UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES
School of Electronics and Computer Science

Non-rigid Registration for Multi-modality Image Fusion Using Prior Shapes

by

Zheng Cui

Primary Supervisor: Dr. Sasan Mahmoodi
Co-supervisors: Dr. Michael Bennett, Dr. Emma Lewis, Dr. Matthew Guy

Internal examiner: Dr. Srinandan Dasmahapatra
External examiner: Prof. Timothy Cootes

Thesis for the degree of Doctor of Philosophy

August, 2018

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES
School of Electronics and Computer Science

Doctor of Philosophy

NON-RIGID REGISTRATION FOR MULTIMODALITY IMAGE FUSION USING PRIOR SHAPES

By Zheng Cui

Chronic obstructive pulmonary disease (COPD) is a chronic lung disease that causes breathing difficulties. One possible course of treatment for severe COPD is lung volume reduction surgery (LVRS), which involves removing, or isolating, the lobe or lobes of the lung that are most affected by the disease. A fusion of the multi-slice computed tomography (MSCT) and ventilation (V) and perfusion (Q) single photon emission computed tomography (SPECT) modalities therefore represents a powerful tool to for COPD analysis and then for guiding the lung resection surgery.

Due to reduced uptake of radioisotope at the location of lesion, the V and Q of a moderate COPD patient delineate photopenic regions, which are normally mis-recognised as part of the background in the target SPECT scan. Non-rigid registration, which lacks displacement constraints, is therefore performed on MSCT scans with excessive deformations. Moreover, considering the low-resolution nature of functional imaging and highly deformable property of lungs, very few published algorithms are able to accommodate current clinical demands. The motivation of this project is to

develop a high-performance, statistical deformation model (SDM)-based non-rigid registration algorithm capable of achieving accurate alignment of lung MSCT and SPECT imaging.

In this project, an innovative similarity registration method for volumetric shapes is proposed at the beginning. The method is based on the characteristic function, and intended to strike a desirable balance between performance and efficiency. Radial moments and spherical coordinate system-based cross-correlation are exploited here to obtain the optimal scaling, rotation and translation parameters within a reasonable time. Moreover, an iterative method is also employed to improve the robustness of the algorithm. Group shapes in the presence of significant noise and lung shapes extracted from a low-dose computed tomography database are employed in the validation experiments.

In order to eliminate the influence of the weighting parameter for the statistical term, a novel MSCT/SPECT registration technique based on a parameter-reduced SDM is proposed in this thesis. The SDM is trained on prior lung shapes. In addition, the multi-channel technique performs V/MSCT and Q/MSCT alignments simultaneously to derive the optimal deformations. Lung MSCT and SPECT imaging data from a real medical database, as well as the 4D extended cardiac-torso phantom, were employed in the experiments. The algorithm proposed here was validated to be capable of preventing excessive deformations, and of achieving accurate registration between the two imaging modalities. The deformations for MSCT/SPECT registration are finally used to warp lobe masks, which are then mapped onto SPECT images for lung lobe/SPECT fusion.

Contents

UNIVERSITY OF Southampton

Research Thesis: Declaration of Authorship

Print name:	**Zheng Cui**

Title of thesis:	**Non-rigid Registration for Multi-modality Image Fusion Using Prior Shapes**

I declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. **This work was done wholly or mainly while in candidature for a research degree at this University;**
2. **Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;**
3. **Where I have consulted the published work of others, this is always clearly attributed;**
4. **Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;**
5. **I have acknowledged all main sources of help;**
6. **Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;**
7. **Either none of this work has been published before submission, or parts of this work have been published as:**

   Z. Cui, S. Mahmoodi, J. Conway, M. Guy, E. Lewis, T. Havelock and M. Bennett, "A Novel Non-Rigid Registration Method Based on Nonparametric Statistical Deformation Model for Medical Image Analysis," accepted in IEEE Nuclear Science Symposium and Medical Imaging Conference, 2017.

   Z. Cui, S. Mahmoodi, M. Guy, E. Lewis, T. Havelock, M. Bennett and J. Conway, "A General Single and Multi-modality Registration Framework for Lung Imaging Analysis Using Statistical Prior Shapes," submitted to Pattern Recognition, 2018.

   Z. Cui, S. Mahmoodi and M. Bennett, "A Robust and High-performance Shape Registration Technique Based on Characteristic Functions," submitted to IEEE International Conference on Image Processing, Applications and Systems, 2018.

Signature:		**Date:**	

Acknowledgement

Firstly, I would like to express my sincere gratitude to my supervisory team, Dr. Sasan Mahmoodi, Dr. Michael Bennett, Dr. Emma Lewis and Dr. Matthew Guy. Under their considerate supervision, I managed to accomplish my Ph.D. research work smoothly and in a timely fashion. In the past three years, I have benefited greatly in terms of not only knowledge and skills but also research ability and academic temperament. Here, I want to give my particular gratitude to my primary supervisor, Dr. Sasan Mahmoodi, who has been a knowledgeable, responsible and dependable supervisor of my studies. He is also warm-hearted and cares for my life in Southampton. I also appreciate the academic guidance and support provided by Michael, Emma and Matt. Their valuable experience and constructive comments helped develop my research ability in the field of medical image analysis. I would also like to acknowledge that the Faculty of Physical Sciences and Engineering, University of Southampton, provided research facilities and full research funding for me.

The MSCT, low-dose CT and SPECT data used in this work were acquired as a part of a study into the application of imaging to the characterisation of the phenotypes of COPD. Written informed consent was given and signed by all subjects. The study was approved by the Southampton and South West Hampshire local research ethics committee (LREC number: 09/H0502/91) and the University Hospital Southampton Foundation Trust Research and Development Department. The study was conducted in the Southampton NIHR Respiratory Biomedical Research Unit. In addition, I sincerely appreciate that Dr. W. Paul Segars from Duke University (U.S.) generously provided the XCAT software for data collection.

Moreover, I would like to give my gratitude to the staff in the Vision, Learning and Control Group as well as my colleagues and friends. They are Prof. Mark Nixon, Prof. Mahesan Niranjan, Dr. Bing Chu, Dr. Wanmu Liu, Mr. Jun Yan Ho, Dr. Daniel Martinho-Corbishley and Mr. Thomas Ladyman. Moreover, I appreciate the fact that two internal examiners, Dr. Srinandan Dasmahapatra and Dr. John Carter, and the external examiner, Prof. Timothy Cootes, offered lots of constructive comments for my upgrade report and final thesis.

Last but not least, I express my endless gratitude to my parents, Mrs. Linchun Cao and Mr. Baoquan Cui, who encouraged me to go abroad and pursue my dreams. Also, I feel very lucky and happy to have met my true love, Ms. Bingchen Guo, at the University of Southampton.

Nomenclature

Abbreviations

CF	characteristic function
COPD	chronic obstructive pulmonary disease
CT	computed tomography
FFD	free-form deformation
FFT	fast Fourier transform
ICP	iterative closest point
KDE	kernel density estimation
KL	Kullback-Leibler
LVRS	lung volume reduction surgery
MDE	mean displacement error
MI	mutual information
MSCT	multi-slice computed tomography
NIP	normalised inner product
NoE	number of dominant eigenmodes
NPP	normal probability plot
NURBS	non-uniform rational B-spline
PCA	principal component analysis
PET	positron emission tomography
Q	perfusion
RMDP	radionuclide multi-modality dosimetry package
ROI	regions of interest
SDF	signed distance function
SDM	statistical deformation model
SNR	signal-to-noise ratio

SPECT	single photon emission computed tomography
SSD	sum-of-squared distance
SUB.	subject
V	ventilation
VD	vector displacement
XCAT	extended cardiac-torso
e.g.	exempli gratia, for example
i.e.	id est, that is (to say)
Ref	the reference method proposed in [61]

Notations for Chapter 2

ϕ	CF of a volumetric shape
Ω	3-D image domain
S	region inside a shape
E	distance between two CFs
ϕ_{tar}	CF of a target shape
ϕ_{src}	CF of a source shape
s	scaling parameter
$\varphi_x, \varphi_y, \varphi_z$	rotation parameters around X, Y and Z axes respectively
$\Delta x, \Delta y, \Delta z$	translations along the X, Y and Z axes respectively
$\overline{\phi_{src}}$	CF of a registered source shape
$\widetilde{\phi_{tar}}$	CF of a target shape after centralisation
$\widetilde{\phi_{src}}$	CF of a source shape after coarse registration
R_c	coarse rotation of the source shape
(x_t, y_t, z_t)	centroid of a target shape
(x_s, y_s, z_s)	centroid of a source shape
M_m^{tar}, M_m^{src}	m-order radial moments of the target and source shapes

M	maximum order of the radial moments
E_{scl}	distance between two radial moments
E_{rot}	distance between two CFs after fine rotation
C	constant
R_f	fine rotation of the source shape
r, θ, φ	radial distance, polar angle and azimuthal angle respectively in spherical coordinate system
$\widetilde{\phi_{s1}}, \widetilde{\phi_{s2}}, \widetilde{\phi_{s3}}$	CFs of the source shapes mapped to spherical coordinate system using the Z, Y and X axes as the zenith direction respectively
$\widetilde{\phi_{t1}}, \widetilde{\phi_{t2}}, \widetilde{\phi_{t3}}$	CFs of the target shapes mapped to spherical coordinate system using the Z, Y and X axes as the zenith direction respectively
η, λ, ξ	spatial frequencies for r, θ, φ respectively in the spherical coordinate system
E_{tran}	distance between two CFs after fine translation
$\widetilde{\Phi_{tar}}, \widetilde{\Phi_{src}}$	Fourier transform of $\widetilde{\phi_{tar}}$ and $\widetilde{\phi_{src}}$
η', λ', ξ'	spatial frequencies for x, y, z respectively in the Cartesian coordinate system
$E_{IP}, E_{Ac}, E_{Ov}, E_{Se}, E_{Sp}$	similarity metrics with respect to normalised inner product, accuracy, overlap, sensitivity and specificity
V_Ω	volume of the image domain
V_c	volume of the common region inside the two shapes
V_o	volume of the background region outside the two shapes
V_t	volume of the region inside the target shape but outside the registered source shape
V_r	volume of the region inside the registered source shape but outside the target shape
E_{HD}	Hausdorff distance between two shapes
i	iteration number

Notations for Chapter 3

X	SDF of a prior shape

$\bar{X}$	SDF of a mean prior shape
N	number of prior shapes
$\boldsymbol{u}$	deformation vectors for non-rigid registration
$\bar{\boldsymbol{u}}$	mean deformation vectors
$\boldsymbol{\mu}$	column vectors reshaped from deformation vectors
$\boldsymbol{F}$	a combined matrix of column vectors
M	number of voxels in a shape matrix
$\boldsymbol{V}, \boldsymbol{\Sigma}$	eigenvectors and eigenvalues respectively
$\boldsymbol{\sigma}$	standard deviations calculated from eigenvalues
$\boldsymbol{u}_{\text{model}}$	deformation vectors built using prior knowledge
$\boldsymbol{\lambda}$	values of coordinates along eigenvectors
E	cost function for SDM-based registration
D	distance term
R	smoothing term
S	statistical term
α, β	weighting factors in a cost function
c	constant
I_{tar}, I_{src}	target and source images respectively
K	the number of eigenvectors used for regularisation
H	Heaviside function
$Sign$	sign function
ε	regularising parameter

Notations for Chapter 5

MI	mutual information between two images
En	Entropy
p	probability

I_{tar}, I_{src}	target and source images respectively
$\boldsymbol{u}$	deformation vectors for non-rigid registration
a, b	greyscale values used in I_{tar} and I_{src} respectively
E_{MI}	inverse MI
V	volume of an image
σ_{tar}, σ_{src}	kernel widths of I_{tar} and I_{src} respectively
K	kernel function
C	constant
$\boldsymbol{\lambda}$	values of coordinates along eigenvectors
S	statistical term
E_{VQ}	multi-channel cost function
E_V	similarity term for MSCT/V registration
E_Q	similarity term for MSCT/Q registration
$\boldsymbol{T}$	ground-truth displacement vectors

To my heavenly father…

Chapter 1

Introduction

1.1 Background

Chronic obstructive pulmonary disease (COPD), including chronic bronchitis and emphysema, is a significant global cause of morbidity and mortality that presents a serious healthcare burden [1]. According to a report from the World Health Organisation, over three million patients died in 2015 due to COPD and the mortality rate has continuously increased in the past few years. Despite many years of research, there are no current pharmacological therapies which significantly alter disease progression. One reason for this is that COPD is not a single disease entity, but a collection of disease phenotypes with a common aetiology and different pathophysiological processes [2].

Nevertheless, early diagnosis and treatment can effectively relieve the symptoms, enhance quality of life and lower the risk of death. Lung activity analysis can be used to identify lobes of the lung that are severely affected and are failing to deliver an appropriate fraction of total lung function. One possible course of treatment for severe COPD is lung volume reduction surgery (LVRS), which involves removing the parts of the lung that are most affected by the disease, often a whole lobe or more. This helps the body to increase the volume of the remaining lobes, thereby improving breathing [3]. In order for this procedure to be effective, it is necessary to know precisely which parts of the lung have been most affected by the disease and therefore which are the best candidates for removal.

Computed tomography (CT) scans are obtained with a combination of an X-ray tube and an arc sensor, which are deployed in a fanbeam geometry on either side of the patient and rotate rapidly together to acquire the cross-sectional images. The multi-slice computed tomography (MSCT) scanner was created to improve imaging resolution (512×512 voxels in the XY plane; 0.5mm between the slices along the Z axis) to support clinical diagnosis. The anatomical planes are illustrated in Figure 1.1. In addition, the MSCT scanner is able to capture multiple slices simultaneously, and then to provide fully three-dimensional imaging of lung structure, which allows the lung lobes and other structural features to be identified [4]. The optimal MSCT parameters for the delineation of the various structural features of the lung require suspended full

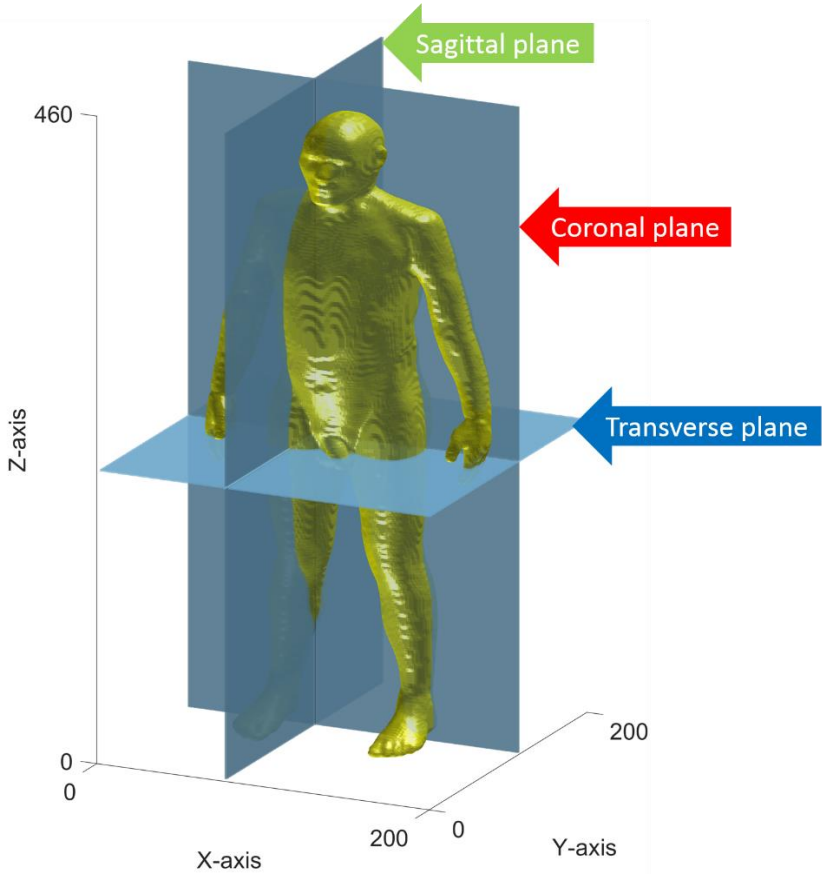


Figure 1.1 *Diagram of anatomical planes. The x-axis runs from left (positive number) to right (0). The y-axis runs from posterior (positive number) to anterior (0). The z-axis runs from head (positive number) to tail (0).*

inspiration, as shown in Figure 1.2. In order to depict the internal structure more clearly, the dynamic range of the lung region is enhanced in Figure 1.2(c) and (d).

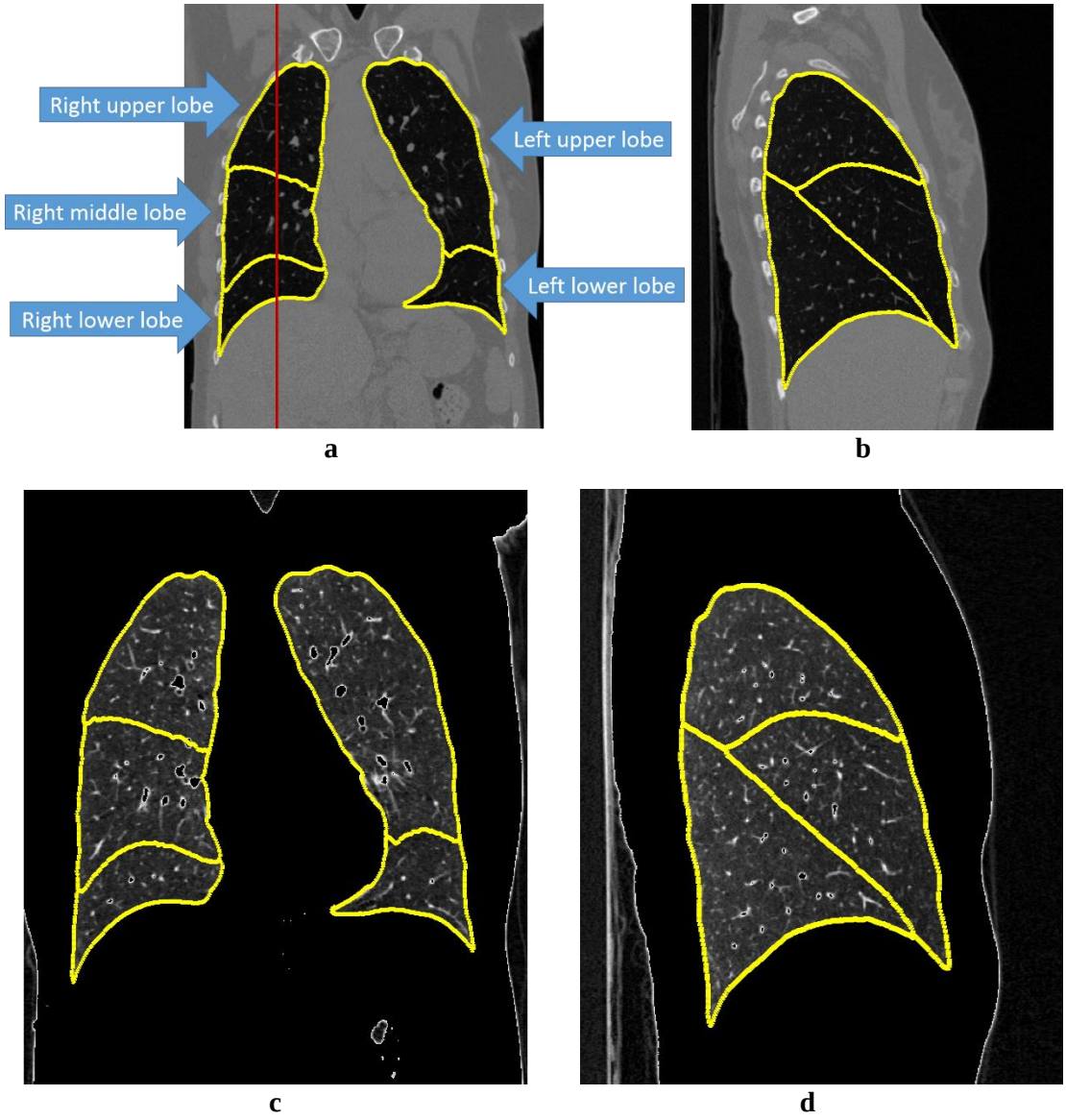


Figure 1.2 *MSCT is displayed in two views with the boundaries of lobes highlighted in yellow. (a) Coronal slice with labels for five lung lobes. (b) Sagittal slice of the right lung. The location of the slice plane is indicated by the red line in (a). (c) Coronal slice of the lung region in high dynamic range. (d) Sagittal slice of the lung region in high dynamic range.*

Nuclear medicine imaging techniques, such as single photon emission computed tomography (SPECT), require the injection of a gamma-emitting radioisotope into the patient. During the decay of the isotope, gamma-ray photons are emitted. The photons travel through a collimator and are detected by gamma cameras which are rotating

around the patient for data acquisition [5]. The planar projection images collected at different angles are then used to reconstruct a true three-dimensional SPECT image to identify the radioactivity of the organ specimen [4]. As the collimator is designed to be in front of the camera to determine the propagating direction of the gamma-rays, the imaging quality (e.g. resolution and sensitivity) is significantly dependent on it. In order to improve the resolution, the range of incident angles is restricted by the collimator, whereas it sacrifices some image sensitivity due to the decrease in the flux (i.e. it allows fewer photons to go through the collimator) [6]. Selection of a collimator needs to trade off between resolution and sensitivity. In addition, SPECT imaging is always subjected to photon attenuation caused by the body tissues, through which the gamma-rays pass before getting to the camera. Another reason for attenuation is Compton scattering, which alters the photon's directions of travel [4]. In this situation, numerous photons which fail to be detected by the camera result in lower signal-to-noise ratio and reduced contrast. On the other hand, the scattering of photons into the wrong detector leads to blurring of the image and loss of spatial resolution. Therefore, low-dose CT/SPECT matching (currently implemented by a CT/SPECT hybrid scanner) is employed for attenuation correction. The CT data forms a map of the different attenuation levels of the tissues in the body, and this map is then used to correct the SPECT image by 'adding photon counts back' or 'subtracting photon counts'.

For specific applications such as lung activity imaging, ventilation (V) and perfusion (Q) SPECT scans employ different medical isotopes and delivery methods to evaluate the circulation of air and blood in the lungs, e.g. the radionuclide Tc-99m (technetium) is injected into the patient for perfusion while Kr-81 (krypton gas) is inhaled by the patient for ventilation. As the overall SPECT scan time is approximate 20 minutes, it is not reasonable to ask the patient to hold his breath throughout. The low-dose CT and

SPECT scans are effectively acquired during free-breathing, which is at a different stage of the breathing cycle compared with the MSCT scans. The SPECT scan allows three-dimensional imaging (voxel size: 4.418mm^3, $128 \times 128 \times 128$ voxels per imaging matrix) of lung function, which makes it possible to identify regions of the lung that are not functioning as they should. As stated in [7], the areas of airway closure in the COPD patients are formed by the presence of emphysematous bulla and cysts. The deflation and compression of lung tissue contribute to the lung destruction and severe vessel and airway narrowing, which cause abnormal ventilation and perfusion within the lungs. SPECT V/Q images for normal and moderate COPD subjects are illustrated in Figure 1.3.

The combination of the structural information from MSCT with the functional information from SPECT represents a powerful tool for the precise identification of regions of the lung that are optimal targets for lung resection surgery. In order to achieve this, it is therefore necessary to develop a tool that is able to compute the deformation required to align the structural and functional imaging information, and to apply these deformable fields to regions and features of the lung that have been identified. The broad requirements for such a tool are that it should be able to achieve sufficiently accurate registration between the imaging modalities, even in cases where the lung imaging data are acquired at a different state of inspiration. Also, a suitable method should strike a balance with the requirements for relatively high performance in terms of processing time, since, to be of any potential clinical value, the software would need to be integrated seamlessly into the radiologist's workflow.

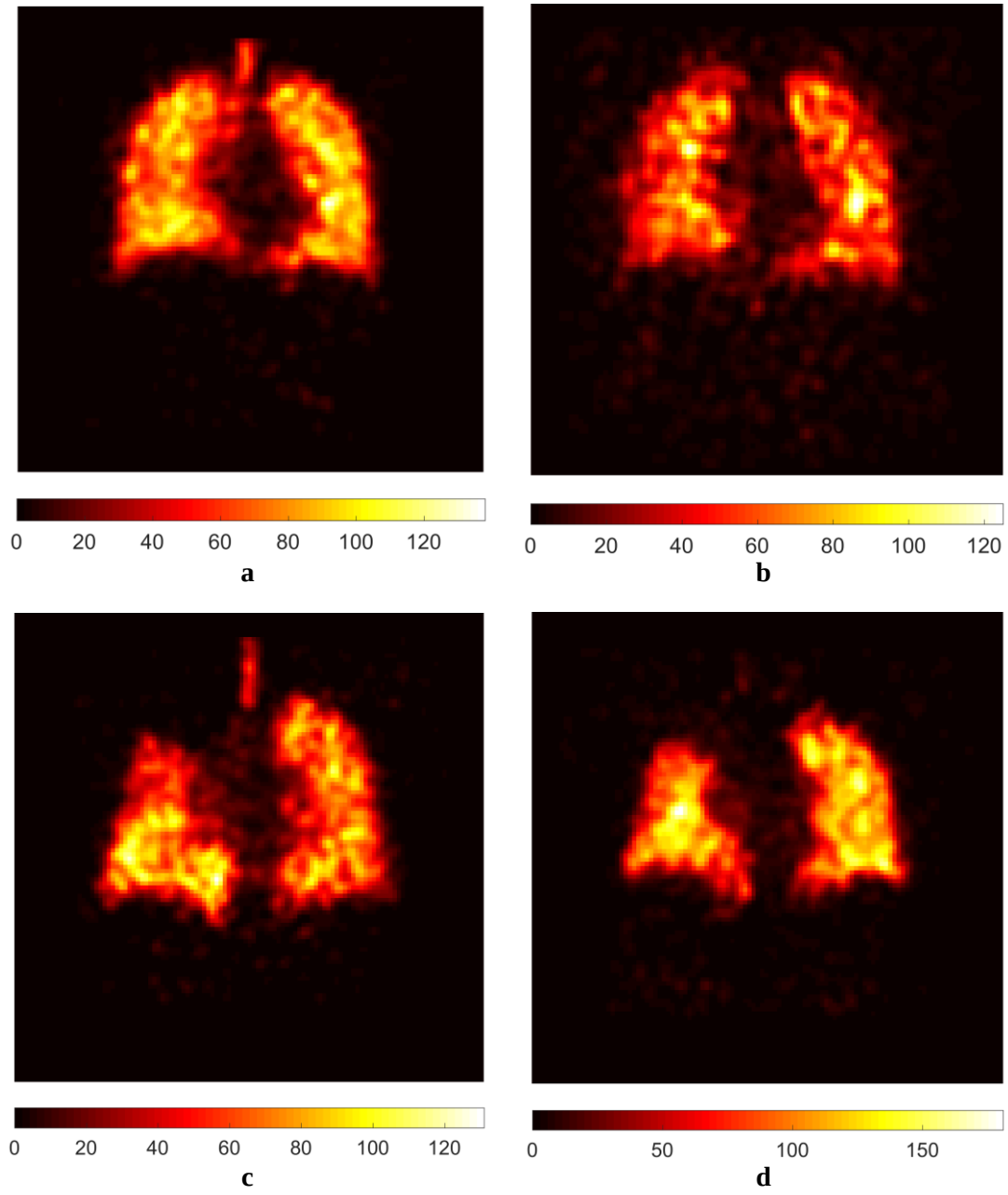


Figure 1.3 *SPECT V and Q scans of normal and COPD subjects are displayed in coronal view. The trachea is visible in V (left subfigures) but not in Q (right subfigures). (a) V image of a normal subject. (b) Q image of a normal subject. (c) V image of a COPD subject. (d) Q image of a COPD subject. Due to severe COPD, the right upper lobe shows photopenic regions in both (c) and (d) in comparison with the normal subject.*

1.2 Previous Work

The hybrid low-dose CT and SPECT scanning technique was first proposed in [8], and then developed by [9] and [10]. A hybrid CT/SPECT-based registration method for brain images was proposed in [11], which employs CT to intervene MRI/SPECT

alignment. The alignment between MRI and SPECT is dependent on the result of CT and MRI registration, since CT and SPECT were assumed to be aligned during data acquisition. This hybrid technique manages to solve the problem of low-dose CT/SPECT registration, whilst it is based on the assumption that the patient remains stationary throughout the scanning process. In fact, motion of some organs (e.g. respiratory, cardiac, digestive and muscular organs) is inevitable and therefore the acquired data are always mismatched.

In order to tackle the multi-modality (e.g. CT/SPECT) registration problem in software, the concept of four-dimensional (4D) imaging was proposed in [12] to represent the variations of three-dimensional imaging with the passage of time. The breathing protocol in 4D imaging can be established by respiratory gating or motion tracking techniques. This method aims to decrease the influence of artefacts. The method proposed in [13] intends to facilitate multi-modality registration by converting SPECT slices into CT-like images using a scatter window. The ITK diffeomorphic Demons method was then used to complete registration work. Furthermore, for the case that the isotope in the adjacent objects is more straightforward to detect, the methods reported in [14] performs registration for tumours based on the spine SPECT imaging. Similar methods were studied in [15] and [16] as well. A registration method for two-dimensional hand X-ray images was proposed in [15]. It begins with rigid registration of the hand skeleton, followed by non-rigid registration of tissues surrounding the bones. The three-step SPECT/CT registration method proposed in [16] performs bone, surface and tissue registration sequentially using an extended Demons algorithm to achieve high performance. Nevertheless, this technique concentrates on the level-set representation of shapes and employs the distance map to calculate transformation parameters, which is based on accurate segmentation. In addition, intensity uncertainty

quantification was used to represent each voxel in [17]. Since the voxel mapping is independent of intensity, it overcomes the shortcoming that fluid-like registration methods (e.g. Demons) are not applicable to multi-modality image registration. A metric learning method was introduced in [18]. Gabor feature vectors were also produced by learning algorithm to boost the performance of cross-modality registration. Recently, a series of sophisticated local descriptors based on the features and textures have been studied in [19] and [20]. Euclidian and Riemannian distances are employed to measure the correspondence similarity respectively.

Nevertheless, as few features can be extracted from SPECT imaging and landmark correspondence is practically unachievable, feature-based methods are not applicable to lung MSCT/SPECT alignment. In addition, lungs have a highly deformable nature and the MSCT and SPECT images are captured at a different state of inspiration. External rigid reference objects (e.g. bones and vessels) hardly contribute to the multi-modality registration. On the other hand, prior knowledge-based methods are considered in the approach proposed in this thesis as non-learning methods only concentrate on the registration similarity, but fail to maintain the structural information during deformation. To the best of my knowledge, very few articles have been published to tackle the complicated lung MSCT/SPECT registration problem with prior knowledge. The motivation of this project is therefore to develop a high-performance, non-rigid registration algorithm capable of achieving accurate alignment of lung MSCT/SPECT imaging.

1.3 Research Overview

In terms of the intra-patient MSCT/SPECT registration, the main challenge is to prevent excessive deformation of the source image (i.e. MSCT), as the target image (i.e. SPECT)

is usually deteriorated by artefacts and outliers (e.g. images containing tumour or lesion tissue can be considered as outliers) [21]. Therefore, the key technique exploited here aims to fuse lung MSCT and SPECT images for individuals based on a statistical deformation model (SDM) [22]. Compared with inter-subject alignment, intra-patient lung imaging registration can be developed using the rules of the respiratory pattern. In the lung database, the imaging data for each patient are comprised of MSCT scans, a pair of SPECT (V and Q) scans and low-dose CT scans (140kV, 2.5mA, 2.6rpm). As prior knowledge for investigating the rules of lung deformation associated with the respiratory pattern is very scarce, inter-subject information extracted from various subjects (e.g. healthy non-smoker, healthy smoker, mild COPD patient and moderate COPD patient) is employed here to train the SDM. Detailed participant information is listed in Table A-3 and A-4 of Appendix A. As the inter-subject information can only provide a global constraint against excessive deformation, rather than intra-patient voxel-wise regularisation, prior lung shapes represented by level-sets are employed for convenience.

The prior lung shapes are pre-processed by the similarity registration method proposed in Chapter 2. In this chapter, the robustness of the registration method is enhanced by an iterative scheme. In each iteration, radial moments, cross-correlation and spherical coordinates are employed to accelerate convergence and avoid local minimum. The high performance in terms of registration accuracy, stability and convergence speed is validated by experiments in the presence of noise, as well as a batch-processing test using the lung database. Compared with state-of-the-art algorithms, the characteristic function (CF)-based method manages to strike a desirable balance between performance and efficiency.

The research into non-rigid registration begins with statistical modelling using principal component analysis (PCA). The shortcomings in previous methods, which make the deformation unreasonable and inflexible, are revealed in Chapter 3. The parameter-reduced SDM is then proposed to end the controversy of the weighting parameter for the statistical term, and to properly regularise the deformable field based on prior shapes. In this chapter, single-modality images are employed to test the proposed algorithm and other state-of-the-art SDM-based methods for comparison. The outstanding performance achieved by the framework proposed here is demonstrated.

In order to enrich the testing data and evaluation methods, a 4D extended cardiac-torso (XCAT) phantom and radionuclide multi-modality dosimetry package (RMDP) are employed in this thesis to generate synthetic lung imaging data and the ground-truth displacement. In Chapter 4, the process of phantom construction is explicitly introduced. It also highlights the procedure of adding artificial defects to normal phantoms, which is used to challenge the learning-based algorithms in the next chapter.

In Chapter 5, the parameter-reduced SDM framework is extended for multi-modality image registration. Instead of the sum-of-squared distance (SSD) used for the evaluation of single-modality alignment, mutual information (MI), in conjunction with gradient-descent optimisation, constitutes the framework for multi-modality registration. In addition, a multi-channel registration method is reported in this chapter to improve the performance of lung MSCT/SPECT registration. It conducts a direct region-to-region registration, which is independent of non-lung landmarks that guide the registration algorithms. Simulations using the actual lung database and synthetic phantoms are performed, and the results are measured with different metrics to demonstrate the improvement achieved by the algorithm proposed here.

The entire research progression of the PhD project is illustrated in Figure 1.4 and summarised in Chapter 6. It highlights the breakthroughs and achievements made by the algorithm proposed in this thesis for medical imaging analysis, and emphasises the applied value of the algorithm with respect to computer vision and pattern recognition. It also states the promising work which is expected to be accomplished in the future to further improve the framework and then aid the diagnosis and treatment of pulmonary disease.

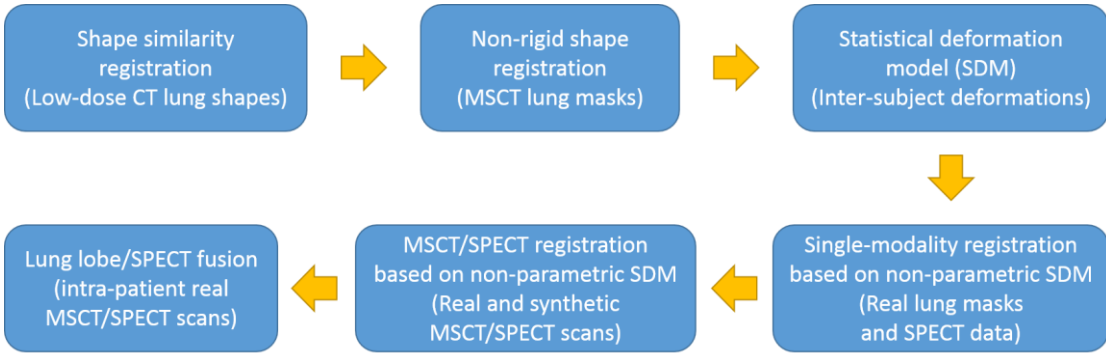


Figure 1.4 *Flow chart of research progression. Each step is described by a milestone, with the relevant data in parentheses.*

Chapter 2

Three-dimensional Shape Similarity Registration

2.1 Introduction

As a necessary initial step for non-rigid registration, similarity registration, which is performed by geometric transformation based on the similarity (e.g. scaling, rotation and translation), should be performed first to create the optimum conditions for the next non-rigid registration step. This helps improve the rate of convergence of the non-rigid registration algorithm and eliminates at least some of the potential local minimum into which the algorithm might otherwise fall. Previous research into shape registration based on particular shape representations (explicit and implicit descriptors, Laplace-Beltrami spectral descriptors and point clouds), advanced optimisation methods (iterative closest point and cross-correlation) and classic similarity measurements (SSD and MI) has been conducted in the past few years.

Feature-based methods are reported in [23] [24] [25], which employ explicit representations to describe key points, shape contour or surface information. Recent research into feature- and correspondence-based shape registration can be found in [26] and [27]. The algorithm proposed in [26] is motivated by retinex theory. It employs an adaptive smoothing method to preserve shape edges and corners. The Hough transform is developed in [27] to reduce the demand for computational memory with the intrinsic Hough and enhance registration accuracy using the minimum-entropy Hough. These methods simplify the calculation of transformation parameters but significantly rely on the feature description and landmark correspondence, which leads to inaccurate

registration when the shapes either contain various sub-shapes or have different Euler characteristics.

Conversely, implicit representations, such as the signed distance function (SDF), concentrate on the description of shape regions. The SDF is frequently employed for shape registration in [28] and [29], since it is a desirable representation for two identical shapes in registration. However, it leads to infinity for the integral of the dissimilarity of two different shapes. Another implicit representation, characteristic function (CF), manages to overcome the shortcomings of SDF and is used in [30] and [31] to perform robust shape registration.

Moreover, another shape representation method known as Laplace-Beltrami spectral descriptors is employed in [32] and [33] for shape analysis. This representation is independent of the shape's orientation and transformation. This advantage is also inherited by the algorithm proposed in this chapter. The recent spectral shape descriptor was proposed in [34]. It employs a neural network to establish the shape spectrum and create learning-based binary shape descriptors. Although these sophisticated spectral descriptors improve the sensitivity and specificity for capturing features, they are significantly dependent on feature correspondence.

Iterative closest point (ICP) was firstly proposed in [35] to optimise the rigid registration based on proper initial pose. It was then developed by [36] [37] [38]. The recent development of ICP is reported in [39], which is based on the point cloud with corresponding points. The transformation is refined iteratively by minimising the distance of geometrically similar groups of points. Similar point-set-based registration methods are studied in [40] and [41] to improve robustness. These aforementioned algorithms boost the processing speed and optimise the similarity metric, but the gradient-descent optimisation method always leads to a local minimum. In [29], cross-

correlation is used to improve the robustness of shape registration, and it is demonstrated to be a possible solution to the local minimum problem. Also, cross-correlation based on the fast Fourier transform (FFT) for shape registration is reported in [42] and [43].

As a similarity metric, SSD is widely used in shape registration as it is straightforward to measure the greyscale difference across the voxels at the same locations [30] [31]. Inspired by information theory, MI is proposed as a statistical registration technique and similarity metric. It is employed in [44] and [45] to calculate the optimal transformation parameters by maximising the MI between different shapes. Moreover, the scale-rotation-translation-compatible distance as an innovative metric is proposed in [46] to achieve scale-invariant registration.

In previous work, the influence of noise has rarely been analysed, but the robustness of algorithms should be attached importance for practical applications. The challenge of the registration for the candidate shapes containing different sub-shapes, named here as group shapes, needs to be overcome. The main objective of this chapter is therefore to propose a robust and high-performance similarity registration method for volumetric shapes. It intends to address the concerns of registration for the group shapes even in the presence of noise. In addition, the desired similarity alignment method is required to achieve high registration accuracy and stability within a reasonable time. In this chapter, the CFs are employed to represent shapes. The technique proposed here is developed using an iterative scheme and radial moments. In each iteration, the cross-correlation based on the spherical coordinate system is employed to accelerate convergence and avoid local minimum. Performance in terms of registration accuracy, stability and rate of convergence is validated by experiments on the group shapes in the presence of noise, as well as a lung shape database.

The contributions in this chapter are as follows: Firstly, an iterative method is introduced to improve the robustness, accuracy and stability of volumetric shape registration, especially for group shapes. Secondly, a method based on the spherical coordinate system is proposed to simplify the calculation of fine rotational angles and to guarantee the global optimum for each fine rotation. Thirdly, the robustness of registration algorithms is analysed in this chapter.

This section is structured as follows: The relevant background materials are introduced in Section 2.2, followed by the methodology with mathematical derivations in Section 2.3. The numerical results are presented in Section 2.4 to validate the improvement of robustness, performance and stability achieved by the algorithm proposed here. Conclusions and future work are finally summarised in Section 2.5.

2.2 Theoretical Basis

2.2.1 Rotation Theorems

In [47], P. Davenport demonstrated that any orientation of a rigid body could be obtained by three rotations around three different non-orthogonal axes. Davenport rotation theorem was then extended to prove that a unique decomposition of a rotation into three rotations around three orthogonal axes is reliable. The decomposition is generalised as the Tait-Bryan rotations and employed in Section 2.3.2.2 to calculate rotation parameters.

2.2.2 Representation of Volumetric Shapes

In this chapter, CFs are used to represent shapes. The CF of a volumetric shape in the image domain Ω is defined as

$$\phi(x,y,z) = \begin{cases} 1 & (x,y,z) \in S \\ 0 & (x,y,z) \in \Omega - S \end{cases} \tag{2.1}$$

where *S* denotes the inside region of the shape.

2.3 Methodology

Compared with rigid registration, which only considers rotation and translation, similarity registration expands the scope of application and is performed sequentially through scaling, rotation and translation. Based on the recent shape registration method proposed in [31], which effectively improves the robustness in a two-dimensional plane, the algorithm proposed here is developed in three-dimensional space for volumetric shape registration. Furthermore, some complicated issues such as computational complexity in three-dimensional space are resolved in this chapter to obtain outstanding performance. In summary, the method is extended as follows:

- Extend the calculation of radial moments for three-dimensional shapes.
- Employ PCA to perform a coarse registration.
- Employ the spherical coordinate system to simplify the calculation of the rotational angle.
- Employ a three-dimensional Fourier transform to calculate the optimal rotation and translation parameters based on a cross-correlation framework.

In this chapter, $\phi_{tar}(x, y, z)$ and $\phi_{src}(x, y, z)$ denote the CFs of the target and source shapes respectively. s is used to denote the scaling parameter. As a rotation can be decomposed into three rotations around three orthogonal axes, φ_z, φ_y and φ_x here represent the fine rotation parameters around the three Cartesian axes, Z, Y and X respectively. Also, Δx, Δy and Δz denote the translations along the X, Y and Z axes respectively. The dissimilarity between registered source shape $\overline{\phi_{src}}$ and target shape ϕ_{tar} can be measured by the SSD.

$$E = \iiint_{\Omega} \left|\phi_{tar}(x,y,z) - \overline{\phi_{src}}(x,y,z)\right|^2 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \tag{2.2}$$

where Ω is the image domain. The optimal transformation parameters can therefore be calculated by minimising the dissimilarity measurement in Eq. (2.2), i.e.

$$\left(\hat{s}, \widehat{\varphi_z}, \widehat{\varphi_y}, \widehat{\varphi_x}, \widehat{\Delta x}, \widehat{\Delta y}, \widehat{\Delta z}\right) = \underset{s,\varphi_z,\varphi_y,\varphi_x,\Delta x,\Delta y,\Delta z}{\operatorname{argmin}} E \tag{2.3}$$

where $\hat{s}, \widehat{\varphi_z}, \widehat{\varphi_y}, \widehat{\varphi_x}, \widehat{\Delta x}, \widehat{\Delta y}, \widehat{\Delta z}$ are the estimated parameters for scaling, rotation and translation respectively.

2.3.1 Coarse Registration

In order to improve the computational efficiency and robustness of the algorithm, PCA-based coarse registration is employed here to find the orthogonal axes and then the rotational angles. The coarse registration starts with alignment of the centroids of the target and source shapes to the origin of the coordinate system. Then, the method proposed here arranges for the first and second main axes of the shapes computed by PCA to coincide with the Z and Y axes (in the Cartesian coordinate system) respectively. The procedure of coarse registration is illustrated in Figure 2.1.

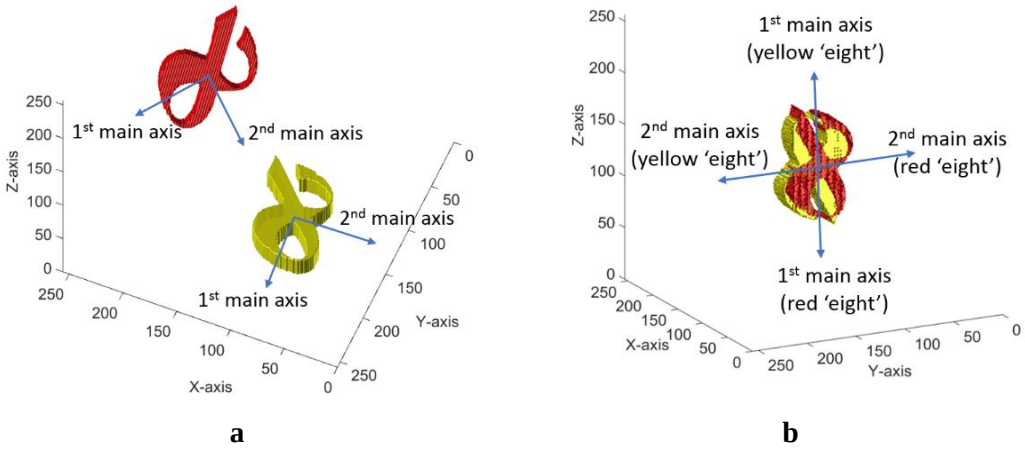
Under some circumstances, the directions of the corresponding main axes might be reversed after coarse registration, as shown in Figure 2.1(b). The flipping problem can be resolved by the subsequent fine rotation, which employs an iterative method as described in Section 2.3.4.

After centralisation and coarse rotation, the ϕ_{tar} and ϕ_{src} are transformed to new poses, i.e.

$$\widetilde{\phi_{tar}}(x,y,z) = \phi_{tar}(x - x_t, y - y_t, z - z_t) \tag{2.4}$$

$$\widetilde{\phi_{src}}(x,y,z) = \phi_{src}(R_c(x - x_s, y - y_s, z - z_s)) \tag{2.5}$$

where R_c denotes the coarse rotation of the source shape using PCA, and (x_t, y_t, z_t) and (x_s, y_s, z_s) are the centroids of the target and source shapes. The coarse registration also aids the shape CF mapping from the Cartesian to the spherical coordinate system in the subsequent calculation of fine rotational angles.



a **b**

Figure 2.1 *The coarse registration is implemented by centralisation and coarse rotation. In order to highlight the coincidence, the coarse registration is presented with the modification that the two centroids are aligned but not translated to the origin of the coordinate system. (a) Initial poses before coarse registration. (b) Coarse registration results.*

2.3.2 Fine Registration

2.3.2.1 Scaling

The difference of radial moments is employed here to derive the scaling parameter. In comparison with geometric moments, which are used to calculate the scaling parameters along the X, Y and Z directions for shear transformation, radial moments are rotation-invariant and simplify the computation by offering a uniform scaling parameter for all directions in three-dimensional space.

For two centralised volumetric shapes, $\widetilde{\phi_{tar}}, \widetilde{\phi_{src}}$, the radial moments are computed by Eq. (2.6) and Eq. (2.7),

$$M_m^{tar} = \iiint_{\Omega} \left(\sqrt{x^2 + y^2 + z^2}\right)^m \widetilde{\phi_{tar}}(x, y, z)\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z \qquad (2.6)$$

$$M_m^{src} = \iiint_{\Omega} \left(\sqrt{x^2 + y^2 + z^2}\right)^m \widetilde{\phi_{src}}(x, y, z)\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \tag{2.7}$$

where *m* denotes the number of moments. The target shape scaled by $1/s$ is written as

$$M_m^{tar/s} = \iiint_{\Omega} \left(\sqrt{x^2 + y^2 + z^2}\right)^m \widetilde{\phi_{tar}}\left(\frac{x}{s}, \frac{y}{s}, \frac{z}{s}\right)\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \tag{2.8}$$

By replacing the variables x/s, y/s and z/s with *X*, *Y* and *Z* respectively, Eq. (2.8) can be rewritten as

$$M_m^{tar/s} = s^{m+3} \iiint_{\Omega} \left(\sqrt{X^2 + Y^2 + Z^2}\right)^m \widetilde{\phi_{tar}}(X, Y, Z)\,\mathrm{d}X\,\mathrm{d}Y\,\mathrm{d}Z = s^{m+3} M_m^{tar} \tag{2.9}$$

The optimal scaling parameter is estimated by minimising the distance between these two radial moments, i.e.

$$\hat{s} = \underset{s}{\operatorname{argmin}}\left(E_{scl}(s)\right) = \underset{s}{\operatorname{argmin}}\left(\sum_{m=0}^{M} |M_m^{src} - s^{m+3} M_m^{tar}|^2\right) \tag{2.10}$$

Then, the derivative of E_{scl} with respect to s should vanish:

$$\frac{dE_{scl}}{ds} = \sum_{m=0}^{M} 2(m+3)[s^{2m+5}(M_m^{tar})^2 - s^{m+2} M_m^{src} M_m^{tar}] = 0 \tag{2.11}$$

It is obvious that Eq.(2.11) is a polynomial equation with respect to s. In fact, only one solution, which is real and positive-valued, can be regarded as the desired scaling parameter. The remaining solutions are either complex or negative real-valued.

2.3.2.2 Rotation

In each iteration, the desired rotational angles can be estimated by minimising the cost function, $E_{rot}(\varphi_z, \varphi_y, \varphi_x)$, which is similar to Eq. (2.2):

$$E_{rot}(\varphi_z, \varphi_y, \varphi_x) = \iiint_{\Omega} \left|\widetilde{\phi_{src}}\right|^2 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z + \iiint_{\Omega} \left|\widetilde{\phi_{tar}}\right|^2 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z - 2\langle \widetilde{\phi_{src}} \cdot \widetilde{\phi_{tar}} \rangle \tag{2.12}$$

where the $\varphi_z, \varphi_y, \varphi_x$ are rotational angles around the Z, Y and X axes, as defined by the Tait-Bryan rotations.

As the first and second terms on the right-hand side of the above equation are independent of the rotational angles, Eq. (2.12) is simplified as

$$E_{rot}(\varphi_z, \varphi_y, \varphi_x) = C - 2\delta \tag{2.13}$$

$$\delta = \iiint_{\Omega} \left(\widetilde{\phi_{tar}}(x,y,z)\widetilde{\phi_{src}}\left(R_f(x,y,z)\right)\right) \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \tag{2.14}$$

where R_f denotes the fine rotation of the source shape.

Therefore, minimising Eq. (2.12) is equivalent to maximising δ with respect to φ_z, φ_y and φ_x:

$$\left(\widehat{\varphi_z}, \widehat{\varphi_y}, \widehat{\varphi_x}\right) = \underset{\varphi_z,\varphi_y,\varphi_x}{\operatorname{argmax}}(\delta) \tag{2.15}$$

As the three rotation parameters have to be derived sequentially, a local minimum problem within the combination of three sequential rotations can be nearly overcome by a reliable computation order, which is from the first to the third main axis [30]. In Section 2.3.1, PCA-based coarse registration was performed to arrange for the first and second main axes of the shapes to coincide with the Z and Y axes respectively. Therefore, the order is determined as Z-Y-X, which is in accordance with from the first to the third main axis.

In order to facilitate the calculation, the shape CFs are mapped to the spherical coordinate system, in which r, θ and φ representing radial distance, polar angle and azimuthal angle respectively are employed to define the locations of voxels. φ is the only variable to compute the rotational angle around an axis. Eq. (2.14) is therefore rewritten based on the spherical coordinate system as

$$\delta(\varphi_z) = \iiint_{\Omega} \widetilde{\phi_{s1}}(r,\theta,\varphi)\widetilde{\phi_{t1}}(r,\theta,\varphi-\varphi_z) r^2 \sin\theta\,\mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi \tag{2.16}$$

$$\delta(\varphi_y) = \iiint_{\Omega} \widetilde{\phi_{s2}}(r,\theta,\varphi)\widetilde{\phi_{t2}}(r,\theta,\varphi-\varphi_y) r^2 \sin\theta\,\mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi \tag{2.17}$$

$$\delta(\varphi_x) = \iiint_{\Omega} \widetilde{\phi_{s3}}(r,\theta,\varphi)\widetilde{\phi_{t3}}(r,\theta,\varphi-\varphi_x)r^2\sin\theta\,\mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi \tag{2.18}$$

where $\widetilde{\phi_{s1}}$, $\widetilde{\phi_{s2}}$ and $\widetilde{\phi_{s3}}$ denote the mapped source shapes using the Z, Y and X axes as the zenith direction respectively. Similarly, $\widetilde{\phi_{t1}}$, $\widetilde{\phi_{t2}}$ and $\widetilde{\phi_{t3}}$ denote mapped target shapes after rotating around the Z, Y and X axes respectively. Eq. (2.16) is used as an example here to complete the following derivation.

Given $\widetilde{\kappa_{s1}}(r,\theta,\varphi)$ to substitute for $\widetilde{\phi_{s1}}(r,\theta,\varphi)r^2\sin\theta$ for convenience, Eq. (2.16) is simplified to

$$\delta(\varphi_z) = \iiint_{\Omega} \widetilde{\kappa_{s1}}(r,\theta,\varphi)\widetilde{\phi_{t1}}(r,\theta,\varphi-\varphi_z)\mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi \tag{2.19}$$

Cross-correlation is employed here to derive the optimal φ_z which maximises $\delta(\varphi_z)$. Given the Fourier transforms of $\widetilde{\kappa_{s1}}(r,\theta,\varphi)$ and $\widetilde{\phi_{t1}}(r,\theta,\varphi)$, denoted by $\widetilde{K_{s1}}(\eta,\lambda,\xi)$ and $\widetilde{\Phi_{t1}}(\eta,\lambda,\xi)$ respectively, the Fourier transform of $\widetilde{\phi_{t1}}(r,\theta,\varphi-\varphi_z)$ is therefore written as

$$\mathcal{F}\left(\widetilde{\phi_{t1}}(r,\theta,\varphi-\varphi_z)\right) = \widetilde{\Phi_{t1}}(\eta,\lambda,\xi)e^{-j\xi\varphi_z} \tag{2.20}$$

where η, λ and ξ are spatial frequencies in the spherical coordinate system. Based on Parseval's theorem, $\widehat{\varphi_z}$ can be calculated by

$$\widehat{\varphi_z} = \underset{\varphi_z}{\operatorname{argmax}}(\delta)$$
$$= \underset{\varphi_z}{\operatorname{argmax}}\left(\frac{1}{2\pi}\iiint \widetilde{K_{s1}}(\eta,\lambda,\xi)\widetilde{\Phi_{t1}}^{*}(\eta,\lambda,\xi)e^{j\xi\varphi_z}\,\mathrm{d}\eta\mathrm{d}\lambda\mathrm{d}\xi\right) \tag{2.21}$$

where * denotes the complex conjugate. In this way, the optimal φ_z can be calculated based on the inverse Fourier transform. The value of φ_z corresponding to the maximum δ can be straightforwardly investigated, as illustrated in Figure 2.2.

Similarly, $\widehat{\varphi_y}$ and $\widehat{\varphi_x}$ are calculated respectively as

$$\widehat{\varphi_y} = \underset{\varphi_y}{\operatorname{argmax}} \left(\frac{1}{2\pi} \iiint \widetilde{K_{s2}}(\eta, \lambda, \xi) \widetilde{\Phi_{t2}}^{*}(\eta, \lambda, \xi) e^{j\xi\varphi_y} \, \mathrm{d}\eta \mathrm{d}\lambda \mathrm{d}\xi \right) \tag{2.22}$$

$$\widehat{\varphi_x} = \underset{\varphi_x}{\operatorname{argmax}} \left(\frac{1}{2\pi} \iiint \widetilde{K_{s3}}(\eta, \lambda, \xi) \widetilde{\Phi_{t3}}^{*}(\eta, \lambda, \xi) e^{j\xi\varphi_x} \, \mathrm{d}\eta \mathrm{d}\lambda \mathrm{d}\xi \right) \tag{2.23}$$

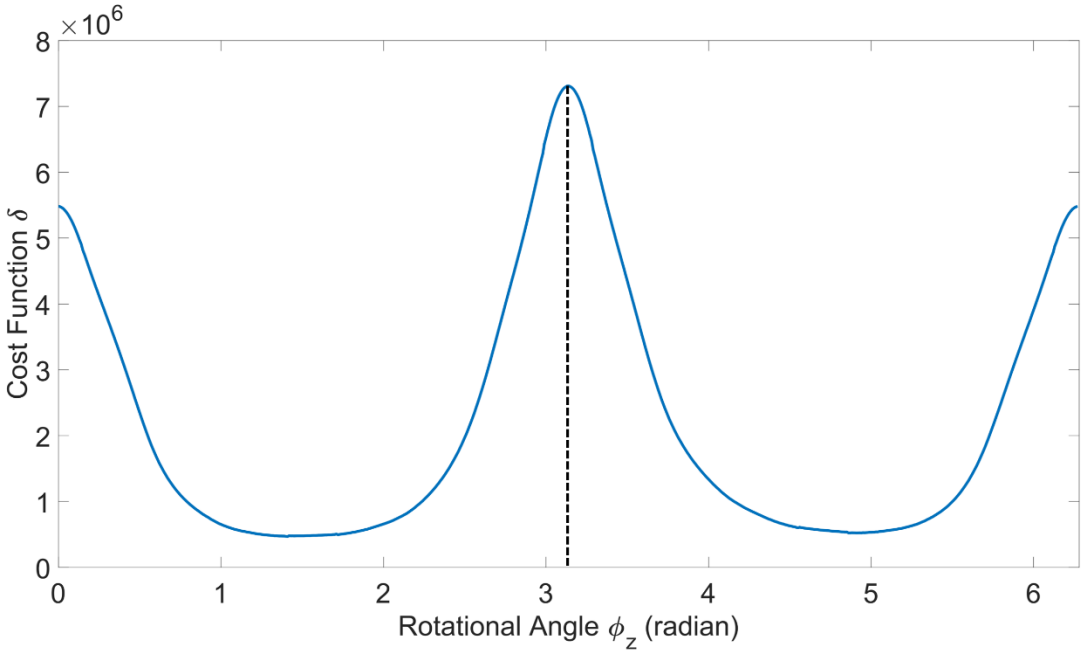


Figure 2.2 *A plot for global cost δ with respect to rotation φ_z. The φ_z corresponding to the maximum of δ is the optimal rotational angle.*

2.3.2.3 Translation

Similar to Eq. (2.12), the first two terms on the right-hand side of Eq. (2.24) are translation-invariant. Minimising Eq. (2.24) is possible by maximising $\langle \widetilde{\phi_{src}} \cdot \widetilde{\phi_{tar}} \rangle$ with respect to $\Delta x, \Delta y$ and Δz:

$$E_{tran}(\Delta x, \Delta y, \Delta z) = \iiint_{\Omega} \left| \widetilde{\phi_{src}} \right|^2 \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z + \iiint_{\Omega} \left| \widetilde{\phi_{tar}} \right|^2 \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z - 2 \langle \widetilde{\phi_{src}} \cdot \widetilde{\phi_{tar}} \rangle \tag{2.24}$$

$$\langle \widetilde{\phi_{src}} \cdot \widetilde{\phi_{tar}} \rangle = \iiint_{\Omega} \widetilde{\phi_{src}}(x, y, z) \widetilde{\phi_{tar}}(x - \Delta x, y - \Delta y, z - \Delta z) \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z \tag{2.25}$$

Given the Fourier transform of $\widetilde{\phi_{src}}(x,y,z)$ and $\widetilde{\phi_{tar}}(x,y,z)$, denoted by $\widetilde{\Phi_{src}}(\eta',\lambda',\xi')$ and $\widetilde{\Phi_{tar}}(\eta',\lambda',\xi')$ respectively, the Fourier transform of $\widetilde{\phi_{tar}}(x-\Delta x, y-\Delta y, z-\Delta z)$ is therefore written as

$$\mathcal{F}\left(\widetilde{\phi_{tar}}(x-\Delta x, y-\Delta y, z-\Delta z)\right) = \widetilde{\Phi_{tar}}(\eta',\lambda',\xi')e^{-j(\eta'\Delta x+\lambda'\Delta y+\xi'\Delta z)} \quad (2.26)$$

where η', λ' and ξ' are spatial frequencies in the Cartesian coordinate system. By using Parseval's theorem, $\widehat{\Delta x}, \widehat{\Delta y}$ and $\widehat{\Delta z}$ are calculated by

$$\left(\widehat{\Delta x}, \widehat{\Delta y}, \widehat{\Delta z}\right) =$$
$$\underset{\Delta x,\Delta y,\Delta z}{\operatorname{argmax}}\left(\frac{1}{2\pi}\iiint \widetilde{\Phi_{src}}(\eta',\lambda',\xi')\,\widetilde{\Phi_{tar}}^{*}(\eta',\lambda',\xi')e^{j(\eta'\Delta x+\lambda'\Delta y+\xi'\Delta z)}\mathrm{d}\eta'\mathrm{d}\lambda'\mathrm{d}\xi'\right) \quad (2.27)$$

where * denotes the complex conjugate.

2.3.3 Evaluation of Similarity Registration

The similarity measurement is employed here to evaluate the performance of algorithms and to determine the iterative progression. In [30], a few evaluation methods were mentioned for measuring the regional similarity. In this chapter, these classic measurements are analysed before use.

Firstly, the normalised inner product is used to evaluate the common region inside two registered shapes and is denoted by

$$E_{IP} = \iiint_{\Omega} \frac{\phi_{tar}(x,y,z)\overline{\phi_{src}}(x,y,z)}{\|\phi_{tar}(x,y,z)\|_2\left\|\overline{\phi_{src}}(x,y,z)\right\|_2}\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \quad (2.28)$$

where $\|\cdot\|_2$ denotes Euclidean norm, and $\phi_{tar}(x,y,z)$ and $\overline{\phi_{src}}(x,y,z)$ are the CFs of the target shape and registered source shape respectively.

Furthermore, accuracy (E_{Ac}), overlap (E_{Ov}), sensitivity (E_{Se}), and specificity (E_{Sp}) are selectable metrics for similarity measurement:

$$E_{Ac} = \frac{V_c+V_o}{V_\Omega} \tag{2.29}$$

$$E_{Ov} = \frac{V_c}{V_c+V_t+V_r} \tag{2.30}$$

$$E_{Se} = \frac{V_c}{V_c+V_t} \tag{2.31}$$

$$E_{Sp} = \frac{V_o}{V_o+V_r} \tag{2.32}$$

V_Ω: the volume of the image domain

V_c: the volume of the common region inside the two shapes

V_o: the volume of the background region outside the two shapes

V_t: the volume of the region inside the target shape but outside the registered source shape

V_r: the volume of the region inside the registered source shape but outside the target shape

Nevertheless, the definition of E_{Ac} and E_{Sp} are involved with V_o, which means the background information inevitably affects the results obtained. In the case where the background occupies a considerable proportion of the whole image domain, the values of E_{Ac} and E_{Sp} are very close to unity, which causes these two measurements to be insensitive to the difference between the shapes being compared. E_{Se} concentrates on the common region in the target shape, but ignores the evaluation with respect to the registered source shape. E_{Se} is therefore inapplicable to the measurement of shape registration.

E_{Ov} is a reliable measurement which is similar to E_{IP}. It concentrates on the foreground information and evaluates the common region within two shapes. An iterative algorithm based on E_{Ov} and E_{IP} manages to maximise the common region and minimise the individual regions simultaneously.

Furthermore, Hausdorff distance (E_{HD}), first proposed in [48], is used to measure the resemblance between two shapes, and the bidirectional Hausdorff distance between ϕ_{tar} and $\overline{\phi_{src}}$ is denoted by

$$E_{HD}(\phi_{tar}, \overline{\phi_{src}}) = \max\left(h(\phi_{tar}, \overline{\phi_{src}}), h(\overline{\phi_{src}}, \phi_{tar})\right) \tag{2.33}$$

where

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \tag{2.34}$$

Since it requires no explicit correspondence between *A* and *B*, Hausdorff distance is employed to robustly evaluate registration against tiny errors that are obtained with feature extraction and descriptors. In terms of shape registration, it is a promising metric and is adopted by this chapter.

In this chapter, normalised inner product (E_{IP}), overlap (E_{Ov}) and Hausdorff distance (E_{HD}) are employed to evaluate performance in the experiments. Higher E_{Ov} and E_{IP} and lower E_{HD} indicate superior registration accuracy.

2.3.4 Iterative Algorithm

In order to obtain the optimal parameters for scaling, rotation and translation, an iterative method is developed in this chapter. In comparison with the non-iterative method proposed in [30], the iterative method can effectively reduce the influence of sequential computation on the transformation parameters and then avoid converging to a local minimum. The registration starts with centralisation and coarse rotation. The optimal registration parameters for scaling, rotation and translation are then calculated iteratively. At the end of each iteration, the similarity of the registered shapes is evaluated by one of the methods mentioned in Section 2.3.3. The iterative process stops

when a certain condition is met. The entire process of registration proceeds as shown in Figure 2.3.

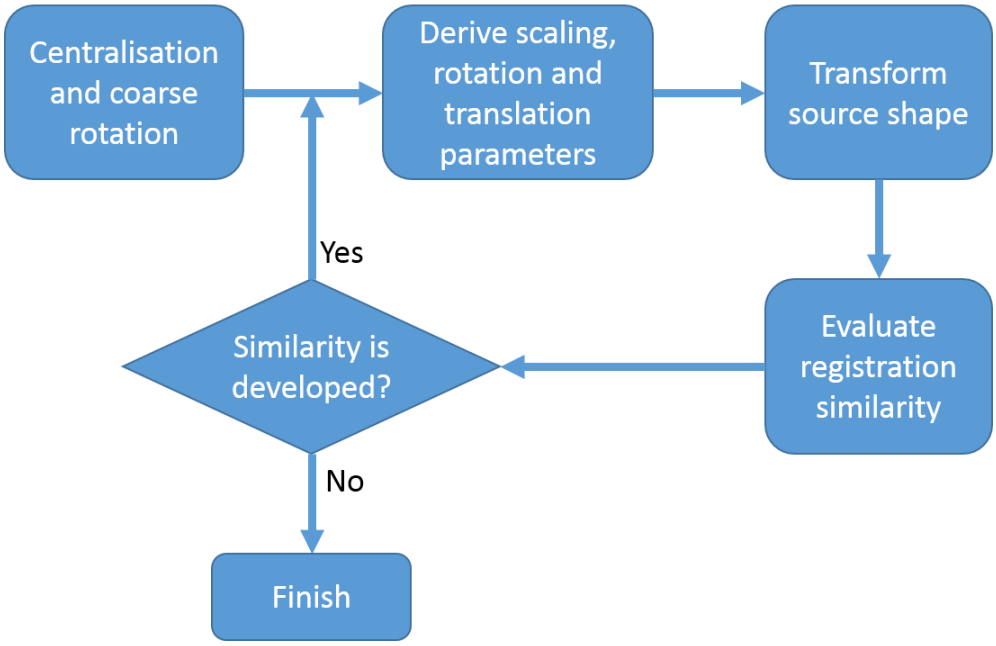


Figure 2.3 *Flow chart of the iterative shape registration procedure.*

The pseudo-code for programming is as follows:

1) *Perform centralisation and coarse rotations*

At iteration i:

2) *Solve the following equation*

$$\sum_{m=0}^{M}(m+3)[{s_i}^{2m+5}(M_m^{tar})^2-{s_i}^{m+2}(M_m^{src})_{i-1}M_m^{tar}]=0$$

3) *Find the following values*

$$\widehat{\varphi_z}_i=\underset{\varphi_z}{\operatorname{argmax}}\left(\frac{1}{2\pi}\iiint\left(\widetilde{K_{s1}}(\eta,\lambda,\xi)\right)_{i-1}\widetilde{\Phi_{t1}}^{*}(\eta,\lambda,\xi)e^{j\xi\varphi_z}\,\mathrm{d}\eta\mathrm{d}\lambda\mathrm{d}\xi\right)$$

$$\widehat{\varphi_y}_i=\underset{\varphi_y}{\operatorname{argmax}}\left(\frac{1}{2\pi}\iiint\left(\widetilde{K_{s2}}(\eta,\lambda,\xi)\right)_{i-1}\widetilde{\Phi_{t2}}^{*}(\eta,\lambda,\xi)e^{j\xi\varphi_y}\,\mathrm{d}\eta\mathrm{d}\lambda\mathrm{d}\xi\right)$$

$$\widehat{\varphi_{x_i}} = \underset{\varphi_x}{\operatorname{argmax}}\left(\frac{1}{2\pi}\iiint \left(\widetilde{K_{s3}}(\eta,\lambda,\xi)\right)_{i-1} \widetilde{\Phi_{t3}}^{*}(\eta,\lambda,\xi)e^{j\xi\varphi_x}\,\mathrm{d}\eta\mathrm{d}\lambda\mathrm{d}\xi\right)$$

$$\left(\widehat{\Delta x_i}, \widehat{\Delta y_i}, \widehat{\Delta z_i}\right)$$
$$= \underset{\Delta x,\Delta y,\Delta z}{\operatorname{argmax}}\left(\frac{1}{2\pi}\iiint \left(\widetilde{\Phi_{src}}(\eta',\lambda',\xi')\right)_{i-1} \widetilde{\Phi_{tar}}^{*}(\eta',\lambda',\xi')e^{j(\eta'\Delta x+\lambda'\Delta y+\xi'\Delta z)}\mathrm{d}\eta'\mathrm{d}\lambda'\mathrm{d}\xi'\right)$$

4) *Transform the source shape which was updated in iteration i-1.*

5) *Evaluate the similarity of the two shapes by E_{IP} or E_{Ov}.*

6) *Repeat steps 2-5, until the value of E_{IP} or E_{Ov} is less than any one of the previous values.*

2.4 Experimental Results and Discussions

In this section, a series of experiments are conducted to evaluate the performance of different registration methods. Before presenting the experimental results, it is necessary to mention some implementation issues at the beginning.

Firstly, M in Eq. (2.11) should be pre-set before calculating the scaling parameter. With the increase of M, the estimated scaling parameter tends to approach the true value, as shown in Figure 2.4. However, when it arrives at a certain value, the accuracy improvement is not substantial but time consumption increases dramatically. Thus, a proper trade-off needs to be made.

Moreover, as the voxels in the Cartesian coordinate system are mapped to the spherical coordinate system to facilitate the calculation of rotational angles, the converted resolutions need to be negotiated. As stated in Section 2.3.2.2, only azimuthal angles are used to manipulate the rotation of shapes. An efficient solution is therefore to lower the resolutions with respect to the r and θ axes whilst enhance the resolution with respect to the φ axis.

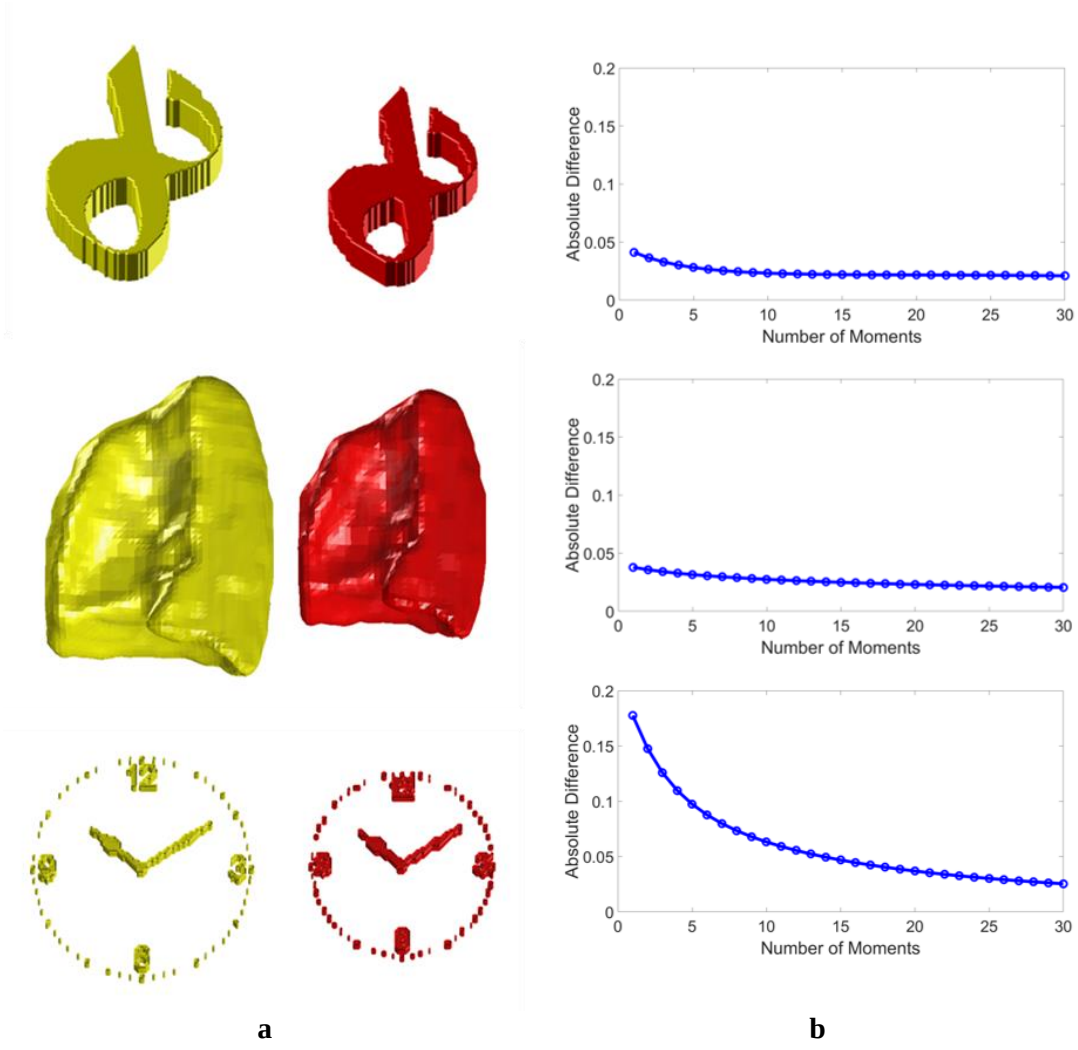


Figure 2.4 *Three pairs of shapes and the corresponding estimated scaling parameters relative to the real values. (a) Three pairs of similar shapes (simple shapes, medical shapes and group shapes) are displayed in yellow (target shapes) and red (source shapes). (b) The absolute difference between real and estimated scaling parameters with the different numbers of moments used.*

The following experiments and comparisons are performed with three methods, two of which are the Matlab built-in function "imregister" and the algorithm proposed in [30], named as Algorithm 1 and Algorithm 2 respectively here. The "imregister" function as a comparable iterative alignment algorithm is included in the Matlab image processing toolbox. The method proposed in [30] is a state-of-the-art CF-based volumetric shape registration technique. It also employs phase-correlation and gradient descent optimisation method to investigate the transformation parameters, whereas it does not

conduct calculation iteratively. Moreover, ICP as a classic method is also included in Section 2.4.2 for comparison.

All the experiments were implemented in Matlab 2017b on a PC workstation with an Intel Core i5 (3.2GHz) processor and 16 GB RAM. In this chapter, the target shapes are depicted in yellow while the source shapes are displayed in red for visual inspection. All the shapes are presented by their surfaces with suitable transparent effects to aid visualization.

2.4.1 Registration for Group Shapes

The experiments in this section were implemented to validate robustness and alignment accuracy for group shapes. The testing candidates are two three-dimensional clocks, as shown in Figure 2.5(a). It is apparent that their dials are different. One is composed of four numerical digits and a discrete frame whilst the other clock is made of twelve Roman numerals and a continuous frame. The numerical digits, Roman numerals, circular frame and hands are regarded as the sub-shapes of the clock shape. The registration is then performed by Algorithm 1, Algorithm 2 and the algorithm proposed here.

These two clocks are initially placed overlapped. The alignment results are depicted in Figure 2.5(b-d). It is obvious that Algorithm 1 fails to register the group shapes since it tends to fall into a local minimum and terminate the registration prematurely, as shown in Figure 2.5(b). Algorithm 2 (shown in Figure 2.5(c)) also fails to register these two group shapes as the non-iterative technique cannot derive the optimal transformation parameters in one attempt. Conversely, the result shown in Figure 2.5(d) significantly demonstrates the power of the iterative scheme and the superior accuracy achieved by the method proposed here.

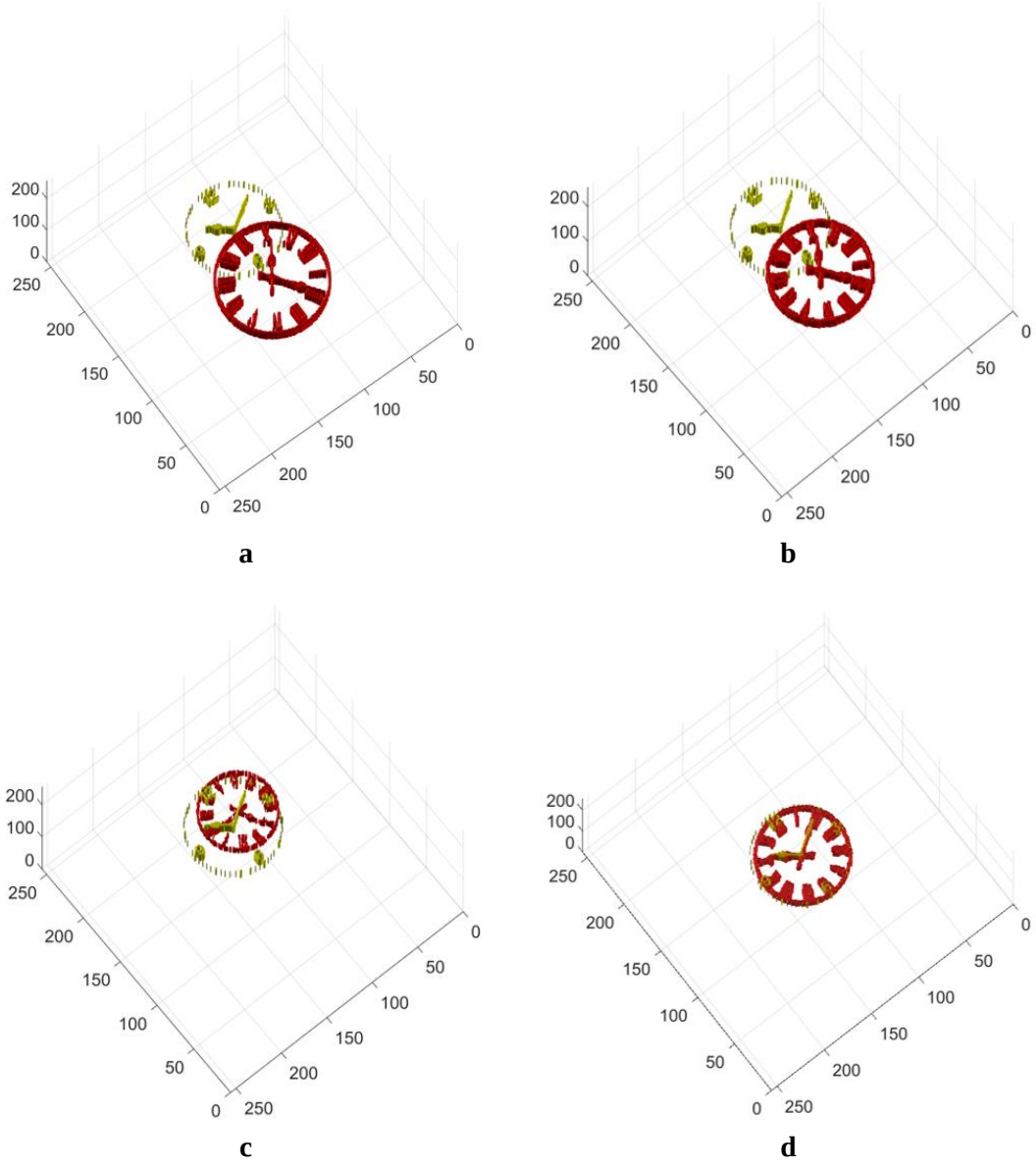


Figure 2.5 *Registration of two clocks without noise. (a) The initial poses of two clock shapes. They are overlapped before registration. (b) The registration result of Algorithm 1. (c) The registration result of Algorithm 2. (d) The registration result of the proposed method.*

Salt and pepper noise, which is applicable to shapes (binary masks), is then added to one of the clock shapes to challenge the robustness of the algorithms. The two clocks are arranged separately without overlap. The results of the two registration tests in the presence of different signal-to-noise ratios (SNRs) are shown in Figure 2.6. It is demonstrated that the algorithm proposed here manages to register the two group shapes robustly.

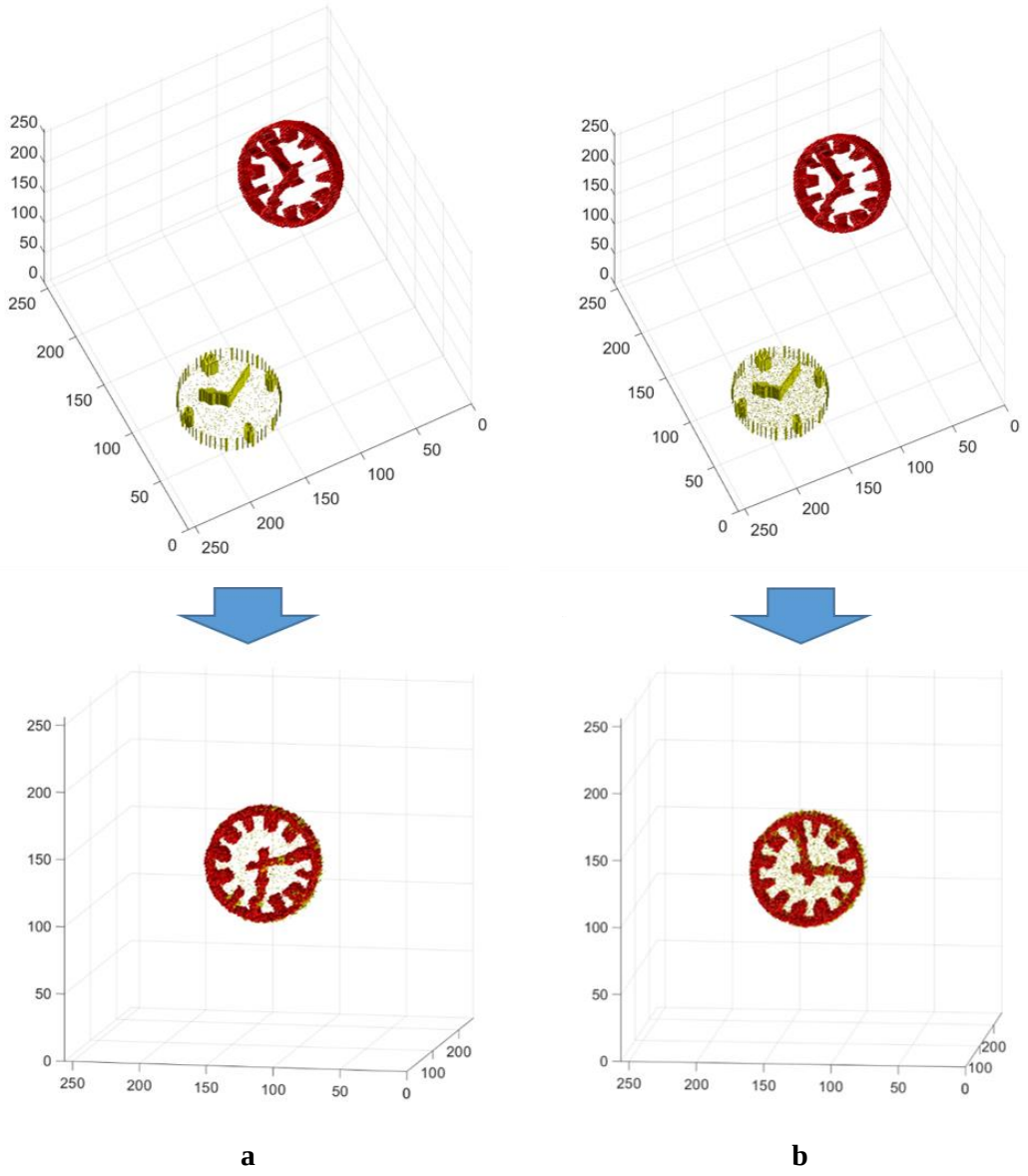


Figure 2.6 *Registration of two clocks in the presence of different SNRs. (a) SNR=20dB. (b) SNR=10dB.*

The following experiment is implemented to further validate the robustness of the different algorithms. As depicted in Figure 2.7, two identical clocks are deployed in the same pose. The target clock shape is deteriorated by adding salt and pepper noise with a specific SNR whilst the source clock shape is initially rotated 180 degrees around the axis which is orthogonal to the dial and passes through the centre of the shape. The error angles (relative to the ground-truth of 180 degrees) after registration are used to measure the alignment performance, especially in terms of rotation.

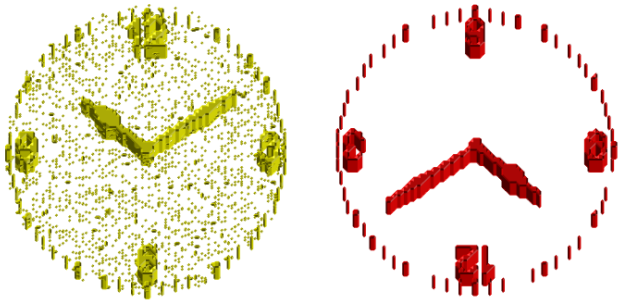


Figure 2.7 *Two group shapes to be registered are displayed separately. The target (on the left-hand side) has noise added to increase the difficulty of registration. The source (on the right-hand side) is initially rotated by 180 degrees.*

For each test mode (i.e. tests with the fixed SNR for the target shape), ten alignment tests with random noise distributions are conducted. The results of eight test modes (SNR= -15, -10, -5, 0, 5, 10, 15, 20dB) are plotted in Figure 2.8. The shortcomings of Algorithm 1 are revealed again in the presence of noise. Since Algorithm 1 is very vulnerable to local minimum, the alignment tends to finish promptly with obtained rotational angles close to zero. Algorithm 2 is able to obtain desirable results when the

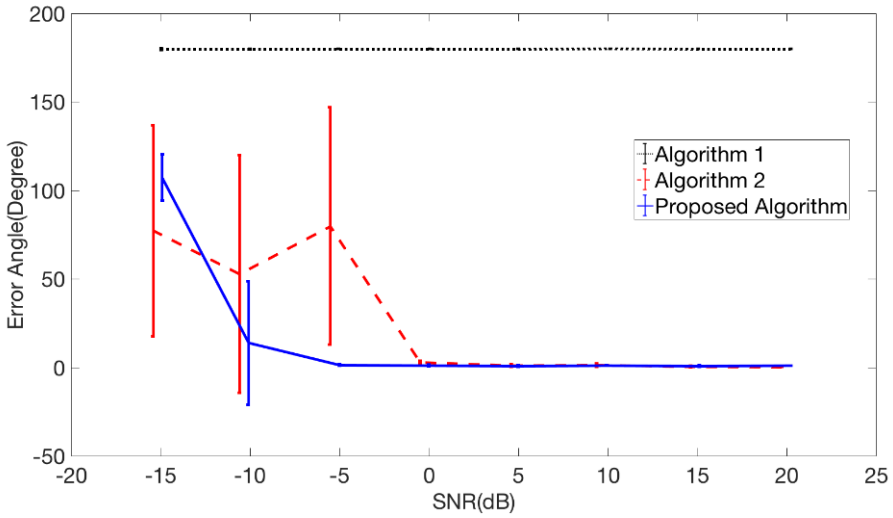


Figure 2.8 *Error angle graph with respect to different SNRs. The error angles of Algorithm 1 and Algorithm 2 are depicted by a black dotted line and red dashed line respectively with their standard deviations. The error angles of the proposed algorithm are displayed with a blue solid line with its standard deviation.*

SNR is not lower than 0dB. It is however not capable of aligning shapes under lower SNR conditions as the robustness is limited by the non-iterative solution. In contrast, the algorithm proposed here with an iterative method outperforms Algorithm 1 and 2 in the presence of noise, even when the SNR is as low as -5dB.

2.4.2 Registration for Lung Shapes in the Database

In this section, the shapes provided by the lung low-dose CT database (subject details refer to Table A-1 and A-2 of Appendix A) are employed to further validate the speed, accuracy and stability of Algorithm 1, Algorithm 2, ICP and the algorithm proposed here. The database contains 33 left and 33 right lung shapes represented by their CFs. In the following experiments, the leave-one-out approach is employed, which means that each subject in the database is sequentially chosen as the target shape and the remaining 32 cases are regarded as the source shapes. Thus, the total number of registration tests performed by each method is 1056 for the left lung and 1056 for the right lung. Furthermore, in order to challenge their abilities against local optima, the initial orientations of the two lung shapes are set randomly. As mentioned in Section 2.3.3, E_{Ov}, E_{IP} and E_{HD} are employed here to evaluate accuracy.

2.4.2.1 Efficiency Evaluation

As an important criterion, the efficiency of registration should be evaluated to guarantee that the algorithm can be applied to the general radiologist's workflow without causing undue delay. The elapsed times using different registration algorithms are displayed in Figure 2.9 and Table 2.1. It is obvious that the CPU time consumed by the method proposed here is less than those of the other two iterative algorithms (Algorithm 1 and ICP are configured to their default settings), since the proposed method considerably simplifies the calculation of scaling and rotation parameters. However, it is admitted that the time consumption of the proposed algorithm (average 3.7 iterations and 4

iterations for left lungs and right lungs respectively) is approximately three times that of Algorithm 2 due to the iterative nature of the method.

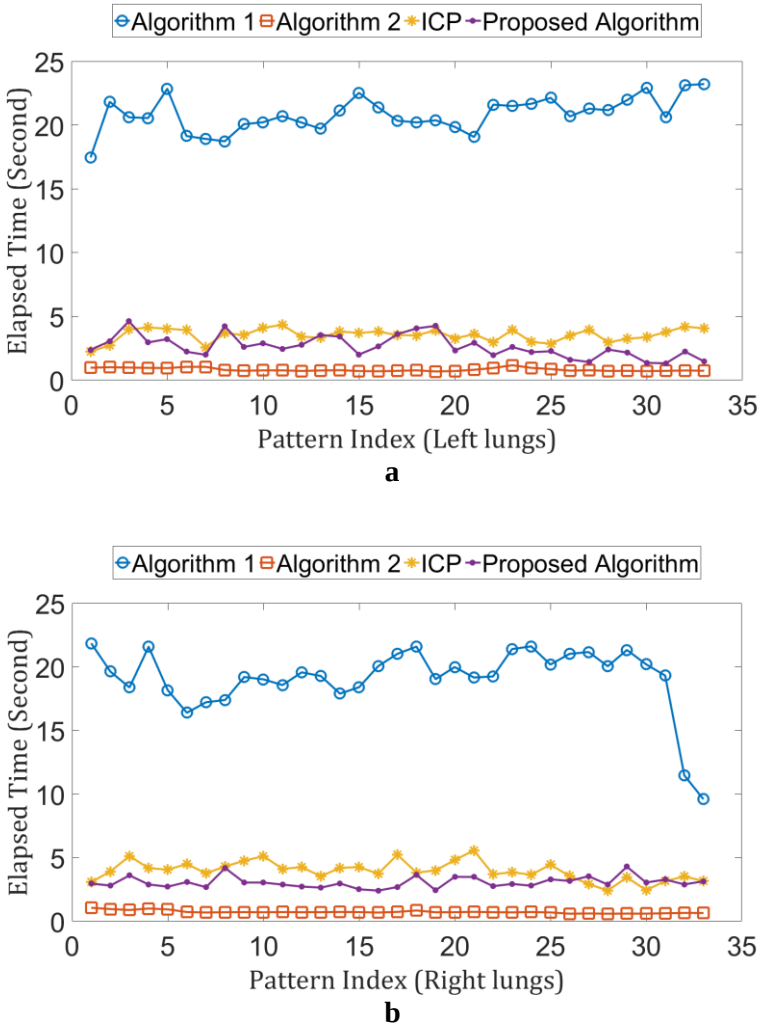


Figure 2.9 *Time consumption of 33 experimental patterns using four algorithms. (a) Average elapsed times of left lung registration. (b) Average elapsed times of right lung registration.*

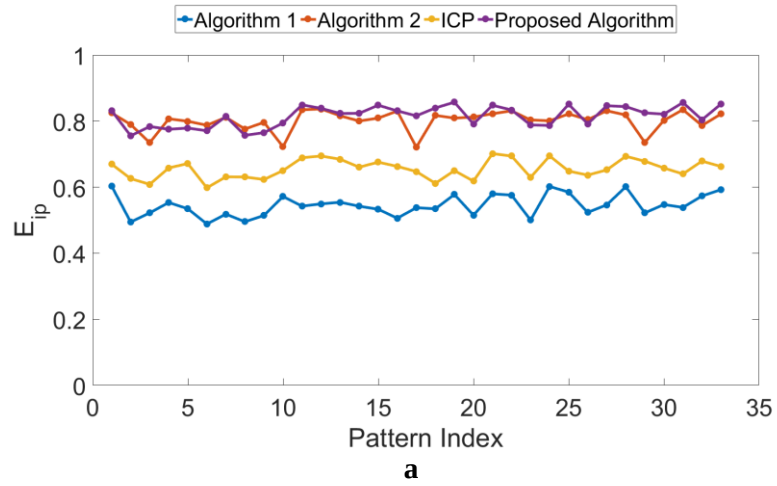
Table 2.1 The average elapsed times using different methods

	Algorithm 1 [sec]	Algorithm 2 [sec]	ICP [sec]	Proposed [sec]
Left lungs	19.08	0.74	3.55	**2.64**
Right lungs	20.83	0.83	3.96	**3.06**

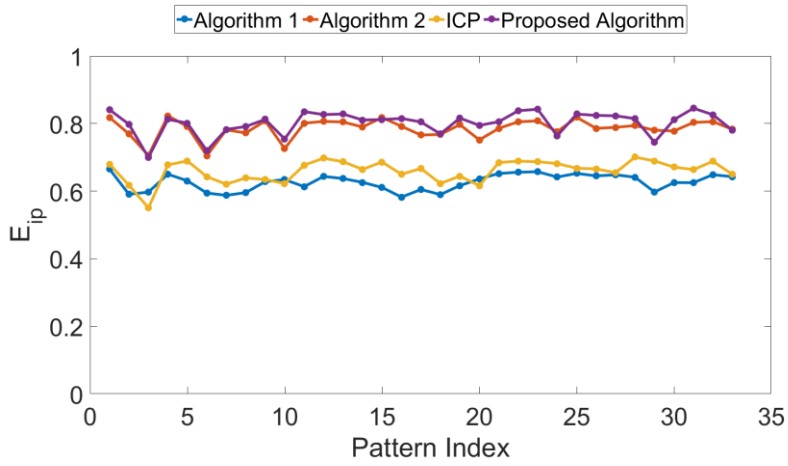
2.4.2.2 Accuracy and Stability Evaluation

The experimental results in terms of registration accuracy are illustrated in Figures 2.10 to 2.12. It is straightforward to see that the accuracies achieved by Algorithm 1 and ICP are much lower than those obtained by the other two methods since Algorithm 1 and ICP always fall into a local minimum when the alignment starts from an unfavourable pose, which is illustrated in Figure 2.13. Therefore, Algorithm 1 and ICP are not competent in shape registration with unfavourable or random initial poses. Conversely, the ability of overcoming the local minimum problem using Algorithm 2 and the proposed algorithm is validated in the batch-processing tests. Furthermore, compared with Algorithm 2, the algorithm proposed here develops the accuracy in the overwhelming majority of patterns.

The average E_{Ov}, E_{IP} and E_{HD} listed in Table 2.2 numerically demonstrate the improved accuracies are up to 7% and 11% for left and right lung alignments respectively. Further investigation about the source of improvement achieved by the proposed method can be performed by comparing the accuracies of Algorithm 2 and Proposed 1st. The "proposed 1st" presents the results obtained by the proposed framework without iterative scheme. A certain of improvement after first iteration achieved by the proposed method can be observed (measured by E_{HD}). Therefore, both spherical coordinate system-based cross-correlation and iterative technique are demonstrated to contribute to the improvement of registration performance.

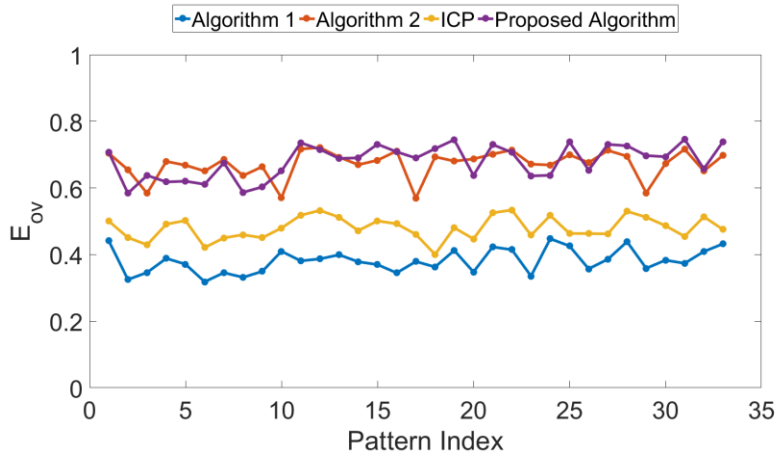


a



b

Figure 2.10 *The average E_{IP} obtained by four algorithms for 33 experimental patterns with random initial poses. (a) Results of left lung registration. (b) Results of right lung registration.*



a

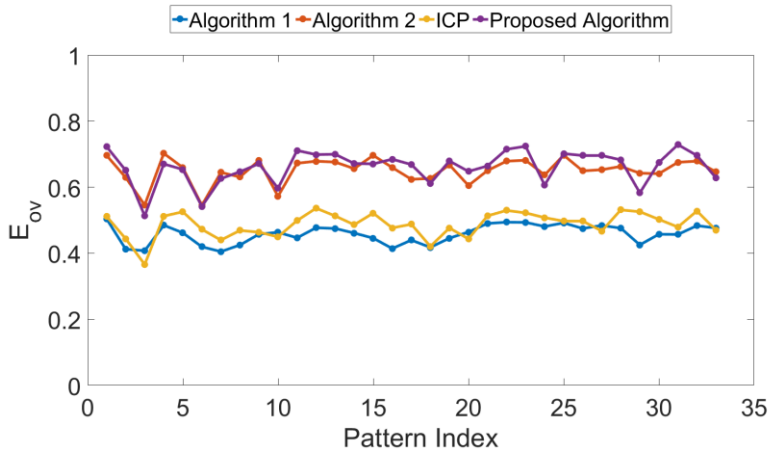


Figure 2.11 *The average E_{Ov} obtained by four algorithms for 33 experimental patterns with random initial poses. (a) Results of left lung registration. (b) Results of right lung registration.*

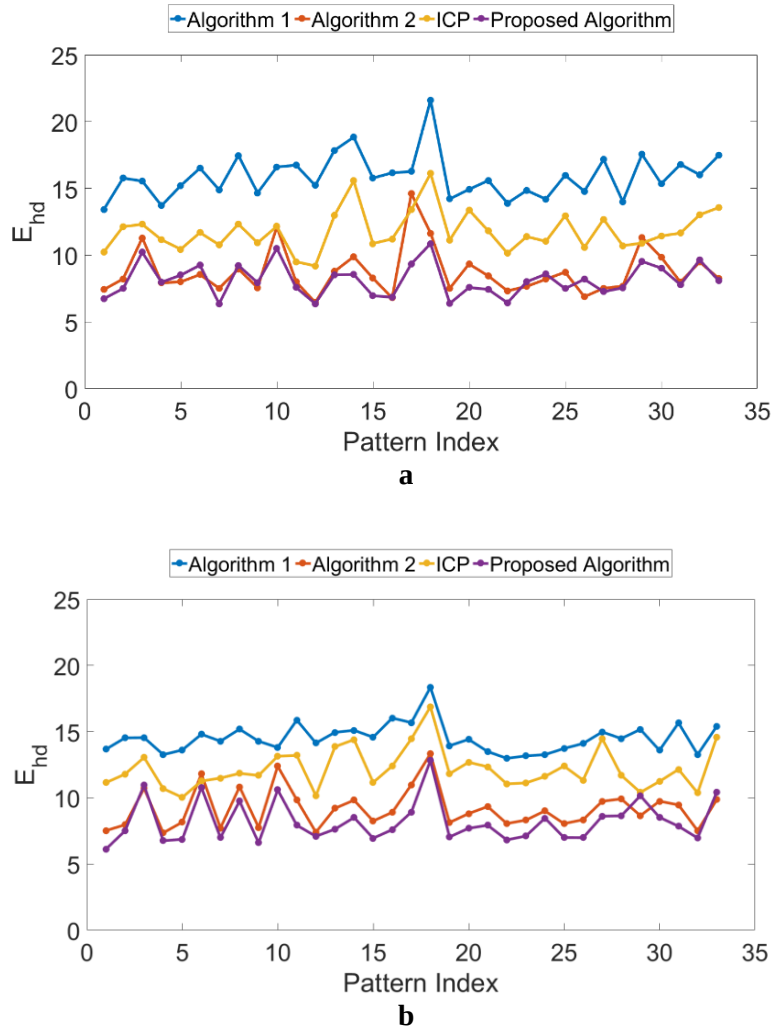
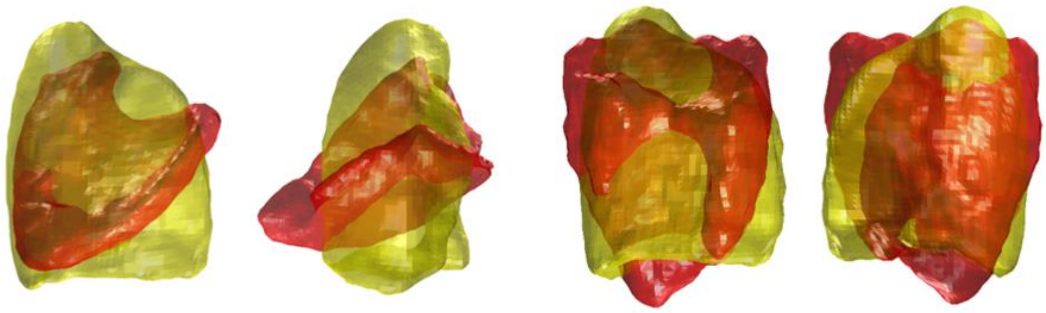


Figure 2.12 *The average E_{HD} obtained by four algorithms for 33 experimental patterns with random initial poses. (a) Results of left lung registration. (b) Results of right lung registration.*



a **b**

Figure 2.13 *An illustration of an unfavourable initial pose and the corresponding registration result using Algorithm 1 or ICP. (a) Two views of an unfavourable initial pose. (b) Two views of the registration result which converges to a local minimum.*

For similarity registration, the variations of lung shapes across different patients are the potential sources of variation in the final results, whereas statistical methods can be employed to further evaluate their performance. Lower standard deviations indicate that the results are insensitive to the variation of cases and rarely subjected to outliers. It is an important criterion for evaluating the performance of a registration technique. The standard deviations in Table 2.2 demonstrate that the algorithm proposed here substantially improves stabilities by up to 61% and 47% (measured by E_{Ov} and E_{IP}) for

Table 2.2 The average E_{Ov}, E_{IP} and E_{HD}, presented as mean±standard deviation (standard error)

		Algorithm 1	Algorithm 2	ICP	Proposed 1st	Proposed
Left	E_{Ov}	0.38±0.13 (0.02)	0.67±0.07 (0.01)	0.48±0.11 (0.02)	0.66±0.04 (0.01)	**0.68±0.04 (0.01)**
	E_{IP}	0.55±0.13 (0.02)	0.80±0.05 (0.01)	0.66±0.09 (0.02)	0.80±0.03 (0.01)	**0.82±0.03 (0.01)**
	E_{HD}	15.91±3.92 (0.68)	8.73±2.84 (0.49)	11.80±3.52 (0.61)	8.51±1.92 (0.33)	**8.13 ±1.77 (0.31)**
Right	E_{Ov}	0.46±0.10 (0.02)	0.65±0.08 (0.01)	0.49±0.11 (0.02)	0.66±0.05 (0.01)	**0.66±0.04 (0.01)**
	E_{IP}	0.63±0.09 (0.02)	0.79±0.06 (0.01)	0.66±0.09 (0.02)	0.80±0.04 (0.01)	**0.80±0.03 (0.01)**
	E_{HD}	14.49±3.35 (0.58)	9.18±2.57 (0.45)	12.18±3.65 (0.63)	8.77±2.11 (0.37)	**8.20±1.82 (0.32)**

left and right lung alignments respectively in comparison with Algorithm 2. The standard error in Table 2.2 can be used to measure the distance between the sample data mean and the true population mean. Lower standard errors indicate that for unknown subjects the superior results obtained by the proposed method is convinced.

2.5 Conclusions and Future Work

In this chapter, a robust and high-performance three-dimensional shape registration algorithm has been proposed. It employs the CFs to represent shapes. The radial moments and cross-correlation are exploited to estimate the scaling and rotation parameters. Also, an iterative method is proposed to improve the overall performance. The entire programming and simulation is implemented in Matlab based on the mathematical derivations.

The method proposed here was validated using two types of shapes: group shapes and lung shapes. The experiments on group shapes demonstrate its superior accuracy and robustness, even in the presence of noise. Over 2100 experiments on the lung shapes provided from a database demonstrate the excellent accuracy, stability and processing speed achieved by the method proposed here. In other words, the proposed similarity registration method is able to find the global optimum and has excellent performance in comparison with all the registration techniques investigated. Also, it strikes a significant balance between speed and performance, offering an innovative solution to the problem of volumetric shape registration.

In future research, a mathematical proof of the proposed algorithm's ability to find the global optimum can be investigated. The registered volumetric shapes can be employed to conduct statistical shape modelling in aid of prior shape-based image registration and segmentation.

Chapter 3

Single-modality Image Registration Using Prior Shapes

3.1 Introduction

3.1.1 Non-rigid Registration

Compared with rigid, similarity and affine registration, non-rigid registration is more complicated as it has a large number of degrees of freedom determined by the number of voxels in the image. In previous work, the transformation matrix for non-rigid registration has been derived using physical models or interpolation theory. A well-known deformable method based on the diffusion model was first proposed in [49], and is known as 'Demons'. The diffusion model is then augmented with an active force to increase the speed of convergence and therefore decrease the number of iterations [50]. On the other hand, B-spline-based Free-Form Deformation (FFD), which uses control points on the lattice to conduct the non-rigid alignment, is proposed in [51]. The FFD initially concentrates on the local deformation, and it is then refined by a hierarchical (global-to-local) method reported in [44]. The global-to-local registration is implemented by global rigid registration followed by local non-rigid alignment, which can obtain high accuracy with less processing time. In addition, since the deformations are conducted on local regions, the CPU and memory costs significantly decrease in accordance with time consumption.

In non-rigid registration, the similarity measurement of two images is prioritised to achieve high accuracy whilst the deformations need to be regularised properly. The diffeomorphic method was developed in [52] [53] [54]. It has been demonstrated that

diffeomorphic methods manage to preserve the image topology using an invertible and differentiable mapping. This method is always employed to align the images involved in substantial deformations, but does not need to be considered if moderate deformation is expected for image alignment. In addition, since the velocity field is dependent on the fluid model, the obtained displacement only concentrates on the location in the field rather than anatomical information. The diffeomorphic method without contour regularisation is therefore undesirable for medical imaging research [53].

3.1.2 Registration Based on Prior Knowledge

Regarding medical imaging analysis, the role of patient-specific prior knowledge is highlighted in [55]. Prior images obtained from previous diagnoses or research provide considerable anatomical information. Prior knowledge-based methods have therefore been studied as a way to enhance the performance of non-rigid registration.

Prior knowledge represented by a joint intensity distribution is proposed in [56] and [57]. Kullback-Leibler (KL) divergence is employed to measure the distance between the prior distribution and the joint intensity distribution of two aligned images. The optimisation process is driven by minimising KL divergence followed by maximising MI across two images. However, the displacement guided by the prior distribution still concentrates on the alignment similarity while neglecting to regularise deformations of source image during registration. In order to address this concern, the statistical deformation model (SDM) was first proposed in [22] and then developed to solve specific medical imaging problems in [58] and [59]. In order to exploit the potential of prior images, the feature-based deformable field is established as an atlas to regularise the registration [60]. A similar technique has recently been developed by [61], and termed as semi-supervised method. In the training process, the initial atlas built by supervised registration is used to guide the unsupervised registration and then combined

with the new atlas for the SDM. Nevertheless, patient movement during data acquisition tends to add artefacts to prior images. If the prior images employed by the SDM are mismatched or of low-quality, the imperfect atlas would mislead subsequent registration. On the other hand, the implicit representation of prior knowledge was investigated in [62] and [63]. The prior shapes were represented by level-sets to delineate the lung region, which considerably reduces manual labour and computational cost. In addition, since almost all the methods suffer from an issue that the number of training subjects is far less than the dimension of deformation representation, shrinkage estimation, singular value decomposition and wavelet-based decomposition have been investigated to tackle the problem [62] [63] [64].

The main contribution in this chapter is to propose a parameter-reduced SDM-based non-rigid registration method. The new model proposed here ends the controversy of the weighting parameter for the statistical term and properly regularises the deformable field based on prior shapes. This chapter is structured as follows: The derivation of new model is stated in Section 3.2. Experimental data collection and necessary pre-processing work is conducted in Section 3.3. The performance of the proposed method is evaluated in Section 3.4 with the necessary comparisons. Conclusions and future work are discussed in the last section.

3.2 Methodology

3.2.1 Statistical Deformation Model

The prior lung shapes are segmented from MSCT scans and represented by their SDFs. Although the size of each lung varies considerably between individuals, the influence made by the scale variance can be eliminated by employing the similarity registration method (proposed in Chapter 2) to pre-process prior shapes at the beginning.

The processed prior shapes in the training set are denoted by $\{X_1, X_2, X_3, \cdots, X_N\}$, where N denotes the number of training shapes. The mean prior shape represented by its SDF can be computed by

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{3.1}$$

As concluded in [65], the mean shape is over-smooth and lacking in lung shape details. It is therefore not recommended to use the mean shape as the target for the SDM. One of the prior shapes which is the closest to the mean shape is chosen as the target for the remaining shapes in the training set.

The non-rigid registration is then performed by the B-spline-based FFD method proposed in [51] to compute the deformations for each prior shape. Here, the deformation $\boldsymbol{u}$ is represented by a three-component vector field (u_x, u_y, u_z) along X, Y and Z directions, which is the function of coordinates (i.e. (x, y, z)). The PCA calculations are performed along X, Y and Z axes separately, whereas the notations in the subsequent derivations are written in vector form for convenience. The combination of deformations are denoted as $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \cdots, \boldsymbol{u}_{N-1}\}$. The mean displacement $\bar{\boldsymbol{u}}$ can be calculated by

$$\bar{\boldsymbol{u}} = \frac{1}{N-1}\sum_{i=1}^{N-1} \boldsymbol{u}_i \tag{3.2}$$

The mean-offset displacement matrices are reshaped to column vectors $\{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3, \cdots, \boldsymbol{\mu}_{N-1}\}$, where

$$\boldsymbol{\mu}_i = \boldsymbol{u}_i - \bar{\boldsymbol{u}} \tag{3.3}$$

Assuming the variability of the displacement obeys the normal distribution, it is then generalised by PCA. The eigen-decomposition of the covariance matrix, $\boldsymbol{\Sigma}_{cov}$, is computed by

$$\boldsymbol{\Sigma}_{cov} = \frac{1}{N-1}\boldsymbol{F}\boldsymbol{F}^{\mathrm{T}} = \boldsymbol{V}_M\boldsymbol{\Sigma}_M{\boldsymbol{V}_M}^{\mathrm{T}} \tag{3.4}$$

where M denotes the number of voxels in the shape matrix. The column vectors in $\boldsymbol{V}_M$ and $\boldsymbol{\Sigma}_M$ are eigenvectors and eigenvalues respectively and $\boldsymbol{F}$ is an $M \times (N-1)$ matrix denoted by

$$\boldsymbol{F} = [\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3, \cdots, \boldsymbol{\mu}_{N-1}] \tag{3.5}$$

As mentioned in the introduction, the issue that $(N-1)$ is far less than M leads to two problems in the calculation. Firstly, for a 128×128×128 three-dimensional image, the magnitude of M is over 10^6. Therefore, it is extremely complicated to calculate the eigenvectors and eigenvalues of the covariance matrix, which is an $M \times M$ unitary matrix. Secondly, even though eigenvectors and eigenvalues are obtained, only the eigenvalues existing in the first $(N-1)$ columns of $\boldsymbol{\Sigma}_M$ are useful, as the other columns are all zeros [66].

Given $\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F}/(N-1)$ instead of $\boldsymbol{F}\boldsymbol{F}^{\mathrm{T}}/(N-1)$, the dimension of the new covariance matrix considerably decreases to $(N-1) \times (N-1)$. The decomposition is rewritten as

$$\frac{1}{N-1}\boldsymbol{F}^{T}\boldsymbol{F} = \boldsymbol{W}_{N-1}\boldsymbol{\Sigma}_{N-1}{\boldsymbol{W}_{N-1}}^{\mathrm{T}} \tag{3.6}$$

where $\boldsymbol{\Sigma}_{N-1} = \mathrm{diag}({\sigma_1}^2, {\sigma_2}^2, \cdots {\sigma_{N-1}}^2)$ contains $(N-1)$ eigenvalues corresponding to $(N-1)$ non-zero eigenvalues in $\boldsymbol{\Sigma}_M$. Eigenvalues are the squared standard deviations. $\boldsymbol{W}_{N-1}$ is then used to calculate the eigenvectors, $\boldsymbol{V}_{N-1}$, as

$$\boldsymbol{V}_{N-1} = \boldsymbol{F}\boldsymbol{W}_{N-1}\boldsymbol{\Sigma}_{norm} \tag{3.7}$$

where $\boldsymbol{\Sigma}_{norm}$ is a diagonal matrix employed to normalise each column of $\boldsymbol{F}\boldsymbol{W}_{N-1}$.

It is therefore possible to represent each deformable field by a dimension-reduced vector $\boldsymbol{\lambda}_{N-1}$, which comprises $(N-1)$ values of coordinates along eigenvectors.

$$\boldsymbol{u}_{\text{model}} = \bar{\boldsymbol{u}} + \boldsymbol{V}_{N-1}\boldsymbol{\lambda}_{N-1} \tag{3.8}$$

3.2.2 Problems in Previous Work

In the interest of capturing statistical information about deformations, the statistical term $S(\boldsymbol{u})$ derived from prior knowledge is added into the cost function $E(\boldsymbol{u})$, together with the similarity metric $D(\boldsymbol{u})$ and smoothing term $R(\boldsymbol{u})$. The general cost function is written as

$$E(\boldsymbol{u}) = (1-\alpha)D(\boldsymbol{u}) + \alpha\big((1-\beta)R(\boldsymbol{u}) + \beta S(\boldsymbol{u})\big) \tag{3.9}$$

where α is used to adjust the weight of regularised terms in the cost function and β is employed to balance the penalties given by smoothing and statistical terms.

In the literature, such as [59] and [63], the distribution of prior deformations is estimated by a multivariate normal distribution:

$$f(\boldsymbol{u}) = c \cdot \exp\left(-\frac{1}{2}(\boldsymbol{u}-\bar{\boldsymbol{u}})^{\mathrm{T}}{\boldsymbol{\Sigma}_{cov}}^{-1}(\boldsymbol{u}-\bar{\boldsymbol{u}})\right) \tag{3.10}$$

where c is a constant and Σ_{cov} is the covariance matrix. The statistical term $S(\boldsymbol{u})$ used in the cost function is represented in association with Eq. (3.10),

$$-\ln\big(f(\boldsymbol{u})\big) \propto \frac{1}{2}(\boldsymbol{u}-\bar{\boldsymbol{u}})^{\mathrm{T}}{\boldsymbol{\Sigma}_{cov}}^{-1}(\boldsymbol{u}-\bar{\boldsymbol{u}}) = S(\boldsymbol{u}) \tag{3.11}$$

Nevertheless, two unsolved issues exist in this well-known framework. Firstly, in order to decrease the cost contribution, $S(\boldsymbol{u})$ penalises any displacement $\boldsymbol{u}$ that is not in accordance with the mean, $\bar{\boldsymbol{u}}$. However, by assuming that the deformations are associated with the normal distribution, any deformation within three standard deviations of the mean is conventionally acceptable. In addition, the empirically

determined weighting parameter β cannot guarantee that the resulting displacement $\boldsymbol{u}$ falls into a reasonable range. In previous research, since it is challenging to investigate a proper weighting parameter which can rigorously confine the displacement, β tends to be assigned roughly according to a specific case.

3.2.3 Parameter-reduced SDM-based Registration

In this part, single-modality image registration is used as an example for the derivation of the parameter-reduced SDM-based function. SSD is used as the similarity metric, and the distance between two images is calculated by

$$D(\boldsymbol{u}) = \int (I_{src}(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})) - I_{tar}(\boldsymbol{x}))^2 \mathrm{d}\boldsymbol{x} \tag{3.12}$$

where I_{tar} and I_{src} are the target and registered source images respectively. The desired transformation is normally derived by minimising $D(\boldsymbol{u})$. The numerical solution is however always ill-posed since the derivation can neither guarantee a unique solution nor avoid twisted deformations. In the previous articles, the second-order derivative of deformations is employed in [67] and [68] as a smoothing constraint, while the sum of squared first-order derivatives reported in [5] and [63] is used to regularise the displacement. The sum of squared first-order derivatives is adopted here for convenience, and is denoted as

$$R(\boldsymbol{u}) = \int |\nabla \boldsymbol{u}(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{x} \tag{3.13}$$

In Eq. (3.12), the displacement $\boldsymbol{u}$ is replaced by a deformation function $S(\boldsymbol{\lambda})$ manipulated by the SDM. The new similarity term is specified in Eq. (3.14).

$$D\big(S(\boldsymbol{\lambda})\big) = \int \Big(I_{src}\big(\boldsymbol{x} + S(\boldsymbol{\lambda})\big) - I_{tar}(\boldsymbol{x})\Big)^2 \mathrm{d}\boldsymbol{x} \tag{3.14}$$

where $S(\boldsymbol{\lambda})$ is a three-component vector and represented as:

$$S(\lambda_x) = \overline{u_x} + \sum_{i=1}^{K} V_{x_i} \cdot S_{reg}\big(\lambda_{x_i}\big) \tag{3.15}$$

$$S(\lambda_y) = \overline{u_y} + \sum_{i=1}^{K} V_{y_i} \cdot S_{reg}\left(\lambda_{y_i}\right) \tag{3.16}$$

$$S(\lambda_z) = \overline{u_z} + \sum_{i=1}^{K} V_{z_i} \cdot S_{reg}(\lambda_{z_i}) \tag{3.17}$$

where

$$\lambda_{x_i} = {V_{x_i}}^{\mathrm{T}} \cdot (u_x - \overline{u_x}) \tag{3.18}$$

$$\lambda_{y_i} = {V_{y_i}}^{\mathrm{T}} \cdot \left(u_y - \overline{u_y}\right) \tag{3.19}$$

$$\lambda_{z_i} = {V_{z_i}}^{\mathrm{T}} \cdot (u_z - \overline{u_z}) \tag{3.20}$$

$$S_{reg}(\lambda_{\mathrm{i}}) = H\left(\left|\frac{\lambda_i}{3\sigma_i}\right|^2 - 1\right) \cdot Sign(\lambda_i) \cdot 3\sigma_i + H\left(1 - \left|\frac{\lambda_i}{3\sigma_i}\right|^2\right) \cdot \lambda_i \tag{3.21}$$

Here, $\boldsymbol{\sigma}_i$ (i.e. ($\sigma_{x_i}, \sigma_{y_i}, \sigma_{z_i}$)) denote the standard deviation, which is the square root of the corresponding eigenvalue. K is the number of eigenvectors used for regularisation. Each $\boldsymbol{\lambda}_i$ (i.e.($\lambda_{x_i}, \lambda_{y_i}, \lambda_{z_i}$)) is regularised to guarantee that all the displacements fall into the range of $3\boldsymbol{\sigma}_i$ (i.e. obeying the Three-sigma Rule of the Gaussian distribution). The Heaviside function $H(\cdot)$ only penalises the $\boldsymbol{\lambda}_i$ which are outside of the aforementioned range. The sign function $Sign(\cdot)$ is employed to retain the sign of each regularised $\boldsymbol{\lambda}_i$. The property of $S_{reg}(\lambda)$ with respect to λ is illustrated in Figure 3.1.

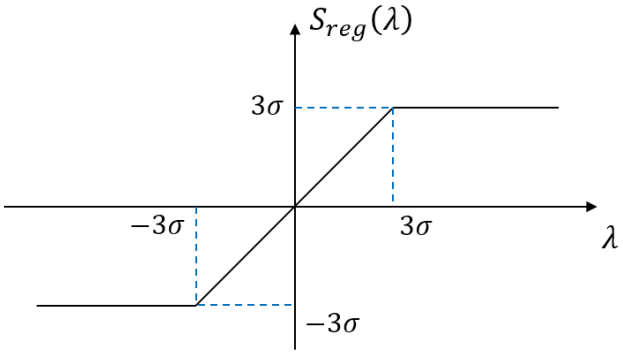


Figure 3.1 *Illustration of the function $S_{reg}(\lambda)$ with respect to λ.*

As stated in Section 3.2.1, the PCA calculations along three directions are conducted independently for convenience. Therefore, the available deformations represented by $(\lambda_x, \lambda_y, \lambda_z)$ are limited in a box region (i.e. $[-3\sigma_x, 3\sigma_x]$, $[-3\sigma_y, 3\sigma_y]$, $[-3\sigma_z, 3\sigma_z]$ along X, Y and Z directions respectively) rather than an ellipsoid.

The cost function is then updated in Eq. (3.22).

$$E(\boldsymbol{\lambda}) = (1-\alpha)D\big(S(\boldsymbol{\lambda})\big) + \alpha R\big(S(\boldsymbol{\lambda})\big) \tag{3.22}$$

where $\boldsymbol{\lambda}$ is a function of coordinates $\boldsymbol{x}$. The expected $\boldsymbol{\lambda}$ is obtained by minimising $E(\boldsymbol{\lambda})$ through a gradient-descent technique:

$$\hat{\boldsymbol{\lambda}} = \underset{\boldsymbol{\lambda}}{\operatorname{argmin}}\big(E(\boldsymbol{\lambda})\big) \tag{3.23}$$

The derivative of $E(\boldsymbol{\lambda})$ with respect to $\boldsymbol{\lambda}$ is computed by using the chain rule:

$$\frac{\partial E(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = \frac{\partial\big((1-\alpha)D(S)+\alpha R(S)\big)}{\partial S} \cdot \frac{\partial S(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} \tag{3.24}$$

Nevertheless, the derivative of $S(\boldsymbol{\lambda})$ does not exist if the argument of $H(\cdot)$ or $Sign(\cdot)$ is singular. $H(\cdot)$ and $Sign(\cdot)$ are regularised by Eq. (3.25) and Eq. (3.26) respectively [69],

$$H(p) = \frac{1}{2}\left(1 + \frac{2}{\pi}\tan^{-1}\frac{p}{\varepsilon}\right) \tag{3.25}$$

$$Sign(q) = \frac{2}{\pi}\tan^{-1}\frac{q}{\varepsilon} \tag{3.26}$$

and their derivatives are computed by

$$H'(p) = \frac{\varepsilon}{\pi(\varepsilon^2+p^2)} \tag{3.27}$$

$$Sign'(q) = \frac{2\varepsilon}{\pi(\varepsilon^2+q^2)} \tag{3.28}$$

where ε is the regularising parameter.

Therefore, the derivative of $S(\boldsymbol{\lambda})$ is computed as

$$\frac{\partial S(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = \sum_{i=1}^{K} \boldsymbol{V}_i \cdot \left(\frac{\varepsilon}{\pi(\varepsilon^2 + \boldsymbol{p}_i{}^2)} \cdot \frac{2\boldsymbol{\lambda}_i}{(3\boldsymbol{\sigma}_i)^2} \cdot Sign(\boldsymbol{\lambda}_i) \cdot 3\boldsymbol{\sigma}_i + H(\boldsymbol{p}_i) \cdot \frac{2\varepsilon}{\pi(\varepsilon^2 + \boldsymbol{\lambda}_i{}^2)} \cdot 3\boldsymbol{\sigma}_i + \frac{\varepsilon}{\pi(\varepsilon^2 + (-\boldsymbol{p}_i)^2)} \cdot \left(-\frac{2\boldsymbol{\lambda}_i}{(3\boldsymbol{\sigma}_i)^2} \right) \cdot \boldsymbol{\lambda}_i + H(-\boldsymbol{p}_i) \right) \tag{3.29}$$

where

$$\boldsymbol{p}_i = \left| \frac{\boldsymbol{\lambda}_i}{3\boldsymbol{\sigma}_i} \right|^2 - 1 \tag{3.30}$$

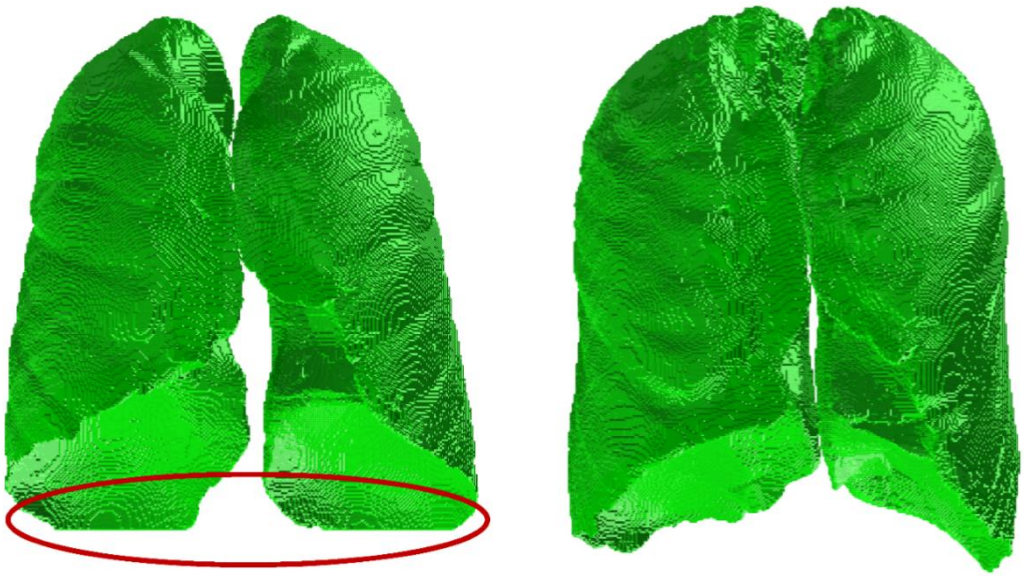
3.3 Data Collection and Pre-processing

3.3.1 Data Collection

In this section, 32 subjects are used for training and testing. The patients' details are listed in Table A-3 and A-4 of Appendix A. The high-resolution lung shapes were segmented using the Apollo software (Vida Diagnostics Inc, Iowa, USA). The voxel dimensions of the MSCT scan are not fixed in the XY plane in order to guarantee that the size of each slice is 512×512, whilst the spacing between the slices (along the Z axis direction) is always fixed to 0.5mm. The size of the SPECT scan is $128 \times 128 \times 128$ voxels and the voxel dimensions are fixed at 4.478 mm^3. In order to unify their dimensions, the MSCT scans can be down-sampled to match the $128 \times 128 \times 128$ voxels of the SPECT scans, but this sacrifices a lot of anatomical information. On the other hand, SPECT data can be up-scaled by using interpolation, but this strategy is not allowed since the value at each voxel represents the number of photons detected at that location. The actual voxel values in a SPECT scan are intended to be preserved and remain unaltered. Furthermore, as the resolution of the SPECT scan is very low, the required amount of interpolation would raise questions about the physical accuracy of the result. Therefore, down-sampling the MSCT scan is the only solution, even though some anatomical information is lost temporarily. Once the registration is completed,

the deformation field can be applied to map the structural information from the original MSCT data, such as the lobes, onto the SPECT image, thus circumventing the loss of anatomical information due to down-sampling.

In the lung database, eight pairs of lung shapes are unavailable for modelling as they are incomplete, as shown in Figure 3.2. Due to the irreparable defects in data acquisition, two edged parts at the bottom of the lungs are clipped. In order to guarantee the performance of the SDM, the eight subjects are not considered in the training process, but left aside as unknown subjects for future testing.



a **b**

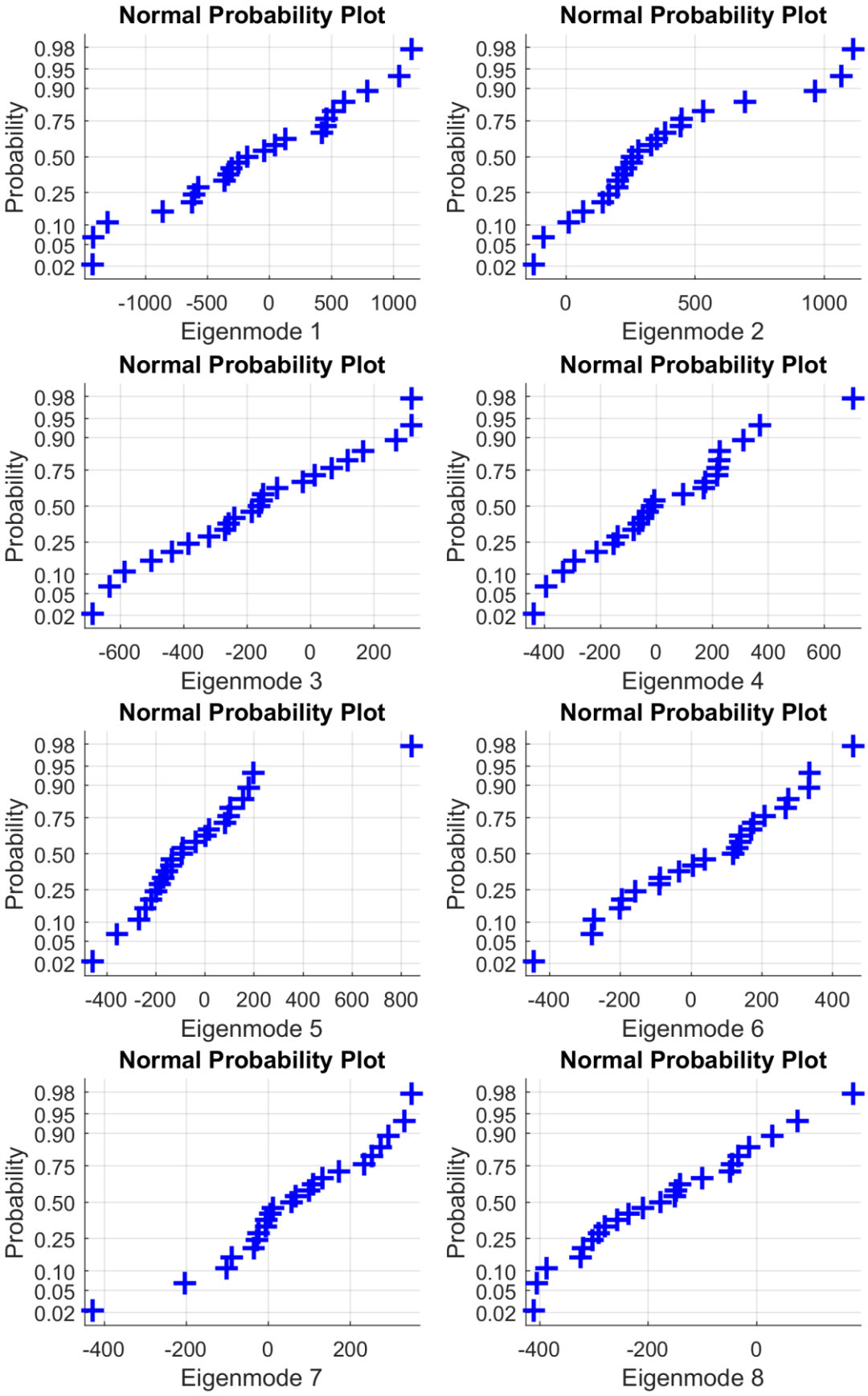
Figure 3.2 *Two pairs of lung shapes in the database. (a) The edged parts at the bottom of the lungs (marked by a red circle) are clipped in the raw CT scans. (b) Completed lungs in the raw CT scans.*

3.3.2 Normality Test

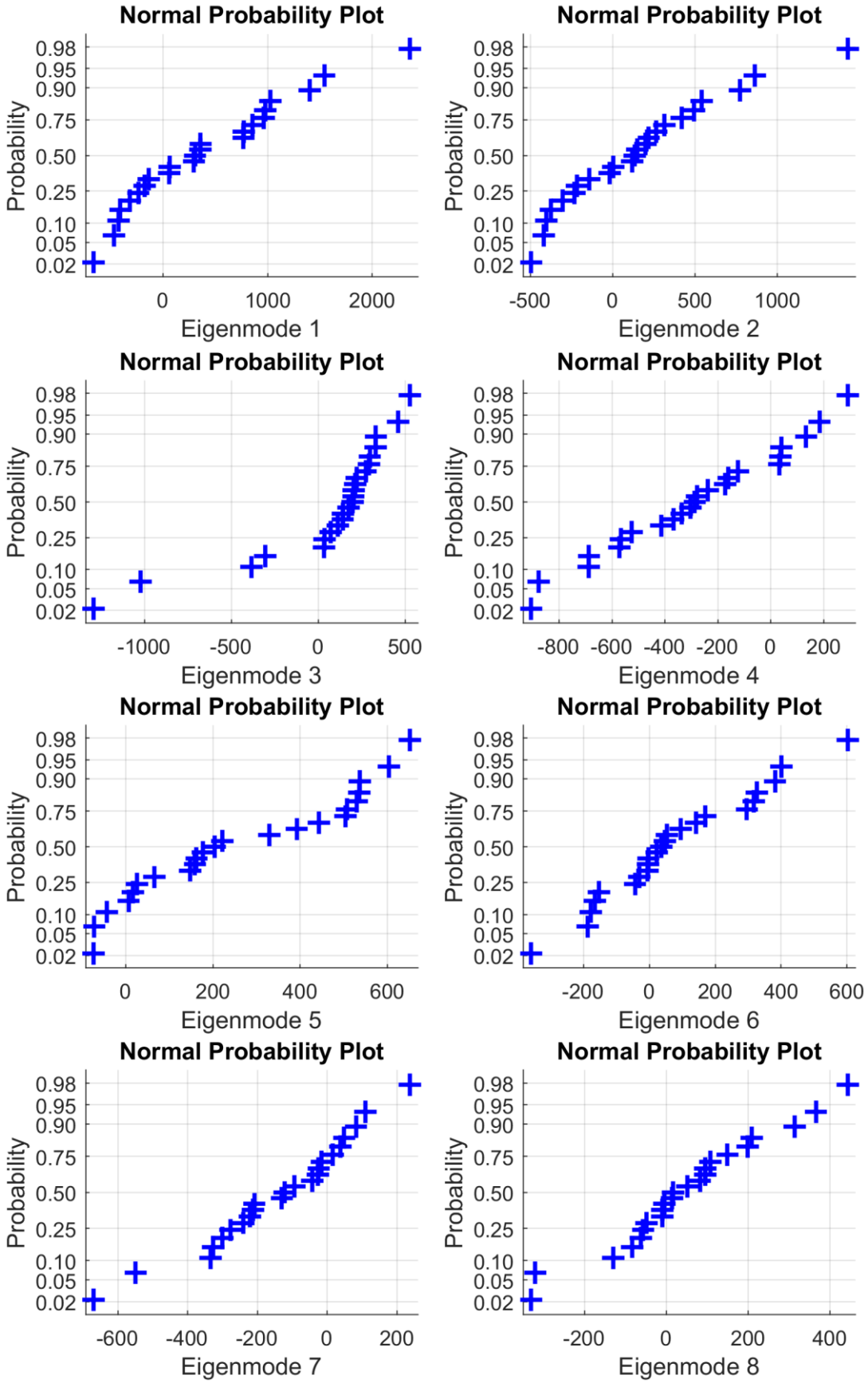
It is known that PCA is applicable only when the data for modelling are normally distributed [70]. In this part, a normality test is performed in 24 pairs of completed lungs, and the results are presented in a normal probability plot (NPP). NPP is widely used as an informal graphical technique to evaluate the normality of a dataset. It

presents a graph for the ordered observations from a dataset against the corresponding percentages from the normal distribution. Generally speaking, the distribution of sample data is approximately regarded as the normal distribution provided that the points form a fairly straight line. Departures from the line indicate the departures from normality. A more strict method to assess the near linearity of the plot is to calculate "correlation coefficient" for the ordered observations, which is equivalent to a formal Gaussianity test such as Shapiro-Wilk test. Here, Matlab built-in function 'normplot' is employed to plot the NPP. The NPPs of values for 24 $\boldsymbol{\lambda}$s corresponding to the eight largest eigenvalues along the X, Y and Z directions are illustrated in Figure 3.3(a-c) respectively. It can be observed that most of data are located in the vicinity of the straight line, apart from some outliers deviating from the line pattern. It is therefore concluded that the variances of the prior shapes follow the normal distribution and they are applicable to PCA modelling.

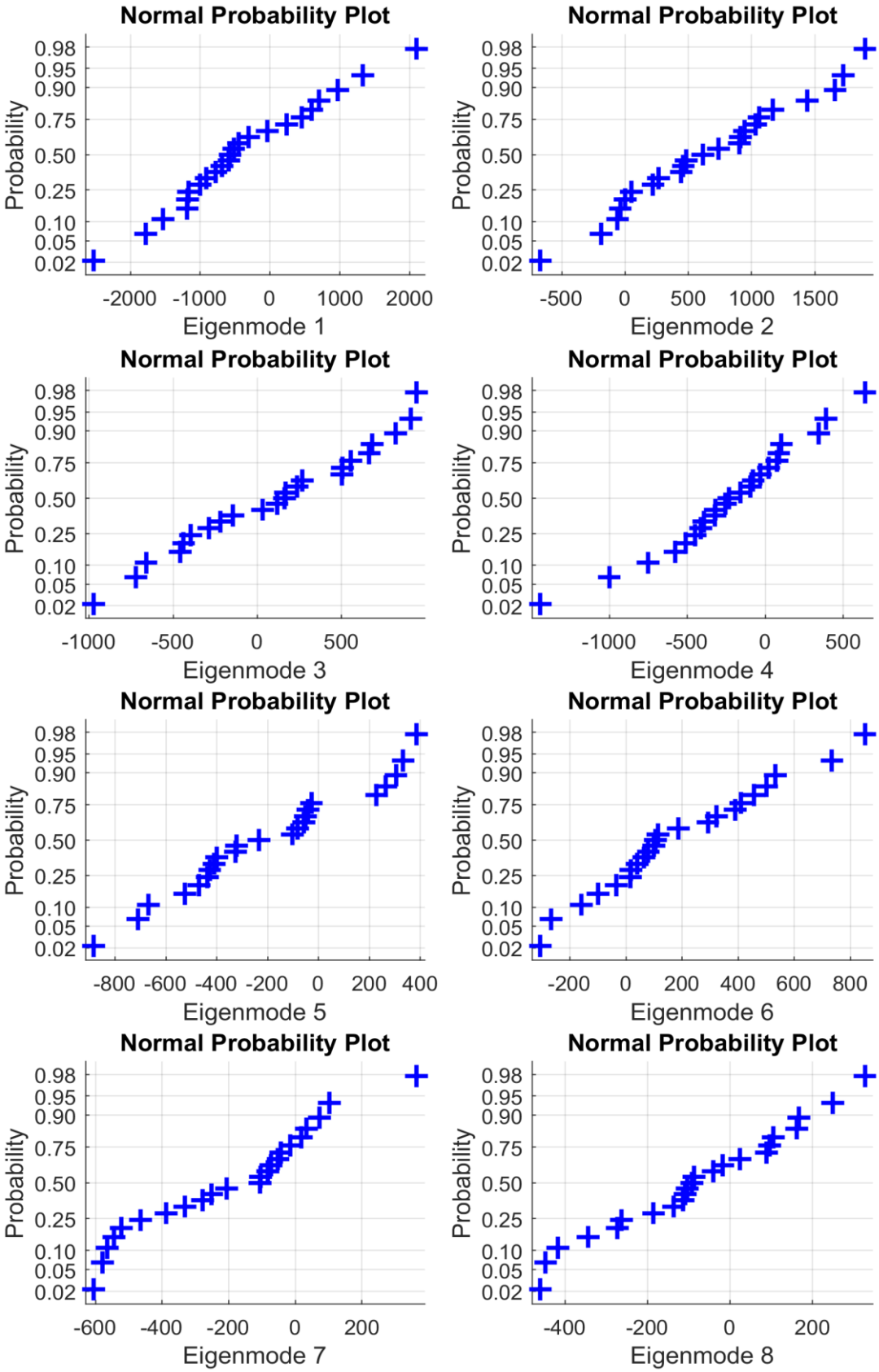
Here, 24 out of the 32 lung shapes are employed to train the SDM whereas the number of eigenmodes used for the SDM is determined by a ratio of the sum of the first K largest eigenvalues to the sum of all the eigenvalues. Provided that the cumulative ratio reaches up to 0.9-0.98, K can be regarded as the number of dominant eigenmodes (NoE) for the SDM [71]. As depicted in Figure 3.4, due to independent calculations along X, Y and Z directions, it is straightforward to observe the difference in these three subfigures. The choice of K needs to satisfy the aforementioned criterion on all three components. Since the cumulative ratios of five eigenvalues along three directions are all above 0.98, it is demonstrated that five or more eigenmodes are adequate to model the deformable field.


Normal Probability Plot
Probability
0.98
0.95
0.90
0.75
0.50
0.25
0.10
0.05
0.02
-1000 -500 0 500 1000
Eigenmode 1
Normal Probability Plot
0 500 1000
Eigenmode 2
Normal Probability Plot
-600 -400 -200 0 200
Eigenmode 3
Normal Probability Plot
-400 -200 0 200 400 600
Eigenmode 4
Normal Probability Plot
-400 -200 0 200 400 600 800
Eigenmode 5
Normal Probability Plot
-400 -200 0 200 400
Eigenmode 6
Normal Probability Plot
-400 -200 0 200
Eigenmode 7
Normal Probability Plot
-400 -200 0
Eigenmode 8

a


Normal Probability Plot
Probability
0.98
0.95
0.90
0.75
0.50
0.25
0.10
0.05
0.02
0
1000
2000
Eigenmode 1
Normal Probability Plot
Probability
-500
0
500
1000
Eigenmode 2
Normal Probability Plot
Probability
-1000
-500
0
500
Eigenmode 3
Normal Probability Plot
Probability
-800
-600
-400
-200
0
200
Eigenmode 4
Normal Probability Plot
Probability
0
200
400
600
Eigenmode 5
Normal Probability Plot
Probability
-200
0
200
400
600
Eigenmode 6
Normal Probability Plot
Probability
-600
-400
-200
0
200
Eigenmode 7
Normal Probability Plot
Probability
-200
0
200
400
Eigenmode 8

b



c

Figure 3.3 *Normal probability plots of the first eight values of 24* ***λ****s corresponding to the eight largest eigenvalues. (a)NPP of the first eight values of 24* ***λ****s along the X direction. (b)NPP of*

*the first eight values of 24 **λ**s along the Y direction. (c)NPP of the first eight values of 24 **λ**s along the Z direction.*

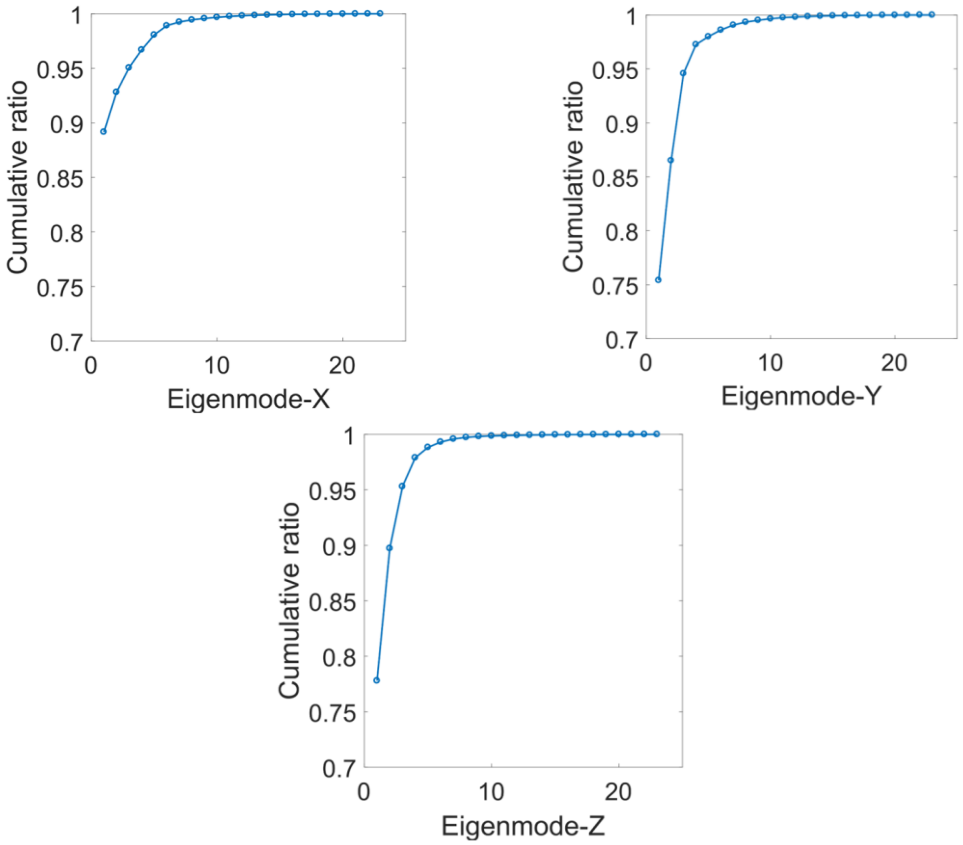


***Figure 3.4** Cumulative ratios for 23 eigenvalues along three directions.*

3.3.3 Transform to Reference Domain

The training work starts with similarity registration to a target prior shape. Therefore, the established SDM is only applicable to the domain of the target prior shape, which is defined as a reference domain here. Before any SDM-based registration, the unknown target and source images, namely the testing images here, are required to be transformed to the reference domain [61]. The transformation parameters can be computed by similarity registration from any testing image to the target prior shape.

3.4 Experimental Results and Discussion

In this section, the proposed parameter-reduced algorithm is validated using lung shapes and single-modality images. The following experiments are divided into two parts. Firstly, the proposed framework and a non-learning method are tested using synthetic defective lung shapes for comparison. Visual inspection is employed to evaluate their performance [72]. Secondly, two SPECT images are used to test two learning-based techniques. Normalised inner product (NIP) is employed for similarity measurement.

3.4.1 Lung Shape Registration

The experiments are performed over eight unknown subjects using a leave-one-out method. Before registration, one lung shape is chosen as the target, from which one of the lobes is manually removed. As shown in Figure 3.5, the right lower lobe of the target shape is removed. The aim is to align the remaining source lung shape to the defective target lung shape using NiftyReg and the method proposed here. As a widely used non-rigid registration method proposed in [73], the code for NiftyReg is freely

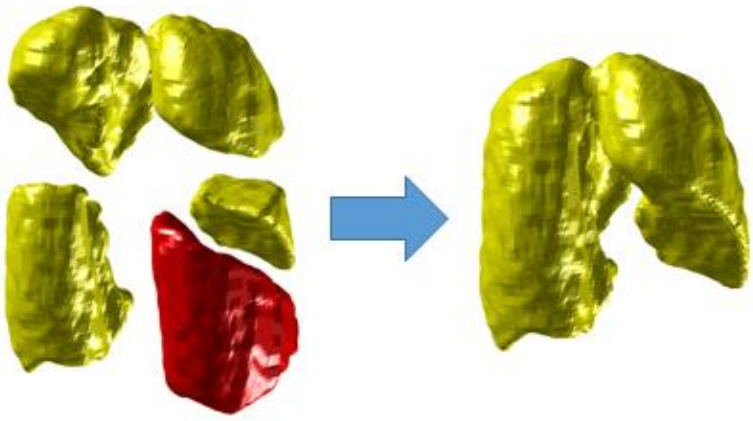
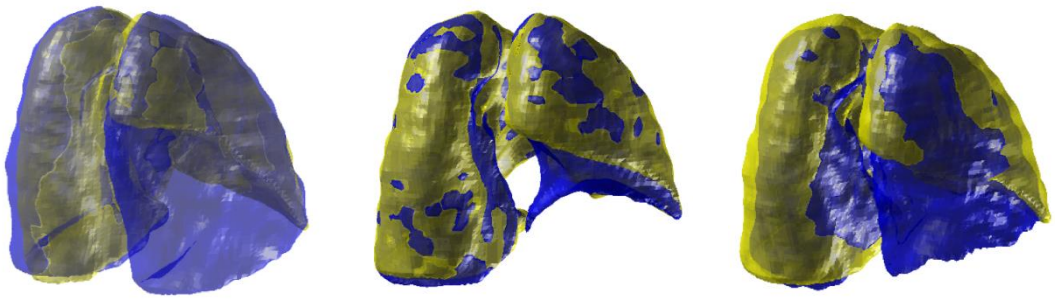


Figure 3.5 *The right lower lobe (highlighted in red) is removed from the target lung shape (shown in yellow) to create a synthetic defective lung shape (shown on the right).*

available. Compared with the proposed method, NiftyReg does not employ prior knowledge to constrain the unreasonable deformations. In the following experiments, NiftyReg is configured using its defaults. The graphical results are illustrated in Figure 3.6. All the lung shapes in this section are shown in posterior view.



a **b** **c**

Figure 3.6 *Lung shapes are shown in yellow for the target and blue for the source. (a) Initial poses of two lung shapes. (b) Registration using the NiftyReg method. (c) Registration using the proposed method with prior knowledge.*

By visual inspection, it is obvious that the SDM-based method manages to constrain excessive deformation. The lung shape is maintained during non-rigid registration. Conversely, the NiftyReg method fails in this test, as it concentrates on minimising the dissimilarity of the two shapes. In order to perform numerical comparisons, the right lower lobe, which is removed from the target lungs before registration, is used to calculate the lobar overlap after registration. The numerical results with respect to the mean lobar overlap of seven tests are listed in Table 3.1. The overlap achieved by the method proposed here is approximately three times higher than that of the NiftyReg method, which demonstrates that the proposed method can prevent excessive deformation and maintain a relatively reasonable lung shape during alignment.

Table 3.1 Average lobar overlap evaluation for eight test subjects using two methods.

	SUB.1	SUB.2	SUB.3	SUB.4	SUB.5	SUB.6	SUB.7	SUB.8
NiftyReg	0.1925	0.1949	0.1127	0.1760	0.1342	0.1207	0.1292	0.1350
Proposed	0.5613	0.5505	0.3353	0.5059	0.3882	0.3305	0.4414	0.3946

3.4.2 Single-modality Image Registration

In this section, two perfusion SPECT images are employed to validate two learning-based methods. The coronal-view slices of two subjects are displayed in Figure 3.7. Compared with a healthy subject (Figure 3.7(a)), it is straightforward to observe the reduced uptake region corresponding to the location of a lesion in Figure 3.7(b), which is a good example to test the performance of learning-based methods. The defective and the normal perfusion SPECT scans are regarded as the target and source images respectively for the following registration.

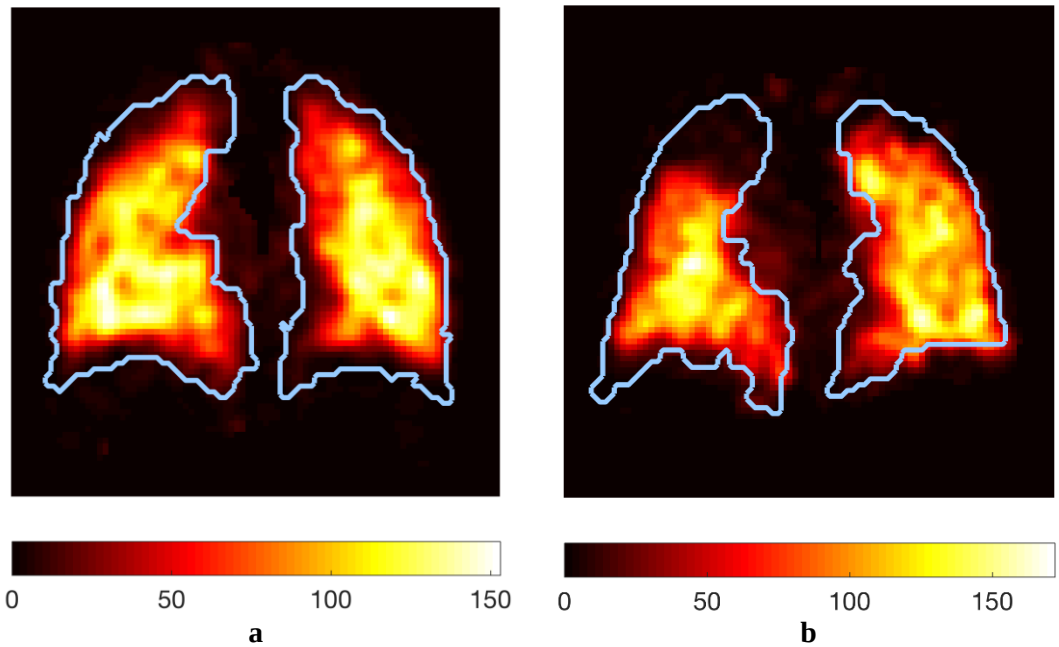
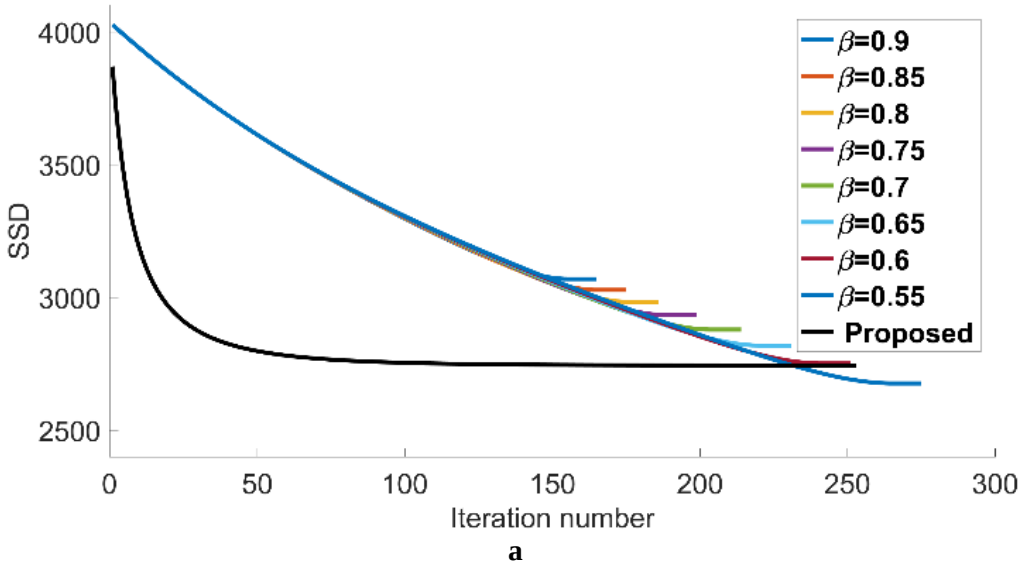
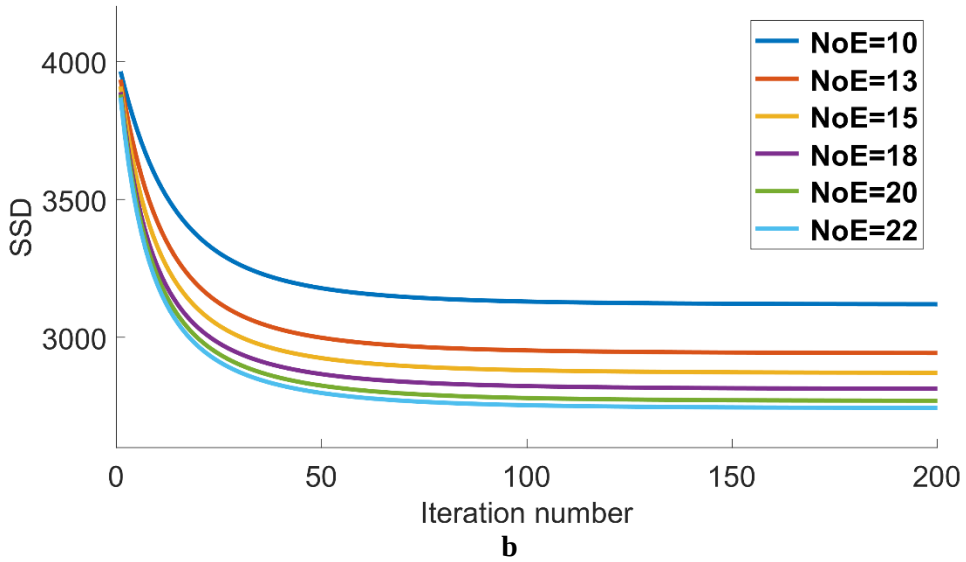


Figure 3.7 *Perfusion SPECT scans of healthy and a moderate COPD subjects. The light blue curves denote the actual lung shape boundaries. (a) SPECT image of the healthy subject. The intensities inside the lung region are bright and relatively uniform. (b) SPECT image of the COPD subject. The intensities in the right upper lung are relatively low because the tissue at this location is not active, resulting in reduced uptake of the radio-isotope.*

The method proposed in [59] is used as a reference method for comparison, named Ref here. The main framework proposed by Ref is stated in Eq. (3.9), where the statistical term is added into the cost function together with the similarity metric and smoothing term. Gradient descent is employed to optimise the deformation. In Eq. (3.9) the parameter α is empirically assigned a moderate value (i.e. 0.5) in both the Ref and proposed methods. Also, it is necessary to mention that β influences the trend of the convergence curve significantly, as shown in Figure 3.8 (a). It is apparent that the



a



b

Figure 3.8 *(a) Convergence curves of the reference method with various β and the proposed method (in black). NoE=22 (b) Convergence curves of the proposed method with various NoE.*

method proposed here converges faster than the Ref method. As the degree of freedom of deformation is dependent on NoE, more displacement flexibility can be obtained with the increase of NoE, which contributes to superior registration accuracy (i.e. lower SSD), as shown in Figure 3.8 (b).

As β varies greatly with different images, the optimal β is not expected to be investigated here. As illustrated in Table 3.2, the SSDs at convergence for the Ref method with β ranging from 0.6 to 0.9 are sampled. These β values produce SSDs that can be compared with the method proposed here. In the case that the two methods achieve the same SSD, the NIP, specified in Eq. (2.28), is employed to evaluate the registration accuracy. It is straightforward to observe that the NIP obtained by the method proposed here is always higher than that achieved by the Ref method. The experimental results demonstrate that the method is more flexible during deformation and therefore obtains superior alignment accuracy.

Table 3.2 Evaluation of registration accuracy using the reference method, with β ranging from 0.6 to 0.9, and the proposed method.

β	SSD at convergence	NIP (Ref)	NIP (Proposed)
0.6	2765	0.8489	**0.8509**
0.65	2833	0.8487	**0.8523**
0.7	2894	0.8492	**0.8521**
0.75	2949	0.8497	**0.8527**
0.8	2999	0.8502	**0.8532**
0.85	3044	0.8505	**0.8537**
0.9	3086	0.8510	**0.8540**

3.5 Conclusions and Future Work

A non-rigid registration method based on SDM is exploited in this chapter. In order to overcome the shortcomings in previous work, a novel parameter-reduced registration method is proposed. In the new cost function, the SDM is encapsulated into the similarity metric and the smoothing term, which avoids using the empirically chosen weighting parameter for the statistical term. The proposed framework is validated to be capable of preventing excessive deformation and of regularising the deformation reasonably for registration of defective images. The experimental results demonstrate that the proposed framework outperforms a similar state-of-the-art method in terms of the convergence rate and registration accuracy.

The method proposed here intends to solve the problem of single-modality image registration. It will be extended for multi-modality image fusion in Chapter 5. In the future, with more prior shapes added into the training set, the SDM will be established with more structural information and will obtain more flexibility to achieve higher performance.

Chapter 4

Synthetic Data Collection

4.1 Introduction

Experimental data are very important for validating the performance of the proposed method. In fact, the lung SPECT scans provided by Southampton General Hospital were acquired without ground-truth boundaries for the lung region. An advanced software is therefore employed to synthesise the perfect intra-patient data for later use. In addition, apart from the result-oriented evaluation methods, such as the MI of the aligned images and the overlap of the aligned regions of interest (ROIs), which have been frequently used before, the process-oriented evaluation, such as the distance to the ground-truth displacement, is also a gold-standard to validate non-rigid registration.

4D extended cardiac-torso (XCAT) phantom was initially developed to provide simulated human anatomy using non-uniform rational B-spline (NURBS) [74]. XCAT was originally named NCAT and created for nuclear medicine research, especially for SPECT and positron emission tomography (PET). It was then extended to the latest version, XCAT, for high-resolution anatomical imaging and advanced simulation of cardiac and respiratory motions, which are close to the scenario of real patients. The XCAT program is designed to produce different outputs in five different modes. According to the specific research purpose, modes 0 and 4 are used to generate body phantom data and the vector displacement (VD) of each voxel on the phantom respectively.

In mode 0, the software is capable of synthesising two physical models: a three-dimensional attenuation phantom with the distribution of attenuation coefficients for a

configured photon energy and a three-dimensional activity phantom with the distribution of emission radionuclide activity for the various organs. Each of these models is imaged by a voxelised phantom matrix with a customised resolution. As the phantom is mathematically defined, no error is introduced even though the phantom is generated at any resolution. The phantom data are stored in raw 32-bit floating point binary files (little endian) with no header. The voxelised phantom can be used in conjunction with analytical or Monte Carlo-based models of the imaging process to synthesise transmission (e.g. X-ray, CT) and emission (e.g. SPECT, PET) imaging data.

In mode 4, the output vectors are constructed by the actual sampled points from the phantom objects and interpolated vectors from these sampled points. The interpolated vectors are computed by averaging neighbouring vectors. By increasing the number of actual sampled points, the accuracy of the vector output is improved accordingly.

XCAT runs with a parameter file which is used to configure the synthesis process. In order to properly initialise the input distribution, the radionuclide multimodality dosimetry package (RMDP) is employed here in cooperation with XCAT. RMDP is intended to model the patient distribution using the International Commission on Radiological Protection lung model, which is used to calculate the absorbed dose in each lung per decay of radionuclide. In addition, RMDP offers an interactive window to add lesions and tumours to the phantom to simulate patients' SPECT imaging. Also, phantom reconstruction is conducted by RMDP to obtain synthetic MSCT and SPECT scans.

In this chapter, the procedure for constructing synthetic data is reported. It specifies the operating modes of XCAT, the configuration of simulated scanning parameters and the reconstruction of synthetic data using RMDP. Furthermore, it illustrates the method for adding artificial defects to the lungs. The structure of this chapter is as follows: The

modes of XCAT used in this chapter and the cooperation with RMDP are briefly introduced in Section 4.1. The procedure for creating the synthetic phantom data using XCAT and RMDP is specified in Section 4.2. The interactive method to add defects to the synthetic data is illustrated in Section 4.3. Conclusions finalise the chapter.

4.2 Phantom Data Processing

4.2.1 Phantom and VD Generation

Before data construction, the breathing cycle is configured over five seconds, starting from maximum inspiration. The attenuation and activity phantoms are sampled with eight frames within one breathing cycle at even intervals, as shown in Figure 4.1. Since each XCAT simulation can only generate one type of activity phantom (synthetic V or Q), two simulations are performed to obtain the raw activity phantoms for synthetic V and Q data. In other words, the parameter files need to be separately configured twice by RMDP according to user-defined simulated scanning parameters, as listed in Table 4.1.

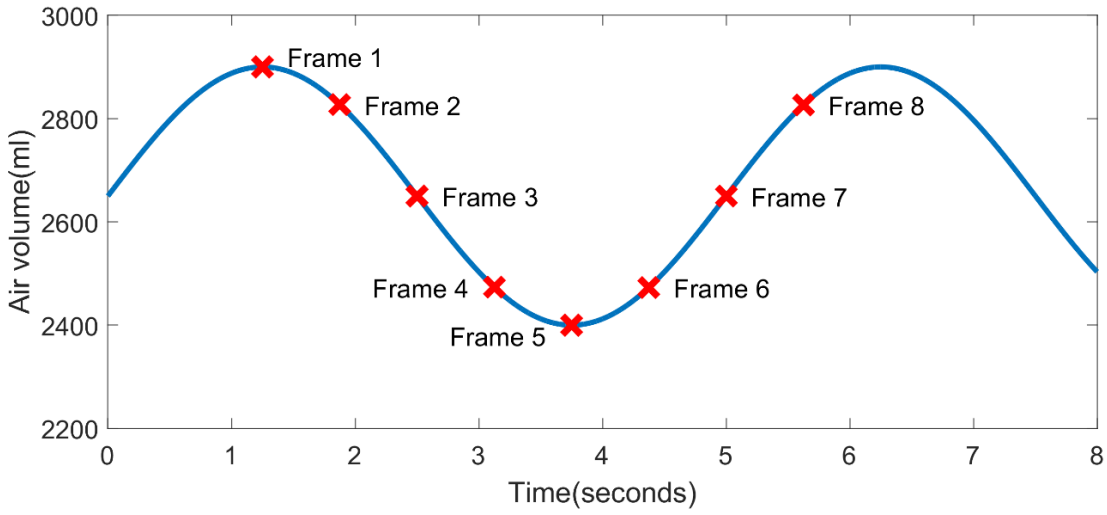


Figure 4.1 *XCAT phantoms are sampled with eight frames within one breathing cycle.*

Table 4.1 User-defined simulated scanning parameters for RMDP.

	Synthetic V	Synthetic Q
Ventilation activity (MBq)	50	0
Perfusion activity (MBq)	0	200
Injection time (min)	0	0
Scan time (min)	10	10
No. of projections	120	120
Seconds per view	100	100
Camera sensitivity (cps/MBq)	120	120

The main differences between the two RMDP configuration parameter files are listed in Table 4.2. It is straightforward to observe that the "lung activity" parameters for synthetic V and Q are the same. In fact, as the lung activity imaging for ventilation SPECT is based on the detection of gamma-ray photons emitted from the airway tree, it is more reasonable to represent real lung activity by "airway tree activity" rather than by "lung activity". Nevertheless, the XCAT airway tree model fails to delineate the terminal branches of the lung tree in detail, as shown in Figure 4.2. "Lung activity" is therefore configured to compensate for the distortion in lung activity presentation.

VD is generated using another mode of XCAT. Each VD file reports the displacement of each imaging voxel from the first frame to one of the subsequent frames. The ground-truth displacement is referred for validation in the next chapter.

Table 4.2 Organ activity parameters calculated by RMDP for XCAT phantom generation.

	Synthetic V	Synthetic Q
Body activity	0.0159945	0.0586195
Liver activity	0.0186639	0.312084
Gall bladder activity	0.000000	0.312084
Lung activity	**0.435841**	0.435841
Stomach wall activity	0.0213334	0.000000
Kidney activity	0.156287	0.000704156
Airway tree activity	0.321665	0.000000

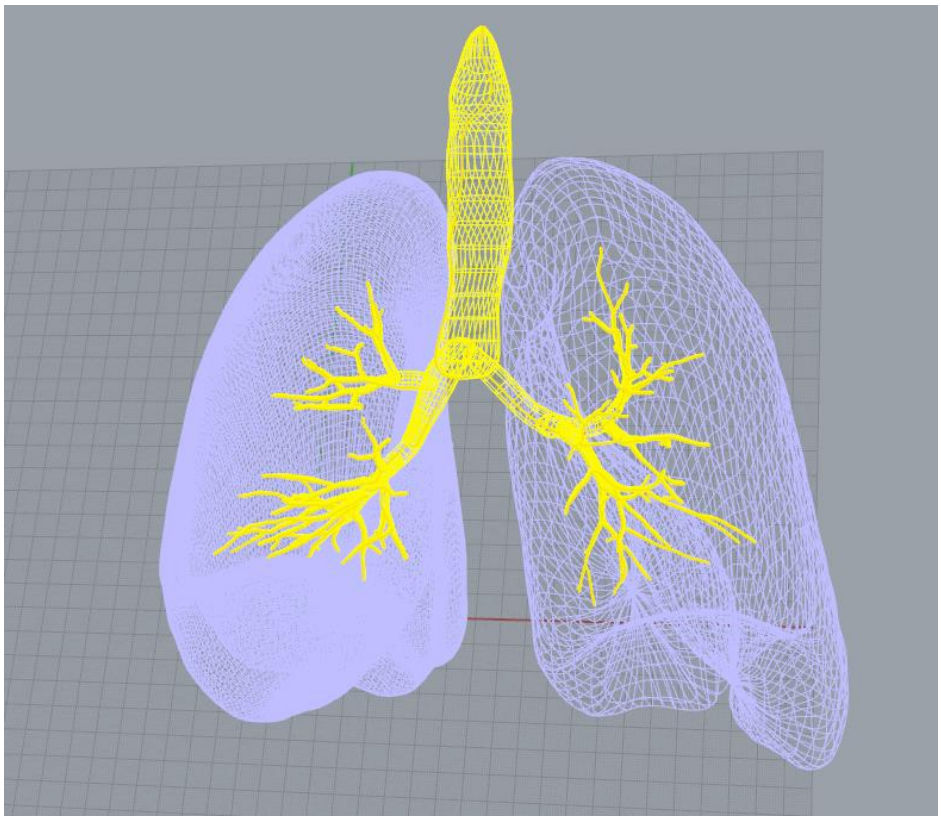
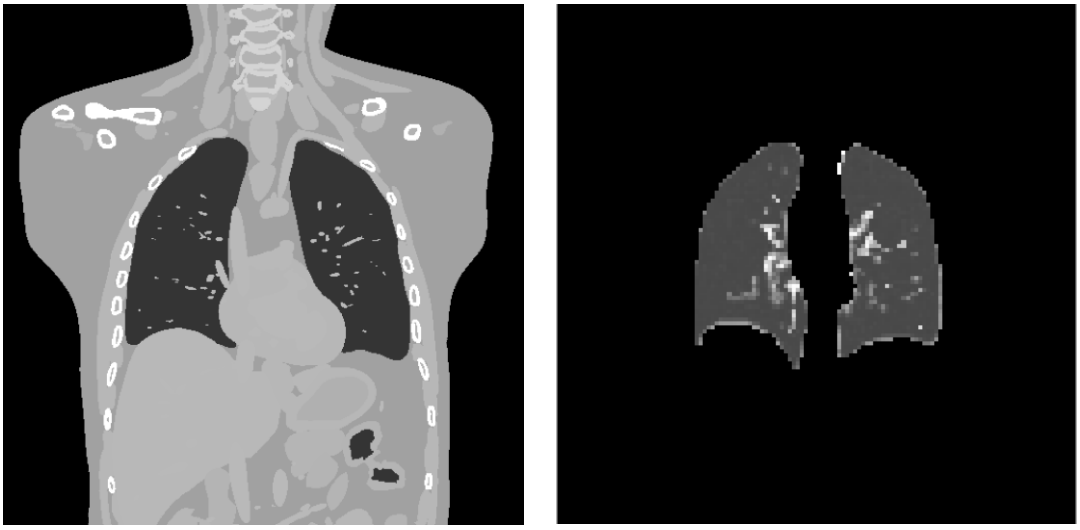


Figure 4.2 *XCAT lung tree model (template) used to generate phantoms. The terminal branches of the lung tree (i.e. bronchus) are rarely displayed, which results in the fact that the ventilation activity phantom has to be compensated by 'lung activity'.*

4.2.2 Phantom Post-processing

4.2.2.1 Synthetic MSCT scans

The attenuation phantom created by XCAT closely resembles the MSCT scan, but it is noise-free, as depicted in Figure 4.3(a). It is reported in [75] that the actual noise characteristics of CT obey the Poisson distribution, where the rate parameter varies with different CT scanners. Here, the Poisson distribution parameter is estimated according to the forthcoming MSCT scans. As shown in Figure 4.4, a part of the background is sampled from each MSCT scan. Assuming the mean noise histogram of all MSCT scans obeys the Poisson distribution, the rate parameter is calculated to be 0.0058 (normalised), which is adopted to add noise to the 'perfect' attenuation phantoms.



a **b**

Figure 4.3 *Synthetic MSCT data. (a) Synthetic attenuation phantom. (b) Synthetic MSCT data after segmentation, down-sampling and transforming to the reference domain.*

In order to enhance the performance of image registration, it is necessary to roughly segment the ROI first of all. As shown in Figure 4.3(a), a normal MSCT scan always contains irrelevant grey-scale information outside the lung region which needs to be removed. The method proposed in [76] can be employed for automatic lung segmentation. Furthermore, the segmented MSCT scans need to be down-sampled to

the same resolution ($128 \times 128 \times 128$) as the SPECT scans. One coronal-view slice with Poisson noise after transforming to the reference domain is shown in Figure 4.3(b).

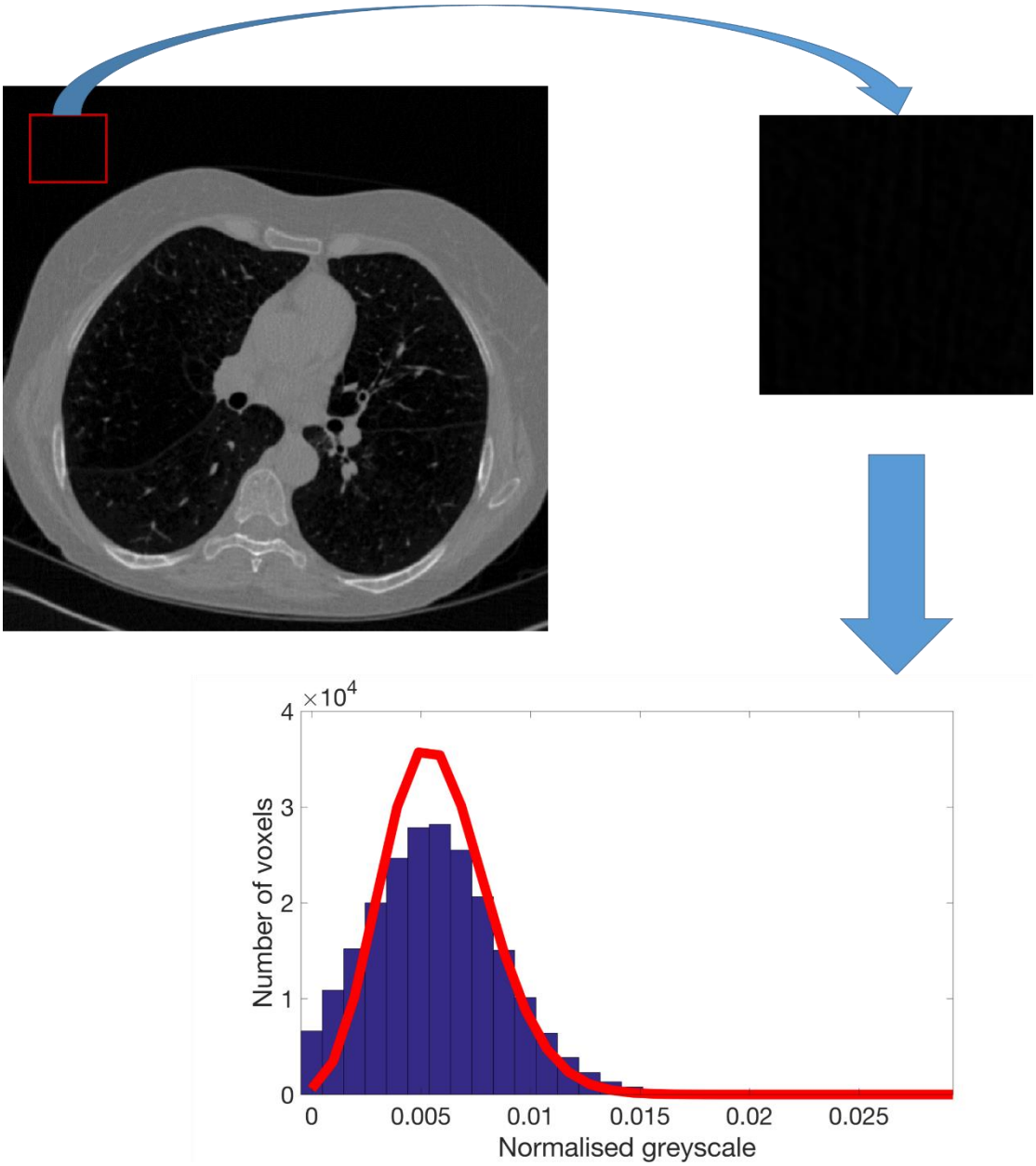


Figure 4.4 *Flow chart for determination of the Poisson distribution parameter. A part of the background is sampled followed by the parameter estimation from the histogram. The red curve indicates the fitted curve of the Poisson distribution.*

4.2.2.2 Synthetic SPECT scans

The activity phantom created by XCAT presents a perfect distribution of activity. As shown in Figure 4.5(a) and (b), the activity is uniformly distributed in each organ. The RMDP is used to post-process the raw data [77]. This dosimetry package incorporates a toolkit to simulate realistic gamma camera image data, derived from this 'perfect' input. Clinically realistic count-rates and system sensitivities are used to apply noise distributions, along with modelling of photon transport, truncation and limited angular sampling of the imaging system. As shown in Table 4.1, the raw phantom is imaged in 120 projections over 360 degrees of rotation, which resembles the actual functional imaging data acquired by two cameras over 60 phases. Then, the applicable real three-dimensional synthetic data (functional imaging matrices) need to be reconstructed from the 120 projections. The reconstructed SPECT images (including Poisson noise) are displayed in Figure 4.5(c) and (d).

Even though the boundaries of the lungs in SPECT imaging are very blurred, the properties of V and Q imaging make the minority of active organs visible (e.g. kidney and trachea in SPECT V and liver in SPECT Q), which greatly decreases the difficulty of segmentation. Coarsely segmented SPECT scans after transforming to the reference domain are shown in Figure 4.6.

4.3 Artificial Defects

The V and Q of a patient with moderate COPD may demonstrate photopenic regions, corresponding to reduced activity, and hence lower ventilation or perfusion, at the location of lesion, as shown in Figure 4.7. These abnormal SPECT images increase the challenge of alignment.

In order to effectively validate the performance of various methods in the next chapter, RMDP is employed to interactively add defects to the obtained phantoms. Here, the 'cold defect' is chosen to add a photopenic region to synthetic V and Q. As shown in Figure 4.8, the torso transmission scans are displayed by three views. By adjusting the panels on the top, it is straightforward to set the locations of defects. Also, the size and activity of the synthetic lesion can be customised by dragging the scroll bars. Synthetic SPECT V and Q with defects are depicted in Figure 4.9.

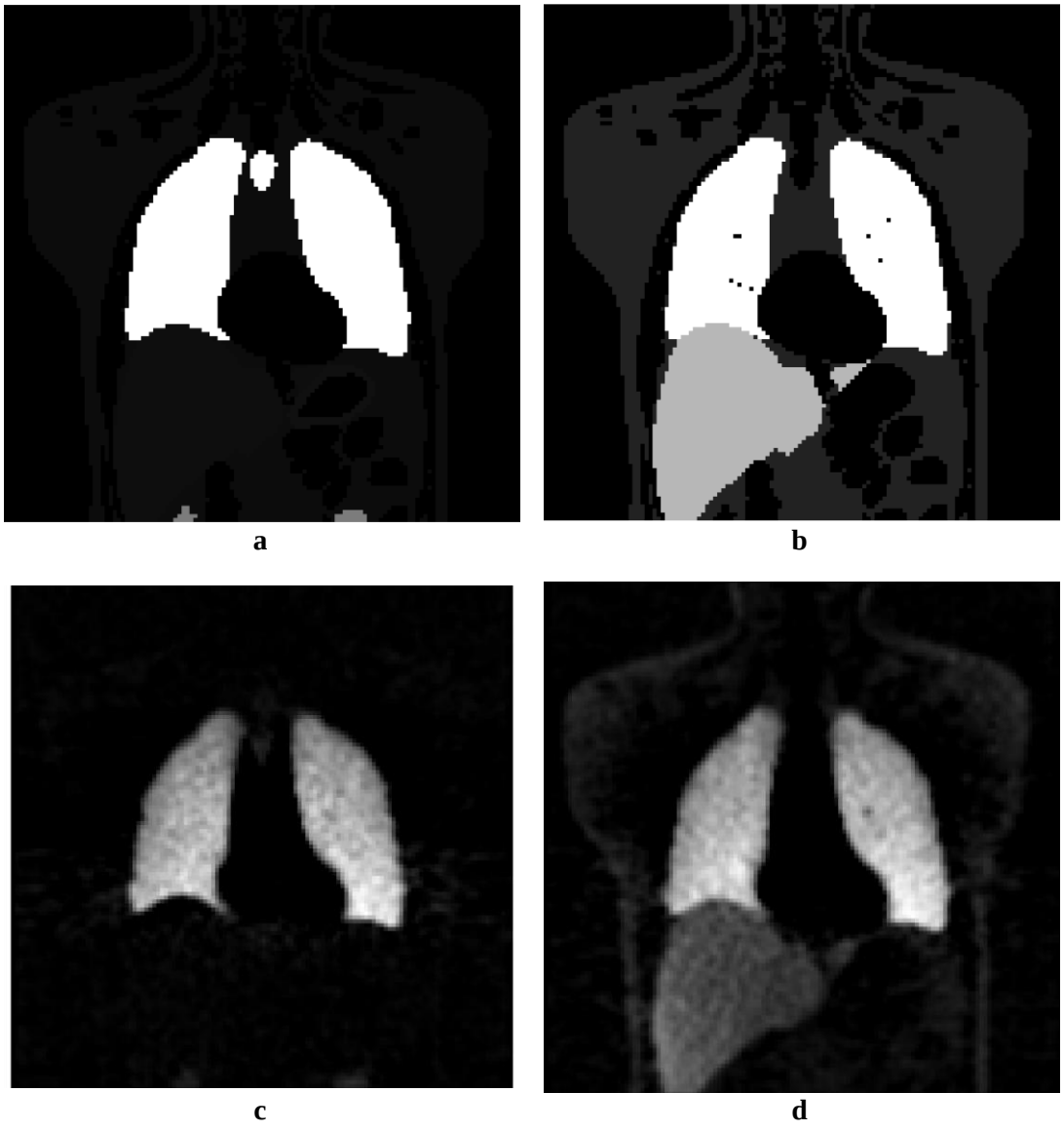
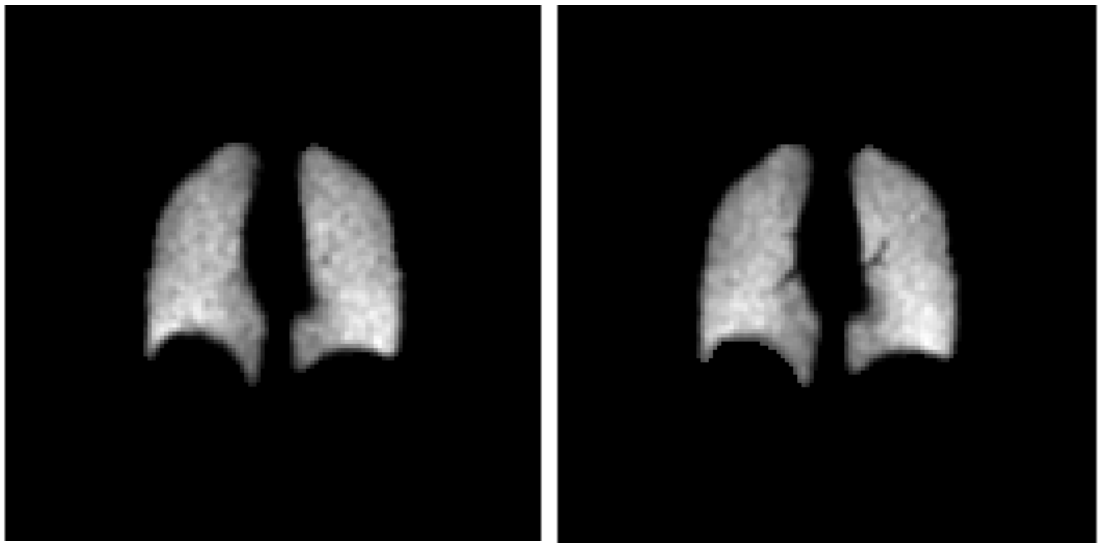
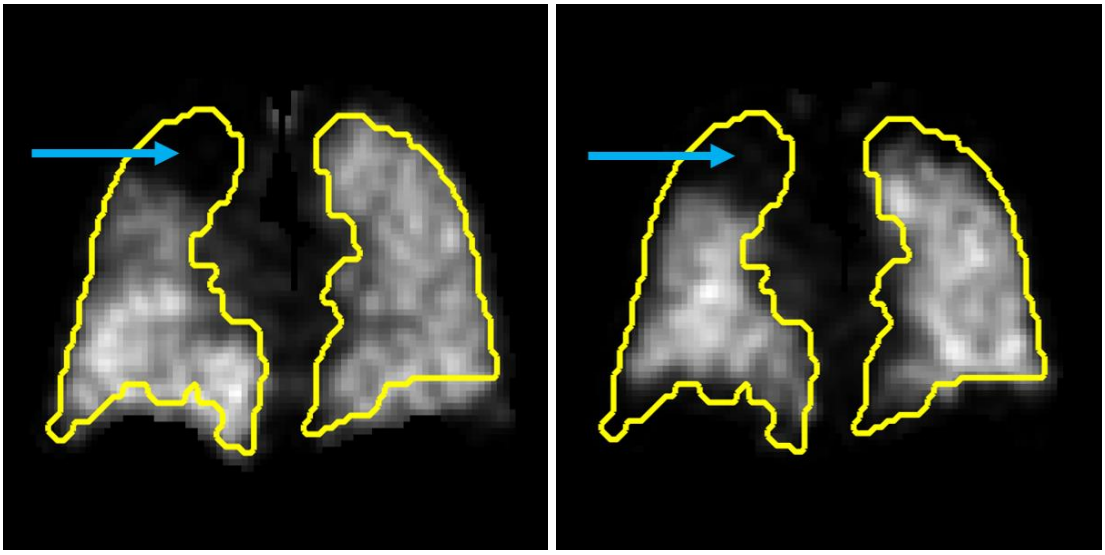


Figure 4.5 *(a) Raw activity phantom for synthetic SPECT V. (b) Raw activity phantom for synthetic SPECT Q. (c) Reconstructed SPECT V image. (d) Reconstructed SPECT Q image.*



a **b**

Figure 4.6 *Segmented results from synthetic SPECT data. (a) Segmented synthetic SPECT V. (b) Segmented synthetic SPECT Q.*



a **b**

Figure 4.7 *Abnormal SPECT V and Q images from a moderate COPD patient. The yellow contours show the boundaries of the lungs. A photopenic region (pointed out by the blue arrow) indicates the location of the lesion. (a) Defective SPECT V image. (b) Defective SPECT Q image.*

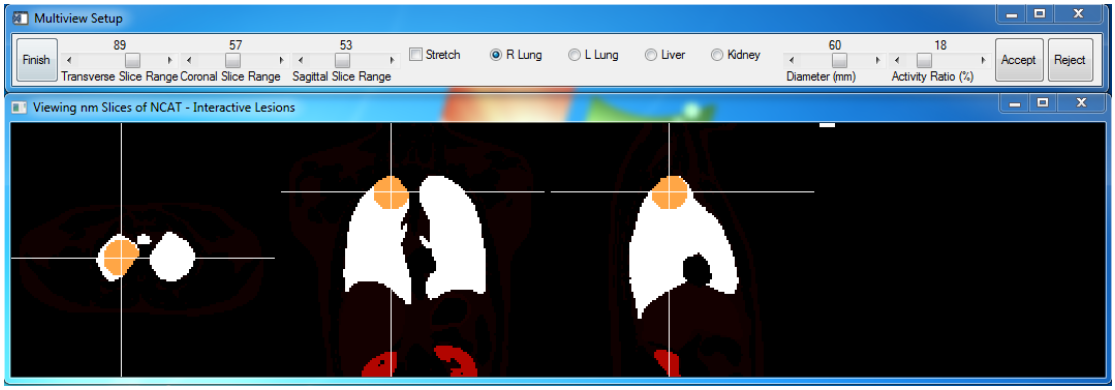
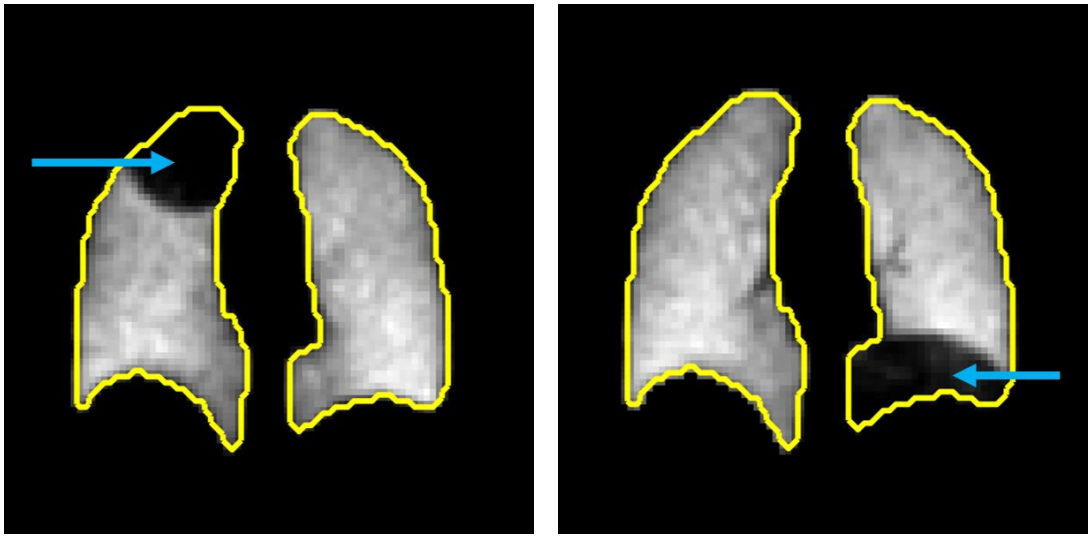


Figure 4.8 *RMDP interactive window is used to add defects to the phantoms.*



a **b**

Figure 4.9 *Abnormal synthetic SPECT V and Q images. The yellow contours show the boundaries of the lungs. A photopenic region (pointed out by the blue arrow) indicates the location of the defects. (a) Defective SPECT V image. (b) Defective SPECT Q image.*

4.4 Conclusions

In order to enrich the experimental data and validation methods, XCAT and RMDP are employed here to synthesise the lung MSCT and SPECT scans and acquire the ground-truth motion vector. By dividing one breathing cycle into eight frames, the intra-patient respiratory pattern is represented by eight phantoms using XCAT. RMDP manages to configure the parameter file, as well as convert the emission phantoms to synthetic data with 'tumour' information at a customised location. As XCAT data were used in

previous articles in aid of research into medical image processing algorithms, it is reasonable to assume that the obtained data can accommodate the needs of validation for the registration algorithms in this thesis.

Chapter 5

Multi-modality Image Fusion Using Prior Shapes

5.1 Introduction

As multi-modality lung imaging is able to provide complementary physiological and pathophysiological information, the alignment of modalities plays an important role in clinical applications [78]. In order to develop the performance of multi-modality registration, recent research articles have concentrated on the analysis of structural information and similarity metrics. A structural representation is developed in [79] using a modified accelerated segment test algorithm. The improved structural representation is considered as another channel in the cost function. Similarly, the method proposed in [80] is based on the spiking cortical model and SSD is used to evaluate the similarity of the structural descriptors extracted from multi-modality images. A traditional deformation estimation method (e.g. free-form deformation) in conjunction with a global optimisation method (e.g. MI) is argued for in [81]. Since the global optimisation fails to precisely guide the local deformation, the hierarchical solution is employed in [81] together with a feature-learning method, which maps the feature representations from the original feature space to a common space. In terms of the similarity metrics, Log-Euclidean and self-similarity metrics are proposed in [82] and [83] respectively. The Log-Euclidean metric is an inversion invariant and similarity invariant metric, and is incorporated into a Gaussian-like penalty function. The self-similarity metric is also based on structural information extracted from the images to be aligned.

MI has been demonstrated to be the best criterion for multi-modality image registration [84]. It was firstly proposed in [85], and then developed with a fluid model [86]. MI is widely used in medical image registration [14] [87] [88]. However, the disadvantages of MI are reported in [51], which argues that MI is dependent on the overlap of the two images and then proposes a modified metric, known as normalised mutual information. Furthermore, weighted mutual information, created with normalised pointwise mutual information and prior knowledge, is investigated in [89] to improve registration accuracy and stability. The latest MI-based multi-modality research can be found in [90] and [91]. In [90], it is argued that normal MI is insensitive to the image with local variations, and the useful spatial information is rarely captured to aid registration. A Harris corner detection-inspired algorithm is therefore studied to analyse the contributions made by each voxel's intensity to the joint probability density function. It is proven that with the method proposed in [90] the alignment performance is boosted, and it can be applied to image-guided surgery. Similarly, the structural features and spatial neighbourhood information are incorporated into feature neighbourhood mutual information to perform accurate registration for clinical diagnosis [91].

In recent publications, very few algorithms concentrate on pulmonary multi-modality image registration, or intend to enhance multi-modality registration using prior knowledge. In this chapter, the parameter-reduced SDM framework reported in Chapter 3 is extended for multi-modality registration. Instead of using SSD for the evaluation of single-modality alignment, MI as a similarity metric, in conjunction with gradient descent optimization, constitutes the framework for multi-modality registration. The experiments are performed using the actual lung database and synthetic phantoms, and the results are evaluated with various metrics.

The contributions of this chapter are as follows: Firstly, a multi-channel registration method is proposed to improve the performance of lung anatomical and functional imaging alignment. It conducts a direct MSCT/SPECT registration, which is independent of any aided features and landmarks. Secondly, the cost function for multi-modality registration is upgraded based on the framework stated in Chapter 3. The parameter-reduced model for multi-modality fusion avoids unnecessary analysis for a weighting parameter and properly regularises the displacement based on the SDM. Thirdly, the synthetic imaging data and the ground-truth VD offered by the 4D extended cardiac-torso phantom together with RMDP are employed to validate the registration methods.

This chapter is structured as follows: The multi-channel registration algorithm is mathematically derived in Section 5.2. Evaluation methods are introduced in Section 5.3, followed by the implementation issues noted in Section 5.4. The experimental results are compared with other state-of-the-art algorithms in Section 5.5 with in-depth discussion. Conclusions are finally written in Section 5.6.

5.2 Methodology

5.2.1 Multi-modality Image Alignment

In this part, MI is employed to measure the similarity and to derive the solution for multi-modality image alignment. If $I_{tar}(\boldsymbol{x})$ and $I_{src}(\boldsymbol{x}+\boldsymbol{u})$ are used to represent the target and registered source images, their MI can be calculated as

$$\mathrm{MI} = En(I_{tar}) + En(I_{src}) - En(I_{tar}, I_{src}) = \iint p(a,b)\log\frac{p(a,b)}{p(a)p(b)}\,\mathrm{d}a\,\mathrm{d}b \quad (5.1)$$

where a and b are the greyscale values of $I_{tar}(\boldsymbol{x})$ and $I_{src}(\boldsymbol{x}+\boldsymbol{u})$ respectively. The inverse MI is regarded as the similarity term and is denoted as

$$E_{MI} = -\iint p(a,b)\log\frac{p(a,b)}{p(a)p(b)}\,\mathrm{d}a\,\mathrm{d}b \tag{5.2}$$

In order to investigate the optimal displacement $\boldsymbol{u}$ which minimises E_{MI}, the derivative with respect to $\boldsymbol{u}$ is calculated as

$$\frac{\partial E_{MI}}{\partial u} = -\iint \frac{\partial}{\partial u}\left(p(a,b)\log\frac{p(a,b)}{p(a)p(b)}\right)\mathrm{d}a\,\mathrm{d}b \tag{5.3}$$

After further investigation,

$$\frac{\partial E_{MI}}{\partial u} = -\iint \left(\frac{\partial p(a,b)}{\partial u}\left(1+\log\frac{p(a,b)}{p(a)p(b)}\right) - \frac{p(a,b)}{p(b)}\cdot\frac{\partial p(b)}{\partial u}\right)\mathrm{d}a\,\mathrm{d}b \tag{5.4}$$

Since

$$\iint \left(\frac{p(a,b)}{p(b)}\cdot\frac{\partial p(b)}{\partial u}\right)\mathrm{d}a\,\mathrm{d}b = \int \frac{\int p(a,b)\mathrm{d}a}{p(b)}\cdot\frac{\partial p(b)}{\partial u}\mathrm{d}b = \frac{\partial \int p(b)\mathrm{d}b}{\partial u} = 0 \tag{5.5}$$

Eq. (5.4) can be simplified as,

$$\frac{\partial E_{MI}}{\partial \boldsymbol{u}} = -\iint \frac{\partial p(a,b)}{\partial \boldsymbol{u}}\cdot\left(1+\log\frac{p(a,b)}{p(a)p(b)}\right)\mathrm{d}a\,\mathrm{d}b \tag{5.6}$$

where the joint probability distribution $p(a,b)$ can be described by a function of $I_{tar}(\boldsymbol{x})$ and $I_{src}(\boldsymbol{x}+\boldsymbol{u})$ using kernel density estimation (KDE) in Eq. (5.7). Compared with histogram estimation, KDE addresses the concern that the estimation is significantly influenced by the bin size.

$$p(a,b) = \frac{1}{V\sigma_{tar}\sigma_{src}}\int K\left(\frac{a-I_{tar}(\boldsymbol{x})}{\sigma_{tar}},\frac{b-I_{src}(\boldsymbol{x}+\boldsymbol{u})}{\sigma_{src}}\right)\mathrm{d}\boldsymbol{x} \tag{5.7}$$

The coefficient V is the volume of the image (i.e. the number of total voxels in I_{tar} or I_{src}). σ_{tar} and σ_{src} are the kernel widths of I_{tar} and I_{src} respectively. Their optimal values can be computed through the modified rule-of-thumb estimator [92], which is described as

$$\sigma = 1.06\times\min\left(\hat{\sigma}(I),\frac{IQR(\mathrm{I})}{1.34}\right)\times V^{-0.2} \tag{5.8}$$

where $\hat{\sigma}(\cdot)$ is the estimated standard deviation and $IQR(\cdot)$ is the interquartile range.

Here a Gaussian kernel is chosen to estimate the distribution and the kernel function $K(\cdot)$ is denoted as

$$K(A,B) = \frac{1}{2\pi}\exp\left(-\frac{1}{2}(A^2+B^2)\right) \tag{5.9}$$

The derivative of p with respect to $\boldsymbol{u}$ is then obtained as

$$\frac{\partial p}{\partial \boldsymbol{u}} = C \cdot \int K\left(\frac{a-I_{tar}(\boldsymbol{x})}{\sigma_{tar}}, \frac{b-I_{src}(\boldsymbol{x}+\boldsymbol{u})}{\sigma_{src}}\right) \cdot \left(-\frac{b-I_{src}(\boldsymbol{x}+\boldsymbol{u})}{{\sigma_{src}}^2}\right) \cdot \left(-\frac{\partial I_{src}(\boldsymbol{x}+\boldsymbol{u})}{\partial \boldsymbol{u}}\right) \mathrm{d}\boldsymbol{x} \tag{5.10}$$

$$C = \frac{1}{V\sigma_{tar}\sigma_{src}} \tag{5.11}$$

By inserting Eq. (5.10) into Eq. (5.6), it is straightforward to derive the following equation:

$$\frac{\partial E_{MI}}{\partial \boldsymbol{u}} = C \cdot \int \iint K\left(\frac{a-I_{tar}(\boldsymbol{x})}{\sigma_{tar}}, \frac{b-I_{src}(\boldsymbol{x}+\boldsymbol{u})}{\sigma_{src}}\right) \cdot \left(1+\log\frac{p(a,b)}{p(a)p(b)}\right) \mathrm{d}a \left(\frac{I_{src}(\boldsymbol{x}+\boldsymbol{u})-b}{{\sigma_{src}}^2}\right) \mathrm{d}b \cdot \left(\frac{\partial I_{src}(\boldsymbol{x}+\boldsymbol{u})}{\partial \boldsymbol{u}}\right) \mathrm{d}\boldsymbol{x} \tag{5.12}$$

Finally, $\boldsymbol{u}$ in Eq. (5.12) is replaced by $S(\boldsymbol{\lambda})$ which is specified in Eq. (3.15)-Eq. (3.17), and Eq. (3.24) is then upgraded for multi-modality registration as follows:

$$\frac{\partial E(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = \frac{\partial\left((1-\alpha)E_{MI}(S)+\alpha R(S)\right)}{\partial S} \cdot \frac{\partial S(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} \tag{5.13}$$

5.2.2 Multi-channel Image Alignment

As different radio-isotopes (Kr-81 for ventilation and Tc-99m for perfusion) can emit gamma-rays at different energies, the signals from each isotope are separated at the receiver. In other words, SPECT V and Q can be acquired at the same time, and the two sets of imaging data are treated as being aligned automatically. Furthermore, two channels can contribute complimentary information in the case of impaired V or Q.

In this chapter, the multi-channel method is used to conduct registration between one down-sampled MSCT scan and two SPECT images simultaneously. The multi-channel cost function, E_{VQ}, derived by multivariate MI for three objects (MSCT, V and Q) can be simplified by the addition of two independent cost functions, E_V and E_Q:

$$E_{VQ} = (1-\alpha)\left[E_V\left(S(\boldsymbol{\lambda})\right) + E_Q\left(S(\boldsymbol{\lambda})\right)\right] + \alpha R\left(S(\boldsymbol{\lambda})\right) \tag{5.14}$$

where E_V and E_Q denote two MI-based similarity terms for MSCT/V and MSCT/Q alignments respectively. The solution for the multi-channel method is therefore written in Eq. (5.15):

$$\frac{\partial E}{\partial \boldsymbol{\lambda}} = \frac{\partial\left((1-\alpha)[E_V(S)+E_Q(S)]+\alpha R(S)\right)}{\partial S} \cdot \frac{\partial S(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} \tag{5.15}$$

5.3 Evaluation of Registration Accuracy

5.3.1 Mask Overlap Metric

Based on accurate segmentation of the ROI, the mask overlap is widely used for registration evaluation. The mask of the source image, $\phi_{src}(\boldsymbol{x})$, is warped according to the deformation computed. The registration similarity can be measured by the normalised inner product (NIP) between the mask of the target image, $\phi_{tar}(\boldsymbol{x})$, and the warped mask of the source image, $\overline{\phi_{src}}(\boldsymbol{x})$. The calculation of NIP is the same as that of E_{IP} in Eq. (2.28).

5.3.2 Vector Displacement Metric

VD is used to evaluate the performance by considering the voxel-wise displacement [93]. In comparison with the mask overlap metric, it provides a more precise metric for evaluation. Given the ground-truth displacement $\boldsymbol{T}$, the mean displacement error (MDE) is calculated as

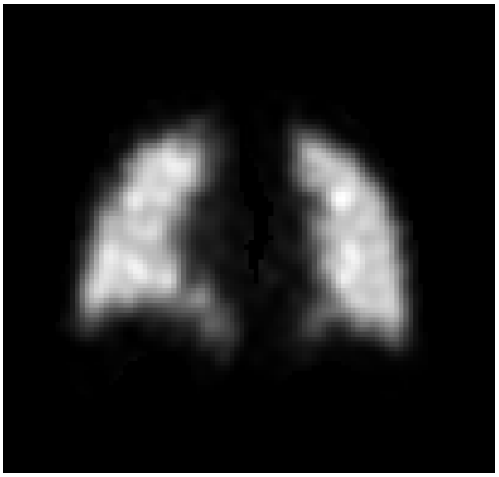
$$\mathrm{MDE} = \frac{\int \| \boldsymbol{T} - \boldsymbol{u} \|_2 \mathrm{d}\boldsymbol{x}}{V} \tag{5.16}$$

5.4 Implementation Issues

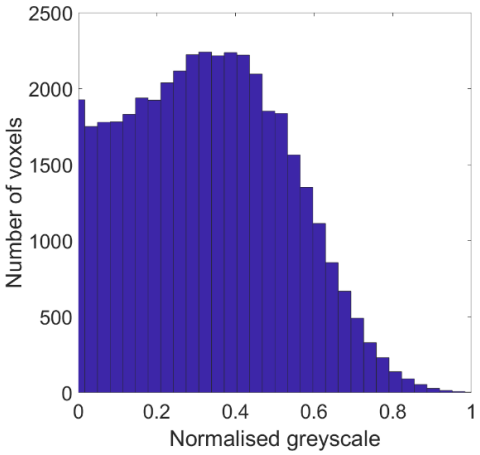
Considering the differences between MSCT and SPECT imaging, it is advisable to reduce their divergence by histogram matching. As shown in Figure 5.1, the histogram of the down-sampled MSCT scan is modified to match that of the target SPECT scan.

The actual MSCT and SPECT images are originally formatted as 16-bit unsigned integers. The maximum greyscale value of MSCT within the lung is around 324 (i.e. typical lung attenuation values are around -700 HU, and the lowest attenuation (that of air) has a value of -1024 HU) and that of SPECT is 200-300. However, employing hundreds of greyscales to calculate MI is computationally expensive and can rarely produce a tremendous improvement. Therefore, the greyscale ranges of MSCT and SPECT data are scaled to $[0,50]$ in practice. Furthermore, the synthetic MSCT represented by the normalised attenuation map also requires to be up-scaled for the computation of MI.

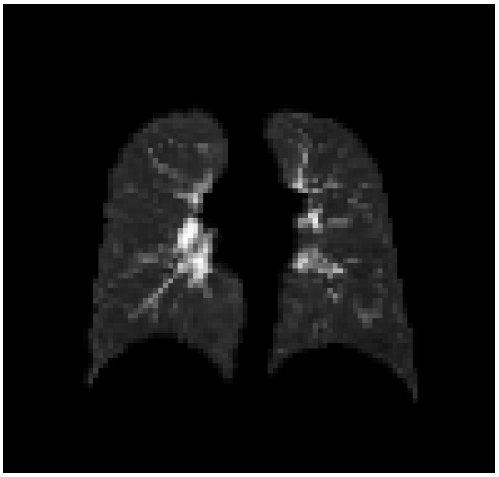
The state-of-the-art method published in [59] is referred to here and named as Ref for convenience. Since the optimal weighting parameter β varies across different scenarios, it should be investigated before starting the following experiments. Here the optimal β is retrieved using the overlap metric. In other words, β is assigned a series of possible values within a possible range to conduct CT/SPECT registration. The value corresponding to the maximum overlap at convergence is adopted for β. In addition, since the parameter α is irremovable in both Ref and the proposed methods and it varies with different cases, it is empirically assigned the same value (i.e. 0.5) for both methods in order to make the influence of α equivalent to these two methods.



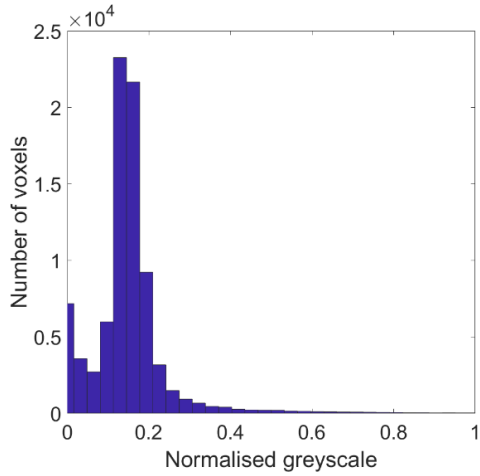
a



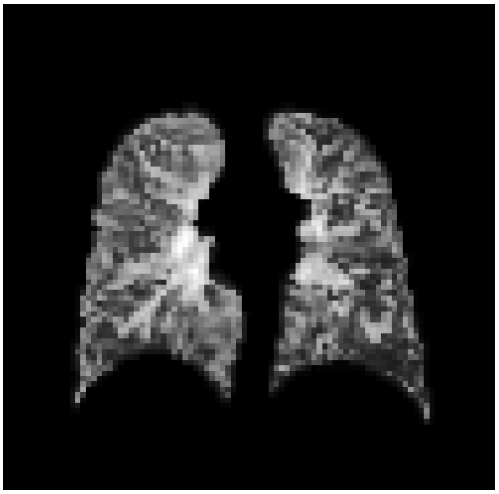
b



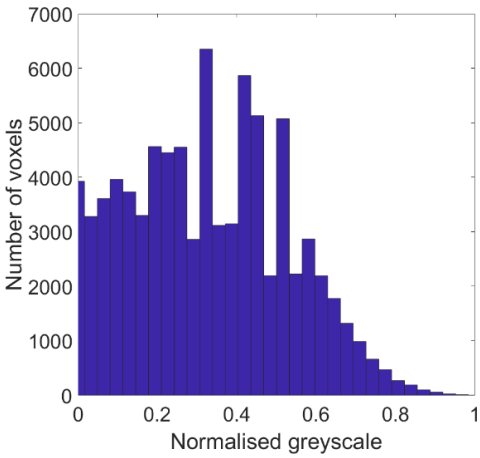
c



d



e



f

Figure 5.1 *Histogram matching before non-rigid registration. (a-b) One coronal-view slice of SPECT V imaging and the corresponding lung region histogram. (c-d) One coronal-view slice of down-sampled MSCT imaging and the corresponding lung region histogram. (e-f) One coronal-view slice of matched down-sampled MSCT imaging and the corresponding lung region histogram.*

In order to avoid local minimum issues associated with gradient descent and to accelerate convergence, the multi-resolution technique is employed here. At the coarsest-scale (lowest-resolution), with irrelevant local minima disappearing local minimum problems can be effectively overcome. Using low-resolution prior to finer scales helps achieve the global optimum [94]. The three resolutions applied here are $32 \times 32 \times 32$, $64 \times 64 \times 64$ and $128 \times 128 \times 128$. The deformations obtained from the lower resolution were up-scaled by cubic-spline interpolation and then regarded as the initial deformations for higher-resolution computation. The number of iterations is pre-set, and the calculation is terminated when they finish.

5.5 Experimental Results and Discussions

5.5.1 Multi-channel Registration Using Medical Imaging Data

In this section, real lung imaging data are employed to validate the accuracy and stability of single-channel and multi-channel registration methods. As introduced before, each subject in the lung database is comprised of MSCT scans, a pair of SPECT (V and Q) scans and low-dose CT scans. SPECT and low-dose CT scans are acquired by a hybrid CT/SPECT scanner, and it is therefore assumed that SPECT and low-dose CT scans have been registered originally. The lung mask of the low-dose CT is segmented manually by clinical experts and regarded as the ground-truth mask for corresponding SPECT V and Q images. The ground-truth lung mask of the MSCT is segmented by the Apollo software (Vida Diagnostics Inc, Iowa, USA).

As stated in Section 3.3.1, eight clipped subjects are chosen as testing data. Moreover, the remaining three moderate COPD subjects in database (Case 1, Case 28 and Case 30) are also used for test. The remaining 21 subjects are employed to train the SDM. In the following experiments, each testing subject's V/Q is set as the target, to which the 11 down-sampled MSCT scans register individually (i.e. MSCT-V, MSCT-Q) and

simultaneously (i.e. MSCT-V&Q). Therefore, the tests are divided in two groups: intra-patient registration and inter-subject registration (namely cross-validation), as shown in Figure 5.2. The total number for V-MSCT, Q-MSCT and V&Q-MSCT alignment tests is 363.

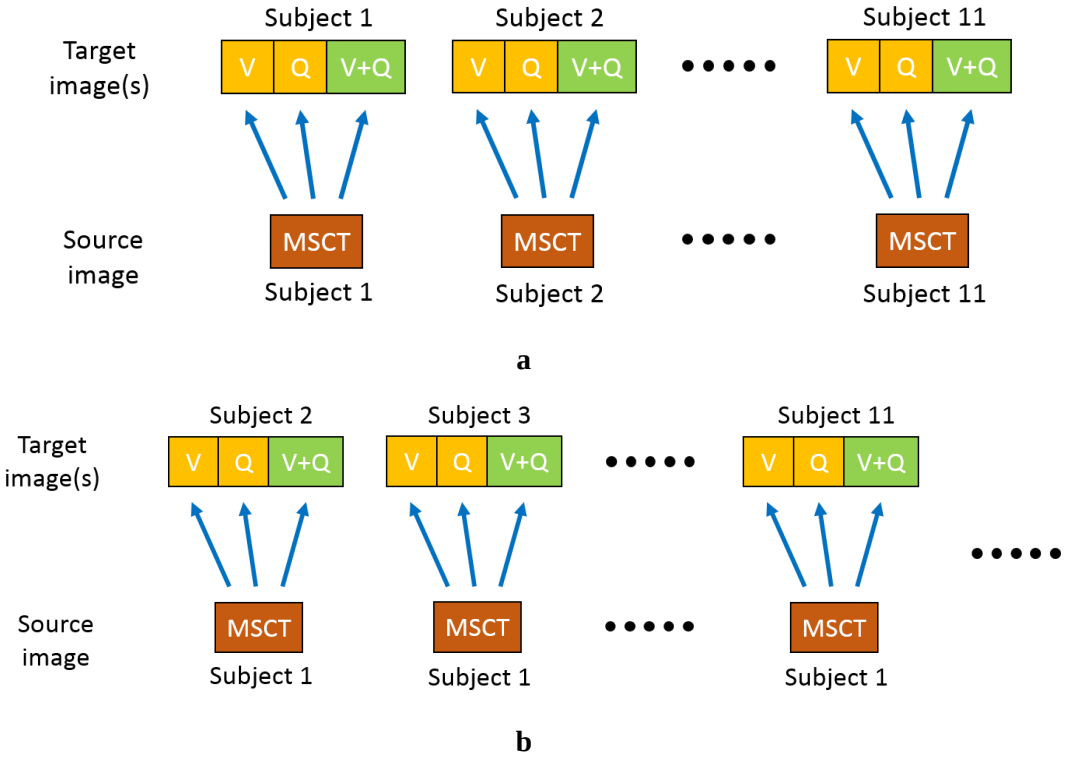


Figure 5.2 *Experimental design with medical imaging data. (a) Diagram of intra-patient registration. (b) Diagram of inter-subject registration using Subject 1 as an example.*

Apart from the Ref method and the method proposed here, NiftyReg presented in [73] is also used for comparison. In single-channel experiments, NiftyReg is properly configured to achieve the best performance. Nevertheless, since NiftyReg cannot perform multi-channel registration, it is not applied to multi-channel simulation.

The intra-patient and inter-subject registration results measured by NIP and average NIP respectively are displayed in Figure 5.3. The average NIP is calculated using the NIPs obtained from the tests with the same target image. In terms of intra-patient registration, since the SDM is employed to constrain the excessive deformation, the Ref

and proposed method apparently outperform NiftyReg in single-channel tests. Compared with the Ref method, the proposed algorithm achieves improved accuracy in both single-channel and multi-channel tests. In terms of the cross-validation, due to a relatively large difference between the target and source images, two learning-based methods' performances slightly decrease under the influence of the regularised deformation. Nevertheless, the superior average NIPs are still obtained by the method proposed here. It is demonstrated that the proposed method manages not only to constrain the excessive deformation but also to provide a desired solution to general registration problems.

For a moderate COPD patient, due to a defective region in the lung imaging, a non-learning method such as NiftyReg inevitably regarded it as background and performed excessive deformations. The intra-patient registration results (for Case 15) using three single-channel (MSCT-Q) methods are exemplified in Figure 5.4(a), (c) and (e). The Ref and the proposed methods manage to constrain the deformations and achieve relatively reasonable results. However, as argued in Section 3.2.2, the Ref method tends to penalise any displacement, and thereby sacrifices flexibility. The drawbacks are illustrated by quivers in Figure 5.4(c). Certain horizontal displacements in the middle part of the right lung can be detected, whereas lesser deformations are expected since the target and source images at this location were almost aligned in their initial poses.

The deformations are then used to map the lobes segmented from the MSCT scans onto the SPECT imaging. The fusion results are depicted in Figure 5.4(b), (d) and (f) for NiftyReg, Ref and the proposed method respectively. Through visual inspection, it is straightforward to observe the excessive deformation caused by NiftyReg and the unreasonable displacement caused by the Ref method. The advantages of the method proposed here are once again demonstrated by lobar lung fusion.

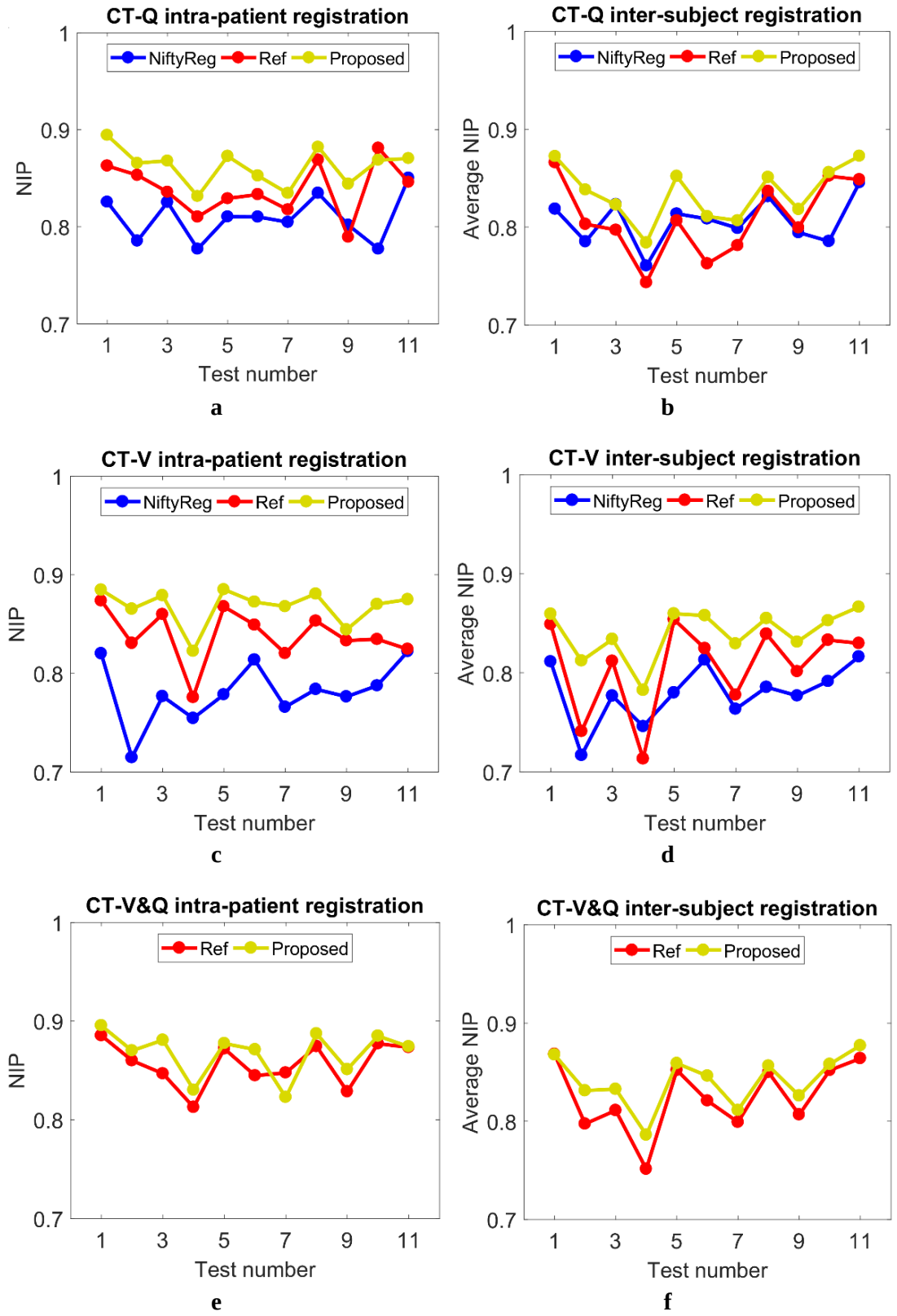


Figure 5.3 *Single-/multi-channel intra-patient/inter-subject registration results are evaluated by NIP (left column) and average NIP (right column).*

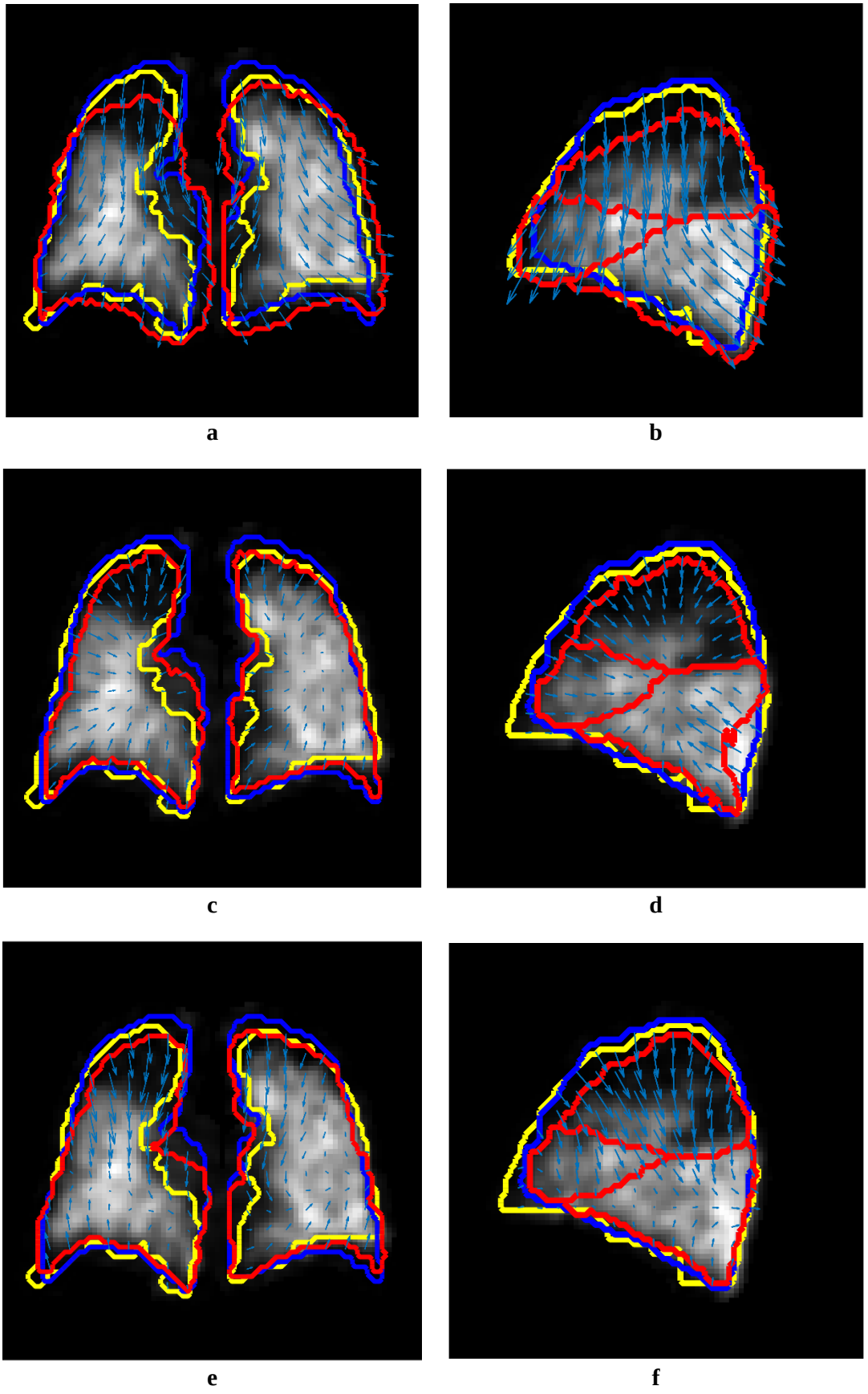


Figure 5.4 *Single-channel registration results from the NiftyReg (Row 1), Ref (Row 2) and proposed (Row 3) methods. Blue contours indicate the ground-truth boundaries of the lung region in the SPECT scan (target). (a,c,e) Coronal view: yellow and red contours represent the*

boundaries of the lung region in an MSCT scan (source) before and after deformation respectively. (b,d,f) Sagittal view: yellow and red contours represent the boundaries of the lung region before deformation and the lung lobes after deformation respectively.

The numerical results for all tests are summarized by the average NIPs and standard deviations in Table 5.1. Compared with Ref, the proposed method provides more flexibility in non-rigid registration, which contributes to the improvement of cross-validation accuracy by up to 4.1% and 1.9% for the single-channel and multi-channel techniques. Due to the more rigorous threshold to constrain the displacement, the alignment stability of the proposed multi-channel method is enhanced by up to 28%. What is more important is that the multi-channel techniques using Ref and the proposed algorithm are able to obtain the improved alignment accuracy by up to 2.2% and 1% respectively, in comparison with either V/CT or Q/CT single-channel registration.

Table 5.1 The average NIP of single-channel and multi-channel registration obtained by different methods are presented as mean ±standard deviation.

		NiftyReg	Ref	Proposed
Intra-patient	V/CT	0.809 ±0.024	0.839 ±0.027	0.862 ±0.020
	Q/CT	0.781 ±0.031	0.839 ±0.027	0.868 ±0.019
	V&Q/CT	N/A	0.857 ±0.023	0.870 ±0.023
Cross-validation	V/CT	0.806 ±0.022	0.809 ±0.050	0.835 ±0.035
	Q/CT	0.762 ±0.029	0.807 ±0.054	0.840 ±0.033
	V&Q/CT	N/A	0.825 ±0.047	0.841 ±0.034

5.5.2 Multi-modality Registration Using Synthetic Data

As stated before, eight-frame phantoms are sampled in one breathing cycle. The MSCT (source image) at the first frame is intended to register the SPECT V and Q (target

images) at each remaining frame using the multi-channel technique. In order to increase the difficulty of registration, three test modes in terms of defective SPECT scans are implemented separately. The three combinations of SPECT images are arranged as:

Mode 1: abnormal SPECT V (Figure 4.9(a)) and complete Q (Figure 4.6(b)),
Mode 2: complete SPECT V (Figure 4.6(a)) and abnormal Q (Figure 4.9(b)),
Mode 3: abnormal SPECT V (Figure 4.9(a))) and abnormal Q (Figure 4.9(b)).

Seven alignment tests are conducted for each mode, and the experimental results are evaluated with mean displacement error (MDE in Eq. 5.12) and normalised inner product (NIP). In this section, Ref and the proposed method are compared in terms of registration accuracy and stability.

The experimental results with respect to the three test modes are depicted in Figure 5.5. The average of MDE and NIP numerically demonstrate the superior registration performance achieved by the proposed method in these three test modes, as it obtains lower MDE and higher NIP. In terms of MDE, the algorithm proposed here decreases the average errors by up to 21%, 18% and 15% for modes 1, 2 and 3 respectively since the new cost function allows more flexibility in deformation. The standard deviation of the registration results demonstrates that the stability of the proposed method is improved by up to 39% and 54% (measured by NIP when the number of dominant eigenmodes is over 5) for modes 1 and 3. By considering the standard errors in Figure 5.5, it is convincing that the proposed algorithm with a certain number of dominant eigenmodes (NoE) manages to constantly achieve high accuracy for the multi-modality image alignment work.

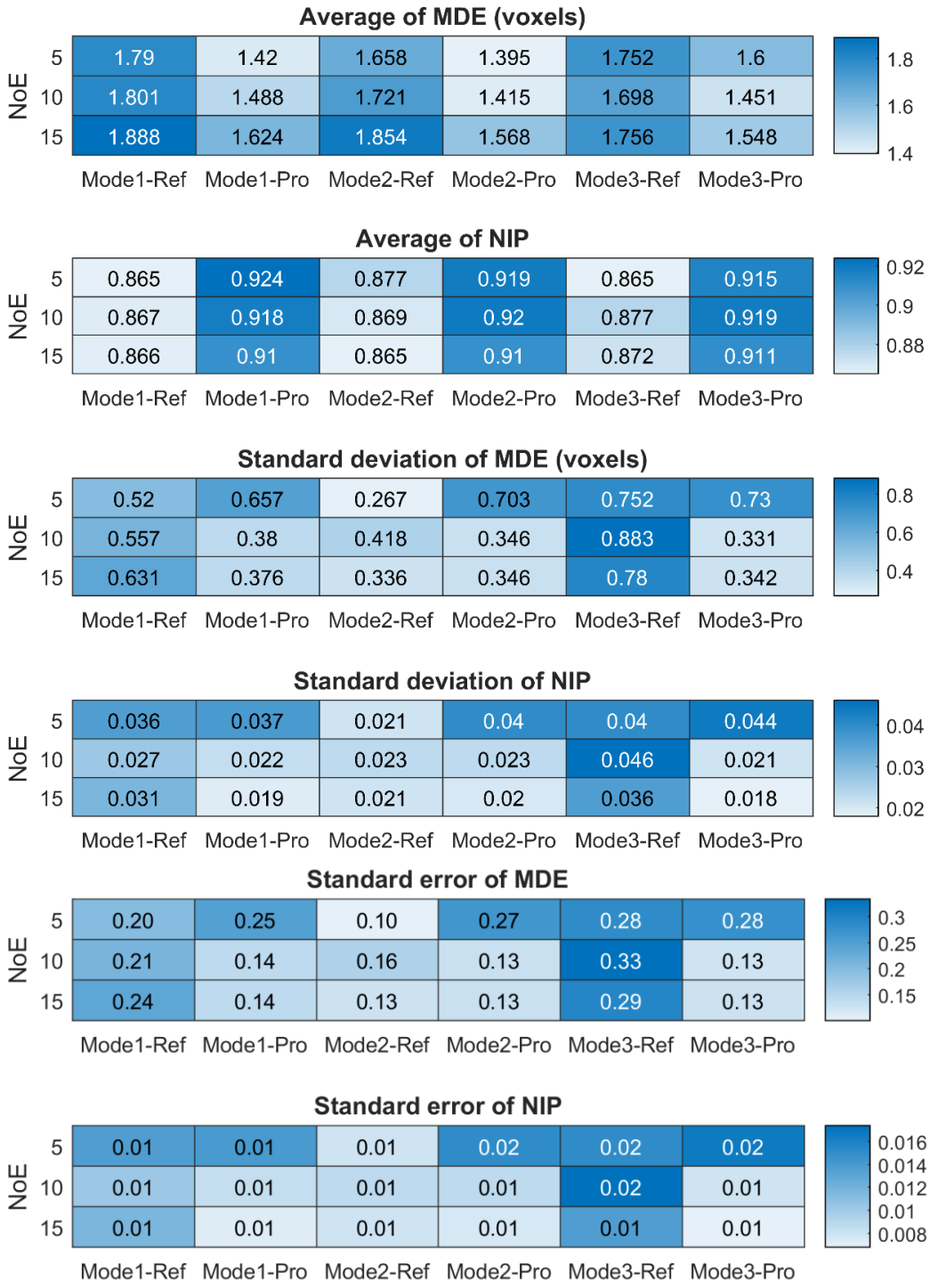


Figure 5.5 *Two algorithms with different NoEs are applied to three test modes. The registration results are evaluated with MDE and NIP.*

With the increase of NoE, the registration stability of the proposed method improves accordingly, which is demonstrated by the standard deviations of MDE and NIP. Conversely, the similar improvement achieved by the Ref method is rarely detected as

this algorithm fails to provide an explicit boundary for deformations. The experiments firmly demonstrate the flaws that the weighting parameter is weak in striking a desirable balance between flexibility and boundary-finding in previous algorithms.

In addition, experimental results with respect to seven different target patterns (sampled at frames 2 to 8 within one breathing cycle) are depicted in Figure 5.6. As illustrated in Figure 4.1, the volumes of the lungs at frames 2 and 8 are the most similar to that at frame 1, whilst the lungs have the largest deformations at frame 5. From the MDE results shown in Figure 5.6, it is straightforward to notice that the method proposed here manages to slightly enhance performance in test no. 1, 2, 6 and 7, whereas it greatly outperforms the Ref method in test no. 4 and 5 (NoE=15 and 10). As clinical SPECT imaging is acquired by tidal breathing, the actual MSCT/SPECT registration is more likely to resemble test no. 4 or 5. Therefore, high performance of the proposed method on large-deformation alignment indicates desirable results in practical application. Also, the NIP results demonstrate that the method proposed here is superior in all registration cases.

Nevertheless, by comparing the average of MDE and NIP along each column of Figure 5.5, it can be noted that the registration performance of the proposed framework slightly decreases with the increment of NoE. As the SDM is trained from the lung shapes extracted from different patients at the state of maximum inspiration, the shortcomings of this inter-subject model are manifested in intra-patient registration. In other words, in the case where five dominant eigenmodes are adequate to globally constrain the deformations, the extra eigenmodes regularised by the inter-subject model may decrease the intra-patient registration accuracy. Nevertheless, comparing the three sub-figures in the left column of Figure 5.6, it can be observed that the performance gap between the two compared methods in test no. 4 and 5 is enlarged with more

eigenmodes employed. Thus, the influence of the inter-subject SDM can be further analysed. The MDE results in Figure 5.6 demonstrate that the negative influence on the algorithm proposed here is more prominent for small-deformation alignment, and considerably reduced for large-deformation registration. It is therefore convincing that, in actual MSCT/SPECT alignment, the registration accuracy will be barely deteriorated with more eigenmodes considered in the SDM, while a greater enhancement to registration stability is obtained.

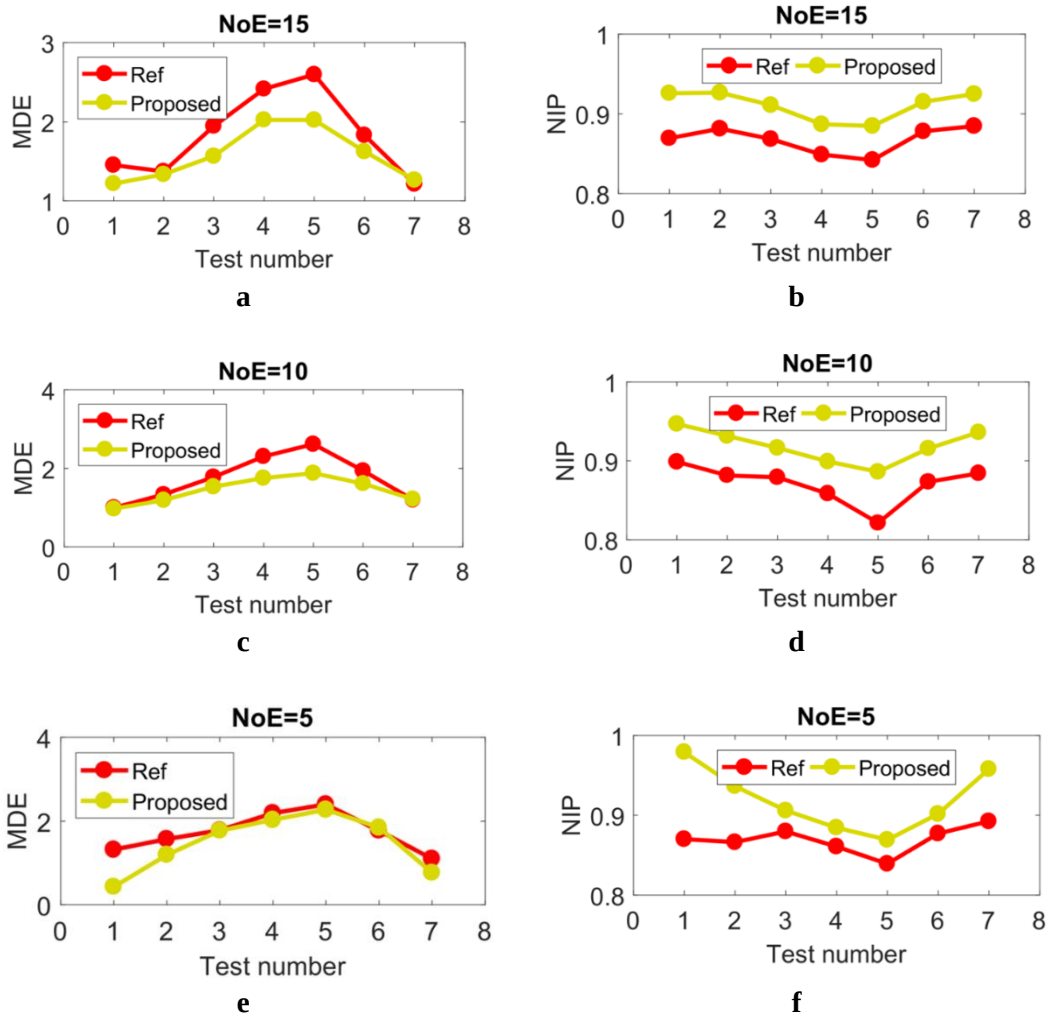


Figure 5.6 *(a,c,e) Registration results evaluated with MDE for different NoEs. (b,d,f) Registration results evaluated with NIP for different NoEs. The results are the average of three test modes.*

5.6 Conclusions

In this chapter, an innovative non-rigid registration method has been proposed for lung MSCT/SPECT image alignment. The novel technique concentrates on the lung region and conducts direct registration through two MSCT/SPECT channels simultaneously. Furthermore, the cost function based on parameter-reduced SDM is extended with MI to address the concerns of multi-modality registration. The proposed framework is validated to be capable of preventing excessive deformation for the registration involved in the images with defects.

The method proposed here and two widely-used non-rigid registration methods are tested using the lung database and synthetic phantoms. The numerical results evaluated by NIP and MDE demonstrate that the multi-channel method is able to improve registration accuracy and decrease variability in comparison with the single-channel framework, as it can take more relevant information into consideration. The proposed method with inter-subject prior knowledge outperforms other similar learning-based and non-learning-based methods investigated here, in terms of inter-subject multi-modality registration, and it is also demonstrated to achieve tremendous improvement with respect to registration accuracy and stability for intra-patient alignment, even in the case where the lung is scanned at a different state of inspiration. The desirable fusion of lung lobes onto SPECT imaging is validated to be achievable by accurate MSCT/SPECT alignment.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

This thesis presents the research work on the topic of multi-modality image fusion for the diagnosis and treatment of COPD. In order to provide a guiding tool for LVRS, a novel multi-modality fusion method was studied and improved. At the beginning, similarity registration was developed to pre-process the prior lung shapes, which are then employed to train the SDM. A learning-based non-rigid registration method was improved using a parameter-reduced function to avoid the adjustment of the weighting parameter for the statistical term. The multi-channel registration method is proposed to conduct direct MSCT/SPECT registration. The deformations are finally employed to map the lobes onto the functional imaging for image fusion.

Compared with non-iterative methods, the iterative similarity registration algorithm proposed in Chapter 2 achieves higher alignment accuracy and stability within a reasonable time. Furthermore, the proposed method is independent of structural correspondence but concentrates on the shape region (i.e. representing shapes using CFs) to save the time and labour of feature extraction. The method proposed here strikes a balance between performance and processing time, which significantly improves efficiency and the scope of application. The proposed robust and high-performance method manages not only to pre-process the prior lung shapes for subsequent use by the SDM, but also to solve universal shape registration problems, even in the presence of noise.

A non-rigid registration technique was studied to achieve high alignment accuracy. In terms of medical image registration, high similarity measured by SSD or MI is not a gold-standard due to underlying excessive deformations. In order to maintain structural information through non-linear transformation, numerous learning-based methods were proposed to address the concerns. Nevertheless, the introduction of the penalised term inevitably leads to an extra-investigation of the corresponding weighting parameter. This shortcoming is overcome thoroughly by means of the parameter-reduced alignment technique proposed in Chapter 3. It was demonstrated that the parameter-reduced SDM-based method can prevent excessive deformation and achieve superior performance.

In order to enrich the experimental data and provide the ground-truth VD of each voxel for the evaluation of registration, XCAT was employed to generate torso phantoms and the displacement of each voxel. In fact, NCAT was considered initially as it is licence-free. However, the VD created by NCAT is incorrect and relevant technical support is no longer available. XCAT was therefore selected to accomplish the synthetic data collection in Chapter 4. The parameter-reduced SDM technique was extended to tackle the multi-modality registration problem based on MI in Chapter 5. The multi-channel method proposed here incorporates complimentary SPECT V/Q imaging information into the cost function to effectively improve the registration accuracy and stability.

The intra-patient registration investigated in this thesis is expected to be regularised by intra-patient prior knowledge. However, due to the limited number of prior shapes, the inter-subject lung shapes were used to train the SDM. The lung shapes were transformed using similarity registration to eliminate the influence of scale and initial orientation. This PCA-based learning procedure is based on the assumption that the variations of lungs are normally distributed. The fusion of lobes onto SPECT images

demonstrated the reliability and validity of the proposed non-rigid registration method using inter-subject knowledge. In summary, this thesis proposes a novel technique to significantly improve the performance of the learning-based registration framework, which is exemplified by the SDM here. Also, multi-modality image fusion lays a solid foundation for lobar lung activity analysis, which can support the formulation of a therapeutic schedule for the COPD patients.

6.2 Future Work

This thesis proposes an algorithm for accurate multi-modality image registration. Considering the clinical application, the limitation of prior lung data for the SDM and the limited research into functional imaging analysis, it is convincing to say that the following three topics will be promising for future investigations.

In terms of clinical application, lung activity analysis is of vital importance, as it can be used to guide LVRS in detail. The image fusion achieved by the method proposed in this thesis is able to map lobe information (e.g. boundaries and fissures) onto SPECT imaging. Automatic classification is expected to be conducted on the SPECT imaging in each lung lobe to identify any lobes that need to be removed. Classification-based diagnosis has become a popular technique in medical image analysis and saves considerable time and efforts from radiologists and surgeons.

As discussed in Section 5.5.2, the SDM trained with inter-subject data cannot perfectly regularise the intra-patient deformation. On one hand, the limited number of patients negatively influences the generalisation of the diversity of lung shapes. On the other hand, each subject only has one frame data within one breathing cycle (i.e. the frame at maximum inspiration), which scarcely contributes to the statistical model of intra-patient deformations. In order to improve the performance of prior knowledge-based

methods in the future, a respiratory model for intra-patient registration estimated by one frame MSCT scan will need to be investigated.

In addition, as feature-based multi-modality registration methods have received lots of attention recently, it is worthwhile to conduct research into the structural information provided by functional imaging. In fact, none of the publications to date have investigated the structural information probably extracted from SPECT imaging as functional imaging is always low-resolution and obtained during tidal breathing. However, there is the potential that structural information such as key points, texture and intensity features may be obtained by a new functional imaging technique or extracted by innovative algorithms.

Appendix A

Case Index and Demographic Data

Table A-1: Case index of the low-dose CT masks used in Chapter 2.

Case Index	Subject ID	Patient State
Case 1	BRUCI 003	Emphysema 2
Case 2	BRUCI 007	Healthy Smoker
Case 3	BRUCI 008	Healthy Smoker
Case 4	BRUCI 009	Healthy
Case 5	BRUCI 014	Healthy
Case 6	BRUCI 015	Healthy
Case 7	BRUCI 016	Healthy Smoker
Case 8	BRUCI 018	Healthy
Case 9	BRUCI 020	Healthy
Case 10	BRUCI 021	Healthy Smoker
Case 11	BRUCI 022	Emphysema 1
Case 12	BRUCI 024	Healthy Smoker
Case 13	BRUCI 027	Healthy Smoker
Case 14	BRUCI 028	Emphysema 1
Case 15	BRUCI 032	Healthy
Case 16	BRUCI 033	Emphysema 2
Case 17	BRUCI 036	Emphysema 2
Case 18	BRUCI 037	Emphysema 1
Case 19	BRUCI 038	Healthy Smoker
Case 20	BRUCI 040	Healthy
Case 21	BRUCI 041	Healthy
Case 22	BRUCI 042	Healthy
Case 23	BRUCI 043	Healthy
Case 24	BRUCI 045	Emphysema 1
Case 25	BRUCI 049	Healthy Smoker
Case 26	BRUCI 051	Emphysema 1
Case 27	BRUCI 062	Emphysema 1
Case 28	BRUCI 072	Emphysema 1
Case 29	BRUCI 076	Emphysema 1
Case 30	BRUCI 081	Emphysema 2
Case 31	BRUCI 085	Emphysema 2
Case 32	BRUCI 089	Emphysema 1
Case 33	BRUCI 090	Healthy Smoker

Healthy: healthy people without smoking habit.

Healthy Smoker: healthy people with smoking habit.

Emphysema 1: unhealthy people with mild COPD.

Emphysema 2: unhealthy people with moderate COPD.

Table A-2: Demographic data for the sample population in Table A-1. FEV_{1pp}, FEV_1/FVC and DLCO are presented as mean ± standard deviation.

Variables	COPD (number=14)	Healthy non-smokers (number =9)	Healthy smokers (number =10)
Gender (number)	F(5) M(9)	F(3) M(6)	F(6) M(4)
Age (year)*	62(15)	47(29)	52(24)
FEV_{1pp} (%)	83.23 ± 14.15	107.61 ± 13.33	107.19 ± 15.14
FEV_1/FVC(%)	60.13 ± 7.44	80.93 ± 3.63	76.26 ± 3.59
DLCO (ml/min/mmHg)	85.79 ± 13.48	96.96 ± 14.12	89.70 ± 13.55

F=Female, M=Male.
FEV1=Forced Expiratory Volume in the first second.
FEV1pp=actual_FEV_1/average_FEV_1.
FVC= Forced Vital Capacity.
DLCO= Diffusing capacity of the Lung for carbon monoxide (CO).

* Values are not normally distributed and are shown as median (range)

Table A-3: Case index of the MSCT/low-dose CT/SPECT images and masks used in Chapter 3 and 5.

Case Index	Subject ID	Patient State
Case 1	BRUCI 003	Emphysema 2
Case 2	BRUCI 007	Healthy Smoker
Case 3	BRUCI 008	Healthy Smoker
Case 4	BRUCI 009	Healthy
Case 5	BRUCI 014	Healthy
Case 6	BRUCI 015	Healthy
Case 7	BRUCI 016	Healthy Smoker
Case 8	BRUCI 018	Healthy Smoker
Case 9	BRUCI 021	Healthy Smoker
Case 10	BRUCI 022	Emphysema 1
Case 11	BRUCI 024	Healthy Smoker
Case 12	BRUCI 027	Healthy Smoker
Case 13	BRUCI 028	Emphysema 1
Case 14	BRUCI 032	Healthy
Case 15	BRUCI 036	Emphysema 2
Case 16	BRUCI 037	Emphysema 1
Case 17	BRUCI 038	Healthy Smoker
Case 18	BRUCI 040	Healthy
Case 19	BRUCI 041	Healthy
Case 20	BRUCI 042	Healthy
Case 21	BRUCI 043	Healthy
Case 22	BRUCI 044	Healthy
Case 23	BRUCI 045	Emphysema 1
Case 24	BRUCI 049	Healthy Smoker
Case 25	BRUCI 051	Emphysema 1
Case 26	BRUCI 062	Emphysema 1
Case 27	BRUCI 076	Emphysema 1
Case 28	BRUCI 081	Emphysema 2
Case 29	BRUCI 084	Emphysema 1
Case 30	BRUCI 085	Emphysema 2
Case 31	BRUCI 089	Emphysema 1
Case 32	BRUCI 090	Healthy Smoker

Healthy: healthy people without smoking habit.

Healthy Smoker: healthy people with smoking habit.

Emphysema 1: unhealthy people with mild COPD.

Emphysema 2: unhealthy people with moderate COPD.

Lung shapes from the highlighted cases were clipped (incomplete) during raw data acquisition.

Table A-4: Demographic data for the sample population in Table A-3. FEV_{1pp}, FEV_1/FVC and DLCO are presented as mean ± standard deviation.

Variables	COPD (number=13)	Healthy non-smokers (number =9)	Healthy smokers (number =10)
Gender (number)	F(4) M(9)	F(2) M(7)	F(6) M(4)
Age (year)*	61(19)	47(29)	52(24)
FEV_{1pp} (%)	85.08 ±14.55	106.71 ±12.22	107.19 ±15.14
FEV_1/FVC(%)	60.69 ±7.46	81.81 ±3.93	76.26 ±3.59
DLCO (ml/min/mmHg)	85.38 ±13.68	79.86 ±11.63	89.70 ±13.55

F=Female, M=Male.
FEV1=Forced Expiratory Volume in the first second.
FEV1pp=actual_FEV_1/average_FEV_1.
FVC= Forced Vital Capacity.
DLCO= Diffusing capacity of the Lung for carbon monoxide (CO).

* Values are not normally distributed and are shown as median (range)

References

[1] D. M. Mannino, "Epidemiology, Prevalence, Morbidity and Mortality, and Disease Heterogeneity," *Chest Journal,* vol. 121, no. 5, pp. 121S-126S, 2002.

[2] M. K. Han, A. Agusti, P. M. Calverley, B. R. Celli, G. Criner, J. L. Curtis, L. M. Fabbri, J. G. Goldin, P. W. Jones, W. MacNee, B. J. Make, K. F. Rabe, S. I. Rennard, F. C. Sciurba, E. K. Silverman, J. Vestbo, G. R. Washko, E. F. Wouters and F. J. Martinez, "Chronic Obstructive Pulmonary Disease Phenotypes: The Future of COPD," *American Journal of Respiratory and Critical Care Medicine,* vol. 182, no. 5, pp. 598-604, 2010.

[3] K. R. Flaherty, E. A. Kazerooni, J. L. Curtis, M. Iannettoni, L. Lange, M. A. Schork and F. J. Martinez, "Short-term and Long-term Outcomes After Bilateral Lung Volume Reduction Surgery: Prediction by Quantitative CT," *Chest,* vol. 119, no. 5, pp. 1337-1346, 2001.

[4] A. J. Patton and G. T. Turkington, "SPECT/CT Physical Principles and Attenuation Correction," *Journal of Nuclear Medicine Techonology*, vol. 36, no. 1, pp. 1-10, March 2008.

[5] B. Eskofier, Non-Rigid 3D SPECT/CT Image Registration, VDM Verlag, 2008.

[6] T. A. Holly, B. G. Abbott, M. Al-Mallah, D. A. Calnon, M. C. Cohen, F. P. DiFilippo, E. P. Ficaro, M. R. Freeman, R. C. Hendel, D. Jain, S. M. Leonard, K. J. Nichols, D. M. Polk and P. Soman, "Single photon-emission computed tomography," *Journal of Nuclear Cardiology*, vol. 17, no. 5, pp. 941-973, 2010.

[7] G. G. King, B. Harris and S. Mahadev, “V/Q SPECT: Utility for Investigation of Pulmonary Physiology,” *Seminars in Nuclear Medicine,* vol. 40, no. 6, pp. 467-473, 2010.

[8] T. F. Lang, B. H. Hasegawa, S. C. Liew, K. J. Brown, S. C. Blankespoor, S. M. Reilly, E. L. Gingold and C. E. Cann, “Description of a Prototype Emission-Transmission Computed Tomography Imaging System,” *Journal of Nuclear Medicine,* vol. 33, no. 10, pp. 1881-1887, 1992.

[9] B. Bybel, R. C. Brunken, F. P. DiFilippo, D. R. Neumann, G. Wu and M. D. Cerqueira, “SPECT/CT Imaging: Clinical Utility of an Emerging Technology,” *Radiographics,* vol. 28, no. 4, pp. 1097-1113, 2008.

[10] F. U. Chowdhury and A. F. Scarsbrook, “The role of hybrid SPECT-CT in oncology: current and emerging clinical applications,” *Clinical Radiology,* vol. 63, no. 3, pp. 241-251, 2008.

[11] Y. Yamaguchi, Y. Takeda, T. Hara, X. Zhou, M. Matsusako, Y. Tanaka, K. Hosoya, T. Nihei, T. Katafuchi and H. Fujita, “Three modality image registration of brain SPECT/CT and MR images for quantitative analysis of dopamine transporter imaging,” in *SPIE Medical Imaging*, San Diego, 2016.

[12] G. Li, D. Citrin, K. Camphausen, B. Mueller, C. Burman, B. Mychalczak, R. W. Miller and Y. Song, “Advances in 4D Medical Imaging and 4D Radiation Therapy,” *Technology in cancer research & treatment,* vol. 7, no. 1, pp. 67-81, 2008.

[13] J. Dey, A. Konik, W. P. Segars and M. A. King, "Multi-Modal Rigid and Non-Rigid Registration for Attenuation Correction in Cardiac SPECT/CT Using Emission Scatter to CT Conversion," in *IEEE Nuclear Science Symposium and Medical Imaging Conference*, Anaheim, 2012.

[14] L. Tang, G. Hamarneh and A. Celler, "Validation of mutual information-based registration of CT and bone SPECT images in dual-isotope studies," *Computer Methods and Programs in Biomedicine*, vol. 92, no. 2, pp. 173-185, November 2008.

[15] M. Á. Martín-Fernández, R. Cárdenes, E. Muñoz-Moreno, R. de Luis-García, M. Martín-Fernández and C. Alberola-López, "Automatic articulated registration of hand radiographs," *Image and Vision Computing*, vol. 27, no. 8, pp. 1207-1222, 2009.

[16] J. W. Suh, D. Scheinost, D. P. Dione, L. W. Dobrucki, A. J. Sinusas and X. Papademetris, "A non-rigid registration method for serial lower extremity hybrid SPECT/CT imaging," *Medical Image Analysis*, vol. 15, no. 1, pp. 96-111, 2011.

[17] I. Reducindo, A. R. Mejía-Rodríguez, E. Arce-Santana, D. U. Campos-Delgado, E. Scalco, G. M. Cattaneo and G. Rizzo, "Multimodal Non-Rigid Registration Methods Based on Demons Models and Local Uncertainty Quantification Used in 3D Brain Images," in *International Symposium on Visual Computing*, Las Vegas, 2014.

[18] F. Michel, M. Bronstein, A. Bronstein and N. Paragios, "Boosted metric learning for 3D multi-modal deformable registration," in *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, Chicago, 2011.

[19] F. Zhu, M. Ding and X. Zhang, “Self-similarity inspired local descriptor for non-rigid multi-modal image registration,” *Information Sciences,* vol. 372, pp. 16-31, 2016.

[20] X. Wan, D. Yu, F. Yang, C. Yang, C. Leng, M. Xu and J. Tian, “A New Region Descriptor for Multi-Modal Medical Image Registration and Region Detection,” in *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* Milan, 2015.

[21] 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.

[22] D. Rueckert, A. F. Frangi and J. A. Schnabel, “Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration,” *IEEE Transactions on Medical Imaging,* vol. 22, no. 8, pp. 1014-1025, 2003.

[23] W.-l. Li, Z.-p. Yin, Y.-a. Huang and Y.-l. Xiong, “Automatic registration for 3D shapes using hybrid dimensionality-reduction shape descriptions,” *Pattern Recognition,* vol. 44, no. 12, pp. 2926-2943, 2011.

[24] G. Li, H. Xie, H. Ning, D. Citrin, J. Capala, R. Maass-Moreno, P. Guion, B. Arora, N. C. Coleman, K. Camphausen and R. W. Miller, “Accuracy of 3D volumetric image registration based on CT, MR and PET/CT phantom experiments,” *Journal of Applied Clinical Medical Physics,* vol. 9, no. 4, pp. 17-36, 2008.

[25] R. J. Schneider, D. P. Perrin, N. V. Vasilyev, G. R. Marx, J. Pedro and R. D. Howe, "Real-time image-based rigid registration of three-dimensional ultrasound," *Medical Image Analysis*, vol. 16, no. 2, pp. 402-414, 2012.

[26] Y. Liu, R. R. Martin, L. De Dominicis and B. Li, "Using retinex for point selection in 3D shape registration," *Pattern Recognition,* vol. 47, no. 6, pp. 2126-2142, 2014.

[27] O. J. Woodford, M.-T. Pham, A. Maki, F. Perbet and B. Stenger, "Demisting the Hough Transform for 3D Shape Recognition and Registration," *International Journal of Computer Vision,* vol. 106, no. 3, pp. 332-341, 2014.

[28] N. Paragios, M. Rousson and V. Ramesh, "Non-rigid registration using distance functions," *Computer Vision and Image Understanding,* vol. 89, no. 2-3, pp. 142-165, 2003.

[29] S. Mahmoodi, M. S. Al-Huseiny and M. S. Nixon, "Similarity Registration for Shapes Based on Signed Distance Functions," in *Lecture Notes in Computer Science of 8th International Symposium*, Rethymnon, 2012.

[30] W. Liu, S. Mahmoodi, T. Havelock and M. J. Bennett, "Robust Similarity Registration Technique for Volumetrc Shapes Represented by Characteristic Functions," *Pattern Recognition,* vol. 47, pp. 1144-1158, 2014.

[31] M. S. Al-Huseiny and S. Mahmoodi, "Shape Registration Using Characteristic Functions," *IET Image Processing,* vol. 9, no. 3, pp. 249-260, March 2015.

[32] M. Reuter, “Hierarchical Shape Segmentation and Registration via Topological Features of Laplace-Beltrami Eigenfunctions,” *International Journal of Computer Vision,* vol. 89, no. 2-3, pp. 287-308, 2010.

[33] R. Litman and A. M. Bronstein, “Learning spectral descriptors for deformable shape correspondence,” *IEEE transactions on pattern analysis and machine intelligence,* vol. 36, no. 1, pp. 171-180, 2014.

[34] J. Xie, M. Wang and Y. Fang, “Learned Binary Spectral Shape Descriptor for 3D Shape Correspondence,” in *IEEE Conference on Computer Vision and Pattern Recognition,* Las Vegas, 2016.

[35] B. J. Paul and N. D. 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.

[36] H. Pottmann, S. Leopoldseder and M. Hofer, “Registration without ICP,” *Computer Vision and Image Understanding,* vol. 95, no. 1, pp. 54-71, 2004.

[37] W. L. Li, Z. P. Yin, Y. A. Hunag and Y. L. Xiong, “Three-dimensional point-based shape registration algorithm based on adaptive distance function,” *IET Computer Vision,* vol. 5, no. 1, pp. 68-76, 2011.

[38] S. Billings and R. Taylor, “Iterative Most Likely Oriented Point Registration,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention,* Boston, 2014.

[39] D. Tihonkih, A. Makovetskii and V. Kuznetsov, “A modified iterative closest point algorithm for shape registration,” in *Applications of Digital Image Processing XXXIX*, San Diego, 2016.

[40] T. Chen, B. C. Vemuri, A. Rangarajan and S. J. Eisenschenk, “Group-wise Point-set Registration Using a Novel CDF-Based Havrda-Charvát Divergence,” *International Journal of Computer Vision*, vol. 86, no. 1, pp. 111-124, 2010.

[41] D. Breitenreicher and C. Schnörr, “Robust 3D object registration without explicit correspondence using geometric integration,” *Machine Vision and Applications*, vol. 21, no. 5, pp. 601-611, 2010.

[42] B. Gutman , Y. Wang, T. Chan, P. M. Thompson and A. W. Toga, “Shape Registration with Spherical Cross Correlation,” in *2nd MICCAI Workshop on Mathematical Foundations of Computational Anatomy*, New-York, 2008.

[43] P. J. Kostelec and D. N. Rockmore, “FFTs on the Rotation Group,” *Journal of Fourier Analysis and Applications*, vol. 14, no. 2, pp. 145-179, 2008.

[44] X. Huang, N. Paragios and D. N. Metaxas, “Shape registration in implicit spaces using information theory and free form deformations,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* , vol. 28, no. 8, pp. 1303-1318, 2006.

[45] L. Wang, X. Gao and Q. Fang, “A Novel Mutual Information-based Similarity Measure for 2D/3D Registration in Image Guided Intervention,” in *International Conference on Orange Technologies*, Tainan, 2013.

[46] M.-T. Pham, O. J. Woodford, F. Perbet, A. Maki, B. Stenger and R. Cipolla, “A new distance for scale-invariant 3D shape recognition and registration,” in *IEEE International Conference on Computer Vision*, Barcelona, 2011.

[47] P. B. Davenport, “Rotations about Nonorthogonal Axes,” *AIAA Journal,* vol. 11, no. 6, pp. 853-857, 1973.

[48] D. P. Huttenlocher, G. A. Klanderman and W. J. Rucklidge, “Comparing images using the Hausdorff distance,” *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 15, no. 9, pp. 850-863, 1993.

[49] 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.

[50] H. Wang, L. Dong, J. O'Daniel, R. Mohan, A. S. Garden, K. K. Ang, D. A. Kuban, M. Bonnen, J. Y. Chang and R. Cheung, “Validation of an accelerated'demons' algorithm for deformable image registration in radiation therapy,” *Physics in medicine and biology,* vol. 50, no. 12, pp. 2887-2905, 2005.

[51] 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.

[52] T. Vercauteren, X. Pennec, A. Perchant and N. Ayache, “Diffeomorphic demons: Efficient non-parametric image registration,” *NeuroImage,* vol. 45, no. 1, pp. S61-S72, 2009.

[53] J. Ashburner, “A fast diffeomorphic image registration algorithm,” *NeuroImage*, vol. 38, no. 1, pp. 95-113, 2007.

[54] T. B. Yeo, M. R. Sabuncu, T. Vercauteren, N. Ayache, B. Fischl and P. Golland, “Spherical Demons: Fast Diffeomorphic Landmark-Free Surface Registration,” *IEEE Transactions on Medical Imaging*, vol. 29, no. 3, pp. 650-668, 2010.

[55] H. Dang, A. S. Wang, M. S. Sussman, J. H. Siewerdsen and J. W. Stayman, “dPIRPLE: a joint estimation framework for deformable registration and penalized-likelihood CT image reconstruction using prior images,” *Physics in medicine and biology*, vol. 59, no. 17, pp. 4799-4825, 2014.

[56] D. Cremers, C. Guetter and C. Xu, “Nonparametric Priors on the Space of Joint Intensity Distributions for Non-Rigid Multi-Modal Image Registration,” in *IEEE Conference on Computer Vision and Pattern Recognition*, New York, 2006.

[57] R. W. So and A. C. Chung, “Multi-modal non-rigid image registration based on similarity and dissimilarity with the prior joint intensity distributions,” in *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, Rotterdam, 2010.

[58] C. Tanner, J. H. Hipwell and D. J. Hawkes, “Using statistical deformation models for the registration of multimodal breast images,” in *Proc. SPIE 7259, Medical Imaging 2009: Image Processing*, 2009.

[59] S. B. Prabu, R. Toth and A. Madabhushi, “A statistical deformation model (SDM) based regularizer for non-rigid image registration: application to registration of

multimodal prostate MRI and histology," in *SPIE Medical Imaging*, Florida, 2013.

[60] L. M. Ellingsen, G. Chintalapani, R. H. Taylor and J. L. Prince, "Robust deformable image registration using prior shape information for atlas to patient registration," *Computerized Medical Imaging and Graphics,* vol. 34, no. 1, pp. 79-90, 2010.

[61] J. A. Onofrey, X. Papademetris and L. H. Staib, "Low-Dimensional Non-Rigid Image Registration Using Statistical Deformation Models From Semi-Supervised Training Data," *IEEE Transactions on Medical Imaging,* vol. 34, no. 7, pp. 1522-1532, 2015.

[62] M. E. Leventon, W. E. L. Grimson and O. Faugeras, "Statistical Shape Influence in Geodesic Active Contours," *IEEE Conference on Computer Vision and Pattern Recognition,* vol. 1, pp. 316-323, 2000.

[63] T. Albrecht, M. Lüthi and T. Vetter, "A statistical deformation prior for non-rigid image and shape registration," in *IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage, 2008.

[64] Z. Xue, D. Shen and C. Davatzikos, "Statistical representation of high-dimensional deformation fields with application to statistically constrained 3D warping," *Medical Image Analysis,* vol. 10, no. 5, pp. 740-751, 2006.

[65] T. Albrecht, "3D statistical shape models of human bones: their construction using a finite element registration algorithm, formulation on Hilbert spaces, and application to medical image analysis," in *PhD thesis,* University of Basel, 2011.

[66] M. E. Leventon, *Statistical models in medical image analysis*, PhD thesis, Massachusetts Institute of Technology, 2000.

[67] J. Lötjönen and T. Mäkelä, "Elastic matching using a deformation sphere," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2001*, Utrecht, 2001.

[68] 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.

[69] T. F. Chan and L. A. Vese, "Active Contours Without Edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266-277, 2001.

[70] W. Liu, *Segmentation of Lungs from Volumetric CT-Scan Images Using Prior Knowledge*, PhD thesis, University of Southampton, 2014.

[71] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3D medical image segmentation: A review," *Medical Image Analysis*, vol. 13, no. 4, pp. 543-563, 2009.

[72] G. Li and R. W. Miller, "Volumetric Image Registration of Multi-modality Images of CT, MRI and PET," in *Biomedical Imaging*, In Tech, 2010, pp. 1-25.

[73] M. Modat, J. McClelland and S. Ourselin, "Lung registration using the NiftyReg package," in *Medical Image Analysis for the Clinic-A Grand Challenge*, Beijing, 2010.

[74] W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes and B. M. Tsui, "4D XCAT phantom for multimodality imaging research," *Medical physics*, vol. 37, no. 9, pp. 4902-4915, 2010.

[75] J. Wang, H. Lu, Z. Liang, D. Eremina, G. Zhang, S. Wang, J. Chen and J. Manzion, "An experimental study on the noise properties of x-ray CT sinogram data in Radon space," *Physics in Medicine & Biology*, vol. 53, no. 12, pp. 3327-3341, 2008.

[76] S. Hu, E. A. Hoffman and J. M. Reinhardt, "Automatic Lung Segmentation for Accurate Quantitation of Volumetric X-Ray CT Images," *IEEE transactions on medical imaging*, vol. 20, no. 6, pp. 490-498, 2001.

[77] M. J. Guy, G. D. Flux, P. Papavasileiou, M. A. Flower and R. J. Ott, "RMDP: A Dedicated Package For 131I SPECT Quantification, Registration and Patient-Specific Dosimetry," *Cancer Biotherapy and Radiopharmaceuticals*, vol. 18, no. 1, pp. 61-69, 2003.

[78] D. L. Chen and P. E. Kinahan, "Multimodality Molecular Imaging of the Lung," *Journal of Magnetic Resonance Imaging*, vol. 32, no. 6, pp. 1409-1420, 2010.

[79] X. Liu, Z. Tang, M. Wang and Z. Song, "Deformable multi-modal registration using 3D-FAST conditioned mutual information," *Computer Assisted Surgery*, vol. 22, no. 1, pp. 295-304, 2017.

[80] F. Zhu, J. Ren, M. Ding and X. Zhang, "The Spiking Cortical Model Based Structural Representations for Non-Rigid Multi-Modal Medical Image

Registration," *Journal of Medical Imaging and Health Informatics,* vol. 7, no. 5, pp. 1001-1004, 2017.

[81] H. Ge, G. Wu, L. Wang, Y. Gao and D. Shen, "Hierarchical Multi-modal Image Registration by Learning Common Feature Representations," in *International Workshop on Machine Learning in Medical Imaging,* 2015.

[82] Q. Liu and H. Leung, "Log-Euclidean metric for robust multi-modal deformable registration," in *20th International Conference on Information Fusion* , Xi'an, 2017.

[83] K. Kasiri, P. Fieguth and D. A. Clausi, "Self-similarity measure for multi-modal image registration," in *IEEE International Conference on Image Processing*, Phoenix, 2016.

[84] H. D. Tagare and M. Rao, "Why Does Mutual-Information Work for Image Registration? A Deterministic Explanation," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 37, no. 6, pp. 1286-1296, 2015.

[85] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging,* vol. 16, no. 2, pp. 187-198, 1997.

[86] 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.

[87] L. Papp, M. Zuhayra, E. Henze and U. Luetzen, “Extended Normalized Mutual Information for Lung SPECT - CT Registration,” in *3rd International Conference on Bioinformatics and Biomedical Engineering*, Beijing, 2009.

[88] C. Guetter, M. Wacker, C. Xu and J. Hornegger, “Registration of Cardiac SPECT/CT Data Through Weighted Intensity Co-occurrence Priors,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007*, Brisbane, Springer, 2007, pp. 725-733.

[89] Y. Lu, R. Liao, L. Zhang, Y. Sun, C. Chefd’hotel and S. H. Ong, “Learning-Based Deformable Registration using Weighted Mutual Information,” in *21st International Conference on Pattern Recognition*, Tsukuba, 2012.

[90] J. Woo, M. Stone and J. L. Prince, “Multimodal Registration via Mutual Information Incorporating Geometric and Spatial Context,” *IEEE Transactions on Image Processing*, vol. 24, no. 2, pp. 757-769, 2015.

[91] P. A. Legg, P. L. Rosin, D. Marshall and J. E. Morgan, “Feature neighbourhood mutual information for multi-modal image registration: an application to eye fundus imaging,” *Pattern Recognition*, vol. 48, no. 6, pp. 1937-1946, 2015.

[92] L. Batina, B. Gierlichs, E. Prouff, M. Rivain, F.-X. Standaert and N. Veyrat-Charvillon, “Mutual Information Analysis: a Comprehensive Study,” *Journal of Cryptology*, vol. 24, no. 2, pp. 269-291, 2011.

[93] J. Jones, E. Lewis, G. Matthew and K. Wells, “A virtual dissection based registration to model patient-specific respiratory motion,” in *IEEE Nuclear Science Symposium and Medical Imaging Conference*, Orlando, 2009.

[94] L. Alvarez, R. Deriche, T. Papadopoulo and J. Sánchez, “Symmetrical Dense Optical Flow Estimation with Occlusions Detection,” *International Journal of Computer Vision* , vol. 75, no. 3, pp. 371-385, 2007.