

Deformable 3D Volume Registration Using Efficient MRFs Model with Decomposed Nodes

Kyong Joon Lee¹, Dongjin Kwon¹, Il Dong Yun², and Sang Uk Lee¹

¹School of EECS, Seoul Nat'l Univ., Seoul, 151-742, Korea

²School of EIE, Hankuk Univ. of F. S., Yongin, 449-791, Korea

kjoon@cvl.snu.ac.kr, djk@cvl.snu.ac.kr, yun@hufs.ac.kr, sanguk@ipl.snu.ac.kr

Abstract

An efficient registration algorithm working on non-rigid 3D objects is presented. We formulate the registration as a discrete labeling problem on MRFs model whose energy can be minimized by optimization techniques in the literature. Due to the huge search range in three-dimensional space, previous approaches produces a vast amount of labels for a node in the MRFs graph. To reduce the number of labels, we decompose a node into three nodes so that the labels in each node represent just one-dimensional displacement. This procedure introduces a factor node with a clique potential of size three, defining ternary interaction between the decomposed nodes. We convert the factor node into pairwise interactions and adopt the tree-reweighted message passing technique, which guarantees the convergence of lower bound of the energy function. In experiments we use clinical and synthetically deformed 3D medical images. Result shows the proposed method enhances computational efficiency without loss of accuracy.

1 Introduction

Non-rigid 3D registration is the process to match the identical volumes that are mismatched in a non-linear or non-uniform manner in two or more views. As the other low-level vision problems such as stereo matching and optical flows, it deals with the most fundamental part of computer vision, *i.e.*, view correspondences. Thus, a plausible method for 3D registration will be of great importance inspiring various fields in the literature, also giving lots of useful applications: the most practical and popular one may be found in medical imaging field.

Conventional approaches utilize a mesh structure superimposed on input volumes. Deformation of the mesh is estimated first, then calculating the rest of voxel transformations is followed. Feature-based methods [7, 6] relies on landmarks extracted from input image pairs. The mesh deformation is controlled by matching the corresponding points and completed by surface interpolation. An answer is given in relatively short time, but may not accurate enough for some applications. In contrast Rueckert *et al.* [8] proposed an image-based approach which searches for pixel-wise similarities and gradually transforms the floating image to align with a fixed image. An energy function is formulated using image matching cost and mesh regularizing cost, and optimized through a classical

technique *i.e.*, a gradient descent algorithm assuming local convexity. Although the algorithm gives a state-of-the-art accuracy, the computation time is tremendous. Furthermore, a simple variation of the cost definition would require that amount of time again, since the optimization depends on the gradient of the energy function.

Recently researchers proposed new approaches reformulating the registration as a discrete labeling problem on MRFs model, making the optimization technique independent from the cost definitions. Glocker *et al.* [1] utilized this strategy and covered inter and intra modality in medical imaging field. They limited search range of similarity matching cost by assuming relatively small local deformation and by indicating specific search directions, *e.g.* eight directions for 2D problems. In this way, they were able to greatly reduce the number of labels in each graph node, however the accuracy might degenerate if applied to the problems with severe deformations. Tang *et al.* [11] applied the graph cuts to the 2D and 3D registration of medical images and compared their algorithm with the previous innovative works such as FFD [8] and DEMON [5]. Even though they showed fine experimental results, however, taking every pixels as individual nodes may cause huge computational complexity, especially in the 3D problem.

Generally, a method to address intensive deformation would demand huge number of labels to cover a possible target object in a high dimensional search range. Thus it would suffer from big computational burden, because the state-of-the-art optimization algorithms increase complexity directly proportional to the number of labels. In the case of 2D registration, the method proposed in [9] could be a good solution. It decomposes a node of a 2D displacement vector into two nodes of 1D vectors to decrease the number of labels. A pairwise potential energy is defined on the interaction between the decomposed nodes. The total energy term can be easily minimized using a presented optimization technique. However, if in a 3D problem, the process would induce higher-order interaction between the decomposed nodes, and directly applying the existing optimization algorithm is not straightforward. In this paper we propose a novel approach to decompose a node in the 3D registration problem by introducing a factor node with higher-order clique potential. We also provide a way to convert the factor graph into a model with simpler interactions and an efficient implementation to optimize it through message-passing based methods. As a result, the proposed algorithm extremely reduce the computational complexity of the 3D registration.

2 Registration as A Discrete Labeling Problem

We introduce the basic framework and review the MRFs formulation for it. Throughout this paper, we refer a fixed volume as *reference* and a floating volume as *input*. Our goal is to find the optimal transformation vector field $\mathcal{T}(\mathbf{x})$, where \mathbf{x} indicates the input volume domain. Assuming local deformation, we superimpose a deformable mesh on the input volume. The overall strategy can be stated as follows: Find the *discrete* displacement vectors for control points, then the pixel-wise deformation field could be calculated accordingly using the FFD model based on cubic B-splines [8].

For the corresponding MRFs model, we start with some definitions. Let \mathcal{G} be an undirected graph, denoting the deformable mesh, with node set \mathcal{V} and edge set \mathcal{E} . A node $s \in \mathcal{V}$ corresponds to a control point in the mesh. Let l_s be a random variable in some discrete sample space $\mathcal{L}_s = \{1, \dots, L\}$, representing the label of the node s . With

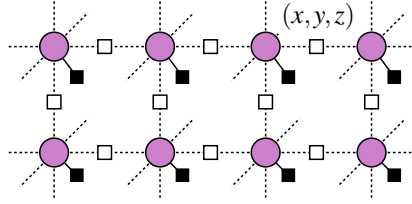


Figure 1: Conventional model. A node represents a control point in the mesh, having a label of 3D displacement vector. Unary potentials are described in black squares while binary potentials in white squares.

all the random variables in each node, we obtain a vector $\mathbf{l} \in \mathcal{L}$ of size $n = |\mathcal{V}|$, where $\mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2 \times \dots \times \mathcal{L}_n$. The goal is restated as finding the best configuration \mathbf{l} that minimizes an energy formulation defined by potential functions as follows,

$$E(\mathbf{l}|\theta) = \sum_{s \in \mathcal{V}} \theta_s(l_s) + \sum_{(s,t) \in \mathcal{E}} \theta_{st}(l_s, l_t). \quad (1)$$

The unary potential function $\theta_s(l_s)$ indicates a data cost defined by similarity measures between input and reference volumes. The binary potential $\theta_{st}(l_s, l_t)$ denotes a smoothness cost computed using the label difference between adjacent nodes s and t . Figure 1 presents this model describing a mesh control points as a node, a unary potentials as a black square, and a binary potential as a white square. The square nodes can be seen as a factor node [3] as well, since factor nodes can represent various type of potential functions, *e.g.*, unary, binary or ternary potentials. We will denote the factor node set as \mathcal{F} to alternatively indicate edges with higher order cliques.

3 Node Decomposition

For the conventional model, we can introduce a function $\mathbf{d} : \mathcal{L}_s \rightarrow \mathbb{R}^3$, mapping each label l_s to a 3D displacement vector $\mathbf{d}(l_s)$. Without loss of generality, we assume each dimension of displacements has the same discrete set of values $\mathcal{D} = \{-D, -D+1, \dots, 0, \dots, D-1, D\}$ where $D \in \mathbb{N}$ controls the displacement width. Then, the model produces labels which is cubic to the displacement width, that is, $|\mathcal{L}| = |\mathcal{D}|^3 = (2D+1)^3$.

The problem is, this label size actually cause heavy computational complexity for discrete optimization algorithms. For example, the approximate computation times for the state-of-the-art methods such as graph cuts and TRW-S are $O(|\mathcal{V}||\mathcal{L}|)$ and $O(|\mathcal{V}||\mathcal{L}|^2)$ respectively, becoming $O(|\mathcal{V}||\mathcal{D}|^3)$ and $O(|\mathcal{V}||\mathcal{D}|^6)$, where the displacement width D usually is not trivial.

We address this challenge by adapting the *decomposed* model [9]. A node $s \in \mathcal{V}$ is decomposed into three nodes s^x , s^y and s^z , constituting new node sets \mathcal{V}^x , \mathcal{V}^y and \mathcal{V}^z , respectively. These node sets construct individual graph structures introducing new edge sets, *i.e.*, $\mathcal{G}^x = (\mathcal{V}^x, \mathcal{E}^x)$, $\mathcal{G}^y = (\mathcal{V}^y, \mathcal{E}^y)$ and $\mathcal{G}^z = (\mathcal{V}^z, \mathcal{E}^z)$. Note that the graph structures are identical to the original graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. As like the original edge set \mathcal{E} , the edge

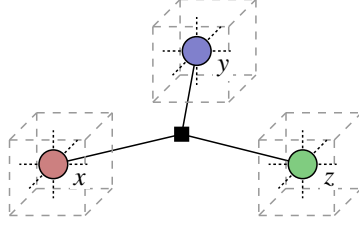


Figure 2: Decomposed model. A factor node is introduced to define the interactions between the *three* decomposed nodes. Note that this factor node is described in a black square, implying it originated from the unary potential in Figure 1.

sets $\mathcal{E}^x, \mathcal{E}^y, \mathcal{E}^z$ define the smoothness cost between neighboring nodes in the individual graphs. We call these pairwise interactions as *intra-volume* edges.

We also define the interaction between the *three* decomposed nodes. A factor node $f \in \mathcal{F}$ is introduced to connect the decomposed nodes as shown in Figure 2. This factor node is defined in a ternary potential function coming from the unary potential of the original graph, e.g., $\theta_{stu}(l_s^x, l_s^y, l_s^z) = \theta_s(l_s)$ where the original label l_s indicates a 3D vector displacement of (l_s^x, l_s^y, l_s^z) . We call this factor node set as *inter-volume* edges. Unary terms for the decomposed nodes are defined as 0, i.e., $\theta_{s^i}(l_{s^i}) = 0$ where $s^i \in \mathcal{V}^x \cup \mathcal{V}^y \cup \mathcal{V}^z$.

The proposed model allows a label in each node to represent just one-dimensional displacement in the Cartesian coordinate. We can modify the mapping function as $d : \mathcal{L}_s \rightarrow \mathbb{R}$ where $s \in \mathcal{V}^x \cup \mathcal{V}^y \cup \mathcal{V}^z$. Energy formulation in (1) is accordingly modified as follows,

$$E(\mathbf{l}|\theta) = \sum_{(s,t) \in \mathcal{E}^x \cup \mathcal{E}^y \cup \mathcal{E}^z} \theta_{st}(l_s, l_t) + \sum_{(s,t,u) \in \mathcal{F}} \theta_{stu}(l_s, l_t, l_u). \quad (2)$$

Consequently, the problem can be reformulated as minimizing the energy function defined with binary and ternary clique potentials.

While we are able to assign just one-dimensional label on each node, the number of nodes increases threefold. Note, however, that this condition is much more favorable since the complexity becomes $O(|\mathcal{V}||\mathcal{D}|)$ and $O(|\mathcal{V}||\mathcal{D}|^3)$ for the graph cuts and the TRW-S.

The MRFs energies computed by same configurations of the product model and the decomposed model are exactly same. However, Werner [14] proves that the solution space of the decomposed model properly contains that of the product model; that is, lower bound of the decomposed model is less tight when solved by LP-relaxation. This can be another trade-off for the simplicity besides the doubled number of nodes, though not serious.

4 Optimization using Tree-Reweighted Message Passing

For optimization, we are not allowed to choose graph cuts-based methods, because the clique potentials defined on inter-volume factor nodes are not included in any type of metric function. Instead we adopt message passing based methods which does not depend on the type of spatial prior. Recently the tree-reweighted message passing (TRW)

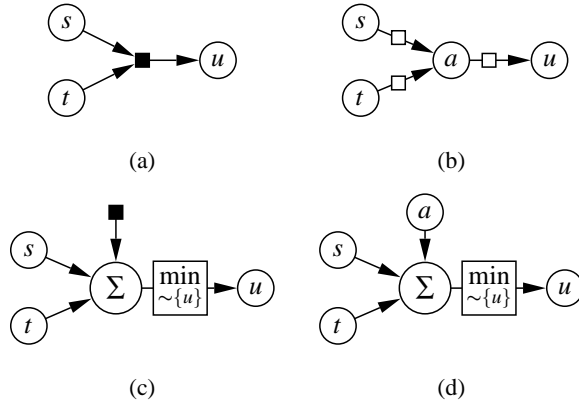


Figure 3: Conversion from a factor graphs to pairwise interactions: (a) A cycle-free factor graph, (b) A graph with pairwise interactions denoted by white squares. The factor node is replaced by an auxiliary node a . (c) A converted graph from (a) to calculate message to the root node u , (d) A converted graph from (b) to calculate message to the root node u .

algorithms [12, 2] are presented in the literature showing excellent performance [10]. The TRW methods decompose a graph into trees where an exact MAP can be calculated by min-sum belief propagation on each tree. This scheme let the methods provide a lower bound for the energy. In particular, the *sequential* TRW (TRW-S) [2] adjusts the message updating schedule and yields a lower bound guaranteed not to decrease. Thus this algorithm is preferable compared to the loopy belief propagation(LBP) of which scheduling is heuristic and convergence is unknown.

4.1 Conversion from a Factor Node to Pairwise Interactions

Since the TRW-S assumes pairwise interactions on edges, we need to reorganize our graph model which involves the higher-order factor nodes. Here, we introduce only conceptual idea of the conversion. Readers are referred to [13, 12] for detailed proof.

In a cycle-free factor graph, passing a message to the root node includes following steps [3]: Replace each variable node with a sum operator. Replace each factor node with a “sum and add by the factor node f ” operator, and between a factor node f and its parent u , insert a $\min_{\sim\{u\}}$ operator, where $\sim\{u\}$ denotes the node set in the graph except the node u . Figure 3(a) and (c) demonstrate this process of conversion from (a) to (c). Now consider a factor graph in Figure 3(b). If we appropriately define the node a and the pairwise factor node (depicted in blank square) between nodes, an equivalent graph (d) would be obtained throughout the process. Having the similar structure, (c) can be easily converted to (d). Consequently, by introducing an auxiliary node a , we can convert a factor graph into a graph with pairwise interactions.

To this end, the pairwise factor node and the auxiliary node a in (b) should be carefully designed. We associate a new label $z \in \mathcal{Z}$ where $\mathcal{Z} = \mathcal{L}_s \times \mathcal{L}_t \times \mathcal{L}_u$, with a unary and a binary potentials as follows

$$\psi_a(z) = \theta_{stu}(l_s, l_t, l_u), \quad (3)$$

$$\psi_{ai}(z, l_i) = \begin{cases} 0 & \text{if } z_i = l_i \\ \infty & \text{otherwise} \end{cases} \quad \forall i \in \{s, t, u\}. \quad (4)$$

Note that the functions force the auxiliary node a to follow the configurations in nodes s , t and u . After the conversion, energy model (2) changes to following

$$E(\mathbf{l}|\theta) = \sum_{(s,t) \in \mathcal{E}^x \cup \mathcal{E}^y \cup \mathcal{E}^z} \theta_{st}(l_s, l_t) + \sum_{a \in \mathcal{V}_{\mathcal{F}}} \psi_a(z) + \sum_{(a,i) \in \mathcal{E}_{\mathcal{F}}} \psi_{ai}(z, l_i). \quad (5)$$

4.2 Efficient Implementation

Although we successfully changed the factor nodes into pairwise interactions, we cannot avoid to include the auxiliary nodes which have vast amount of label space. Due to the $\min_{\sim\{u\}}$ operator in message passing, the auxiliary node can induce $O(|\mathcal{D}|^6)$ of time complexity as well as $O(|\mathcal{D}|^4)$ of memory space to store the message. However, the storage for ψ_{ai} actually is not needed since it never adds a value on a message. We can ignore all summations if $z_i \neq x_i$. Furthermore, as the auxiliary node should follow the configurations of nodes s and t when they are fixed, the search space of $\min_{\{a\}}$ operator is reduced to $O(|\mathcal{D}|)$ which is same as the ordinary nodes. Thus, the complexity becomes $O(|\mathcal{D}|^2)$ as like general pairwise TRW-S, by just modifying logics in the implementation.

5 Experiments

Medical images include good examples of deformable objects in 3D volumes. We conduct experiments on volume data of liver CT images. The test CT volumes are scanned from five patients, indexed by i , at arterial(A_i) and portal phases(P_i), which means before and after contrast medium injection. The volumes have the size of $256 \times 256 \times S_z$ where S_z ranges from 70 to 170. The mesh spacing is equally set to 16 pixels for x , y , z coordinates. All experiments are performed on an Intel Pentium 4, 3.6 GHz machine with 3.12 GB RAM.

Data Cost. Since the optimization process is independent from the definition of data cost, we can use various types of data cost, *e.g.*, the SSD(Sum of Squared Difference), the NCC(Normalized Cross Correlation) or the NMI(Normalized Mutual Information). For simplicity, we chose the SSD with matching cubes of $16 \times 16 \times 16$ where the search space covers $15 \times 15 \times 15$. Note that the number of labels for a node would be 3,375 in the product model, while in the proposed model, only 15 labels are needed for each node.

Smoothness Cost. Previous methods [1, 11] usually adopt the *linearly* penalizing function for the smoothness cost, for example,

$$\theta_{st}(x_s, x_t) = \lambda_{st} \min \{ |\mathbf{d}(x_s) - \mathbf{d}(x_t)|, T \}. \quad (6)$$

We empirically found a convex prior with higher order polynomial showed better results. It assigns substantially low cost on tolerable label difference and gives more flexibility on mesh deformation.

$$\theta_{st}(x_s, x_t) = \lambda_{st} \min \left\{ \left| \frac{1}{\alpha} (d(x_s) - d(x_t)) \right|^n, T \right\}, \quad (7)$$

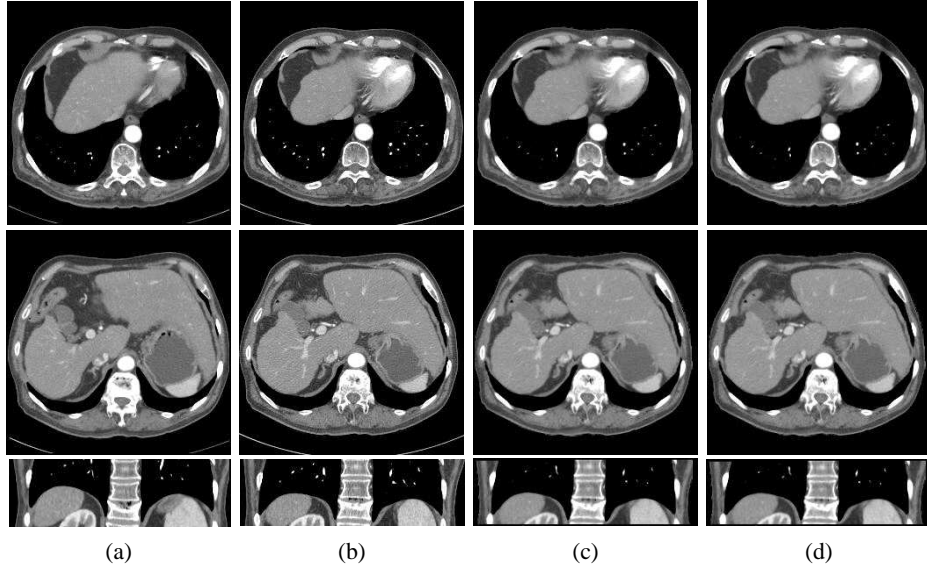


Figure 4: Registration result in synthetic data set. Top and middle row images show two different slices for *axial* section while bottom row shows *coronal* section. (a) Input volume A_1^{syn} , (b) Reference volume A_1 , (c) Result from the conventional model, (d) Result from the proposed model.

where $n \gg 1$ and α sets the boundary in which the function assigns almost zero-cost for small displacement difference. For convenience in implementation, we introduce an alternative definition with two linear functions [9]. As we set $c_1 \gg c_2$ in accordance with n , (8) approximates (7).

$$\theta_{sr}(x_s, x_t) = \lambda_{sr} \min \left\{ \max \{c_1(|d_i(x_s) - d_i(x_t)| - \alpha), c_2|d_i(x_s) - d_i(x_t)|\}, T \right\}. \quad (8)$$

5.1 Synthetically Deformed Data

The ground truth data for 3D registration generally is not available. For quantitative analysis we synthetically deformed corresponding reference volumes from the clinical data, obtaining ground truth vector field to calculate exact errors. For a pair of arterial and portal volumes, we manually put landmarks on some common characteristic features and matches them. Then the TPS(Thin-Plate Spline) interpolation generates new volumes for each pair, providing pixel-wise displacement vector fields. From the five volume pairs, ten synthetic volumes are generated and denoted as A_i^{syn} and P_i^{syn} . Note that A_i^{syn} will be almost identical to P_i (and vice versa) except the modality if the landmarks are sufficiently extracted.

We have tested four models characterized by the node decomposition and the smoothness cost. The proposed model uses the decomposed model with the cost function (8). The conventional model can be considered as a product model with the smoothness cost function (6) as in [1]. Figure 4 shows the registration results from these two models. Difference images are also given in Figure 5 where the background color demonstrates no

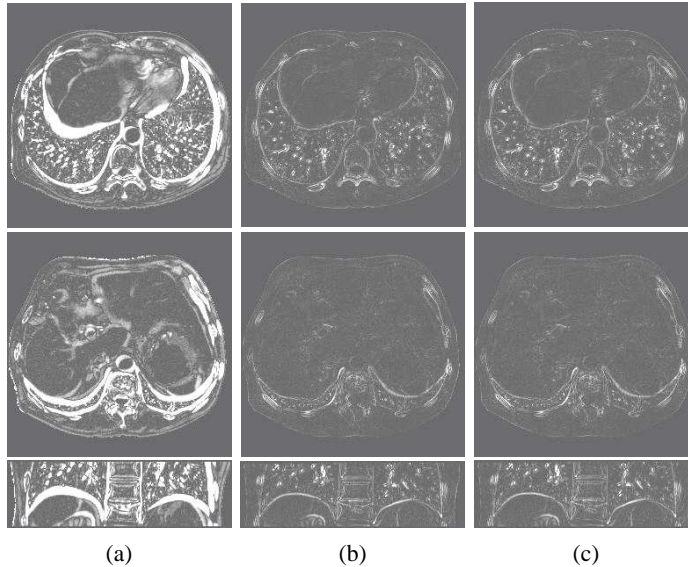


Figure 5: Intensity difference images in synthetic data set A_1 and A_1^{syn} . Bright areas indicates big differences. (a) Initial difference, (b) After registration using the conventional model and (c) the proposed model.

difference. As shown in Figure 5(b) and (c), both of the methods successfully removed white areas except some object boundaries and some details in *lung* section, which is smaller than a mesh resolution. Actually, all the four models present similar performance, at least in visual sense. Note that no particular qualitative degeneracy is found in the result from the decomposed model, compared to the product model.

Now we turn to the quantitative analysis. Table 1 shows intensity and vector differences with elapsed times in average of the data sets. We assume, without loss of generality, RMS errors implies overall accuracy while max errors indicates flexibility of the models. As can be seen, the proposed smoothness cost outperforms the conventional smoothness cost, also converging faster as well. In spite of the weak relaxation [14] in the optimization technique, we can hardly find worse performance of the decomposed model in terms of accuracy and flexibility. On the other hand, convergence time takes 5,000 ~ 10,000 seconds for the product models, while it only takes 60 ~ 90 seconds for the proposed models. Note that the precalculation time for the data cost is not included, which takes several minutes for all models.

5.2 Clinical Data

In clinical tests, we transformed the arterial phase A_i to register the portal phase volume P_i [4]. The registration result using the proposed model is denoted as P_i^{reg} . A malignant tumor of liver shows different contrast characteristic to normal liver tissues after the contrast medium perfusion. Thus the intensity is inherently (and deliberately) different in two phases even in the same organs, as shown in Figure 6(a) and (b). Compared with the initial difference in Figure 6, the difference after registration in Figure 6(d) actually

Table 1: Statistical results on synthetically deformed data sets

	initial	product		decomposed	
		sc (6)	sc (8)	sc (6)	sc (8)
i rms	200.85	48.53	46.82	48.75	47.33
i max	2051.00	1381.00	1348.20	1379.25	1367.38
v rms	5.70	1.44	1.26	1.43	1.31
v max	15.74	12.57	11.63	12.52	11.96
time	37.16	9504.46	5086.47	85.73	65.08

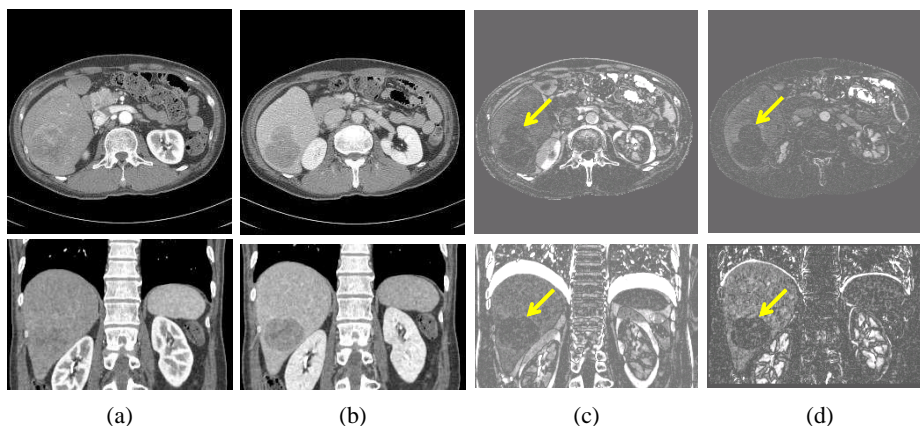


Figure 6: Registration result in clinical data using proposed method: (a) Arterial phase A_4 , (b) Portal phase P_4 , (c) Signed difference image $A_4 - P_4$ before registration, (d) Signed difference image $A_4 - P_4^{reg}$ after registration. The registration result actually shows better visual contrast for a liver cancer as indicated by arrows.

provides enhanced visibility (as intended) for a *hypervascular hepatocellular carcinoma*, *i.e.*, liver cancer. The result also shows our method works fine with the images of different modality, suggesting a plausible application for medical diagnosis.

6 Conclusion

In this paper, we have presented an efficient algorithm for the registration of deformable 3D objects. Our approach formulates the problem with discrete labeling on MRFs model and reduces the number of labels by the decomposed model. The energy function is optimized by the modified TRW-S handling higher order cliques. In the experimental results, we have shown the proposed model significantly enhanced the computational efficiency of the problem, without losing accuracy and robustness in both qualitative and quantitative sense. Still, the required time to calculate the data cost in three dimensional search space is tremendous and should be addressed in the future.

Acknowledgements

This work was supported by the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government(MOST) (R01-2007-000-11425-0).

References

- [1] B. Glocker, N. Komodakis, N. Paragios, G. Tziritas, and N. Navab. Inter and Intra-modal Deformable Registration: Continuous Deformations Meet Efficient Optimal Linear Programming. In *IPMI*, 2007.
- [2] V. Kolmogorov. Convergent Tree-Reweighted Message Passing for Energy Minimization. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(10):1568–1583, 2006.
- [3] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger. Factor graphs and the sum-product algorithm. *IEEE Trans. on Information Theory*, 47(2):498–519, 2001.
- [4] D. Kwon, I. D. Yun, K. H. Lee, and S. U. Lee. Efficient Feature-Based Nonrigid Registration of Multiphase Liver CT Volumes. In *BMVC*, 2008.
- [5] X. Pennec, P. Cachier, and N. Ayache. Understanding the "demon's algorithm": 3d non-rigid registration by gradient descent. In *MICCAI*, pages 597–605, 1999.
- [6] J. Pilet, V. Lepetit, and P. Fua. Real-Time Non-Rigid Surface Detection. In *CVPR*, 2005.
- [7] K. Rohr. Image Registration Based on Thin-Plate Splines and Local Estimates of Anisotropic Landmark Localization Uncertainties. In *MICCAI*, pages 1174–1183, 1998.
- [8] 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 Trans. Medical Imaging*, 18(8):712–721, 1999.
- [9] A. Shekhovtsov, I. Kovtun, and V. Hlaváč. Efficient MRF Deformation Model for Non-Rigid Image Matching. In *CVPR*, 2007.
- [10] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother. A Comparative Study of Energy Minimization Methods for Markov Random Fields. In *ECCV*, 2006.
- [11] T. W. H. Tang and A. C. S. Chung. Non-rigid image registration using graph-cuts. In *MICCAI*, 2007.
- [12] M. J. Wainwright, T. Jaakkola, and A. S. Willsky. MAP estimation via agreement on trees: message-passing and linear programming. *IEEE Trans. on Information Theory*, 51(11):3697–3717, 2005.
- [13] Y. Weiss and W. T. Freeman. On the optimality of solutions of the max-product belief-propagation algorithm in arbitrary graphs. *IEEE Trans. on Information Theory*, 47(2):736–744, 2001.
- [14] T. Werner. A linear programming approach to max-sum problem: A review. *IEEE Trans. Pattern Anal. Machine Intell.*, 29(7):1165–1179, Jul. 2007.