A General Framework in Single and Multi-modality Registration for Lung Imaging Analysis Using Statistical Prior Shapes

Zheng Cui[1,*], Sasan Mahmoodi[1], Matthew Guy[2], Emma Lewis[3], Tom Havelock[4], Michael Bennett[4], Joy Conway[4]

[1] School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, United Kingdom

[2] Department of Imaging Physics, University Hospital Southampton NHS Foundation Trust, Southampton, SO16 6YD, United Kingdom.

[3] Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford GU2 7JP, United Kingdom

[4] Southampton NIHR Respiratory Biomedical Research Unit, University Hospital Southampton NHS Foundation Trust, Southampton SO16 6YD, United Kingdom

[*]Email: tcxzz1988@yahoo.com Tel: +44 (0)2380593284

Abstract: Background and Objective: A fusion of multi-slice computed tomography (MSCT) and single photon emission computed tomography (SPECT) represents a powerful tool for chronic obstructive pulmonary disease (COPD) analysis. In this paper, a novel and high-performance MSCT/SPECT non-rigid registration algorithm is proposed to accurately map the lung lobe information onto the functional imaging. Such a fusion can then be used to guide lung volume reduction surgery. Methods: The multi-modality fusion method proposed here is developed by a multi-channel technique which performs registration from MSCT scan to ventilation and perfusion SPECT scans simultaneously. Furthermore, a novel parameter-reduced function is also proposed to avoid the adjustment of the weighting parameter and to achieve a better performance in comparison with the literature. Results: A lung imaging dataset from a hospital and a synthetic dataset created by software are employed to validate single- and multi-modality registration results. Our method is demonstrated to achieve the improvements in terms of registration accuracy and stability by up to 23% and 54% respectively. Our multi-channel technique proposed here is also proved to obtain improved registration accuracy in comparison with single-channel method. Conclusions: The fusion of lung lobes onto SPECT imaging is achievable by accurate MSCT/SPECT alignment. It can also be used to perform lobar lung activity analysis for COPD diagnosis and treatment.

Key words: Multi-modality image fusion, Parameter-reduced method, Statistical modelling, Non-rigid registration

1. Introduction

1.1 Background

Chronic obstructive pulmonary disease (COPD) is a significant global cause of morbidity and mortality that presents a serious health care burden. One possible course of treatment for severe COPD is lung volume reduction surgery (LVRS), which involves removing some parts of the lung that are most affected by the disease. This treatment increases the volume of the remaining lung lobes to improve breathing in COPD patients. 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. Multi-slice computed tomography (MSCT) provides three-dimensional imaging of lung structure to allow the identification of the lung lobes. Single photon emission computed tomography (SPECT) allows lung activity to be measured using nuclear medicine imaging and makes it possible to identify regions of the lung that are not functioning as they should. The combination of the structural information from MSCT with the functional information from SPECT therefore represents a powerful tool to address the concerns of diseased lobe recognition. As the fusion from these two modalities can be performed directly on both MSCT and SPECT imaging, CT/SPECT registration has received a lot of attention in the past few years.

1.2 Previous Work

The method proposed in [1] intends to facilitate the registration by converting SPECT slices into CT-like images using scatter window. For the case that the isotope in the adjacent objects is more straightforward to detect, the method reported in [2] performs registration for tumors based on the spine SPECT imaging. Similar methods are studied in [3] and [4] as well. In addition, intensity uncertainty quantification is used to represent each voxel in [5]. Since the voxel mapping is independent of intensity, it overcomes the shortcoming that fluid-like registration methods are not applicable to multi-modality image registration. Recently, a series of sophisticated local descriptors based on the features and textures have been studied in [6] and [7]. Euclidian and Riemannian distances are employed to measure the correspondence

similarity respectively. Mutual information (MI) is demonstrated to be the best criterion for multi-modality image registration [8]. MI is widely used in medical image registration [2] [9]. Weighted mutual information established by normalized pointwise MI and prior knowledge is investigated in [10] to develop the registration accuracy and stability. Furthermore, structural feature-aided methods based on MI are proposed in [11] and [12] to improve registration performance.

Since the target image (i.e. SPECT) is usually deteriorated by artefacts and outliers (e.g. images depicted with the presence of tumors or lesion tissue can be considered as outliers) [13], the main challenge on the MSCT/SPECT registration is to prevent the excessive deformation of source image (i.e. MSCT). Prior knowledge-based methods are therefore studied to address the concern. Motion models in respiratory imaging are studied in [14] and [15], which can be used to guide the deformation during registration. However, the respiratory motion model requires a certain number of high-resolution scans within one breathing cycle and manual feature label on CT scans. Statistical deformation model (SDM) is firstly proposed in [16] and then developed to solve specific medical imaging problems [17]. In order to exploit the potential information of prior images, a technique was proposed in [18], 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 new atlas for SDM. In contrast, the implicit representation of prior knowledge is investigated in [19]. The prior shapes represented by level-set concentrate on the delineation of lung region and considerably reduces manual labor and computational cost.

1.3 Problems in SDM-based Work

The deformation $\boldsymbol{u}$ is normally represented by a three-component vector (u_x, u_y, u_z) along the directions of X, Y and Z axes. In the interest of capturing statistical information about deformation $\boldsymbol{u}$, 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 $R(\boldsymbol{u})$ [19]. The conventional cost function is written as

$$E(\boldsymbol{u}) = (1-\alpha)D(\boldsymbol{u}) + \alpha[(1-\beta)R(\boldsymbol{u}) + \beta S(\boldsymbol{u})] \quad (1)$$

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

In the literature (e.g. see [17] and [19]), the distribution of prior deformations is approximated by a multivariate normal distribution:

$$f(\boldsymbol{u}) = c \cdot exp\left(-\frac{1}{2}(\boldsymbol{u}-\overline{\boldsymbol{u}})^T {\Sigma_{cov}}^{-1}(\boldsymbol{u}-\overline{\boldsymbol{u}})\right) \quad (2)$$

where c is a constant and Σ_{cov} is the covariance matrix of prior deformations.

Since $ln[f(u)]$ is proportional to $-(u-\bar{u})^T {\Sigma_{cov}}^{-1}(u-\bar{u})$, the statistical term $S(u)$, which is supposed to have a minimum, is represented in association with (2), i.e.:

$$S(\boldsymbol{u}) = (\boldsymbol{u}-\overline{\boldsymbol{u}})^T {\Sigma_{cov}}^{-1}(\boldsymbol{u}-\overline{\boldsymbol{u}}) \quad (3)$$

Nevertheless, two unsolved issues always exist in this well-known framework and similar prior knowledge-based algorithms. At first, in order to minimize the cost contribution, $S(\boldsymbol{u})$ penalizes any displacement $\boldsymbol{u}$ that is not in accordance with the mean, $\overline{\boldsymbol{u}}$. However, by assuming that the deformations are associated with 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 $\boldsymbol{u}$ 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 specific circumstances.

1.4 Research Overview

As few features can be extracted from SPECT imaging and the landmark correspondence is practically unachievable, feature-based methods are inapplicable to the lung MSCT/SPECT alignment. In addition, lungs have the highly deformable property and their MSCT and SPECT images are captured at a different state of inhalation (i.e. SPECT scans are performed at tidal breathing, whilst the optimal parameters for MSCT for the delineation of the various structural

features of the lung requires suspended full inhalation). To the best of our knowledge, very few articles have been published to tackle the complicated lung MSCT/SPECT registration problem with prior knowledge and the problem described in Section 1.3. The motivation of this paper is therefore to develop a learning-based registration algorithm capable of capturing prior deformation information and achieving high-performance alignment of lung MSCT/SPECT imaging with less parameters.

In this paper, we substantially extend our previous work [20] to complete multi-modality image registration. We perform the simulations through lung dataset provided by Southampton General Hospital and synthetic phantoms created by 4D extended cardiac-torso (XCAT) and radionuclide multimodality dosimetry package (RMDP). The registration results are measured by different metrics to demonstrate the superior registration accuracy and stability achieved by the algorithm proposed here.

The contributions of this paper are as follows:

- Firstly, we propose a multi-channel registration method to improve the performance of lung anatomical and functional imaging alignment. Our method conducts a direct MSCT/SPECT registration which is independent of aided features and landmarks.
- Secondly, a novel cost function is proposed here for multi-modality registration. The parameter-reduced model avoids the unnecessary adjustment for a weighting parameter.
- Thirdly, a new statistical representation is presented to properly regularize the displacement based on the SDM.
- Finally, the synthetic imaging data and the ground-truth vector displacement offered by 4D XCAT together with RMDP are employed here to validate registration methods.

This paper is structured as follows: The mathematical derivations of our method proposed here are presented in Section 2. The experimental data are pre-processed in Section 3 followed by some implementation issues reported in Section 4. The experimental results achieved by our method are compared with those of the algorithms related to the classical framework (specified

in (1)) in Section 5 with in-depth discussion. Conclusions and future work are finally presented in Section 6.

2. Methodology

The non-rigid registration method proposed here starts with SDM which is briefly introduced in Section 2.1, followed by our cost function proposed in Section 2.2. Here we present mathematical derivations for single- and multi-modality image registration. The extended multi-channel technique proposed here is also clarified in this section. The metrics for validation are stated in Section 2.3.

2.1 Statistical Deformation Model

In this paper, the lung shapes as prior knowledge are extracted from various patients (e.g. healthy non-smoker, healthy smoker, mild and moderate COPD patients) to train the SDM. As the inter-subject information can only provide the global constraint against excessive deformations, rather than specific intra-patient voxel-wise regularization, prior lung shapes represented by level-sets are employed here for convenience. Although the size of lung varies considerably among individuals, the influence made by scale variations can be eliminated by employing a registration method which takes scaling transformation into consideration. Here we used the method proposed in [21]. It conducts similarity registration which is involved in scaling, rotation and translation transformations, to initially register prior shapes.

The B-spline-based non-rigid registration is then performed to compute the deformations for each prior shape. The deformation variability of prior shapes is generalized by principal component analysis, i.e.:

$$\boldsymbol{u}_{model} = \bar{\boldsymbol{u}} + \boldsymbol{V}\boldsymbol{\lambda} \tag{4}$$

where $\bar{\boldsymbol{u}}$ and $\boldsymbol{V}$ are the mean displacement and the eigenvectors of SDM respectively. The deformation can be represented by a dimension-reduced vector $\boldsymbol{\lambda}$, which comprises the values of coordinates along eigenvectors.

2.2 Parameter-reduced SDM Based Registration

In this section, the mathematical derivations of the parameter-reduced framework are clarified. Our method is intended to solve both single-modality (e.g. CT-CT or SPECT-SPECT) and multi-modality (e.g. CT-SPECT) registration problems, which are stated in Section 2.2.1 and 2.2.2 respectively. SPECT image alignment, which employs MI as a metric, can be performed using the framework proposed in Section 2.2.2. Multi-channel technique is stated in Section 2.2.3 for further improvement in performance.

2.2.1 Single-modality image alignment

Sum of squared distance (SSD) is used as the similarity metric, and the distance between two images is calculated by

$$D(\boldsymbol{u}) = \int (I_{src}(\boldsymbol{x}+\boldsymbol{u}) - I_{tar}(\boldsymbol{x}))^2 d\boldsymbol{x} \tag{5}$$

where I_{tar} and I_{src} are the target and registered source images respectively.

Here we replace $\boldsymbol{u}$ in (5) with a novel function $S(\boldsymbol{\lambda})$ manipulated by SDM to represent the regularized displacement. Our proposed statistical term, $S(\boldsymbol{\lambda})$, employed in a similarity term is presented in (6)-(9).

$$D\big(S(\boldsymbol{\lambda})\big) = \int \Big(I_{src}\big(\boldsymbol{x}+S(\boldsymbol{\lambda})\big) - I_{tar}(\boldsymbol{x})\Big)^2 d\boldsymbol{x} \tag{6}$$

where $S(\boldsymbol{\lambda})$ is a three-component vector $\Big(S(\lambda_x), S\big(\lambda_y\big), S(\lambda_z)\Big)$ and represented as:

$$S(\boldsymbol{\lambda}) = \overline{\boldsymbol{u}} + \sum_{i=1}^{K} \boldsymbol{V} \cdot S_{reg}(\boldsymbol{\lambda}_i) \tag{7}$$

In the above equations, $\boldsymbol{V}$ and $\overline{\boldsymbol{u}}$ have been computed in (4). Also K is the number of eigenvectors used for regularization and

$$S_{reg}(\boldsymbol{\lambda}_i) = H\left(\left|\frac{\lambda_i}{3\boldsymbol{\sigma}_i}\right|^2 - 1\right) \cdot Sign(\boldsymbol{\lambda}_i) \cdot 3\boldsymbol{\sigma}_i + H\left(1 - \left|\frac{\lambda_i}{3\boldsymbol{\sigma}_i}\right|^2\right) \cdot \boldsymbol{\lambda}_i \tag{8}$$

where

$$\boldsymbol{\lambda}_i = \boldsymbol{V}_i{}^T \cdot (\boldsymbol{u} - \overline{\boldsymbol{u}}) \tag{9}$$

Here, T indicates transpose. $\boldsymbol{\sigma}_i$ (i.e. ($\sigma_{x_i}, \sigma_{y_i}, \sigma_{z_i}$)) denotes a vector representing the standard deviation along X, Y and Z axes. The standard deviation is also considered as the square root of the corresponding eigenvalue. Each $\boldsymbol{\lambda}_i$ (i.e. ($\lambda_{x_i}, \lambda_{y_i}, \lambda_{z_i}$)) is regularized 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 penalizes those $\boldsymbol{\lambda}_i$ which are outside of the aforementioned range. The sign function $Sign(\cdot)$ is employed to retain the sign of each regularized $\boldsymbol{\lambda}_i$. $S_{reg}(\lambda)$ with respect to λ is shown in Fig. 1.

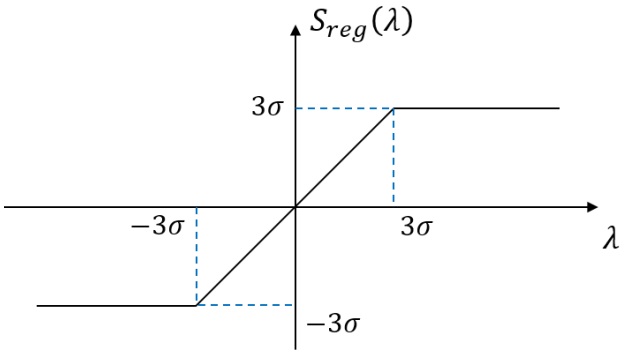


Fig. 1. Illustration of $S_{reg}(\lambda)$ with respect to λ.

Then, we use (10) to clarify $S(\boldsymbol{\lambda})$:

$$S(\boldsymbol{\lambda}) = \begin{cases} \overline{\boldsymbol{u}} + \sum_{i=1}^{K} \boldsymbol{V}_i \cdot Sign(\boldsymbol{\lambda}_i) \cdot 3\boldsymbol{\sigma}_i & if \quad \boldsymbol{\lambda}_i > 3\boldsymbol{\sigma}_i \; or \; \boldsymbol{\lambda}_i < -3\boldsymbol{\sigma}_i \\ \overline{\boldsymbol{u}} + \sum_{i=1}^{K} \boldsymbol{V}_i \cdot \boldsymbol{\lambda}_i & if \quad -3\boldsymbol{\sigma}_i \leq \boldsymbol{\lambda}_i \leq 3\boldsymbol{\sigma}_i \end{cases} \tag{10}$$

The smoothing term is rewritten in (11), and our cost function proposed here for single-modality registration is presented in (12).

$$R\big(S(\boldsymbol{\lambda})\big) = \int |\nabla S(\boldsymbol{\lambda})|^2 d\boldsymbol{x} \tag{11}$$

$$E_{sgl}(\boldsymbol{\lambda}) = (1-\alpha) D\big(S(\boldsymbol{\lambda})\big) + \alpha R\big(S(\boldsymbol{\lambda})\big) \tag{12}$$

The expected $\boldsymbol{\lambda}$ can be obtained by minimizing $E_{sgl}(\boldsymbol{\lambda})$ through a gradient descent technique:

$$\hat{\boldsymbol{\lambda}} = \underset{\boldsymbol{\lambda}}{argmin} \left(E_{sgl}(\boldsymbol{\lambda}) \right) \tag{13}$$

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

$$\frac{\partial E_{sgl}(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = \frac{\partial\left((1-\alpha)D(S)+\alpha R(S)\right)}{\partial S} \cdot \frac{\partial S(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} \tag{14}$$

Nevertheless, the derivative of $S(\boldsymbol{\lambda})$ does not exist if $H(\cdot)$ or $Sign(\cdot)$ possess singularities. In order to regularize $H(\cdot)$ and $Sign(\cdot)$, we employ (15) and (16) to represent $H(\cdot)$ and $Sign(\cdot)$ respectively [22],

$$H(p) = \frac{1}{2}\left(1 + \frac{2}{\pi} tan^{-1}\frac{p}{\varepsilon}\right) \tag{15}$$

$$Sign(q) = \frac{2}{\pi} tan^{-1}\frac{q}{\varepsilon} \tag{16}$$

and their derivatives are calculated by

$$H^{'}(p) = \frac{\varepsilon}{\pi(\varepsilon^2+p^2)} \tag{17}$$

$$Sign^{'}(q) = \frac{2\varepsilon}{\pi(\varepsilon^2+q^2)} \tag{18}$$

where ε is the regularizing parameter.

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

$$\frac{\partial S(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = \sum_{i=1}^{N} \boldsymbol{V}_i \cdot \left(\frac{\varepsilon}{\pi(\varepsilon^2+\boldsymbol{p}_i{}^2)} \cdot \frac{2\boldsymbol{\lambda}_i}{(3\boldsymbol{\sigma}_i)^2} \cdot Sign(\boldsymbol{\lambda}_i) \cdot 3\boldsymbol{\sigma}_i + H(\boldsymbol{p}_i) \cdot \frac{2\varepsilon}{\pi\left(\varepsilon^2+\boldsymbol{\lambda}_i{}^2\right)} \cdot 3\boldsymbol{\sigma}_i + \frac{\varepsilon}{\pi(\varepsilon^2+(-\boldsymbol{p}_i)^2)} \cdot \left(-\frac{2\boldsymbol{\lambda}_i}{(3\boldsymbol{\sigma}_i)^2}\right) \cdot \boldsymbol{\lambda}_i + H(-\boldsymbol{p}_i)\right) \tag{19}$$

where

$$\boldsymbol{p}_i = \left|\frac{\boldsymbol{\lambda}_i}{3\boldsymbol{\sigma}_i}\right|^2 - 1 \tag{20}$$

2.2.2 Multi-modality image alignment

In this section, a new parameter-reduced cost function is proposed for multi-modality image alignment. MI is conventionally employed to measure the similarity. Supposing that we use $I_{tar}(\boldsymbol{x})$ and $I_{src}(\boldsymbol{x}+\boldsymbol{u})$ to represent the target and registered source images, their MI can be calculated as

$$MI = H(I_{tar}) + H(I_{src}) - H(I_{tar}, I_{src}) = \iint p(a,b)\, log\frac{p(a,b)}{p(a)p(b)}\, da\, db \tag{21}$$

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,

$$E_{MI} = -\iint p(a,b)\, log\, \frac{p(a,b)}{p(a)p(b)}\, da\, db \tag{22}$$

Here we used gradient descent to investigate the optimal displacement $\boldsymbol{u}$ which minimizes E_{MI}. For the MI term, the derivative with respect to $\boldsymbol{u}$ is calculated as:

$$\frac{\partial E_{MI}}{\partial \boldsymbol{u}} = -\iint \frac{\partial p(a,b)}{\partial \boldsymbol{u}} \cdot \left(1 + log\, \frac{p(a,b)}{p(a)p(b)}\right) da\, db \tag{23}$$

In appendix A, the derivation for (23) is presented. The joint probability distribution $p(a,b)$ in (23) is described here by a function of $I_{tar}(\boldsymbol{x})$ and $I_{src}(\boldsymbol{x}+\boldsymbol{u})$ using kernel density estimation [23] in (24).

$$p(a,b,\boldsymbol{u}) = \frac{1}{V\sigma_{tar}\sigma_{src}} \int K\left(\frac{a-I_{tar}(\boldsymbol{x})}{\sigma_{tar}}, \frac{b-I_{src}(\boldsymbol{x}+\boldsymbol{u})}{\sigma_{src}}\right) d\boldsymbol{x} \tag{24}$$

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, which can be computed through the modified rule-of-thumb estimator [24]. Here we choose a Gaussian kernel to estimate the distribution and the kernel function $K(\cdot)$ is denoted as,

$$K(A,B) = \frac{1}{2\pi} exp\left(-\frac{1}{2}(A^2 + B^2)\right) \tag{25}$$

The derivative of $p(\mathrm{a},\mathrm{b},\mathbf{u})$ with respect to $\boldsymbol{u}$ is calculated by (26).

$$\frac{\partial p(a,b,\boldsymbol{u})}{\partial \boldsymbol{u}} = C \cdot \int K\left(\frac{a-I_{tar}(\boldsymbol{x})}{\sigma_{tar}}, \frac{b-I_{src}(\boldsymbol{x}+\boldsymbol{u})}{\sigma_{src}}\right) \cdot \left(-\frac{b-I_{src}(\boldsymbol{x}+\boldsymbol{u})}{{\sigma_{src}}^2}\right) \cdot \left(-\frac{\partial I_{src}(\boldsymbol{x}+\boldsymbol{u})}{\partial \boldsymbol{u}}\right) d\boldsymbol{x} \tag{26}$$

where

$$C = \frac{1}{V\sigma_{tar}\sigma_{src}} \tag{27}$$

By using (26) into (23), it is straightforward to write the following equation:

$$\frac{\partial E_{MI}(\boldsymbol{u})}{\partial \boldsymbol{u}} = C \cdot \int \iint K\left(\frac{a-I_{tar}(\boldsymbol{x})}{\sigma_{tar}}, \frac{b-I_{src}(\boldsymbol{x}+\boldsymbol{u})}{\sigma_{src}}\right)\left(1 + log\frac{p(a,b)}{p(a)p(b)}\right) da \left(\frac{I_{src}(\boldsymbol{x}+\boldsymbol{u})-b}{{\sigma_{src}}^2}\right) db \cdot \left(\frac{\partial I_{src}(\boldsymbol{x}+\boldsymbol{u})}{\partial \boldsymbol{u}}\right) d\boldsymbol{x} \quad (28)$$

Given the original cost function for multi-modality image alignment as:

$$E_{mul}(\boldsymbol{u}) = (1-\alpha)E_{MI}(\boldsymbol{u}) + \alpha R(\boldsymbol{u}) \quad (29)$$

we substitute $S(\boldsymbol{\lambda})$ proposed in (7) for $\boldsymbol{u}$ in (29) to arrive at (30).

$$E_{mul}(\boldsymbol{\lambda}) = (1-\alpha)E_{MI}\big(S(\boldsymbol{\lambda})\big) + \alpha R\big(S(\boldsymbol{\lambda})\big) \quad (30)$$

The derivative of the cost function with respect to $\boldsymbol{\lambda}$ is calculated as

$$\frac{\partial E_{mul}(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = \frac{\partial\big((1-\alpha)E_{MI}(S)+\alpha R(S)\big)}{\partial S} \cdot \frac{\partial S(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} \quad (31)$$

2.2.3 Multi-channel image alignment

As different radio-isotopes (Kr-81 for ventilation and Tc-99m for the perfusion) can emit gamma-rays at different energies, the signals from each isotope are separated at the receiver. In other words, SPECT ventilation (V) and perfusion (Q) can be acquired at the same time by the same scanner. V and Q are therefore treated as being aligned automatically. In this section, we propose a multi-channel alignment method, which contributes complimentary information in the case of impaired V or Q. Alignment is conducted 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:

$$E_{VQ} = (1-\alpha)\big[E_V\big(S(\boldsymbol{\lambda})\big) + E_Q\big(S(\boldsymbol{\lambda})\big)\big] + \alpha R\big(S(\boldsymbol{\lambda})\big) \quad (32)$$

where E_V and E_Q denote two MI-based similarity terms for MSCT/V and MSCT/Q alignments respectively. The solution for multi-channel method is therefore written in (33):

$$\frac{\partial E_{VQ}}{\partial \boldsymbol{\lambda}} = \frac{\partial\big((1-\alpha)[E_V(S)+E_Q(S)]+\alpha R(S)\big)}{\partial S} \cdot \frac{\partial S(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} \quad (33)$$

By employing gradient descent technique for optimization, the optimal $\boldsymbol{\lambda}_j$ can be obtained by iterative computation:

$$\boldsymbol{\lambda}_j = \boldsymbol{\lambda}_{j-1} - \frac{\partial E}{\partial \boldsymbol{\lambda}} \cdot \Delta t \tag{34}$$

where Δt is the step length to control the iterative rate for numerical solution. Small values for Δt rate may lead to slow convergence. On the other hand, large values of Δt may cause instability (divergence). Here, we have found that $\Delta t = 0.01$ leads to desirable results.

2.3 Evaluation of Registration Accuracy

In order to evaluate registration accuracy, two methods are introduced in this section. Based on the segmentation of the ROI using the method proposed in [25], the registration similarity can be measured by the normalized inner product (NIP) between the mask of the target image, $\phi_{tar}(\boldsymbol{x})$, and the warped mask of the source image, $\overline{\phi_{src}}(\boldsymbol{x})$.

$$NIP = \int \frac{\phi_{tar}(\boldsymbol{x})\,\overline{\phi_{src}}(\boldsymbol{x})}{\|\phi_{tar}(\boldsymbol{x})\|_2 \left\|\overline{\phi_{src}}(\boldsymbol{x})\right\|_2} d\boldsymbol{x} \tag{35}$$

where $\|\cdot\|_2$ denotes Euclidian norm.

The vector displacement (VD) is used to evaluate the performance by the voxel-wise displacement [26]. In comparison with the mask metric, VD provides a more precise metric for evaluation. Given the ground-truth displacement $\boldsymbol{T}$, the mean displacement error (MDE) is calculated as:

$$MDE = \frac{\int \|\boldsymbol{T}-\boldsymbol{u}\|_2 d\boldsymbol{x}}{V} \tag{36}$$

In summary, lower MDE and higher NIP indicate superior registration accuracy.

3. Data Collection

Experimental data are very important for validating performance of registration methods. Apart from the lung dataset provided by Southampton General Hospital used here, an advanced

simulation software (4D XCAT) is also employed here to synthesize the intra-patient data and provide ground-truth VD for non-rigid registration validation.

3.1 Phantom Data Processing

3.1.1 Phantom and VD generation using XCAT

4D XCAT phantom is developed for high-resolution anatomical imaging and advanced simulation of cardiac and respiratory motions, which are close to the scenario of real patients [27]. In order to obtain 3D synthetic MSCT and SPECT data for subsequent evaluation, a few of time points (i.e. image frames) are pre-selected in a breathing cycle. At each time point, the transient 3D images are sampled. Here, we configure the breathing cycle with 5 seconds, which starts from the maximum inhalation. The 3D attenuation and activity phantoms are sampled by eight frames within one cycle at even intervals, as shown in Fig.2. The air volume is an approximate value of people's lung capacity for the cycle presentation purpose only. The remaining parameters are properly configured by RMDP for phantom generation.

The VD is generated using another mode of XCAT. Each VD file reports the voxel movements from the first frame to one of the remaining frames, which is regarded as the ground-truth displacement for later validation.

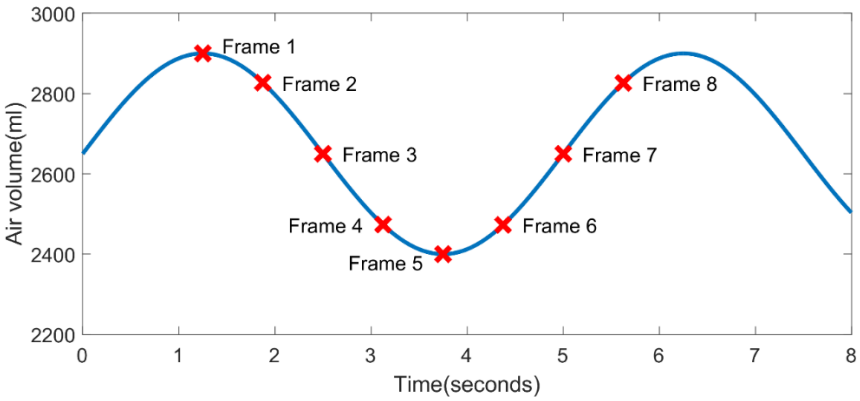


Fig. 2. XCAT phantoms are sampled by eight frames within one breathing cycle.

3.1.2 Phantom post-processing using RMDP

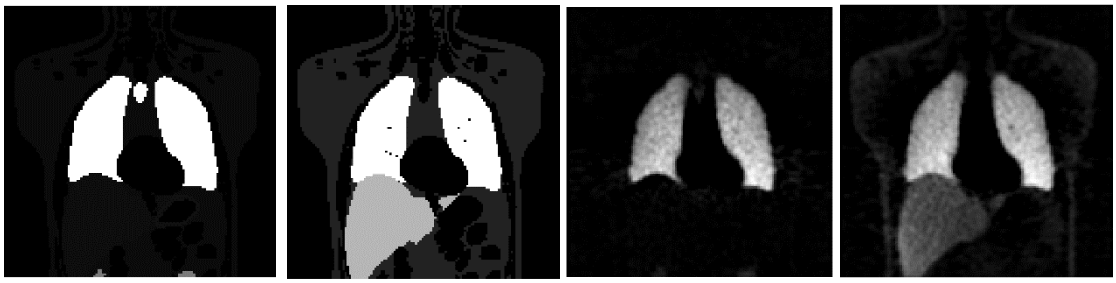
The attenuation image created by XCAT is produced by mono-energy (140 keV) photons, which is converted to for display by $\mu \times 1000$ numerically, where μ is the attenuation coefficient in cm^{-1}. Even though the attenuation image is produced based on mono-energy X-rays, RMDP takes the narrow beam (mono-energy) X-ray and turns them into a realistic distribution. Weighted x-ray spectrum sampling at 16 locations (16-140 keV) allows for Beam Hardening and realistic Hounsfield Unit calculation [28]. Fanbeam geometry is included to simulate truncation artefacts present in the data [29]. Furthermore, as the synthetic MSCT images are noise free, we added some noise to the phantom. It is reported in [30] that the actual noise characteristics of CT obeys Poisson distribution, where the rate parameter varies with different CT scanners. Therefore, we sampled the background regions of the forthcoming MSCT dataset obtained from the same CT scanner and calculated the mean of their greyscales as the rate parameter. Here, we practically set the Poisson distribution parameter to be 0.0058.

In this paper, RMDP is used to post-process the 3D emission and transmission data created by XCAT [31]. RMDP is originally developed to handle a high- and multi-energy emitter such as I-131, which means that RMDP has sufficient tools to deal with a low-energy mono-energetic emitter (e.g. Tc-99m) and a medium energy emitter (e.g. Kr-81). In order to optimize the SPECT acquisition parameters, the RMDP combining the XCAT and a fast simulation of the GE Infinia II gamma camera detector has been developed and validated experimentally [32]. Using this tool, the sensitivity and specificity of Tc-99m Ventilation-Perfusion SPECT versus planar scans can be evaluated and the optimal SPECT parameters can be determined. Furthermore, multiple spherical ‘defects’ are added interactively to the lung perfusion activity distributions, simulating Pulmonary Emboli (PE). The size, severity (partial or complete obstruction) and position of each PE are controlled by the user to allow user-control, visual feedback, of activity concentrations (kBq/cc) in 42 organs and regions. Initial settings follow those given by ICRP 53 [33], allowing for user-specified injection and scan times of perfusion and ventilation studies.

In our work, two necessary simulations were performed to obtain the activity phantoms for synthetic V and Q data. The XCAT parameter files were separately configured by RMDP according to user-defined simulated scanning parameters, as listed in Table 1. The activity phantom created by the XCAT presents a perfect distribution of activity as the activity is uniformly distributed in each organ, as shown in Fig. 3(a) and (b). The synthetic SPECT images are reconstructed using Filtered Back-Projection (FBP) with a Hanning filter (cut-off 0.9) and displayed in Fig. 3(c) and (d).

Table 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



a **b** **c** **d**

Fig. 3. Synthetic data created by XCAT and RMDP.
(a) Raw activity phantom for synthetic SPECT V.
(b) Raw activity phantom for synthetic SPECT Q.
(c) Reconstructed synthetic SPECT V image.
(d) Reconstructed synthetic SPECT Q image.

3.2 Experimental Data Pre-processing

The experimental data including the real medical data and synthetic images need to be pre-processed before non-rigid registration. The following three steps are applied to synthetic images whilst steps 1 and 3 are implemented on the real medical dataset.

3.2.1 Segmentation and down-sampling

In order to enhance the performance of image registration, it is necessary to segment the ROI initially. Furthermore, the segmented MSCT scans need to be down-sampled to the same resolution ($128 \times 128 \times 128$) as the SPECT scans. The boundaries of lungs in SPECT imaging are very blurred. However, 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, as shown in Fig. 3(c) and (d)), which greatly decreases the difficulty of segmentation.

3.2.2 Artificial defects

The SPECT 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 Fig 4(a) and (b). These abnormal SPECT images increase the challenge of alignment. In order to effectively validate the performance of various methods in Section 5, we employ RMDP to interactively add defects into generated phantoms. The synthetic abnormal SPECT V and Q are depicted in Fig. 4(c) and (d).

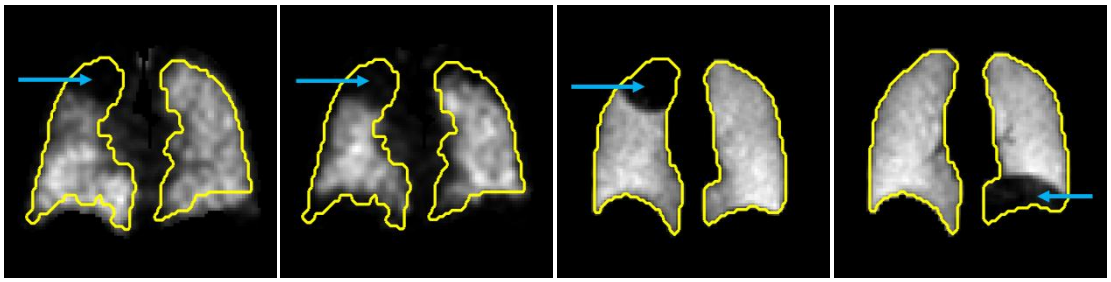


Fig. 4. Abnormal SPECT V and Q images. The yellow contours present the boundaries of lungs. The photopenic region (pointed by blue arrow) indicates the location of defects.
(a) Abnormal SPECT V image from a moderate COPD patient.
(b) Abnormal SPECT Q image from a moderate COPD patient.
(c) Abnormal SPECT V image synthesized by XCAT phantom.
(d) Abnormal SPECT Q image synthesized by XCAT phantom.

3.2.3 Transform to reference domain

As stated in Section 2.1, the training process starts with the similarity registration for prior shapes. The SDM is trained based on the deformations obtained by non-rigid registration from registered shapes to the target shape. Therefore, the SDM can only be applied to the coordinate system, in which the unknown images are aligned to the target shape under similarity registration. The coordinate system associated with the target shape is referred to as the reference domain here. In other words, before any SDM-based registration, unknown target and source images, namely the testing images here, are required to be transformed to the reference domain [18]. The transformation parameters can be computed by the similarity registration from any testing image to the target prior shape.

4. Implementation Issues

In order to obtain improved results and accelerate convergence, the multi-resolution technique is employed here. 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. Moreover, it is advised to reduce MSCT and SPECT imaging difference in terms of histogram by histogram matching.

In order to compare the results of our proposed method with those of the similar state-of-the-art algorithms in the literature, the method in [17] is referred here and named as Ref method for convenience. The main framework of the Ref method is described to (1), where $D(\boldsymbol{u})$ is evaluated by SSD and inverse MI for single- and multi-modality registration measurements respectively. Compared with our method proposed here, Ref method considers prior knowledge as an additional independent term in the cost function, which requires further investigation and manual adjustment for the weighting parameter β. The shortcomings of parameter increment have been reported in Section 1.3. In contrast, the prior term $S(\boldsymbol{\lambda})$ is defined as the variable of $D\big(S(\boldsymbol{\lambda})\big)$ and $R\big(S(\boldsymbol{\lambda})\big)$ in our proposed work, which simplifies the calculation and also leads to a reasonable and competitive result.

For Ref method, since the optimal weighting parameter β varies across different scenarios, the optimal value is fairly investigated and set for the following experiments. In addition, the parameter α always presents in both Ref and our proposed methods to balance two terms. The optimal α slightly vary with different data and a desired balance contributes to high registration accuracy. With respective to α, a large value gives rise to less flexibility whilst a small value leads to weak regularized ability. Here, it is empirically assigned the same value (i.e. 0.5) for a fair comparison.

5. Experimental Results and Discussions

Here, we employ 21 out of 32 lung shapes to establish SDM whereas the number of eigenvectors used for SDM is determined by a ratio of the sum of 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 eigenvectors (NoE) for SDM [34]. In our dataset, since the cumulative ratios of five eigenvalues along three directions are all above 0.98, it is demonstrated that five or more eigenvectors are adequate to model the deformable field.

5.1 Single-modality Registration Using Lung Shapes

The experiments are performed over 11 test subjects using a leave-one-out strategy. Before registration, one subject is chosen as the target, from which one of the lobes is manually removed (i.e. the right lower lobe of the target lung shape is removed, as shown in Fig. 5). Each of the remaining 10 source lung shapes is aligned to the defective target shape using Ref and the method proposed here.

In order to conduct numerical comparisons, the removed lobe is added back to the target lung shape in the subsequent evaluation. NIP, computed by using (35), is used here to measure alignment accuracy. The results with respect to the average NIPs for 11 groups are listed in Table 2. It can be observed that the accuracy achieved by the method proposed here is always higher than that of the Ref method, which demonstrates that our parameter-reduced method

manages to reasonably maintain the lung shape during non-rigid alignment and then obtain superior performance.

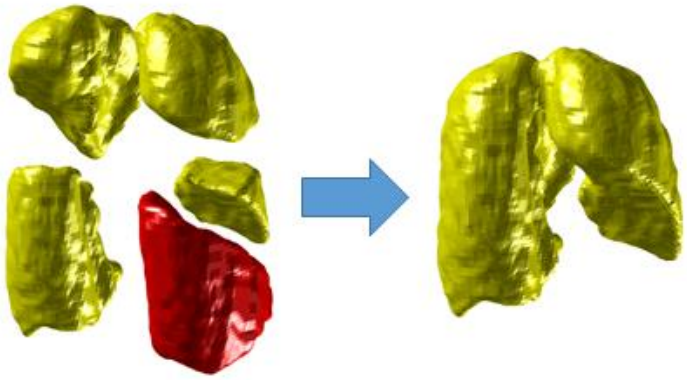


Fig. 5. The right lower lobe (highlighted in red) is removed from the target lung shape (colored in yellow) to create a synthetic defective lung shape (shown on the right-hand side).

Table 2. Average NIP for 11 test groups.

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10	Group 11
Ref	0.82	0.82	0.79	0.79	0.81	0.82	0.80	0.80	0.84	0.84	0.83
Proposed	0.86	0.85	0.85	0.83	0.85	0.85	0.85	0.83	0.88	0.86	0.86

Here, we also present three image examples to visually evaluate their accuracies in Fig. 6. The background image is depicted by signed distance function for visual convenience. The photopenic region at the bottom of the right lung indicates the manually removed lobe, which easily causes over-deformation by large gradient. The ref method intends to avoid over-deformation by increasing the value of optimal β, which sacrifices the flexibility in deformation and leads to fast convergence. In contrast, the framework proposed in our paper not only maintains a reasonable lung shape but also achieves higher accuracy in non-rigid registration.

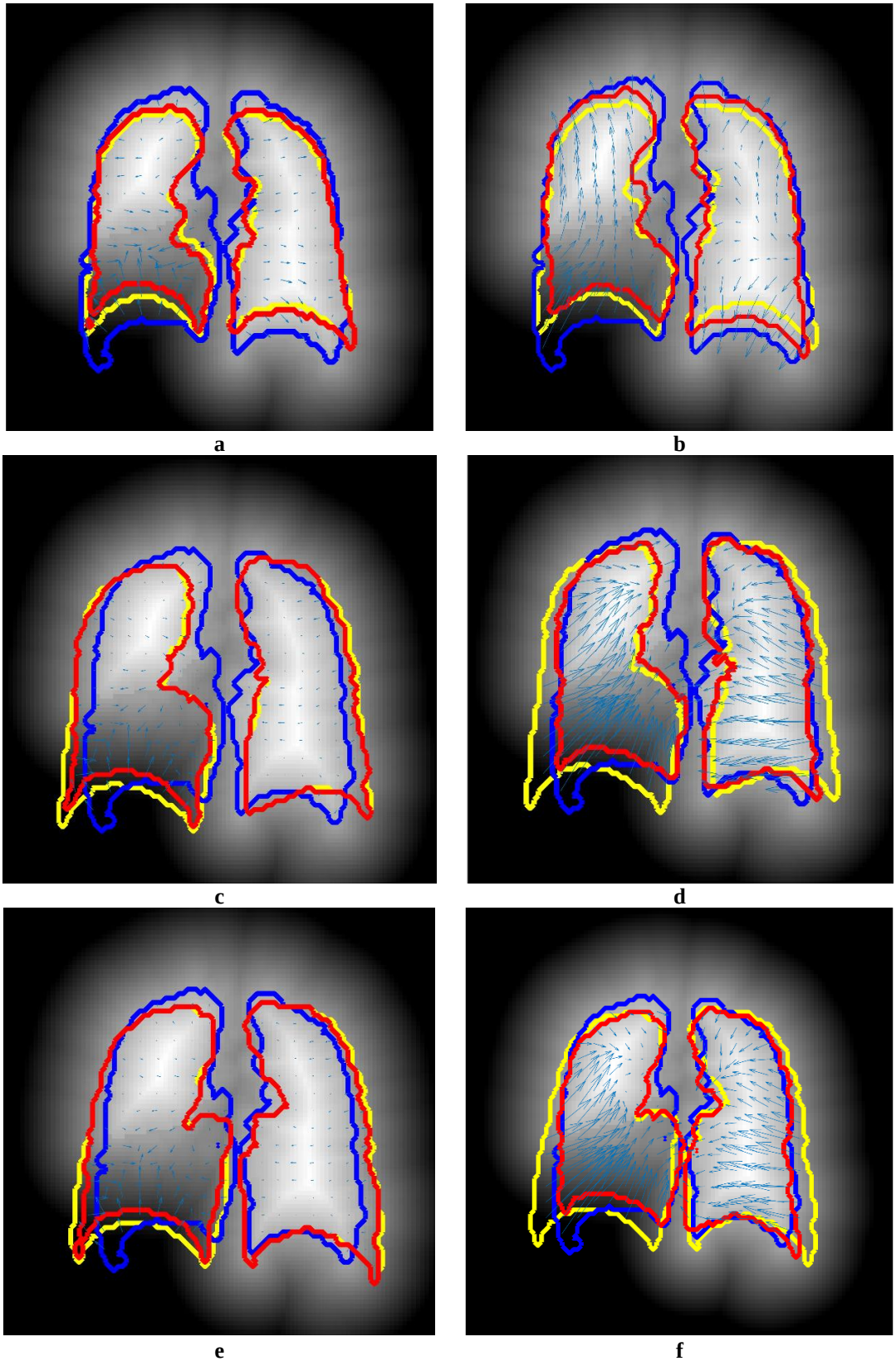


Fig. 6. Registration results for the Ref method (left column) and our method (right column). Blue contours indicate the ground-truth contours of the lung shape (target shape: subject 1). Blue arrows indicate the deformable field in non-rigid registration. Yellow and red contours represent the contours of the lung shape before and after registration respectively. (a,b) Source shape: Subject 2. (c,d) Source shape: Subject 4. (e,f) Source shape: Subject 8.

5.2 Multi-modality Registration Using Medical Imaging Data

In our medical dataset, MSCT scans, a pair of SPECT (V and Q) scans and low-dose CT scans are available for each subject. The low-dose CT scans and SPECT images for each patient were sequentially acquired by the same equipment on the same visit, where patients were instructed to remain still during the scan. In this study, we have therefore assumed that low-dose CT lung masks segmented manually by clinical experts are an approximate for the ground-truth masks of the corresponding SPECT V and Q images. The ground-truth lung mask of MSCT is segmented by Apollo (Vida Diagnostics Inc, Iowa, USA).

In the following experiments, each test 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, namely single-channel registration) and simultaneously (i.e. MSCT-V&Q, namely multi-channel technique). The experiments are divided in two groups: intra-patient registration (for clinical purpose) and inter-subject registration (for validation purpose, namely cross-validation), as shown in Fig. 7. The total number for V-MSCT, Q-MSCT and V&Q-MSCT alignment experiments is 363.

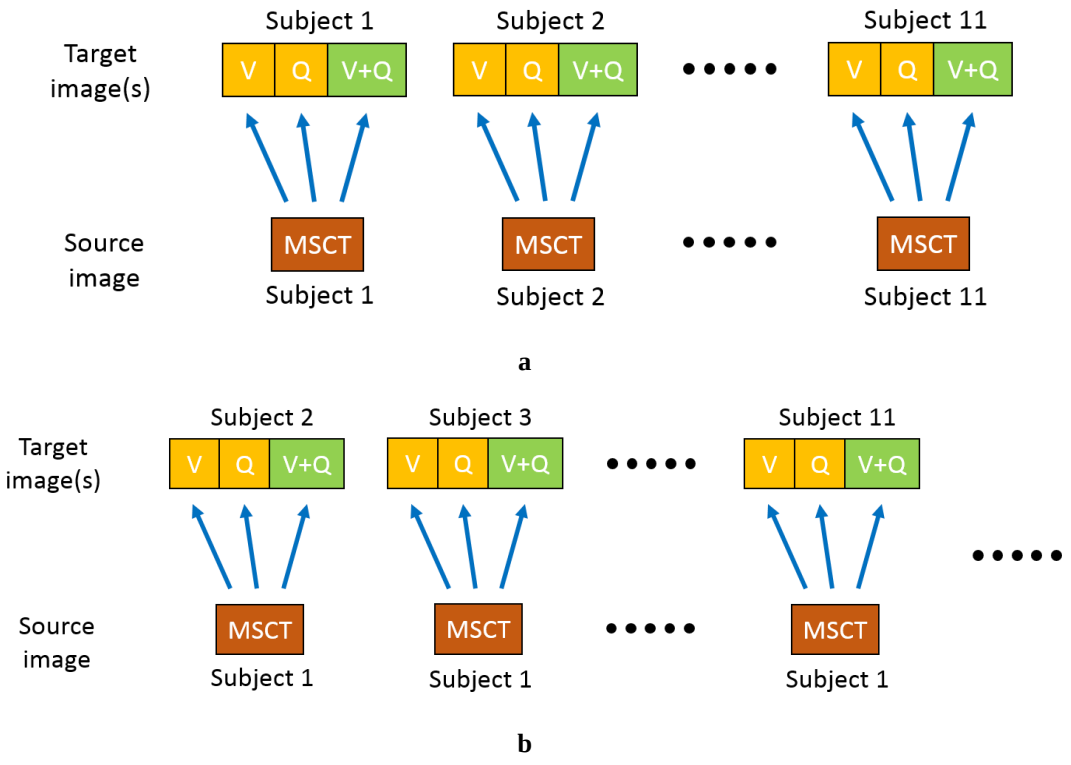


Fig. 7. Experimental design with medical imaging data. (a) Diagram of intra-patient registration. (b) Diagram of inter-subject registration using Subject 1 as an example.

Apart from the Ref method and our method proposed here, NiftyReg presented in [35] is also used for comparison. Compared with our method, NiftyReg does not use any prior knowledge, with a cost function denoted in (29). 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 Fig. 8. The average NIP is calculated using the NIPs obtained from the experiments with the same target image. In terms of intra-patient registration, since the SDM is employed to constrain the excessive deformation, the Ref and our method apparently outperform NiftyReg in single-channel experiments. Compared with the Ref method, the algorithm proposed here achieves improved accuracy in both single-channel and multi-channel experiments. In terms of the cross-validation, due to a relatively large difference between the target and source images, the performances of two learning-based methods slightly decrease under the influence of the regularized deformation. Nevertheless, the outstanding average NIPs are still obtained by the method proposed here. It is demonstrated that our method manages not only to constrain the excessive deformation but also to provide a desired registration solution.

For a moderate COPD patient, a defective region in the lung imaging is inevitably regarded as background by a non-learning method such as NiftyReg, followed by excessive deformations in non-rigid registration. A typical intra-patient registration example in coronal view using single-channel (MSCT-Q) method is presented in Fig. 9(a), (c) and (e). The Ref and our methods manage to constrain the deformations and achieve relatively reasonable results. However, as argued in Section 1.3, the Ref method tends to penalize any displacement, and thereby sacrifices flexibility. The drawbacks are illustrated by quivers in Fig. 9(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.

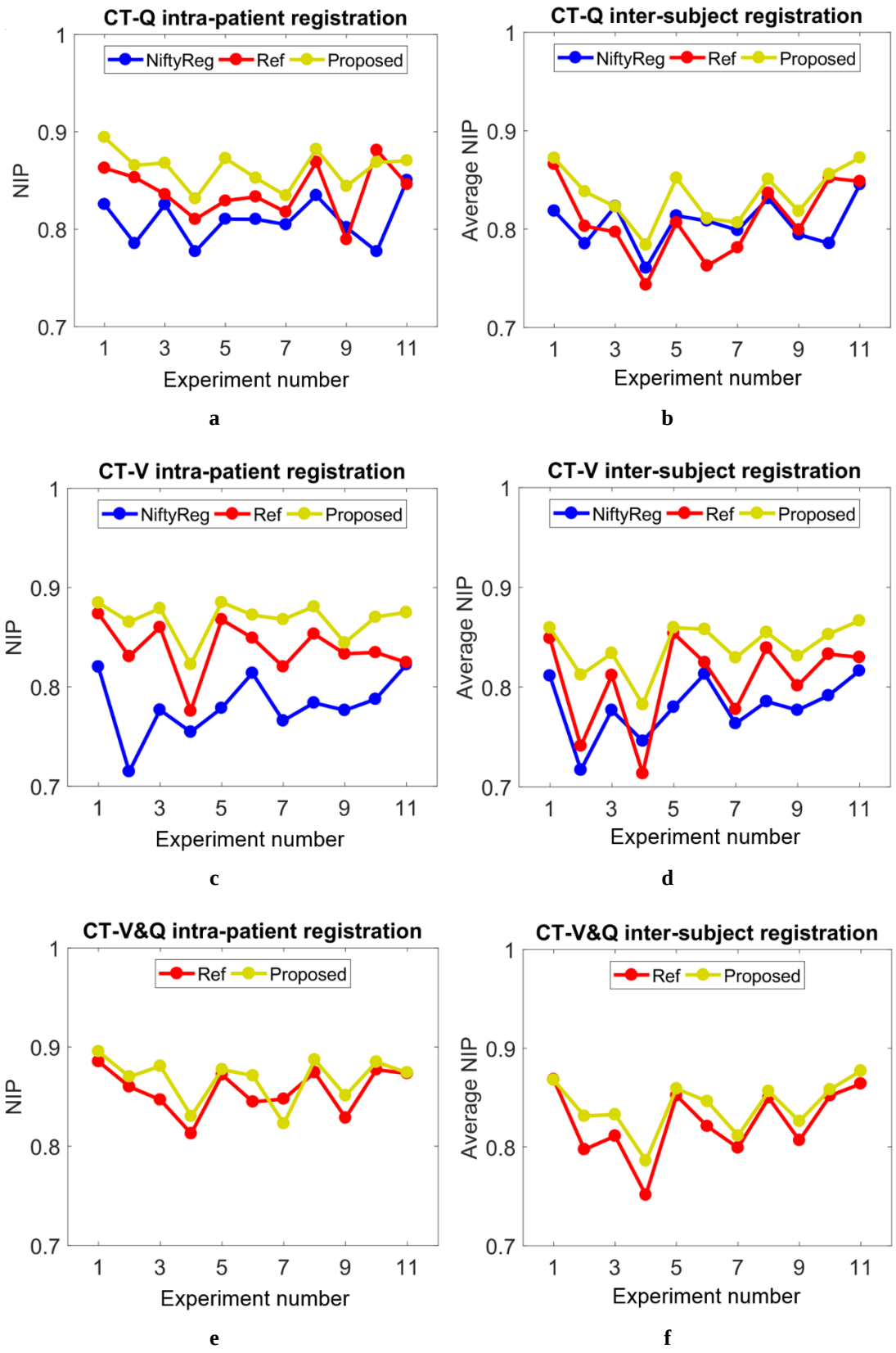


Fig. 8. Registration results are evaluated by NIP (left column) and average NIP (right column). (a,c) Single-channel intra-patient registration results. (b,d) Single-channel inter-subject registration results. (e,f) Multi-channel registration results.

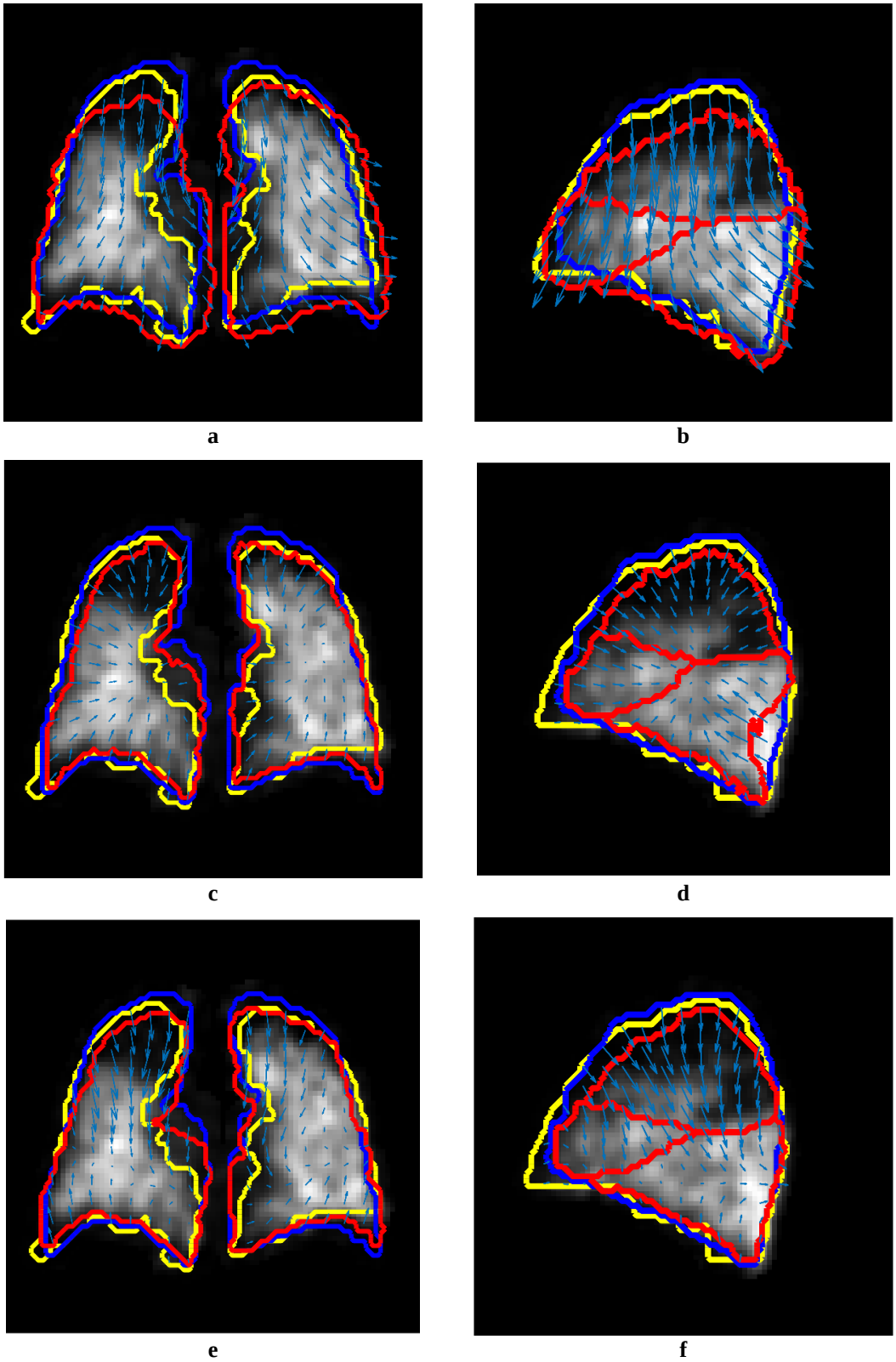


Fig. 9. Single-channel registration results for the NiftyReg (Row 1), Ref (Row 2) and proposed (Row 3) methods with accuracies 0.805, 0.818 and 0.835 against the initial 0.810. Blue contours indicate the ground-truth boundaries of the lung region in the SPECT scan (target). Blue arrows indicate the deformable field in non-rigid registration. (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 deformations are then used to map the lobes segmented from the MSCT scans onto the SPECT imaging. The fusion results are depicted in Fig. 9(b), (d) and (f) for NiftyReg, Ref and our 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.

The numerical results for all tests are summarized by the average NIPs and standard deviations in

Table 3. Compared with Ref, the method proposed here 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 our multi-channel method is enhanced by up to 28%. What is more important is that the multi-channel techniques using Ref and our algorithms are able to obtain improved alignment accuracies in comparison with either V/CT or Q/CT single-channel registration.

Table 3. 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

5.3 Multi-modality Registration Using Synthetic Data

As stated in Section 3.1.1, 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 by a multi-channel technique. In order to generate synthetic data challenging for registration algorithms, three scenarios are considered here separately. The three scenarios are arranged as:

Scenario1: normal SPECT V and abnormal Q,
Scenario 2: abnormal SPECT V and normal Q,
Scenario 3: abnormal SPECT V and abnormal Q.

We conduct seven aligning experiments for each scenario and evaluate the results with MDE and NIP. In this section, Ref and the method proposed here are compared in terms of registration accuracy and stability. Three scenarios are illustrated in Fig. 10 with multi-channel registration results for visual inspection. Here, we use two synthetic SPECT images together at frame 3 as a target. In order to present them simultaneously, we fuse the images from two channels in Fig.10 by scaling the greyscale of the sum of two images to [0,1]. This channel fusion is for illustration purposes only. Our method demonstrates to be closer to the ground-truth and capable of avoiding over-deformation in the scenario with defects.

The averages of MDE and NIP listed in Fig. 11 numerically demonstrate the superior registration performance achieved by the method proposed here in the three experimental scenarios as it obtains lower MDE and higher NIP. In terms of MDE, our algorithm decreases the average errors by up to 21%, 18% and 15% for scenarios 1, 2 and 3 respectively. The standard deviation of NIP demonstrates that the registration stability is improved by up to 39% and 54% for scenarios 1 and 3. With the increase of the NoE, the registration stability of our method improves accordingly, which is demonstrated by lower standard deviations of MDE and NIP. In contrast, the improved stability is rarely achieved by Ref method as this algorithm fails to provide an explicit boundary for deformations. The experimental results reveal the flaws in previous algorithms that the weighting parameter can neither provide a desirable flexibility nor contribute to boundary-finding during non-rigid registration.

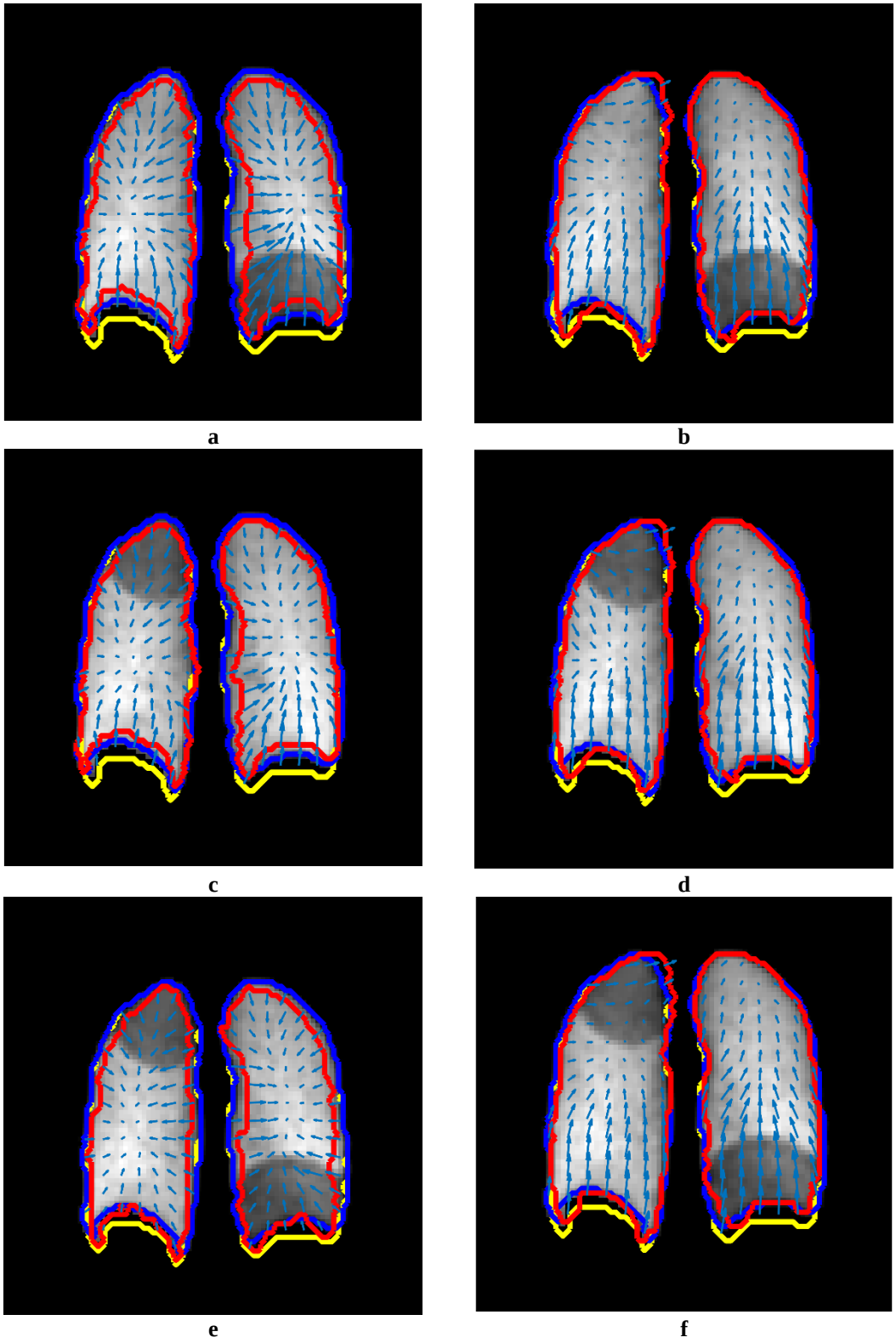


Fig. 10. Multi-channel registration results for the Ref method (left column) and our method (right column). Blue contours indicate the ground-truth boundaries of the lung region in the SPECT scan (target). Blue arrows indicate the deformable field in non-rigid registration. Yellow and red contours represent the boundaries of the lung region in an MSCT scan (source) before and after deformation respectively. For the purpose of illustrations, two target channels are fused as background in each sub-figure. (a,b) Scenario 1: normal SPECT V and abnormal Q. (c,d) Scenario 2: abnormal SPECT V and normal Q. (e,f) Scenario 3: abnormal SPECT V and abnormal Q.

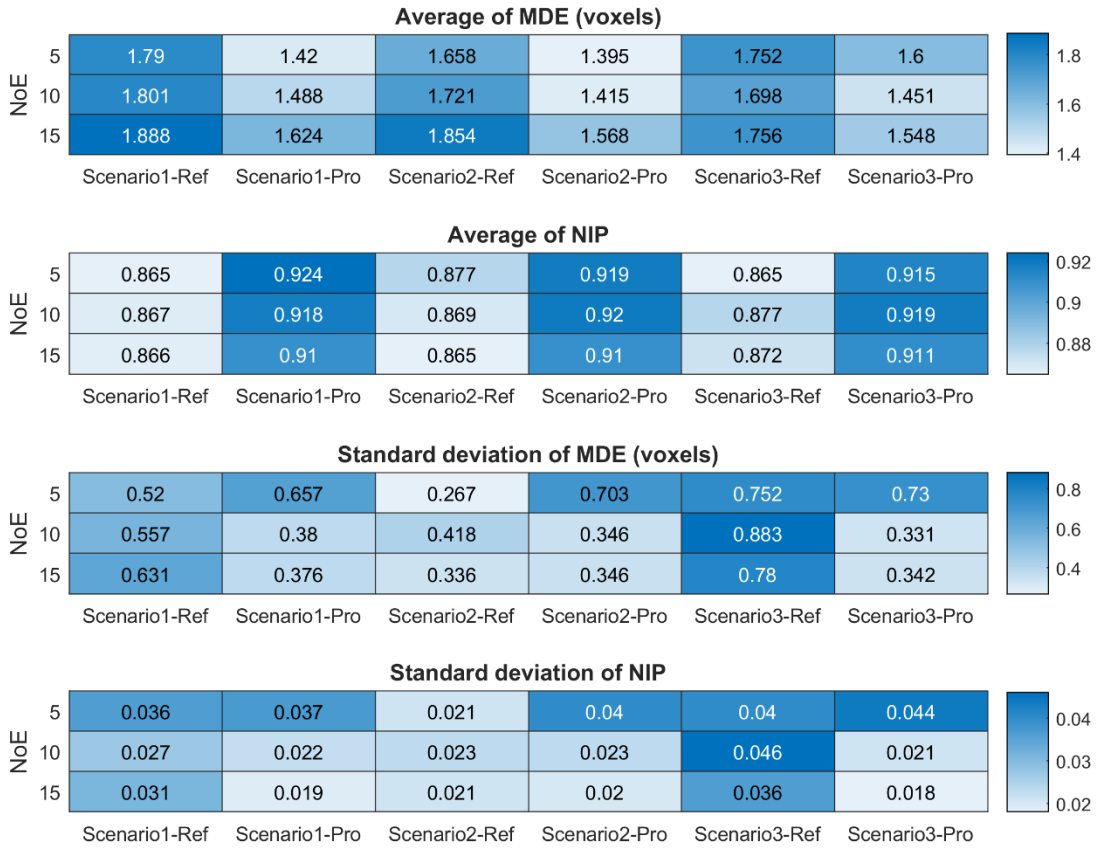


Fig. 11. Two algorithms with different NoEs are applied to three experimental scenarios. The registration results are evaluated with MDE and NIP.

In addition, the experimental results with respect to seven different target patterns (sampled at Frame 2 to 8) are depicted in Fig. 12. As illustrated in Fig.2, the volumes of lungs at Frame 2 and 8 are the most similar to that at Frame 1 whilst lungs have the largest deformations at Frame 5. According to the MDE results shown in Fig. 12, it is straightforward to notice that our method proposed here manages to slightly enhance the performance in experiment 1, 2, 6, 7 whereas it greatly outperforms Ref method in experiment 4 and 5 (NoE=15 and 10). As the clinical SPECT imaging is acquired by tidal breathing, the actual MSCT/SPECT registration is more likely to resemble experiment 4 or 5. Therefore, the high performance of our method on large-deformation alignment indicates desirable results in practical application. Also, the NIP results demonstrate that our method outperforms Ref method in all registration cases.

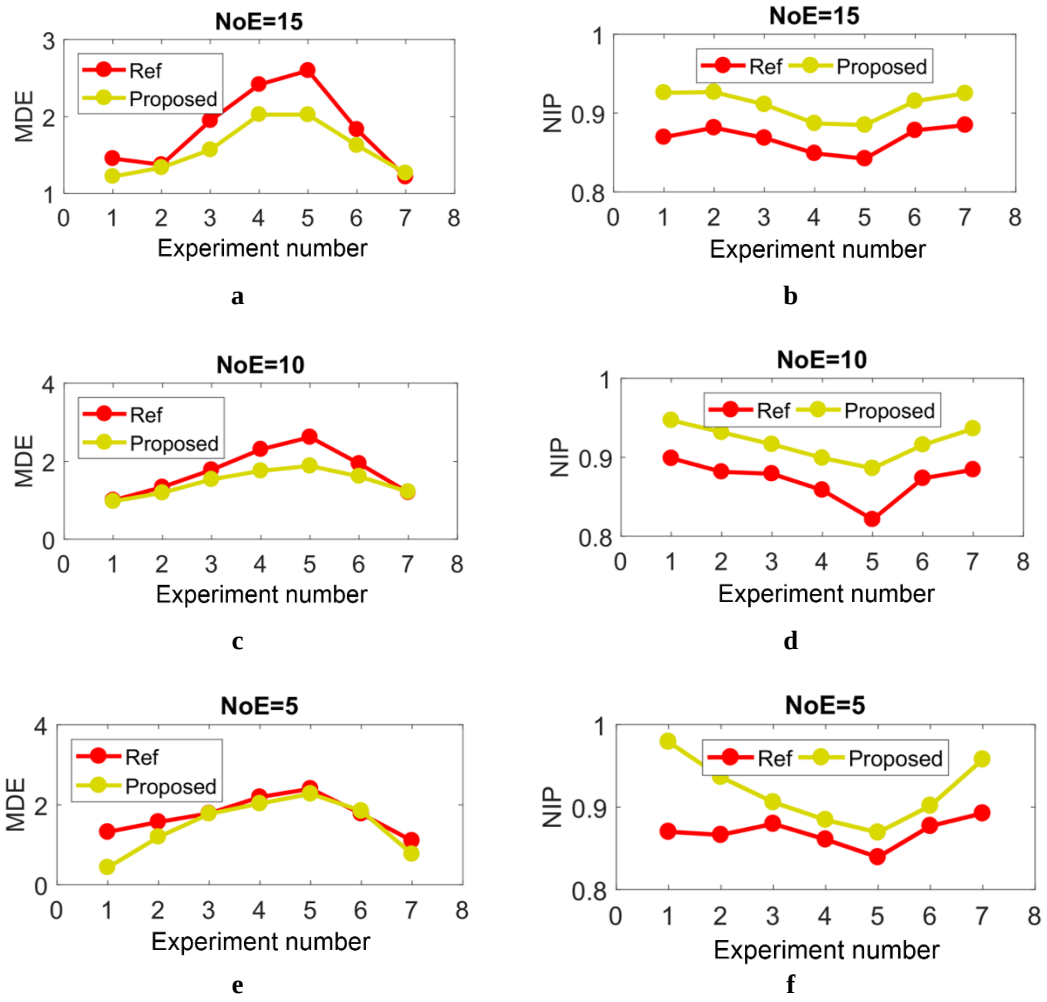


Fig. 12. (a,c,e) Registration results evaluated with MDE for different NoEs. (b,d,f) Registration results evaluated with NIP for different NoEs. The results are the average of three experimental scenarios.

6. Conclusions and Future Work

In this paper, a novel non-rigid registration method is proposed for lung MSCT/SPECT imaging alignment. Our method concentrates on the lung region and conducts direct registration through two MSCT/SPECT channels simultaneously. Furthermore, we also propose a novel parameter-reduced cost function, which addresses the concerns of investigating the optimal weighting parameter for statistical term and therefore provides more flexible displacement and solid boundary. The real lung dataset is employed here to train SDM and conduct experiments to evaluate the performance of multi-modality image fusion. The XCAT and RMDP are used to generate the synthetic imaging data and the ground-truth VD for further validation.

The method proposed here and two widely-used non-rigid registration methods are compared here by real medical imaging data and synthetic phantoms. The numerical results evaluated by NIP and MDE metrics demonstrate that the multi-channel method is able to improve registration accuracy and stability in comparison with single-channel framework as it can take more relevant information into consideration. Moreover, our proposed parameter-reduced cost function is validated to be capable of preventing the excessive deformation in the registration involved in defective imaging and able to outperform other similar learning or non-learning-based methods in terms of registration accuracy and stability. The fusion of lung lobes onto SPECT imaging is achievable by accurate MSCT/SPECT alignment.

In future research, with more intra-patient imaging data incorporated for SDM, the higher aligning performance can be achieved. Furthermore, the image fusion achieved here can be used to perform lobar lung activity analysis which enables the precise identification of the lung regions that are optimal targets for lung resection surgery.

Appendix A

The derivative of E_{MI} with respect to $\boldsymbol{u}$ is calculated as

$$\frac{\partial E_{MI}}{\partial \boldsymbol{u}} = -\iint \frac{\partial}{\partial \boldsymbol{u}} \left(p(a,b)\, log \frac{p(a,b)}{p(a)p(b)} \right) da\, db \tag{A-1}$$

As $p(a)$ is independent of $\boldsymbol{u}$, whereas $p(a,b)$ and $p(b)$ are dependent on the deformations during registration, (A-1) can be written as:

$$\frac{\partial E_{MI}}{\partial \boldsymbol{u}} = -\iint \left(\frac{\partial p(a,b)}{\partial \boldsymbol{u}} \left(1 + log \frac{p(a,b)}{p(a)p(b)} \right) - \frac{p(a,b)}{p(b)} \cdot \frac{\partial p(b)}{\partial \boldsymbol{u}} \right) da\, db \tag{A-2}$$

Since

$$\iint \left(\frac{p(a,b)}{p(b)} \cdot \frac{\partial p(b)}{\partial \boldsymbol{u}} \right) da\, db = \int \frac{\int p(a,b) da}{p(b)} \cdot \frac{\partial p(b)}{\partial \boldsymbol{u}} db = \frac{\partial \int p(b) db}{\partial \boldsymbol{u}} = 0 \tag{A-3}$$

(A-2) can be simplified as,

$$\frac{\partial E_{MI}}{\partial \boldsymbol{u}} = -\iint \frac{\partial p(a,b)}{\partial \boldsymbol{u}} \cdot \left(1 + log \frac{p(a,b)}{p(a)p(b)} \right) da\, db \tag{A-4}$$

which is the same as (23).

Acknowledgment

The CT/SPECT data used in this work were acquired as a part of a study into the application of imaging to the characterization of the phenotypes of COPD. The 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. Hereby, we sincerely appreciate that Dr. W. Paul Segars from Duke University generously provided the XCAT software for data collection and algorithm validation, which enriches the experimental results and evaluation methods.

Conflict of interest statement

None.

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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 Berlin Heidelberg, 2007, pp. 725-733.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] D. Chen, H. Xie, S. Zhang and L. Gu, "Lung respiration motion modeling: a sparse motion field presentation method using biplane x-ray images," *Physics in Medicine & Biology*, vol. 62, no. 19, p. 7855–7873, 2017.

[15] M. Wilms, I. Y. Ha, H. Handels and M. P. Heinrich, "Model-Based Regularisation for Respiratory Motion Estimation with Sparse Features in Image-Guided Interventions," in *Medical Image Computing and Computer-Assisted Intervention*, Athens, 2016.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] Z. Cui, S. Mahmoodi, J. Conway, M. Guy, E. Lewis, T. Havelock and M. Bennett, "A novel non-rigid registration method based on non-parametric statistical deformation model for medical image analysis," in *IEEE Nuclear Science Symposium and Medical Imaging Conference*, Atlanta, 2017.

[21] Z. Cui, S. Mahmoodi and M. Bennett, "A Robust and High-performance Shape Registration Technique Using Characteristic Functions," in *IEEE conference on Image Processing Applications and Systems*, Inria Sophia Antipolis, 2018.

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

[23] 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.

[24] 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.

[25] S. Dai, K. Lu, J. Dong, Y. Zhang and Y. Chen, "A novel approach of lung segmentation on chest CT images using graph cuts," *Neurocomputing*, vol. 168, no. 6, pp. 799-807, 2015.

[26] 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.

[27] 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.

[28] R. Birch, M. Marshall and M. G. Ardran, "Catalogue of spectral data for diagnostic X-rays," in *Hospital Physicists' Association*, London, 1979.

[29] A. J. Britten, M. Crotty, H. Kiremidjian, A. Grundy and E. J. Adam, "The addition of computer simulated noise to investigate radiation dose and image quality in images with spatial correlation of statistical noise: an example application to X-ray CT of the brain," *The British journal of radiology,* vol. 77, no. 916, pp. 323-328, 2004.

[30] 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.

[31] 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.

[32] J. Scuffham, *Quantitative NM Inkjet Phantom System and its Application to Lung Imaging,* MSc Dissertation, University of Surrey, 2016.

[33] Radiation Dose to Patients from Radiopharmaceuticals, ICRP Publication 53, 1988.

[34] 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.

[35] 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.