Radiotherapy and Oncology 127 (2018) 332–338

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Online adaptive radiotherapy

Fast and robust adaptation of organs-at-risk delineations from planning scans to match daily anatomy in pre-treatment scans for online-adaptive radiotherapy of abdominal tumors

Vikas Gupta [a,b,c,*], Yibing Wang [a], Alejandra Méndez Romero [a], Andriy Myronenko [d], Petr Jordan [d], Calvin Maurer [d], Ben Heijmen [a], Mischa Hoogeman [a]

[a] Department of Radiation Oncology, Erasmus MC Cancer Institute, Groene Hilledijk 301, Rotterdam 3075 EA, The Netherlands; [b] Division of Cardiovascular Medicine, Department of Medical and Health Sciences, Linköping University; [c] Center for Medical Imaging and Visualization (CMIV), Department of Medical and Health Sciences, Linköping University, Sweden; and [d] Accuray Incorporated, 1310 Chesapeake Terrace, Sunnyvale, CA 94089, USA

ARTICLE INFO

Article history:
Received 8 April 2016
Received in revised form 5 February 2018
Accepted 12 February 2018
Available online 8 March 2018

Keywords:
Online adaptive radiotherapy
Cyberknife
SBRT
Image registration
Liver
Dose sparing

ABSTRACT

Purpose: To validate a novel deformable image registration (DIR) method for online adaptation of planning organ-at-risk (OAR) delineations to match daily anatomy during hypo-fractionated RT of abdominal tumors.

Materials and methods: For 20 liver cancer patients, planning OAR delineations were adapted to daily anatomy using the DIR on corresponding repeat CTs. The DIR's accuracy was evaluated for the entire cohort by comparing adapted and expert-drawn OAR delineations using geometric (Dice Similarity Coefficient (DSC), Modified Hausdorff Distance (MHD) and Mean Surface Error (MSE)) and dosimetric (D_{max} and D_{mean}) measures.

Results: For all OARs, DIR achieved average DSC, MHD and MSE of 86%, 2.1 mm, and 1.7 mm, respectively, within 20 s for each repeat CT. Compared to the baseline (translations), the average improvements ranged from 2% (in heart) to 24% (in spinal cord) in DSC, and 25% (in heart) to 44% (in right kidney) in MHD and MSE. Furthermore, differences in dose statistics (D_{max}, D_{mean} and $D_{2\%}$) using delineations from an expert and the proposed DIR were found to be statistically insignificant ($p > 0.01$).

Conclusion: The validated DIR showed potential for online-adaptive radiotherapy of abdominal tumors as it achieved considerably high geometric and dosimetric correspondences with the expert-drawn OAR delineations, albeit in a fraction of time required by experts.

Recent studies demonstrated that re-planning based on daily acquired computed tomography (CT) scans can improve dose delivery in stereotactic body radiotherapy (SBRT) of liver and pancreatic tumors [1–5]. Online-adaptive SBRT, including online re-planning or verification of computed dose statistics of organs-at-risk (OARs), is being increasingly employed in radiotherapy as well as proton therapy centers with the use of in-room CT-on-rails or MR imaging. Such a setup, however, requires accurate and fast delineations of the relevant organs [1,6]. Manual delineation of OARs in a repeat CT is tedious and time-consuming, and is therefore not suitable for online adaptive strategies, where time is a major limiting factor. Automatic OAR delineation using deformable image registration (DIR) has the potential to obtain adapted delineations from the planning CT within a permissible time-frame [1,7].

DIR of abdominal organs has proven to be particularly challenging due to large day-to-day deformations and moderate soft-tissue contrast in CT images [8]. As a result, sparse literature exists on the application of DIR algorithms, especially, for abdominal radiotherapy. Brock et al. [9] assessed the accuracy of multiple DIR algorithms on data from different anatomical sites including abdomen. Liu et al. [10] evaluated a free-form DIR algorithm based on calculus of variations [11] using a physical deformable abdominal phantom. Hoffman et al. [12] tested a commercially available DIR tool, Velocity AI (Velocity Medical Systems, Atlanta, GA), on abdominal CTs of 5 patients. Most of the proposed methods demonstrated an overall improvement in accuracy, but also showed significantly high errors. In addition, the validation was performed either on phantoms [10] or limited clinical data [9,12], thus lacking sufficient evidence for their application in a clinical setting.

* Corresponding author at: Department of Radiation Oncology, Erasmus MC Cancer Institute, Groene Hilledijk 301, Rotterdam 3075 EA, The Netherlands.
E-mail address: v.gupta@erasmusmc.nl (V. Gupta).

https://doi.org/10.1016/j.radonc.2018.02.014
0167-8140/

The aim of this study is to comprehensively evaluate a fast DIR method for online-adaptive radiotherapy of abdominal tumors, especially with in-room CT on rails as the imaging device. The proposed DIR method enables fast and efficient adaptation of OAR delineations from planning to repeat CT scans. By adopting a multi-resolution scheme, it avoids local minima and achieves fast convergence. As the entire process is also unsupervised, it has the potential for online automated application. Compared to the existing studies, we evaluated it with the clinical data from a sufficiently large patient group of 20 liver cancer patients. The proposed DIR was validated by comparing automatically obtained delineations with the expert drawn gold-standard delineations.

Materials and methods

Patient characteristics and CT acquisition

20 patients (Table 1), diagnosed with liver metastases or hepatocellular carcinoma (HCC) and previously treated with liver SBRT, were included in this retrospective study. The planning CT image acquisition and fractionated treatments involved the positioning of patients in a stereotactic body frame (SBF) (Elekta Instrument AB, Sweden) with abdominal compression to reduce respiratory-induced tumor motion. The tumor and surrounding OARs were delineated in the large planning CTs (including lungs) that were enhanced with contrast in only 7 patients. For rest of the patients, only short (spanning only liver) planning CTs were acquired with contrast. The tumor delineation from such short CTs were transported to the larger non-contrast planning CTs that were used for OAR delineations. Only large planning CTs were used in this study. On each treatment day, prior to dose delivery, a contrast-enhanced CT scan was acquired to establish the position of the tumor in the SBF. This position was used to calculate the setup correction vector for the alignment of the tumor and the treatment beams. The dataset of each patient was acquired using Siemens SOMATOM Sensation Open CT scanner (Erlangen, Germany), and included 1 planning CT and 3 or 6 repeat CT scans with slices of 512×512 pixels, thickness of 2.5 mm and in-plane pixel size of 0.98–1.27 mm.

Table 1
Patient characteristics ($N = 20$).

Gender	
Female	7
Male	13
Age (years)	
Mean ± Standard deviation	65.9 ± 11.3
Median	65
Patient diagnosis	
Liver Metastases (Liver Mets)	17
Hepatocellular carcinoma (HCC)	3
Patients with contrast agent in	
Planning CT	7
Repeat CTs	20
Patients with	
3 treatment fractions	16
6 treatment fractions	4

Gold-standard OAR delineations

Only those OARs that had dose limiting constraints in our institute's treatment protocol were included in this study. Consequently, duodenum (partial small bowel), esophagus, heart, kidneys, liver, spinal cord, and stomach constituted the OAR set that was delineated *de novo* in each repeat CT by one of the observers in the 3-member expert panel. Prior to their acceptance as gold-standard, OAR delineations were extensively reviewed and modified (if required) by the panel to ensure a reliable anatomical representation. Manual delineations of esophagus, stomach and duodenum proved to be the most challenging, primarily due to the ambiguities related to their junctions. To investigate the effect of junction uncertainties on the accuracy measurements, we created another OAR i.e., ESD (Esophagus–Stomach–Duodenum), by combining the existing delineations of these organs.

OAR adaptation in repeat CTs

Translation, rigid alignment and the proposed DIR (Accuray Incorporated, Sunnyvale, CA, USA) were used to automatically adapt OAR delineations from the planning to the repeat CTs (Fig. 1). Translation for each OAR was obtained by utilizing the clinical setup correction required for the alignment of the tumor and treatment beams. Rigid alignment, i.e. an alignment including rotations and translations, was also evaluated [13–15]. Consequently, for each pair of planning and repeat CTs, we obtained 3 different transformation fields corresponding to translation, rigid alignment and the proposed DIR. Such fields, when applied to the planning delineations, resulted in automatically segmented or DIR-adapted OARs in the repeat CTs. The subsequent section provides more detail on the proposed DIR method.

Proposed DIR method

A novel multi-organ and intensity based DIR method was evaluated. A multi-organ method allows simultaneous deformation of multiple organs instead of using an organ specific deformation method. The DIR software evaluated is a standalone version of

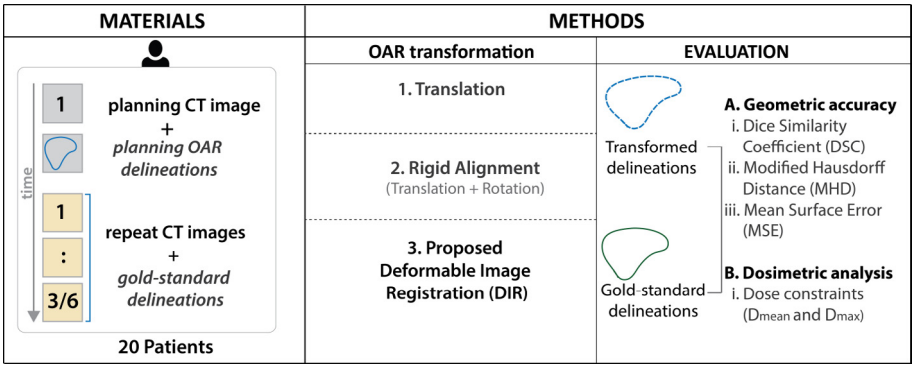

Fig. 1. A schematic illustration of materials and methods. The material column shows the composition of patient data and corresponding delineations. The method column presents the different OAR transformation methods used in this study along with the evaluation strategy that illustrates how the transformed (or automatically adapted) and gold-standard delineations were compared using geometric and dosimetric measures.

the software available in Precision™ Treatment Planning System (Accuray Incorporated, Sunnyvale, CA). With input as a pair of image volumes that are already aligned rigidly in an unsupervised manner, it estimates the global deformation field by optimizing the local normalized cross-correlation coefficient (NCC), where local refers to the use of a patch of voxels instead of the entire image for the computation of NCC. It results in robust image matching even in the presence of local intensity inhomogeneities. If I_r and I_m are the patches from reference and moving images, respectively, the NCC for the corresponding patch is computed as:

$$NCC(I_r, I_m) = \frac{\sum(I_r - \bar{I_r})(I_{m'} - \bar{I_{m'}})}{\sqrt{\sum(I_r - \bar{I_r})^2 (I_{m'} - \bar{I_{m'}})^2}}, \quad (1)$$

where, $I_{m'}$ refers to the moving patch deformed using the proposed DIR and $\bar{I_{m'}}$ refers to its average intensity. A free-form non-rigid transformation is employed opposed to parametric transformation such as a B-spline transformation. It employs a multi-resolution coarse-to-fine (hierarchical) scheme using a smoothed image and at each level there are several iterations. During optimization, the new displacement fields are iteratively estimated, which are then smoothed using a Gaussian filter. This is analogous to the use of a regularization that ensures that neighboring voxels are mapped to similar location by the deformation field, and hence, avoids unrealistic morphology of OAR delineations.

As the computation of 3D deformation fields for registering two abdominal CTs is computationally expensive, its parallelization using GPU (NVIDIA GeForce GTX TITAN) led to drastic improvements in computational efficiency. Each 3D image underwent isotropic Gaussian smoothing and subsampling up to 2 times, providing 3 resolution levels from fine to coarse. This multi-resolution scheme allowed faster processing at coarse resolutions, avoided local minima, and decoupled misalignments at finer scales. At each resolution, the closest match was obtained by iterative updates of the deformation field. We obtained an optimal combination of resolution levels and the respective number of iterations by following a procedure described in the subsequent text.

Parameter selection for the DIR method

Different combinations of resolution levels and corresponding iterations were assessed and ranked on the basis of normalized mutual information (NMI) and modified Hausdorff Distance (MHD). The aim was to find the combination that provided highest NMI and lowest MHD. To this end, NMI determined the best 2 combinations for each resolution, and then, MHD ascertained the best overall combination from this set (Appendix A). In case of a tie, the parameter combination with least processing time was selected as the winner.

Each combination discussed earlier was tested using leave-10-out experiments. Data for each experiment consisted of 20 patients, equally partitioned into mutually exclusive and randomized training and test populations. Over-fitting was avoided by this partitioning as the best parameters from the training data were validated on the unseen test data. We repeated this experiment 5 times to mitigate the impact of a particular random choice of patients in training or test population.

Evaluation methods

Geometric accuracies of the proposed method were compared with translation, i.e. the clinical setup correction, and with rigid transformation, considering expert-drawn delineations as the reference or gold-standard. The differences between DIR-adapted and gold-standard delineations were quantified using three different metrics: (a) Dice Similarity Coefficient (DSC), (b) MHD and (c) Mean Surface Error (MSE). DSC [16] is a commonly used metric to report the extent of spatial overlap between two segmentations. Its values range between 0 (no overlap) and 1 (complete overlap). The MHD [17] is the maximum 3-dimensional symmetric distance between two delineations. It is a variant of the classical Hausdorff distance [18] that offers better discriminatory properties even in the presence of noise. MSE corresponds to the mean surface distance between the closest points on two 3-dimensional delineation sets.

To test if the proposed method performs as well as an expert in meeting clinical dose constraints, the dose–volume histograms (DVH) for each OAR in a repeat CT were re-computed using DIR-adapted and gold-standard delineations. This was accomplished by first aligning the planned dose distribution to the repeat CT using the clinically applied setup vector. D_{mean}, $D_{\max}$ and $D_{2\%}$ were then computed from these DVHs for the auto-segmented and gold-standard OAR delineations. Among these metrics, $D_{\max}$ values show the worst-case test for the use of DIR for detecting constraint violations. Although, recalculation of the dose distribution is preferable over its realignment, we do not expect significant differences due to the high number of beam directions used and the relatively homogeneous anatomical region.

Inter-observer variability

As discussed earlier, manual delineation of OARs on each repeat CT is prone to inter-observer variations i.e., even when manual delineation would be feasible in clinical practice for online-adaptive radiotherapy, the resulting delineations will vary around the gold-standard. It is therefore imperative that the accuracy of the DIR-adapted delineations does not differ considerably from the accuracy of the manual delineations with respect to the gold-standard. To this end, 2 patients were randomly selected (patients 17 and 18). For these patients, the esophagus, duodenum and stomach were delineated by two additional observers (obsA and obsB), since these are the most challenging OARs for delineation. To avoid any bias, each observer made independent delineations *de novo,* with the respective delineations from the planning CT serving as the common reference. Geometric distances (MHD, MSE) and overlap measure (DSC) were computed by comparing contours from obsA, obsB, and the proposed DIR with the gold-standard. Difference vector for each evaluated metric was computed using the following equation:

$$\text{Diff}_{i,j} = \frac{1}{2}\left\{ d(i,j)_{obsA|GS} + d(i,j)_{obsB|GS} \right\} - d(i,j)_{DIR|GS} \quad (2)$$

where i and j correspond to the organ and CT image, respectively, for which the metric, d, is being computed. GS and DIR in Eq. (2) refer to the gold-standard and the proposed DIR, respectively.

Results

Parameter tuning

An overall reduction of 16% in MHD and 63% in processing time was achieved when iterations at multiple resolutions were employed as opposed to the iterations at the original resolution. Summarized results from these experiments are shown in Appendix A.

Geometric accuracy

The geometric accuracies of translation (baseline), rigid alignment and the proposed DIR method for all the OARs across the entire cohort are summarized in Table 2. Considerable

Table 2
Geometric accuracy measured with different metrics for all OARs in the entire cohort.

	DSC			MHD (mm)			MSE (mm)		
	Mean ± SD	Median	IQR	Mean ± SD	Median	IQR	Mean ± SD	Median	IQR
Translation	0.78 ± 0.05	0.78	0.75–0.81	3.21 ± 0.94	2.98	2.71–3.76	2.87 ± 0.67	2.78	2.45–3.19
Rigid Alignment	0.8 ± 0.04 ($p \gg 0.01$)	0.80	0.77–0.83	2.95 ± 0.94	2.91 ($p \gg 0.01$)	2.32–3.38	2.62 ± 0.57	2.52 ($p \gg 0.01$)	2.16–2.97
Proposed DIR	0.86 ± 0.03	0.88 ($p \ll 0.01$)	0.83–0.89	2.06 ± 0.89	1.86 ($p \ll 0.01$)	1.53–2.14	1.72 ± 0.43	1.62 ($p \ll 0.01$)	1.47–1.77

Abbreviations: DSC: Dice Similarity Coefficient, MHD: Modified Hausdorff Distance, MSE: Mean Surface Error; SD: Standard Deviation, IQR: Inter-quartile range; *p*-values were obtained using a 2-sided Wilcoxon signed rank test on individual values for each OAR of every patient, and show the results of comparison with the corresponding values for baseline (translation). Significant tests between rigid alignment and DIR for all three measures revealed significant improvements ($p \ll 0.01$).

improvements compared to rigid alignment indicate that the addition of only rotation proved insufficient in recovering day-to-day anatomical changes. To determine whether these improvements were significant when compared with the baseline, a 2-sided Wilcoxon signed ranked test was performed with a conservative level of 1% significance, and its results are reported in the parentheses for respective metrics in Table 2. As shown by the *p*-values, only marginal improvements were achieved by rigid alignment, whereas, with the proposed DIR method, improvements were statistically significant. Of particular importance is the consistency of the proposed method in achieving high spatial correspondence across all organs and patients, as indicated by the evaluated metrics (DSC, MHD, and MSE) for the entire cohort.

Similar trends were observed when metrics for each organ were analyzed. The overall improvements for all organs ranged from 2% (heart) to 24% (spinal cord) in DSC, and 25% (heart) to 44% (right kidney) in MHD and MSE. For heart, kidneys, liver, spinal cord, and stomach, DSC above 0.9 was achieved. Appendix B displays improvements in DSC for each patient. The mean MHD for all organs, except duodenum and stomach, was below 2 mm as shown in Fig. 2. MHD corresponding to ESD in Fig. 2 indicates minor improvements in mean, when compared to duodenum and stomach. MSE follows the same pattern for these OARs (Appendix B). These improvements are also reflected in the visual illustration (Fig. 3) of typical kidney and stomach delineations with the compared methods.

Dosimetric analysis

We compared the dose statistics of all the OARs with both gold-standard and DIR delineations to investigate if the potential dose constraint violations can be automatically detected. As shown in

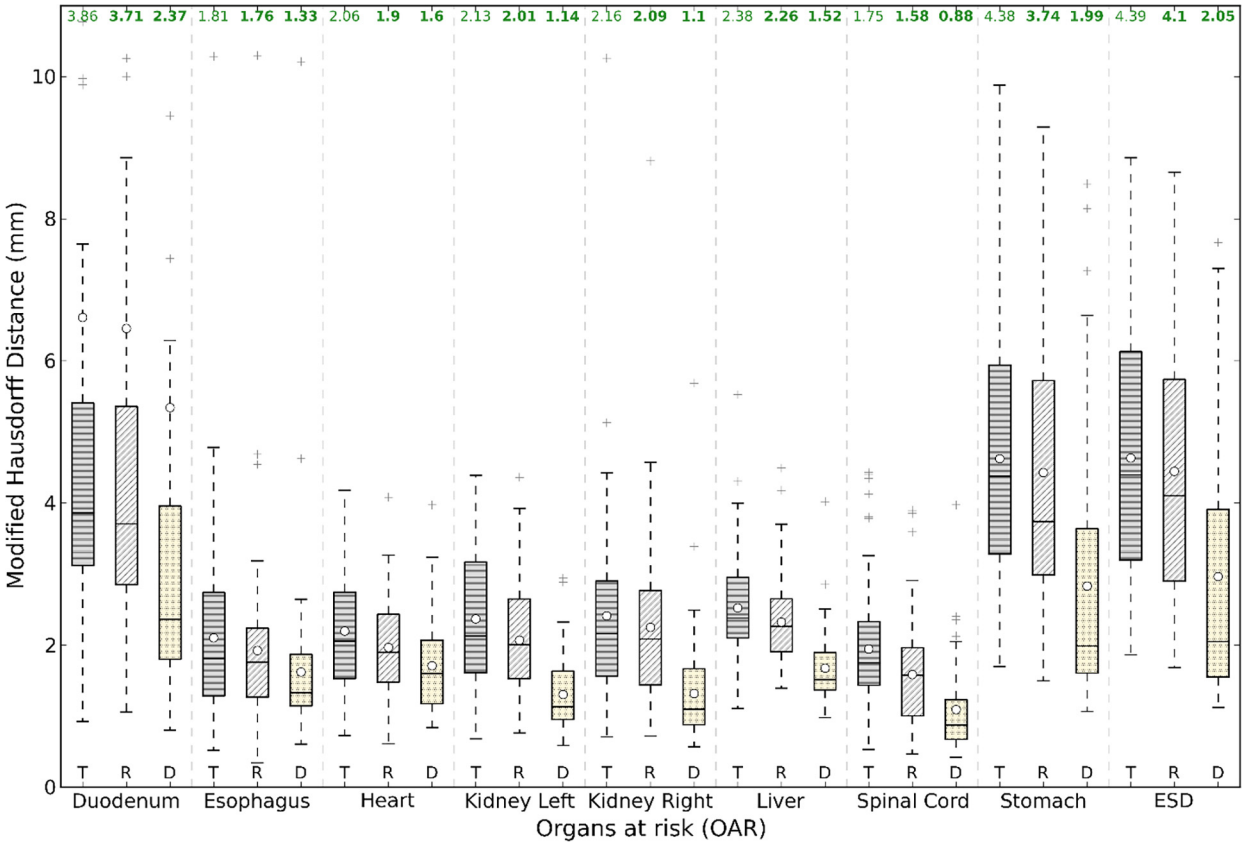

Fig. 2. Distribution of Modified Hausdorff Distances (mm) for each OAR across the entire cohort. The numbers at the top of each box represent median values, and the circles represent mean values. The whiskers are 1.5 times the inter-quartile range from the top/bottom of the box, and values outside this range are plotted individually as outliers ('+'). Significantly lower spread is achieved with the DIR as shown by the 25th and 75th percentiles. Corresponding ranges for translation and rigid alignment show a large variation. *Abbreviations:* ESD = Joined Esophagus, Stomach, and Duodenum; T = Translation, R = Rigid alignment, D = Deformable image registration.

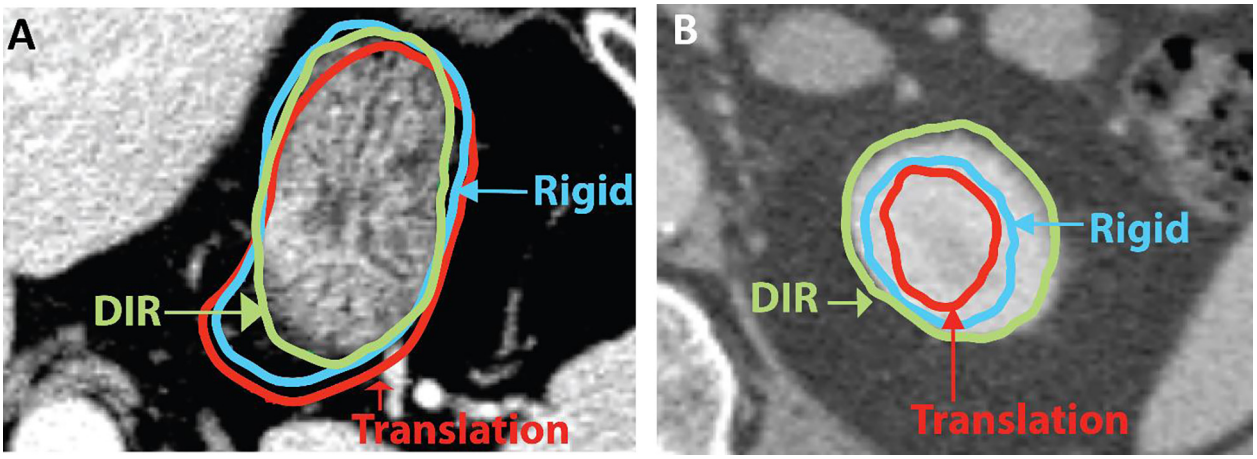

Fig. 3. Examples of (A) Stomach and (B) Left Kidney delineations obtained using translation, rigid alignment and deformable image registration (DIR). Both translation and rigid alignment fail to follow the complex geometric changes in abdominal anatomy. In contrast, DIR provides a close approximation of the daily anatomy in stomach and left kidney.

Table 3
Comparison of OAR dose (Gy) obtained with Gold-standard (GS) and deformed delineations (proposed DIR) for the entire cohort.

OAR	D_{mean}			D_{max}			$D_{2\%}$		
	GS	DIR	Signed difference (GS-DIR)	GS	DIR	Signed difference (GS-DIR)	GS	DIR	Signed difference (GS-DIR)
Duodenum	2.4 ± 0.4	2.4 ± 0.2	0.1 ± 0.3	10.9 ± 3.1	11.5 ± 1.7	−0.6 ± 2.9	6.7 ± 1.6	7.1 ± 1.1	−0.4 ± 1.5
Esophagus	5.9 ± 0.6	6.1 ± 0.5	−0.1 ± 0.5	15.0 ± 1.6	14.8 ± 1.1	0.3 ± 1.1	12.2 ± 1.3	12.3 ± 1.0	−0.2 ± 0.8
Heart	3.0 ± 0.3	3.1 ± 0.3	−0.1 ± 0.1	19.6 ± 1.3	19.5 ± 1.3	0.1 ± 0.9	11.5 ± 0.8	11.6 ± 0.8	−0.1 ± 0.4
Kidney Left	1.5 ± 0.1	1.5 ± 0.1	0.0 ± 0.0	7.0 ± 0.7	7.0 ± 0.7	0.0 ± 0.2	4.7 ± 0.4	4.7 ± 0.4	0.0 ± 0.1
Kidney Right	3.7 ± 0.4	3.5 ± 0.4	0.1 ± 0.1	21.3 ± 1.9	21.4 ± 1.5	−0.1 ± 0.9	16.8 ± 1.3	15.3 ± 1.3	0.2 ± 0.6
Liver	15.8 ± 0.3	16.1 ± 0.4	−0.3 ± 0.1	65.3 ± 0.4	65.3 ± 0.4	0.0 ± 0.0	59.2 ± 0.6	59.6 ± 0.5	−0.4 ± 0.4
Spinal Cord	3.3 ± 0.3	3.4 ± 0.3	−0.1 ± 0.1	10.7 ± 0.5	11.1 ± 0.6	−0.3 ± 0.2	8.6 ± 0.5	8.8 ± 0.5	−0.2 ± 0.1
Stomach	4.7 ± 0.3	4.7 ± 0.2	0.0 ± 0.2	17.9 ± 2.2	16.9 ± 1.2	1.0 ± 1.6	11.4 ± 0.5	11.5 ± 0.5	−0.1 ± 0.2

the Table 3, the standard deviations in dose differences were minimal for the majority of organs, except the D_{max} for stomach and duodenum. A 2-sided Wilcoxon signed rank test showed insignificant differences ($p > 0.01$) for all OARs. Furthermore, no trend was observed, which shows the absence of any systematic bias in the proposed method.

Inter-observer variability

Differences with respect to the gold-standard were computed using observer and DIR-adapted delineations for each metric (Eq. (2)). On average, the differences in MHD and MSE were found to be less than 0.53 mm, while the difference in DSC measure was 3%. In total, 113 ± 30 min were required by each observer to delineate esophagus, duodenum and stomach in 6 repeat CTs.

Processing time and implementation details

Translation and rigid alignment required 7–10 s to compute the transformation field, while the DIR required 10–12 s. Including the transformation of planning delineations, a new set of delineations were obtained for each repeat CT in less than 20 s with the proposed DIR. It is worth noting that while the DIR was optimized for speed using a GPU, no optimization was performed for transforming the planning delineations. With an optimized implementation of the complete framework, a further reduction can be achieved in the overall processing time. All the experiments were performed using an in-house image analysis framework and MATLAB (version R 2014b) on a system with Intel i-7 3.6 GHz quad core processor, 32 GB RAM and GeForce GTX Titan (Nvidia).

Discussion

Online-adaptive SBRT in the current setup as well as with the increasingly popular in-room CT-on-rails requires accurate and fast tracking of daily anatomy. The proposed DIR method provides an effective tool to achieve this goal. As discussed earlier, several DIR methods have been evaluated either for other anatomical sites or with relative sparse data [8–12]. To the best of our knowledge, this is the first comprehensive study for validating a DIR method using both geometric and dosimetric measures on such a large clinical dataset from the abdominal region. Our results showed that the validated DIR achieved high spatial and dosimetric correspondences with the gold-standard. Such correspondences were achieved despite large intensity differences caused by the absence of a contrast agent between the planning and repeat CT images in 13 out of 20 patients. Considerable improvements in mean (10% in DSC, 36% in MHD, and 40% in MSE) were observed when compared to clinical setup corrections (translation) for all the OARs. In contrast, the existing methods have been shown to provide average errors of 0.8–5.9 mm in clinical data (9), 3.7 ± 1.8 mm (10) in software phantom and more than 4 mm in 41% of the cohort consisting of 5 patients with abdominal tumor (12).

Geometric accuracy alone, however, is not the only requirement for online-adaptive treatments. Stability and processing time are also critical. The proposed method's stability with respect to parameter tuning was demonstrated by the insignificant variations (σ = 0.176 mm) in MHD, when the same parameter combinations were used for randomized patient samples. As a result, a predetermined set of optimal parameters can be used for the entire cohort as opposed to time-intensive parameter tuning for each patient [13].

The inter-observer agreements provided further indication of the proposed method's feasibility in a clinical setting. The results showed that GS vs DIR errors were close to the inter-observer errors, even for the most challenging organs and the overall performance will be similar to those of multiple observers, albeit within a fraction of time used by each observer (2 h vs <1 min). In addition, while the manual segmentation is prone to inter-observer variability, the auto-segmentation method produces the same output over multiple executions, using a predetermined set of parameters for the entire cohort.

Manual delineation of OARs is essential, yet practically infeasible, for optimal dose delivery during online and offline adaptive treatment or re-planning based on daily anatomy [6]. DIR can be a viable alternative [1,6], as indicated also by the dosimetric analyses of the proposed method. The majority of OARs (Table 3) demonstrated trivial differences when compared to the gold-standard. Except for stomach and duodenum, 95% (2σ) of the dose differences are within 0–2.2 Gy, and hence, can be compensated for by artificially lowering the dose constraints to take into account the inaccuracy in computing the dose statistics. Artificially lowering the threshold is meant to alert the clinician that there may be a constraint violation and he/she can verify if this is indeed the case. Similar compensation will, however, require that constraints be lowered by 5.8 Gy and 3.2 Gy for duodenum and stomach, respectively. Such clinically impractical requirement (especially for duodenum) makes it imperative that manual check and correction of DIR-adapted delineations are performed for these organs before their use in dose computation. For other organs, however, the evaluation showed significant improvements, and hence, it may not be necessary to edit their delineations.

The current study has two major limitations. First, since OAR intensities differ significantly in planning and repeat CT images of a few patients, both, gold-standard and DIR-adapted delineations are affected. The uncertainties associated with the former were mitigated using the rigorous reviews by 3-member expert panel. DIR-adapted delineations for the non-contrast data however led to significant geometric differences ($0.05 > p > 0.01$) at 5% significance level (Appendix B). Further analyses revealed that such differences resulted not only from the absence of contrast, but also from the unusually large OAR deformations in 3 patients (patients 16, 18, 20). Additional anatomical distortions occurred in these patients due to: (a) the failure of abdominal compression as all of them were extremely thin and (b) underlying liver diseases. In addition, patient 18 had gall bladder stones that severely affected the surrounding anatomy. As shown in Appendix B, major improvements were achieved when these patients were removed from the evaluation.

Second, the proposed DIR is especially susceptible to errors in duodenum and stomach (Fig. C1 in Appendix C), which often undergo complex deformations. The indistinct junctions (esophagus–stomach, stomach–duodenum), which are not clear even for the expert clinician, add further complexity. We investigated whether the ambiguity with respect to organ junctions can be avoided if they were combined. MHD values for ESD (Fig. 2) indicated that although the average and spread were reduced marginally, significant improvements in overall results could not be achieved using such an approach and observation that can be attributed to the larger size of duodenum and stomach along with their corresponding outliers.

The lack of comparison with existing studies may be considered another limitation. However, comparing different methods in a consistent way is a challenging task. It requires tuning of DIR parameters to compare the accuracy conditional to the same objectives, e.g. speed. Therefore, this falls outside the scope of the current work, but is highly recommended for future work.

One of the potential approaches to overcome current limitations could involve region specific regularization to counter large and complex deformations, especially, in stomach and duodenum. Such solutions will be investigated in our future work. Furthermore, as the current DIR employs an intensity-based matching, the superior contrast of MR imaging can be utilized for improving the registration of soft-tissue organs (mainly, stomach and duodenum).

Conclusion

We validated a novel DIR method that adapts planning OAR delineations according to anatomy-of-the-day in abdominal region within 20 s. Results of its application on clinical data from 20 liver cancer patients demonstrated its high spatial and dose correspondence to the gold-standard. With its high accuracy, robustness and short processing time (20 s), the validated DIR has potential to enable online-adaptive SBRT of abdominal tumors, especially in increasingly popular setups that employ in-room CT or MR imaging to track daily anatomical changes.

Funding support

Partial funding for this work was provided by Accuray Incorporated as part of a research collaboration with Erasmus MC Cancer Institute.

Conflict of interest

Andriy Myronenko, Petr Jordan, and Calvin Maurer are employed at Accuray Incorporated, 1310 Chesapeake Terrace, Sunnyvale, CA 94089, USA.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.radonc.2018.02.014.

References

[1] Leinders SM, Breedveld S, Méndez Romero A, et al. Adaptive liver stereotactic body radiation therapy: automated daily plan reoptimization prevents dose delivery degradation caused by anatomy deformations. Int J Radiat Oncol Biol Phys 2013;87:1016–21.
[2] Ahunbay EE, Peng C, Chen G-P, et al. An on-line replanning scheme for interfractional variations. Med Phys 2008;35:3607.
[3] Ahunbay EE, Kimura B, Liu F, et al. Comparison of various online strategies to account for interfractional variations for pancreatic cancer. Int J Radiat Oncol Biol Phys 2013;86:914–21.
[4] Liu F, Erickson B, Peng C, et al. Characterization and management of interfractional anatomic changes for pancreatic cancer radiotherapy. Int J Radiat Oncol Biol Phys 2012;83:e423–9.
[5] Li XA, Qi XS, Pitterle M, et al. Interfractional variations in patient setup and anatomic change assessed by daily computed tomography. Int J Radiat Oncol Biol Phys 2007;68:581–91.
[6] Méndez Romero A, Zinkstok RT, Wunderink W, et al. Stereotactic body radiation therapy for liver tumors: impact of daily setup corrections and day-to-day anatomic variations on dose in target and organs at risk. Int J Radiat Oncol Biol Phys 2009;75:1201–8.
[7] Janssens G, Orban de Xivry J, Fekkes S, et al. Evaluation of nonrigid registration models for interfraction dose accumulation in radiotherapy. Med Phys 2009;36:4268.
[8] Pace DF, Aylward SR, Niethammer M. A locally adaptive regularization based on anisotropic diffusion for deformable image registration of sliding organs. IEEE Trans Med Imaging 2013;32:2114–26.
[9] Brock KK. Results of a multi-institution deformable registration accuracy study (MIDRAS). Int J Radiat Oncol Biol Phys 2010;76:583–96.
[10] Liu F, Hu Y, Zhang Q, et al. Evaluation of deformable image registration and a motion model in CT images with limited features. Phys Med Biol 2012;57:2539–54.
[11] Lu W, Chen M-L, Olivera GH, et al. Fast free-form deformable registration via calculus of variations. Phys Med Biol 2004;49:3067–87.

[12] Hoffmann C, Krause S, Stoiber EM, et al. Accuracy quantification of a deformable image registration tool applied in a clinical setting. J Appl Clin Med Phys 2014;15:237–45.
[13] Fortunati V, Verhaart RF, Angeloni F, et al. Feasibility of Multimodal Deformable Registration for Head and Neck Tumor Treatment Planning. Int J Radiat Oncol Biol Phys 2014.
[14] Andronache A, Krayenbuehl J, Szekely G, et al. Hierarchical enhanced non-rigid registration for target volume correction and propagation for adaptive external beam radiotherapy of carcinoma of the prostate. J Appl Clin Med Phys 2013;14:222–30.
[15] Fabri D, Zambrano V, Bhatia A, et al. A quantitative comparison of the performance of three deformable registration algorithms in radiotherapy. Z Med Phys 2013;23:279–90.
[16] Dice LR. Measures of the amount of ecologic association between species. Ecology 1945;26:297–302.
[17] Dubuisson M, Jain AK. A modified Hausdor distance for object matching. Int Conf Pattern Recognit 1994:566–8.
[18] Hausdorff F. Set theory. Am Math Soc 1957.