Investigating CT to CBCT image registration for head and neck proton therapy as a tool for daily dose recalculation

Guillaume Landry[a)]
Department of Medical Physics, Ludwig-Maximilians-University, Munich D85748, Germany
and Department of Radiation Oncology, Ludwig-Maximilians-University, Munich D81377, Germany

Reinoud Nijhuis
Department of Radiation Oncology, Ludwig-Maximilians-University, Munich D81377, Germany

George Dedes and Josefine Handrack
Department of Medical Physics, Ludwig-Maximilians-University, Munich D85748, Germany

Christian Thieke
Department of Radiation Oncology, Ludwig-Maximilians-University, Munich D81377, Germany

Guillaume Janssens and Jonathan Orban de Xivry
ICTEAM, Université Catholique de Louvain, Louvain-La-Neuve B1348, Belgium

Michael Reiner
Department of Radiation Oncology, Ludwig-Maximilians-University, Munich D81377, Germany

Florian Kamp and Jan J. Wilkens
Department of Radiation Oncology, Technische Universität München, Klinikum rechts der Isar,
Munich D81675, Germany and Physik-Department, Technische Universität München,
Garching D85748, Germany

Chiara Paganelli, Marco Riboldi, and Guido Baroni
Dipartimento di Elettronica Informazione e Bioingegneria, Politecnico di Milano, Milan 20133, Italy

Ute Ganswindt and Claus Belka
Department of Radiation Oncology, Ludwig-Maximilians-University, Munich D81377, Germany

Katia Parodi
Department of Medical Physics, Ludwig-Maximilians-University, Munich D85748, Germany

Purpose: Intensity modulated proton therapy (IMPT) of head and neck (H&N) cancer patients may be improved by plan adaptation. The decision to adapt the treatment plan based on a dose recalculation on the current anatomy requires a diagnostic quality computed tomography (CT) scan of the patient. As gantry-mounted cone beam CT (CBCT) scanners are currently being offered by vendors, they may offer daily or weekly updates of patient anatomy. CBCT image quality may not be sufficient for accurate proton dose calculation and it is likely necessary to perform CBCT CT number correction. In this work, the authors investigated deformable image registration (DIR) of the planning CT (pCT) to the CBCT to generate a virtual CT (vCT) to be used for proton dose recalculation.
Methods: Datasets of six H&N cancer patients undergoing photon intensity modulated radiation therapy were used in this study to validate the vCT approach. Each dataset contained a CBCT acquired within 3 days of a replanning CT (rpCT), in addition to a pCT. The pCT and rpCT were delineated by a physician. A Morphons algorithm was employed in this work to perform DIR of the pCT to CBCT following a rigid registration of the two images. The contours from the pCT were deformed using the vector field resulting from DIR to yield a contoured vCT. The DIR accuracy was evaluated with a scale invariant feature transform (SIFT) algorithm comparing automatically identified matching features between vCT and CBCT. The rpCT was used as reference for evaluation of the vCT. The vCT and rpCT CT numbers were converted to stopping power ratio and the water equivalent thickness (WET) was calculated. IMPT dose distributions from treatment plans optimized on the pCT were recalculated with a Monte Carlo algorithm on the rpCT and vCT for comparison in terms of gamma index, dose volume histogram (DVH) statistics as well as proton range. The DIR generated contours on the vCT were compared to physician-drawn contours on the rpCT.
Results: The DIR accuracy was better than 1.4 mm according to the SIFT evaluation. The mean WET differences between vCT (pCT) and rpCT were below 1 mm (2.6 mm). The amount of voxels passing 3%/3 mm gamma criteria were above 95% for the vCT vs rpCT. When using the rpCT contour set to derive DVH statistics from dose distributions calculated on the rpCT and vCT the differences, expressed in terms of 30 fractions of 2 Gy, were within [−4, 2 Gy] for parotid glands (D_{mean}), spinal

cord ($D_{2\%}$), brainstem ($D_{2\%}$), and CTV ($D_{95\%}$). When using DIR generated contours for the vCT, those differences ranged within [−8, 11 Gy].

Conclusions: In this work, the authors generated CBCT based stopping power distributions using DIR of the pCT to a CBCT scan. DIR accuracy was below 1.4 mm as evaluated by the SIFT algorithm. Dose distributions calculated on the vCT agreed well to those calculated on the rpCT when using gamma index evaluation as well as DVH statistics based on the same contours. The use of DIR generated contours introduced variability in DVH statistics. ○

Key words: adaptive radiotherapy, cone-beam CT, deformable image registration, head and neck cancer, intensity modulated proton therapy

1. INTRODUCTION

The observation that intensity modulated radiation therapy (IMRT) of head and neck (H&N) lesions shows reduced toxicity[1] due to better dose conformity and sparing of organs at risk (OAR) than conformal therapy suggests that proton therapy, with its even higher dose conformity, may be beneficial to patients. Intensity modulated proton therapy (IMPT) planning studies support this hypothesis by demonstrating improved sparing of OARs compared to IMRT.[2,3]

The large anatomical changes which may be observed during fractionated H&N radiation therapy[4] suggest that the use of repeated computed tomography (CT) scanning may be essential during IMRT (Ref. 5) and IMPT (Ref. 6) for evaluation of potential dose distribution degradation. As the positioning of the Bragg peak is sensitive to the patient's three dimensional stopping power ratio distribution as estimated by a planning CT scan, the use of a CT-on-rail inside the treatment room, as investigated in the context of passively scattered proton therapy of the prostate,[7] would offer the optimal data for proton dose recalculation. However, such installations are not common, require a couch movement between imaging and treatment, and increase the complexity of the treatment workflow.

Imaging directly at the treatment position, with the use of gantry-mounted cone beam CT (CBCT), is considered for proton therapy setup corrections[8] and is currently being introduced by vendors. The lower quality of CBCT images[9] and high scatter-to-primary photon ratio[10] render them unsuitable for direct calculation of proton therapy dose distributions, as opposed to photon therapy[11] where uncorrected CBCT images may be used within certain tolerance levels.

Preliminary studies have investigated CT number correction for CBCT images by image processing alone.[12] It may however be possible to obtain accurate stopping power estimates by performing deformable image registration (DIR) between the planning CT (pCT) and CBCT. This has been explored by several groups in the context of adaptive photon therapy for photon dose recalculation and automatic contour propagation.[13–18] A recent phantom study from our group suggested that the approach may provide stopping power distributions which are equivalent to those obtained by a planning CT with proton range differences below 2%.[19] However, that study was restricted to the evaluation of the stopping power distribution and did not explore the use of the vector field for contour propagation for a more profound analysis.

In this work, we performed an investigation of CT to CBCT DIR using clinical patient data from 6 IMRT H&N cancer patients where CBCT and replanning CT scans were acquired. We aimed at generating data sets allowing accurate proton dose recalculation and dose volume histogram (DVH) parameter evaluation but did not investigate dose accumulation. We evaluated the geometrical accuracy of the DIR using a corresponding-feature identification algorithm. Furthermore, CT number accuracy was evaluated by comparison to the planning CT and replanning CT. A dosimetric evaluation was performed by comparing dose distributions recalculated on the results of CT to CBCT DIR and on the corresponding replanning CT. Finally, deformed contours were compared against independent physician delineation on the replanning CT scans as well as on the deformed CT scans.

2. MATERIALS AND METHODS

2.A. Clinical data

Datasets from 6 IMRT H&N cancer patients containing a contoured planning CT scan (pCT), weekly CBCT scans, and a contoured replanning CT (rpCT) scan were employed in this study. For some patients, an additional, later rpCT was available; however, we only made use of one rpCT per patient in this study. We focused on the CBCT scans closest in time to the rpCT. Contouring was performed by the same radiation oncologist. The cases were selected using the following criteria: no metal artefacts in the CTV, and full CTV coverage by the CBCT field of view (FOV). The planning CT scanner was a Toshiba Aquilion LB (Toshiba Medical Systems, the Netherlands) and images were reconstructed on a $1.074 \times 1.074 \times 3$ mm grid. The CBCT scanner was the on-board imager of an Elekta Synergy Linac equipped with XVI R4.5 (Elekta, Sweden) and images were reconstructed on a $1 \times 1 \times 1$ mm grid. The clinical protocol for H&N patients used 18.3 mA s at 100 kV with CTDI$_{vol}$ of 0.6 mGy. No beam shaping filter was used and the S20 collimation was employed. Table I presents the characteristics of the patient datasets employed in this study.

2.B. Registration

A Morphons algorithm,[20,21] implemented in the REGGUI DIR package,[22] was chosen in this study to perform 3D image

TABLE I. Characteristics of the patient datasets employed in this study. The times between planning and replanning CT acquisition Δt_{rpCT} and between planning CT and CBCT acquisition Δt_{CBCT} are reported.

	Age	Sex	Tumor site	TNM stage	Δt_{rpCT} (days)	Δt_{CBCT} (days)
Pat1	65	M	Larynx	pT2pN0M0	51	50
Pat2	54	F	Hypopharynx, esophagus	cT4cN2M0	39	41
Pat3	71	M	Larynx	pT1bN0M0	34	35
Pat4	87	M	Hypopharynx	cT2cN2bM0	33	34
Pat5	49	M	Nasopharynx	cT2cN2bM0	40	40
Pat6	42	M	Larynx	pT2bpN1M0	44	41

registration between the pCT and CBCT. The Morphons algorithm is a nonparametric intensity-based registration method with similarities to the Demons algorithm.[23] It is applied in a scale-space scheme. Starting from a coarse scale, local phases at that scale in both images are estimated using quadrature filters. Phase differences between the two images are computed from the product between filter outputs and the optimal displacement is found by minimizing a $L2$-norm on all possible directions. The dense vector field is regularized at each iteration through a normalized convolution with a Gaussian kernel weighted by a value expressing the certainty the method has on the local field estimation. After convergence at each scale, the field is up-sampled and the method is applied at a new scale with better resolution. A thorough description of the method can be found in Janssens *et al.*[22] The present study used eight scales with ten iterations for the six coarsest scales and two iterations for the last two higher scales, and a Gaussian regularization filter of 1.2 voxels standard deviation.

Initially designed to follow bony structures in the case of surgery,[20] the algorithm has since then been tested and validated in the context of radiotherapy for dose deformation purposes in phantoms[24] and to estimate breathing motion in 4D-CT of the lungs with and without contrast enhancement.[22] The latter investigations highlighted that the Morphons algorithm is more robust to changes in intensity than the Demons algorithm thanks to its use of the local phase difference as metric. This is the reason for the choice of the method in the context of CT-CBCT registration.

The method employed here was the same as in Landry *et al.*[19] An automatic rigid translation registration was employed to align the CBCT and pCT, mimicking an online position correction protocol. Prior to rigid registration, the CBCT was cropped to eliminate the conical sections at the inferior end of the FOV. Rotations were not permitted to replicate the limited degrees of freedom of clinical alignment procedures. Following rigid registration, the pCT was resampled to the CBCT image grid which has a higher resolution. DIR of the resampled pCT to the CBCT was subsequently performed using Morphons. The procedure yielded a deformed pCT in the coordinate system and grid spacing of the CBCT. We called this image the virtual CT, as in Peroni *et al.*[15] As this image is in the CBCT coordinate system, we add the CBCT subscript (vCT$_{\text{CBCT}}$). In general, images in the CBCT coordinate system and grid spacing have the CBCT subscript in this paper. The corresponding rpCT, usually taken within 1–2 days of the CBCT, was aligned with the CBCT using a rigid translation registration, yielding the resampled rpCT$_{\text{CBCT}}$, for direct evaluation of the vCT$_{\text{CBCT}}$.

The deformation vector field resulting from DIR of the pCT$_{\text{CBCT}}$ to the CBCT$_{\text{CBCT}}$ is defined in the coordinate system and grid spacing of the CBCT, and not directly applicable to the pCT in its own coordinate system and grid spacing. Thus, the vector field was resampled onto the original pCT grid to allow deformation of the original pCT. In regions outside the CBCT FOV, the vector field was zero; in practice, this is equivalent to stitching the original pCT to the vCT outside the CBCT FOV. We applied the vector field to the contours as well to provide a DIR-contoured virtual CT in the coordinate system and grid spacing of the pCT (vCT$_{\text{pCT}}$). In general, in this paper, images in the pCT coordinate system and grid spacing have the pCT subscript. The vCT$_{\text{pCT}}$ can be used interchangeably with the pCT$_{\text{pCT}}$ to evaluate the dosimetric impact of anatomical changes by recalculating the dose distribution of a plan, without modification to that plan. To evaluate the vCT$_{\text{pCT}}$, the rpCT and its contours were also rigidly aligned to it, yielding a rpCT$_{\text{pCT}}$.

2.C. Deformation accuracy evaluation

The accuracy of the registration was first evaluated in the CBCT coordinate system using the scale invariant feature transform (SIFT) algorithm[25–27] described in Paganelli *et al.*,[28] implemented in PLASTIMATCH.[29] The algorithm compares two images and identifies points corresponding to matching features on each image. The vCT$_{\text{CBCT}}$, pCT$_{\text{CBCT}}$, rpCT$_{\text{CBCT}}$, and CBCT were intercompared in terms of the distances between matching features.

To ensure CT numbers were not altered by the deformation, the mean CT numbers inside the contours of the parotid glands, spinal cord, and lower mandible from the pCT$_{\text{pCT}}$, vCT$_{\text{pCT}}$, and rpCT$_{\text{pCT}}$ were compared.

To evaluate the accuracy of proton range estimation, the CT numbers were converted to water equivalent stopping power, and the water equivalent thickness (WET) was calculated for projections at 0° starting from the patient's skin surface down to a normal plane coinciding with the central axis of the image. The 2D WET distributions from the pCT$_{\text{CBCT}}$ and vCT$_{\text{CBCT}}$ were compared to that of the rpCT$_{\text{CBCT}}$ by means of gamma analysis with criteria 3 mm/3 mm WET. We chose 3 mm as it corresponds to the pCT slice thickness. We employed a signed-gamma implementation where the gamma index takes the sign of the dose difference.[30]
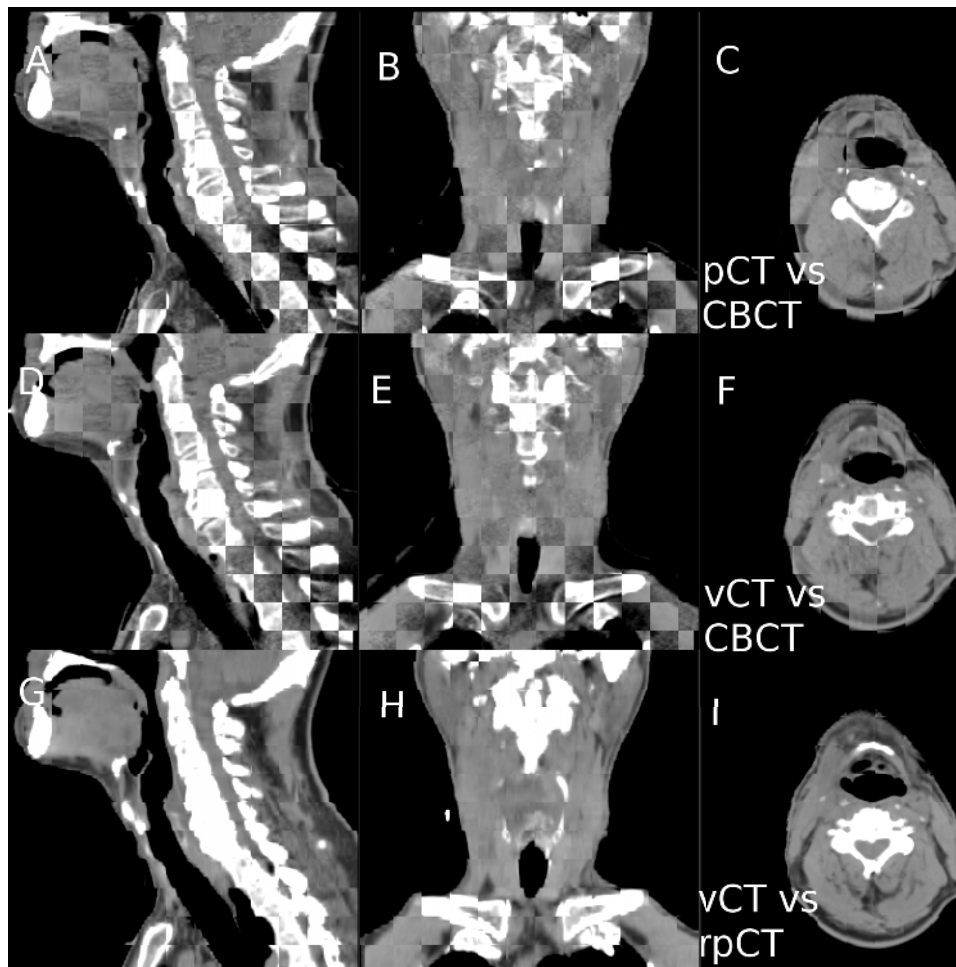
FIG. 1. Checkerboard comparison of (A), (B), and (C), the pCT_{CBCT} vs $CBCT_{CBCT}$, (D), (E), and (F) vCT_{CBCT} vs $CBCT_{CBCT}$, and (G), (H), and (I) vCT_{CBCT} vs $rpCT_{CBCT}$ for Pat1.

2.D. Structure deformation

The validity of the DIR generated contours for the vCT_{pCT} was evaluated by comparing them to contours generated by a trained radiation oncologist. Both the vCT_{pCT} and the corresponding $rpCT_{pCT}$ were delineated manually and used as reference to evaluate the DIR generated contours. We focused on the following structures: CTVs, PTVs, parotid glands, mandible, spinal cord, and brainstem. Contours were compared using 3D Slicer in terms of the dice similarity coefficient (DSC), the mean, and 95% range of the distribution of Hausdorff distance (d) between the two contour sets, and center of mass (COM) shifts for the parotid glands, GTV, and CTVs. As the physician-drawn contours may extend beyond the treatment area and CBCT FOV, we cropped all datasets to the area of validity of the vCT_{pCT} (which corresponds to the CBCT FOV). This was relevant for the brainstem and the spinal canal but did not affect other structures.

2.E. Dosimetric evaluation

Simultaneous integrated boost (SIB) IMPT plans with two dose levels (1.8 and 2 Gy per fraction, 30 fractions) were generated using a CERR-based[31] research treatment planning system (TPS) featuring a proton pencil beam algorithm.[32] The plans consisted of 4 beams at gantry angles of 30°, 60°, 300°, and 330° on the International Electrotechnical Commission scale and were optimized using the pCT_{pCT} and corresponding structures. When optimizing the plan, we considered the parotid glands (mean dose <25 Gy or <32 Gy if overlapping with PTV), spinal cord, and brainstem (each maximum dose as $D_{2\%} < 54$ Gy).

As our TPS did not allow recalculation of plans on modified anatomy, Monte Carlo (MC) dose recalculations were performed by exporting the proton fluence from the TPS to a GEANT 4 dose calculation engine where the CT scans were imported and converted into mass density and elemental composition using lookup tables established for the scanner used in this study. The MC code scored dose to water on the CT voxel grid by weighting energy deposits with the relative mass stopping power on the fly. The number of protons varied between 20×10^6 and 60×10^6, corresponding to 0.2% of the total number of protons, depending on the PTV volume. Doses were recalculated for the pCT_{pCT}, vCT_{pCT}, and $rpCT_{pCT}$ and DVH statistics for the CTV, PTV ($D_{95\%}$), and OARs (D_{mean} for parotid and $D_{2\%}$ for brain stem and spinal cord) were extracted. The contours obtained from DIR were used for the vCT_{pCT} DVH analysis together with the physician-drawn

TABLE II. Median and interquartile range (in parenthesis) of the distribution of distances between SIFT-identified feature pairs for the six possible pairings of the pCT_{CBCT}, vCT_{CBCT}, $rpCT_{CBCT}$, and $CBCT_{CBCT}$ for the six patients investigated in this study. The last two rows compare body contour volumes and COM between the pCT_{CBCT} and the $rpCT_{CBCT}$. Boldface table entries relate to the vCT_{CBCT} vs $CBCT_{CBCT}$ comparison.

	Pat1	Pat2	Pat3	Pat4	Pat5	Pat6
	Distance in mm					
vCT vs CBCT	**1.4(1.7)**	**1.0(1.4)**	**1.0(1.4)**	**1.0(1.4)**	**1.0(1.4)**	**1.0(1.4)**
pCT vs CBCT	2.2(1.9)	5.4(3.1)	3.0(1.9)	3.0(2.0)	2.2(1.6)	1.4(1.2)
pCT vs vCT	2.4(1.6)	5.4(2.8)	3.0(2.0)	2.9(2.0)	1.7(1.0)	1.4(1.1)
rpCT vs CBCT	2.0(1.4)	2.4(1.6)	1.4(0.8)	2.8(1.3)	1.4(1.2)	1.7(0.9)
rpCT vs vCT	2.2(1.8)	2.4(1.1)	1.4(1.2)	2.4(1.3)	1.0(0.7)	1.4(1.2)
rpCT vs pCT	2.2(2.3)	3.7(2.7)	3.6(3.9)	2.2(1.9)	2.2(1.6)	1.4(1.0)
ΔV pCT rpCT body contour (%)	0.3	−8	−4	0.6	−5	−3
COM shift rpCT pCT (mm)	0.3	1.9	3.2	1.8	1.5	0.1
ΔV vCT rpCT body contour (%)	−2	−0.2	−0.2	0.2	1.3	0.7
COM shift rpCT vCT (mm)	0.5	0.5	1.3	1.5	0.6	0.3

$rpCT_{pCT}$ contours following rigid registration of the $rpCT_{pCT}$ to the vCT_{pCT}. Additionally, the contours from physician delineation of the vCT_{pCT} were also considered. The dose distributions from MC recalculations obtained using the vCT_{pCT} and pCT_{pCT} were compared to that obtained from the $rpCT_{pC}$ using gamma analysis with 2%/2 mm and 3%/3 mm criteria. In this case, the global dose criteria are relative to 2 Gy and voxels with dose >20% were considered.
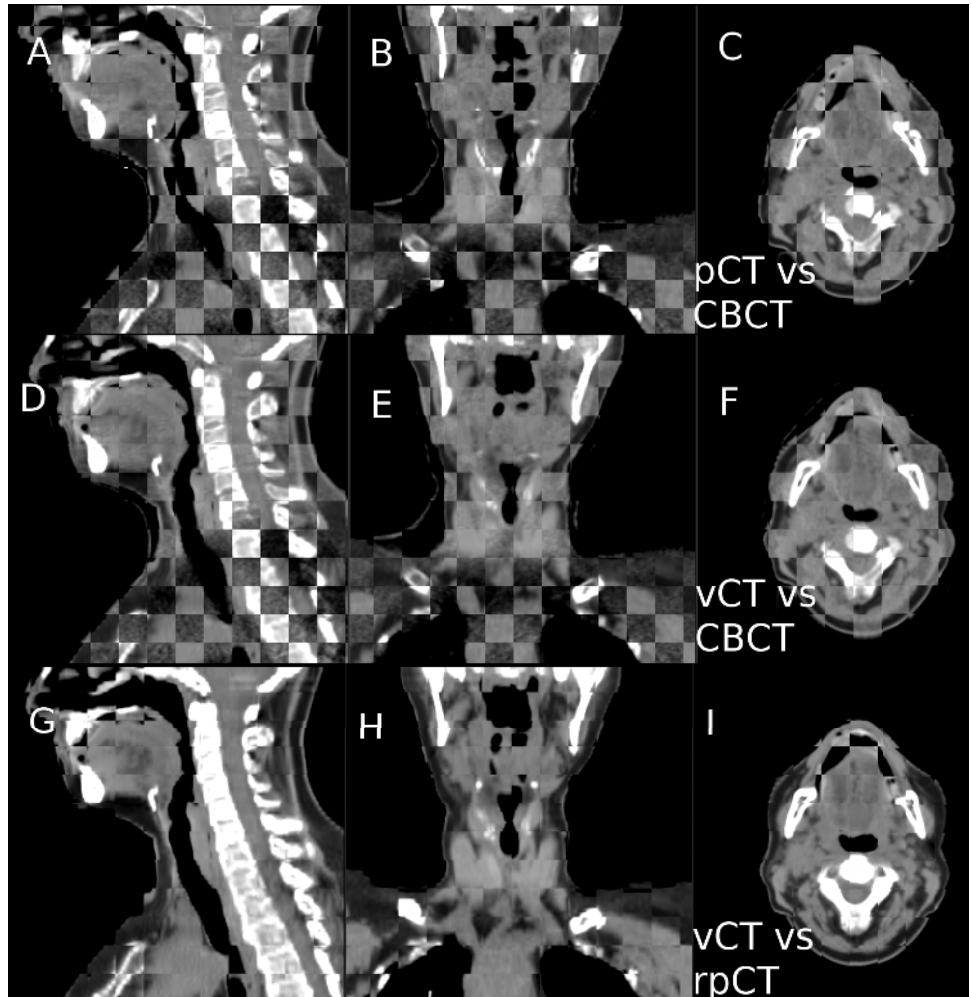

FIG. 2. Checkerboard comparison of (A), (B), and (C), the pCT_{CBCT} vs $CBCT_{CBCT}$, (D), (E), and (F) vCT_{CBCT} vs $CBCT_{CBCT}$, and (G), (H), and (I) vCT_{CBCT} vs $rpCT_{CBCT}$ for Pat2.

TABLE III. Percentage of pixels passing the gamma criterion 3 mm/3 mm WET and the WET difference (mean and standard deviation) of the pCT_{CBCT} and vCT_{CBCT} compared to the $rpCT_{CBCT}$.

| | % pixel passing | | Δ WET (mm) | |
	pCT	vCT	pCT	vCT
Pat1	78	96	2.0 ±3.9	−0.7 ± 2.9
Pat2	69	89	2.6 ± 4.7	0.6 ± 3.5
Pat3	73	86	1.5 ± 5.1	0.9 ± 4.0
Pat4	77	95	−2.0 ± 4.0	−0.4 ± 3.3
Pat5	90	93	1.2 ± 3.8	0.5 ± 3.4
Pat6	77	88	−0.9 ± 4.6	0.0 ± 3.8

TABLE IV. Percentage of pixels passing the gamma tests and the range difference (mean ± standard deviation) of the pCT_{pCT} and vCT_{pCT} compared to the $rpCT_{pCT}$ for the single beam range evaluation.

| | % pixel passing 2 mm/2 mm range | | % pixel passing 3 mm/3 mm range | | Δrange (mm) | |
	pCT	vCT	pCT	vCT	pCT	vCT
Pat1	88	96	97	99	−1.1 ± 2.9	0.5 ± 2.5
Pat2	78	86	90	95	−0.8 ± 7.4	0.3 ± 6.5
Pat3	82	90	94	97	−1.1 ± 4.0	0.1 ± 2.9
Pat4	82	88	93	96	−0.6 ± 4.3	0.2 ± 3.9
Pat5	87	90	95	96	−0.4 ± 4.4	−0.2 ± 4.0
Pat6	87	88	95	96	−0.1 ± 3.4	−0.7 ± 3.6

To evaluate the proton range, single field dose distributions were generated using the 30° beam only. The proton range was defined as the distance between the patient surface derived from the CT images and the distal 80% isodose, using the 1.8 Gy prescription of the SIB plan. Two dimensional proton range maps in beam-eye-view (BEV) from the pCT_{pCT}, $rpCT_{pCT}$, and vCT_{pCT} were compared using gamma evaluation with criteria of 2 mm/2 mm range and 3 mm/3 mm range.

3. RESULTS

Figure 1 presents the results of the deformation for Pat1 in the CBCT frame of reference using a checkerboard comparison between the $CBCT_{CBCT}$, vCT_{CBCT}, pCT_{CBCT}, and $rpCT_{CBCT}$. For that case, the replanning CT was acquired 51 days after the planning CT and the CBCT a day before the replanning CT. We observed that at several locations, the patient's skin was in different positions in the pCT_{CBCT} and the $CBCT_{CBCT}$ [Figs. 1(A)–1(C)]; this has been improved in the vCT_{CBCT} [Figs. 1(D)–1(F)]. For this patient, no noticeable weight loss was observed, as further discussed in the Sec. 3.A and Table II, and the differences are mostly attributable to posture differences. The intensity differences between the pCT_{CBCT} or vCT_{CBCT} and the $CBCT_{CBCT}$ are visible [Figs. 1(A)–1(F)], especially in the thorax area; however, no effort was made to adjust the window and level to match intensities. The qualitative agreement between the vCT_{CBCT} and the $rpCT_{CBCT}$ appears acceptable [Figs. 1(G)–1(I)] and no intensity degradation is observed in the thorax, as opposed to the $CBCT_{CBCT}$. The agreement between $rpCT_{CBCT}$ and

the vCT_{CBCT} appeared worst in the shoulder area where positioning differences between the CT scanner table and the treatment couch are possible.

Figure 2 shows similar results for Pat2 with replanning CT acquired 39 days after the planning CT and the CBCT 2 days later. In that case, we observed larger anatomical changes caused by patient weight loss as visible in Figs. 2(A)–2(C). The vCT_{CBCT} compares well to the $CBCT_{CBCT}$ in the checkerboard comparison [Figs. 2(D)–2(F)] except in the mouth where the tongue was depressed at the time of CBCT acquisition as opposed to the pCT. The DIR algorithm generally failed to correct for the filling or emptying of cavities such as the mouth or sinuses. This discrepancy is also observed in the vCT vs rpCT comparison of Figs. 2(G)–2(I). For Pat2, the alignment of the $rpCT_{CBCT}$ and vCT_{CBCT} in the thorax area appears worse than for Pat1 [Fig. 2(H)], potentially due to larger posture differences between the $rpCT_{CBCT}$ and $CBCT_{CBCT}$. However, in the neck area, where the high dose CTV is located, the agreement between vCT_{CBCT} and $rpCT_{CBCT}$ appears acceptable. The results shown in Figs. 1 and 2 provide a qualitative overview of deformation accuracy for two example cases; in Secs. 3.A and 3.B, the accuracy is systematically quantified for all six patients.

3.A. Deformation accuracy

Table II presents the median distances between corresponding features as identified by the SIFT algorithm between the pCT_{CBCT}, vCT_{CBCT}, $rpCT_{CBCT}$, and $CBCT_{CBCT}$ for each
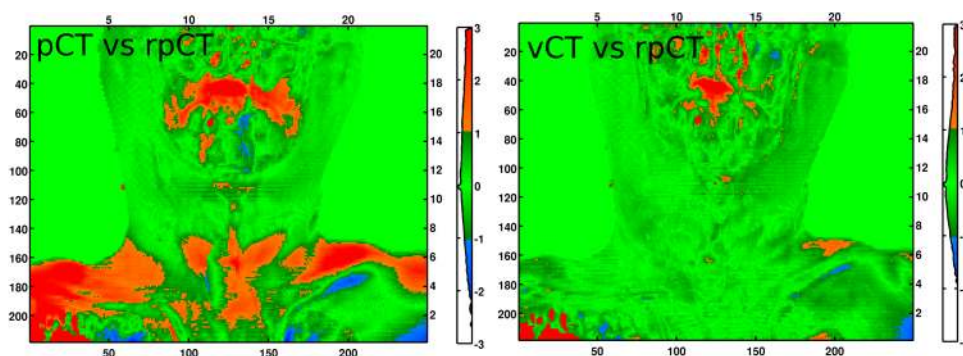

FIG. 3. Gamma index of the WET difference between (*left*) pCT_{CBCT} and $rpCT_{CBCT}$ and (*right*) vCT_{CBCT} and $rpCT_{CBCT}$ using 3 mm/3 mm range for Pat4.
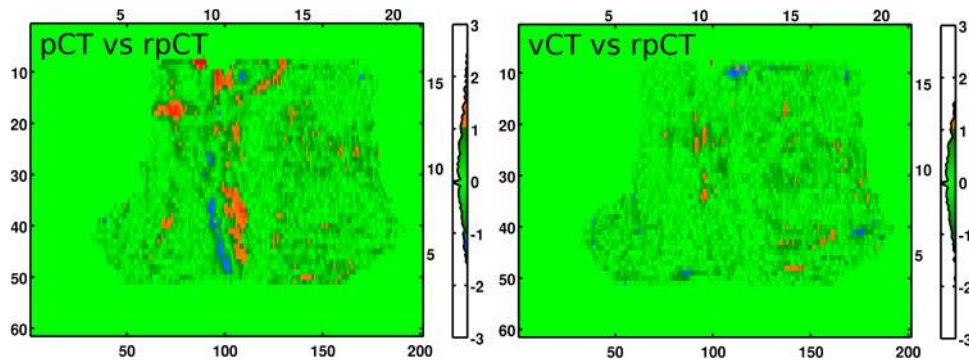
Fig. 4. Gamma index of the range difference for a single 30° beam between (*left*) pCT_{pCT} and $rpCT_{pCT}$ and (*right*) vCT_{pCT} and $rpCT_{pCT}$ using 3 mm/3 mm range for Pat3. The colorbar also shows a histogram of the gamma index distribution including regions outside the field.

patient. The vCT_{CBCT} vs $CBCT_{CBCT}$ present the lowest feature distances. The DIR reduced median feature distances between the pCT_{CBCT} and the $CBCT_{CBCT}$ in the range of [1.4, 5.4 mm] down to [1.0, 1.4 mm] for vCT_{CBCT} vs $CBCT_{CBCT}$, which is of the order of the CBCT pixel size. The distances between the pCT_{CBCT} and the $CBCT_{CBCT}$ were similar to those between the pCT_{CBCT} and the vCT_{CBCT} within 0.5 mm.

Not all $rpCT_{CBCT}$ agreed with either the vCT_{CBCT} or $CBCT_{CBCT}$ in terms of feature distances; the best agreement is for Pat3, Pat5, and Pat6 with 1.4 mm median feature distance. The worst agreement is for Pat4 (2.8 mm) and Pat2 (2.4 mm). This could indicate that the $rpCT_{CBCT}$ is not a reliable reference for all cases due to different

postures at each imaging couch for patients with weight loss where the immobilization device is rendered less efficient. According to the SIFT features, Pat2 and Pat3 present the largest difference between pCT and rpCT. We have estimated these differences by calculating the volume change between the body contours between planning and replanning CTs (pCT_{CBCT} vs $rpCT_{CBCT}$). We also estimated volume changes between the vCT_{CBCT} and $rpCT_{CBCT}$. Additionally, we looked at COM shifts between the contour pairs. For this estimation, we restricted the analysis to the region above the shoulders and below the ears to isolate weight loss effects. These results are at the lower part of Table II. We observed that the largest SIFT distances correspond to cases with both
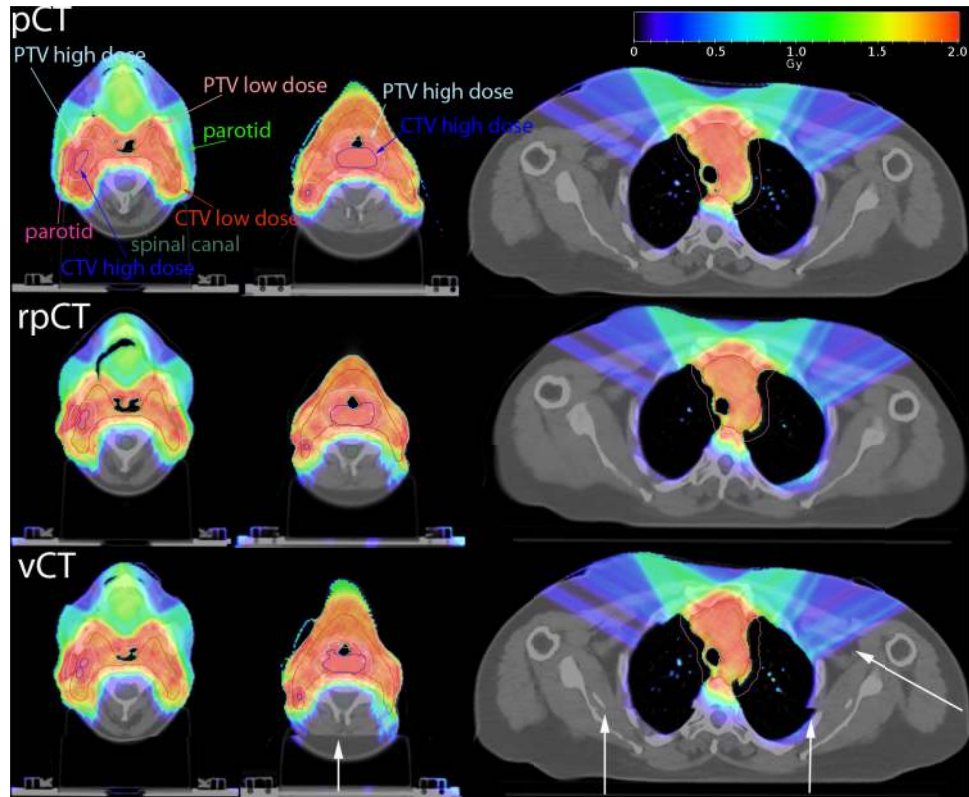

Fig. 5. Dose distributions calculated on the pCT_{pCT}, $rpCT_{pCT}$, and vCT_{pCT} based on proton fluence optimized on the pCT_{pCT} for Pat2. Three different slices are shown (head, neck, upper thorax). White arrows indicate where the vCT was stitched to the pCT due to the limited CBCT FOV. Relevant structures are indicated on the pCT. The CBCT used for the vCT and the rpCT was acquired 2 days apart. Dose in air and lung was masked here for visualization purposes.
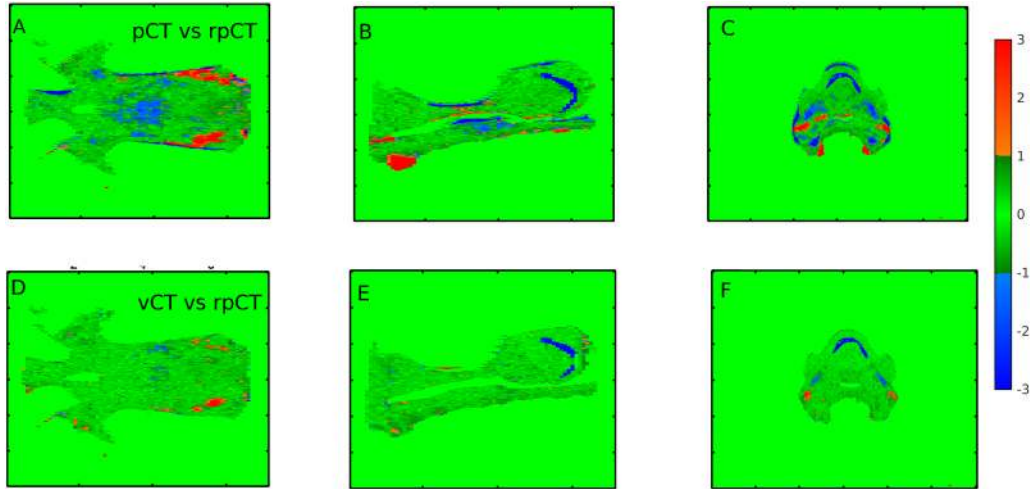
FIG. 6. Gamma index using 3%/3 mm for Pat2. (A)–(C) Comparison of the pCT$_{pCT}$ and the rpCT$_{pCT}$. (D)–(E) Comparison of the vCT$_{pCT}$ and the rpCT$_{pCT}$. The same slices are shown for A–C and D–F. Only doses above 20% of the prescription dose are considered in the analysis.

volume changes and COM shifts after rigid registration of the rpCT$_{CBCT}$ and pCT$_{CBCT}$ to the CBCT$_{CBCT}$. Large volume changes make the registration task more difficult due to the lowered correspondence between the images and a higher likelihood of posture differences. There was generally good agreement between the vCT$_{CBCT}$ and the rpCT$_{CBCT}$ in terms of body contour volume and COM.

The evaluation of the mean CT numbers of the vCT$_{CBCT}$ within the parotid glands, spinal cord, and mandible contours obtained from the DIR presented differences of less than 5 HU when compared to those from the pCT$_{CBCT}$. These differences are within the standard deviation of the distribution of CT numbers of about 15 HU. The results of the WET evaluation are presented in Table III. For all patients, we observed a lower mean WET difference and a higher percentage of pixels passing the gamma evaluation for the vCT$_{CBCT}$ vs rpCT$_{CBCT}$ evaluation than for the pCT$_{CBCT}$ vs rpCT$_{CBCT}$ comparison.

Figure 3 presents 2D gamma index distributions for Pat4 using 3 mm/3 mm WET criteria where the amount of pixels passing is increased when comparing the vCT$_{CBCT}$ to the rpCT$_{CBCT}$ instead of the pCT$_{CBCT}$. Most of the differences between the vCT$_{CBCT}$ and the rpCT$_{CBCT}$ appear in the nasal and oral cavities. As the CBCT is not acquired at the same time

as the rpCT (1 day difference for Pat4), differences in nasal cavity filling or tongue position may cause discrepancies. A nasal cavity which was filled with mucus at the time of planning and empty on the CBCT images would appear filled in the vCT as the Morphons algorithm failed to correct for this. However, when planning IMPT, 0° beams passing through the nasal and oral cavities are avoided thus minimizing the impact of this discrepancy.

3.B. Dosimetric evaluation

In Table IV, the single 30° field range comparison of the vCT$_{pCT}$ vs rpCT$_{pCT}$ and pCT$_{pCT}$ vs rpCT$_{pCT}$ is presented. In general, similar trends as for Table III are observed; an increase of the percentage of pixels passing the gamma evaluation and a reduction of the mean range difference when going from pCT$_{pCT}$ to vCT$_{pCT}$. However, for this test, the improvements are somewhat less than those observed in Table III. Figure 4 shows the 2D gamma index distributions for Pat3 using 3 mm/3 mm range criteria. We observe that the mismatch of airways between the pCT$_{pCT}$ and rpCT$_{pCT}$ causes some overshoot (gamma index > 1), while different shoulder positioning causes undershoot (gamma index < −1). The amount of pixels passing, which was determined only in the field area, is improved with the vCT$_{pCT}$.

Figure 5 presents dose distributions calculated on the pCT$_{pCT}$, rpCT$_{pCT}$, and corresponding vCT$_{pCT}$ for Pat2. We observed that the dose distribution was altered due to markedly different anatomy between the pCT and rpCT. The vCT captures the general anatomical changes well; however, the mouth opening observed in the rpCT is not reproduced in the vCT. The mouth opening was also observed in the CBCT.

Figure 6 presents the gamma index for the comparison for Pat2 of the four field dose distributions obtained using the vCT$_{pCT}$ and the pCT$_{pCT}$ compared to the rpCT$_{pCT}$. The gamma criteria were 3%/3 mm in this case. The passing rate for the pCT$_{pCT}$ was 87%, while for the vCT$_{pCT}$, it was 96%. The largest discrepancies for the vCT$_{pCT}$ appeared along the beam

TABLE V. Percentage of voxels passing the gamma tests when comparing the 4-beams dose distributions obtained from the pCT$_{pCT}$ and vCT$_{pCT}$ compared to those obtained from the rpCT$_{pCT}$. The dose criterion is relative to 2 Gy. Only voxels with dose above 20% of the prescription dose were considered.

	% voxels passing 2%/2 mm		% voxels passing 3%/3 mm	
	pCT	vCT	pCT	vCT
Pat1	74	89	90	97
Pat2	75	86	87	96
Pat3	78	89	91	98
Pat4	84	83	94	95
Pat5	83	84	94	95
Pat6	87	88	96	96

TABLE VI. Comparison of contours from DIR with those drawn by a physician on the vCT$_{pCT}$ and rpCT$_{pCT}$ using DSC, mean Hausdorff distance d between contours, and 95% range in parenthesis as well as COM shifts for selected contours.

Pat2		Parotid left	Parotid right	Brain stem	Spinal canal	Mandible	Low dose CTV	High dose CTV	Low dose PTV	High dose PTV
Physician vCT	DSC	0.91	0.89	0.93	0.90	0.90	0.93	0.88	0.95	0.91
	d (mm)	0.5(2.3)	0.75(3.0)	1.0(4.6)	0.4(1.4)	0.5(1.8)	0.8(3.0)	0.6(2.4)	1.0(3.2)	0.9(3.0)
	COM shift (mm)	1.8	2.2					1.5		
Physician rpCT	DSC	0.88	0.91	0.90	0.85	0.85	0.93	0.82	0.94	0.88
	d (mm)	0.8(2.9)	0.7(2.4)	1.1(4.6)	0.7(2.1)	0.7(2.7)	0.9(3.1)	1.2(3.0)	1.2(3.7)	1.61(3.1)
	COM shift (mm)	2.9	0.9					2.3		

TABLE VII. Comparison of contours from DIR with those drawn by a physician on the vCT$_{pCT}$ and rpCT$_{pCT}$ using DSC, mean Hausdorff distance d between contours, and 95% range in parenthesis as well as COM shifts for selected contours. All quantities are averaged over the six patients. Both mean and standard deviation σ, in parenthesis, are reported.

Average over six patients		Parotid left	Parotid right	Brain stem	Spinal canal	Mandible	GTV	Low dose CTV	High dose CTV	Low dose PTV	High dose PTV
Physician vCT	Mean DSC	0.88(0.03)	0.88(0.01)	0.90(0.04)	0.86(0.04)	0.90(0.05)	0.76(0.16)	0.92(0.02)	0.92(0.03)	0.94(0.01)	0.93(0.01)
	Mean d (mm)	0.7(0.2)	0.7(0.1)	1.5(0.3)	0.6(0.1)	0.5(0.1)	0.9(0.4)	1.0(0.3)	0.8(0.3)	1.2(0.4)	1.1(0.3)
	Mean COM shift (mm)	1.8(1.7)	1.7(0.6)				1.8(0.6)		1.4(0.8)		
Physician rpCT	Mean DSC	0.86(0.03)	0.88(0.03)	0.91(0.02)	0.81(0.03)	0.89(0.04)	0.76(0.09)	0.91(0.03)	0.90(0.04)	0.94(0.01)	0.92(0.03)
	Mean d (mm)	0.9(0.3)	0.8(0.1)	1.4(0.2)	0.9(0.3)	0.6(0.2)	1.1(0.2)	1.2(0.3)	1.2(0.4)	1.4(0.4)	1.4(0.3)
	Mean COM shift (mm)	2.4(1.1)	1.2(0.7)				2.4(0.3)		2.0(1.3)		

path beyond the tongue which was depressed in the rpCT$_{pCT}$ but not in the pCT$_{pCT}$. The deformation tool generally failed to fill or empty cavities such as the mouth or sinuses. This issue and potential correction methods are further discussed in Sec. 4 of this paper.

Table V summarizes the results of the gamma comparison of 4-field dose distributions for the six patients using both 2%/2 mm and 3%/3 mm criteria. The amount of voxels passing the gamma evaluation for 2%/2 mm (3%/3 mm) criteria ranged from 74%(87%) to 87%(94%) when comparing the pCT$_{pCT}$ and the rpCT$_{pCT}$, while it ranged from 83%(87%) to 89%(98%) when comparing the vCT$_{pCT}$ to the rpCT$_{pCT}$. For Pat4, Pat5, and Pat6, no difference was observed between the two comparisons suggesting that for those cases, the dose distributions calculated on the pCT$_{pCT}$ and vCT$_{pCT}$ are similar to those calculated on the rpCT$_{pCT}$. This is consistent with the Δ range results shown in Table IV. For the other patients, the vCT$_{pCT}$ vs rpCT$_{pCT}$ showed better results.

Table VI presents the DSC, mean d, and COM shifts for Pat2 where the contours from the CT to CBCT DIR are compared to those from the physician drawn on the vCT$_{pCT}$ and on the rpCT$_{pCT}$. In this case, the rpCT$_{pCT}$ was acquired 2 days before the CBCT scan. For this particular patient, no GTV was delineated and the high dose CTV was considered equivalent to the GTV in terms of COM shifts. When comparing against the physician-drawn contours on the rpCT$_{pCT}$(vCT$_{pCT}$), we observe DSC > 0.80(0.88) for all structures. Table VII presents the same data as Table V averaged over the six patients considered in this study. The results are similar to those observed in Table VI. For three patients, a GTV was delineated and we observed lower average DSC and larger variation. A recent study by Mencarelli *et al.* using implanted fiducial markers questioned the ability of DIR to track tumor changes in the H&N region.[33]

Table VIII presents DVH statistics for Pat2 in terms of the total planned dose estimated on the pCT$_{pCT}$, rpCT$_{pCT}$, and vCT$_{pCT}$ using the three sets of contours mentioned above. For Pat2, the mean dose to the parotids increased by 7 and 9 Gy between the planning and replanning CT and the spinal canal saw a large increase of 19 Gy, which is however below the 54 Gy constraint established at the planning stage. The numbers presented here do not correspond to the cumulative

dose throughout the treatment; it is however easier to interpret them in terms of 30 fractions instead of dose per fraction.

The differences between DVH statistics for OARs from dose distributions calculated on the vCT are dependent on the contour set used for the evaluation. When using the contours from the rpCT, the differences between DVH statistics computed on the vCT$_{pCT}$ dose distribution and those computed on the rpCT$_{pCT}$ dose distribution are below 1 Gy for all structures except the spinal canal. The spinal canal shows a larger difference, potentially due to the higher sensitivity of maximum dose metrics such as $D_{2\%}$. This suggests that the spatial distribution of the dose calculated on the vCT$_{pCT}$ is in good agreement with that calculated on the rpCT$_{pCT}$ when evaluated with DVH statistics.

For the parotid glands, variations of D_{mean} of -3.9 to 4.2 Gy are observed depending on the contour set employed for evaluation of the dose distribution from the vCT$_{pCT}$. Both CTVs show differences of $D_{95\%}$ of less than 1 Gy. As expected, the PTV is more sensitive with differences ranging from -3.3 to 2.4 Gy due to the sharp dose gradients.

Table IX presents similar data as Table VIII averaged over the six patients in this study. In general, OAR DVH statistics increased from the pCT$_{pCT}$ to the rpCT$_{pCT}$, while for the targets, they were decreased. The IMRT margins employed in this study yielded maximum CTV reduction of $D_{95\%}$ of 4 Gy. When using the same contours to compare the dose distributions computed on the rpCT$_{pCT}$ and vCT$_{pCT}$, the differences were not larger than -2 Gy for the CTVs and -4 Gy for OARs. When the vCT was delineated by the physician, this range was similar; however, the parotid glands showed differences of -4 to 4 Gy as opposed to -1 to 1 Gy when using the rpCT$_{pCT}$ contours. The DIR contours showed the largest differences of DVH statistics when compared to the rpCT$_{pCT}$ with differences ranging from -8 to 11 Gy for the brain stem and spinal canal and -5 to 2 Gy for parotids. The CTV differences were similar to those from the rpCT$_{pCT}$ and physician contours.

Figure 7 shows the time evolution of DVH statistics for Pat4 over the course of treatment. A vCT$_{pCT}$ was generated for each available CBCT and dose distributions were recalculated on each vCT$_{pCT}$ as well as on two rpCT$_{pCT}$. Pat4 had negligible doses to the spinal canal and brain stem; thus, we present

TABLE VIII. DVH parameters for the initial plan calculated with MC on the pCT$_{pCT}$ and recalculated on the rpCT$_{pCT}$ and vCT$_{pCT}$. For the vCT$_{pCT}$, three sets of contours are employed: those from the rpCT$_{pCT}$ following rigid registration to the vCT$_{pCT}$, those drawn by the physician on the vCT$_{pCT}$, and those obtained from DIR of the pCT contours. The 2 Gy fraction doses were multiplied by 30 for interpretation. High dose PTVs should receive 60 Gy and low dose PTV should receive 54 Gy.

Pat2		pCT	rpCT	Δ rpCT, pCT	Δ vCT (rpCT contours), rpCT	Δ vCT (physician contours), rpCT	Δ vCT (DIR contours), rpCT
Parotid left	D_{mean} (Gy)	15.3	24.3	9	0	−3.9	−3.6
Parotid right	D_{mean} (Gy)	27.6	34.5	6.9	−0.3	4.2	0
Brain stem	$D_{2\%}$ (Gy)	0.3	0.3	0	0	0.3	0
Spinal canal	$D_{2\%}$ (Gy)	4.2	23.7	19.5	−4.2	−4.8	−7.5
CTV low dose	$D_{95\%}$ (Gy)	52.8	51.6	−1.2	0.3	0	0
CTV high dose	$D_{95\%}$ (Gy)	60.6	57.3	−3.3	0.6	0.6	0.3
PTV low dose	$D_{95\%}$ (Gy)	46.8	43.5	−3.3	−0.9	−3.3	−2.1
PTV high dose	$D_{95\%}$ (Gy)	58.5	55.2	−3.3	0.6	1.2	2.4

TABLE IX. Average difference of DVH parameters for the initial plan calculated with MC on the pCT_{pCT} and recalculated on the $rpCT_{pCT}$ and vCT_{pCT}. For the vCT_{pCT}, three sets of contours are employed: those from the $rpCT_{pCT}$ following rigid registration to the vCT_{pCT}, those drawn by the physician on the vCT_{pCT}, and those obtained from DIR. The 2 Gy fraction doses were multiplied by 30 for interpretation. vCT comparisons are against the rpCT. The numbers in brackets indicate the minimum and maximum over the six patients.

		Δ rpCT, pCT	Δ vCT (rpCT contours), rpCT	Δ vCT (physician contours), rpCT	Δ vCT (DIR contours), rpCT
Parotid left	D_{mean} (Gy)	2 [−2,9]	0 [−1,0]	0 [−4,1]	−1 [−5,2]
Parotid right	D_{mean} (Gy)	1 [−2,7]	0 [0,1]	0 [−2,4]	−1 [−2,0]
Brain stem	$D_{2\%}$ (Gy)	0 [3,0]	0 [0,2]	0 [0,1]	2 [0,11]
Spinal canal	$D_{2\%}$ (Gy)	3 [−1,20]	−1 [−4,0]	0 [−5,2]	0 [−8,1]
CTV low dose	$D_{95\%}$ (Gy)	−1 [−4,0]	−1 [−2,0]	0 [−2,2]	0 [−2,0]
CTV high dose	$D_{95\%}$ (Gy)	0 [−3,1]	0 [−2,1]	0 [−2,1]	0 [−2,0]
PTV low dose	$D_{95\%}$ (Gy)	−5 [−11,0]	−1 [−5,1]	0 [−4,2]	1 [−5,5]
PTV high dose	$D_{95\%}$ (Gy)	−3 [−7,1]	0 [−1,1]	1 [−2,5]	−1 [−2,4]

results for the high and low dose CTVs and one of the parotid glands. We detected an increase in parotid gland D_{mean} over the course of treatment using the vCT; this was confirmed by the D_{mean} obtained using the corresponding rpCT.

4. DISCUSSION

The accuracy of the Morphons algorithm applied to CT to CBCT DIR in this study was of the order of 1.0–1.4 mm when evaluated by comparison of matching features as identified by the SIFT algorithm. Paganelli et al.[34] reported a group median (median of distribution of medians) feature distance for 20 patients of 0.66 mm between vCT_{CBCT} and $CBCT_{CBCT}$ ranging from 0.4 to 1.1 mm. For the $rpCT_{CBCT}$ vs $CBCT_{CBCT}$ comparison, they reported a group median feature distance 0.66 mm ranging from 0.4 to 1.7 mm. Our results for the $rpCT_{CBCT}$ vs $CBCT_{CBCT}$ evaluation were slightly worse than theirs, with median distances ranging from 1.4 to 2.8 mm. This is potentially due to the quality of the rigid registration between $rpCT_{CBCT}$ and $CBCT_{CBCT}$ which was performed in this study for evaluation purposes but which had no impact on the vCT_{CBCT}. We employed a translation registration to replicate clinical reality, while they employed a registration allowing rotations. It is difficult to establish whether the larger distances between features are due to poor rigid registration accuracy or posture differences b etween t he r pCT_{CBCT} and $CBCT_{CBCT}$ caused by less effective p atient immobilization due to weight loss. Our investigation covered volumes which included the shoulder area where posture changes are difficult to correct with rigid registration.

Our SIFT results were comparable to the results of the Hausdorff d istance w hen c omparing c ontours d rawn o n the reference $rpCT_{pCT}$ and those obtained from DIR, with average Hausdorff distance ranging from 0.6 to 1.4 mm depending on the structure investigated. Our DSC was higher than those of Peroni et al.[15] who reported median DSC of about 0.8 for the parotid glands and the mandible and 0.55 for the GTV. Similar results were obtained for COM shifts for the parotid glands and GTV. Our GTV results however are based on only three patients. As plan evaluation is generally performed on OARs

and CTVs, the poorer results of GTV deformation should not be seen as an obstacle to the vCT method.

The $vCT_{pCT/CBCT}$ performed better than a rigidly aligned $pCT_{pCT/CBCT}$ when considering proton-specific metrics such as WET differences, single beam range differences, and 4-field dose distribution comparison to the $rpCT_{pCT/CBCT}$. When comparing dose distributions using the contours drawn on the rpCT to evaluate DVH statistics, the results from $rpCT_{pCT}$ and vCT_{pCT} were similar within −2 to 2 Gy for the parotids, brain stem, and CTVs. Deviations of up to −4 Gy were observed for the spinal canal when using $D_{2\%}$, most likely due to the sensitivity of near-maximum metrics. This supports our hypothesis that the vCT_{pCT} is a reliable method to obtain the dose distribution of the day using CBCT imaging. However, in clinical practice, contoured $rpCT_{pCT}$ scans are not available at each fraction and either physician drawn-contours on the vCT_{pCT} or DIR-derived contours must be used to extract DVH statistics. In that scenario, the reliability of
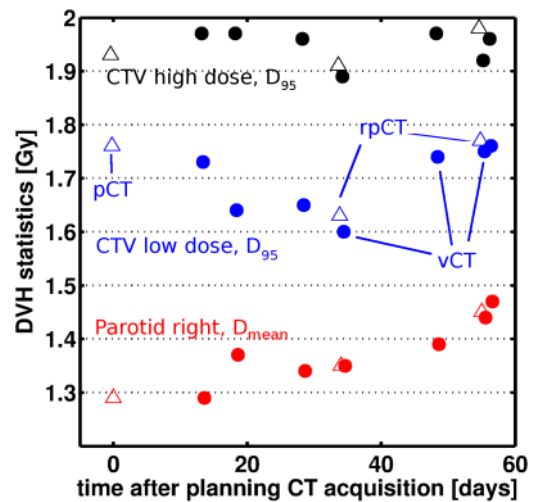

FIG. 7. Time evolution of DVH indices for the high and low dose CTVs (D_{95}) and right parotid gland (D_{mean}) for Pat4. Triangles correspond to DVH indices calculated from diagnostic quality CT images (pCT_{pCT} and $rpCT_{pCT}$), while circles are from vCT_{pCT}. DVH indices are expressed in dose per fraction.
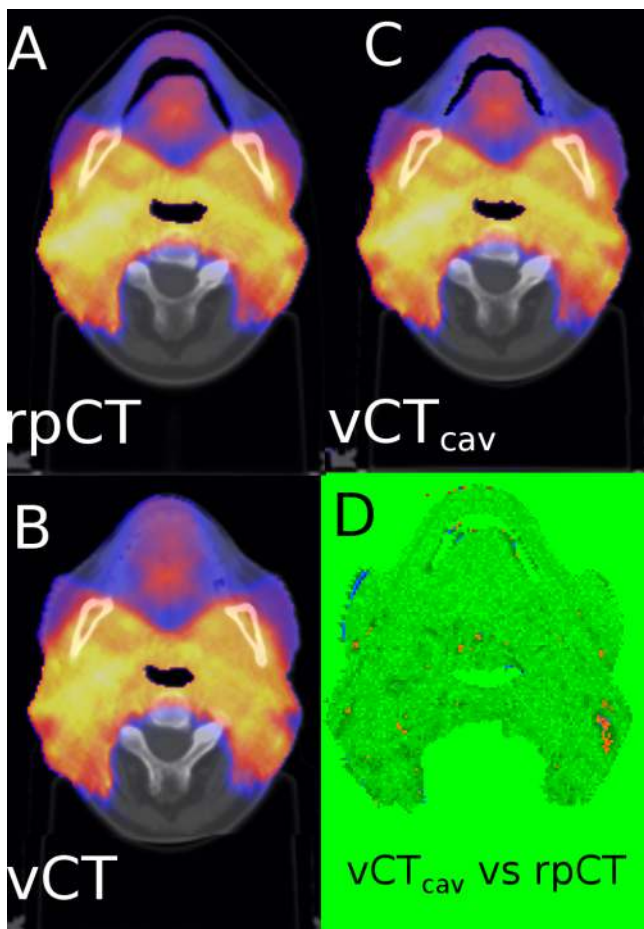
FIG. 8. Dose distribution calculated and overlaid on the (A) $rpCT_{pCT}$, (B) uncorrected vCT_{pCT}, and (C) cavity-corrected $vCT_{cav,pCT}$. (D) gamma comparison using 3%/3 mm between the dose distributions corresponding to (C) and (A). (D) should be directly compared to Fig. 6(F).

the method was slightly worse for OARs, especially when using DIR generated contours. However, for all scenarios, the CTV $D_{95\%}$ was recovered within ±2 Gy. These uncertainties suggest that the vCT_{pCT} method would be useful as a treatment monitoring tool at facilities equipped with a CBCT scanner but should not replace the acquisition of a new planning CT and the construction of a new patient immobilization mask once it has been decided to adapt the IMPT plan. It may also be interesting to investigate action levels based on metrics other than DVH statistics such as the distance between SIFT features between the pCT and CBCT (Ref. 34) or by performing gamma evaluation comparison of the dose distributions from the pCT and vCT. Whether the vector fields obtained when generating the vCT_{pCT} may be used as a basis for dose accumulation was outside the scope of this work and deserves special attention and validation.

We made use of replanning CT scans acquired within 3 days of the CBCT scans as reference in this study. The relatively short time scale minimizes chances of anatomy variation between the two scans. However, the repositioning of the patient for each scan and the patient weight loss since the design of the immobilization mask may reduce the validity of the replanning CT scan as a reference due to potential posture differences.

The validity of our vCT_{pCT} is currently limited to the CBCT FOV and we relied on the pCT_{pCT} for regions outside of it. Proposed methods to tackle this problem include performing DIR using CBCT projections,[17] acquiring the CBCT using two rotations coupled to a table shift[35] or deforming the pCT by continuity of the vector field at the edge of the CBCT FOV.[18] This had no impact on our evaluation since our IMPT beams were not coming from beneath the table or through the shoulders. However, this could be an issue for modalities such as arc therapy or for CTVs extending beyond the CBCT FOV.

The issue of cavities being filled in the pCT and empty in the CBCT and *vice versa* deserves special attention. Figure 5 shows that the Morphons algorithm failed to correct these situations; a possible solution would be to identify these regions based on the difference of the vCT_{CBCT} and $CBCT_{CBCT}$ using thresholds and naively replacing the CT numbers of the vCT_{pCT} by either those of air or water. Preliminary investigation supports that this is feasible, as shown in Fig. 8. In that case, the vCT_{pCT} was corrected using the method described above to yield a $vCT_{cav,pCT}$. We applied the correction to the head region only as the intensity differences of the thorax induced erroneous corrections. We observed that the underestimated hotspot in Fig. 8(B) appears to be correctly reproduced when correcting for cavities [Fig. 8(C)]. The gamma evaluation of vCT_{pCT} vs $rpCT_{pCT}$, restricted to the head region, yielded 85%(94%) passing pixels, while $vCT_{cav,pCT}$ vs $rpCT_{pCT}$ yielded 91%(97%) when using 2%/2 mm (3%/3 mm). The method was tested on several other cases and found to be reliable and deserving of further investigation.

Our TPS did not allow patching fields in the superior–inferior direction; for this reason, we employed the four field plans for the entire PTV. To avoid uncertainties related to cavities, it would be advantageous to use parallel opposed beams from above the shoulders to the top of the PTV and to use one or two anterior fields to cover sections of the PTV located below the shoulders/base of neck.

5. CONCLUSION

In this work, we aimed at generating stopping power distributions based on CBCT imaging in the context of adaptive proton therapy. The dose distributions calculated on the images resulting from deformable image registration of the planning CT to the CBCT (virtual CT) were comparable to those calculated on corresponding replanning CT when considering DVH statistics using the same set of contours. The use of DIR generated contours introduced larger variability and would most likely require physician correction. When using the replanning CT as reference, using the virtual CT instead of a rigidly registered planning CT, showed improved results for all metrics considered in this work. The virtual CT may be a useful treatment monitoring and plan adaptation decision making tool for centers equipped with gantry mounted CBCT scanners. However, the generation of new treatment plan should be done on a replanning CT acquired at a diagnostic quality CT scanner.

In future work, we will explore the performance of the proposed approach in further anatomical locations and consider the integration of the tools of relevance to our workflow in a single platform.

ACKNOWLEDGMENTS

[a] Author to whom correspondence should be addressed. Electronic mail: g.landry@lmu.de

[1] M. K. Kam *et al.*, "Prospective randomized study of intensity-modulated radiotherapy on salivary gland function in early-stage nasopharyngeal carcinoma patients," J. Clin. Oncol. 25, 4873–4879 (2007).

[2] C. B. Simone II *et al.*, "Comparison of intensity-modulated radiotherapy, adaptive radiotherapy, proton radiotherapy, and adaptive proton radiotherapy for treatment of locally advanced head and neck cancer," Radiother. Oncol.: J. Eur. Soc. Ther. Radiol. Oncol. 101, 376–382 (2011).

[3] T. A. van de Water, A. J. Lomax, H. P. Bijl, M. E. de Jong, C. Schilstra, E. B. Hug, and J. A. Langendijk, "Potential benefits of scanned intensity-modulated proton therapy versus advanced photon therapy with regard to sparing of the salivary glands in oropharyngeal cancer," Int. J. Radiat. Oncol., Biol., Phys. 79, 1216–1224 (2011).

[4] J. L. Barker *et al.*, "Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system," Int. J. Radiat. Oncol., Biol., Phys. 59, 960–970 (2004).

[5] W. Wang *et al.*, "Clinical study of the necessity of replanning before the 25th fraction during the course of intensity-modulated radiotherapy for patients with nasopharyngeal carcinoma," Int. J. Radiat. Oncol., Biol., Phys. 77, 617–621 (2010).

[6] A. C. Kraan *et al.*, "Dose uncertainties in IMPT for oropharyngeal cancer in the presence of anatomical, range, and setup errors," Int. J. Radiat. Oncol., Biol., Phys. 87, 888–896 (2013).

[7] A. Trofimov *et al.*, "Interfractional variations in the setup of pelvic bony anatomy and soft tissue, and their implications on the delivery of proton therapy for localized prostate cancer," Int. J. Radiat. Oncol., Biol., Phys. 80, 928–937 (2011).

[8] S. Park, M. Cho, and H. Kim, "On-board CBCT/CBDT for image-guided proton therapy: Initial performance evaluation," Int. J. Radiat. Oncol., Biol., Phys. 75, S595–S596 (2009).

[9] M. Stock, M. Pasler, W. Birkfellner, P. Homolka, R. Poetter, and D. Georg, "Image quality and stability of image-guided radiotherapy (IGRT) devices: A comparative study," Radiother. Oncol.: J. Eur. Soc. Ther. Radiol. Oncol. 93, 1–7 (2009).

[10] J. H. Siewerdsen and D. A. Jaffray, "Cone-beam computed tomography with a flat-panel imager: Magnitude and effects of x-ray scatter," Med. Phys. 28, 220–231 (2001).

[11] Y. Yang, E. Schreibmann, T. Li, C. Wang, and L. Xing, "Evaluation of on-board kV cone beam CT (CBCT)-based dose calculation," Phys. Med. Biol. 52, 685–705 (2007).

[12] E. Bentefour, S. Tang, S. Both, G. Chen, and H. Lu, "TU - A - 204B - 02: On the potential of CBCT for range verification in proton therapy," Med. Phys. 37, 3370 (2010).

[13] T. Zhang, Y. Chi, E. Meldolesi, and D. Yan, "Automatic delineation of on-line head-and-neck computed tomography images: Toward on-line adaptive radiotherapy," Int. J. Radiat. Oncol., Biol., Phys. 68, 522–530 (2007).

[14] Q. Wu, Y. Chi, P. Y. Chen, D. J. Krauss, D. Yan, and A. Martinez, "Adaptive replanning strategies accounting for shrinkage in head and neck IMRT," Int. J. Radiat. Oncol., Biol., Phys. 75, 924–932 (2009).

[15] M. Peroni *et al.*, "Automatic segmentation and online virtualCT in head-and-neck adaptive radiation therapy," Int. J. Radiat. Oncol., Biol., Phys. 84, e427–e433 (2012).

[16] X. Zhen, X. Gu, H. Yan, L. Zhou, X. Jia, and S. B. Jiang, "CT to cone-beam CT deformable registration with simultaneous intensity correction," Phys. Med. Biol. 57, 6807–6826 (2012).

[17] X. Zhen, H. Yan, L. Zhou, X. Jia, and S. B. Jiang, "Deformable image registration of CT and truncated cone-beam CT for adaptive radiation therapy," Phys. Med. Biol. 58, 7979–7993 (2013).

[18] C. Veiga *et al.*, "Toward adaptive radiotherapy for head and neck patients: Feasibility study on using CT-to-CBCT deformable registration for 'dose of the day' calculations," Med. Phys. 41, 031703 (12pp.) (2014).

[19] G. Landry *et al.*, "Phantom based evaluation of CT to CBCT image registration for proton therapy dose recalculation," Phys. Med. Biol. 60, 595–613 (2015).

[20] H. Knutsson and M. Andersson, "Morphons: Paint on priors and elastic canvas for segmentation and registration," in *Image Analysis*, edited by H. Kalviainen, J. Parkkinen, and A. Kaarna (Springer, Berlin Heidelberg, Germany, 2005), pp. 292–301.

[21] A. Wrangsjo, J. Pettersson, and H. Knutsson, "Non-rigid registration using morphons," in *Image Analysis*, edited by H. Kalviainen, J. Parkkinen, and A. Kaarna (Springer, Berlin Heidelberg, Germany, 2005), pp. 501–510.

[22] G. Janssens, L. Jacques, J. O. d. Xivry, X. Geets, and B. Macq, "Diffeomorphic registration of images with variable contrast enhancement," J. Biomed. Imaging 2011, 1–12.

[23] J. P. Thirion, "Image matching as a diffusion process: An analogy with Maxwell's demons," Med. Image Anal. 2, 243–260 (1998).

[24] G. Janssens, J. O. de Xivry, S. Fekkes, A. Dekker, B. Macq, P. Lambin, and W. van Elmpt, "Evaluation of nonrigid registration models for inter-fraction dose accumulation in radiotherapy," Med. Phys. 36, 4268–4276 (2009).

[25] D. G. Lowe, "Object recognition from local scale-invariant features," in *1999. The Proceedings of the Seventh IEEE International Conference on Computer Vision* (IEEE, Corfu, Greece, 1999).

[26] D. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. Comput. Vision 60, 91–110 (2004).

[27] W. Cheung and G. Hamarneh, "n-SIFT: n-Dimensional scale invariant feature transform," Image Process., IEEE Trans. 18, 2012–2021 (2009).

[28] C. Paganelli, M. Peroni, G. Baroni, and M. Riboldi, "Quantification of organ motion based on an adaptive image-based scale invariant feature method," Med. Phys. 40, 111701 (12pp.) (2013).

[29] J. A. Shackleford, N. Kandasamy, and G. C. Sharp, "On developing B-spline registration algorithms for multi-core processors," Phys. Med. Biol. 55, 6329–6351 (2010).

[30] L. C. G. G. Persoon, M. Podesta, W. J. C. van Elmpt, S. M. J. J. G. Nijsten, and F. Verhaegen, "A fast three-dimensional gamma evaluation using a GPU utilizing texture memory for on-the-fly interpolations," Med. Phys. 38, 4032–4035 (2011).

[31] J. O. Deasy, A. I. Blanco, and V. H. Clark, "CERR: A computational environment for radiotherapy research," Med. Phys. 30, 979–985 (2003).

[32] S. Schell and J. J. Wilkens, "Advanced treatment planning methods for efficient radiation therapy with laser accelerated proton and ion beams," Med. Phys. 37, 5330–5340 (2010).

[33] A. Mencarelli, S. R. van Kranen, O. Hamming-Vrieze, S. van Beek, C. R. Nico Rasch, M. van Herk, and J. J. Sonke, "Deformable image registration for adaptive radiation therapy of head and neck cancer: Accuracy and precision in the presence of tumor changes," Int. J. Radiat. Oncol., Biol., Phys. 90, 680–687 (2014).

[34] C. Paganelli *et al.*, "Scale invariant feature transform in adaptive radiation therapy: A tool for deformable image registration assessment and re-planning indication," Phys. Med. Biol. 58, 287–299 (2013).

[35] D. Yang, H. H. Li, S. M. Goddu, and J. Tan, "CBCT volumetric coverage extension using a pair of complementary circular scans with complementary kV detector lateral and longitudinal offsets," Phys. Med. Biol. 59, 6327–6339 (2014).