Deformable mesh registration for the validation of automatic target localization algorithms

Scott Robertson, Elisabeth Weiss, and Geoffrey D. Hugo^{a)}

Department of Radiation Oncology, Virginia Commonwealth University, Richmond, Virginia 23298

(Received 20 December 2012; revised 17 May 2013; accepted for publication 29 May 2013; published 18 June 2013)

Purpose: To evaluate deformable mesh registration (DMR) as a tool for validating automatic target registration algorithms used during image-guided radiation therapy.

Methods: DMR was implemented in a hierarchical model, with rigid, affine, and B-spline transforms optimized in succession to register a pair of surface meshes. The gross tumor volumes (primary tumor and involved lymph nodes) were contoured by a physician on weekly CT scans in a cohort of lung cancer patients and converted to surface meshes. The meshes from weekly CT images were registered to the mesh from the planning CT, and the resulting registered meshes were compared with the delineated surfaces. Known deformations were also applied to the meshes, followed by mesh registration to recover the known deformation. Mesh registration accuracy was assessed at the mesh surface by computing the symmetric surface distance (SSD) between vertices of each registered mesh pair. Mesh registration quality in regions within 5 mm of the mesh surface was evaluated with respect to a high quality deformable image registration.

Results: For 18 patients presenting with a total of 19 primary lung tumors and 24 lymph node targets, the SSD averaged 1.3 ± 0.5 and 0.8 ± 0.2 mm, respectively. Vertex registration errors (VRE) relative to the applied known deformation were 0.8 ± 0.7 and 0.2 ± 0.3 mm for the primary tumor and lymph nodes, respectively. Inside the mesh surface, corresponding average VRE ranged from 0.6 to 0.9 and 0.2 to 0.9 mm, respectively. Outside the mesh surface, average VRE ranged from 0.7 to 1.8 and 0.2 to 1.4 mm. The magnitude of errors generally increased with increasing distance away from the mesh.

Conclusions: Provided that delineated surfaces are available, deformable mesh registration is an accurate and reliable method for obtaining a reference registration to validate automatic target registration algorithms for image-guided radiation therapy, specifically in regions on or near the target surfaces. © 2013 American Association of Physicists in Medicine. [<http://dx.doi.org/10.1118/1.4811105>]

Key words: image-guided radiation therapy, deformable registration, localization, lung cancer

I. INTRODUCTION

Image-guided radiotherapy has facilitated substantial improvements in target localization accuracy.¹ Many studies have explored the use of direct target alignment to reduce treatment margins and improve the overall precision of radiotherapy delivery.^{2–4} For lung cancer alone, several groups have demonstrated the feasibility and benefit of automatic or semiautomatic soft-tissue-based rigid registration for localizing primary tumors.^{5,6} The ability of these registration algorithms to provide fast, reproducible target alignment makes them attractive for patient positioning. However, large anatomical changes during treatment and the limited image quality of onboard imaging systems can impact the accuracy of automated rigid target registration.⁷ As a result, interest in deformable and semideformable registration during image-guided radiation therapy is mounting, with efforts including the multiple region of interest registration technique for head-and-neck tumors⁸ and model-based registration for liver tumors.⁹ In locally-advanced non-small cell lung cancer, targets are often dynamic, demonstrating substantial shape and volume change throughout conventionally fractionated radiotherapy.^{10,11} Development and implementation of deformable registration to guide online tumor localization would also be warranted in this site.

During algorithm development and before clinically implementing deformable registration algorithms for the purpose of target localization, an important preliminary step involves their validation. This includes establishing residual error that can be used to determine appropriate margins. A tumor is considered to be well-localized during treatment if its surface can be positioned inside the same target from the planning CT. Validation must, therefore, be capable of identifying potential misalignments in the tumor surface resulting from automatic registration. A common strategy involves physician delineation of targets on both planning and localization images. By propagating the target contour through registration from one image to the other, the propagated and delineated contours can be compared as an assessment of registration quality.

An appropriate measure must be selected for comparing propagated and delineated contours that provides information on the accuracy of the registration algorithm. In this application, strategies such as volume overlap³ are not useful for comparing the contours, as they provide no information on spatial error that can be used for margin formulation. Surface border error^{12,13} and other distance-to-surface methods provide a measure of spatial accuracy. However, because manually delineated contours lack point-to-point correspondence

FIG. 1. Starting with the initially unregistered meshes, deformable mesh registration proceeds through three sequential steps: (1) a rigid similarity transform that includes an isotropic scaling parameter, (2) an affine transform, and (3) a B-spline deformable transform.

information, distance-to-surface metrics measure radial errors between surfaces, and cannot identify exact distance between corresponding points on the surfaces (see Fig. 6, for an example). Identification of a distribution of corresponding intensity features or landmarks between the two registered images may solve this problem, but seems superfluous if the physician-delineated target contours are already available on both images. Furthermore, direct comparison of surfaces with a distance-to-surface metric results in high sensitivity to the high-frequency component of manual delineation error. For example, delineation on axial slices is known to result in substantial slice-to-slice variability, which leads to rough surface contours. In the case of perfect registration, the rough surface contours would result in the distance-to-surface metric being on the order of the slice-to-slice contour variability. This issue makes it difficult for distance-to-surface metrics to distinguish between registration accuracy and delineation error.

To solve these problems, we propose to use deformable mesh registration (DMR) algorithm to align surface meshes from physician-delineated targets, and to parameterize the resulting transform with a cubic B-spline algorithm. This basis was selected (1) so that the transform can be estimated near as well as on the surface, and (2) so that the coarseness of the B-spline grid can be controlled to reduce the effect of manual delineation error on the registration. The primary purpose of this study is to evaluate the accuracy of an in-house DMR algorithm using meshes with known vertex-to-vertex correspondence. The accuracy of this tool is assessed by comparison against a known transform from intensity-based deformable image registration (DIR). Evaluation is conducted for the target surface as well as for regions near the surface. We also assess the ability of DMR to align physician-delineated surface meshes where the vertex-to-vertex correspondence is unknown.

II. METHODS

A deformable mesh registration tool was developed for the nonrigid alignment of a pair of surface meshes. The accuracy of this tool was assessed using physician-delineated tar-

gets as well as deformably propagated targets generated from a known transform. Deformable mesh registration was then compared against this known deformation to evaluate the registration accuracy both on and near the mesh surface.

II.A. Deformable mesh registration

An in-house DMR tool was implemented using the Insight Segmentation and Registration Toolkit (Ref. 14) (v3.20) to consist of three transforms optimized sequentially, as shown in Fig. 1. First, a rigid similarity transform provided translational and rotational degrees of freedom in addition to a single uniform scaling parameter. This scaling promoted a more robust initial alignment for regressing target structures than rigid registration without scaling. The next step involved an affine transform with three-dimensional scaling and shearing parameters in addition to any further translational and rotational corrections. Finally, a deformable B-spline transform fine-tuned the registration and yielded the final mesh alignment.

The similarity between reference and moving meshes was computed using the symmetric Euclidean distance. Briefly, the magnitude of all nearest-neighbor distances (NND) was calculated in both directions: from reference mesh vertices to the moving mesh and from moving mesh vertices to the reference mesh. This accounted for the orientation of both surfaces relative to each other, as opposed to the unidirectional Euclidean distance commonly employed in point set registrations.¹⁵ Although this did not guarantee symmetry in the final forward and inverse transforms, preliminary tests demonstrated that the symmetric Euclidean distance did increase the consistency between these transforms. These distances were minimized during the registration process using the Levenberg-Marquardt algorithm, which was applied separately for each of the three sequential transforms.¹⁶

II.B. Physician-delineated and deformably propagated surface meshes

Computed tomography (CT) images were acquired once per week throughout radiotherapy treatment for 18 lung

cancer patients using an active breath-hold technique. Four to seven weekly CTs were available for each patient. Refer to Weiss *et al.* for details on the study population and to Glide-Hurst *et al.* regarding the imaging protocol.^{10,17} One patient presented with two primary tumors, and 15 patients presented with between 1 and 4 involved lymph nodes, for a total of 19 tumors and 24 involved lymph nodes in this study. Treatment targets were contoured on all images by a physician and converted to surface meshes using the Pinnacle³ treatment planning system (version 8.1, Philips Medical Systems, Fitchburg, WI). Primary tumors ranged in size from 0.6 to 377.3 cm³ (median: 50.8 cm³) and contained between 414 and 3097 mesh vertices (median: 2350). The lymph node targets ranged from 0.4 to 17.8 cm³ (median: 2.2 cm³) and contained 266–2170 mesh vertices (median: 448). Meshes contained an average of 50 vertices per cm² surface area, which provided sufficiently accurate registrations without demanding extensive run times. Registrations were performed on the Pinnacle workstation and required between 5 and 25 min to compute all three transforms.

The first image (labeled “week 1”) generally corresponded to the planning CT, so all targets from the week 1 CT were designated as reference meshes for deformable registration. Corresponding weekly tumors and involved nodes designated as moving meshes. All weekly images were initially registered to the planning CT by manually aligning bony anatomy. A total of 93 primary tumor registrations and 119 lymph node registrations were then performed using these surface meshes. For patients presenting with multiple primary tumors and/or multiple metastatic lymph nodes, each individual target structure was contoured and registered separately.

Physician-delineated meshes marked the ground truth for this study, but no correspondence existed between the vertices from reference and moving meshes, which precluded the calculation of mesh registration errors. To provide a known correspondence for initial validation of the deformable mesh registration, a second set of moving meshes was generated by deformably propagating the reference meshes to all other weekly images using a known transform. The known transform was obtained from previous intensity-based deformable image registration studies in this dataset¹⁸ using the small deformation, inverse consistent linear elastic (SICLE) algorithm.¹⁹ Briefly, SICLE uses a cost function composed of a sum of squared distances similarity metric, a linear elasticity penalty term, an inverse consistency penalty term, and a Fourier basis to represent the transform. For two patients, resolving atelectasis throughout treatment prohibited accurate image registration with SICLE, so their treatment targets were excluded from the analysis of deformably propagated meshes. Thus, 104 and 135 primary tumor and lymph node meshes were included in 86 and 112 registrations, respectively.

II.C. Analysis

II.C.1. Mesh registration accuracy

Registrations were first performed for the deformably propagated meshes. Because the correspondence of vertices

was preserved after applying the known transform, it was possible to compute the distance from each vertex on the moving mesh to its known location on the reference mesh. These distances provided a measure of the vertex registration error (VRE). Prior to and after each step in the registration process, the average, root-mean-square (RMS), and maximum VRE were tabulated. The mean and standard deviation for each of these statistics was then computed separately over all primary tumor registrations and all lymph node registrations.

Next, registrations were performed for the physician-delineated meshes. In this case, no correspondence existed between vertices of reference and moving meshes, precluding the calculation of VRE. The quality of mesh alignment was instead computed using the symmetric surface distance (SSD), a metric very closely related to the symmetric Euclidean distance utilized in DMR.²⁰ Prior to and after each step in the registration, the average SSD was given by the mean of all nearest-neighbor vertex distances in both directions (reference mesh vertices to the moving mesh as well as moving mesh vertices to the reference mesh); the RMS SSD was given by the square root of the mean of the squared nearest-neighbor distances in both directions; and the maximum SSD was given by the single largest nearest-neighbor distance in either direction (also known as the Hausdorff distance). As with VRE, the mean and standard deviation for each of these statistics was computed separately over all primary tumor registrations and all lymph node registrations. Note that the SSD did not constitute a true error metric. This measure only indicated how closely the vertices of two surface meshes were aligned.

II.C.2. Mesh registration regional consistency

In order for mesh registration to provide a useful validation for target localization from intensity-based registration algorithms, the mesh registration tool must be accurate not just on the surface but near it as well, since presumably a deformable target localization algorithm will use intensity information near the surface to drive the registration. The following methods were designed to extend the VRE analysis to regions surrounding the surface of target meshes. After deformable registration of propagated meshes, the resulting B-spline transform was sampled at the center of all voxels on or near the surface of the reference mesh. The resulting deformation vectors were compared against the known intensity-based deformation that was used to generate these propagated meshes. At the surface of the meshes, the two deformation fields were expected to demonstrate accuracy similar to the VRE measured above. Therefore, the goal of this analysis was to quantify the accuracy of deformable mesh registration near, but not on, the surface.

To accomplish this goal, the reference mesh was converted to a binary surface mask to limit the analysis to only those voxels within a narrow band on the mesh surface. A distance map was then used to compute the proximity of all neighboring voxels to the binary mask, measured in an integer number of voxels.^{14,21} This allowed the binary mask to be expanded (positive voxels) or contracted (negative voxels) for analysis

FIG. 2. Average, RMS, and maximum vertex registration errors between reference and *deformably propagated* surface meshes at each step in the registration process. Error bars represent one standard deviation across all target registrations.

of additional bands of image voxels which varied with proximity to the surface of the reference mesh. Expansions and contractions of up to 5 voxels were considered in this study, corresponding to distances of up to 10 mm from the mesh surface for $1.2 \times 1.2 \times 2.0 \text{ mm}^3$ voxels. On a voxel-by-voxel basis within each narrow band, the difference vectors were computed between the deformation fields from DMR and the known transform. The absolute mean, RMS, and maximum of the magnitude of the difference vectors were tabulated for each propagated mesh registration. These statistics were then averaged over all primary tumor registrations and all lymph node registrations.

To further evaluate the regional consistency of DMR with respect to the known deformation, a subvolume was defined around each voxel, spanning a diameter of 1–17 voxels. All difference vectors within a given subvolume were averaged to obtain a measure of the local agreement between DMR and the known deformation. This analysis was conducted as a function of distance from the mesh surface, with the absolute mean and maximum magnitude of difference vectors recorded for each propagated mesh registration.

III. RESULTS

III.A. Mesh registration accuracy

Figure 2 shows the VRE for the deformably propagated target meshes in their initial alignment and following each transform in the registration process. Errors for the unregistered primary tumor averaged $7.2 \pm 4.5 \text{ mm}$, but individual vertices demonstrated maximum errors as large as $11.9 \pm 7.2 \text{ mm}$. The rigid transform reduced the average and maximum errors to 2.3 ± 1.6 and $5.3 \pm 3.5 \text{ mm}$, respectively, with additional improvements provided by the affine and B-spline transforms. After registration, the final average VRE was $0.8 \pm 0.7 \text{ mm}$ (range: 0.0–1.5 mm), with maximum errors of $3.2 \pm 2.9 \text{ mm}$ (range: 0.0–10.1 mm). Lymph node registrations

demonstrated a similar trend. Average VRE for this target was $5.9 \pm 3.9 \text{ mm}$ prior to registration and $0.8 \pm 0.5 \text{ mm}$ after the rigid transform. Final average and maximum VRE were $0.2 \pm 0.3 \text{ mm}$ (range: 0.0–0.8 mm) and $0.8 \pm 1.1 \text{ mm}$ (range: 0.0–3.4 mm), respectively.

For comparison between VRE and SSD metrics, we also computed the SSD for the deformably propagated target meshes. After the final B-spline transform, the average and maximum SSD were 0.7 ± 0.5 and $2.0 \pm 1.5 \text{ mm}$ for the primary tumor, and 0.2 ± 0.2 and $0.6 \pm 0.7 \text{ mm}$ for lymph nodes.

For physician-delineated meshes, Fig. 3 summarizes the SSD throughout the registration process. The average primary tumor SSD was $4.1 \pm 2.8 \text{ mm}$ prior to registration, with maximum SSD of $15.2 \pm 7.9 \text{ mm}$. After deformable mesh registration, the B-spline transform resulted in an average SSD of $1.3 \pm 0.5 \text{ mm}$ (range: 0.1–9.0 mm) and a maximum SSD of $4.5 \pm 1.9 \text{ mm}$ (range: 1.2–9.9 mm). For the lymph nodes, the average SSD prior to registration was $3.2 \pm 2.1 \text{ mm}$, with a maximum SSD of $8.4 \pm 4.4 \text{ mm}$. The B-spline transform resulted in an average SSD of $0.8 \pm 0.2 \text{ mm}$ (range: 0.1–0.9 mm), with a maximum SSD of $2.1 \pm 0.8 \text{ mm}$ (range: 1.3–3.7 mm).

III.B. Mesh registration regional consistency

Figure 4 highlights the results of a voxel-by-voxel comparison of the transform vectors from deformable mesh registration and the known deformation field. At the surface of primary lung tumors, the magnitude of transform vector differences averaged $0.6 \pm 0.5 \text{ mm}$ and demonstrated maximum differences of $2.5 \pm 2.5 \text{ mm}$. These differences were noted to increase with increasing distance from the mesh. Inside the reference mesh (represented by negative voxel distances), differences averaged 0.6–0.9 mm. Average vector differences outside the mesh surface ranged from 0.7 to 1.8 mm, with maximum differences ranging from 3.3 to 8.2 mm. For lymph

FIG. 3. Average, RMS, and maximum symmetric surface distances between reference and *physician-delineated* meshes at each step in the registration process. Error bars represent one standard deviation across all target registrations.

node registrations, the average and maximum transform vector differences were 0.2 ± 0.3 and 0.6 ± 1.0 mm at the surface of these targets. Average vector differences ranged from 0.2 to 0.9 mm inside the lymph node surface and 0.2–1.4 mm outside the surface. Maximum errors of 1.1 ± 0.8 and 5.5 ± 6.1 mm were computed at a distance of five voxels inside and outside the mesh surface, respectively.

Figure 5 demonstrates the effect of local averaging within regions of varying diameter. On the surface of primary tumor meshes, the average magnitude of differences between DMR and the known deformation ranged from 0.5 to 0.6 mm. At a distance of five voxels, average differences were 0.5–0.9 mm inside the mesh and 1.6–1.8 mm outside the mesh. In both cases, the differences were generally reduced as the region diameter increased. The maximum differences for primary tumor registrations ranged from 2.4 to 2.8 mm at the mesh sur-

face, from 1.3 to 2.0 mm five voxels inside the mesh surface, and from 6.7 to 8.2 mm five voxels outside the mesh surface. Similar trends were noted for lymph node registrations. Average differences of 0.2–0.5, 0.2–0.9, and 0.9–1.4 mm were observed on the mesh surface, five voxels inside the mesh, and five voxels outside the mesh, respectively. Corresponding maximum errors for the lymph nodes were 0.5–1.2, 0.2–1.1, and 2.7–5.5 mm. At a distance of five voxels inside or outside the mesh surface, differences were again reduced as the region diameter increased.

IV. DISCUSSION

This study aimed to evaluate deformable mesh registration as a validation tool for deformable and semideformable image registration algorithms for automatic target localization

FIG. 4. Average, root-mean-square, and maximum of the magnitude of the differences between transform vectors from deformable mesh registration and the known deformation, plotted as a function of distance (in voxels) from the fixed mesh surface. Positive distances represent outward displacements from the surface mesh.

FIG. 5. Average (top row) and maximum (bottom row) of the regional differences between deformable mesh registration and a known deformation for primary tumor meshes (left column) and lymph node meshes (right column). Results were assessed as a function of both region size and the distance from the center voxel of a given region to the mesh surface.

in image-guided radiation therapy. The developed DMR algorithm demonstrated vertex registration errors for the deformably propagated meshes that were less than the voxel size for images in this study. Primary tumor registrations contributed larger errors than lymph node registrations. This was because the primary tumors were generally larger and more irregularly shaped than the lymph nodes, and thus underwent greater shape and volume change (by absolute distance and volume). Results in this study are relatively consistent with intensity-based registrations, which often demonstrate accuracy on the order of the voxel size or less. This is also less than the range of typical action levels for patient setup corrections, where treatment couch shifts less than 2–3 mm are generally ignored.^{2,22,23}

Using DMR, intensity-based registration algorithms for online target localization could be validated on and near the target surface by comparison against the DMR transform. It should be noted that DMR was not intended for online validation during image-guided radiotherapy (e.g., as a “second check” of automated registration). Rather, this validation tool was designed for use during the development and validation of intensity-based target localization algorithms to be implemented during online guidance. Because DMR relies solely on the alignment of surface mesh vertices, this approach provides a method of validation that is independent of local intensity values and that is expected to be more robust than DIR against large deformation and tissue change.

Figure 6 highlights differences between DMR and a nearest-neighbor distance (NND) validation standard. NND is a distance-to-surface measure that depends strongly on the initial orientation between meshes from the planning and on-treatment images. Each distance represents the shortest possible path from the on-treatment mesh to the planning mesh, which often poorly represents changes in the tumor surface throughout treatment. In contrast, DMR results in better tumor surface alignment by registering the entire distribution of mesh vertices through a series of transforms. Furthermore, because each transform parameterizes the final mesh alignment, transform vectors can be computed at small distances away from the mesh surface (Fig. 4). DMR can therefore validate DIR algorithms in larger image regions.

A DMR-based validation technique may also be applied to other types of image registration. One particularly suitable application is block-matching registration, in which the rigid alignment of a series of small blocks of image content yields an effective global registration of the entire image.^{24,25} In this case, the expected transform for an individual block placed near the surface of a contoured target could be predicted by DMR. If all blocks are registered consistently with the corresponding DMR transforms, then it is likely that the global image registration will adequately capture the contoured target volume. Similar methods may be extended to the validation of rigid and affine transforms as well.

Because the deformably propagated meshes were generated from intensity-based registration of the weekly CTs,

FIG. 6. (a)–(c): Intensity-based DIR between planning and on-treatment images. (d) The DIR transform is extracted at the surface of the on-treatment mesh and compared against possible validation standards, including (e) NND between the planning and on-treatment surface meshes, and (f) DMR. Differences between DIR, NND, and DMR are shown in (g)–(h).

these meshes were reasonably similar to the physician-delineated meshes for the patients in this study. Exceptions were noted for two patients with resolving atelectasis, which was why these patients were excluded from VRE-related studies and limited to SSD-based analysis. The final VRE results were thus expected to provide an estimate of registration errors for physician-delineated meshes, which lacked a known vertex correspondence. Differences between the VRE and SSD were analyzed by computing both metrics for artificially deformed meshes. As expected, slight reductions were noted for the average SSD, indicating that the nearest-neighbor vertices were generally located very close to the actual corresponding vertices. However, substantially larger differences were noted between the maximum VRE and maximum SSD.

This result highlights the potential for large registration errors parallel to the target surface, where nearest-neighbor vertices may be somewhat removed from the true corresponding vertices. The maximum SSD of physician-delineated meshes may, therefore, be somewhat underestimated. When applying DMR as a validation tool, the average SSD should be computed as the more reliable measure, specifically by considering the transform vectors from a distribution of points throughout the entire target surface.

The SSD for physician-delineated meshes was actually found to be somewhat larger than the VRE of the artificially deformed meshes, the reason for which may be twofold. First, although contouring uncertainties may impact the initial target from the planning CT, no additional contouring variability

was present for the remaining artificially deformed meshes. In contrast, the physician-delineated meshes were all created independently, and contouring variability between targets could have potentially contributed to larger SSD results. The second reason likely stemmed from the fact that the known deformation in this study only approximated the true underlying anatomic and pathological changes for patients in this study. Because the physician-delineated meshes were identified on the actual patient images, these meshes likely incorporated a greater degree of irregularity that was actually present in the weekly CTs. These mesh irregularities could have also potentially contributed to larger SSD results.

The final DMR transform, as well as the magnitude of the SSD metric, has some dependence on the density of points in the surface mesh. In this work, targets contained an average of 50 vertices per cm^2 (range: 12–120) along their surface, which was selected to balance registration accuracy and execution time. Increasing the density of mesh vertices would have likely had a negligible influence on the final DMR transform and on the SSD reported in this study. However, decreasing the number of mesh points could potentially introduce a slight degree of variability in the final transform and could also increase the SSD by a small amount. The extent of this effect would depend on how sparsely distributed the mesh vertices were. In general, we anticipate a very small impact on the DMR transform for all but the most sparsely sampled vertex distributions.

The focus of this study involved the application of DMR to validate intensity-based registration of local image subvolumes. For image voxels at the surface of the propagated meshes, the transform vectors from deformable mesh registration demonstrated submillimeter agreement with a known deformation field. The average magnitude of vector differences was found to be consistent with the average VRE, as expected. Furthermore, these discrepancies remained relatively constant inside the surface of target meshes. Larger discrepancies, however, were noted outside the surface of these meshes. This indicates that while DMR may be a reliable validation tool on or just inside the target surface, uncertainties can become much more substantial at distances farther outside these volumes. It was possible to reduce uncertainties by averaging the result of DMR over localized regions, but as evident in Fig. 5, large uncertainties persisted at greater distances outside the mesh surface. The scope of DMR as a validation tool should thus be limited to regions near the target surface. Registration accuracy in these regions is one of the primary concerns for target localization, which stems from the importance of orienting the target entirely within the intended radiation fields. For other applications such as dose mapping and adaptive radiotherapy tasks, the requirement for accurate surface alignment is necessary but may not be sufficient.

Because DMR relies on target contours, this method is also inherently limited by contouring variability. The surface of primary lung tumors is generally well-visualized against the surrounding lung parenchyma, improving the reliability of contours in these regions. Larger variability may be expected for tumors extending into the mediastinum or bordering atelectasis. Systematic contouring errors along these parts of

the tumor surface would lead directly to local systematic errors in the deformable mesh transform, consequently leading to larger discrepancies between DMR and DIR in these regions. Because of the reduced contrast in the mediastinum or near atelectasis, however, discrepancies between the two algorithms could also stem from uncertainties in the intensity-based registration. Operator judgment may be necessary to resolve the true source of discrepancies in this case.

Contouring variability with high spatial frequency, such as slice-to-slice variability, is not anticipated to substantially detract from the final DMR transform. This is because B-spline registration provides a locally smooth transform in which the effect of purely random fluctuations should cancel out. Relatively, few B-spline control points are used in this study to promote a smooth transform and avoid overfitting the mesh models. This effect can be observed in the rightmost panel of Fig. 1, where the high-frequency “roughness” of the contours has not been smoothed out by the registration.

As a final limitation, deformable mesh registration cannot be considered a “ground truth” for registration error because of potential contouring variability noted above as well as the small but not inconsequential residual registration error in DMR itself. Instead, in lieu of a ground truth, DMR is intended to serve as a standard metric for comparing different intensity-based registrations at or near the tumor surface (Fig. 6). DMR provides a reasonable approximation of the true underlying changes in anatomy as it is based on the physician’s clinical judgment and potentially other sources of clinical information (e.g., PET/CT). In general, a DIR algorithm in better agreement with DMR is more likely to reflect the underlying anatomical changes at the target surface than another algorithm with worse agreement. This logic may also be applied to algorithm development, where two different versions of a code could both be compared against DMR to determine which better localizes the target surface.

Future work will involve the application of DMR toward development of robust automatic registration tools for target localization. Additionally, we hope to explore the application of DMR in other disease sites with similar challenges of target shape change and regression, including head and neck and gynecological cancers.

V. CONCLUSION

A deformable mesh registration tool has been implemented for the purpose of validating the quality of target localization from automatic image registration in image-guided radiation therapy. The subvoxel accuracy of this tool was reasonable for the specific application of target localization, with errors less than the typical action levels of modern patient setup protocols. Although this tool is limited to validating registrations of the image subvolumes encompassed by the given meshes, its applicability extends to both rigid (including affine) and nonrigid registration problems.

ACKNOWLEDGMENTS

This work was partially supported by National Institutes of Health (NIH) R01CA166119. It relied on the Insight

Segmentation and Registration Toolkit (ITK), an open source software developed as an initiative of the U.S. National Library of Medicine and available at www.itk.org. None of the authors has any actual or potential conflicts of interest.

^{a)} Author to whom correspondence should be addressed. Electronic mail: gdhugo@vcu.edu; Telephone: 804-628-3457; Fax: 804-628-0271.

¹M. van Herk, "Different styles of image-guided radiotherapy," *Semin. Radiat. Oncol.* **17**, 258–267 (2007).

²J.-P. Bissonnette, T. G. Purdie, J. A. Higgins, W. Li, and A. Bezjak, "Cone-beam computed tomographic image guidance for lung cancer radiation therapy," *Int. J. Radiat. Oncol., Biol., Phys.* **73**, 927–934 (2009).

³N. Mohammed, L. Kestin, I. Grills, C. Shah, C. Glide-Hurst, D. Yan, and D. Ionascu, "Comparison of IGRT registration strategies for optimal coverage of primary lung tumors and involved nodes based on multiple four-dimensional CT scans obtained throughout the radiotherapy course," *Int. J. Radiat. Oncol., Biol., Phys.* **82**, 1541–1548 (2012).

⁴E. S. Worm, A. T. Hansen, J. B. Petersen, L. P. Muren, L. H. Præstegaard, and M. Høyer, "Inter- and intrafractional localisation errors in cone-beam CT guided stereotactic radiation therapy of tumours in the liver and lung," *Acta Oncol.* **49**, 1177–1183 (2010).

⁵A. R. Yeung, J. G. Li, W. Shi, H. E. Newlin, A. Chvetsov, C. Liu, J. R. Palta, and K. Olivier, "Tumor localization using cone-beam CT reduces setup margins in conventionally fractionated radiotherapy for lung tumors," *Int. J. Radiat. Oncol., Biol., Phys.* **74**, 1100–1107 (2009).

⁶J. Higgins, A. Bezjak, K. Franks, L. W. Le, B. C. Cho, D. Payne, and J.-P. Bissonnette, "Comparison of spine, carina, and tumor as registration landmarks for volumetric image-guided lung radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **73**, 1404–1413 (2009).

⁷J.-J. Sonke and J. Belderbos, "Adaptive radiotherapy for lung cancer," *Semin. Radiat. Oncol.* **20**, 94–106 (2010).

⁸S. van Beek, S. van Kranen, A. Mencarelli, P. Remeijer, C. Rasch, M. van Herk, and J.-J. Sonke, "First clinical experience with a multiple region of interest registration and correction method in radiotherapy of head-and-neck cancer patients," *Radiat. Oncol.* **94**, 213–217 (2010).

⁹K. K. Brock, M. Hawkins, C. Eccles, J. L. Moseley, D. J. Moseley, D. A. Jaffray, and L. A. Dawson, "Improving image-guided target localization through deformable registration," *Acta Oncol.* **47**, 1279–1285 (2008).

¹⁰C. K. Glide-Hurst, E. Gopan, and G. D. Hugo, "Anatomic and pathologic variability during radiotherapy for a hybrid active breath-hold gating technique," *Int. J. Radiat. Oncol., Biol., Phys.* **77**, 910–917 (2010).

¹¹G. Lim *et al.*, "Tumor regression and positional changes in non-small cell lung cancer during radical radiotherapy," *J. Thorac. Oncol.* **6**, 531–536 (2011).

¹²G. D. Hugo, E. Weiss, A. Badawi, and M. Orton, "Localization accuracy of the clinical target volume during image-guided radiotherapy of lung cancer," *Int. J. Radiat. Oncol., Biol., Phys.* **81**, 560–567 (2011).

¹³S. P. Robertson, E. Weiss, and G. D. Hugo, "Localization accuracy from automatic and semi-automatic rigid registration of locally-advanced lung cancer targets during image-guided radiation therapy," *Med. Phys.* **39**, 330–341 (2012).

¹⁴L. Ibanez, W. Schroeder, L. Ng, and J. Crates, *The ITK Software Guide*, 2nd ed. (Kitware, Inc., Clifton Park, NY, 2005).

¹⁵Z. Zhang, "Iterative point matching for registration of free-form curves and surfaces," *Int. J. Comput. Vis.* **13**, 119–152 (1994).

¹⁶B. Zitová and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.* **21**, 977–1000 (2003).

¹⁷E. Weiss, S. P. Robertson, N. Mukhopadhyay, and G. D. Hugo, "Tumor, lymph node, and lymph node-to-tumor displacements over a radiotherapy series: Analysis of interfraction and intrafraction variations using active breathing control (ABC) in lung cancer," *Int. J. Radiat. Oncol., Biol., Phys.* **82**, e639–e645 (2012).

¹⁸A. M. Badawi, E. Weiss, W. C. Sleeman IV, and G. D. Hugo, "Classifying geometric variability by dominant eigenmodes of deformation in regressing tumours during active breath-hold lung cancer radiotherapy," *Phys. Med. Biol.* **57**, 395–413 (2012).

¹⁹G. E. Christensen, J. H. Song, W. Lu, I. El Naqa, and D. A. Low, "Tracking lung tissue motion and expansion/compression with inverse consistent image registration and spirometry," *Med. Phys.* **34**, 2155–2163 (2007).

²⁰T. Heimann *et al.*, "Comparison and evaluation of methods for liver segmentation from CT datasets," *IEEE Trans. Med. Imaging* **28**, 1251–1265 (2009).

²¹P.-E. Danielsson, "Euclidean distance mapping," *Comput. Graph. Image Process.* **14**, 227–248 (1980).

²²G. R. Borst, J.-J. Sonke, A. Betgen, P. Remeijer, M. van Herk, and J. V. Lebesque, "Kilo-voltage cone-beam computed tomography setup measurements for lung cancer patients: First clinical results and comparison with electronic portal-imaging device," *Int. J. Radiat. Oncol., Biol., Phys.* **68**, 555–561 (2007).

²³I. S. Grills, G. D. Hugo, L. L. Kestin, A. P. Galerani, K. K. Chao, J. Wloch, and D. Yan, "Image-guided radiotherapy via daily online cone-beam CT substantially reduces margin requirements for stereotactic lung radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **70**, 1045–1056 (2008).

²⁴M. Söhn, M. Birkner, Y. Chi, J. Wang, D. Yan, B. Berger, and M. Alber, "Model-independent, multimodality deformable image registration by local matching of anatomical features and minimization of elastic energy," *Med. Phys.* **35**, 866–878 (2008).

²⁵U. Malsch, C. Thieke, P. E. Huber, and R. Bendl, "An enhanced block matching algorithm for fast elastic registration in adaptive radiotherapy," *Phys. Med. Biol.* **51**, 4789–4806 (2006).