IEEE Transactions on Medical Imaging

Local Metric Learning in 2D/3D Deformable Registration With Application in the Abdomen

Journal:	*Transactions on Medical Imaging*
Manuscript ID:	TMI-2013-0702
Manuscript Type:	Full Paper
Date Submitted by the Author:	25-Sep-2013
Complete List of Authors:	Zhao, Qingyu; UNC, CS Pizer, Stephen; UNC, CS Chou, Chen-Rui; UNC, CS Mageras, Gig; Memorial Sloan-Kettering, Medical Physics
Keywords:	X-ray imaging and computed tomography < Imaging modalities, Lung < Object of interest, Abdomen < Object of interest, Registration < General methodology, Image-guided treatment < General methodology, Motion compensation and analysis < General methodology, Probabilistic and statistical methods < General methodology
Specialty/Area of Expertise:	Radiation treatment, Registration, esp. 2D/3D, Regression, Abdomen

SCHOLARONE™ Manuscripts

For Review

Local Metric Learning in 2D/3D Deformable Registration With Application in the Abdomen

Qingyu Zhao, Chen-Rui Chou, Gig Mageras, and Stephen Pizer

***Abstract*—In image-guided radiotherapy (IGRT) of disease sites subject to respiratory motion, soft tissue deformations can affect localization accuracy. We describe the application of a method of 2D/3D deformable registration to soft tissue localization in abdomen. The method, called Registration Efficiency and Accuracy through Learning a Metric on Shape (REALMS), is designed to support real-time IGRT. In a previously developed version of REALMS, the method interpolated 3D deformation parameters for any credible deformation in a deformation space using a single globally-trained Riemannian metric for each parameter. We propose a refinement of the method in which the metric is trained over a particular region of the deformation space, such that interpolation accuracy within that region is improved. We report on the application of the proposed algorithm to IGRT in abdominal disease sites, which is more challenging than in lung because of low intensity contrast and non-respiratory deformation. We introduce a rigid translation vector to compensate for non-respiratory deformation, and design a special region-of-interest around fiducial markers implanted near the tumor to produce a more reliable registration. Both synthetic data and actual data tests on abdominal datasets show that the localized approach achieves more accurate 2D/3D deformable registration than the global approach.**

***Index Terms*—IGRT, 2D/3D registration, abdomen, radiation oncology.**

I. Introduction

THE goal of the Image-Guided Radiation Therapy (IGRT) process is to localize tumors and target organs in 3D at treatment time. The advent of so-called on-board imaging systems mounted on medical linear accelerators makes possible the acquisition of 2D planar images during radiation treatment for monitoring internal patient motion. This has motivated the use of 2D/3D registration for target localization and patient positioning just prior to treatment, and for tracking target motion during treatment delivery [1]. A challenge in lung and abdominal disease sites subject to respiratory motion is that organ deformation may occur, thus requiring incorporation of deformation in the registration process. In this context, the patient's treatment-time 3D deformations can be computed by performing image registration between the treatment-time on-board planar images (X-ray) and the treatment-planning 3D image (CT). Chou et al. [2] recently introduced a real-time 2D/3D deformable registration method, called Registration Efficiency and Accuracy through Learning Metric on Shape (REALMS). At treatment-planning time REALMS learns a Riemannian metric that measures the distance between two projection images. At treatment time it interpolates the patient's 3D deformation parameters using a kernel regression with the learned distance metric. REALMS can locate the target in less than 10 ms at treatment time, which shows potential to support real-time registration. However, the previously reported method approximates the Riemannian metric by using a linear regression over the global deformation space between the projection image intensity differences and the deformation parameter differences. Therefore, the accuracy highly depends on how well this relationship can fit into a global linear model.

We describe an improvement scheme for REALMS using local metric learning. The global deformation space is divided into several local subspaces, and a local Riemannian metric is learned in each of these subspaces. At treatment time it first determines into which subspace the deformation parameters fall, and it then interpolates the deformation parameters within the subspace using the local metric and local training deformation parameters. Local metric learning makes REALMS more accurate by fitting a better linear relationship between the projection differences and the parameter differences in each subspace to yield a good local metric.

In this paper, we investigate this localized REALMS with several abdominal IGRT cases. Registration with abdominal images is more challenging than in lung for two reasons. First, in addition to respiratory deformations during the patient's breathing cycle, there are other deformations in the abdomen, such as digestive deformations, between planning time and treatment time, which makes the learned metric inappropriate for treatment-time registration. Second, the formation of the deformation space depends on accurate deformable 3D/3D registration among planning-time RCCTs (Respiratory-Correlated CTs), but a challenge is the low intensity contrast in the abdomen. We propose several methods in this paper that show promise in dealing with these problems. To our knowledge, this study represents the first attempt at 2D/3D deformable registration with abdominal image sets.

The rest of the paper is organized as follows: In section 2, we discuss some of the background research work (alternative methods) for 2D/3D registration. In section 3, we describe the interpolative scheme and metric learning in the REALMS framework. In section 4, we introduce the localized approach to make REALMS more accurate. In section 5, we describe some specific situations for localized REALMS in the abdomen, including an updated deformation model for the composition of respiratory deformation and digestive deformation. Finally in section 6, we discuss the results of synthetic tests and real tests on abdominal cases.

Q. Zhao, C. Chou and S. Pizer are with the Computer Science Department at the University of Carolina, Chapel Hill.

G. Mageras is with Memorial Sloan-Kettering Cancer Center.

For Review Only

II. Related Work

A number of 2D/3D registration methods [3], [4], [5], [6] were designed to optimize over a 2D/3D geometric transformation that minimize an objective function consisting of an image mismatch term and a regularization term. Traditional optimization-based methods are computationally expensive due to the evaluation of the function's Jacobian in each optimization iteration. With GPU parallelization recent optimization-based 2D/3D registration methods [7], [8] are able to localize the tumor within 1s assuming rigid patient motion. For non-rigid motion in lung and abdomen, in order to lower the number of deformation parameters and produce reasonable deformations, a common approach is to adopt a deformation model based on Principal Component Analysis (PCA) [9] from the patient's respiration-correlated CT (RCCT), consisting of a set of (typically ten) 3D CT images over one breathing cycle. The PCA model can efficiently represent the deformation with only a few deformation parameters and corresponding eigenvectors. Li et al. [10], [11] used a gradient-descent optimization scheme on GPUs to find the optimal PCA scores that minimized the difference between DRRs (Digitally-Reconstructed Radiographs) and on-board projections. However, the image mismatch term often introduces a non-convex objective function which can be trapped in local minima. In order to avoid local minima and to reduce the registration time, a bootstrap-like approach [10], [11] was adopted, and the optimizations were initialized by registration results from previous time points.

Other methods have used neural networks to model rigid [12], [13], [14], or non-rigid transformations [15] and to achieve efficient computation at registration time. However, they can not support both rigid and non-rigid 2D/3D registration.

Chou et al. [16], [17] recently proposed a regression-based approach, CLARET, to estimate deformation parameters. CLARET first learns a linear regression between projection intensities and deformation parameters at planning time. Then at treatment time the learned regression is iteratively applied to refine the deformation parameters. However, CLARET still involves computationally demanding production of DRRs in each registration iteration. Reference [18] introduced a localized version of CLARET that could avoid generating DRRs at treatment time. It uses graph cuts to separate the whole deformation space and learns a linear regression independently in each division. At treatment time it uses a decision forest to classify a target projection into a partition, and the learned regression for that training partition is non-iteratively applied to the target projection image. Even though localized CLARET achieves a large improvement in registration speed at treatment time, it is still slower than REALMS due to the voting procedure of the decision forest in each registration.

III. REALMS Framework

A. Deformation Modeling at Planning Time

As described in [2], REALMS first models the deformation with a shape space, in which each deformation is formed by a linear combination of basis deformations calculated through PCA analysis. For deformations due to respiration, we use a set of RCCT images $\{J_\tau|\tau = 1, 2, ..., 10\}$ of the patient at planning time that records a cyclically varying target area. Chou et al. [2] computed a Fréchet mean image $\bar{J}$ as well as the diffeomorphic deformations ϕ from $\bar{J}$ to each image J_τ and performed a statistical analysis based on the 10 deformations. Here we use an alternative approach, in which a reference image I_{ref} is chosen from the 10 RCCTs and a set of deformations ϕ between the other 9 images $\{J_\tau|\tau = 1, 2, ..., 9\}$ and I_{ref} are computed via an LDDMM (Large Deformation Diffeomorphic Metric Mapping) framework described in Lorenzen et al. [19]. I_{ref} is chosen as the closest phase to the average phase of the breathing cycle represented by the RCCTs according to the position of a fiducial marker implanted in the patient. Using the calculated diffeomorphic deformation set $\{\phi_\tau|\tau = 1, 2, ..., 9\}$, our method finds a set of linear deformation basis vectors ϕ^i_{pc} by PCA analysis. The scores c^i_τ on each ϕ^i_{pc} yield ϕ_τ in terms of these basis vectors.

$$T(\mathbf{c}_\tau) = \phi_\tau = \overline{\phi} + \sum_{i=1}^{9} c^i_\tau \cdot \phi^i_{pc} \tag{1}$$

We have found that 3 eigenmodes are sufficient to capture more than 95% of the total variation. Then we let the first 3 scores form the 3-dimensional parametrization $\mathbf{c} = (c^1, c^2, c^3)$. This 3-tuple $\mathbf{c}$ forms a 3D parameter space (also referred to as the deformation space or the shape space).

B. Treatment-time Registration

At treatment time, REALMS uses kernel regression (eq. 2) to interpolate the patient's three 3D deformation parameters $\mathbf{c} = (c^1, c^2, c^3)$ separately from the on-board projection image $\Psi(\theta)$. Each parameter is interpolated as a weighted sum of N deformation parameters from a set of N training projection images with training $\mathbf{c}$ values $\{\mathbf{P}(I_{ref} \circ T(\mathbf{c}_\gamma); \theta)|\gamma = 1, 2, ...N\}$ simulated at planning time. P simulates the DRRs with deformation parameters $\mathbf{c}_\gamma$ according to the treatment-time imaging geometry, e.g., the projection angle θ.

In the treatment-time registration, each deformation parameter c^i in $\mathbf{c}$ can be estimated with the following kernel regression:

$$c^i = \frac{1}{K} \sum_{\gamma=1}^{N} c^i_\gamma \cdot e^{\frac{-d^2_{i,\gamma}}{(\beta^i)^2}} \tag{2}$$

where K is used for normalizing the overall weight, $d^2_{i,\gamma}$ denotes the squared distance between the on-board projection image and $\mathbf{P}(I_{ref} \circ T(\mathbf{c}_\gamma); \theta)$. Moreover, we use a Riemannian metric tensor $\mathbf{M}^i$ to determine $d^2_{i,\gamma}$ using eq. 4.

$$\Delta I = \Psi(\theta) - \mathbf{P}(I_{ref} \circ T(\mathbf{c}_\gamma); \theta) \tag{3}$$

$$d^2_{i,\gamma} = \Delta I^T M^i \Delta I \tag{4}$$

C. Metric Learning at Treatment-planning Time

At treatment-planning time, REALMS learns a metric tensor $\mathbf{M}^i$ with a corresponding kernel width β^i for the patient's ith deformation parameter c^i. To limit the number of degrees of freedom in the metric tensor, we structure $\mathbf{M}^i$ as a rank-1

For Peer Review

matrix formed by a basis vector a^i: $\mathbf{M}^i = a^i a^{i^T}$. We assume that this basis vector presents a linear relationship between the intensity difference between two training DRRs and their parameters' distance in the deformation space (eq. 6). Therefore, REALMS learns a^i by applying a linear regression (eq. 7) between the projection difference matrix $R = (r_1, r_2, ...r_N)^T$ and the parameter difference matrix $\Delta C^i = (c_1^i, c_2^i, ..., c_N^i)^T$. In particular, the projection difference vector is

$$r_\gamma = \mathbf{P}(I_{ref} \circ T(\mathbf{c}_\gamma); \theta) - \mathbf{P}(I_{ref}; \theta) \quad (5)$$

$$\Delta C^i = R \cdot a^i \quad (6)$$

$$a^i = (R^T R)^{-1} R^T \Delta C^i \quad (7)$$

In such a case, the squared distance function $d^2_{i,\gamma}$ denotes the weighted squared distance between two projection images (eq. 8).

$$d^2_{i,\gamma} = \Delta I^T M^i \Delta I = (a^{i^T} \cdot \Delta I)^T (a^{i^T} \cdot \Delta I) \approx (c^i - c^i_\gamma)^2 \quad (8)$$

Chou et al. have pointed out that the basis vector a^i can be refined by using a Leave-One-Out (LOO) training strategy after the linear regression. In our work, the learned regression is directly used as the basis vector. We select the optimal kernel width β^i from a number of candidates that minimizes synthetic tests' error (described in 6.1).

IV. Localized REALMS

A. Localized Interpolation Strategy

The accuracy of kernel regression highly depends on the quality of the learned metric $\mathbf{M}^i$. Traditional (global) REALMS uses linear regression to learn the underlying basis vector a^i. Therefore, a^i uses a linear model to approximate the relationship between intensity difference and parameter difference in the overall shape space. Usually this approximation is not sufficiently accurate because the linear property does not hold for the entire shape space. However, if the basis vector is tailored for a particular region of the shape space, the approximation will be more accurate in that region. Moreover, the patient's 3D deformation parameter $\mathbf{c}$ is more related to nearby parameter values in the shape space. Therefore, we introduce a localized interpolation method based on REALMS. Localized REALMS will place the N training deformation parameters closer to the target parameters' position, so that both a robust linear regression and an accurate interpolation will be achieved in that local region.

In the localized REALMS approach, we first divide the global shape space $\mathbf{S}^0$ into several first level subspaces $\mathbf{S}^1_k$ $(k = 1, 2, ..., S)$, where k denotes the k^{th} subspace, and S is the number of subspaces. At treatment-planning time, each subspace has its own N deformation parameters sampled and corresponding DRRs generated. Then a local metric $\mathbf{M}^i_k$ is learned for the k^{th} subspace. $\mathbf{M}^i_k$ is still decomposed by a basis vector a^i_k, which is learned by a local linear regression from its own deformation parameters and their corresponding DRRs. The treatment-time registration first decides to which subspace the deformation parameters $\mathbf{c}$ belong. If $\mathbf{c}$ locates in $\mathbf{S}^1_k$, an interpolation is applied using its local metric $\mathbf{M}^i_k$ and local training DRRs to yield a finer result. (Fig. 1)

B. Subspace Division and Selection

There are many ways to divide global space into subspaces. Chou et al. [18] used normalized graph cuts to separate the shape space and trained a decision forest for deciding into which subspace a target projection image should be classified. In our work, we uniformly divide the global shape space by considering that the global shape space can be viewed as a 3 dimensional box. The length in each dimension is $6\sigma_i$ (from $-3\sigma_i$ to $3\sigma_i$), where σ_i is the standard deviation of the scores on the principal component ϕ^i_{pc}. We take 3 subintervals in each dimension: $-3\sigma_i$ to $0\sigma_i$, $-1.5\sigma_i$ to $1.5\sigma_i$, $0\sigma_i$ to $3\sigma_i$. The combinations of subintervals of 3 dimensions yield 27 overall subspaces. Each of these subspaces is also a 3D box, where the length in each dimension is half of the global shape space. Fig. 2 shows an example of a 2D shape space constructed by the first two principal components of deformation. In this 2D case, there are 9 subspaces in total. Each of them is a 2D square. Red and green squares are two examples subspaces.

To choose the subspace to which the target parameters belong, we use global REALMS to compute a first round interpolation over the whole space. The estimated parameter values indicate a position in the global 3D shape space. Then, localized REALMS chooses the subspace with its centre closest to the initial estimated position. (Fig. 2)

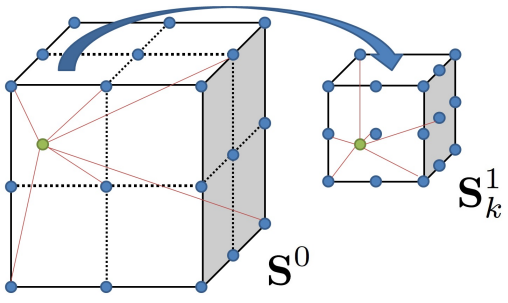


Fig. 1. Global REALMS first determines the subspace to which the parameters belong. Next, localized REALMS does a finer interpolation using local training parameters and the local metric. Green dots denote target parameters. Blue dots denote training parameters.

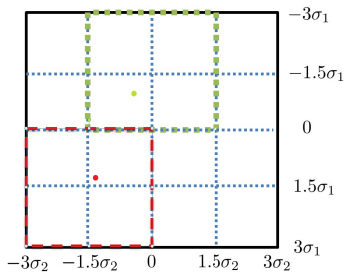


Fig. 2. Subspaces division in 2D: The global shape (shown as black) space consisting of the first two principal components is a 2D square. The edges of all subspaces are along black and blue lines. Global REALMS determines the initial position of deformation parameters. Localized REALMS chooses the corresponding subspace. Two examples are shown as red and green.

For Peer Review

Localized REALMS can be extended to a multi-scale approach, in which the level of localization can be added if an i^{th} level subspace $\mathbf{S}_k^i$ is divided into several $(i+1)^{th}$ level subspaces $\mathbf{S}_{k'}^{i+1}$ $(k' = 1, 2, ..., N')$. In this case, a metric tensor should be learned respectively for every subspace on every level. The interpolation determined in $\mathbf{S}_k^i$ will identify which subspace $\mathbf{S}_{k'}^{i+1}$ to choose on the next level for applying a further local interpolation.

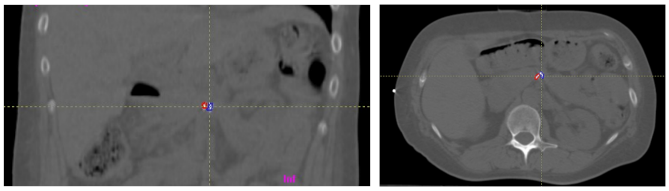
V. Localized REALMS in the Abdomen

The findings by Chou et al. [2] indicate that REALMS performs reasonably well in lung. However, there are additional challenges posed by registration in abdomen. We describe these special problems and propose several methods to address them.

A. Non-respiratory Deformation

In lung IGRT, the deformation between treatment time and planning time images is almost entirely respiratory deformation. However in the abdomen, there also exists non-respiratory deformation caused by the digestive system. Changes in filling of the stomach, duodenum and bowel will cause deformation of these organs. There can also be changes of gas filling in these same organs. These factors will induce non-respiratory deformations of nearby target organs, such as the pancreas. Therefore in the abdomen, the digestive process may deform the target organ to different positions for each treatment, and respiration will impose further cyclic deformations. However, since the learning stage and the interpolation stage in REALMS are both based on the intensity difference caused only by respiratory deformation, REALMS is not credible for finding these non-respiratory deformations. In other words, REALMS alone only learns and produces deformations modeled by the parameter space formed from the patient's respiratory variation.

To deal with this problem, we first assume that only respiratory deformation occurs during treatment and that non-respiratory deformation only happens between planning time and treatment time. In the long term, we can address this problem by taking a new reference CT at the beginning of treatment and apply the old deformation space based on the new reference. However, for the time being, we act as if the non-respiratory deformation is a 3D rigid translation vector with respect to the abdominal region around the target organ. The rationale is to first estimate a translation vector at the beginning of treatment time registration, and we translate the planning reference image I_{ref} as well as the deformation principal components $(\{\phi_{pc}^i | i = 1, 2, .., 9\}, \overline{\phi})$ by the vector to compensate for non-respiratory deformation. Following rigid translation correction, the remaining deformations at treatment time with respect to the new I_{ref} are all respiratory deformations and can be determined by REALMS. Therefore, at treatment registration, the intensity difference ΔI is still generated using eq. 5, where I_{ref} now is the translated reference image, and $T(\mathbf{c}_\gamma)$ generates deformations from the translated deformation principal components.



(a) (b)

Fig. 3. Two fiducial markers delineated by red and blue outlines. (a) A coronal slice of the reference CT. (b) An axial slice of the reference CT.

We make use of fiducial markers to compute the rigid translation vector. These are small radiopaque markers (Visicoil, RadioMed Corp., Bartlett, TN), 20mm long $\times$ 1mm diameter, which are implanted in or near the target abdominal organ at planning time (Fig. 3). The markers curl up and take on irregular shapes upon implantation. Their high contrast with respect to surrounding tissues allows them to be easily located in CT, projection images and even in coarse CT reconstructions from projection images. Therefore, we can compute the translation vector from the positional difference of a fiducial marker between planning time and treatment time.

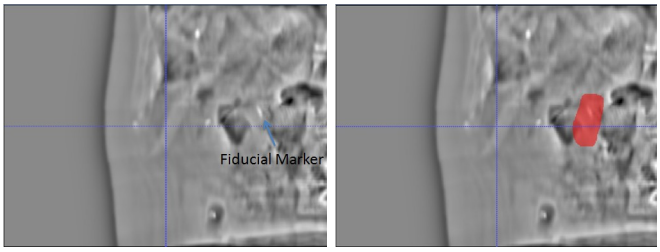
In our application, the treatment time images are obtained with a kilovoltage cone-beam CT (CBCT) system [20] mounted on a medical linear accelerator (TrueBeam, Varian Medical Systems, Palo Alto, CA). Immediately prior to the start of radiation treatment, a CBCT scan is acquired, consisting of 660 projection images over a 360 arc around the patient while simultaneously recording respiration with a position monitor placed on the patients abdomen (Real-time Position Management System, Varian Medical Systems). The projection images are sorted into 10 respiratory bins and each bin is reconstructed [21] (Varian iTools version 1.0.32.0), yielding a sequence of 10 preRT-CBCT images over the patients respiratory cycle. We extract fiducial positions from each of these reconstructions and compute a mean fiducial marker position for treatment time. In addition, we locate the fiducial positions in the 10 RCCT images and compute a mean fiducial position at planning time. Finally, the translation vector is the difference between the two mean fiducial positions.

B. Contrast Issues

Because soft tissue contrast in CT images of the abdomen has lower contrast than in lung, it is more challenging to achieve accurate deformable registration between RCCT images, which in turn affects the accuracy of the deformation space. To address this problem, the LDDMM 3D/3D registration uses a narrower intensity window (-100 to +100 HU), which encompasses the soft tissue intensity range of abdominal structures.

The robustness of REALMS highly depends on the intensity difference information between DRR and projection image caused by deformation. However there are two factors which adversely affect projection images in abdomen. First, because of the similar intensities of soft tissue organs, it is not possible to reliably localize them in the projection images or determine their deformations. Second, movement of gas pockets in the abdomen induces variable high contrast intensity patterns

For Peer Review



(a) (b)

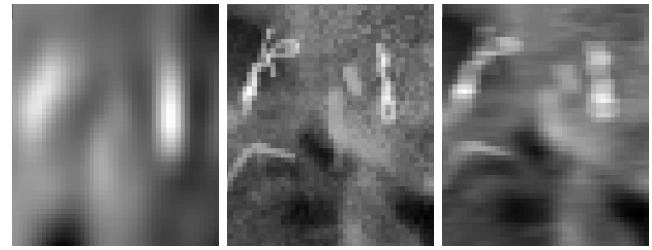
Fig. 4. (a) A DRR generated from the mean CT (deformed from the reference CT by $\mathbf{c} = (0, 0, 0)$). (b) Red area indicates the region-of-interest for computing intensity differences between DRR and projection image.

in the projection images, which yields inconsistent intensity differences between planning time and treatment time. Our solution is to focus our region of interest (ROI) on one or more fiducial markers visible in the images (Fig. 4). Because the fiducial is usually very distinct in the images, and soft tissue deformation is implied by its displacement, we use the intensity change in a small region around the fiducial to indicate the overall deformation in the surrounding soft tissue organs. In our experiments, we used a 4cm $\times$ 6cm square ROI including the fiducial marker, and perform REALMS learning and interpolation only using the pixels in that area.

C. DRRs vs. On-board Radiographs

REALMS requires comparable intensities between DRRs and real projection images. To account for variations caused by x-ray scatter that produces inconsistent projection intensities, we normalize both training DRRs and the on-board projection image $\Psi(\theta)$. In particular, we use the localized Gaussian normalization introduced in [22], which has shown promise in removing the undesired scattering artifacts. In addition, the histogram matching scheme described in [16] is used to produce intensity-consistent radiographs.

Due to the different image resolutions, a fiducial marker usually has very distinct appearances in DDRs and on-board projections. DRRs are simulated from CTs, in which the optical resolution (the minimum distance at which two fiducial markers can be distinguished in the image) is approximately 3mm, while an on-board radiograph has an approximate optical resolution of 1mm in the projection plane (measured manually over 3 patient datasets). On the other hand, CT has an anisotropic pixel size of 1mm$\times$1mm$\times$2mm (2mm in axial direction), while a radiograph has an approximate pixel size of 0.3mm$\times$0.3mm. Consequently the fiducial markers appear more blurred in DRRs than in on-board projection images. We use a simple blurring technique on the latter to yield an image with comparable resolution to the DRR (Fig. 5). We convolve the on-board projection image with an anisotropic mean kernel with η_y rows and η_x columns:



(a) (b) (c)

Fig. 5. (a) Fiducial appearance in a DRR. (b) Fiducial appearance in a cone-beam projection before convolution. (c) Fiducial appearance in a cone-beam projection after convolution with a blurring kernel.

$$\begin{pmatrix} \frac{1}{\eta_y \times \eta_x} & \frac{1}{\eta_y \times \eta_x} & \cdots \\ \frac{1}{\eta_y \times \eta_x} & \cdots & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix}$$

where η_x and η_y is the blurring width in each dimension. For each dimension, the blurring width is the product of the ratio between optical resolutions and the ratio between pixel sizes in that dimension.

In our case:

$$\eta_x = \frac{3mm}{1mm} \times \frac{1mm}{0.3mm} \tag{12}$$

$$\eta_y = \frac{3mm}{1mm} \times \frac{2mm}{0.3mm} \tag{13}$$

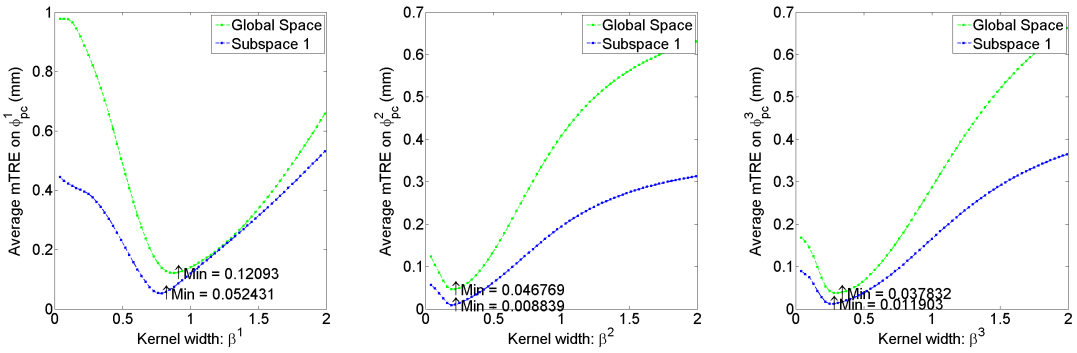
For computational convenience with a discretized image, we set $\eta_x = 9$, and $\eta_y = 15$.

VI. Results

A. Evaluation Measurement

We tested localized REALMS on 3 patients datasets in abdomen. In each dataset, segmentations of the pancreas, duodenum and fiducial markers were drawn on the reference RCCT image by a medical physicist experienced in radiation treatment planning. In addition, each patient dataset included a reconstructed CBCT at end expiration (EE) with corresponding segmentations drawn by the same observer. The CBCT at end expiration was acquired by respiration gating of the on-board imaging system, such that rotation of the gantry and acquisition of projection images occurred within a gate centered at expiration and encompassing approximately 25% of the respiratory cycle [20]. The 2D/3D registration is evaluated by first determining the deformation parameters of the reference CT and then deforming the reference segmentation by these parameters. The quality of the registration is measured between the estimated segmentation and the target (manual or synthetically deformed) segmentation. The evaluation tests made use of two types of measurement. When the ground truth deformation is known for every voxel of the target segmentation, such as in the synthetic data tests (section 6.2), we use mTRE (mean Target Registration Error) to evaluate the average deformation error over all the voxels between the estimated and target segmentation. When the ground truth deformation of the target is not known, such as in tests of the actual data (section 6.3), we alternatively measure the registration error by the magnitude of the 3D target centroid difference (TCD), which is the centroid distance between two segmentations. Note that TCD may overestimate the error

For Peer Review



(a) (b) (c)

Fig. 6. Average mTREs over 30 test cases induced by the interpolation error of the (a) first, (b) second, (c) third deformation parameter versus different kernel widths. Green: 30 test cases sampled in the global parameter space and interpolated from global training images. Blue: 30 test cases sampled in the first subspace and interpolated from local training images.

Dataset#	REALMS Method	ϕ_{pc}^1 (mm)	ϕ_{pc}^2 (mm)	ϕ_{pc}^3 (mm)	Overall (mm)
1	Global	0.110 ± 0.081	0.027 ± 0.024	0.036 ± 0.032	0.101 ± 0.078
1	Localized	0.037 ± 0.030	0.009 ± 0.007	0.011 ± 0.009	0.032 ± 0.027
2	Global	0.275 ± 0.246	0.215 ± 0.164	0.174 ± 0.178	0.350 ± 0.278
2	Localized	0.069 ± 0.063	0.030 ± 0.026	0.036 ± 0.075	0.080 ± 0.088
3	Global	0.210 ± 0.193	0.062 ± 0.061	0.070 ± 0.073	0.101 ± 0.093
3	Localized	0.048 ± 0.042	0.012 ± 0.012	0.011 ± 0.012	0.023 ± 0.022

TABLE I
AVERAGE MTRES OVER 128 TEST CASES PROJECTED ONTO THE EACH DEFORMATION PRINCIPAL COMPONENT.

For Peer Review

because of inconsistencies in manual segmentations between reference CT and target CBCT.

B. Synthetic Data Tests

The tests of synthetic data made use of the RCCT datasets. As stated in section 4.2, the global shape space was divided into 27 subspaces, from which 343 training images were generated for the global shape space and each subspace. For each patient, an ROI was independently selected based on the fiducial position and the projection angle, which can best reveal the fiducial marker in the projection images. DRRs (dimension: 512×384) were generated from synthetically deformed target CTs and were used to represent on-board cone-beam projection images. The target CTs were synthetically deformed from the reference CT I_{ref} by normally distributed random samples of the 3-dimensional deformation parameters $\mathbf{c} = (c^1, c^2, c^3)$. The quality of the registration was measured by using the average mTRE of the target organ (duodenum) over all the test cases.

We first investigated whether interpolation in a local region yielded more accurate registration. Thirty DRRs were tested for each subspace as well as the global space. As shown in Fig. 6, interpolation in a subspace yielded a lower mTRE than did global interpolation for each candidate kernel width.

Next, the kernel width resulting in the least mTRE was selected separately for each subspace and global space, and used to test global REALMS and localized REALMS on an independent set of 128 images sampled from the global shape space. As shown in Table I, localized REALMS reduced the global REALMS registration error by an average of 74.2%. Further, global REALMS showed a 97% success rate in choosing the correct subspace for the local interpolation.

C. Actual Data Tests

Tests of actual data consisted of registration of the reference CT with actual on-board cone-beam projection images at the EE (End-Expiration) phase in the 3 patient datasets. The rigid translation and resolution normalization described in Section 5 were carried out prior to deformable registration. The same training methodology was used as in the synthetic data tests. The learned metrics and the selected kernel widths were used to estimate deformation parameters. The quality of the registration was measured by the magnitude of 3D Target Centroid Differences (TCDs) of the target organ (duodenum) between the REALMS-estimated segmentation and the manual segmentations in the reconstructed CBCT at end expiration.

As shown in Table II, localized REALMS outperformed global REALMS in all cases. Localized REALMS reduced the registration error of global REALMS by an average of 30% in average. The TCDs in the first two patient cases were approximately 2mm, consistent with the accuracy needed for

Dataset#	Initial (mm)	Global (mm)	Localized (mm)	Time (ms)
1	6.901	6.826	2.475	25.83
2	7.321	2.58	1.961	25.93
3	6.797	6.216	6.098	34.01

TABLE II
MAGNITUDE OF TARGET CENTROID DIFFERENCES OF THE DUODENUM BEFORE AND AFTER REGISTRATION FOR 3 PATIENT DATASETS.

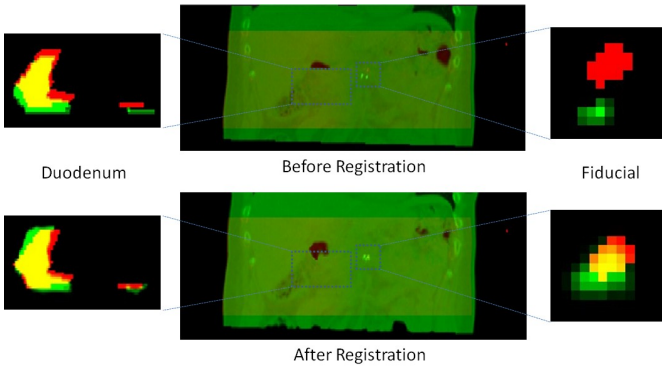


Fig. 7. Top left: Image overlay of a coronal slice of the duodenum segmentation before registration. Segmentation from the reference CT is in green and from reconstructed CBCT in red. Yellow is the overlapping region. Bottom left: Green denotes the duodenum segmentation deformed by the estimated deformation parameters. Top center: Image overlays of the reference CT (green) and the reconstructed CBCT (red). Bottom center: Images overlays of the estimated CT (green) and the reconstructed CBCT (red). Top and bottom right: Coronal image overlays of the fiducial marker segmentation.

clinical application.

Fig. 7 shows an example (dataset 2) image overlays of the reference CT and target (CBCT) segmentations. The duodenum and the fiducial marker are more closely aligned following registration. We note that in dataset 3, localized REALMS achieves only a small improvement in TCD of the duodenum (Table II). Inspection of the fiducial marker following registration (Fig. 8) shows that it is correctly aligned (TCD: 1.19mm) whereas the duodenum alignment is improved only in the superior-inferior direction (i.e. vertically in the images). This indicates that the deformation shape space derived from the RCCT has captured only part of the remaining deformation in the CBCT after rigid correction. There remains a significant displacement of the duodenum in the patients left-right direction (5.8mm) which may have been induced by the digestive deformation, e.g. caused as changes in stomach or bowel filling. In this particular dataset, the rigid

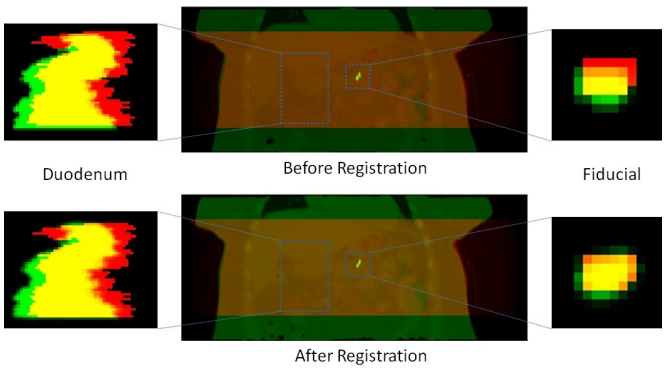


Fig. 8. Top left: Image overlay of a coronal slice of the duodenum segmentation before registration. Segmentation from the reference CT is in green and from reconstructed CBCT in red. Yellow is the overlapping region. Bottom left: Green denotes the duodenum segmentation deformed by the estimated deformation parameters. Top center: Image overlays of the reference CT (green) and the reconstructed CBCT (red). Bottom center: Images overlays of the estimated CT (green) and the reconstructed CBCT (red). Top and bottom right: Coronal image overlays of the fiducial marker segmentation.

translation based on mean fiducial position difference may not be sufficiently accurate for modeling duodenal deformation. The duodenum is the portion of the small bowel that connects to the stomach, whereas the fiducial marker is implanted in the pancreas, an organ adjacent to the duodenum. Inspection of the reference CT and CBCT images showed a larger amount of stomach and bowel gas in in the CBCT than in the CT (data not shown). Since the fiducial marker is approximately 5 cm distant from the duodenum and adjacent to a different part of the stomach, it is likely subject to different gas-induced deformations than is the duodenum. We did not attempt to incorporate such deformations into the REALMS shape space in this study.

VII. Conclusions

In this work, we have presented a new algorithm for improving the accuracy of a previously developed 2D/3D deformable registration method, called REALMS. Rather than carrying out a kernel regression with a globally trained metric as in the prior method, the improved algorithm divides the deformation space into several subspaces and learns a local metric in each subspace. After determining the subspace to which the deformation parameters belong, each parameter is interpolated using the local metric from local training samples. We evaluated the performance of our proposed method in patient abdominal image sets, which posed additional challenges relative to previously test image sets such as lung, caused by low soft tissue contrast in the abdominal images and the presence of non-respiratory deformations caused by changes in the contents of digestive organs. Several methods were introduced for processing abdominal images. First, to account for digestive changes, a 3D rigid translation was applied between the planning time and treatment time images by aligning to the respiration-averaged position of an implanted fiducial marker as a surrogate for soft tissue positions. Second, to facilitate 2D/3D deformable registration in low-contrast projection images, differences between the DRR and target projection images was computed within an ROI around the fiducial. Third, a convolution of projection images with a blurring kernel was applied to yield fiducial appearances consistent with those in the lower-resolution DRRs. Evaluation of synthetic data and actual data show that the proposed localized method improves registration accuracy when accounting for respiratory deformation alone. Two of the three actual data cases demonstrate the methods potential utility for clinical use in abdominal IGRT. However, the methods accuracy can be limited when the proposed fiducial-based 3D rigid translation is insufficient to approximate the digestive deformation. Future work will include modeling of the digestive deformation between planning time and treatment time, updating the deformation space and the reference image at treatment setup time and evaluating the method on more abdominal patient datasets.

VIII. Acknowledgements

This work was partially supported by National Cancer Institute Award Numbers R01-CA126993 and R01-CA126993-02S1. The content is solely the responsibility of the authors

For Peer Review

and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

REFERENCES

[1] P. Markelj, D. Tomazevic, B. Likar, and F. Pernus, "A review of 3D/2D registration methods for image-guided interventions," *Medical Image Analysis*, vol. 16, no. 3, pp. 642–661, 2012.

[2] C. Chou and S. Pizer, "Real-time 2D/3D deformable registration using metric learning," *MCV: Recognition Techniques and Applications in Medical Imaging*, vol. 7766, pp. 1–10, 2012.

[3] D. Russakoff, T. Rohlfing, and C. Maurer, "Fast intensity-based 2D-3D image registration of clinical data using light fields," *In Proc. 9th IEEE Int. Conf. Computer Vision*, vol. 1, pp. 416–422, 2003.

[4] D. Russakoff, T. Rohlfing, K. Mori, D. Rueckert, A. Ho, J. R. Adler, and C. R. Maurer, "Fast generation of digitally reconstructed radiographs using attenuation fields with application to 2D-3D image registration," *IEEE Trans. Med. Imag.*, vol. 24, no. 11, pp. 1141–1454, 2005.

[5] A. Khamene, P. Bloch, W. Wein, M. Svatos, and A. F. Sauer, "Automatic registration of portal images and volumetric CT for patient positioning in radiation therapy," *Medical Image Analysis*, vol. 10, no. 1, pp. 96–112, 2006.

[6] R. Munbodh, D. A. Jaffray, D. J. Moseley, Z. Chen, a. P. C. J. P. S. Knisely, and J. S. Duncan, "Automated 2D-3D registration of a radiograph and a cone beam CT using line-segment enhancement," *Medical Physics*, vol. 33, no. 5, pp. 1398–1411, 2006.

[7] H. Furtado, C. Gendrin, C. Bloch, J. Spoerk, S. A. Pawiro, C. Weber, M. Figl, M. Stock, D. Georg, H. Bergmann, and W. Birkfellner, "Real-time 2D/3D registration for tumor motion tracking during radiotherapy," *in Proc. SPIE 8314, Medical Imaging 2012*, 2012.

[8] C. Gendrin, H. Furtado, C. Weber, C. Bloch, M. Figl, S. A. Pawiro, H. Bergmann, M. Stock, G. Fichtinger, D. Georg, and W. Birkfellner, "Monitoring tumor motion by real time 2D/3D registration during radiotherapy," *Radiotherapy and oncology*, vol. 102, no. 2, pp. 274–280, 2012.

[9] Q. Zhang, A. Pevsner, A. Hertanto, Y. Hu, K. Rosenzweig, C. Ling, and G. Mageras, "A patient-specific respiratory model of anatomical motion for radiation treatment planning," *Medical Physics*, vol. 34, no. 12, pp. 4772–4781, 2007.

[10] R. Li, J. Lewis, X. Jia, X. Gu, M. Folkerts, C. Men, W. Song, and S. Jiang, "3D tumor localization through real-time volumetric x-ray imaging for lung cancer radiotherapy," *Medical Physics*, vol. 38, no. 5, pp. 2783–2794, 2011.

[11] R. Li, X. Jia, J. H. Lewis, X. Gu, M. Folkerts, C. Men, and S. B. Jiang, "Realtime volumetric image reconstruction and 3D tumor localization based on a single x-ray projection image for lung cancer radiotherapy," *Medical Physics*, vol. 37, no. 6, pp. 2822–2826, 2010.

[12] S. Banks and W. Hodge, "Accurate measurement of three-dimensional knee replacement kinematics using single-plane fluoroscopy," *IEEE Trans. Biomed. Eng.*, vol. 43, no. 6, pp. 638–649, 1996.

[13] L. Freire, A. Gouveia, and F. Godinho, "FMRI 3D registration based on fourier space subsets using neural networks," *in Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 5624–5627, 2010.

[14] J. Zhang, Y. Ge, S. H. Ong, C. K. Chui, S. H. Teoh, and C. H.Yan, "Rapid surface registration of 3D volumes using a neural network approach," *Image and Vision Computing*, vol. 26, no. 2, pp. 201–210, 2008.

[15] M. Wachowiak, R. Smolikova, J. Zurada, and A. Elmaghraby, "A supervised learning approach to landmark-based elastic biomedical image registration and interpolation," *in : Neural Networks, 2002. IJCNN '02. Proceedings of the 2002 International Joint Conference on*, vol. 2, pp. 1625–1630, 2002.

[16] C. Chou, B. Frederick, X. Liu, G. Mageras, S. Chang, and S. Pizer, "CLARET: A fast deformable registration method applied to lung radiation therapy," *Fourth International (MICCAI) Workshop on Pulmonary Image Analysis*, pp. 113–124, 2011.

[17] C. Chou, B. Frederick, G. Mageras, S. Chang, and S. Pizer, "2D/3D image registration using regression learning," *Computer Vision and Image Understanding*, vol. 117, no. 9, pp. 1095–1106, 2013.

[18] C. Chou and S. Pizer, "Local regression learning via forest classification for 2d/3d deformable registration," *in Proc. In MICCAI workshop on Medical Computer Vision (MCV)*, 2013.

[19] P. Lorenzen, M. Prastawa, B. Davis, G. Gerig, E. Bullitt, and S. Joshi, "Multi-modal image set registration and atlas formation," *Medical Image Analysis*, vol. 10, no. 3, pp. 440–451, 2006.

[20] R. E. Kincaid, E. D. Yorke, K. A. Goodman, A. Rimner, A. J. Wu, and G. S. Mageras, "Investigation of gated cone-beam CT to reduce respiratory motion blurring," *Medical Physics*, vol. 40, no. 7, 2013.

[21] L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *J. Opt. Soc. Amer. A*, vol. 1, no. 6, pp. 612–619, 1984.

[22] P. Cachier and X. Pennec, "3D non-rigid registration by gradient descent on a gaussianwindowed similarity measure using convolutions," *in Proc. Mathematical Methods in Biomedical Image Analysis*, pp. 182–189, 2000.

For Peer Review