Lung 4D CT Image Registration Based on High-Order Markov Random Field

Peng Xue, Enqing Dong, *Member, IEEE,* and Huizhong Ji

***Abstract*—To solve the problem that traditional image registration methods based on continuous optimization for large motion lung 4D CT image sequences are easy to fall into local optimal solutions and lead to serious misregistration, a novel image registration method based on high-order Markov Random Field (MRF) is proposed. By analyzing the effect of the deformation field constraint of the potential functions with different order cliques in MRF model, energy functions with high-order cliques form are designed separately for 2D and 3D images to preserve topology of the deformation field. In order to preserve the topology of the deformation field more effectively, it is necessary to apply a smooth term and a topology preservation term simultaneously in the energy function and use logarithmic function to impose a penalty on the Jacobian matrix with high-order cliques in the topology preservation term. For the complexity of the designed energy function with high-order cliques form, Markov Chain Monte Carlo (MCMC) algorithm is used to solve the optimization problem of the designed energy function. To address the high computational requirements in lung 4D CT image registration, a multi-level processing strategy is adopted to reduce the space complexity of the proposed registration method and promotes the computational efficiency. In the DIR-lab dataset with 4D CT images and the COPD (Chronic Obstructive Pulmonary Disease) dataset with 3D CT images, the average target registration error (TRE) of our proposed method can reach 0.95mm respectively.**

***Index Terms*—Image Registration, 4D CT, Markov Random Field, Topology Preservation, Multi-level Processing Strategy**

I. Introduction

IMAGE registration is one of the basic research methods of digital image processing, and it has been widely used in many popular fields such as computer vision and pattern recognition. Because the deformation field obtained by registering the 4D CT image sequences of the lung can effectively describe the relative motion of the corresponding tissue structure between the image sequences, and thus can reflect the motion characteristics of the tumor or other anatomical features, therefore, the quality of the image registration directly affects the accuracy of lung motion modeling. Due to the influence of heart beats and respiratory movements, the local intensity inhomogeneity of the lung 4D CT images and the large deformation of the fine textures are caused. In this case, some traditional continuous optimization-based image registration methods are easy to fall into local optimal solutions [1], and obtain unacceptable results, which are not suitable for the registration of lung images.

The image registration method based on Markov Random Field (MRF) model is a non-rigid registration method using discrete optimization. The key of the method is to choose a suitable discrete optimization algorithm for an ill-posed, nonconvex optimization problem with several million degrees-of-freedom according to different forms of energy functions. For MRF-based image registration method, according to whether the high-order clique structure is included in the energy function, it can be divided into a low-order MRF-based image registration method and a high-order MRF-based image registration method [2]. In general, the low-order MRF-based image registration method only considers the energy function based on pairwise interactions between the variables of the field. Therefore, most of the modern discrete optimization algorithms such as graph cuts [3]-[5] and message passing [6], [7] are used to solve the optimization problem. Among them, the most representative optimization algorithm is the graph cuts algorithm proposed in [3], which converts the extreme value problem of energy function into the maximum flow-minimum cut problem. On the basis of [3], Tang *et al.* [8] first applied the graph cuts algorithm to medical image registration. In [8], the image registration problem was converted into a label problem in the MRF since the set of labels for each point was defined as a set of displacement vectors, finally, the graph cuts algorithm could be used to solve the optimization problem in the energy function. With the mature of the convex optimization theory, Komodakis *et al.* [5] proposed the Primal-Dual schema (namely, FastPD) based on the Lagrangian multiplier method in 2007. On the basis of [5], Glocker *et al.* [9] applied the FastPD algorithm to lung image registration successfully, and proposed a multi-scale incremental approach based on B-spline interpolation strategy. Apart from [9], Heinrich *et al.* [10] proposed a global optimization algorithm based on Minimum Spanning Tree (MST). And then, Heinrich *et al.* [11] also proposed to use hyper-labels to directly estimate the density change of lung tissue and the local lung ventilation.

This work was supported by the Fundamental Research Funds for the Central Universities (China), National Natural Science Foundation of China under Grant 81671848 and 81371635. Shandong Key Research and Development Program (Major Science and Technology Innovation Project). Natural Science Foundation of Shandong Province (ZR2017PF009). (Corresponding author: Enqing Dong.)

The authors are with the Department of Mechanical, Electrical and Information Engineering, Shandong University, Weihai 264209, China (e-mail: xuepeng2016@126.com; enqdong@sdu.edu.cn; jihuizhong131@qq.com)

In high-order MRF-based image registration, the energy functions are composed of complex high-order cliques above two-element cliques. Although the high-order cliques can impose constraints more effectively on the deformation field and can further improve the registration accuracy, but these modern discrete optimization algorithms mentioned above can't solve the energy functions with high-order cliques. In response to the above problem, in 2007, Kohli *et al.* [12] extended the applicable scope of the graph cut algorithm and applied it to a class of p^n Potts models with high-order cliques. In the same year, Rother *et al.* [13] proposed the Quadratic Pseudo-Boolean Optimization (QPBO) algorithm to improve the graph cut algorithm, so that QPBO algorithm can solve the non-submodular optimization problems. With the introduction of the QPBO algorithm, in order to solve the high-order MRF-based image registration problem with topology preservation term, in 2011, Cordero-Grande *et al.* [14] proposed an image registration framework based on QPBO algorithm. However, at the same time, the problem of low computational efficiency of the QPBO algorithm was found. After that, Cordero-Grande *et al.* [15] proposed a 2D image registration method that could maintain the topology by using parameter estimation and Markov Chain Monte Carlo (MCMC) based optimization algorithm. According to literature [15], MCMC-based optimization algorithm can effectively solve the optimization problem of the energy function with three-element cliques, and in the optimization process, since the MCMC algorithm only needs to update the labels of the current optimized control point, and does not need to update the labels of other control points in the same high-order cliques at the same time, so it has great potential for optimizing the energy function with more complex cliques. However, there is a problem of slow optimization.

In lung CT image registration, in order to solve the problem of large motion of fine textures in the lung, Han *et al.* [16] proposed an image registration method based on robust 3D SURF (Speeded Up Robust Features) descriptors for feature detection and matching. By introducing structural information into the MI (Mutual Information) metric, the SeSaMI (Self Similarity-MI) measure proposed by Rivaz *et al.* [17] can also solve the problem of large motion in lung image. Apart from [17], Zheng *et al*. [18] proposed a novel framework to simultaneously segment and register lung and tumor in serial CT data. In [18], Zheng *et al*. used a B-Spline based nonrigid transformation to model the lung deformation while imposing rigid transformation on the tumor to preserve the volume and the shape of the tumor. Soliman *et al.* [19] proposed a registration methodology to co-align successive 3D CT scans of a segmented lung object with a given prototype, the proposed method in [19] involves global and local alignment steps, and the local alignment step can maintain the lung anatomy effectively. For lung images with local intensity inhomogeneity, Normalized Gradient Fields (NGF) [20], [21] or Modality Independent Neighborhood Descriptors (MIND) [22], [23] are usually used. In addition, including spatial information by combining mutual information with a term based on the image gradient proposed by Pluim *et al.* in [24] can also be used to solve the problem of local intensity inhomogeneity. Based on the classical optical flow algorithm, Castillo *et al.* [25] proposed a compressed optical flow algorithm that can model lung compression, which can solve the problem of local intensity inhomogeneity in lung 4D CT images. After that, Cao *et al.* [26] proposed an accurate inverse-consistent symmetric optical flow method for lung 4D CT registration, which further improved the registration accuracy.

In recent related researches, Vishnevskiy *et al.* [27] proposed an isopTV method that used an isotropic total variation regularization term to constrain the deformation field. This method works best in the DIR-lab dataset with 4D CT images. In the same year, Rühaak *et al.* [28] proposed a method for estimate large motion in lung CT by integrating regularized keypoint correspondences into dense deformable registration. In [28], the lung parenchyma is extracted and the detection of keypoint correspondences enables robustness against large deformations by jointly optimizing over a large number of potential discrete displacements, the dense continuous registration achieves sub-voxel alignment with smooth transformations. Although Rühaak *et al.* method is best in the COPD dataset, it is based on accurate lung parenchymal image extraction, so there is no comparability between this method and other methods that don't require lung parenchymal extraction. In 2018, Eppenhof *et al.* [29] used 3D convolutional neural networks for 4D CT lung image registration, although the registration time can be greatly shortened while ensuring a certain registration accuracy, due to the limited amount of training data for network training, the registration accuracy has no advantage compared with other methods.

In order to further improve the accuracy of lung 4D CT image registration, a high-order MRF-based 4D CT image registration method (HO-MRF) is proposed in this paper. Based on high-order MRF, an energy function with high-order cliques is designed to maintain the topology of deformation field and MCMC algorithm is used to solve the optimization problem of the energy function. In view of the fact that the lung 4D CT image has many voxels and large local motion range, an effective multi-level processing strategy is adopted to reduce the complexity of the algorithm and improve the computational efficiency. The remainder of this paper is structured as follows. Section II introduces the construction of high-order MRF registration model and analyzes the role of the potential function based on the clique structure model in deformation field constraints. Then, we respectively designed energy functions with high-order cliques that can maintain the topology for 2D and 3D images. Section III introduces the MCMC algorithm and discuss how to make parameters selection. In section IV, a 2D model image and a 3D model image are taken as examples to verify the validity of the topology preservation term in the energy function of the design for different deformations. Section V describes the multi-level processing strategy of the lung 4D CT image registration in the form of detailed flow chart, and carries out the experiment on the DIR-lab dataset with the target registration error (TRE) as the evaluation index. Comparing the proposed method (HO-MRF) with several current best methods, the effectiveness

of the proposed method is verified. Finally, some conclusions and discussions are presented in section VI, together with the existing problems and further research ideas.

II. High-Order MRF Registration Model

A. General Form of The MRF Model

MRF theory provides a convenient and consistent way of modeling context-dependent entities such as image pixels and correlated features and it has been widely used in many fields such as image registration [8-15], image segmentation [30] and medical image diagnosis [31], [32]. In image registration, consider N-dimensional target image I: $\Omega \to \mathbb{R}^N$ and N-dimensional moving image J: $\Omega \to \mathbb{R}^N$, for each point $p \in \Omega$ (with spatial location x_p), there is a set of labels $\mathbf{L}$, which corresponds to N-dimensional discrete displacements $\mathbf{d} = (d_1, d_2, d_3 \ldots d_N) \in \mathbf{L}$. Therefore, an energy function in image registration can be expressed as the following form

$$E = \sum_{p \in \Omega} \left| I(x_p) - J(x_p + \mathbf{D}(x_p)) \right| + \omega \cdot R(\mathbf{D}) \quad (1)$$

where $\mathbf{D}$ is the deformation field composed of $\mathbf{d}$ and $\mathbf{D} = \sum_{p \in \Omega} \mathbf{d}_p$, $I(x_p)$ represents the intensity of the target image at point p, $J(x_p + \mathbf{D}(x_p))$ represents the intensity of the point p on the moving image after the deformation field is applied, R as a constraint term is usually a function of the deformation field $\mathbf{D}$, ω is a weight parameter, which determines the influence of the constraint term on the whole energy function.

In MRF, the constraint term R is usually determined by the three elements of the neighborhood system $\mathcal{N}$, the cliques $\mathcal{C}$ and the potential functions $\mathcal{V}$ in the random field. In a 2D image, a common neighborhood system is a 4-neighborhood system as shown in Fig. 1(a) or 8-neighborhood system as shown in Fig. 1(b). Accordingly, in a 3D image, there are 6-neighobrhood system, 18-neighborhood system and 26- neighborhood system. On this basis, we can define a clique as a subset of points according to the set of points in the neighborhood. According to the size of the clique (the number of points p included in the clique), a one-element clique, a two-element clique and a three-element clique can be defined as:

$$\begin{aligned} \mathcal{C}_1 &= \{p \,|\, p \in \Omega\} \\ \mathcal{C}_2 &= \{(p,q) \,|\, p \in \Omega, q \in \mathcal{N}_p\} \\ \mathcal{C}_3 &= \{(p,q,r) \,|\, p \in \Omega, q, r \in \mathcal{N}_p\} \end{aligned} \quad (2)$$

where $\mathcal{N}_p$ (light blue dots in Fig.1) represents the neighborhood of the point p (dark blue dots in Fig. 1) and Fig. 1(c)-Fig. 1(e) show the structure of each clique in the 8-neighborhood system of the 2D image respectively.

In MRF, the potential functions $\mathcal{V}$ are the core factor that determines the constraints. According to the Hammersley-Clifford theorem, the energy function E can be represented by the potential functions $\mathcal{V}$. Usually, a potential function is generally defined as a function of cliques, and its purpose is to calculate the cost of the corresponding potential function according to different order cliques. In image registration, the potential functions with one-element cliques represent a data term, the potential functions with two-element cliques and above are used to constrain the deformation field. The role of various potential functions with different cliques in image registration will be described in detail below.

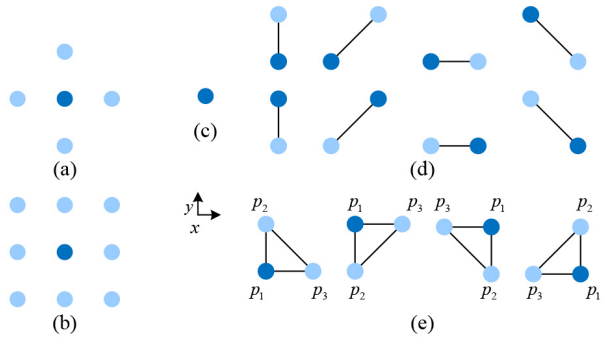


Fig. 1. Neighborhood system and clique categories in 2D image. (a) 4-neighborhood system. (b) 8-neighborhood system. (c) One-element clique (d) Two-element cliques. (e) Three-element cliques.

B. Potential Functions with Different Cliques

1) Potential Functions with One-element Cliques

In MRF, a potential function with one-element cliques (also called data term) is usually used to measure the similarity of control points in two images. In the lung 4D CT image registration, due to the compression of the lungs and the local intensity inhomogeneity of the lung 4D CT images, some traditional similarity metrics, such as Sum of Absolute Differences (SAD) and Sum of Squared Differences (SSD), are not suitable for registration. To solve the above problems, we evaluated the following three image metrics [1]: SAD, Local Correlation Coefficient (LCC) [33] and MIND. The evaluation by using the above three metrics shows that the MIND metric is more suitable for the registration framework proposed in this paper

$$\mathcal{V}_1(x_p) = \frac{1}{|R|} \sum_{r \in R} \left| \mathrm{MIND}(I_p, x_p, r) - \mathrm{MIND}(J_p, x_p, r) \right| \quad (3)$$

where R defines the search region of the modality independent neighborhood descriptor.

2) Potential Functions with Two-element Cliques

In MRF, since the set of labels for each point is defined as a set of the displacement vectors, so labels can represent the displacement vectors. In order to make the deformation field satisfy a certain smoothness, the potential functions with two-element cliques are usually used to apply the smoothing constraint to the deformation field. In the lung image, the deformation field is not necessarily continuous because of the respiratory motion, so the following truncation model in [5] is used in this paper

$$\mathcal{V}_2(\mathbf{d}_p, \mathbf{d}_q) = \min(r \cdot \left\| \mathbf{d}_p - \mathbf{d}_q \right\|_2, n) \quad (4)$$

where $\mathbf{d}_p$, $\mathbf{d}_q$ represent the labels of point p and point q respectively. $\|\bullet\|_2$ represent the vector ℓ_2 - norm, n is the

[1] We provide detailed evaluations of various similarity metrics in the online *Supplementary Materials.*

maximum cost of the potential function $\mathcal{V}_2$ and r is a linear increasing coefficient.

3) Potential Functions with High-Order Cliques

Since the high-order cliques can describe the complex relationship between the labels, it can not only constrain the smoothness, but also maintain the topology of the deformation field, so that the deformation can be better constrained. The purpose of lung 4D CT image registration is to obtain the deformation field of the lung respiratory movement and model the respiratory motion according to the deformation field. Therefore, by introducing a topology preservation term with high-order cliques in the energy function, not only the accurate estimation of the deformation field can be achieved, but also the accuracy of the registration can be improved. In 2D images, the topology preservation term is built according to three-elements cliques, as shown in Fig. 1(e). Note that the 4 three-element cliques defined in Fig. 1(e) are analogous to the four corners from which the Jacobian is computed in [30]. Therefore, we can get the Jacobian determinants[2] according to the different forms of the three-element cliques in Fig. 1(e)

$$
\begin{aligned}
J^{ff}\left(\mathbf{d}_{p_1},\mathbf{d}_{p_2},\mathbf{d}_{p_3}\right)&=\left(1+\mathbf{d}_{p_{2x}}-\mathbf{d}_{p_{1x}}\right)\left(1+\mathbf{d}_{p_{3y}}-\mathbf{d}_{p_{1y}}\right)-\left(\mathbf{d}_{p_{3x}}-\mathbf{d}_{p_{1x}}\right)\left(\mathbf{d}_{p_{2y}}-\mathbf{d}_{p_{1y}}\right)\\
J^{bf}\left(\mathbf{d}_{p_1},\mathbf{d}_{p_2},\mathbf{d}_{p_3}\right)&=\left(1+\mathbf{d}_{p_{1x}}-\mathbf{d}_{p_{2x}}\right)\left(1+\mathbf{d}_{p_{3y}}-\mathbf{d}_{p_{1y}}\right)-\left(\mathbf{d}_{p_{3x}}-\mathbf{d}_{p_{1x}}\right)\left(\mathbf{d}_{p_{1y}}-\mathbf{d}_{p_{2y}}\right)\\
J^{fb}\left(\mathbf{d}_{p_1},\mathbf{d}_{p_2},\mathbf{d}_{p_3}\right)&=\left(1+\mathbf{d}_{p_{2x}}-\mathbf{d}_{p_{1x}}\right)\left(1+\mathbf{d}_{p_{1y}}-\mathbf{d}_{p_{3y}}\right)-\left(\mathbf{d}_{p_{1x}}-\mathbf{d}_{p_{3x}}\right)\left(\mathbf{d}_{p_{2y}}-\mathbf{d}_{p_{1y}}\right)\\
J^{bb}\left(\mathbf{d}_{p_1},\mathbf{d}_{p_2},\mathbf{d}_{p_3}\right)&=\left(1+\mathbf{d}_{p_{1x}}-\mathbf{d}_{p_{2x}}\right)\left(1+\mathbf{d}_{p_{1y}}-\mathbf{d}_{p_{3y}}\right)-\left(\mathbf{d}_{p_{1x}}-\mathbf{d}_{p_{3x}}\right)\left(\mathbf{d}_{p_{1y}}-\mathbf{d}_{p_{2y}}\right)
\end{aligned}
\tag{5}
$$

where, in the isosceles right-angled triangle (the length of the right angle is the unit distance) formed by the clique elements p_1, p_2, p_3 (as shown in Fig. 1(e)), $\mathbf{d}_{p_1}, \mathbf{d}_{p_2}, \mathbf{d}_{p_3} \in \mathbf{D}$ respectively represents the displacement vector of the right-angled vertex, the displacement vector of the vertical right-angled vertex, and the displacement vector of the horizontal right-angled vertex, Equation (5) is the Jacobian determinant of the invariance of the constrained topology we need. According to (5), the topological constraint requires four vertices, and the values of the Jacobian determinants corresponding to the four vertices are all greater than or equal to 0.

According to the characteristics of MCMC algorithm, if we use the step function in [15] to impose a penalty on the topology preservation term, it will produce an unreasonable solution. In addition, the step function in [15] has serious problems that cannot distinguish the two topological preservation forms in Fig. 2. It can be easily seen from Fig. 2 that Fig. 2(b) and Fig. 2(c) obtained by two kinds of deformations of the grid composed of four-element cliques in Fig. 2(a) respectively satisfy the topological invariance. However, the values of the Jacobian determinants of Fig. 2 (b) and Fig. 2 (c) which have different deformations are actually different. If we use the step function as the potential function expression without distinction, then there is no difference in the value of the Jacobian determinants that distinguishes between different deformations, so it will be easy to give a wrong solution. Therefore, according to the characteristics of the MCMC algorithm, in order to avoid the excessive cost of the topology preservation term composed of high-order cliques, we propose to use logarithm function to impose penalties on large deformations. Therefore, for 2D image, the topology preservation term can be written as follows:

$$
\mathcal{V}_3\left(\mathbf{d}_{p_1},\mathbf{d}_{p_2},\mathbf{d}_{p_3}\right)=\begin{cases}s\cdot\log\left(J\left(\mathbf{d}_{p_1},\mathbf{d}_{p_2},\mathbf{d}_{p_3}\right)+1\right), & J\left(\mathbf{d}_{p_1},\mathbf{d}_{p_2},\mathbf{d}_{p_3}\right)\geq 0,\\ m, otherwise.\end{cases}
\tag{6}
$$

where s is a linear increasing coefficient, m is the cost value when the topology is not preserved and satisfies $\max(k\cdot\log(J(\mathbf{d}_{p_1},\ \mathbf{d}_{p_2},\ \mathbf{d}_{p_3})+1)) \ll m < 1$.

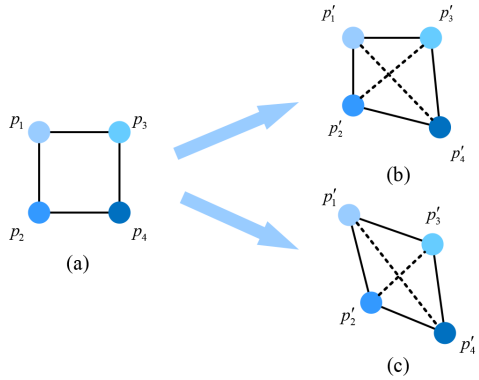


Fig. 2. Two forms of Topology preservation in 2D image. (a) Grid without deformation. (b) After a small deformation of the grid. (c) After a large deformation of the grid.

For 3D images, the corresponding topological constraint is mainly composed of an eight-element clique shown in Fig. 3(a). Calculating the eight-element clique constraint requires at least 64 Jacobian determinants, which imposes a huge computation burden in practice. Therefore, according to [34], we constrain the topology by using 8 Jacobian matrices corresponding to the eight four-element cliques as shown in Fig. 3(b). In this way, not only the topology in the 3D image can be effectively maintained, but also the amount of calculation can be effectively reduced. The derivation of its potential function is similar to that of a 2D image. Therefore, we can express a potential function with four-element cliques as follows:

$$
\begin{aligned}
&\mathcal{V}_4\left(\mathbf{d}_{p_1},\mathbf{d}_{p_2},\mathbf{d}_{p_3},\mathbf{d}_{p_4}\right)\\
&=\begin{cases}s\cdot\log\left(J\left(\mathbf{d}_{p_1},\mathbf{d}_{p_2},\mathbf{d}_{p_3},\mathbf{d}_{p_4}\right)+1\right), & J\left(\mathbf{d}_{p_1},\mathbf{d}_{p_2},\mathbf{d}_{p_3},\mathbf{d}_{p_4}\right)\geq 0,\\ m, otherwise.\end{cases}
\end{aligned}
\tag{7}
$$

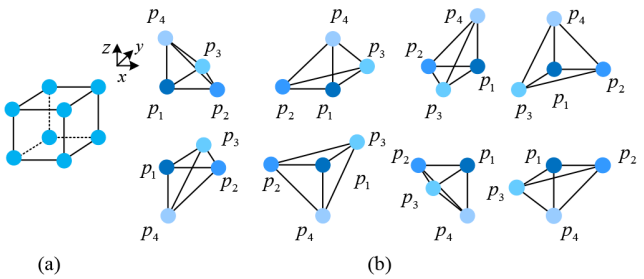


Fig. 3. Four-element cliques and eight-element clique in 3D image. (a) Eight-element clique. (b) Four-element cliques.

C. High-Order MRF Registration Model

In high-order MRF registration model, the energy function is usually composed of three-element cliques or higher-order

[2] We provide detailed derivation of Jacobian determinants for 2D and 3D images in the online *Supplementary Materials*.

cliques. According to the above analysis of different potential functions, in order to ensure that the deformation field can maintain the topology while having a certain smoothness, in the energy function, it is necessary to add not only a smoothing constraint term corresponding to two-element cliques potential functions, but also it is necessary to add a topology preservation term with three-element cliques potential functions or the four-element cliques potential functions. Therefore, for 2D image registration, an energy function with the highest clique is a three-element clique can be expressed as follows:

$$E = \sum_{p\in\mathcal{C}_1} \mathcal{V}_1(x_p) + \sum_{p,q\in\mathcal{C}_2} \mathcal{V}_2(d_p, d_q) + \sum_{p,q,r\in\mathcal{C}_3} \mathcal{V}_3(d_p, d_q, d_r) \quad (8)$$

In the above equation, $\mathcal{C}_i$ represents a set of all *i*-element cliques, $\mathcal{V}_i$ is the potential function corresponding to *i*-element cliques. Correspondingly, in the 3D image space, an energy function with the highest clique is a four-element clique can be expressed as follows:

$$E = \sum_{p\in\mathcal{C}_1} \mathcal{V}_1(x_p) + \sum_{p,q\in\mathcal{C}_2} \mathcal{V}_2(d_p, d_q) + \sum_{p,q,r,s\in\mathcal{C}_4} \mathcal{V}_4(d_p, d_q, d_r, d_s) \quad (9)$$

III. MRF Optimization Algorithm and Parameters Selection

A. MCMC Optimization Algorithm

In the MRF-based image registration framework, different energy models often need to be optimized by the corresponding optimization algorithm. According to [1], the discrete optimization algorithms based on MRF can be divided into the following three categories: Message passing method [6], [7], graph cuts method [3]-[5] and method based on Monte Carlo simulation [35]. Usually, in the low-order MRF, although the message passing method and the graph cuts method have a faster convergence speed and more accurate solution than MCMC-based method, but they cannot handle the optimization problem of the energy function with high-order cliques as (8) and (9). According to the MCMC algorithm, it only needs to update the labels of the current optimized control point, and does not need to update that of other points at the same time, so it can be used to optimize the energy functions in the form of high-order cliques in equation (8) and equation (9).

B. Parameters Selection

In the MCMC algorithm, the selection of parameters for each term in the energy function must satisfy a certain criterion. Choosing inappropriate parameters may result unsatisfactory results. In the process of optimization by using MCMC algorithm, the data term consisting of one-element cliques is the main deformation driving factor, however, the regular term with two-element cliques and other high-order cliques are the secondary deformation driving factor. Therefore, in the parameter selection, the cost value of the regular term does not exceed the cost value of the data term. According to this criterion, it is more appropriate to select 0.3 for the smoothing parameter *n* and select 0.5 for the topology preservation parameter *m*.

C. Labels Selection

In image registration based on the MRF model, the sampling method of the deformation field also has a great influence on the registration result. According to [26], the sampling method can be divided into two types: dense sampling and sparse sampling. Dense sampling refers to densely selecting displacement vectors in all dimensions. Unlike dense sampling, sparse sampling only selects the displacement vectors along the direction of the coordinate axis, therefore, it has the advantage of reducing the complexity of the optimization problem and significantly improving the efficiency of the algorithm. However, since the displacement vector is only selected along the coordinate axes, it is inevitable that the optimal solution will be missed, which in turn will result in a decrease in registration accuracy. Therefore, in practice, the selection of sampling methods needs to make a compromise between registration efficiency and registration accuracy.

IV. Model Analysis

A. Experimental Analysis of Topology Preservation in 2D Deformation Field

We use the following two 2D images to discuss the role of the potential functions with one-element cliques, two-element cliques and three-element cliques corresponding to the data term, smooth term and topology preservation term in image registration respectively. The main concern here is the effect of the potential functions with high-order cliques on the topological structure of the image deformation field.

In our experiments[3], we select Fig. 4(a) as the moving image, Fig. 4(b) is the target image. The size of both images is 128×128, the smoothing parameter *n* and the topology preservation parameter *m* are set to 0.3 and 0.5 respectively. First, we only rely on the similarity metric (data term) to register the moving and target image. After registration the deformation field is distorted and folded (Fig. 4(c)). Fig. 4(d) is the deformation field after registration by using the data term and the smoothing term, the smoothness of the deformation field is significantly improved compared to the deformation field after registration using only the data term. However, the partial enlarged image (Fig. 4(e)) of Fig. 4(d) shows that the deformation field of the image is still folded, and the topology cannot be maintained. Fig. 4(f) is the deformation field after adding the topology preservation term in addition to the data term and the smoothing term in the energy function. Comparing Fig. 4(e) with Fig. 4(g), the folding phenomenon of the deformation field after the application of the potential functions with three-element cliques disappears obviously, which proves that the potential functions with three-element cliques are introduced into the energy function as the topology preservation term can maintain the topology of the deformation effectively.

B. Experimental Analysis of Topology Preservation in 3D Deformation Field

In order to explain the role of the potential functions with high-order cliques of the energy functions in the topological

[3]The related codes are provided in https://github.com/4DCTreg/HOMRF

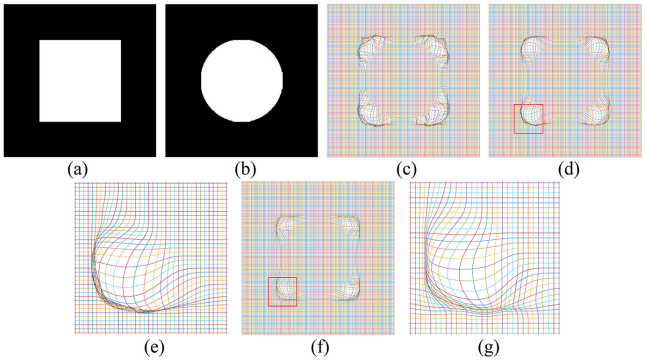


Fig. 4. Experimental results of the topology preservation for 2D deformation field. (a) Moving image. (b) Target image. (c) The deformation field after registration only based on data term. (d) The deformation field after registration based on data term and smoothing term. (e) Partial enlarged image of the red box area in Fig. 4(d). (f) The deformation field after registration based on data term, smoothing term and topology preservation term. (g) Partial enlarged image of the red box area in Fig. 4(f).

invariance of the 3D image deformation field more detailed, we use the following 3D model image experiment to further verify. The size of the moving (Fig. 5(a)) and the target image (Fig. 5(b)) are both $33\times33\times33$, The radius of the dark blue sphere target in the moving image is 10 and the size of the dark blue cube target in the target image is $21\times21\times21$. In both images, the intensity value of the target body is 1 and the intensity value of the external space is 0. In order to facilitate the observation of the deformation field in the 3D image, we select the deformation fields of the 7th to 9th slice in the data volume for continuous observation (as shown in Fig. 5(c) to Fig. 5(e)).

Without any constraint, the deformation field (see Fig. 5(c)) loses its basic smoothness and appears to be severely distorted. Although the deformation field (Fig. 5(d)) obtained by applying only the potential function with two-element cliques as a smooth term in the energy function has a certain smoothness, it still has a folding phenomenon, which is insufficient to maintain the topology of the deformation field. Fig. 5(e) is the deformation field obtained by reintroducing the potential function with four-element cliques in (9) as a topology preservation term in the energy function, which is compared with Fig. 5(c) and Fig. 5(d) respectively. It is significantly improved in both smoothness and topology maintenance.

V. Experiments

A. Multi-Level Processing Strategy

Since the lung 4D-CT image has high resolution and large range of local anatomical features, so it is difficult to directly register on the original resolution. To solve this problem, a multi-resolution processing strategy [27] or multi-level processing strategy [5] are used to reduce the computational complexity. The multi-level processing strategy used in this paper is actually a multi-resolution strategy in the time domain/space domain and Fig. 6 shows the detailed flow of the multi-level strategy.

In the multi-level processing strategy, we can set the grid spacing as $\lfloor 1/k^l \rfloor$ in each level, where l is the current level, $k \in (0 \sim 1]$ is the sampling coefficient of the grid, when

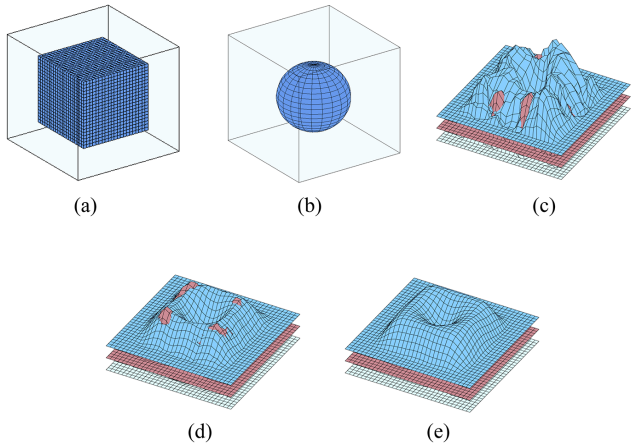


Fig. 5. Experimental results of the topology preservation for 3D deformation field (a) Moving image. (b) Target image. (c) The 7th-9th slice deformation field after registration only based on data term. (d) The 7th-9th slice deformation field after registration based on data term and smoothing term. (e) The 7th-9th slice deformation field after registration based on data term, smoothing term and topology preservation term.

$k=1$, the current grid spacing is equal to one pixel spacing. The sampling coefficient can be determined according to the actual size and characteristics of the image in the experiment. It should be noted that in the division of the grid map, the grids should be uniformly distributed and avoid the overlap of the pixels between the grids. In addition, the selection of the up-sampling method for the deformation field needs to be selected according to the image features, here, we use a linear interpolation method.

B. 4D CT DIR Dataset

In this section, we evaluated our proposed HO-MRF method using DIR-lab dataset [4] that represents three-dimensional abdominal time series in respiratory motion. The DIR-lab dataset consists of 10 different sequences labeled 4D CT1-4D CT10, each of which is further divided into ten respiratory phase sequences from T00 to T90, where the maximum inspiratory phase (T00) and maximum the expiratory phase (T50) provides 300 expert landmarks. The average voxel resolution of the dataset is $1\times1\times2.5mm^3$. In our experiment, we select the image corresponding to the maximum inspiratory phase (T00) of each case as the moving image, the image corresponding to the maximum expiratory phase (T50) as the target image for registration.

In order to select a suitable multi-level processing strategy, we calculate the average TRE of 10 cases in the 4D CT dataset and COPD dataset under different levels. As can be seen from Fig. 7, it is optimal to divide the images in the 4D CT dataset and COPD dataset into 5 levels and 6 levels respectively. The sampling coefficient k of the grid is set to 0.6, the smoothing term parameter n and the topological preservation term parameter m in the energy function are set to 0.3 and 0.5 respectively. For improving the registration accuracy, the displacement vector is densely sampled to form a label set. The dense sampling range is defined to be $\mathbf{L}=\{0,\pm2,\cdots\pm12\}^3$ voxels ($|\mathbf{L}|=2197$) for the

[4] https://www.dir-lab.com/index.html

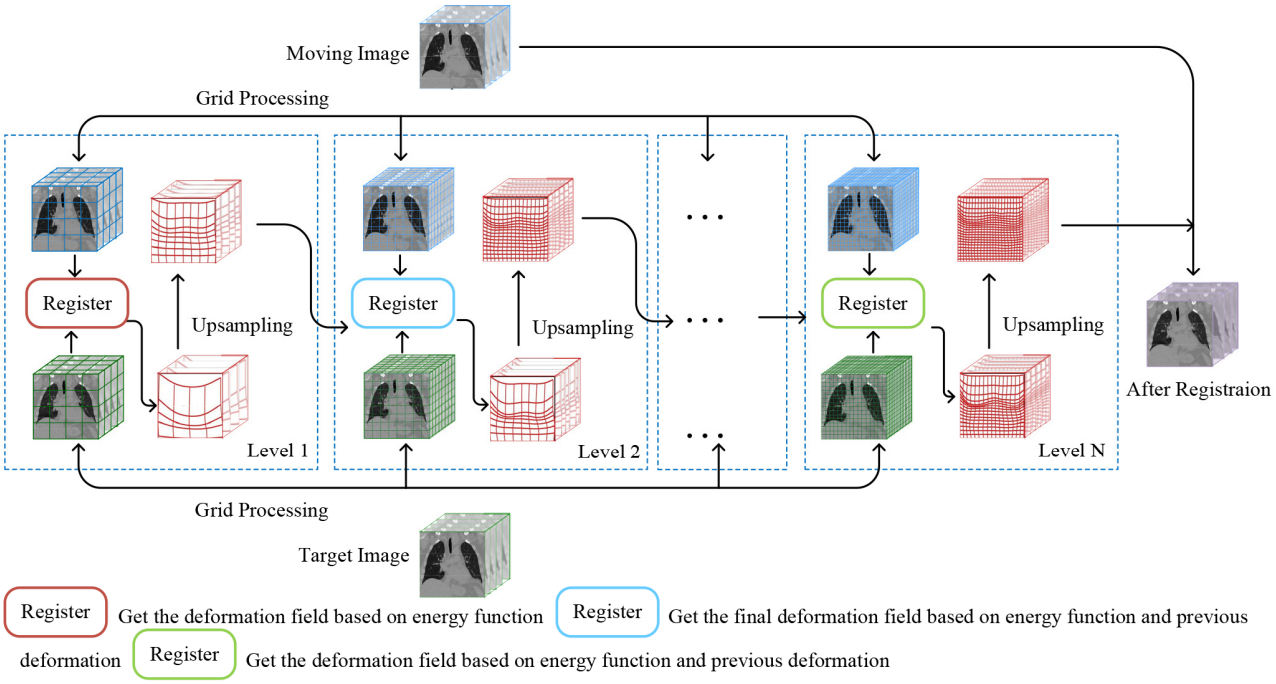


Fig. 6. The basic framework for multi-level processing strategy

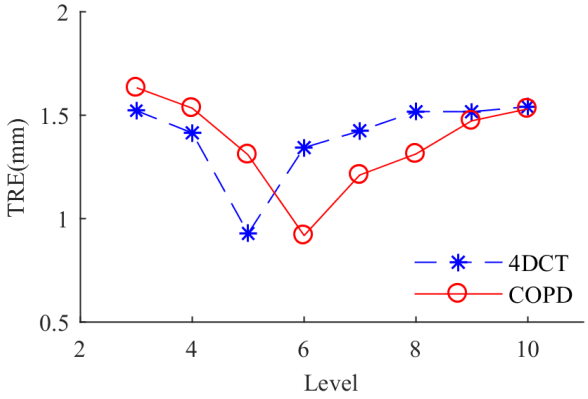


Fig. 7. Sensitivity of TRE to levels of HO-MRF method in 4D-CT and COPD dataset.

coarsest level (for datasets (1-5), with smaller deformation, the sampling range is decreased to $\mathbf{L}_{max} = 8$). At each subsequent level, the rang is halved.

Fig. 8 shows the lung CT coronal overlay image of the proposed HO-MRF method and the isopTV method for the fourth case in the 4D CT dataset. The overlay image is obtained by superimposing the T00 phase image or the registered T00 phase image and the T50 phase image. In the overlay image, the red portion and the green portion respectively indicate image under registration and image over registration. If the difference between the two images is small, the overlay image is dominated by gray. Otherwise, more red or green portions appear in the image. As shown in Fig. 8(a), there is a great difference between T00 phase image and the T50 phase image, in particular, there is a large displacement in the lower part of the lung, and there are many red regions in the figure, indicating severe under registration. Compared with Fig. 8(a), the overlay images of Fig. 8(b) and Fig. 8(c) obtained by HO-MRF method and isopTV method respectively are obviously improved, but there are still over registration and under registration. Comparing the HO-MRF method with the isoPTV method, the

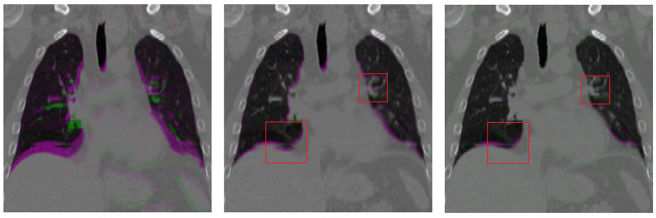
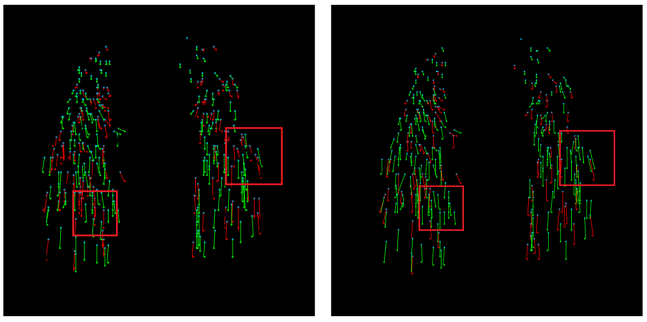


Fig. 8. Coronal overlay images for case 4 of 4D CT dataset. (a) Non-Registered overlay image. (b) isopTV overlay image. (c) HO-MRF overlay image.

under-registered red region and the over-registered green region in Fig. 8(c) are significantly less than those in Fig. 8(b) respectively. For obvious areas, see the corresponding red frame area marked in the Fig. 8.

In order to show the validity of the registration more intuitively, we take case 4 in the DIR-lab dataset as an example, and give the schematic diagrams of registration vector displacement error of the isopTV method (Fig. 9(a)) and the proposed HO-MRF method (Fig. 9(b)) respectively. The blue dots in the figure indicate the landmarks in the T50 phase, the red dots and the green dots represent the landmarks in the T00 phase, and the lines connecting the registered points and the blue points are used as the displacement vectors of the landmarks. The range of the registration displacement vector error is marked by different colors, wherein the green point or the green line represents a registration displacement vector error of less than 1.5 mm between the registered point and the corresponding point in the T00 phase, and a red dot or a red line indicates that the registration displacement vector error is greater than 1.5 mm. If only a blue dot appears in the figure, the landmark has no displacement. Fig. 9 shows that for these two methods, the deformation of the registered landmarks basically maintains a smooth consistency, and the displacement vectors of the landmarks in the red squares marked in the two images are compared, the proposed HO-MRF method has better



(a) (b)

Fig. 9. The schematic diagrams of registration vector displacement errors for case 4 of 4D CT dataset. (a) isopTV method. (b) HO-MRF method.

registration accuracy in the right lung than the isopTV method.

We also use the TRE indicator to further evaluate the effect of image registration. The TRE is defined as the Euclidean distance between the coordinate of the expert landmarks and the coordinate of the landmarks registered by the deformation field transformation. The smaller the distance, the better the registration effect of the method. Table I lists average TREs for various methods and percentage of negative Jacobian determinants for proposed methods in the DIR-lab dataset. In order to comprehensively analyze the performance of the proposed method, several essential methods published in the DIR-lab dataset are used for comparative analysis. In addition, average TREs and percentage of negative Jacobian determinants of the proposed method with high-order cliques (HO-MRF) and without high-order cliques (LO-MRF) are listed for further analysis. Among all methods in Table I, the proposed method (HO-MRF 0.95mm) is inferior to the optimal pTVreg [27] method (0.92mm) and sub-optimal NGF [21] method (0.94mm). The MRF-based method (deeds [11]) can be used as a baseline to measure the validity of the registration. By comparison in DIR-lab dataset, the registration results of the proposed HO-MRF method is better than deeds method to except for the case1. Compared with the optimal pTVreg method, among the results of the proposed HO-MRF method, one case is equivalent and the two cases are superior. Compared with the suboptimal NGF method, the proposed HO-MRF method is superior to the NGF method in three cases. In average TREs of all methods for 10 cases, the average TRE of the HO-MRF method is 0.95mm, which is merely 0.03mm worse than the best method (pTVreg) and 0.01mm worse than the sub-optimal mask-free method (NGF). Comparing the HO-MRF method with the LO-MRF method, the introduction of high-order cliques in energy function can reduce the number of negative Jacobian determinants and improve the registration accuracy effectively.

C. 3D Breath-Hold CT DIR COPD Dataset

Compared to the 4D CT DIR-lab dataset, the COPD dataset has a larger amplitude of motion (23.46 mm) and a larger data volume image. This dataset contains 10 cases labeled COPD1~COPD10, and each case has 300 landmarks in both the expiratory phase and the inspiratory phase. The average resolution of the dataset is $0.6\times0.6\times2.5mm^3$, and the average image size is $512\times512\times120$ voxels. Similarity, we select the maximum inspiratory phase of each case as the moving image (T00), and the maximum expiratory phase is registered as the target image (T50).

In order to solve the problem that it takes too long to directly register large-scale images in the COPD dataset, the 3D Scale Invariant Feature Transform (SIFT) descriptor proposed in [41] is used to perform feature-level coarse registration, the purpose of which is to extract only the deformation field of large deformation. Table II shows that after SIFT pre-registration, the TRE of each case in COPD dataset decreases significantly, but the results (7.03mm) are still unsatisfactory and need further processing. On this basis, the deformation field extracted by the 3D SIFT descriptor is maximally meshed to obtain a deformed field grid map. Secondly, the average value of the displacement vectors of all the pixels in each grid of the deformed field grid map is calculated as the displacement vector of the grid control point, and the displacement vectors of all grid control points are used as the initial label values of the MCMC optimization algorithm. Finally, we use the same approach as the DIR-lab dataset for post-registration. In the post-registration, the dense sampling range is defined to be $\mathbf{L}=\{0,\pm2,\cdots\pm20\}^3$ voxels ($|\mathbf{L}|=9261$) for the coarsest level, at each subsequent level, the rang is halved. After registration, the average TRE is further

TABLE I

AVERAGE TREs FOR VARIOUS METHODS AND PERCENTAGE OF NEGATIVE JACOBIAN DETERMINANTS FOR PROPOSED METHODS IN 4D CT DATASET

Methods	TREs and percentage of negative Jacobian determinants for 4D CT cases										Mean
	case1	case2	case3	case4	case5	case6	case7	case8	case9	case10	
No regis.	4.01±2.9	4.65±4.1	6.37±4.2	9.42±4.8	7.10±5.1	11.10±7.0	11.59±7.9	15.16±9.1	7.82±4.0	7.63±6.5	8.52
cEPE[36]	0.80±0.9	0.77±0.9	0.92±1.1	1.22±1.2	1.21±1.5	0.90±1.0	0.98±1.0	1.16±1.5	1.00±1.0	0.99±1.3	1.00
NLR[37]	0.77±0.9	0.78±0.9	0.92±1.1	1.27±1.3	1.11±1.5	0.91±1.0	0.86±1.0	1.03±1.2	0.97±0.9	0.87±1.0	0.95
LMP[38]	**0.74±0.9**	0.78±0.9	0.91±1.1	1.24±1.3	1.17±1.5	0.90±1.0	0.87±1.0	1.04±1.2	0.98±1.0	0.89±1.0	0.95
SGM3D[39]	0.76±0.9	0.72±0.9	0.94±1.1	1.24±1.3	1.15±1.4	0.90±1.0	0.89±1.0	1.13±1.4	0.91±0.9	0.83±0.9	0.95
NGF[21]	0.76±0.9	0.79±0.9	0.93±1.1	1.27±1.3	**1.07±1.5**	0.90±1.0	0.85±1.0	1.03±1.2	0.94±0.9	0.86±1.0	0.94
aTV[40]	0.79±0.9	0.74±0.9	0.95±1.1	**1.21±1.2**	1.24±1.5	0.96±1.0	0.97±1.0	1.17±1.5	1.03±1.1	1.00±1.0	1.03
deeds[11]	0.80±0.7	0.86±0.7	1.14±0.8	1.71±1.7	1.77±1.7	1.88±1.4	2.21±2.3	2.78±3.1	1.35±0.8	1.50±1.4	1.60
isopTV[27]	0.76±0.9	0.77±0.9	**0.90±1.1**	1.24±1.3	1.12±1.4	0.85±0.9	0.80±1.3	1.34±1.9	0.92±0.9	0.82±0.9	0.95
pTVreg[27]	0.80±0.9	0.77±0.9	0.92±1.1	1.30±1.3	1.13±1.4	**0.78±0.9**	**0.79±0.9**	**1.00±0.9**	0.91±1.0	**0.82±1.0**	**0.92**
LO-MRF	0.84±1.0	0.74±1.0	1.12±1.1	1.35±1.3	1.28±1.6	1.05±0.9	1.04±1.1	1.23±1.3	1.08±1.2	1.01±1.3	1.07
(P.N.J.D)	0.11%	1.07%	1.08%	0.28%	0.28%	0.20%	2.43%	1.3%	0.87%	1.3%	0.89%
HO-MRF	0.80±0.9	**0.70±0.9**	0.96±1.1	1.28±1.3	1.13±1.5	0.95±0.9	0.94±1.1	1.01±1.3	**0.90±1.1**	0.88±1.3	0.95
(P.N.J.D)	0.10%	0.48%	0.88%	0.25%	0.27%	0.15%	1.01%	1.2%	0.73%	0.92%	0.60%

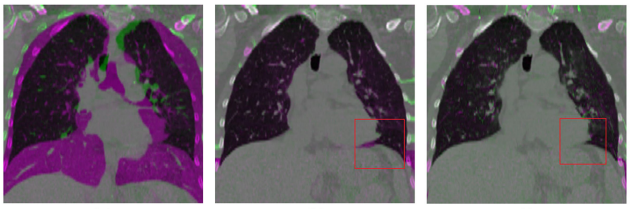
P.N.J.D.=Percentage of Negative Jacobian Determinants

TABLE II
AVERAGE TREs FOR VARIOUS METHODS AND PERCENTAGE OF NEGATIVE JACOBIAN DETERMINANTS FOR PROPOSED METHODS IN COPD DATASET

Methods	TREs and percentage of negative Jacobian determinants for COPD cases										Mean
	case1	case2	case3	case4	case5	case6	case7	case8	case9	case10	
No regis.	25.90±11.6	21.77±6.5	12.29±6.4	30.90±13.5	30.90±14.1	28.32±9.2	21.66±7.7	25.57±13.6	14.84±10.0	22.48±10.7	23.46
NLR[37]	1.33±1.6	2.34±2.9	1.12±1.1	1.54±1.6	1.39±1.4	2.08±3.0	1.10±1.3	1.57±2.1	0.99±1.3	1.42±1.4	1.49
LMP[38]	1.21±1.5	1.97±2.4	1.06±1.0	1.64±1.8	1.46±1.5	1.34±1.7	1.16±1.5	1.54±2.6	0.99±1.3	1.39±1.5	1.38
SGM3D[39]	1.22±2.7	2.48±3.8	1.01±0.9	2.42±3.6	1.93±3.2	1.45±2.4	1.05±1.4	1.16±1.8	0.81±0.7	1.28±1.3	1.48
MILO[42]	0.93±0.9	1.77±1.9	0.99±0.9	1.14±1.0	1.02±1.2	0.99±1.1	1.03±1.1	1.31±1.8	0.86±1.1	1.23±1.3	1.13
MRF[42]	1.00±0.9	1.62±1.8	1.00±1.1	1.08±1.1	0.96±1.1	1.01±1.3	1.05±1.1	1.08±1.2	0.79±0.8	1.18±1.3	1.08
DIS-CO[28]	0.79±0.8	**1.46±2.3**	0.84±0.8	0.74±0.9	0.71±0.8	**0.64±0.8**	0.79±0.9	**0.77±0.9**	**0.62±0.67**	0.86±0.9	**0.82**
isopTV[27]	0.77±0.7	2.22±2.9	0.82±1.5	0.85±0.9	0.77±0.8	0.86±1.9	**0.74±1.1**	0.81±1.8	0.83±1.2	0.92±0.9	0.96
pTVreg[27]	**0.71±0.8**	1.89±3.6	**0.77±0.8**	**0.68±0.7**	**0.71±0.8**	0.66±1.2	0.75±0.9	0.78±1.6	0.63±1.0	**0.85±0.9**	0.84
SIFT[41]	6.58±5.9	8.87±7.4	2.87±2.8	18.26±21.5	6.33±4.5	3.58±3.2	4.25±3.4	5.93±6.7	3.84±4.3	9.79±7.2	7.03
LO-MRF	1.15±1.0	2.23±1.0	1.12±1.1	1.07±1.3	1.05±1.6	0.98±0.9	1.04±1.1	1.05±1.3	0.95±1.2	1.01±1.3	1.17
(P.N.J.D)	0.58%	3.07%	1.08%	2.59%	0.15%	1.03%	1.45%	0.56%	0.86%	2.37%	1.37%
HO-MRF	0.92±0.9	1.90±3.8	0.83±1.3	0.80±0.9	0.88±0.9	0.87±1.5	0.83±1.0	0.81±1.7	0.80±1.1	0.90±0.9	0.95
(P.N.J.D)	0.46%	2.45%	0.68%	1.25%	0.09%	0.47%	1.09%	0.49%	0.63%	1.25%	0.89%

reduced (1.17mm for LO-MRF method, 0.95 mm for HO-MRF method), which indicates that SIFT pre-registration can effectively process large-scale images in COPD datasets.

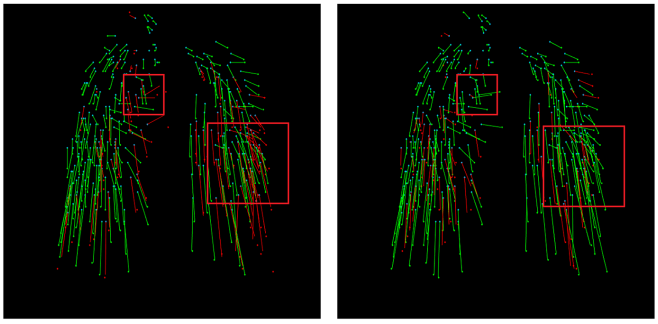
In this paper, considering the comparability between methods, the best method for the extraction process without lung parenchyma is selected for comparison. For this reason, only the isopTV algorithm [27] is selected, and the DIS-CO method [28] is not used. Therefore, similarly, we respectively give the overlay image of the isopTV algorithm and the proposed HO-MRF method for the fourth case in the COPD dataset. Fig. 10(a) shows that there is a large deformation when unregistered, and the under-registered area is relatively large. From Fig. 10(b) and Fig. 10(c), except for the obvious under registration (red area) in the lower mark box in Fig. 10(b), there is no large range of under registration and over registration area, but the background color of Fig. 10(b) is light red, which indicates that the isopTV algorithm has under registration problem. Correspondingly, the background color of Fig. 10(c) is light green, which indicates the proposed method has some over-registration. In general, the overall morphology of the lung image after registration is similar, and the overlap between the ribs and the lung parenchyma of the black background is significantly increased.



(a) (b) (c)

Fig. 10. Coronal overlay images for case 4 of COPD dataset. (a) Non-Registered overlay image. (b) isopTV overlay image. (c) HO-MRF overlay image.

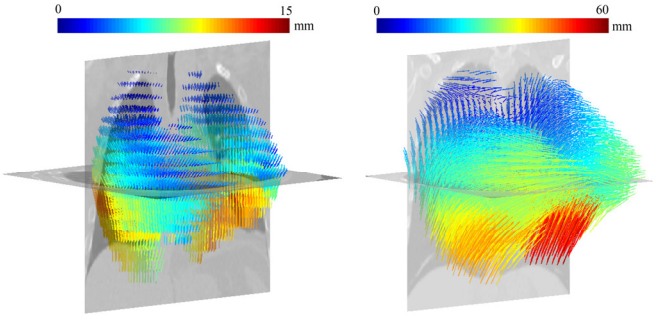
Similarity, we take the case 4 in the COPD dataset as an example to give schematic diagrams of the registration vector displacement errors of the isopTV method (Fig. 11(a)) and the proposed method (Fig. 11(b)). The representation in the Fig.11 is similar to Fig. 9. Fig. 11 shows that although the case has large deformation characteristics, the deformation of the landmarks registered by the two methods maintains a good smooth consistency. The green vector lines in the red frame marked in the Fig.11(b) are more than the red vector lines, which means that the proposed method is better than isopTV method in the registration accuracy.



(a) (b)

Fig. 11. The schematic diagrams of registration vector displacement errors for case 4 of COPD dataset. (a) isopTV method. (b) HO-MRF method.

In addition, displacement fields of case 4 estimated by HO-MRF method for the DIR-lab dataset and COPD dataset are provided in Fig.12(a) and Fig.12(b) respectively. Fig.12 shows that the displacement at the lower lobe of lung is significantly greater than the displacement at the upper lobe of lung, and the COPD dataset has a larger displacement than the DIR-lab dataset.



(a) (b)

Fig. 12. Displacement fields of case 4 in 4D CT dataset and COPD dataset. (a) 3D vector field visualization of case 4 in 4D CT dataset. (b) 3D vector field visualization of case 4 in COPD dataset.

In order to comprehensively analyze the performance of the proposed method, the TRE indicators of several current optimal algorithms in the COPD data set are given in Table II. In the average TREs of all methods for 10 cases, the proposed method (0.95mm) is inferior to the optimal DIS-CO [28] method

(0.82mm) and the sub-optimal pTVreg method (0.84mm). Compared with the same type of MRF method, the proposed method is superior to the MRF method except for case 2. Similarly, by comparing the proposed LO-MRF method and the proposed HO-MRF method, the HO-MRF method has less negative Jacobian determinants and higher registration accuracy than LO-MRF method, which means that the introduction of high-order cliques into energy function can effectively maintain the topological structure of the deformation field and improve the registration accuracy.

VI. Discussions and Conclusions

The image registration based on MRF is an effective non-rigid method, and image registration based on high-order MRF is a very promising and challenging research work. The key question is how to build an energy function that describes large deformations based on high-order cliques. For this reason, according to the characteristics of the deformation field that the topological structure needs to be kept constant in the image registration, the energy functions with high-order form of the topological structure of the deformation field are constructed for 2D and 3D images respectively. By comparing and analyzing the topological structure of the 2D and 3D image model deformation fields, we can find that compared with the low-order MRF registration algorithm with only smooth constraint in the energy function, although the energy function designed in this paper is more complicated, its ability to maintain the topological structure of the deformation field is also stronger, indicating that the design of the energy function with high-order cliques is worthy of further study.

In lung 4D CT image, due to the influence of heart beats and respiratory movements, large motion of small features and sliding motions between organs are caused [19], [27]. So, it is difficult to estimate the non-smooth motion fields. In this case, the deformation field obtained by commonly-used registration methods with smooth term will be distorted and folded, thus affecting the registration accuracy. In order to accurately estimate the sliding motions between organs, we adopt truncation model for potential functions with two-element cliques, but the registration results are not satisfactory (mean TREs of 1.07 mm for DIR-lab dataset, 1.17 mm for COPD dataset). On this basis, we introduce high-order cliques in energy function to prevent the folding phenomenon of the deformation field. According to our experiments, we can find that introducing high-order cliques can improve the registration accuracy (mean TREs of 0.95 mm for DIR-lab and COPD dataset respectively) and reduce the number of negative Jacobian determinants effectively.

Compared with the existing popular discrete optimization algorithm based on MRF, the traditional MCMC algorithm can solve the energy function with high-order clique form designed in the paper effectively. However, the smoothing term parameter in the energy function needs to be artificially set, this way will cause subjective human errors. Therefore, in future research, we can use a parameter estimation method to select the appropriate parameter value. At the same time, since the MCMC algorithm used in this paper is optimized based on stochastic sampling. Compared to the spatial complexity $O(P\times\log_2 L)$ of the graph cuts algorithm, the MCMC algorithm needs at least $O(P\times L)$ bits to store the result, where p is the number of the points and L is the number of labels for each point. In terms of time complexity, the time complexity of MCMC algorithm and graph cuts algorithm can be written as $O(P\times niter)$ and $O(P\times L)$, where $niter \gg L$. In the case of using the same set of labels, the MCMC algorithm needs to have higher space complexity and time complexity than other algorithms (such as graph cuts and FastPD). In [44], in order to estimate the kinetic parameters in MBF (Myocardial Blood Flow), by comparing MCMC algorithm with VB (Variational Bays) algorithm, Saillant *et al.* [41] gave a VB algorithm that can effectively improve the convergence rate. In subsequent studies, we can use the method further enhance the convergence of the designed method.

In this paper, we apply the multi-level processing strategy to 4D CT lung image registration very well, and express the multi-level processing flow with a clear and concise frame diagram. Parallel computing can be used to reduce the computation time when calculating the cost values of the data terms of the grid control points. However, since the cost values are calculated by using a simple averaging method for the pixels in a grid, the differences of the pixels in the grid are not considered. In the further research, we can consider using the B-spline strategy proposed in [45] to improve the accuracy of registration.

Considering that the DIR-lab dataset provides enough landmarks only in the two phases of the maximum expiratory phase and the maximum inspiratory phase, this paper only uses the images corresponding to the two phases in the 4D CT images for registration. In order to fully utilize the other phase information in the 4D CT images to continuously correct the registration process, in the future work, dynamic compensation of each phase image sequence can be used to estimate the deformation field of the whole lung respiratory motion, and further improve the registration accuracy of the 4D CT images. In view of the large amount of data, complex structure and large deformation of the lung 4D CT images, in the future research, the method proposed in [34] can be used to first preprocess the images by lung parenchymal extraction and region segmentation, and coarse registration and sub-regional registration are performed separately.

References

[1] J. H. Kappes, B. Andres, F. A. Hamprecht, C. Schnorr, S. Nowozin, D. Batra, S. Kim, B. X. Kausler, T. Kroger, J. Lellmann, N. Komodakis, B. Savchynskyy, and C. Rother, "A Comparative Study of Modern Inference Techniques for Structured Discrete Energy Minimization Problems," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR),* 2015, pp. 1328–1335.

[2] S. Z. Li, "*Markov Random Field Modeling in Image Analysis,*" Springer-Verlag, 2001.

[3] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell,* vol. 23, no. 11, pp. 1222–1239, Nov. 2001.

[4] D. Freedman and P. Drineas, "Energy minimization via graph cuts: Settling what is possible," *in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2005, pp. 939–946.

[5] N. Komodakis, and G. Tziritas, "Approximate labeling via graph cuts based on linear programming," *IEEE Trans. Pattern Anal. Mach Intell,* vol. 29, no. 8, pp. 1436–1453, Aug. 2007.

[6] M. J. Wainwright, T. S. Jaakkola, and A. S. Willsky, "MAP estimation via agreement on trees: message-passing and linear programming," *IEEE Trans. Inform. Theory,* vol. 51, no. 11, pp. 3697–3717, Nov. 2005.

[7] V. Kolmogorov, "Convergent tree-reweighted message passing for energy minimization," *IEEE Trans. Pattern Anal. Mach Intell,* vol. 28, no. 10, pp. 1568–1583, Oct. 2006.

[8] Tang. T W. H, Chung A C S. "Non-rigid image registration using graph-cuts," in *Proc. Int. Conf. Med. Image Com8669put. Comput.-Assist. Intervent,* 2007, pp. 916–924.

[9] Glocker. B, Komodakis. N, Tziritas. G, Navab. N, Paragios N, "Dense image registration through MRFs and efficient linear programming," *Med. Image Anal,* vol. 12, no. 6, pp. 731–741, Dec. 2008.

[10] M. P. Heinrich, M. Jenkinson, M. Brady, J. A. Schnabel, "Globally optimal deformable registration on a minimum spanning tree using dense displacement sampling," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent,* 2012, pp. 115–122.

[11] M. P. Heinrich, M. Jenkinson, M. Brady, and J. A. Schnabel, "MRF-Based Deformable Registration and Ventilation Estimation of Lung CT," *IEEE Trans. Med Imag,* vol. 32, no. 7, pp. 1239–1248, Jul. 2013.

[12] P. Kohli, M. P. Kumar, and P. H. S. Torr, "P3 & beyond: solving energies with higher order cliques," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR),* 2007, pp. 1950–1957.

[13] C. Rother, V. Kolmogorov, V. Lempitsky, M. Szummer, "Optimizing binary MRFs via extended roof duality," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR),*2007, pp. 1784-1791.

[14] L. Cordero-Grande, G. Vegas-Sanchez-Ferrero, P. Casaseca de la Higuera, and C. Alberola-Lopez, "Topology-Preserving Registration: A Solution via Graph Cuts," in *Proc. 14th Int. Workshop Combinatorial Image Anal.,* Madrid, Spain, May 2011, Lect. Notes Comp. Science, pp. 420–431.

[15] L. Cordero-Grande, G. Vegas Sanchez Ferrero, P. Casaseca de la Higuera, and C. Alberola-Lopez, "A Markov Random Field Approach for Topology-Preserving Registration: Application to Object-Based Tomographic Image Interpolation," *IEEE Trans. Image Process,* vol. 21, no. 4, pp. 2047–2061, Apr. 2012.

[16] X. Han, "Feature-constrained Nonlinear Registration of Lung CT Images," In *Proc*. *Med Image Anal Clinic*. 2010, pp.137–146.

[17] H. Rivaz, Z. Karimaghaloo, and D. L. Collins, "Self-similarity weighted mutual information: A new nonrigid image registration metric," *Med. Image. Anal.,* vol. 18, no. 2, pp. 343–358, Feb. 2014.

[18] Y. Zheng, K. Steiner, T. Bauer, J. Yu, D. Shen, and C. Kambhamettu, "Lung nodule growth analysis from 3D CT data with a coupled segmentation and registration framework," in *Proc. IEEE 11th Int. Conf. Comp. Vision*. 2007, pp. 2861–2867

[19] A. Soliman, F. Khalifa, N. Dunlap, B. Wang, M. Abou. El-Ghar, A. El-Baz, "An iso-surfaces based local deformation handling framework of lung tissues," in *Proc 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI).* 2016, pp. 1253–1256.

[20] W. Lugo, and J. Seguel, "A Fast and Accurate Parallel Algorithm for Genome Mapping Assembly Aimed at Massively Parallel Sequencers," in *Proc. ACM Conf. Bioinfor.* 2015, pp. 574–581.

[21] L. Koenig, J. Ruehaak, "A fast and accurate parallel algorithm for non-linear image registration using normalized gradient fields," *in Proc. IEEE Int. Symp. Biomed. Imag. (ISBI)*, 2014, pp. 580–583.

[22] M. P. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F. V. Gleeson, S. M. Brady and J. A. Schnabel, "MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration," *Med. Image. Anal.,* vol. 16, no. 7, pp. 1423–1435, Oct. 2012.

[23] S. Reaungamornrat, T. De Silva, A. Uneri, S. Vogt, G. Kleinszig, A. J. Khanna, J.-P. Wolinsky, J. L. Prince, and J. H. Siewerdsen, "MIND Demons: Symmetric Diffeomorphic Deformable Registration of MR and CT for Image-Guided Spine Surgery," *IEEE Trans. Med Imag.* vol. 35, no. 11, pp. 2413–2424, Nov, 2016.

[24] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Image Registration by Maximization of Combined Mutual Information and Gradient Information," *IEEE Trans. Med Imag* vol. 19, no. 8, pp. 809–814, Aug. 2000.

[25] E. Castillo, R. Castillo, B. White, J. Rojo, and T. Guerrero, "Least median of squares filtering of locally optimal point matches for compressible flow image registration," *Physics In Medicine And Biology,* vol. 57, no. 15, pp. 4827–4843, Aug. 2012.

[26] Z. Cao, E. Dong, Q. Zheng, W. Sun, and Z. Li, "Accurate inverse-consistent symmetric optical flow for 4D CT lung registration," *Biomed. Signal Process. And Control,* vol. 24, pp. 25–33, Feb. 2016.

[27] V. Vishnevskiy, T. Gass, G. Szekely, C. Tanner, and O. Goksel, "Isotropic Total Variation Regularization of Displacements in Parametric Image Registration," *IEEE Trans. Med Imag.,* vol. 36, no. 2, pp. 385–395, Feb. 2017.

[28] J. Ruehaak, T. Polzin, S. Heldmann, I. J. A. Simpson, H. Handels, J. Modersitzki, and M. P. Heinrich, "Estimation of Large Motion in Lung CT by Integrating Regularized Keypoint Correspondences into Dense Deformable Registration," *IEEE Trans. Med Imag.* vol. 36, no. 8, pp. 1746–1757, Aug. 2017.

[29] K. A. J. Eppenhof, and J. P. W. Pluim, "Pulmonary CT Registration through Supervised Learning with Convolutional Neural Networks," *IEEE Trans. Med. Imag*., vol. 38, no. 5, pp 1097–1105, May. 2019.

[30] A. B. Ashraf, S. C. Gavenonis, D. Daye, C. Mies, M. A. Rosen and D. Kontos, "A Multichannel Markov Random Field Framework for Tumor Segmentation with an Application to Classification of Gene Expression-Based Breast Cancer Recurrence Risk," *IEEE Trans. Med Imag,* vol. 32, no. 4, pp. 637–648, Apr. 2013.

[31] A. Shaffie, A. Soliman, L. Fraiwan, M. Ghazal, F. Taher, N. Dunlap, B. Wang, V. van Berke, R. Keynton, A. Elmaghraby, and A. El-Baz, "A Generalized Deep Learning-Based Diagnostic System for Early Diagnosis of Various Types of Pulmonary Nodules," *Technol. Cancer. Res. & Treatment,* vol.17, pp. 1–9, Sep. 2018.

[32] A. Shaffie, A. Soliman, M. Ghazal, F. Taher, N. Dunlap, B. A. Wang, A. Elmaghraby, G. Gimel'farb, A. El-Baz, "A New Framework for Incorporating Appearance and Shape Features of Lung Nodules for Precise Diagnosis of Lung Cancer," in *Pro. 24th IEEE Int. Conf. Imag. Process. (ICIP),* 2017, pp. 1372–1376.

[33] X. Sun, N. P. Pitsianis, and P. Bientinesi, "Fast computation of local correlation coefficients," in *Proc. Advanced Signal Processing Algorithms, Architectures, and Implementations*, 2008.

[34] B. Karacali, and C. Davatzikos, "Estimating topology preserving and smooth displacement fields," *IEEE Trans. Med. Imag.,* vol. 23, no. 7, pp. 868–880, Jul. 2004.

[35] J. Besag, "On the Statistical Analysis of Dirty Pictures," *Royal Statistical Society Series B-Methodological,* vol. 48, no. 3, pp. 259–302, 1986.

[36] S. Hermann, and R. Werner, "High Accuracy Optical Flow for 3D Medical Image Registration Using the Census Cost Function," in *Pro Imag. Video Tech*., 2014, pp. 23–35.

[37] J. Ruehaak, S. Heldmann, T. Kipshagen, and B. Fischer, "Highly Accurate Fast Lung CT Registration," in *Proc. SPIE*, 2013, pp. 86690Y-1–86690Y-9.

[38] T. Polzin, J. Rühaak, R. Werner, J. Strehlow, S. Heldmann, H. Handels, and J. Modersitzki, "Combining Automatic Landmark Detection and Variational Methods for Lung CT Registration", in *Proc. 5th Int. Workshop Pulmonary Imag. Anal*., 2013, pp. 85–96.

[39] S. Hermann, "Evaluation of Scan-Line Optimization for 3D Medical Image Registration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2014, pp. 3073–3080.

[40] V. Vishnevskiy, T. Gass, G. Szekely, and O. Goksel, "Total Variation Regularization of Displacements in Parametric Image Registration," *in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.,* 2014, pp. 211–220.

[41] E. Murray, J. H. Cho, D. Goodwin, T. Ku, J. Swaney, S.-Y. Kim, H. Choi, Y.-G. Park, J.-Y. Park, A. Hubbert, M. McCue, S. Vassallo, N. Bakh, M. P. Frosch, V. J. Wedeen, H. S. Seung, and K. Chung, "Simple, Scalable Proteomic Imaging for High-Dimensional Profiling of Intact Systems," *Cell*, vol. 163, no. 6, pp. 1500–1514, Dec. 2015.

[42] E. Castillo, R. Castillo, J. Martinez, M. Shenoy, and T. Guerrero, "Four-dimensional deformable image registration using trajectory modeling," *Phys. Med. Biol.*, vol. 55, no. 1, pp. 305–327, Jan. 2010.

[43] M. P. Heinrich, H. Handels, and I. J. A. Simpson, "Estimating Large Lung Motion in COPD Patients by Symmetric Regularised Correspondence Fields," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent,* 2015, pp. 338–345.

[44] A. Saillant, I. Armstrong, V. Shah, S. Zuehlsdorff, C. Hayden, J. Declerck, K. Saint, M. Memmott, M. Jenkinson, and M. A. Chappell, "Assessing reliability of myocardial blood flow after motion correction with dynamic PET using a Bayesian framework," *IEEE Trans. Med Imag.*, vol. 38, no. 5, pp. 1216–1226, May. 2019.

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