A Symmetric 4D Registration Algorithm for Respiratory Motion Modeling

Huanhuan Xu and Xin Li

School of Electrical Engineering and Computer Science, Louisiana State University
xinli@lsu.edu

Abstract. We propose an effective 4D image registration algorithm for dynamic volumetric lung images. The registration will construct a deforming 3D model with continuous trajectory and smooth spatial deformation, and the model interpolates the interested region in the 4D (3D+T) CT images. The resultant non-rigid transformation is represented using two 4D B-spline functions, indicating a forward and an inverse 4D parameterization respectively. The registration process solves these two functions by minimizing an objective function that penalizes intensity matching error, feature alignment error, spatial and temporal non-smoothness, and inverse inconsistency. We test our algorithm for respiratory motion estimation on public benchmarks and on clinic lung CT data. The experimental results demonstrate the efficacy of our algorithm.

Keywords: 4D Image Registration; Respiratory Motion Modeling.

1 Introduction

Image registration is important in image-guided radiotherapy management. For example, in lung cancer radiotherapy, it can establish the correspondences among the 4D (sequential volume) CT images. This correspondence can be used to build a motion estimation model that describes the movement and deformation of organs during breathing cycles. We can generally classify the algorithms for dynamic volumetric image registration into two categories.

Pairwise 3D Registration. Given a sequence of volume images, the conventionally popular registration approach is through the *pairwise 3D registration*, which computes a set of mapping functions f_{ij} between image i and image j. These registrations $\{f_{ij}\}$ can be interpolated to obtain a deforming volumetric model. 3D registration algorithms often approximate natural deformation between two shapes through minimizing certain physical deformation energies [1]or geometric smoothness [2,3]. Pairwise 3D registrations have two general limitations. One is its sensitivity to the selection of the reference frame, especially when describing a motion sequence undergoing large deformations. More importantly, the second limitation of 3D piecewise registration is its lack of control on the smoothness of the resultant motion trajectory. High quality 3D image registration may provide accurate inter-image matching, but in the temporal

K. Mori et al. (Eds.): MICCAI 2013, Part II, LNCS 8150, pp. 149–156, 2013.

dimension, the motion/deformation composed from the individually computed 3D matching is often not smooth and thus less physically natural.

4D Registration. Registration across sequential images can be solved in a 4D space directly. This can avoid the bias caused by the selection of a predetermined reference frame, and can directly enforce both spatial and temporal smoothness of the transformation to indicate more physically natural deformations. This approach, called the *4D Registration* [2,4,5], attracts a lot of attentions recently. Metz et al. [2] solved a 4D registration by reducing intensity matching errors on a common domain; their algorithm minimizes a non-linear and non-smooth optimization with many local minima, which requires a good initial guess to get desirable matching. Geometric information such as feature correspondence can guide the optimization to avoid many undesirable local minima. Wu et al. [4] suggested a 4D registration framework utilizing both image intensity and feature guidance, and solve the registration on a refined implicit domain for lung image data. They also proposed a groupwise registration scheme [6] by iteratively resolving feature correspondence and thin-plate spline deformation, which demonstrates high matching accuracy for brain image registration. This method, however, is relatively expensive. Xu and Li [7] also integrate feature guidance in 4D-image registration to improve the matching performance. To model the respiratory motion which is nearly periodic, the algorithms of [2,7] use a geometric constraint that enforces the average deformation to be identity. This constraint assumes the inhale and exhale phases are symmetric and the temporal samplings on these two phases are uniform, which are often not the case.

Main Contributions. This paper introduces a symmetric 4D registration model, which represents the sequential transformations using two 4D spline functions. The respiratory motion of the lung tumors/organs from the temporal CT scans can be described by a continuously deforming parametric geometry. Our experiments show that this algorithm results in significant performance improvement from existing methods in matching accuracy and trajectory smoothness.

2 Algorithm

2.1 Problem Statement and Overview

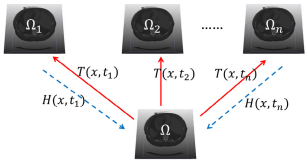

Fig. 1. Mapping Model

Given sequential volume images $I_1, I_2, \ldots, I_\Gamma$, where each image $I_i(\mathbf{x}) : \Omega_i \to \mathcal{R}, \mathbf{x} \in \Omega_i \subset \mathcal{R}^3$ is a 3D intensity function[1], we want to compute a temporally deforming 3D model $T(\mathbf{x}, t) : \Omega \times \mathcal{R} \to \mathcal{R}^3, \Omega \subset \mathcal{R}^3$ that correlates all the input images, as illustrated in Figure 1. A point $\mathbf{x} \in \Omega_i$ in I_i is correlated with a point $\mathbf{x}'$ in I_j by $\mathbf{x}' = T(T^{-1}(\mathbf{x}, t_i), t_j)$. Then, a continuous 4D deforming image $I(\mathbf{x}, t)$ can be constructed using the intensity function defined in the first image I_1, namely, $I(\mathbf{x}, t) = I_1(T(x, t_1))$.

[1] For sequential CT scans, their parametric domains Ω_i simply overlay in $\mathcal{R}^3$.

To obtain this deforming parametric geometry and the deforming image, we need to explicitly compute two 4D functions: (1) a forward 4D parameterization T, spatially defined on a common parametric domain, $T : \Omega \times \mathcal{R} \to \mathcal{R}^3$, and (2) its inverse mapping $H = T^{-1} : \mathcal{R}^3 \times \mathcal{R} \to \Omega$ which maps coordinate space of the deforming images $\Omega_i (\subset \mathcal{R}^3) \times \mathcal{R}$ to the common domain. To model the nonrigid freeform deformations of human organs during respiratory cycles, we use 4D B-spline functions to approximate these two transformations T and H, through which both the spatial and temporal smoothness can be formulated easily. The B-spline approximation for T can be formulated as:

$$T(\mathbf{y}) = \mathbf{x} + \sum_{\mathbf{y}_k \in N_{\mathbf{y}}} p_k \beta^r (\mathbf{y} - \mathbf{y}_k), \tag{1}$$

where $\mathbf{y} = (\mathbf{x}, t)$, $\mathbf{y}_k$ is a knot on the parametric domain $\Omega \times \mathcal{R}$; $\beta^r(\cdot)$ is the r-th order multidimensional B-spline polynomial (here we take $r = 3$); $\mathbf{p}_k \in \mathcal{R}^3$ are B-spline control points to be solved, and $N_{\mathbf{y}}$ denotes $\mathbf{y}$'s neighboring local support regions where the basis functions are nonzero. The knots $\mathbf{y}_k$ are defined on a 4D regular grid, uniformly overlaid the 4D image.

Because the inverse of B-spline transformation cannot be derived in close-form, we explicitly approximate this inverse mapping using another B-spline transformation H using a same formulation to eq (1). Then with T and H, a transformation F^{ij} from any frames i to j can be composed as

$$F^{ij}(\mathbf{x}) = T(H(\mathbf{x}, t_i), t_j), \mathbf{x} \in \Omega_i. \tag{2}$$

The entire 4D registration problem is formulated as an optimization on T and H that minimizes an objective function:

$$E = E_I + \alpha E_F + \lambda E_S + \rho E_C, \tag{3}$$

where E_I measures the intensity matching error, E_F measures the feature alignment, E_S measures the spatial and temporal smoothness of the deformation, E_C measures the inverse consistency, and α, λ, ρ are weighting factors.

Intensity Matching Error. With the assumption that the corresponded points have the same intensity, the registration should minimize the intensity differences of corresponded points. We can derive the intensity difference between corresponded points in any pair of images I_i and I_j taken in time t_i and t_j. For any point $\mathbf{x} \in \Omega_i$ in time t_i, its corresponding location in time t_j can be composed by H and T. The accumulated difference between $I_i(\mathbf{x})$ and the intensity of its corresponding coordinate in t_j can be formulated as:

$$\tilde{E}_I = \frac{1}{|S||\Gamma|^2} \sum_{t_i \in \Gamma} \sum_{t_j \in \Gamma} \sum_{\mathbf{x} \in S_i} (I_j(T(H(\mathbf{x}, t_i), t_j)) - I_i(\mathbf{x}))^2, \tag{4}$$

where S_i is the sets of spatial voxel coordinates in each Ω_i and for $\forall i, |S| = |S_i|$. Simultaneously solving both T and H is expensive. We first solve a forward parameterization T, then iteratively, fix the parameterization in one direction and optimize the other (see Section 2.2 for the complete algorithm).

To solve the initial forward parameterization T without knowing H, we formulate the reduction of intensity error by minimizing the intensity variance:

$$T_I = \frac{1}{|S||\Gamma|} \sum_{\mathbf{x} \in S} \sum_{t \in \Gamma} (I_t(T(\mathbf{x}, t)) - \bar{I}(\mathbf{x}))^2, \tag{5}$$

where $\bar{I}(\mathbf{x})$ is the average intensity value follows the forward parameterization: $\bar{I}(\mathbf{x}) = \frac{1}{|\Gamma|} \sum_{t \in \Gamma} I_t(T(\mathbf{x}, t))$. $S \subset \Omega$ are the spatial voxel coordinates (e.g. coordinates of all the pixels) and $\Gamma \subset \mathcal{R}$ contains the temporal coordinates indexing temporal sample images. After obtaining the initial T, we iteratively optimize H and T by minimizing:

$$E_I = T_I + \tilde{E}_I. \tag{6}$$

Feature Alignment Error. The intensity term has many local minima. Geometric features can help effectively avoid many undesirable solutions. We extract feature points using a slightly modified 3D SIFT algorithm [8], then compute a set of consistently corresponded feature points $\{p_{ij}\}$ across the entire sequence of images, where p_{ij} indicates the i-th feature point on time t_j, where $i = 1, \ldots, N, j = 1, \ldots, |\Gamma|$.

Each consistently corresponded feature point has a parametric coordinate $m_i, i = 1, \ldots, N$ in Ω, which is mapped to the feature p_{it} in image I_t at time t. The feature correspondence in the forward parameterization should penalize the deviation of $T(m_i, t)$ from p_{it}:

$$T_F = \frac{1}{N|\Gamma|} \sum_{t \in \Gamma} \sum_{i=1}^{N} ||p_{it} - T(m_i, t)||^2, \tag{7}$$

For the inverse map H, the variance of $H(p_{ij}, j)$ should be minimized:

$$H_F = \frac{1}{N|\Gamma|} \sum_{i=1}^{N} \sum_{t \in \Gamma} ||H(p_{it}, t) - \bar{H}(p_{i*})||, \tag{8}$$

where $\bar{H}(p_{i*}) = \frac{1}{|\Gamma|} \sum_{t \in \Gamma} H(p_{i,t}, t)$ is the average coordinates of the i-th feature p_{i*}. Finally, the entire feature alignment error is:

$$E_F = T_F + H_F. \tag{9}$$

Deformation and Motion Smoothness. The transformation (hence both parameterizations T and H) should be spatially and temporally smooth. The 2nd-order derivatives of the B-spline transformation functions can be derived as the smoothness energy to minimize:

$$\begin{aligned} E_S &= T_S + H_S; \\ T_S &= \frac{1}{|S||\Gamma|} \sum_{\mathbf{x} \in S} \sum_{t \in \Gamma} (||\frac{\partial^2 T}{\partial \mathbf{x}^2}||_F^2 + ||\frac{\partial^2 T}{\partial t^2}||^2 + 2||\frac{\partial^2 T}{\partial \mathbf{x} \partial t}||_F); \\ H_S &= \frac{1}{|S||\Gamma|} \sum_{\mathbf{x} \in S_i} \sum_{t \in \Gamma} (||\frac{\partial^2 H}{\partial \mathbf{x}^2}||_F^2 + ||\frac{\partial^2 H}{\partial t^2}||^2 + 2||\frac{\partial^2 H}{\partial \mathbf{x} \partial t}||_F). \end{aligned} \tag{10}$$

Inverse Consistency means the matching between two frames I_i and I_j should be symmetric and bijective. Namely, the matching from I_i to I_j (composed by T and H) is one-to-one and also consistent with that from I_j to I_i. In 3D pairwise registration, inverse consistency also means the selection of reference doesn't affect the matching result [9]. This consistency can be achieved by making the composition of T and H to be as near-identity as possible:

$$E_C = \frac{1}{|S||\Gamma|}\sum_{\mathbf{x}\in S_i}\sum_{t\in\Gamma}||T(H(\mathbf{x},t),t)-\mathbf{x}||^2 + \frac{1}{|S||\Gamma|}\sum_{\mathbf{x}\in S}\sum_{t\in\Gamma}||H(T(\mathbf{x},t),t)-\mathbf{x}||^2, \tag{11}$$

2.2 Solving the Optimization

Simultaneously solving T and H reduces to a very expensive optimization problem. We develop an iterative algorithm to seek for the optimal solution. During each iteration, T (or H) is solved using a gradient-based optimization method algorithm proposed in [2], which uses a stochastic sampling strategy to reduce the computational cost. With the B-spline representation we derive the derivatives of E_F, E_S, E_C explicitly, and we use the finite difference approximation to get the derivatives of E_I.

We first solve a forward parameterization T by minimizing $E = T_I + \alpha T_F + \rho T_S$ from equations (5,7,10), then with T fixed, we solve its inverse parameterization H by minimizing the entire objective function E in equation 3. Then iteratively, we fix one parameterization and revise its inverse parameterization, until the energy reduction is smaller than a threshold. This optimization algorithm is formulated as follows.

1) Compute an initial forward parameterization T by minimizing $T_I + \alpha T_F + \rho T_S$;
2) Fix T, and solve H by minimizing E;
3) Fix H, and solve T by minimizing E;
4) If E converges, STOP; otherwise GOTO 2).

3 Experimental Results

We implement our registration model via a multi-resolution strategy and use linear interpolation in the spatial domain for the derivation of intensity values for any point not on a grid. Our algorithm was implemented in C++ using an Intel Xeon X5570 @2.93 GHz, 8GB RAM. In our experiments, we set the weight factors in eq-(3) as $\alpha = 0.1, \lambda = 0.5, \rho = 0.5$.

3.1 Experiments Using Public Datasets

We perform 4D registration using our algorithm on two public benchmark datasets: POPI [10] and DIR-lab [11]. The dataset from POPI has one 4D CT

series including ten 3D volume images ($482 \times 360 \times 141$ pixels) representing ten different phases of one breathing cycle. We also select five datasets from the DIR-lab dataset (Case-1 to Case-5) where lardmarks are available. Each dataset contains 6 sequential volume images. This CT pixel unit can be converted to real physical space units millimeter by multiplying a scaling factor (recorded in the image header file). Consistent landmarks are also available in the benchmark to measure the accuracy of the registration. Denoting the landmarks on frame-t as $Q_t = \{q_{t,1}, q_{t,2}, \ldots, q_{t,n}\}$, the registration accuracy with respect to frame-r can be measured by a *Mean Target Registration Error* (MTRE):

$$D_r = \frac{1}{n|\Gamma|} \sum_{t \in \Gamma} \sum_{q_{r,i} \in Q_r} ||F^{rt}(q_{r,i}) - q_{t,i}||, \tag{12}$$

where F^{rt} is the transformation between frames r and t, composed by the forward and inverse parameterizations following equation (2).

Unlike existing 4D parameterization methods that solve mappings in two directions separately, our model uses a symmetric objective function that can be optimized with guaranteed convergence. Fig. 2 shows the convergence of the energy E (eq-(3)) when parameterizing the POPI dataset. Our algorithm converges in 4 iterations.

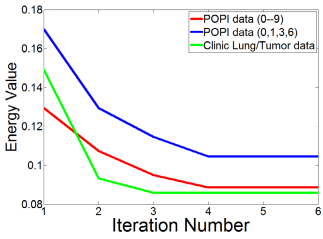

Fig. 2. Convergence of E

We compared our registration results with existing 3D pairwise registration [2] and 4D registration [2,7] algorithms using the benchmark data from POPI and DIR-lab. The results are documented in Table 1 and Table 2. On the POPI dataset, we evaluated the registration using the matching error of the consistent 40 landmarks. On each of DIR-lab datasets, both maximum inhale and exhale phases possess 300 landmarks, whose matching errors were used to evaluate the registration accuracy.

Table 1. The landmark predication error D_i and its standard deviation σ_i (in mm) of i^{th} time frame on the POPI-data [10]. $\bar{D}$ is the average MTRE.

	$D_1(\sigma_1)$	$D_2(\sigma_2)$	$D_3(\sigma_3)$	$D_4(\sigma_4)$	$D_5(\sigma_5)$	$D_6(\sigma_6)$	$D_7(\sigma_7)$	$D_8(\sigma_8)$	$D_9(\sigma_9)$	$D_0(\sigma_0)$	$\bar{D}$
3D Reg.[2]	3.6(2.3)	2.3(1.8)	2.1(1.7)	2.2(2.0)	2.4(2.3)	2.9(2.4)	2.8(2.3)	2.1(1.7)	2.1(1.5)	2.7(2.1)	2.5
4D Reg.[2]	3.8(2.3)	2.6(2.0)	2.2(1.8)	2.2(2.0)	2.5(2.2)	2.9(2.3)	2.8(2.3)	2.2(1.8)	2.2(1.5)	2.8(2.2)	2.6
4D Reg.[7]	2.1(1.6)	1.8(1.5)	1.6(1.3)	1.6(1.2)	2.1(1.4)	2.4(1.7)	2.1(1.6)	1.7(1.0)	1.6(1.2)	1.9(1.6)	1.9
Our 1^{th} iter.	1.9(1.4)	1.6(1.2)	1.6(1.3)	1.8(1.5)	2.0(1.7)	2.0(1.7)	2.0(1.6)	1.6(1.2)	1.7(1.1)	2.2(1.6)	1.9
Our 2^{th} iter.	1.1(0.8)	1.2(0.9)	1.3(0.9)	1.2(0.8)	1.5(0.9)	1.5(1.0)	1.4(1.1)	1.1(0.6)	1.2(0.8)	1.2(0.8)	1.3
Our 3^{th} iter.	1.1(0.8)	1.2(0.9)	1.3(0.9)	1.2(0.8)	1.5(0.9)	1.5(1.0)	1.4(1.1)	1.1(0.6)	1.2(0.8)	1.2(0.8)	1.3

Our algorithm results in significantly smaller MTRE errors, i.e., better landmarks prediction accuracy. Furthermore, in practice, the breathing cycles are often sampled in a non-perfectly uniform manner: the exhale is longer than the inhale. In this case, the geometric constraints adopted in [2] and [7], which enforces the average movement of each point during the entire cycle to be identity,

Table 2. The landmark predication error and its standard deviation $D_i(\sigma_i)$ (in mm) for the registration of DIR-LAB 4D dataset: $i = 1$ to 5 for Case-1 to Case-5

Method/Results:	$D_1(\sigma_1)$	$D_2(\sigma_2)$	$D_3(\sigma_3)$	$D_4(\sigma_4)$	$D_5(\sigma_5)$
3D Reg[2]	2.03(1.09)	0.72(0.44)	0.99(0.71)	1.14(0.81)	1.64(1.70)
4D Reg[2]	2.12(1.09)	0.92(0.61)	1.39(0.93)	1.44(0.96)	1.85(1.69)
4D Reg[7]	1.58(0.99)	0.70(0.57)	0.79(0.55)	0.91(0.75)	1.41(1.36)
Our Reg	1.28(0.76)	0.56(0.34)	0.59(0.43)	0.69(0.49)	1.10(0.94)

will not be correct. The geometric smoothness cost (eqn 10) suggested in our framework is a more robust description against this non-uniformity.

3.2 Motion Modeling of Our Clinical Lung Tumor Scans

We also build a 4D parameterization to describe the lung/tumor deformation from clinic CT scans. We first perform image segmentation and construct finite element mesh models [12], then use the 4D mapping to compute its deformation. Fig. 3 illustrates a few snapshots of this tracking. (a) shows the surface contour segmented from frame-1 and (b) shows the deformed contour on frame-6; (c) and (d) show the color-encoded displacement fields of our deformable model; (e) illustrates the matching error measured by Hausdorff distance. This illustrated matching, between the maximum inhalation (I_6) and maximum exhalation status (I_1) which undergoes a largest deformation, infers the maximum matching errors during the respiratory cycles. See our video for better visualization.

(a)S_1 from I_1 (b) $F^{1,6}(S_1)$ on I_6 (c) I_1 (d) $d(F^{1,6}(I_1), I_1)$ (e) $H(F^{1,6}(I_1), I_6)$

Fig. 3. Lung/Tumor Tracking via a Deforming Surface Geometry. (a, b) show the alignment of iso-contours and the scanned images. (d) shows the color-coded displacement field of $F^{1,6}(I_1)$ from I_1 in (c); (e) visualizes the Hausdorff distance from the deformable model to the scan.

4 Conclusion

We propose an effective 4D registration algorithm for dynamic volume images. The 4D parameterization is represented using two coupled B-spline functions and solved by minimizing an objective function E measuring intensity matching, feature alignment, spatial and temporal smoothness, and transitive inverse-consistency. Compared with existing 3D and 4D registration models, this algorithm has unique advantages in matching dynamic volume image sequences that undergo relatively big nonrigid deformation and/or are non-uniform in the

temporal dimension. To minimize E, we alternatively optimize the forward and inverse parameterizations T and H, which iteratively refines each other in a symmetric manner. Our experimental results demonstrate that this computational model greatly improves the registration accuracy of existing methods.

Acknowledgements. This work is supported in part by LA-Board of Regents (BOR) RCS LEQSF(2009-12)-RD-A-06, LA-BOR LEQSF-EPS(2013)-PFUND-312, and IBM Faculty Award. The clinic data are from Amit Sawant of UT Southwestern Medical Center.

References

1. Gorbunova, V., Sporring, J., Lo, P., Loeve, M., Tiddens, H., Nielsen, M., Dirksen, A., Bruijne, M.: Mass preserving image registration for lung CT. Medical Image Analysis 16(4), 786–795 (2012)
2. Metz, C., Klein, S., Schaap, M., van Walsum, T., Niessen, W.: Nonrigid registration of dynamic medical imaging data using nD + t B-splines and a groupwise optimization approach. Medical Image Analysis 15, 238–249 (2011)
3. Heinrich, M., Jenkinson, M., Brady, S., Schnabel, J.: Mrf-based deformable registration and ventilation estimation of lung CT. IEEE Trans. on Med. Imag. (2013)
4. Wu, G., Wang, Q., Lian, J., Shen, D.: Estimating the 4D respiratory lung motion by spatiotemporal registration and building super-resolution image. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part I. LNCS, vol. 6891, pp. 532–539. Springer, Heidelberg (2011)
5. Xue, Z., Wong, K., Wong, S.: Joint registration and segmentation of serial lung CT images for image-guided lung cancer diagnosis and therapy. Computerized Medical Imaging and Graphics 34(1), 55–60 (2010)
6. Wu, G., Wang, Q., Jia, H., Shen, D.: Feature-based groupwise registration by hierarchical anatomical correspondence detection. Human Brain Mapping 33(2), 253–271 (2012)
7. Xu, H., Li, X.: Consistent feature-aligned 4d image registration for respiratory motion modeling. In: Int. Symp. on Biom. Imaging, pp. 580–583 (2013)
8. Xu, H., Chen, P., Yu, W., Sawasnt, A., Iyengar, S., Li, X.: Feature-aligned 4D spatiotemporal image registration. In: Int. Conf. on Patt. Recog., pp. 2639–2642 (2012)
9. Christensen, G., Johnson, H.: Consistent image registration. IEEE Transaction on Medical Imaging 20(7), 568–582 (2001)
10. Vandemeulebroucke, J., Sarrut, D., Clarysse, P.: Point-validated pixel-based breathing thorax model. In: Int. Conf. on the Use of Comp. in Rad. Therapy (2007)
11. Castillo, R., Castillo, E., Guerra, R., Johnson, V.E., McPhail, T., Garg, A.K., Guerrero, T.: A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. Phy. in Med. & Bio. 54, 1849–1870 (2009)
12. Iyengar, S., Li, X., Xu, H., Mukhopadhyay, S., Balakrishnan, N., Sawant, A., Iyengar, P.: Toward more precise radiotherapy treatment of lung tumors. IEEE Computer 45, 59–65 (2012)