Evaluating geometrical accuracy of image registration methods in SPECT guided radiation therapy

Lingshu Yin [1,2,*]; Lisa Tang [3]; Ghassan Hamarneh [3]; Anna Celler [4]; Sergey Shcherbinin [4]; Tsien-Fei Fua [5]; Hannah Carolan [5]; Anna Thompson [5]; Mitchell Liu [6]; Cheryl Duzenli [1,2]; Brad Gill [2]; Finbar Sheehan [5]; John Powe [7]; Daniel Worsley [7]; Vitali Moiseenko [1,2];

(1) Physics and Astronomy, University of British Columbia, Canada (2) Medical Physics, Vancouver Cancer Centre, BC Cancer Agency, Canada (3) Medical Image Analysis Lab, Simon Fraser University, Canada (4) Department of Radiology, University of British Columbia (5) Radiation Oncology, Vancouver Cancer Centre, BC Cancer Agency (6) Radiation Oncology, Fraser Valley Cancer Centre, Canada (7) Nuclear Medicine, Vancouver General Hospital, Canada

Abstract: We evaluated six methods for CT-CT registration for SPECT-guided radiotherapy treatment planning: affine registration with manual marking of corresponding control points either on the skin or interior to the lung, automatic rigid registration and three deformable registration methods (B-Spline, Demons and Level set). According to the registration results, SPECT images obtained from a hybrid SPECT/CT were warped and aligned with the RTTP CT. Visual and quantitative evaluations suggested that the use non-rigid registration algorithms achieved higher CT-CT registration accuracy than the other registrations. However, the warped SPECT images from Demons or Level set became clinically invalid due to the excessive degrees of freedom in these deformation fields. We conclude that the use of intensity-based metrics alone is not sufficient to drive registration; the importance of regularizing the deformation field should not be underplayed; and it is essential to examine a variety of accuracy metrics for proper evaluation of image registration.

1. Introduction

Radiation therapy (RT) plays an important role in the treatment of patients with lung cancer. The goal of radiation therapy is to eradicate cancer cells while sparing normal tissues. The objective of RT treatment planning (RTTP) is to provide a patient-specific balance between these two goals. In lung cancer RT, it is particularly important to minimize dose to organ at risks (OAR) such as normal lung, spinal cord, and esophagus.

Marks et al. [1] have demonstrated that the incorporation of perfusion Single Photon Emission Tomography (SPECT) imaging using 99mTc macroaggregated albumin (MAA)

* *Corresponding author*: Lingshu Yin, M.Sc, Medical Physics, Vancouver Cancer Centre, 600 West 10th Ave, BC, V5Z 4E6, Canada. E-mail: lyin@bccancer.bc.ca

into RTTP may treatment-related morbidity for patients with lung cancer. SPECT can accurately capture the 3D distribution of lung perfusion. Image registration of perfusion SPECT and RT planning CT images can allow treatment planners to identify sites with high or low perfusion with respect to the morphological information from CT [2-7]. Tighter radiation dose constraints are then placed on the high perfusion volume in order to improve the sparing of functional parts of the lung.

In previously reported clinical studies, SPECT-CT registrations were carried out via point-based registration and rigid registration. Some of these studies [2-4] made use of skin fiducial markers and others were based on anatomical and physiological landmarks within the lung [5, 6]. While these point-based registration approaches have been adopted in most clinics, they require manual identification of landmarks, which is time-consuming. More importantly, their accuracies are not well established. Accordingly, the work presented in this paper aims at 1) introducing and evaluating several automatic non-rigid registration methods in the context of RTTP of the lung; and 2) comparing the accuracies of these techniques against the clinically-adopted point-based image registration methods.

2. Methods

Six registration methods were tested to register the hybrid CT and RTTP CT images from ten lung cancer patients. Institutional ethics approval was obtained prior to recruiting patients.

2.1 Image acquisition

Ten patients (6 Males aged: 50±10yrs, 4 females 40±10yrs) recently diagnosed with lung cancer and to be treated with radiation therapy were consented to have a perfusion SPECT scan for the purpose of this study. In order to localize the tumor volume and setup radiation fields, a conventional planning CT scan (reconstructed to resolution of 512x512x80 voxels, each in size of 0.96x0.96x5mm^3) was performed with patients in the treatment position. Radio-opaque markers were placed on the skin as a reference for subsequent SPECT/CT scan. Treatment planning and dose calculations were performed on this CT set. Thus, it was used as the *fixed CT image* in the image registration.

After acquiring the planning CT scan, a perfusion-SPECT/CT scan was performed using Infinia-Hawkeye SPECT-CT camera (GE Healthcare), with the patients in the treatment position. This generated a pair of inherently registered images: a SPECT image (which we labeled as the *moving SPECT image*) (128*128*128 voxels, each 4.42*4.42*4.42mm^3) and a low resolution CT image (labeled *moving CT image*) (256*256*40 voxels, each 2.21*2.21*10mm^3).

2.2 Image registration

In all registrations, the *moving CT image* was transformed to bring it into correspondence with the *fixed CT image*. The registration-derived transformation and/or deformation field were applied to the *moving SPECT image* to yield the *warped SPECT image* which would be co-registered to the *fixed CT image* and can be used for SPECT-guided RTTP.

We employed and compared the following six registration methods to evaluate their performance and accuracy for the task. The first three methods have been used in previous SPECT-guided RT studies.

Skin control point based registration (referred to as the 'Skin' method hereafter): This method was adopted to mimic the work in [2-4], where the moving CT image was unavailable from a stand-alone SPECT scanner. For the SPECT scan, small flasks filled with 99mTc were placed in the same locations on the skin as the radio-opaque markers placed in the CT acquisition. For registration, eight to twelve pairs of control points were manually chosen on the skin contour of the fixed and moving CT image respectively. The 3D affine transform minimizing the distance between two sets of corresponding control points was calculated as a closed-form solution to this least-squares problem.

Lung control point based registration (Lung): Eight to twelve pairs of control points were manually chosen on landmarks within the lung volume in the fixed and moving CT images. The 3D affine transform is calculated as done in the previous method.

Rigid registration (Rigid): The fixed CT image and moving CT image were transformed rigidly to minimize the negative value of the mutual information (MI) metric between the two images, using the implementation by Mattes et al. [8].

To assess the potential benefit of the use of non-rigid registration, we also examined the accuracies of the following automatic and deformable registration algorithms. In all three, we investigated registration parameters that control the level of regularization enforced on the displacement field. Based on initial experiments, we restricted the regularization weight to a range that allows for successful optimization of the image similarity metric and at the same time does not enforce too excessive regularization.

B-Spline registration (B-Spline): We employed a gradient-descent based optimizer to minimize the negative value of MI using [8]. The combined use of B-Spline regularization and MI was chosen due to its proven suitability for registering CT lung studies [9].

Diffeomorphic Demons registration (Demons): This is an optical-flow like technique which has been applied to the registration of sequential CT images of a lung cancer study and qualitatively evaluated in [10]. We used the diffeomorphic version in [11].

Level set registration (Level set): This method aligns images by modeling deformations as the level set motion [12]. The equations of motion are similar to those from the original Demons' formulation, but it is relatively more efficient than the original Demons. We thus selected it to examine the tradeoff between computational efficiency and performance.

2.3 Registration evaluation

In the literature, validation of image registrations can be done quantitatively or qualitatively. One commonly adopted quantitative validation approach is the calculation of target registration error (TRE) using two sets of corresponding landmarks identified in the registered images by clinical experts [13]. This approach becomes inadequate for our application since it is challenging to identify a set of landmarks along the lung surface in a reproducible manner [13]. Another practice commonly adopted in the clinic and in [10,12,14] is the calculation of some image metrics that compare the similarity between the registered images. While they may reflect the success of the optimization procedure used for registration, these metrics do not convey the validity of the obtained spatial transformations. Consequently, as we will show later in the paper, neither of these quantitative evaluation measures provides a reliable measure of the true registration accuracy.

Similarly, qualitative validation, which involves experts' subjective evaluations of the registered images in forms of difference images (examples are shown in Fig 1), color-overlay or checkerboard fusion have failed to reflect the validity of the displacement fields derived by some of the deformable registration methods we tested. The validity of the displacement field, while often not reported in validation studies, is of particular importance in our study because singularities in the field (e.g. crossing of displacement vectors of two adjacent voxels) would lead to false or implausible activity distribution, which in turn produces an incorrect functional lung volume for RTTP.

Accordingly, rather than merely computing TRE, image similarity measures, or performing subjective evaluations for assessing the geometric accuracy of alignment, we also computed several other criteria to examine the effect of registration on the SPECT images, despite that the registration was done on the CT images. These criteria include examining the level of "preservation" of SPECT activity counts and the plausibility of deformation field, and comparison of the histograms of the SPECT images before and after registration. The different evaluation criteria adopted are summarized in sections 2.3.1-4.

2.3.1. Image intensity based (CT-based) metrics

We computed the following intensity based metrics as done by Ursher et al. [14] for validation of CT-CT registrations: RMS_{int} (Root-Mean-Square of Intensity Differences), MAD_{int} (Median-Absolute-Deviation of Intensity Differences) and MID_{int} (Maximum Intensity Differences).

2.3.2 Expert review based on anatomical landmarks

After the completion of image registration, two radiation oncologists (T.F. and H.C.) were asked to identify lung landmarks on the fixed CT image and the moving CT image. Window and intensity level settings were adjusted to visualize the edges of the lung. The TRE was computed using the selected landmarks.

2.3.3 Preservation of SPECT counts (SPECT-based metric)

The intensity histogram of moving SPECT image and warped SPECT image were calculated. Then, we calculated the histogram difference of the SPECT intensities ($HIST_{SPECT}$) as the sum of absolute volume difference in each bin:

$$HIST_{SPECT} = \sum_i \left| V^i_{moving} - V^i_{warped} \right| \quad \textbf{(1)}$$

where V^i_{moving} and V^i_{warped} are the volume of ith bin in the intensity histogram of the moving SPECT image and warped the SPECT image, respectively. We adopted this metric to examine the difference between the *moving SPECT images* and the warped SPECT images. It assesses the amount of changes made to the original SPECT's intensity distribution caused by the application of the displacement field. Although a smaller value of this metric does not necessarily suggest a better registration, a larger value does indicate that more SPECT activity counts have been affected by registration.

2.3.4 Singularities in deformation field

The determinant of the Jacobian of the displacement field gives information about local volume changes. To assess the amount of singularities in the deformation field, we computed JAC_{disp}, the percentage of negative values in the Jacobian determinant as done in [14]. Ideally this measure should be 0.

3. Results

Results of the six registration methods tested averaged over ten patients are shown in Table 1. For the deformable registration algorithms, we repeated registrations over a selected range of registration parameters and the one that yielded the lowest JAC_{disp} and TRE was chosen for evaluation. From Table 1, we can see a marked improvement in RMS_{int}, MAD_{int} and MID_{int} from manual to automatic rigid registration, then automatic rigid to non-linear registration. By visual inspection, the intensity differences between the registered CT images are getting smaller for different registration methods as shown in Fig. 1. The reduction in these differences is consistent with the reduction of TRE and intensity metrics (Table 1). However, based on visual examination of the warped SPECT, we found that those generated from Level set and Demons registrations were not clinically realistic when compared with the original SPECT images. Examples are shown in Fig. 1 where false SPECT counts (local maxima in intensity values in highlighted regions) can be seen. This was largely due to the singularities (folding, etc.) in the deformation field (Fig. 2 and Table 1). We also examined the level of preservation in the activity counts of SPECT; values of the SPECT-based metric are reported in Table 1. The averaged values of selected CT-based and SPECT-based metrics were normalized to the same scale and are shown in Fig. 3. It is clear that values of CT-based and SPECT-based metrics increased in opposite directions. This suggested that while the degree of deformation had increased as a result of the improved alignment between the registered CT images in the cases of Demons and Level set registrations, it had increased to such an extent that the *warped SPECT images* became clinically invalid. Thus, CT-based metrics have failed to reflect the validity of these registrations.

Table 1. Averaged results of the metrics used in this study.

	Skin	Lung	Rigid	B-Spline	Demons	Level set
RMS_{int} (HU)	230.77	205.60	158.01	123.39	79.68	63.68
MAD_{int} (HU)	44.5	40.3	34	29.1	19.2	16.9
MID_{int} (HU)	577.4	507.5	358.5	258	167.9	110.5
Anatomy based metric (mm)	5.629	3.8065	3.2765	1.902	1.116	0.684
$HIST_{SPECT}$ (mm^3)	7225	7832	6444	9594	18572	15483
JAC_{disp} (%)	0	0	0	0.2571	37.88	24.69

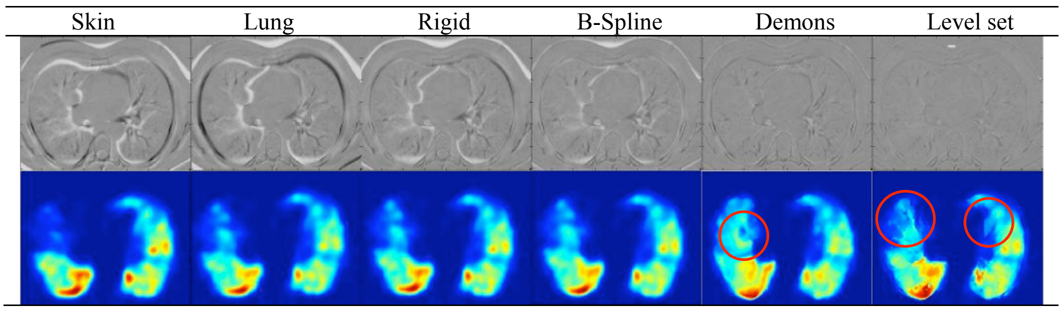

Fig. 1. Above: Slices of the difference image computed between the *moving CT image* and the *fixed CT image* after each registration. Below: Slices of the SPECT images warped with the deformation field from each registration. Circles indicated where SPECT signals are clinically invalid.

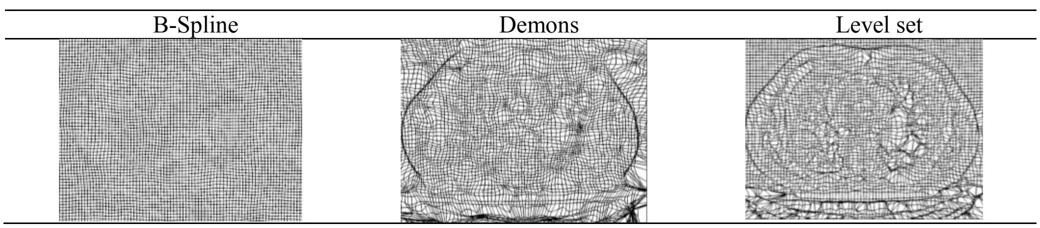

Fig. 2. Slices of the deformation fields obtained from B-Spline, Demons and Level set registrations. These slices correspond to those shown in Fig. 1.

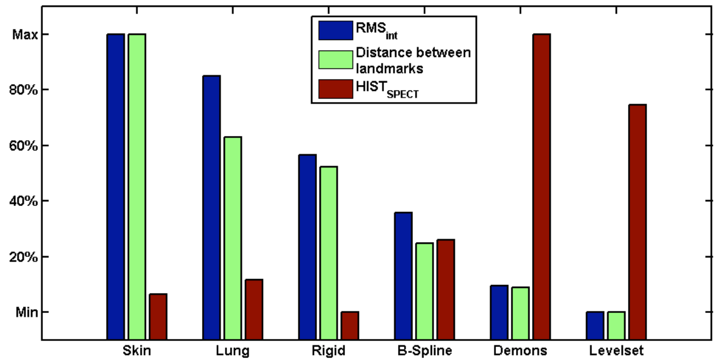

Fig. 3. Normalized average values of selected CT-based and SPECT-based metrics.

4. Discussions and Conclusions

In this study, we compared the accuracies of six CT-CT registrations, three of which, Skin, Lung and Rigid, have been previously reported [2-7]. Of these three methods, rigid registration was demonstrated to perform the best based on intensity metrics and anatomical landmarks. In the two control point-based registration methods, the skin method was found inferior to the lung method. This suggests that lung control points can better illustrate the deformation of the lung.

We further introduced the use of three automated registration methods, B-Spline, Demons and Level set, for image registration in SPECT guided RT to correct for possible non-rigid deformations that may exhibit between the CT scans. Improvements in both the image intensity based metrics and anatomy based metrics were observed in all three. However, upon applying the deformation field obtained from Demons and Level set to the perfusion SPECT, clinically unrealistic SPECT activity would be generated within the lung regions. Neither intensity-based metrics nor anatomy-based metrics could reflect this problem because they do not account for the disruptions in the connectivity of image voxels, especially in homogenous regions in CT, due to singularities in the obtained deformation fields. This suggests that metrics derived from the CT images (i.e. based on control point or image similarity) for driving the process of registering these images do not necessarily reflect the validity of the derived solution needed for SPECT-CT registration. Consequently, as illustrated in this work, in lieu of inspecting the obtained displacement fields and warped SPECT images, calculation of the $HIST_{SPECT}$ and JAC_{disp} measures served as a useful sanity check on the registration solutions. This also implies that special care must be taken when designing the regularization term in the cost function used for registration. Addressing this shall be the next stage of our work.

In summary, we have shown that image registrations done in current SPECT guided RTTP can be effectively improved via combined use of automatic rigid and B-Spline registration algorithms. Furthermore, we have demonstrated and explained in this work the importance of employing complementary evaluative measures when assessing the accuracies of registrations, and in our problem, how such approach has informed us of the serious impact invalid deformation fields have on the design of patients' treatments due to the false activity they can induce on SPECT.

References

1. Marks, L.B., Sherouse, G.W., Munley, M.T., et al.: Incorporation of functional status into dose-volume analysis. Med.Phys. 26(2), pp. 196-199 (1999)
2. Christian, J., Partridge, M., Nioutsikou, E., et al. The incorporation of SPECT functional lung imaging into inverse radiotherapy planning for non-small cell lung cancer. Radiother.Oncol. 77(3), pp. 271-277 (2005)
3. Lavrenkov, K., Christian, J., Partridge, M., et al.: A potential to reduce pulmonary toxicity: the use of perfusion SPECT with IMRT for functional lung avoidance in radiotherapy of non-small cell lung cancer. Radiother.Oncol. 83(2), pp. 156-162 (2007)
4. Lavrenkov, K., Singh, S., Christian, J., et al.: Effective avoidance of a functional SPECT-perfused lung using intensity modulated radiotherapy (IMRT) for non-small cell lung cancer (NSCLC): An update of a planning study. Radiother.Oncol. 91(3), pp. 349-352 (2009)
5. McGuire, S., Zhou, S., Marks, L., et al.: A methodology for using SPECT to reduce intensity-modulated radiation therapy (IMRT) dose to functioning lung. Int.J.Radiat.Oncol.Biol.Phys. 66(5), pp. 1543-1552 (2006)
6. Shioyama, Y., Jang, S., Liu, H., et al. Preserving functional lung using perfusion imaging and intensity-modulated radiation therapy for advanced-stage non-small cell lung cancer. Int.J.Radiat.Oncol.Biol.Phys. 68(5), pp. 1349-1358 (2007)
7. Shen, S., Spencer, S., Bender, L., et al.: Incorporating SPECT Functional Lung Images in Routine Treatment Planning for Lung Cancer. Med.Phys. 35(6), pp. 2876 (2008)
8. Mattes, D., Haynor, D., Vesselle, H., et al.: Non-rigid multimodality image registration. Medical Imaging 2001: Image Processing, pp. 1609-1620 (2001)
9. De Craene, M., D'Aische, A., Macq, B., et al.: Incorporating Metric Flows and Sparse Jacobian Transformations in ITK. The Insight Journal (2006), http://hdl.handle.net/1926/183
10. Wang, H., Dong, L., O'Daniel, J., et al. Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy. Phys.Med.Bio. 50, pp. 2887-2905 (2005)
11. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Diffeomorphic Demons Using ITK's Finite Difference Solver Hierarchy. The Insight Journal (2007), http://hdl.handle.net/1926/510
12. Vemuri, B.C., Ye, J., Chen, Y., et al.: A level-set based approach to image registration. Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis, pp. 86--93, IEEE Press New York (2000)
13. Fitzpatrick, J.M., Hill, D., Shyr, Y., et al.: Visual assessment of the accuracy of retrospective registration of MR and CT images of the brain, Trans. Med. Imag. 17(4) pp. 571-585 (1998)
14. Urschler, M., Kluckner, S., Bischof, H.: A framework for comparison and evaluation of nonlinear intra-subject image registration algorithms. The Insight Journal (2007), http://hdl.handle.net/1926/561