration algorithm

Although not shown in the diagram, the use of an *ex vivo* CT scan of the resected specimen was also tested, as suggested by some authors. However, no useful information could be extracted from this and, therefore, the strategy was abandoned. Further details are provided in the following sections.

5.2 Materials and Methods

Some of the materials and methods used in this chapter, particularly tissue dissection and 3D tumour reconstruction, are identical to the ones explored in Section 4.2, and therefore are not detailed again in this section. Similarly, all image

registration algorithms mentioned later in this chapter are reviewed in Section 3.3.2.

5.2.1 Patient Recruitment

The use of human tissue for research purposes is tightly regulated by government and regulatory authorities. Additionally, ethical considerations need to be taken into account, such as seeking informed consent from the tissue donor, ensuring anonymity of the samples, or finding a balance between patient inconvenience versus potential benefits of the research. Our study protocol was examined by a NHS Research Ethics Committee*, and was given favourable ethical opinion in February 2017 (reference 17/WS/0027).

Between January and June 2018, nine patients with suspected primary non-small-cell lung cancer, referred for radical lobectomy, were recruited from the Golden Jubilee National Hospital (Clydebank, UK). Potential subjects were screened to check whether they fulfilled the study's inclusion criteria:

- adult patient referred for radical lobectomy with curative intent
- cancerous tissue concentrated in one single tumour cluster
- main volume of the tumour located within the lung, that is, not spread over the pulmonary pleura and/or airways
- lesion's major axis longer than 30 mm

Suitable patients were approached by a research nurse, who handed them a Participant Information Sheet containing the details of the research project, including the implications for the participants. The nurse was also in charge of seeking informed consent, for which the patients had to initial and sign a form, acknowledging and agreeing to the terms of the study. These two forms are included in Appendix B.

Of the nine recruited patients, one had to be excluded at a later stage, as the lesion was found to be benign after resection.

*West of Scotland REC 4, Clinical Research and Development, NHS Greater Glasgow and Clyde, Glasgow (UK)

5.2.2 Lung Tissue Processing

Recruited patients proceeded to have radical surgery with curative intent. Diseased pulmonary lobe was removed using either open or video-assisted thoracoscopic (keyhole) surgery. Immediately after resection, the specimen was transferred to the pathology laboratory. Next, the lobe was put in a formaldehyde bath for 48 hours. This process is known as *tissue fixation*, which is standard procedure in pathology for three reasons: it prevents the tissue from biological decay, by halting all ongoing biochemical processes; it eliminates any infectious agents; and it increases the stiffness of the lobe, which ultimately facilitates its slicing. Finally, the sample was rinsed with water, to remove any formaldehyde residue.

The specimen was subsequently prepared for enhanced dissection. An agar solution at 4% by weight was injected, while in liquid form, into the tissue. This procedure was carried out to inflate the lung lobe to its approximate *in vivo* size. Once set, the sample was placed into the slicing rig's tissue container, and embedded in agar at 4% concentration. The mixture was then left to solidify overnight. At this point, some of the samples were scanned *ex vivo* on a CT machine (see Section 5.2.3). Next, the specimen was dissected at 5 mm intervals using the slicing rig, as described in Section 4.2.3. Each cross-section was photographed and stored for further processing. After slicing, tumour-containing slices were segmented into approximate 2×2 cm squares by an experienced pathologist, and placed into plastic cassettes, as shown in Figure 5.2. These samples were embedded in paraffin wax, cut into 10 μm-thick sections using a microtome, and stained with hematoxylin and eosin dyes. Finally, they were placed on microscope glass slides, for ease of handling and inspection.

5.2.3 Image Acquisition

PET/CT scans

A free-breathing PET/CT scan is one of the most common tests performed on patients suspected of suffering from lung cancer. The images obtained show the anatomical and metabolic fingerprints of the cancerous lesion. Referred cases were scanned at the Beatson West of Scotland Cancer Centre (Glasgow, UK) in a

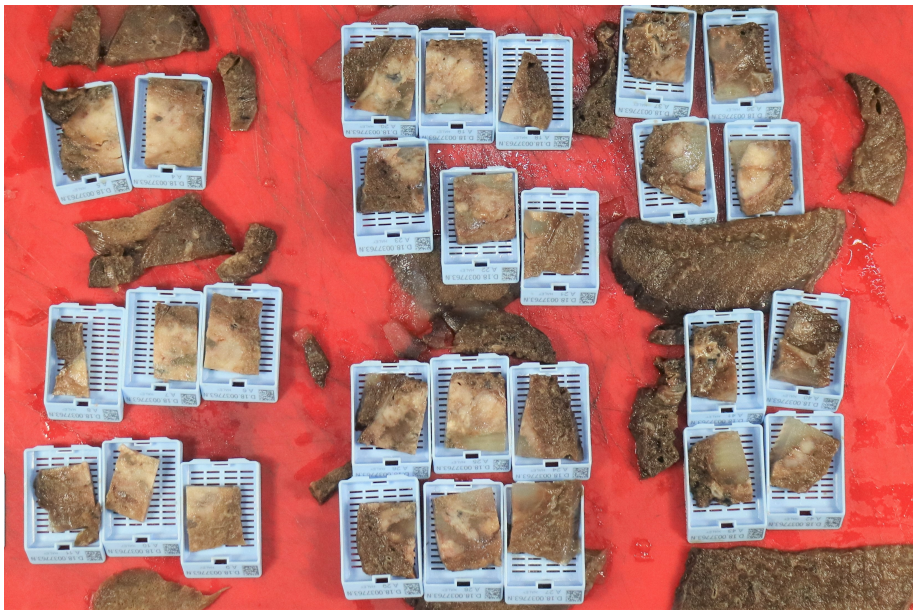


Figure 5.2: Lung tissue is segmented and placed into plastic cassettes for histology processing

combined PET/CT machine*. Slice thickness was set at 2.5 mm for the CT and 3.27 mm for the PET. After obtaining informed consent from the participants, their corresponding image sets were anonymised and downloaded from the organisation's Picture Archiving and Communication System (PACS)† in DICOM format.

An *ex vivo* CT of the excised tissue was performed immediately before pathological dissection. The specimen, which had been previously inflated and embedded in agar inside the slicing rig's tissue container, was imaged at the West of Scotland PET/CT Centre in a CT simulator scanner‡, with slice thickness set at 2 mm. Images were stored in DICOM format.

Gross pathology photographs

Gross photographs of the specimen cross-sections were taken using the same set-up described in Section 4.2.3. The digital camera was focused prior to slicing by placing a fine, high-contrast pattern at the same position as the cutting plane, and adjusting the focus accordingly. Once the calibration grid was in focus, the camera was set to manual focus mode. Also, the shutter was triggered with a

*GE Discovery 690/710, GE Healthcare, Chicago (US)
†Carestream Vue PACS, Carestream Health, Rochester (US)
‡Brilliance Big Bore CT scanner, Philips, Amsterdam (Netherlands)

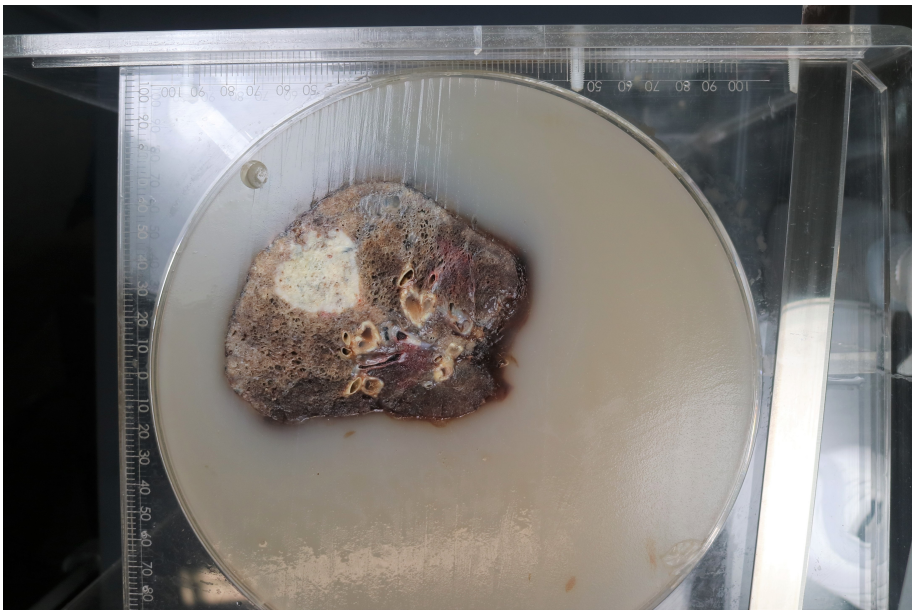


Figure 5.3: Cross-sectional gross pathology photograph taken on the slicing rig, showing the laser-engraved ruler on the top and left-hand side of the frame

remote controller, to avoid touching the camera and therefore minimising movement between takes.

Once the slice of interest was exposed, it was dabbed with ethanol to remove any traces of formaldehyde and to improve its contrast. A ruler, laser-engraved on the slicing rig frame, was also included in the scene, to help establish a relationship between pixel size and real-world units. This is shown in Figure 5.3. Images were captured and stored in both RAW and JPG formats.

Histology slides

Histology glass slides were digitised using a high-resolution slide scanner*. Images were acquired at a resolution of 96,262×162,822 pixels (40x magnification), each pixel corresponding to a 250nm×250nm square area in real-world units. Due to the extremely large size of these datasets, a subsampled (1x magnification) version was used for registration purposes. Images were anonymised and downloaded in PNG format.

*Philips Ultra Fast Scanner (v1.8.6524), Philips, Amsterdam (Netherlands)

5.2.4 Image Segmentation

Gross pathology photographs

Gross pathology images were visually inspected and segmented by an experienced consultant pathologist. For each cross-sectional photograph, the perimeter of the lesion was manually delineated using ImageJ software [89]. Resulting binary masks were subsampled by a factor of 2 in each direction and exported in PNG format. A 3D model of the tumour was generated from the masks using shape-based interpolation with a cubic spline kernel, which minimises local shape reconstruction errors, as explained in Section 4.3.3. Shape extrema were extrapolated with the curvature-based method described in Section 4.2.5.

PET/CT scans

PET/CT scans, taken prior to surgery, were exported from PACS to a commercial radiotherapy planning software* with segmentation capabilities. An experienced consultant oncologist visually inspected and manually segmented the tumour on the CT scans, using a lung window visualisation setting (-1350 to 150 Hounsfield Units). Attenuation-corrected PET scans were segmented in 3D Slicer [86] by thresholding at 50% of the tumour maximum SUV, as suggested in [90]. In both cases, resulting volumes were exported in STL format.

In three of the eight cases processed, the patients' CT volumes were independently segmented by three different experienced thoracic consultants, using the same settings detailed above.

5.3 Results and Discussion

5.3.1 Histology to Gross Pathology Registration

Moving Least Squares with affine deformation was used to register block-face gross pathology photographs of the specimens with their corresponding histology slides. Common anatomical landmarks, such as blood vessels, airways or edges, were visually identified on each pair of images. A subset of those was utilised as

*Eclipse^{TM}Treatment Planning System, Varian Medical Systems, Palo Alto (US)

control points for the algorithm, and some of the remaining were used to evaluate the quality of the registration using Target Registration Error (TRE) measure.

In order to calculate the optimal parameters to minimise registration error, three different slices, I3187(12), S3510(10) and F3203(3), were selected, each showing a characteristic feature: I3187(12) presents a branched section of the lesion, surrounded by large, hollow spaces in the lung tissue (emphysema); slice S3510(10) is mainly comprised of solid tumour, which shows up as an intense pink region in the hematoxylin and eosin-stained histology image; and case F3203(3) contains a cancerous lesion with a hollow, necrotic core. These are shown in Figure 5.4. A total of 18 common points were chosen for each pair of images. For each run of the algorithm, $n = \{4, 6, 8, 10, 12, 14\}$ landmarks were randomly selected, to act as control points, and another 4 were used for evaluation. Also, three different α values (see Equation 3.38) were tested. This parameter controls the attraction force a control point exerts over its neighbours. For each set of images, the algorithm was executed 10 times for every configuration of n and α. Histology slides were deformed to match their corresponding gross pathology photograph.

Results are shown on Table 5.1. As expected, TRE decreased as the number of common landmarks n increased. Note that the experiments were capped at 18 landmarks (14 control points and 4 for TRE calculation). The reasoning behind this choice comes from the observation of multiple histology and gross pathology pairs of images. It was noted that the vast majority of cases shared between 15 and 20 unique, common landmarks, easily indentifiable by an observer without specialist pathology training. Any higher number produces redundant points that do not contribute to the overall result. Moreover, their manual selection becomes more time-consuming.

In all three cases, points that corresponded to the corners of the histology slide were included in the landmark cohort, in order to maximise overlap between both images. Whenever, due to random selection, one or more of the corner locations were not included in the pool of control points, registration error was noticeably higher, especially in cases when these corner locations were used as evaluation points. Visual inspection of such cases showed a consistent underwarping around the missed corners. As the algorithm lacked an anchor point around the missed corner points, the deformed image was not stretched in those directions. Therefore, a larger n produced better results not only because of the increased number

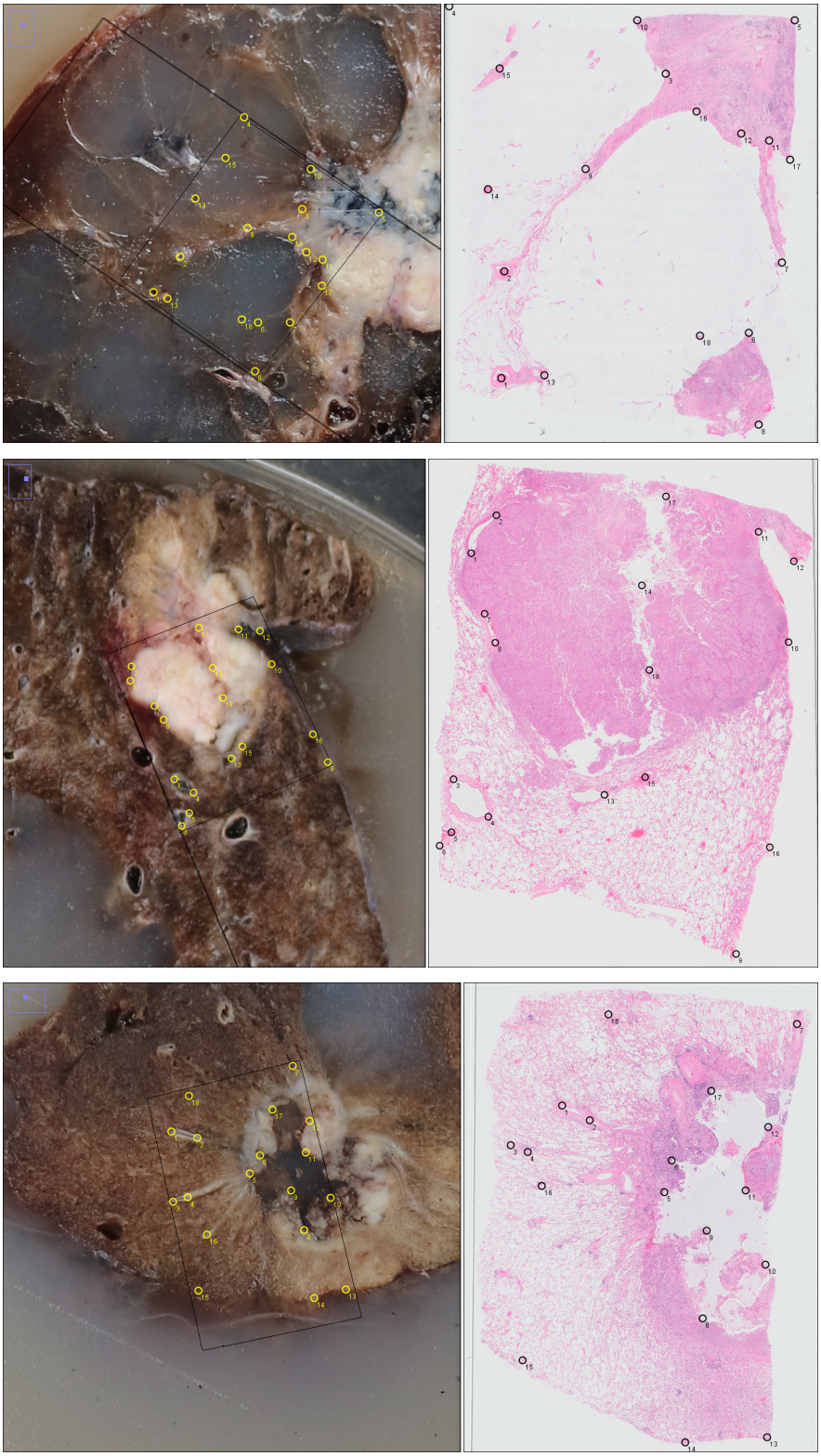


Figure 5.4: Cases used to assess histology to gross pathology registration accuracy (top: I3187(12); middle: S3510(10); bottom: F3203(3)). 18 landmarks were selected on each pair of images, marked with yellow/black circles. Numbers indicate correspondences

	TRE (mm)					
	$\alpha = \mathbf{1}$		$\alpha = \mathbf{1.5}$		$\alpha = \mathbf{2}$	
$\boldsymbol{n}$	**Mean**	**Std dev**	**Mean**	**Std dev**	**Mean**	**Std dev**
4	9.06	2.88	9.79	3.11	10.05	2.64
6	8.43	2.01	8.37	2.48	8.98	3.01
8	7.79	1.86	8.03	1.82	7.98	2.41
10	4.56	1.02	4.69	1.11	4.93	1.06
12	2.34	0.56	2.31	0.40	2.40	0.51
14	1.37	0.23	1.39	0.24	1.36	0.22

Table 5.1: Target Registration Error (TRE) mean and standard deviation for different values of α as a function of the number of landmarks n

of enforced correspondences, but also because the corner points had a greater probablility of being selected.

For smaller n, the algorithm produced more significant errors for larger α values. The higher the magnitude of α, the greater the attraction force of the landmarks closer to the point being transformed. Particularly, TRE increased when landmarks were sparsely distributed on the plane, leading to large image overstretching in those areas with no control points. However, for larger n, the choice of α had little influence on the overall registration. In such cases, overstretching was minimised due to the elevated number of spatial constraints, and consequently distortion was low.

As expected, best performance was observed for $n = 14$, when some control points were placed at the corners of the scene, independently of α. Warped histology images were overlaid to their corresponding block-face pathology photographs at 50% transparency, to allow visual comparison of both datasets. An example is shown in Figure 5.5.

5.3.2 Gross Pathology 3D Tumour Reconstruction

Pathology cross-sectional photographs were taken after the lung specimen was sliced at 5 mm-thick intervals. For each image, tumour outline was manually delineated by an experienced pathologist (CD), using ImageJ [89]. Resulting binary masks were subsampled by a factor of 2 in each direction, to reduce the computational complexity and execution time of the subsequent interpolation al-

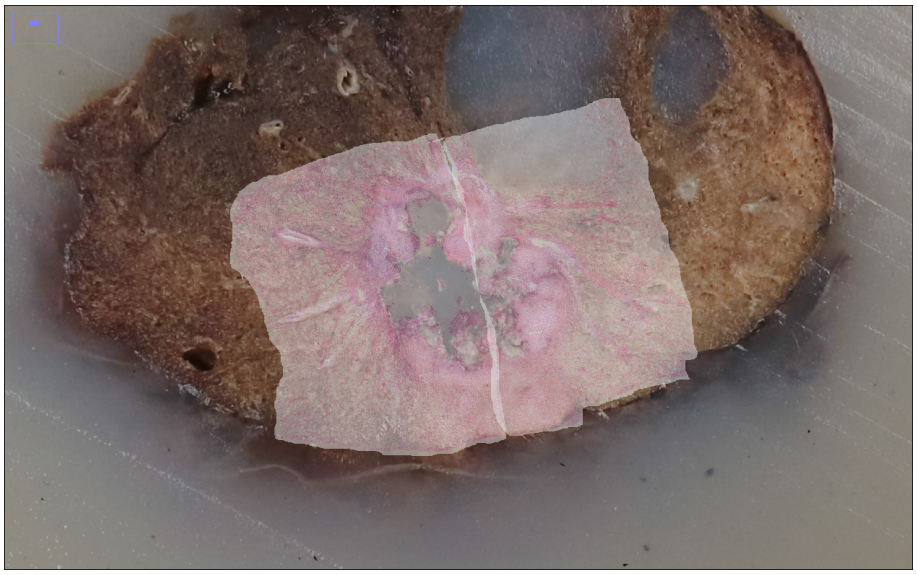


Figure 5.5: Gross pathology photograph with two registered histology slides overimposed

gorithm. Cubic spline interpolation was applied to the data, and extrema regions were smoothed with the curvature-based method described in Section 4.2.5.

An example is shown in Figure 5.6: six binary masks, corresponding to the tumour cross-sections extracted from block-face photographs (left); and 3D model obtained after interpolating the data (right). The narrow crests that can be seen at the top and bottom of the volume are created by the curvature-based smoothing algorithm, as the extrapolated data follows the gradient's rate of change to generate extrema. This behaviour can be explained by the steep slope found along one of the cross-sectional major axes, combined with the slow-changing features observed perpendicularly.

The accuracy of the resulting reconstructions could not be quantitatively evaluated, due to the lack of ground truth data. However, the protocol had been previously tested on synthetic phantoms, as discussed in Chapter 4.

5.3.3 PET/CT Tumour Segmentation

Similarly, tumour regions were segmented from PET/CT scans and imported into Matlab* and converted to matrix form. Voxel size was matched to their

*Matlab R2017b, The MathWorks Inc., Natick (US)

Figure 5.6: Binary masks corresponding to the gross pathology tumour cross-sections (left) and 3D model obtained after interpolation (right)

corresponding pathology models. For each case, PET, CT and pathology tumour volumes were compared, as shown in Table 5.2.

		Volume (mm^3)		
Patient ID	**Location**	**Pathology**	**CT**	**PET**
G3397	RL	31,779	36,920	83,851
N5398	RU	6,990	11,747	17,266
F3203	LL	34,107	45,012	91,557
J5258	LL	74,520	93,444	139,082
S3510	RU	14,233	17,215	22,474
I3187	RU	39,828	43,510	59,301
R3473	RU	13,281	13,572	18,707
S6454	RL	21,646	23,073	34,569

Table 5.2: Tumour volumes obtained from pathology, CT and PET modalities. The second column indicates the anatomical location of the lesion (RL - right lower lobe; RU - right upper lobe; LL - left lower lobe)

In line with the results discussed in Chapter 4, CT volumes were larger than their pathology counterparts. This behaviour can be explained by the standard delineation procedures followed by radiation oncologists, and also by the coarser resolution of CT scans, which hinders segmentation precision. Also, PET volumes were, on average, 60% larger than their corresponding CT segmentations, following the trend reported in [91]. This result is a consequence of the slow image acquisition procedure in PET. Breathing-induced chest movement affects the PET signal in two ways: a) metabolically active regions appear larger on the resulting images; and b) maximum SUV is underestimated. Deep-inspiration, breath-hold techniques are not feasible, due to the long acquisition times. Le-

sions situated in the lower lobes exhibit more pronounced volume mismatches between PET and CT, due to their close proximity to the diaphragm, which means they are more affected by respiratory motion [92]. One possible way of reducing these motion artefacts would be to use time-gated (4D) PET scanning protocols. In this modality, a movement tracking device is placed on the patient's chest, which allows decorrelation of the image signal from the periodic breathing-induced oscillations. However, 4D PET is not a standard diagnostic procedure for lung cancer patients.

After reviewing the results in conjunction with the clinical team, it was concluded that the PET signal obtained was too distorted. The resulting smooth volumes lacked any distinctive features for them to be registered accurately to their corresponding pathology models. For this reason, PET data was given no further consideration for the purposes of this study.

In addition, to analyse the inter-rater variability of manual tumour segmentation on CT scans, three radiation oncologists independently delineated the lesions of three patients*: J5258, S3510 and S6454. For each patient, Dice coefficient (see Equation 3.49) was used to calculate the overlap between pairs of segmentations, as shown in Table 5.3. The purpose of this experiment was to establish a registration accuracy reference value, against which to compare the subsequent pathology to CT tumour registrations. Results show that CT lesion segmentations of the same patient differ, on average, by 10%, depending on the operator. Therefore, the reciprocal of this divergence value can be used as a pseudo-upper bound when assessing multimodal registration accuracy.

Lung specimen *ex vivo* CT scan

The use of an *ex vivo* CT scan of the resected specimen as a bridging modality between pathology and pre-operative CT was investigated, following the results reported in [48–52, 54]. Specimens from patients G3397 and N5398 were CT scanned after resection. An example is shown in Figure 5.7. Resulting images were analysed by a consultant thoracic oncologist, who concluded that the tumour region could not be distinguished from its surroundings. Upon comparison with their corresponding pre-operative scans, it was observed that the absorption pro-

*It should be noted that manual segmentation of tumours is a highly time-consuming task. For this reason only three cases were evaluated. However, region overlap was computed on every CT slice for each pair of consultants, therefore figures shown in Table 5.3 represent the average over all pairwise calculations, for each patient.

	Dice coefficient			
Patient ID	**C1-C2**	**C1-C3**	**C2-C3**	**Avg.**
J5258	0.894	0.952	0.905	0.917
S3510	0.878	0.892	0.920	0.897
S6454	0.879	0.941	0.892	0.904

Table 5.3: Overlap between pairs of CT manual tumour segmentations, calculated using Dice coefficient, performed by three independent consultants (C1, C2 and C3). Column labelled C1-C2 shows the overlap between segmentations performed by consultants 1 and 2, and so on. The last column displays the average Dice coefficient for each patient

file of the *in vivo* tumours was very similar to the properties of the agar used for inflating and embedding the tissue. After discussion with a consultant pathologist, the possibility of adding a contrast agent during the inflation stage was excluded, as it could potentially change the chemical properties of the specimen and thus affect its clinical value.

The main difference between the tissue used in our study, in comparison to that of the surveyed literature, is its mechanical properties. Lung specimens need to be injected with an inflating agent, such as paraffin or agar, to prevent the tissue from collapsing. In contrast, prostate and esophagus are both rigid organs which do not need a supporting medium, and therefore their pre- and post-resection radiological properties remain very similar. For this reason, it was concluded that an *ex vivo* CT scan of the lung specimen would not provide any additional information to the registration process, and the strategy was therefore abandoned.

5.3.4 Pathology to CT Tumour Volume Registration

Reconstructed STL tumour models from pathology gross photographs and CT were imported into Matlab and converted to binary volumetric images, preserving their correspondence to real-world units. A naive pre-registration step was applied to the data, which consisted of aligning their centres of mass, to increase spatial overlap. Next, rigid registration was performed between both volumes, using gradient descent as the optimiser and sum of squared element-wise differences as metric. The pathology model was used as reference image. Optimiser parameter values used for this task are shown in Table 5.4.

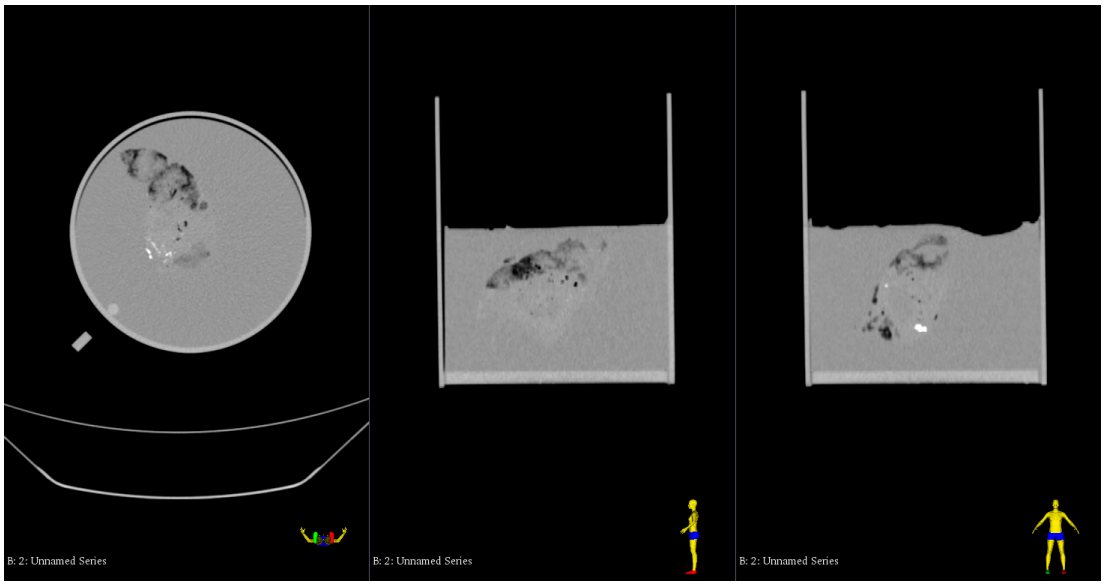


Figure 5.7: Transversal, sagittal and coronal slices of a lung specimen *ex vivo* CT scan, corresponding to patient G3397

Parameter	Value
Convergence threshold (ϵ)	1×10^{-5}
Max. iterations (n)	1000
Initial step size (γ_0)	0.1
Step relaxation factor (r)	0.75
Min. step size (γ_{min})	1×10^{-5}

Table 5.4: Gradient descent parameters used for rigid registration of pathology and CT tumour models

Registered volumes were visually inspected, and their overlap was calculated using the Dice method. Nevertheless, despite relatively large Dice coefficients (between 0.79 and 0.84), further analysis of the results uncovered potential feature mismatches in some of the cases. For instance, in patient S6454, the basal end of its pathology tumour model presented a well-defined crest. After transformation, however, this feature was aligned with the apical region of the CT model.

To further assess the quality of the registrations, common anatomical fiducials that could be visualised on both modalities were added to the models. The most easily distinguishable feature that was present in most cases was the primary and secondary bronchi, as suggested in [57]. These structures were segmented in the pathology and CT images of patient S6454, and added to their corresponding tumour volumes. Next, the models were manually aligned, using as reference the position of the bronchial segments. Although the final registration was only ap-

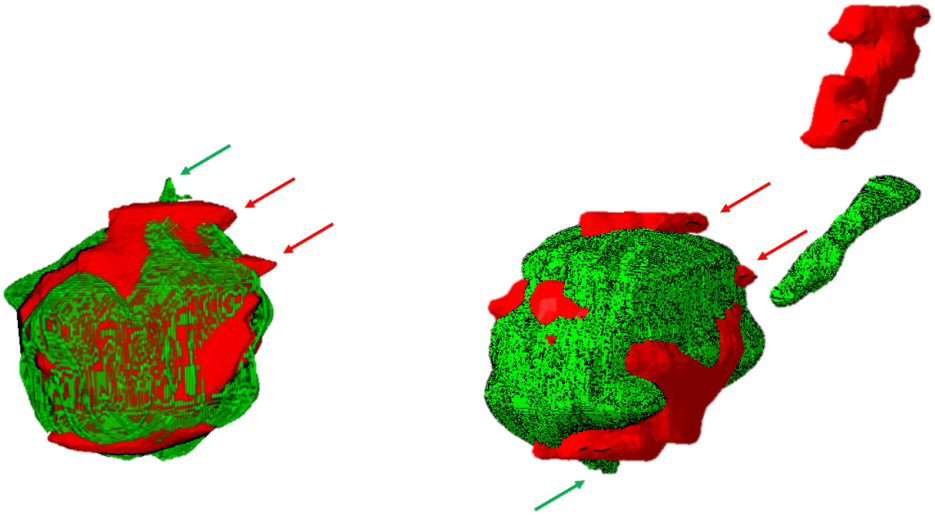


Figure 5.8: Pathology (green) to CT (red) tumour registration results for patient S6454. On the left, only the tumour shape was used to compute optimal alignment. Green arrow indicates the crest feature on the pathology model, and red arrows mark two salient features of the CT volume. On the right, bronchial segments are used to guide the transformation. The crest on the pathology model (green arrow) is aligned with the basal region of the CT shape

proximate, it was clear that the crest feature in the pathology model was matching the basal region of the CT volume, as initially expected. This is illustrated in Figure 5.8: green arrows indicate the position of the crest in the pathology model. On the left, only the tumour shape was used to compute optimal alignment. On the right, shapes were manually registered using the bronchus as spatial reference.

These findings proved that the optimisation metric used (mean square error) was not convex, with many local minima that led to sub-optimal registration results. Therefore, algorithm initialisation was found to be a key step in the process to guarantee best shape alignment, especially in those cases where the tumour models were quasi-spherical. For that purpose, bronchial regions were also included as anatomical fiducials in the lesion segmentations of two more patients: G3397 and S3510. These two cases were selected because the airway tree was clearly visible on both modalities.

Following the results published by Dura *et al.* [78], pre-registration of shapes was performed by aligning their Minimal Bounding Box (MBB) bases. Firstly, for each case, the MBB algorithm was run on a subset of points situated on the outermost layer of the volume, to reduce its computational time. Figure 5.9 shows an example of the minimal bounding boxes corresponding to the pathology

(green) and CT (red) tumour models of patient S6454, calculated over a 5% random sample of their outermost layer of voxels.

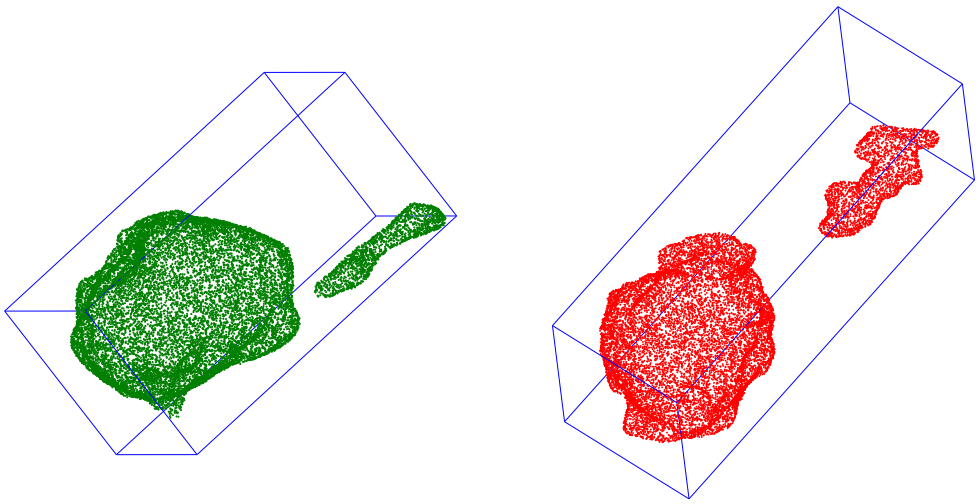


Figure 5.9: Minimal bounding boxes corresponding to the pathology (green) and CT (red) tumour models of patient S6454, calculated over a 5% random sample of their outermost layer of voxels

Secondly, for each model, the direction of its MBB longest side was arbitrarily taken as the x axis, the second longest as the y axis, and the shortest as the z axis [78]. To resolve sign ambiguity, a vector a was defined, joining the centres of mass of the tumour-only model (without the external fiducials) and the tumour-plus-bronchus model. The idea behind this reasoning was that a will always point towards the bronchus. Next, the inner product between the x axis unit vector $\hat{i}$ and a was calcuated for each direction of $\hat{i}$. The positive direction of the x axis was defined as the one giving the largest inner product $\hat{i} \cdot a$ (i.e. most closely aligned to a). The sign for axes y and z was subsequently defined using the left hand rule. Therefore, the MBB orthonormal bases for the pathology and CT tumour models were defined as $R_p = \{\hat{i}_p, \hat{j}_p, \hat{k}_p\}$ and $R_c = \{\hat{i}_c, \hat{j}_c, \hat{k}_c\}$, respectively. The relative orientation S of the pathology model basis R_p with respect to the CT model basis R_c is defined by:

$$S = R_p^{-1} R_c \tag{5.1}$$

Hence, the CT tumour model was projected onto the pathology MBB basis using Equation 5.2, where $\mathbf{p}$ is any point of the original CT dataset, and $\mathbf{p}'$ is its transformed counterpart:

$$\mathbf{p}' = S^{-1}\mathbf{p} \tag{5.2}$$

Finally, the centres of mass of the pathology and rotated CT models were aligned. Fine shape registration was subsequently performed, using the same metric and optimisation method described above. Resulting rotation matrix F was stored. It should be noted that the translational part of the registration was not considered. This is because the registration is performed on a subset of the image data that contains tumour information. However, clinicians usually inspect the whole CT volume, to visualise the lesion in its anatomical context. As the registration results were displayed side-by-side, rather than overimposed, the relative position of the CT tumour with respect to its pathological counterpart was not relevant.

Dice coefficients were calculated for the alignment results with MBB-based pre-registration, and compared to the previous approach with naive pre-registration, as shown in Table 5.5. In two of the cases, the overlap decreased with MBB-based pre-registration, although being qualitatively superior. These results confirm that Dice coefficient is not sufficient to assess registration performance in this particular problem, and therefore other metrics are needed. In this instance, visual inspection of common anatomical landmarks was used. However, mathematical, quantitative alternatives would provide more consistent results.

	Dice coefficient	
Patient ID	**Naive pre-reg.**	**MBB pre-reg.**
G3397	0.7993	0.7864
S3510	0.7881	0.7922
S6454	0.8398	0.8209

Table 5.5: Dice coefficient comparison, obtained from pathology to CT tumour model registration tasks, for different patients, using naive and MBB-based pre-registration

5.3.5 Visualisation on CT Imaging Data

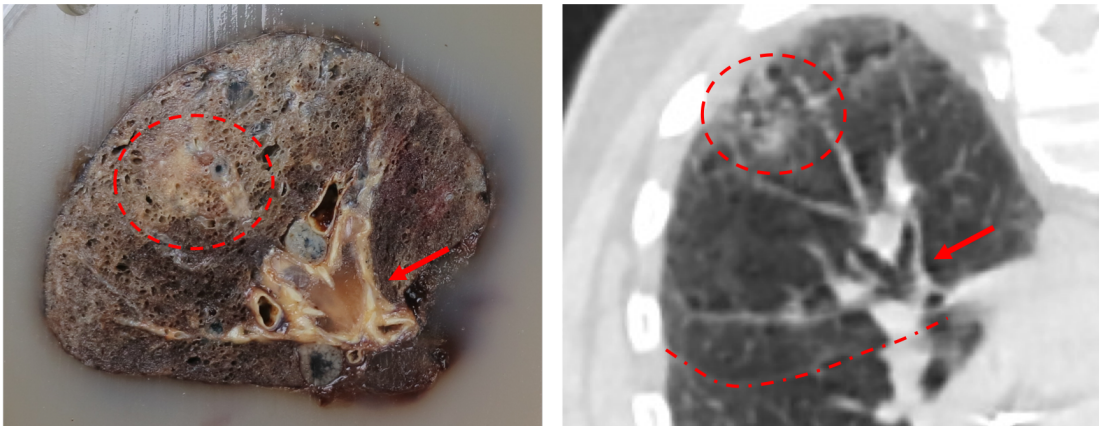
Overall transformation matrix $T = F^{-1}S^{-1}$, obtained from the registration between the pathology and CT binary tumour models, was applied to the CT imaging data using 3D Slicer [86]. Relative position along the z axis (pathology slicing direction) was manually adjusted until visual correspondence between pathology and CT slices was achieved. CT volume was resliced at 5 mm intervals, to match the slice separation of the pathology dataset. An example is shown in Figure 5.10, where two registered pairs of pathology and CT images are displayed.

Finally, using the method described in Section 5.3.1, microscopic slides were overimposed to the gross pathology photographs, thus achieving the main goal of the project: the registration of pre-operative CT scans and post-operative histopathology maps in lung cancer patients.

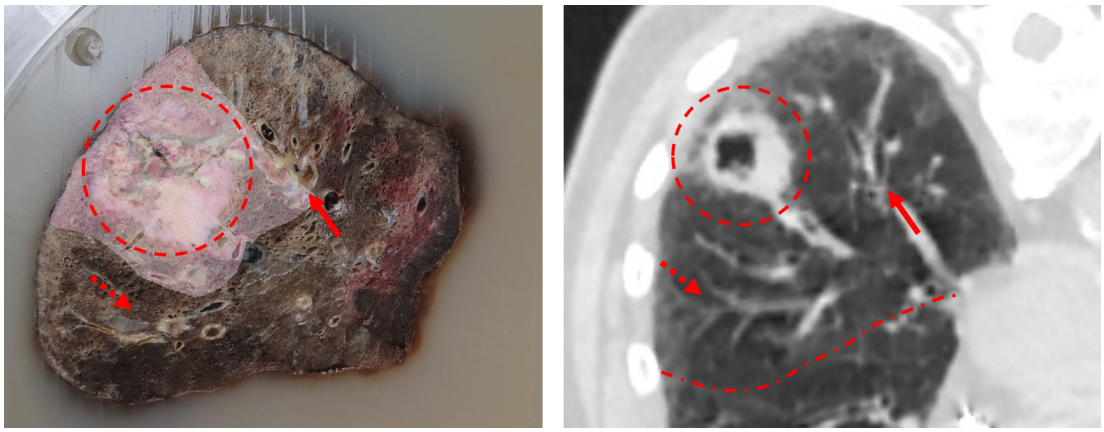
5.4 Conclusions

Lung tumour multimodality registration was carried out using a five-step approach, to tackle the different challenges previously identified. The use of binary tumour models, to perform shape-based registration, proved insufficient to determine the correct orientation of the volumes, despite the fact that high Dice overlap coefficients were reported. Anatomical fiducials had to be included in the lesion segmentations to improve the clinical significance of the registration results.

Moreover, PET models were found to be severely affected by breathing movement, which resulted in large volume overestimation and smooth, featureless shapes. For these reasons, PET images were not included in the final results. Similarly, in all cases, CT tumour volumes were larger than their pathology segmentations by 22% on average, in line with the results obtained in Chapter 4.



(a) Slice #1



(b) Slice #5

Figure 5.10: Two registered pairs of pathology (left column) and CT images (right column) from patient S6454. (a) Tumour boundary can be visualised as a slight change of colour in the pathology photograph, which corresponds to an area of ground glass opacity in the CT slice (dashed circle). Trifurcation of the main bronchus can be clearly seen on both modalities (solid arrow). Dot-dashed line indicates the boundary of the pulmonary lobe. (b) Solid tumour with necrotic, hollow core (dashed circle). Airway structures can be observed on both images, such as the oblique cross-sections of a group of tertiary bronchi (solid arrow) and a branched bronchiole (dotted arrow). Dot-dashed line indicates the boundary of the pulmonary lobe. Histology sections are overlaid on the pathology block-face photograph

Chapter 6

Clinical Validation of Results

This chapter focuses on the validation of the lung image registration results presented in Chapter 5. Due to the lack of ground truth, a subjective evaluation was performed by two experienced clinicians, using a blind, randomised, two-step approach. The first section covers the rationale behind this testing method. Next, the two experiments used for validation are described. Finally, in the last section, results are discussed.

6.1 Introduction

The performance of many medical image processing algorithms, such as lesion segmentation and registration, are usually assessed against the manual output obtained from expert clinicians, which acts as ground truth. However, in some circumstances, there is no gold standard against which to compare the solution delivered by an automatic computer method. The lack of a well-defined procedure included in the standards of care, or the high dimensionality of a particular dataset, are examples of why ground truth results are not always available.

In our case, the registration of pre-operative CT scans and post-operative pathology photographs is a procedure not performed as a standard clinical procedure, in part due to the inherent difficulties derived from the manual handling of two virtual 3D shapes on a two-dimensional medium, such as a computer screen. Therefore, the generation of a ground truth, even when defined by experienced thoracic consultants, would be complicated and unreliable.

To overcome this, we designed two tests that would allow an expert clinician to blindly compare and rate their manual pathology to CT alignments against the algorithm output, thus providing a consistent, qualitative measure of the automatic method's performance. These experiments are described in the following section.

6.2 Experiment Design

6.2.1 Manual Registration

A pathology consultant (CD) and a thoracic oncology consultant (SH) were independently asked to perform, to the best of their abilities and within the limitations of the software tools used, a manual registration of the pathology and CT tumour datasets. For this purpose, they were provided with a dual-screen computer: on the left monitor, they were shown the pathology photograph stack, which they could navigate using the mouse central wheel. On the right display, the corresponding CT volume was displayed in 3D Slicer [86], which allowed the user to rotate the dataset around its three main axes. Similarly, the CT stack could be navigated along its z axis using the mouse wheel. A screenshot showing the set-up used in 3D Slicer is displayed in Figure 6.1.

After familiarising themselves with the software controls, the consultants were asked to rotate the CT volume until visual correspondence with the pathology cross-sections was achieved. No time limit was enforced, and they performed the tasks independently, to avoid inter-operator bias. Once satisfied with the level of accuracy obtained, they notified the test instructor, who recorded the rotation and offset applied to the CT data. This process was repeated for three different patients: G3397, S3510 and S6454.

The purpose of this first test was to determine the feasibility of performing a manual multimodality regstration at baseline, that is, without the use of any registration algorithms. It was used to evaluate the user-friendliness of the software, time consumption, and registration accuracy. Also, it would allow the expert consultants to blindly review their own manual registrations, as described in Section 6.2.2.

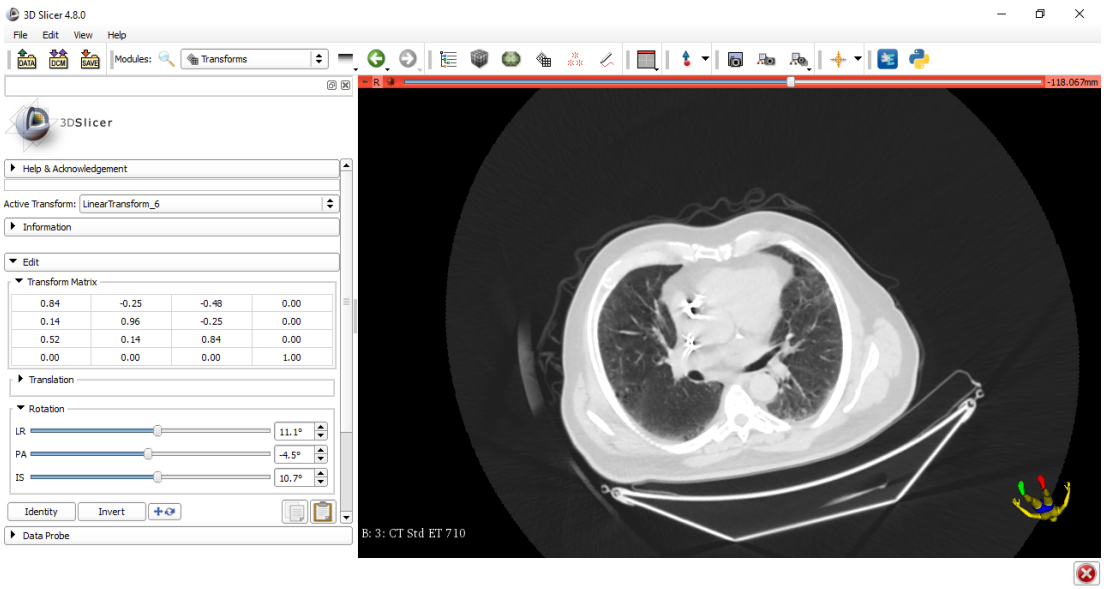


Figure 6.1: Screenshot showing the set-up in 3D Slicer used to rotate the CT volume. Operators can interact with the sliders shown on the left-hand side of the screen, which control the degree of rotation around the three main anatomical axes: left-right (LR), anteroposterior (PA) and craniocaudal (also known as inferior-superior, IS). The resulting transform matrix is also displayed

In addition, an external, non-expert operator was asked to repeat the test described above for the same three cases. This added result was used to increase the number of rating samples, and therefore potentially reduce the raters' bias.

6.2.2 Qualitative Rating of Registration Results

The test instructor collated the registration outputs from the expert, non-expert and automatic registrations, and arranged the pathology photographs next to their matched CT cross-sections. A sample dataset is shown in Figure 6.2, corresponding to the algorithm registration output for patient ST3510 (see Appendix A for a complete list of registration results). Three days later, both consultants (CD and SH) were given the anonymised registration results, grouped by patient. Each patient folder contained four cases, labelled A to D. They were informed that each of the four cases could be any of the following: the algorithm output, a non-expert manual registration, their own manual registration, or their fellow consultant's manual registration.

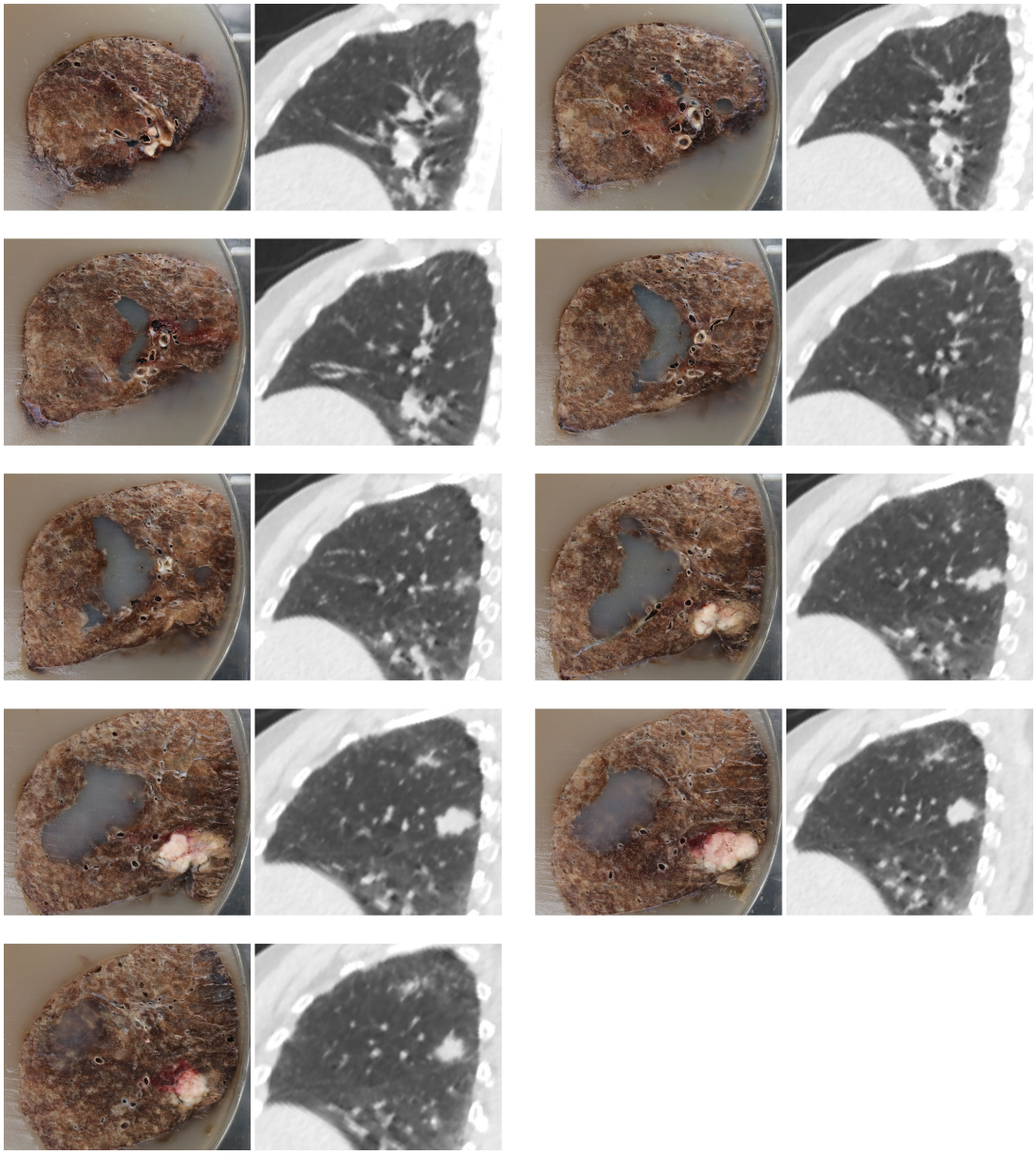


Figure 6.2: Sample dataset used for rating registration results, containing nine pairs of registered pathology (left) and CT cross-sectional images (right). Images correspond to the registration output generated by our algorithm, for patient ST3510.

A scoring rubric was also provided to the raters, with a grading scale ranging from 0 to 3. The scores were based on the number of anatomical correspondences between both datasets:

- **0**: no match between pathology and CT
- **1**: partial match, where most of the tumour mass coincides spatially on both pathology and CT
- **2**: good match, where tumour mass and at least one other feature (airways, tumour shape, lobe contour) coincide spatially on both pathology and CT
- **3**: best match, where tumour mass and at least two other features (airways, tumour shape, lobe contour) coincide spatially on both pathology and CT

Both experts, with no knowledge of the registration methods applied, graded all cases following the above rubric. Results are discussed in the next section.

6.3 Results and Discussion

Scores obtained from the assessment of registration outcomes, assigned by two experienced consultants (CD and SH), are summarised in Table 6.1. Several trends can be observed in the data: first, both raters agreed that the developed method consistently outperforms manual registrations in patients G3397 and S6454, where the algorithm output achieved the highest possible score. This result is encouraging, as it indicates that the method developed has the potential to deliver superior results, compared to the baseline standard of manual registration.

Patient S3510 also provides an interesting case. The algorithm output was rated as a no-match by both consultants. However, each expert graded their colleague's registration as a no-match, while assigning a 2 out of 3 score to their own. Upon inspection of the registered datasets for that patient, a large pocket of agar was observed in many of the pathology cross-sections (see Figure 6.2). These artefacts can appear when the lung is affected with emphysema, which creates large, hollow spaces within the tissue. When agar is injected as a fixation medium, it tends to acumulate in these cavities, thus distorting the shape of the lobe. Our hypothesis is that the deformation introduced an additional confounding factor for the raters, which potentially demonstrates that, in complex cases, experts may be unable to reach consensus on whether a given alignment is *correct* or *incorrect*.

	Registration score			
Patient ID	**Algorithm**	**Non-Exp**	**Exp-CD**	**Exp-SH**
G3397	3	1	1	2
S3510	0	1	2	0
S6454	3	2	1	2

(a) CD results

	Registration score			
Patient ID	**Algorithm**	**Non-Exp**	**Exp-CD**	**Exp-SH**
G3397	3	0	0	1
S3510	0	1	0	2
S6454	3	3	2	2

(b) SH results

Table 6.1: Scores obtained from blind assessment of registration outcomes, assigned by experienced consultants CD and SH. For each patient, four different cases were evaluated: algorithm output, non-expert manual registration, expert (CD) manual registration, and expert (SH) manual registration. Scores range from 0 to 3, as defined in the evaluation rubric described in Section 6.2.2

That said, patient G3397 also shows large agar pockets in some of the block-face photographs. However, it is suspected that the presence of a main bronchus and blood vessels in the images, along with the tumour, helped experts to establish a relationship between both scenes, regardless of distortion.

Another relevant observation is that in all cases, non-expert manual registrations are at least as good as one of the expert outputs. Two mutually-exclusive readings can be made from this: either manual registration is a trivial task, that can be performed by non-trained operators; or it is a complex undertaking, open to misinterpretation and prone to error. The lack of expert consensus shown in case S3510, added to the inherent difficulty experienced by the operators when carrying out the manual transformations described in Section 6.2.1, support the latter hypothesis.

In terms of data variance, both graders scored all cases consistently, with inter-rater differences of 1 point at most. These small deviations can be explained by the lack of a quantitative scoring rubric, which potentially leaves the definition of a grade open to interpretation. The only exception to this trend was observed in

the experts' ratings for patient S3510, which could be attributed to the apparent complexity of this case.

Finally, it is worth commenting on the user experience the operators had with the manual registration software, 3D Slicer. Regardless of their level of expertise, all three individuals spent between 5 and 7 minutes per patient. However, two main usability issues were reported: firstly, the lack of a 3D representation of the tumour in both modalities; and secondly, the fact that pathology cross-sections could not be overimposed to the CT. Regarding the first issue, the registration of a stack of 2D pathology photographs with a volumetric CT scan effectively becomes a 2D-to-3D registration, which is more complex to perform compared to 3D-to-3D. The impossibility of overlaying pathology data onto the CT images was a limitation of the software used. These two factors reflect the inherent difficulty of the task, and may explain why none of the expert registrations achieved the highest score.

6.4 Conclusions

The purpose of these tests was two-fold: to investigate the feasibility of performing manual multimodality registrations with off-the-shelf, open-source software; and to provide a framework for validating and rating these manual alignments against the output of our proposed method. Operators encountered several difficulties when trying to manually register pathology and CT data, due to the lack of a 3D perspective of the scene and the impossibility of overlaying both modalities.

Blind scoring of different registration outputs revealed that the proposed algorithm consistently outperforms manual methods in 2 out of 3 cases. The no-match obtained in case S3510 could be explained by the presence of additional deformations in the pathology dataset, supported by the fact that both expert graders disagreed on the quality of their own manual registrations.

Chapter 7

Conclusion and Future Work

The scope of this project was to develop a robust imaging framework to enable the registration of *in vivo* and post-operative scans of lung cancer patients. A review of the literature identified several studies pursuing the same objective. However, three main shortcomings were found across all reviewed articles: firstly, inconsistencies in the pathological dissection of the lung specimen could potentially result in non-parallel cross-sections, introducing error in the posterior tumour model reconstruction and registration. Secondly, tumour volume reconstruction from pathology photographs was disregarded in most cases, or otherwise built using cross-sectional area projection. Lastly, registration of pre- and post-operative modalities was either performed manually, or assumed co-planar. With these limitations in mind, a comprehensive data acquisition and processing protocol was designed.

Initial experiments were carried out on synthetic tumour phantoms. The fundamental advantage of this approach was the existence of a ground truth against which to compare the shape reconstruction algorithms. Physical dissection of the phantom was performed on a bespoke soft tissue slicing rig, to observe the relationship between slice thickness and overall reconstructed shape. Similarly, virtual slicing of the model allowed us to investigate the outcomes of different interpolation functions in terms of shape and volume recovered.

In the next phase of the study, nine lung cancer patients who were due for radical lobectomy were recruited. Their resected tissue was sliced at 5 mm intervals using the slicing rig, and each cross-section photographed. Tumour area was delineated on each image by experienced consultants, and virtual 3D models of the lesions

were reconstructed. Likewise, pre-operative PET/CT scans were segmented. Unfortunately, PET images could not be used, as they were severely distorted by respiratory artefacts. The registration of pathology and CT models was guided by the shape of the tumour, combined with external anatomical fiducials. A pre-registration step (based on their minimal bounding box) was used prior to performing rigid registration. Moreover, histopathology slides were matched to their block-face photographs by applying a non-linear registration algorithm.

Validation of the results was a complex task, with no ground truth to evaluate the outcomes of the proposed method against. Dice coefficient was calculated to assess the registrations, but proved to be insufficient as a quality measure. To investigate further and establish a baseline standard, two experienced thoracic consultants and a non-expert performed manual registrations between pathology and CT for three different patients. Subsequently, the two experts were asked to rate all registrations, generated manually and by the proposed algorithm, without knowing which method was used in each case. Results showed that in two out of three cases both clinicians scored the algorithm's output as better than their own manual alignments.

This proof-of-concept study was conceived to ascertain whether meaningful registrations could be carried out between *in vivo* functional imaging and histopathology maps. While there are still difficulties to overcome, some of the results presented earlier are promising, suggesting that the proposed approach outperforms expert manual registration, and thus demonstrating its feasibility.

In terms of future work, the next logical step would be to include functional imaging into the registration workflow. As discussed in Chapter 5, the current standard of care for lung cancer patients is to use conventional, free-breathing PET/CT combined scans. Although faster and simpler, respiratory motion blurs the signal, which causes the total tumour volume to be grossly overestimated. One potential solution to this problem would be to use respiratory-gated 4D PET, where breathing movement is cancelled out by tracking the patient's chest motion. At the time of writing, two new 4D PET/CT scanning protocols were being deployed and trialled within NHS Greater Glasgow and Clyde. The main advantage of using combined scans is that both PET and CT signals are already registered by the scanner. For this reason, adding the functional imaging modality to our method is straightforward. Unfortunately, we could not fulfil this requirement due to time constraints. As said before, respiratory-gated PET/CT

was not the standard procedure for lung cancer patients. To acquire this modality, subjects would have to be re-scanned using the new protocol, in addition to their standard diagnostic scan. As PET/CT imaging involves ionising radiation, it would have been necessary to apply to the relevant NHS Research Ethics Committee for a substantial protocol amendment, and prove that the benefits obtained from the new modality outweighed the risks for the patient. Due to the limited timespan of the project, that was not possible.

Another unanswered research question is whether the macroscopic tumour segmentation matches the microscopic extent of the disease. Currently, our reconstruction method relies on a consultant pathologist (CD) manually tracing the contours of the lesion on the block-face photographs. The decision to include or exclude a region within the tumour boundary is based on colour information (lung carcinomas are normally whiter than their surroundings) and their clinical experience. However, in some cases, the chromatic contrast between healthy and diseased tissue is not clear, and tumour segmentation becomes ambiguous. Further research work would determine the precision of the macroscopic delineation by comparing it against the lesion shown on histopathology images. The latter, due to their high resolution, would allow the pathologist to trace the tumour contour at cellular level, and could act as ground truth for testing different segmentation routines.

Appendix A

Lung Image Registration Results

This appendix displays the registration outcomes for patients G3397 (Figure A.2), S3510 (Figure A.3) and S6454 (Figure A.4), generated manually and by the proposed algorithm. Rows within the figures correspond to the slice number. Columns are arranged as follows, from left to right: pathology block-face photograph; CT volume registered by our method; CT volume manually registered by non-expert; CT volume manually registered by consultant 1 (CD); and CT volume manually registered by consultant 2 (SH). For clarification purposes, Figure A.1 shows a diagram with the arrangement of the registration results within Figures A.2, A.3 and A.4.

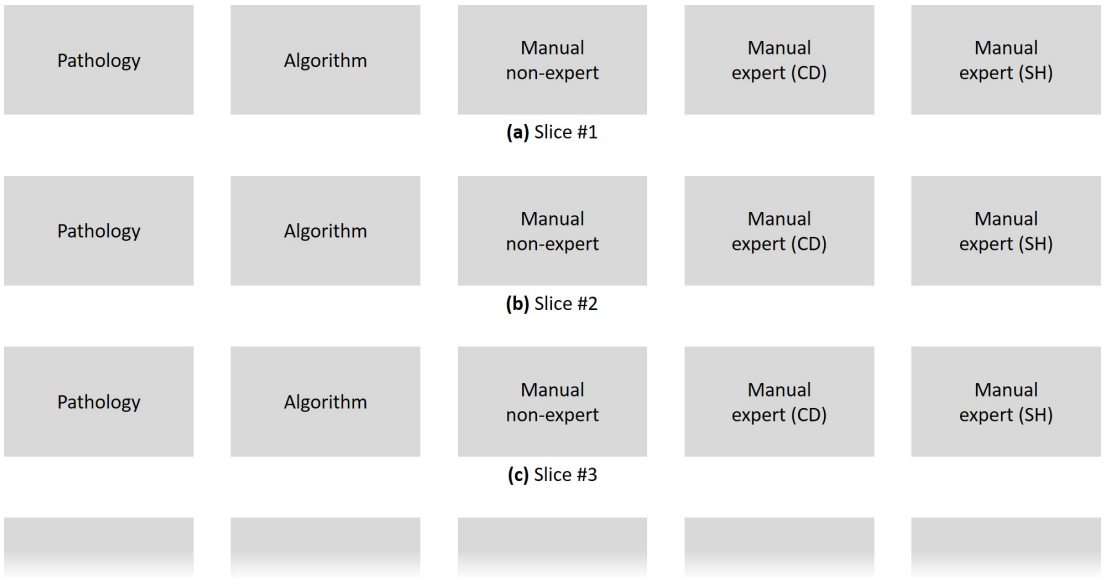
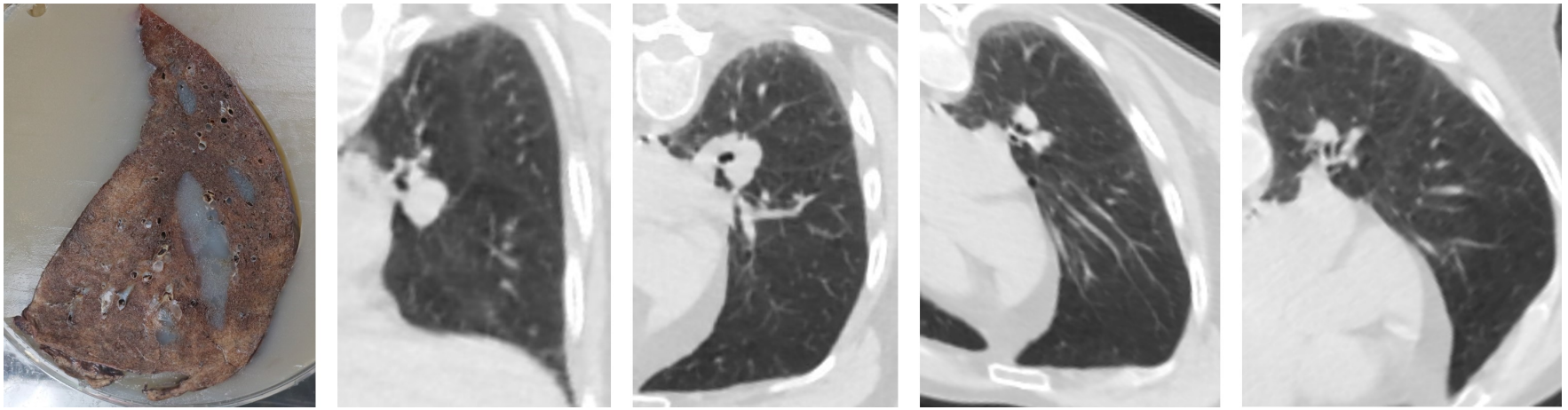
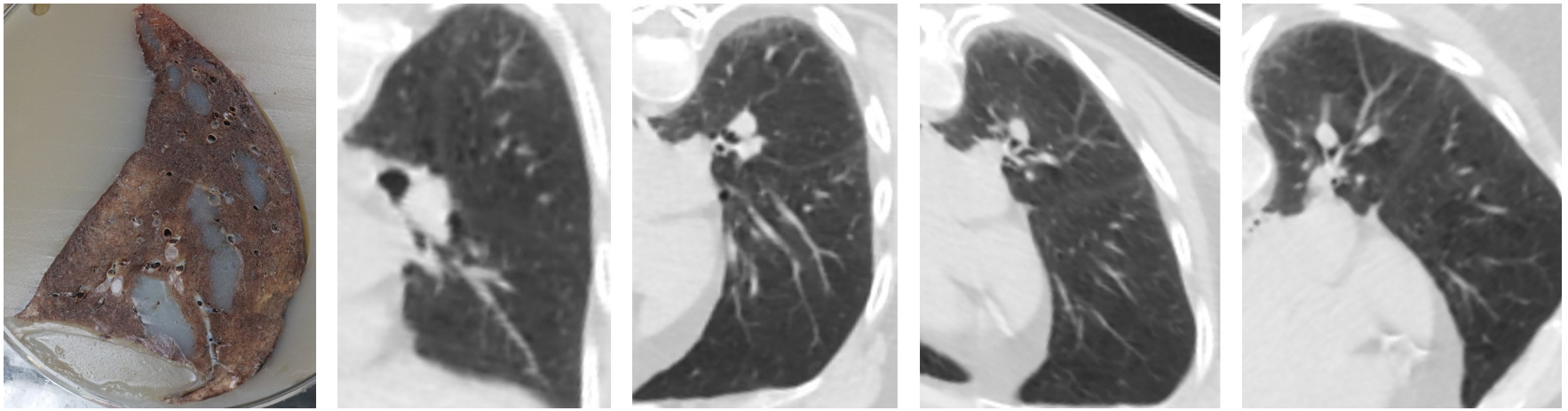


Figure A.1: Diagram showing the arrangement of the registration results

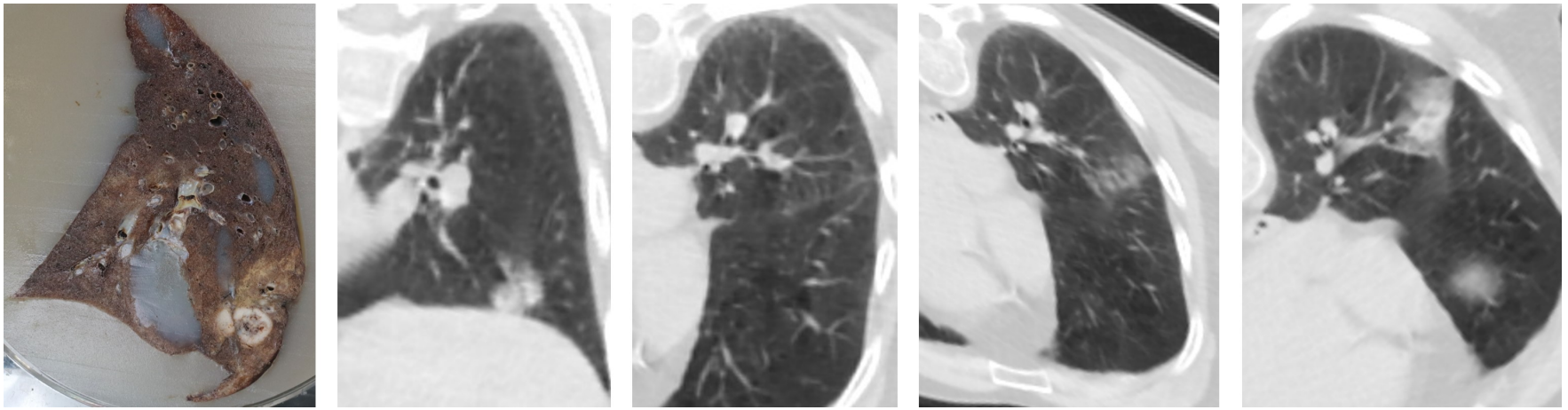


(a) Slice #1

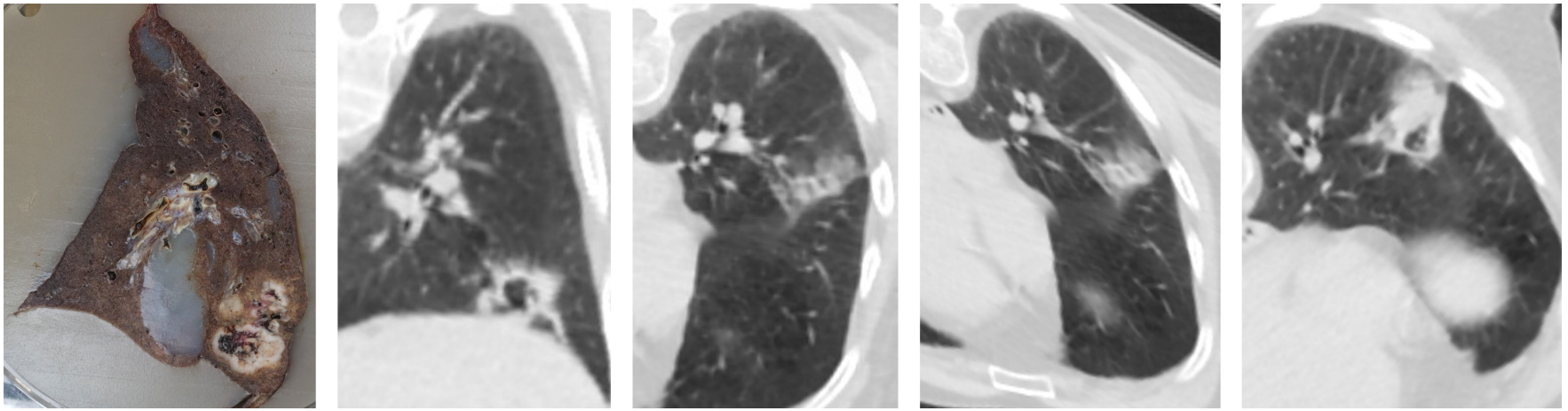


(b) Slice #2

Figure A.2: Registration outcomes for patient G3397. From left to right: pathology photograph; algorithm output; non-expert manual registration; expert (CD) manual registration; and expert (SH) manual registration

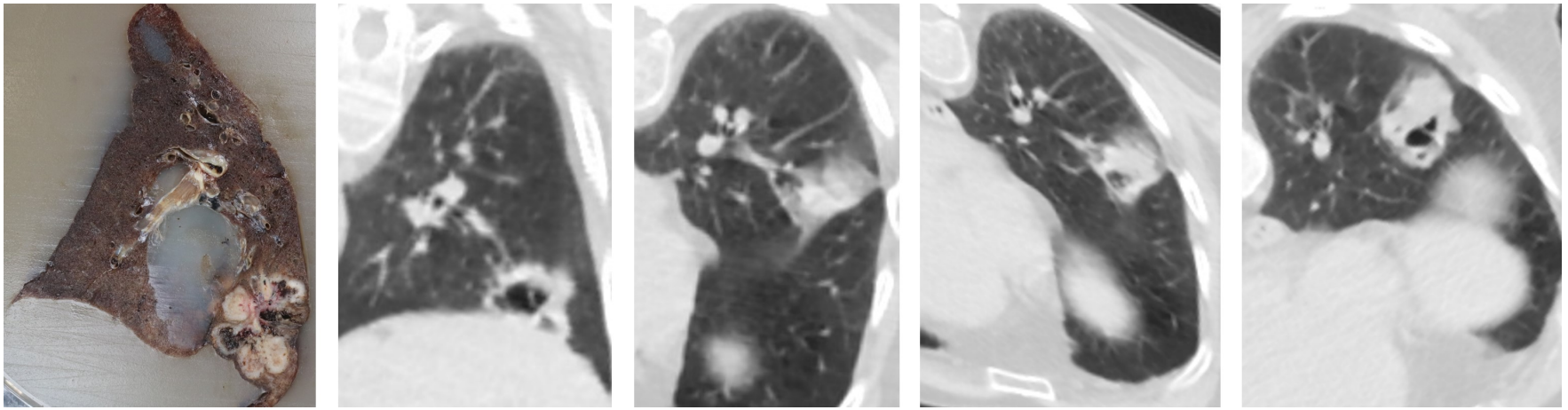


(c) Slice #3

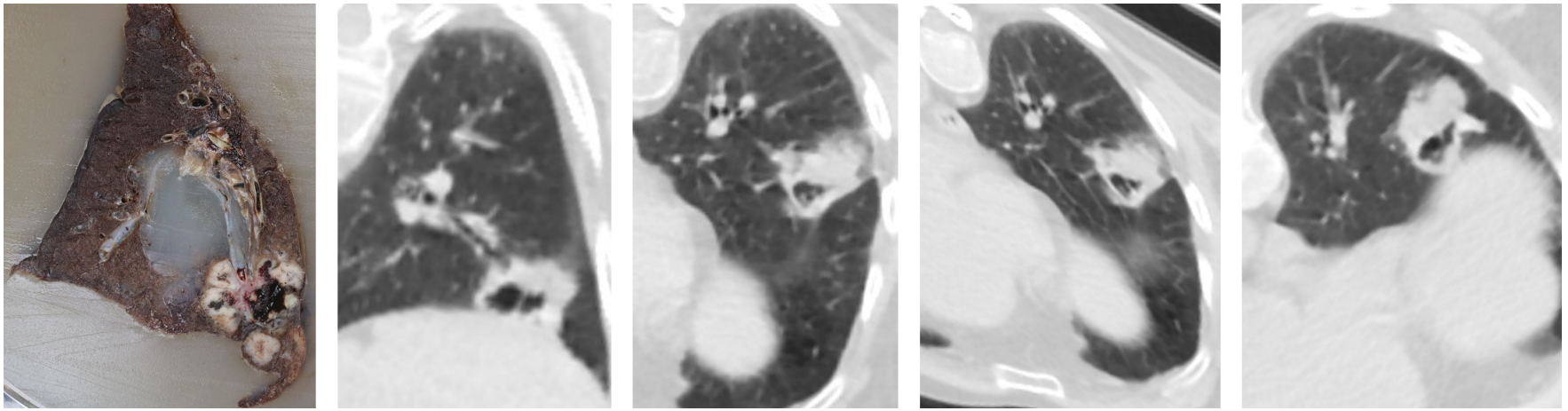


(d) Slice #4

Figure A.2 (contd.): Registration outcomes for patient G3397. From left to right: pathology photograph; algorithm output; non-expert manual registration; expert (CD) manual registration; and expert (SH) manual registration

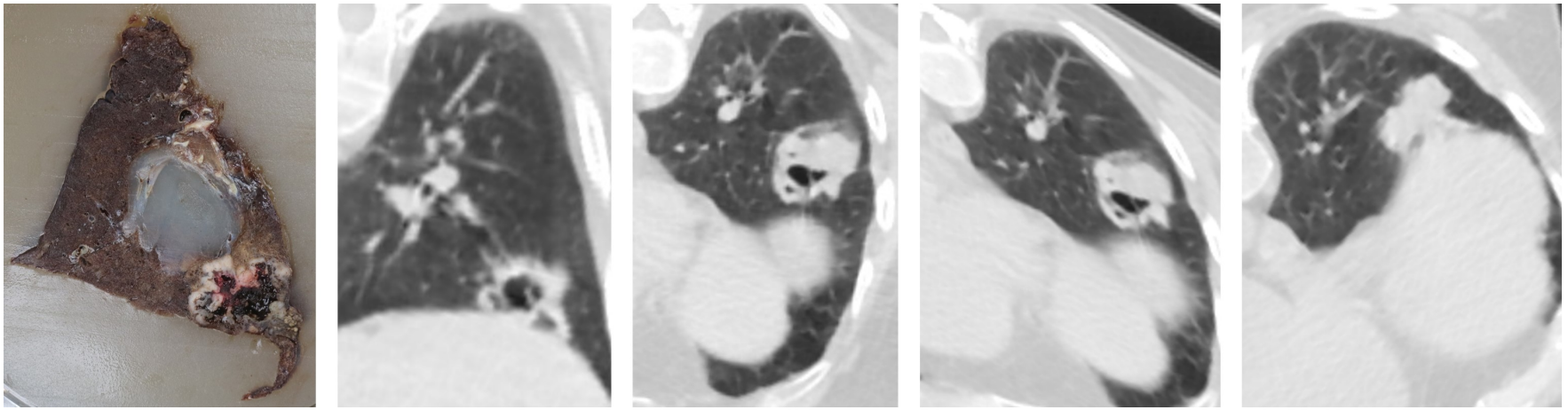


(e) Slice #5

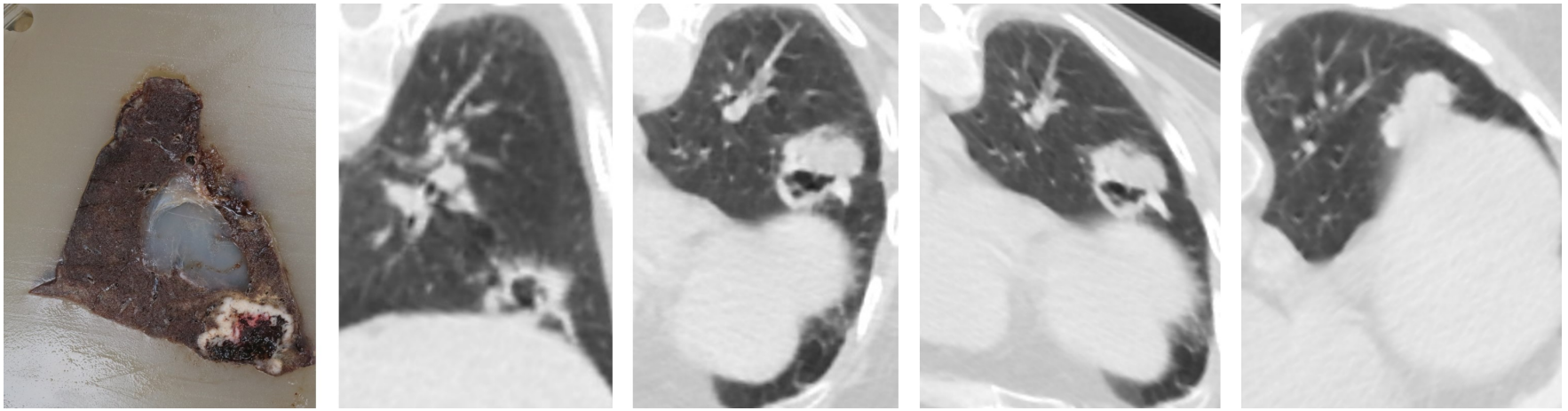


(f) Slice #6

Figure A.2 (contd.): Registration outcomes for patient G3397. From left to right: pathology photograph; algorithm output; non-expert manual registration; expert (CD) manual registration; and expert (SH) manual registration

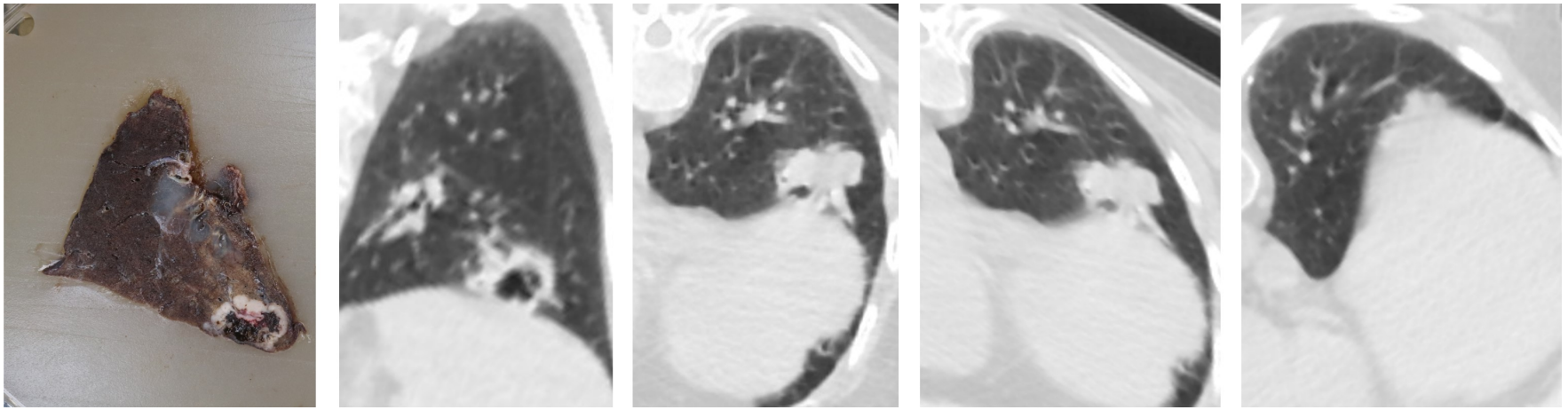


(g) Slice #7

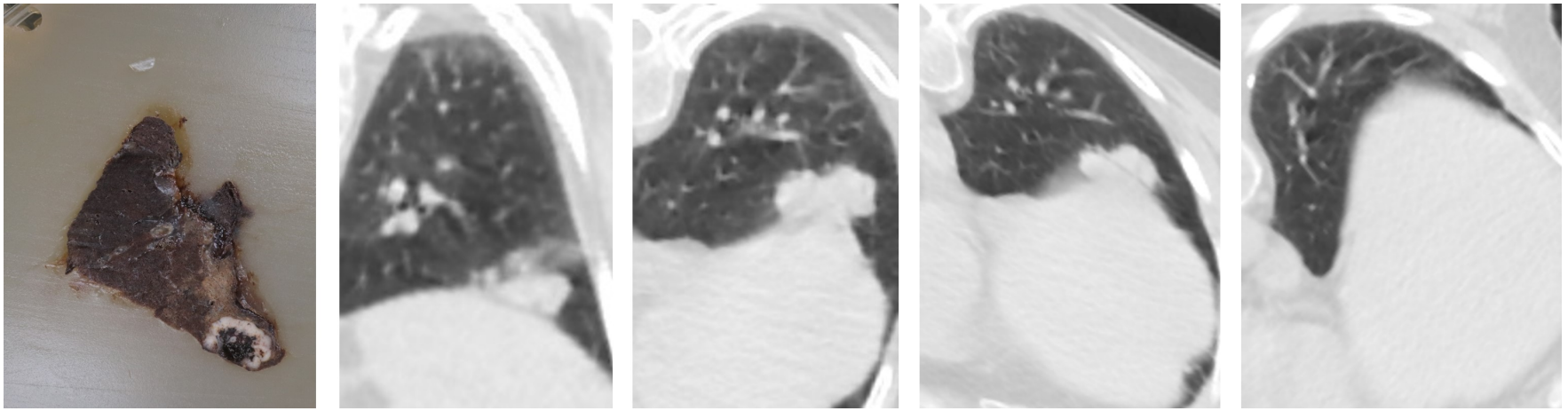


(h) Slice #8

Figure A.2 (contd.): Registration outcomes for patient G3397. From left to right: pathology photograph; algorithm output; non-expert manual registration; expert (CD) manual registration; and expert (SH) manual registration

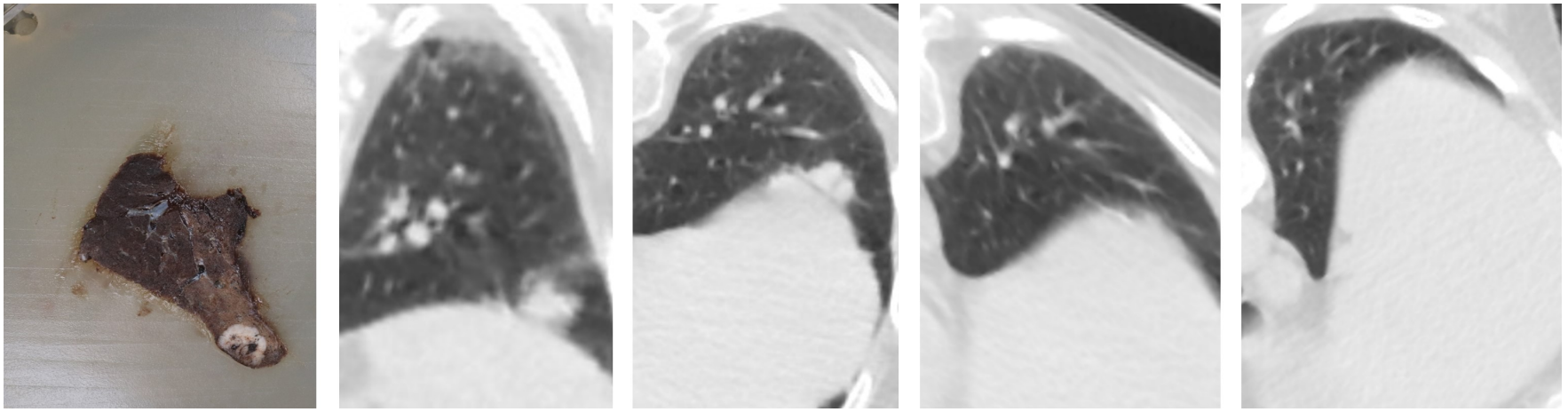


(i) Slice #9



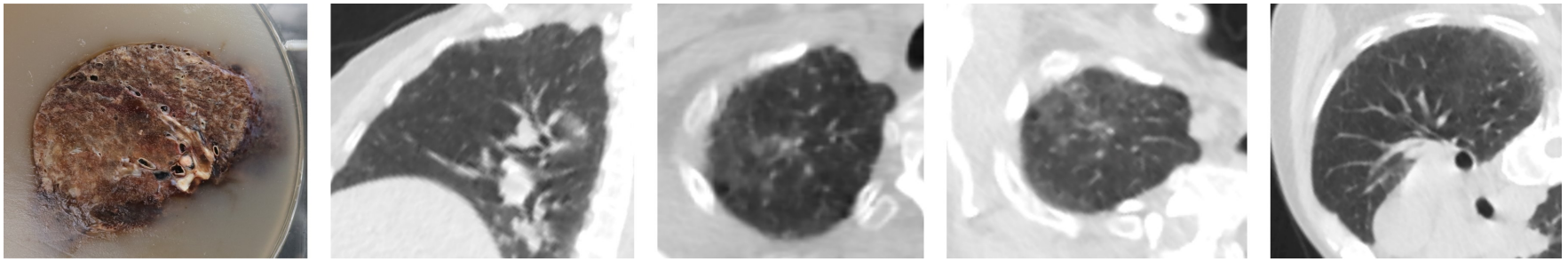
(j) Slice #10

Figure A.2 (contd.): Registration outcomes for patient G3397. From left to right: pathology photograph; algorithm output; non-expert manual registration; expert (CD) manual registration; and expert (SH) manual registration

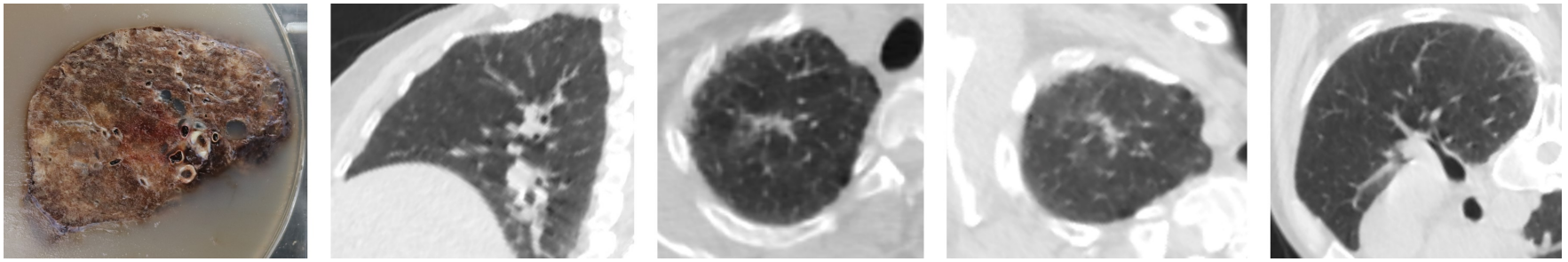


(k) Slice #11

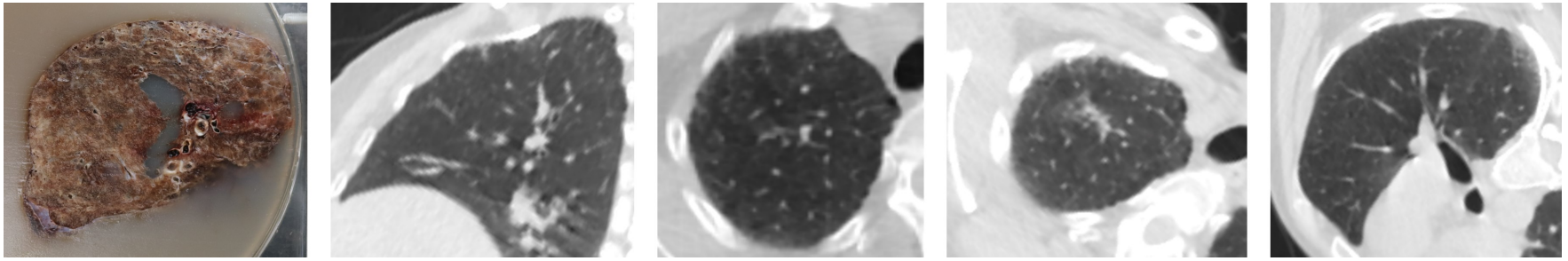
Figure A.2 (contd.): Registration outcomes for patient G3397. From left to right: pathology photograph; algorithm output; non-expert manual registration; expert (CD) manual registration; and expert (SH) manual registration



(a) Slice #1

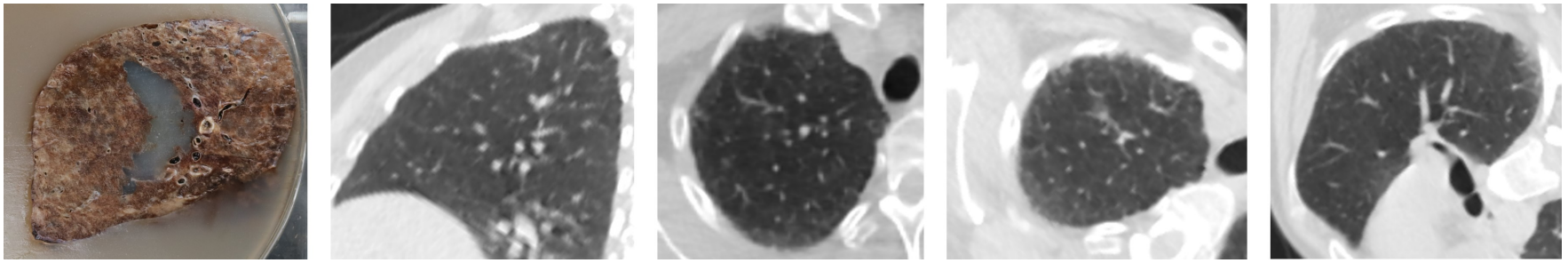


(b) Slice #2

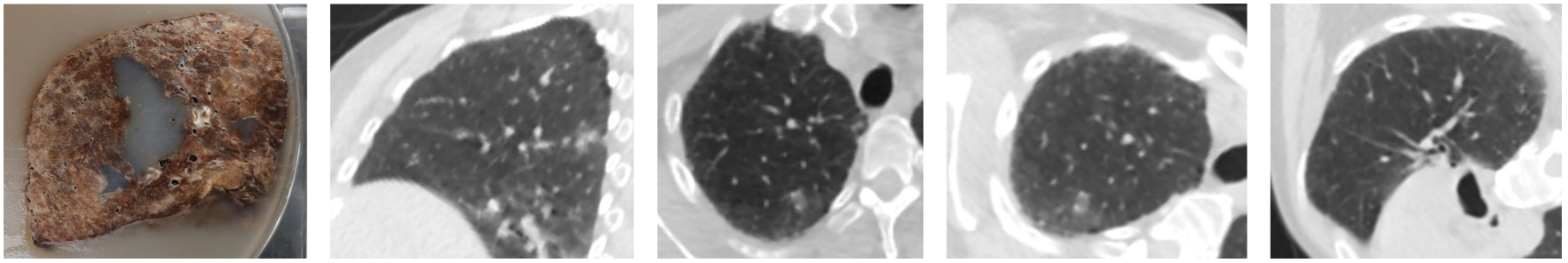


(c) Slice #3

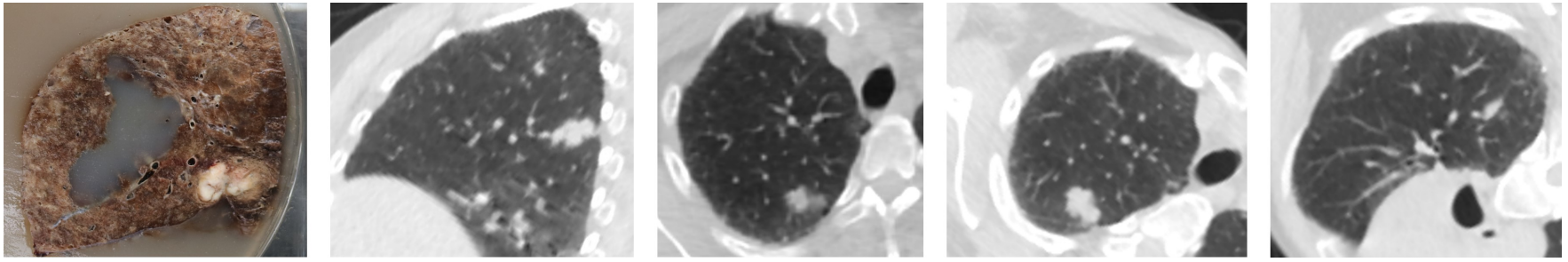
Figure A.3: Registration outcomes for patient S3510. From left to right: pathology photograph; algorithm output; non-expert manual registration; expert (CD) manual registration; and expert (SH) manual registration



(d) Slice #4

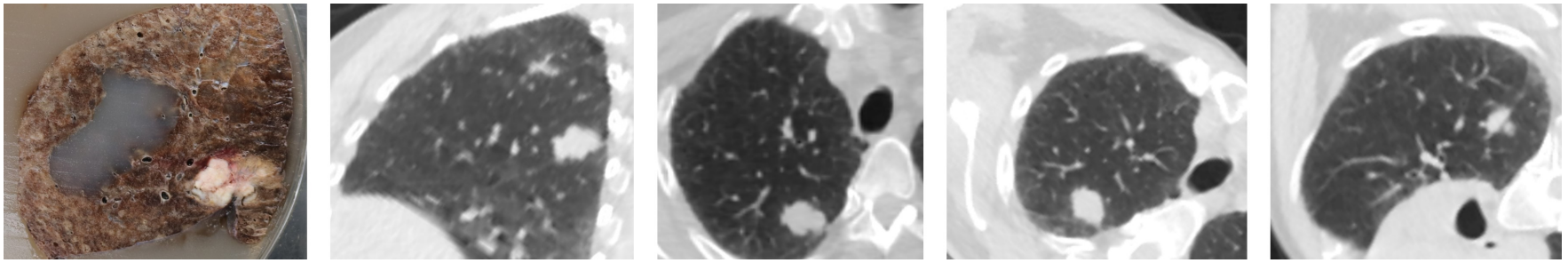


(e) Slice #5

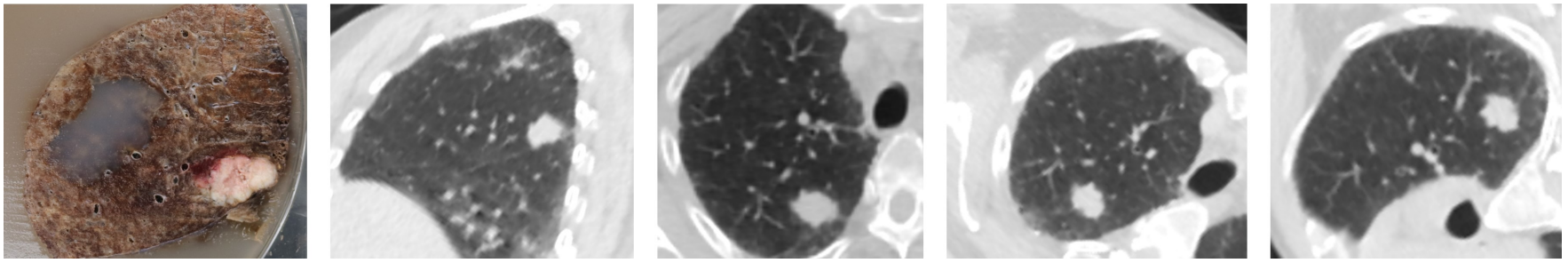


(f) Slice #6

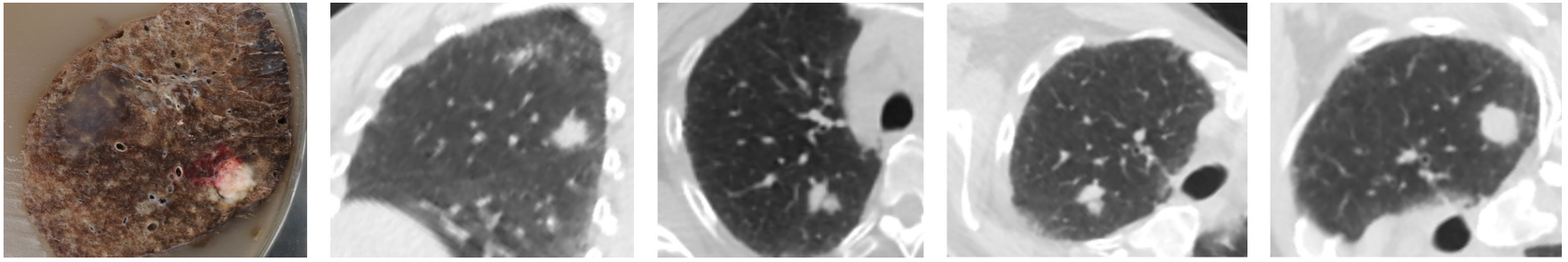
Figure A.3 (contd.): Registration outcomes for patient S3510. From left to right: pathology photograph; algorithm output; non-expert manual registration; expert (CD) manual registration; and expert (SH) manual registration



(g) Slice #7

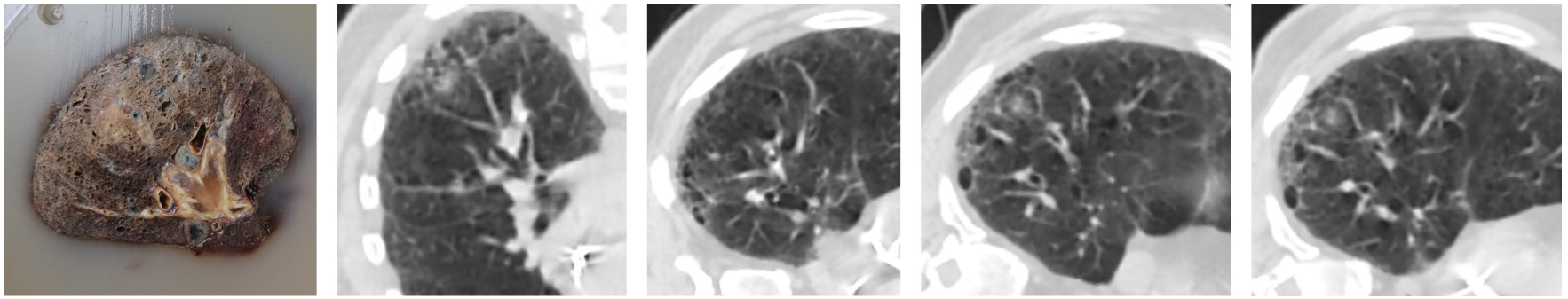


(h) Slice #8

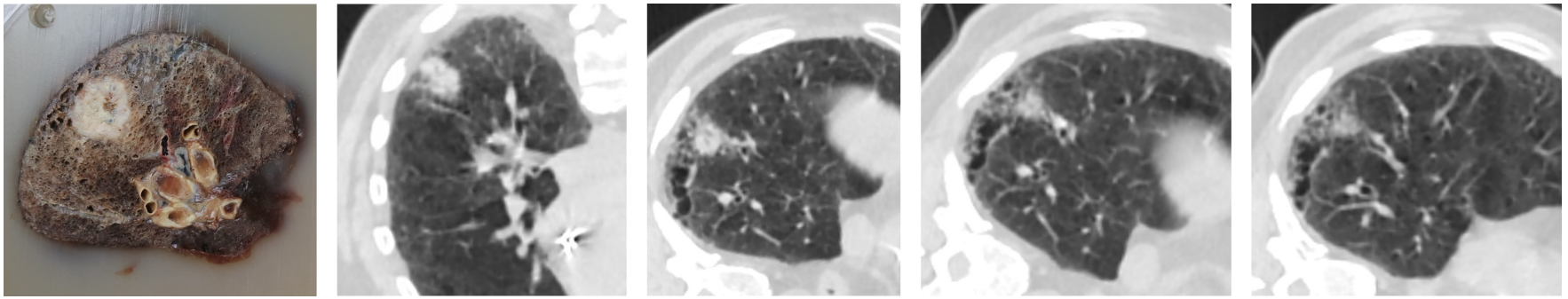


(i) Slice #9

Figure A.3 (contd.): Registration outcomes for patient S3510. From left to right: pathology photograph; algorithm output; non-expert manual registration; expert (CD) manual registration; and expert (SH) manual registration

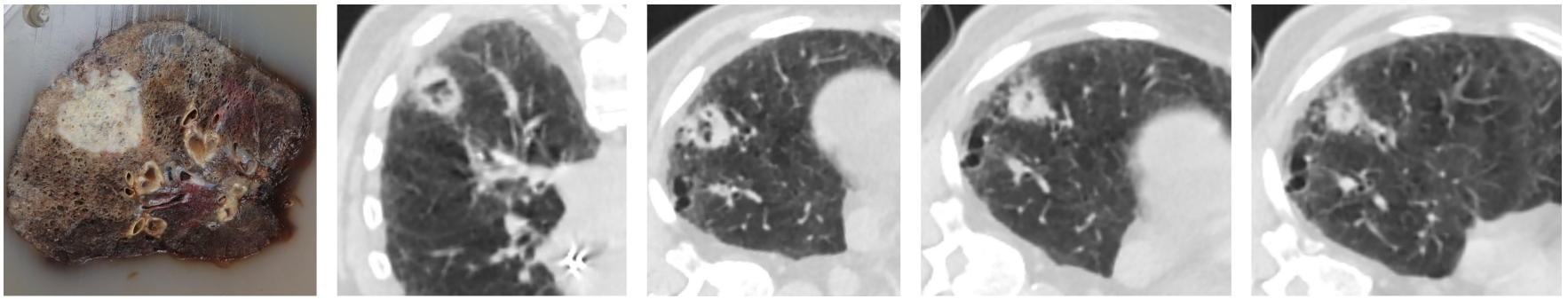


(a) Slice #1

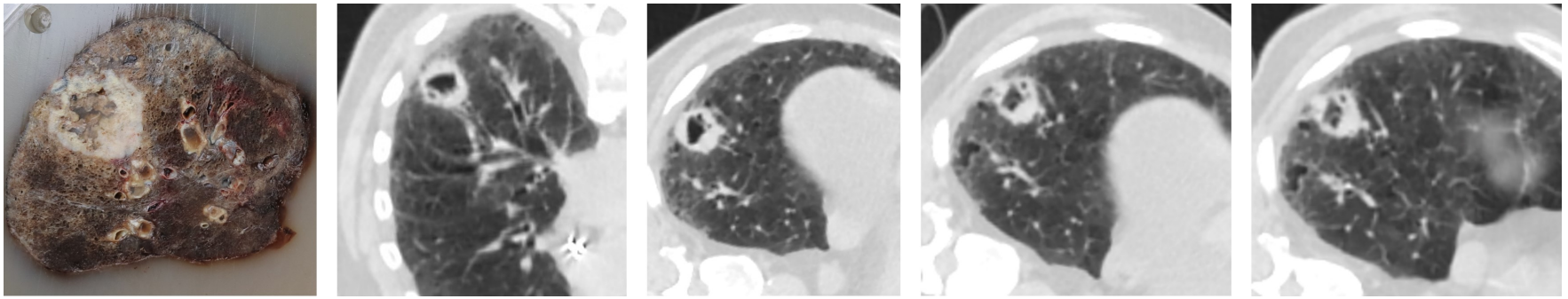


(b) Slice #2

Figure A.4: Registration outcomes for patient S6454. From left to right: pathology photograph; algorithm output; non-expert manual registration; expert (CD) manual registration; and expert (SH) manual registration

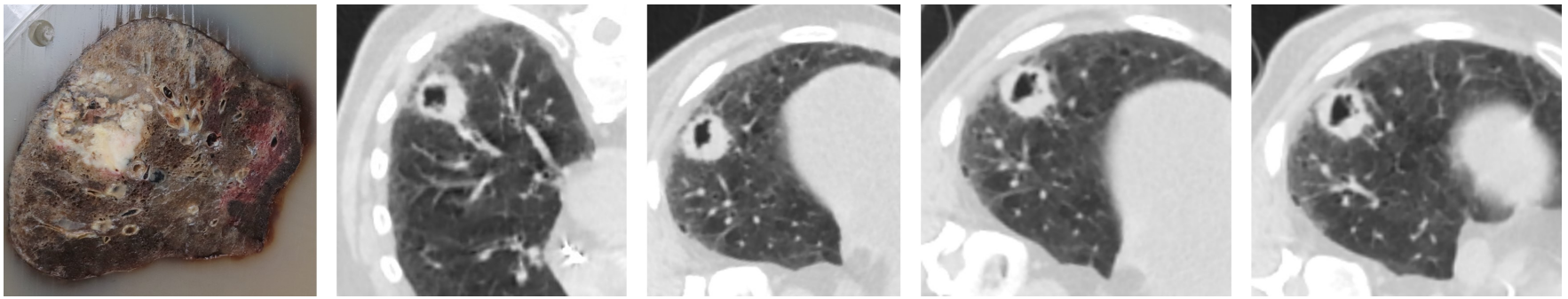


(c) Slice #3

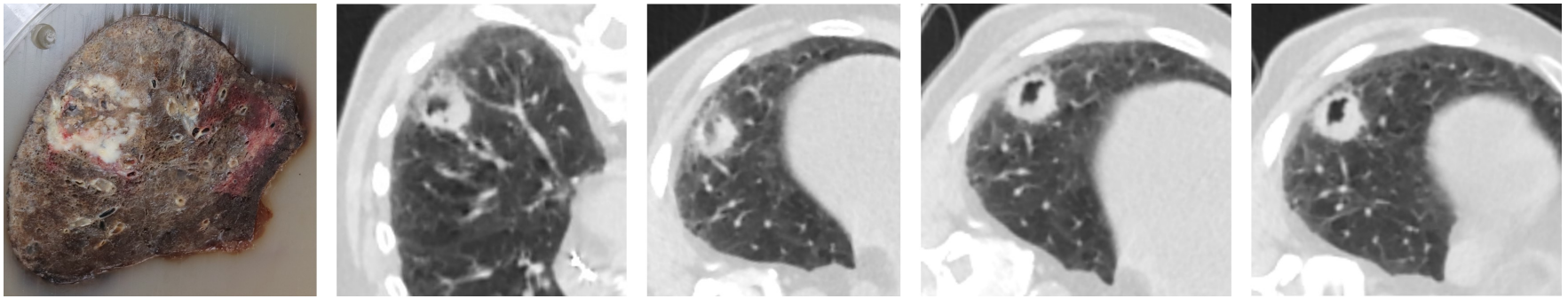


(d) Slice #4

Figure A.4 (contd.): Registration outcomes for patient S6454. From left to right: pathology photograph; algorithm output; non-expert manual registration; expert (CD) manual registration; and expert (SH) manual registration

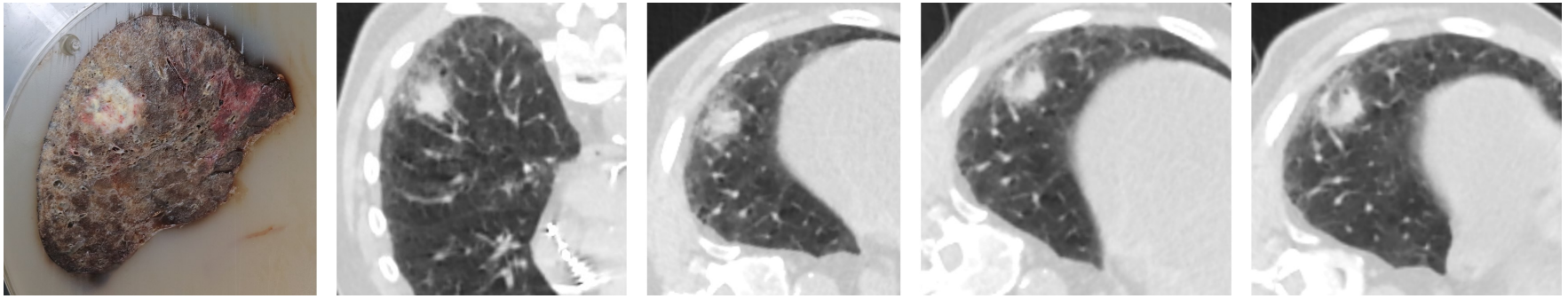


(e) Slice #5

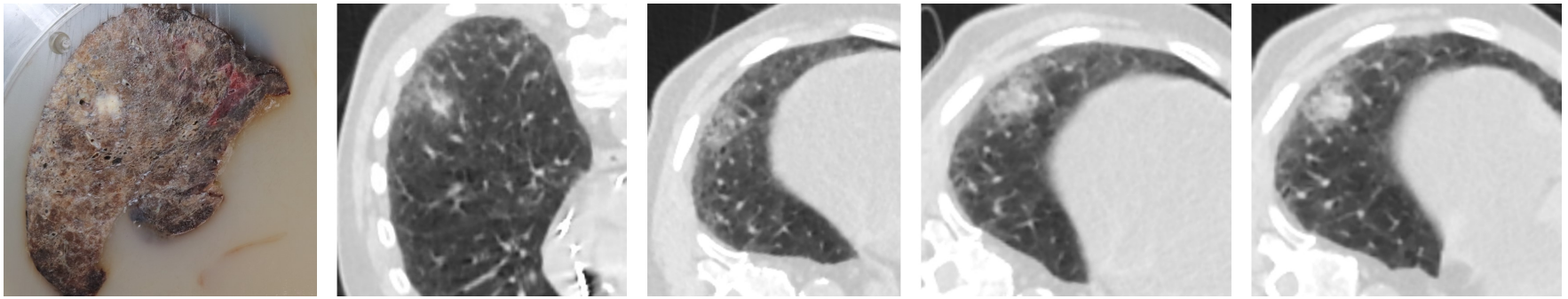


(f) Slice #6

Figure A.4 (contd.): Registration outcomes for patient S6454. From left to right: pathology photograph; algorithm output; non-expert manual registration; expert (CD) manual registration; and expert (SH) manual registration



(g) Slice #7



(h) Slice #8

Figure A.4 (contd.): Registration outcomes for patient S6454. From left to right: pathology photograph; algorithm output; non-expert manual registration; expert (CD) manual registration; and expert (SH) manual registration

Appendix B

Participant Information Sheet and Consent Form

This appendix includes the Participant Information Sheet and Consent Form used in the patient recruitment process described in Section 5.2.1.

PARTICIPANT INFORMATION SHEET

Investigation of new medical imaging techniques for improving accuracy in lung cancer treatment

You are being invited to take part in a research project. This document contains information regarding the implications for the patient and the potential benefits for the wider community. Please take time to read the following information, and feel free to discuss it with others. **Please be assured that your clinical standard of care and future treatment will not be affected in any way regardless of your decision**.

Why have I been chosen to participate?

For the purpose of our research study, we need adult patients who have been diagnosed with lung cancer and are due to undergo surgery for their tumour.

What should I expect if I choose to take part?

You will not be asked to do anything extra, and the study will not change your treatment or diagnosis. We only wish to perform additional scans on the lung tumour after it has been removed. Additionally, we need to collect information from your records.

What will happen to my extracted tissue if I choose to take part?

Following any surgery, the tissue is sent to the pathologist for dissection and analysis, to give important information about the tumour. The study will follow this normal pathway, but there will be some additional steps which will look at the tissue in a more detailed way. The tissue will also be scanned before dissection and during dissection to try and give us more information about the tumour - this is an additional step which does not normally take place. It will not interfere with the normal pathology work we do on your specimen. This additional step is expected to provide information to help us improve the accuracy of the scans that have been taken before the operation. We can then compare them with the scans and the pathology information we take after the tumour has been removed.

What are the benefits of taking part?

It is not likely that the study will benefit you directly. Our aim is to build a system that allows lung tumour scans to be used more accurately in the future. This should help with treatments and therapy that rely on these scans. You are welcome to request project updates and/or relevant publications by contacting any of the investigators.

Do I have to take part?

No, and be assured that your clinical standard of care will not be affected in any way regardless of your decision.

Will my personal information be kept confidential?

Your records will only be accessed by clinical staff (e.g. doctors, nurses) who are part of your care team. Additionally, clinical research regulatory authorities might also need to access your records for research audit purposes.

Contact information:

[redacted]

If you wish to take part in the study you will be asked to read and sign a Consent Form. You will be provided with a copy of this Participant Information Sheet and the signed Consent Form for you to keep.

Centre Number:
Study Number:
Participant ID num. for this trial:

CONSENT FORM

Title of Project: PET-PATH: Investigation of new medical imaging techniques for improving lung cancer treatment

Name of Principal Investigators: [redacted]

Please initial box

1. I confirm that I have read the information sheet dated ..... (version .....) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. ☐
3. I understand that relevant sections of my medical notes and data collected during the study may be looked at by individuals from regulatory authorities or from NHS GG&C, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐
4. I understand that the information collected about me will be used to support other ethically approved research in the future, and may be shared anonymously with other researchers. ☐
5. I agree to take part in the above study. ☐

_____	_____	_____
Name of Participant	Date	Signature
_____	_____	_____
Name of Person taking consent	Date	Signature

When completed: 1 for participant; 1 for researcher site file; 1 (original) to be kept in medical notes.

References

[1] Cancer Research UK, "Lung cancer." https://www.cancerresearchuk.org/about-cancer/lung-cancer, accessed June 2019.

[2] National Health Service UK, "Lung cancer." https://www.nhs.uk/conditions/lung-cancer/, accessed June 2019.

[3] T. Coroller, V. Agrawal, V. Narayan, Y. Hou, P. Grossmann, S. Lee, R. Mak, and H. Aerts, "Radiomic phenotype features predict pathological response in non-small cell lung cancer," *Radiotherapy and Oncology*, vol. 119, no. 3, pp. 480–486, 2016.

[4] J. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[5] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *IEEE Transactions on Medical Imaging*, vol. 32, no. 7, pp. 1153–1190, 2013.

[6] F. P. M. Oliveira and J. M. R. S. Tavares, "Medical image registration: a review," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 17, no. 2, pp. 1–21, 2012.

[7] C. Broit, *Optimal registration of deformed images*. PhD thesis, University of Pennsylvania, 1981.

[8] C. Davatzikos, "Spatial transformation and registration of brain images using elastically deformable models.," *Computer Vision and Image Understanding*, vol. 66, no. 2, pp. 207–222, 1997.

[9] D. Alexander, J. Gee, and R. Bajcsy, "Elastic matching of diffusion tensor MRIs," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 233–250, 1999.

[10] R. Bajcsy and S. Kovačič, "Multiresolution elastic matching," *Computer Vision, Graphics, and Image Processing*, vol. 46, no. 1, pp. 1–21, 1989.

[11] B. Marami, S. Sirouspour, S. Ghoul, J. Cepek, S. R. H. Davidson, D. W. Capson, J. Trachtenberg, and A. Fenster, "Elastic registration of prostate MR images based on estimation of deformation states," *Medical Image Analysis*, vol. 21, no. 1, pp. 87–103, 2015.

[12] G. E. Christensen, S. C. Joshi, and M. I. Miller, "Volumetric transformation of brain anatomy," *IEEE Transactions on Medical Imaging*, vol. 16, no. 6, pp. 864–877, 1997.

[13] E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens, "A viscous fluid model for multimodal non-rigid image registration using mutual information," *Medical Image Analysis*, vol. 7, no. 4, pp. 565–575, 2003.

[14] J. P. Thirion, "Image matching as a diffusion process: an analogy with Maxwell's demons.," *Medical Image Analysis*, vol. 2, no. 3, pp. 243–260, 1998.

[15] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Symmetric log-domain diffeomorphic registration: A demons-based approach," *Lecture Notes in Computer Science*, vol. 5241, pp. 754–761, 2008.

[16] N. D. Cahill, J. A. Noble, and D. J. Hawkes, "A Demons algorithm for image registration with locally adaptive regularization," *Lecture Notes in Computer Science*, vol. 5761, pp. 574–581, 2009.

[17] A. Gooya, G. Biros, and C. Davatzikos, "Deformable registration of glioma images using EM algorithm and diffusion reaction modeling," *IEEE Transactions on Medical Imaging*, vol. 30, no. 2, pp. 375–390, 2011.

[18] L. Han, J. H. Hipwell, B. Eiben, D. Barratt, M. Modat, S. Ourselin, and D. J. Hawkes, "A nonlinear biomechanical model based registration method for aligning prone and supine mr breast images," *IEEE Transactions on Medical Imaging*, vol. 33, no. 3, pp. 682–694, 2014.

[19] V. Rajagopal, A. Lee, J. H. Chung, R. Warren, R. P. Highnam, M. P. Nash, and P. M. F. Nielsen, "Creating Individual-specific Biomechanical Models of the Breast for Medical Image Analysis," *Academic Radiology*, vol. 15, no. 11, pp. 1425–1436, 2008.

[20] Y. Hu, H. U. Ahmed, Z. Taylor, C. Allen, M. Emberton, D. Hawkes, and D. Barratt, "MR to ultrasound registration for image-guided prostate interventions," *Medical Image Analysis*, vol. 16, no. 3, pp. 687–703, 2012.

[21] S. K. Kyriacou, C. Davatzikos, S. J. Zinreich, and R. N. Bryan, "Nonlinear elastic registration of brain images with tumor pathology using a biomechanical model," *IEEE Transactions on Medical Imaging*, vol. 18, no. 7, pp. 580–592, 1999.

[22] O. Clatz, M. Sermesant, P.-Y. Bondiau, H. Delingette, S. K. Warfield, G. Malandain, and N. Ayache, "Realistic Simulation of the 3D Growth of Brain Tumors in MR Images Coupling Diffusion with Mass Effect," *IEEE Transactions on Medical Imaging*, vol. 24, pp. 1334–1346, 2005.

[23] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast MR images.," *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712–721, 1999.

[24] T. Rohlfing, C. R. Maurer, D. A. Bluemke, and M. A. Jacobs, "Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint," *IEEE Transactions on Medical Imaging*, vol. 22, no. 6, pp. 730–741, 2003.

[25] M. Sdika, "A fast nonrigid image registration with constraints on the Jacobian using large scale constrained optimization," *IEEE Transactions on Medical Imaging*, vol. 27, no. 2, pp. 271–281, 2008.

[26] F. L. Bookstein, "Principal Warps: Thin-Plate Splines and the Decomposition of Deformations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 6, pp. 567–585, 1989.

[27] F. Bookstein, "Thin-plate splines and the atlas problem for biomedical images," *Information Processing in Medical Imaging*, vol. 511, pp. 326–342, 1991.

[28] C. R. Meyer, J. L. Boes, B. Kim, P. H. Bland, K. R. Zasadny, P. V. Kison, K. Koral, K. A. Frey, and R. L. Wahl, "Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations," *Medical Image Analysis*, vol. 1, no. 3, pp. 195–206, 1997.

[29] S. Schaefer, T. McPhail, and J. Warren, "Image deformation using moving least squares," *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 533–540, 2006.

[30] K. J. Friston, a. P. Holmes, K. J. Worsley, J.-P. Poline, C. D. Frith, and R. S. J. Frackowiak, "Statistical parametric maps in functional imaging: A general linear approach," *Human Brain Mapping*, vol. 2, no. 4, pp. 189–210, 1995.

[31] J. Ashburner and K. J. Friston, "Nonlinear spatial normalization using basis functions," *Human Brain Mapping*, vol. 7, no. 4, pp. 254–266, 1999.

[32] A. Roche, X. Pennec, G. Malandain, and N. Ayache, "Rigid registration of 3-D ultrasound with MR images: a new approach combining intensity and gradient information.," *IEEE Transactions on Medical Imaging*, vol. 20, no. 10, pp. 1038–1049, 2001.

[33] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, "Automated multi-modality image registration based on information theory," *Information Processing in Medical Imaging*, vol. 3, no. 6, pp. 263–274, 1995.

[34] W. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-Modal Volume Registration by Maximisation of Mutual Information," *Medical Image Analysis*, vol. 1, no. 1, pp. 35–51, 1996.

[35] J. P. W. Pluim, J. B. A. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: A survey," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 986–1004, 2003.

[36] D. Rueckert, M. J. Clarkson, D. L. G. Hill, and D. J. Hawkes, "Non-rigid registration using higher-order mutual information," *Proceedings of SPIE Medical Imaging*, vol. 5, pp. 438–447, 2000.

[37] C. Studholme, C. Drapaca, B. Iordanova, and V. Cardenas, "Deformation-based mapping of volume change from serial brain MRI in the presence of local tissue contrast change," *IEEE Transactions on Medical Imaging*, vol. 25, no. 5, pp. 626–639, 2006.

[38] T. Lindeberg, "Feature Detection with Automatic Scale Selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 79–116, 1998.

[39] T. Kadir and J. M. Brady, "Scale, Saliency and Image Description," *International Journal of Computer Vision*, vol. 45, no. 2, pp. 83–105, 2001.

[40] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[41] M. Leordeanu and M. Hebert, "A spectral technique for correspondence problems using pairwise constraints," in *Proceedings of IEEE International Conference on Computer Vision*, pp. 1482–1489, 2005.

[42] P. Besl and H. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.

[43] X. Pennec, P. Cachier, and N. Ayache, "Understanding the "Demon's Algorithm": 3D non-rigid registration by gradient descent," in *Proceedings of International Conference on Medical Image Computing And Computer-Assisted Intervention*, pp. 597–605, 1999.

[44] S. Klein, M. Staring, and J. P. W. Pluim, "Evaluation of Optimization Methods for Nonrigid Medical Image Registration Using Mutual Information and B-Splines," *IEEE Transactions on Image Processing*, vol. 16, no. 12, pp. 2879–2890, 2007.

[45] J. Ashburner and K. J. Friston, "Diffeomorphic registration using geodesic shooting and Gauss-Newton optimisation," *NeuroImage*, vol. 55, no. 3, pp. 954–967, 2011.

[46] D. Salas-Gonzalez, J. Górriz, J. Ramírez, A. Lass, and C. Puntonet, "Improved Gauss-Newton optimisation methods in affine registration of SPECT brain images," *Electronics Letters*, vol. 44, pp. 1291–1292, 2008.

[47] A. Kirov and L. Fanchon, "Pathology-validated PET image data sets and their role in PET segmentation," *Clinical and Translational Imaging*, vol. 2, pp. 253–267, 2014.

[48] T. Puri, A. Chalkidou, R. Henley-Smith, A. Roy, P. Barber, T. Guerrero-Urbano, R. Oakley, R. Simo, J.-P. Jeannon, M. McGurk, E. Odell, M. O'Doherty, and P. Marsden, "A method for accurate spatial registration of PET images and histopathology slices," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 5, no. 64, 2015.

[49] R. Garcia-Parra, D. Wood, R. Shah, J. Siddiqui, H. Hussain, H. Park, T. Desmond, C. Meyer, and M. Piert, "Investigation on tumor hypoxia in resectable primary prostate cancer as demonstrated by ^{18}F-FAZA PET/CT utilizing multimodality fusion techniques," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 38, no. 10, pp. 1816–1823, 2011.

[50] H. Park, C. Meyer, D. Wood, A. Khan, R. Shah, H. Hussain, J. Siddiqui, J. Seo, T. Chenevert, and M. Piert, "Validation of automatic target volume definition as demonstrated for ^{11}C-Choline PET/CT of human prostate cancer using multimodality fusion techniques," *Academic Radiology*, vol. 17, no. 5, pp. 614–623, 2010.

[51] H. Park, M. Piert, A. Khan, R. Shah, H. Hussain, J. Siddiqui, and C. Meyer, "Registration methodology for histological sections and *in-vivo* imaging of human prostate," *Academic Radiology*, vol. 15, no. 8, pp. 1027–1039, 2008.

[52] Y. Zhan, Y. Ou, M. Feldman, J. Tomaszeweski, C. Davatzikos, and D. Shen, "Registering histologic and MR images of prostate for image-based cancer detection," *Academic Radiology*, vol. 14, no. 11, pp. 1367–1381, 2007.

[53] M. Axente, J. He, C. Bass, G. Sundaresan, J. Zweit, J. Williamson, and A. Pugachev, "An alternative approach to histopathological validation of PET imaging for radiation therapy image-guidance: A proof of concept," *Radiotherapy and Oncology*, vol. 110, pp. 309–316, 2014.

[54] C. Meyer, B. Ma, L. Kunju, M. Davenport, and M. Piert, "Challenges in accurate registration of 3D medical imaging and histopathology in primary prostate cancer," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 40, pp. 72–78, 2013.

[55] T. Ohnishi, Y. Nakamura, T. Tanaka, T. Tanaka, N. Hashimoto, H. Haneishi, T. Batchelor, E. Gerstner, J. Taylor, M. Snuderl, and Y. Yagi, "Deformable image registration between pathological images and MR image via an optical macro image," *Pathology - Research and Practice*, vol. 212, no. 10, pp. 927–936, 2016.

[56] H. R. Roth, K. Nagara, H. Oda, M. Oda, T. Sugiyama, S. Nakamura, and K. Mori, "Multi-scale image fusion between pre-operative clinical CT and X-ray microtomography of lung pathology," in *Proceedings of International Forum on Medical Imaging*, 2017.

[57] J. Stroom, H. Blaauwgeers, A. van Baardwijk, L. Boersma, J. Lebesque, J. Theuws, R.-J. van Suylen, H. Klomp, K. Liesker, R. van Pel, C. Siedschlag, and K. Gilhuijs, "Feasibility of pathology-correlated lung imaging for accurate target definition of lung tumours," *International Journal of Radiation Oncology*, vol. 69, no. 1, pp. 267–275, 2007.

[58] J. van Loon, C. Siedschlag, J. Stroom, H. Blaauwgeers, R.-J. van Suylen, J. Knegjens, M. Rossi, A. van Baardwijk, L. Boersma, H. Klomp, W. Vogel, S. Burgers, and K. Gilhuijs, "Microscopic disease extension in three dimensions for non-small-cell lung cancer: development of a prediction model using pathology-validated positron emission tomography and computed tomography features," *International Journal of Radiation Oncology*, vol. 82, no. 1, pp. 448–456, 2012.

[59] M. Dahele, D. Hwang, C. Peressotti, L. Sun, M. Kusano, S. Okhai, G. Darling, M. Yaffe, C. Caldwell, K. Mah, J. Hornby, L. Ehrlich, S. Raphael, M. Tsao, A. Behzadi, C. Weigensberg, and Y. C. Ung, "Developing a methodology for three-dimensional correlation of PET-CT images and whole-mount histopathology in non-small-cell lung cancer," *Current Oncology*, vol. 15, no. 5, pp. 62–69, 2008.

[60] K. Wu, Y. Ung, J. Hornby, M. Freeman, D. Hwang, M. Tsao, M. Dahele, G. Darling, D. Maziak, R. Tirona, K. Mah, and C. S. Wong, "PET CT thresholds for radiotherapy target definition in non-small-cell lung cancer: how close are we to the pathologic findings?," *International Journal of Radiation Oncology*, vol. 77, no. 3, pp. 699–706, 2010.

[61] J. Yu, X. Li, L. Xing, D. Mu, Z. Fu, X. Sun, X. Sun, G. Yang, B. Zhang, X. Sun, and C. Ling, "Comparison of tumor volumes as determined by pathologic examination and FDG-PET/CT images of non-small-cell lung cancer: a pilot study," *International Journal of Radiation Oncology*, vol. 75, no. 5, pp. 1468–1474, 2009.

[62] X. Meng, X. Sun, D. Mu, L. Xing, L. Ma, B. Zhang, S. Zhao, G. Yang, F.-M. Kong, and J. Yu, "Noninvasive evaluation of microscopic tumor extensions using standardized uptake value and metabolic tumor volume in non-small-cell lung cancer," *International Journal of Radiation Oncology*, vol. 82, no. 2, pp. 960–966, 2012.

[63] M. Wanet, J. A. Lee, B. Weynand, M. D. Bast, A. Poncelet, V. Lacroix,

E. Coche, V. Gregoire, and X. Geets, "Gradient-based delineation of the primary GTV on FDG-PET in non-small cell lung cancer: A comparison with threshold-based approaches, CT and surgical specimens," *Radiotherapy and Oncology*, vol. 98, pp. 117–125, 2011.

[64] A. Schaefer, Y. J. Kim, S. Kremp, S. Mai, J. Fleckenstein, H. Bohnenberger, H.-J. Schaefers, J.-M. Kuhnigk, R. Bohle, C. Ruebe, C.-M. Kirsch, and A. Grgic, "PET-based delineation of tumour volumes in lung cancer: comparison with pathological findings," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 40, pp. 1233–1244, 2013.

[65] J. Y. Fu, Y. L. Wan, T. Y. Huang, C. F. Wu, Y. H. Liu, M. J. Hsieh, Y. C. Wu, and C. Y. Wu, "Correlation between image characteristics and pathologic findings in non small cell lung cancer patients after anatomic resection," *PLoS ONE*, vol. 13, no. 10, p. e0206386, 2018.

[66] C. H. Park, T. H. Kim, S. Lee, D. H. Moon, and H. S. Park, "Correlation between maximal tumor diameter of fresh pathology specimens and computed tomography images in lung adenocarcinoma," *PLoS ONE*, vol. 14, no. 1, p. e0211141, 2019.

[67] OpenStax College (2013), "Anatomy and Physiology." Available at: `http://cnx.org/content/col11496/latest/`, accessed June 2019.

[68] P. Dendy and B. Heaton, *Physics for diagnostic radiology*. Boca Raton: Taylor and Francis Group, 3rd ed., 2012.

[69] S. Vaseghi, *Advanced Digital Signal Processing and Noise Reduction*. Chichester: John Wiley and Sons, 4th ed., 2000.

[70] J. Lambers, "Piecewise polynomial interpolation." Lecture notes, MAT 460/560, University of Southern Mississipi. Available at: `https://www.math.usm.edu/lambers/mat460/fall09/lecture20.pdf`, accessed July 2019.

[71] G. Herman, J. Zheng, and C. Bucholtz, "Shape-based interpolation," *IEEE Computer Graphics and Applications*, vol. 12, no. 3, pp. 69–79, 1992.

[72] S. P. Raya and J. Udupa, "Shape-based interpolation of multidimensional objects," *IEEE Transactions on Medical Imaging*, vol. 9, no. 1, pp. 32–42, 1990.

[73] X. Huang, J. Ren, G. Guiraudon, D. Boughner, and T. M. Peters, "Rapid dynamic image registration of the beating heart for diagnosis and surgical navigation," *IEEE Transactions on Medical Imaging*, vol. 28, no. 11, pp. 1802–1814, 2009.

[74] A. Goshtasby, *Image Registration. Principles, Tools and Methods*. London: Springer-Verlag, 1st ed., 2012.

[75] J. Phillips, "Iterative Closest Point and Earth Mover's Distance." Lecture notes, CPS296.2 Geometric Optimization, Duke University. Available at: `https://www2.cs.duke.edu/courses/spring07/cps296.2/scribe_notes/lecture24.pdf`, accessed July 2019.

[76] K. Arun, T. Huang, and S. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 5, pp. 698–700, 1987.

[77] I. Jolliffe, *Principal Component Analysis*. New York: Springer-Verlag, 2nd ed., 2002.

[78] E. Dura and J. Domingo, "Alternative method for binary shape alignment of non-symmetrical shapes based on minimal enclosing box," *Electronics Letters*, vol. 48, no. 22, pp. 1401–1402, 2012.

[79] J. Korsawe, "Minimal Bounding Box Toolbox." Available at: `https://www.mathworks.com/matlabcentral/fileexchange/18264-minimal-bounding-box`, accessed October 2019.

[80] J. Schindelin, "Moving Least Squares registration (ImageJ)." Available at `https://github.com/fiji/VIB/blob/master/src/main/java/Moving_Least_Squares.java`, accessed April 2019.

[81] J. M. Fitzpatrick and J. B. West, "The distribution of target registration error in rigid-body point-based registration," *IEEE Transactions on Medical Imaging*, vol. 20, no. 9, pp. 917–927, 2001.

[82] A. Zijdenbos, B. Dawant, R. Margolin, and A. Palmer, "Morphometric analysis of white matter lesions in MR images: method and validation," *IEEE Transactions on Medical Imaging*, vol. 13, no. 4, pp. 716–724, 1994.

[83] G. Rote, "Computing the minimum Hausdorff distance between two point sets on a line under translation," *Information Processing Letters*, vol. 38, pp. 123–127, 1991.

[84] M. Ashby and K. Johnson, *Materials and Designs: the Art and Science of Material Selection in Product Design*. Burlington: Elsevier, 1st ed., 2002.

[85] M. Nagura, T. Hamano, and H. Ishikawa, "Structure of poly(vinyl alcohol) hydrogel prepared by repeated freezing and melting," *Polymer*, vol. 30, pp. 762–765, 1989.

[86] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. Aylward, J. Miller, S. Pieper, and R. Kikinis, "3D Slicer as an Image Computing Platform for the Quantitative Imaging Network," *Magnetic Resonance Imaging*, vol. 30, no. 9, pp. 1323–1341, 2012.

[87] R. Adams and L. Bischof, "Seeded Region Growing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 6, pp. 641–647, 1994.

[88] Health and Safety Executive UK, *The Control of Substances Hazardous to Health (COSHH) Regulations 2002*. London: The Stationery Office, 2013.

[89] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis," *Nature Methods*, vol. 9, pp. 671–675, 2012.

[90] F. van Velden, G. Kramer, V. Frings, I. Nissen, E. Mulder, A. de Langen, O. Hoekstra, E. Smit, and R. Boellaard, "Repeatability of Radiomic Features in Non-Small-Cell Lung Cancer [^{18}F]FDG PET/CT Studies: Impact of Reconstruction and Delineation," *Molecular Imaging and Biology*, vol. 18, pp. 788–795, 2016.

[91] C. Liu, L. Pierce, A. Alessio, and P. Kinahan, "The impact of respiratory motion on tumor quantification and delineation in static PET/CT imaging," *Physics in Medicine and Biology*, vol. 54, pp. 7345–7362, 2009.

[92] A. Chi and N. Nguyen, "4D PET/CT as a strategy to reduce respiration motion artifacts in FDG-PET/CT," *Frontiers in Oncology*, vol. 4, no. 8, 2014.