Title: 4D dose reconstruction for scanned proton therapy using liver 4DCT-MRI

Shortened running title: 4DCT-MRI for proton dose reconstruction (max 40 characters)

Authors: MSc. Kinga Bernatowicz^{1,2}, PhD Marta Peroni², PhD Rosalind Perrin², Prof
Dr med Damian C. Weber^{1,3} and Prof Dr Anthony J. Lomax^{1,2}

¹ Department of Physics, ETH Zürich, 8092 Zürich, Switzerland

² Proton Therapy Center, Paul Scherrer Institute, 5232 Villigen PSI, Switzerland

³ Department of Radiation Oncology, University Hospital of Zürich, 8092 Zürich, Switzerland

Correspondence address: Kinga Bernatowicz, Paul Scherrer Institute, WBBA/007, 5232 Villigen PSI, Switzerland; Phone +41 56 310 50 16 Fax: +41 56 310 21 99; kinga.bernadowicz@psi.ch

Acknowledgment: This work is supported by the SNF grant 320030-1493942-1. Authors would like to acknowledge MGH Boston for sharing liver 4DCT data. Many thanks to ETH Zurich and University Basel for collaboration on 4DMRI image acquisition and processing.

Summary

4DCT-MRI can simulate many 4DCT data sets by combining patient CT with motion extracted from 4DMRI studies. 4DCT-MRI images have been generated using different motion extraction approaches and compared with 4DCT data of liver patients. This method extends the capabilities of motion modeling for accurate 4D dose calculations by accounting for realistic and variable motion patterns. 4DCT-MRI enables plan robustness studies to different motion conditions and 4D dose reconstruction of delivered plan for patient-specific QA.

Abstract

Purpose/Objective(s): 4DCT-MRI is an image processing technique for simulating many 4DCT data sets from a static reference CT and motions extracted from 4DMRI studies performed on either volunteers or patients. In this work, different motion extraction approaches have been tested on six liver cases, and a detailed evaluation of the 4DCT-MRI in comparison with 4DCT has been performed.

Methods and Materials: 4DCT-MRI has been generated using two approaches. First using motion extracted from 4DMRI as being 'most similar' to 4DCT of the same patient (subject-specific) and second, with the 'most similar' motion obtained from a motion library derived from 4DMRI liver studies of 13 healthy volunteers (population-based). The resulting 4DCT-MRI and 4DCT's have then been compared using scanned proton 4D dose calculations (4DDC).

Results: Dosimetric analysis showed that $93\pm 8\%$ of points inside the CTV agreed between 4DCT and subject-specific 4DCT-MRI (gamma analysis, 3%/3mm). The population-based approach however showed lower dosimetric agreement with only $79\pm 14\%$ points in the CTV reaching the 3%/3mm criteria.

Conclusions: 4DCT-MRI extends the capabilities of motion modeling for dose calculations by accounting for realistic and variable motion patterns, which can be directly employed in clinical research studies. We have found that the subject-specific liver modeling appears more accurate than the population-based approach. The former is particularly interesting for clinical applications, such as improved target delineation and 4D dose reconstruction for patient-specific QA to allow for inter- and/or intra-fractional plan corrections.

Introduction

4DCT imaging is widely used in radiotherapy, but does not represent inter-cycle respiratory motion variability, may contain artifacts when acquired under conditions of irregular breathing (Yamamoto et al 2008, Pan and Sun 2007, XXX) and can lead to target delineation errors and systematic treatment uncertainties (Persson et al 2010). In particular, artifacts can have a substantial effect on the accuracy of particle therapy treatments, where the calculated range of the particles may be adversely affected. Various motion mitigation techniques have been developed, to either improve breathing regularity or limit motion (Langner et al 2008, Mittauer et al 2015, Damkjær et al 2013), at the cost of patient comfort and relying on patient compliance. Radiation therapy under free breathing nevertheless remains a relevant goal and can even be an advantage for rescanned particle therapy, which is dependent on sufficient statistical averaging (XXX), or if tumor tracking can be performed (XXX, Eley et al 2014).

Studies show that 4D-MRI imaging is capable of resolving the irregular characteristics of organ motion (XXX, Dinkel et al 2009). Two 4D-MRI methods have been developed so far, both of which do not rely on an external surrogate for image sorting, but on a navigator slice (XXX), or on image content only (Paganelli et al 2015). While being a promising technique for motion evaluation, MR imaging alone does not provide density information for accurate dose calculations, with significant dose differences being reported compared to CT based calculations for photons when replacing segmented MR anatomy with bone and water-equivalent tissue (Eilertsen et al 2008).

The conversion of MRI to density data can be done using either voxel-based or anatomy-based segmentation methods. In the first case, voxels are directly characterized by different tissue classes, but discriminating between air and bone is still a major difficulty (Catana et al 2010). On the other hand, anatomy-based methods match reference density datasets (either the patient's CT or a MRI-CT atlas) using deformable registration to a new patient (Dowling et al 2012). Such methods were recently extended to 4D using motion models derived from 4DCT (Ehrhardt et al 2011) or 4D-MRI (XXX, Marx et al 2014). The latter approach is particularly interesting, as it extends the capabilities

of 4DCT imaging by also being able to represent multiple breathing cycles without extra imaging dose.

In this work, we expand on the study already published by XXX, where motions extracted from 4D-MRI of volunteers were used to deform a single phase CT data in order to generate what we call 4DCT-MRI images. 4DCT-MRI is compatible with dose calculation algorithms, but can also represent multiple, variable breathing cycles. Clinical applications of this approach include pre-treatment quality assurance procedures to test the robustness of a particular mitigation technique to realistic breathing motions and reconstruction of delivered ‘4D dose of the day’ based on the subject-specific motion model. As such, two different motion extraction approaches have been tested, subject-specific and population-based, on six liver patient cases. The results have been compared to 4DCT using three different metrics: (1) spatial accuracy of the images and extracted motions, (2) comparison of Hounsfield Units (HU) and proton range, and (3) scanned proton 4D dose calculations (4DDC).

Methods and Materials

4DCT-MRI concept

The concept of 4DCT-MRI has been previously proposed by XXX, and involves an originating single phase CT that is deformed using separate motion vector fields to form one or more simulated 4DCT data sets. The advantage of this is that the two information streams (density and motion) can be obtained from modalities that are best suited for these purposes. For instance, CT data is the modality of choice for imaging density, whereas motion can be more conveniently imaged using 4D-MRI - a no-dose imaging technique, which allows for the acquisition of motion data over multiple breathing cycles of patients and/or volunteers (XXX).

Data sources

CT data. For one patient, both 4DCT (SI motion 10 mm) and 4DMRI images have been analyzed to demonstrate the potential of subject-specific 4DCT-MRI. The reference CT was defined as the end-

inhale phase of 4DCT. As this patient was treated at our institute for a non-liver tumor, a simulated target volume was drawn in the right liver lobe ($CTV = 38\text{cm}^3$) for 4DDC.

In addition, 4DCT-MRI data has been derived for 5 liver patients with different tumor sizes [range: $70\text{-}403\text{cm}^3$; mean: 191cm^3] and motion characteristics [range: $9\text{-}17\text{mm}$; mean: 13mm]. 4DCT studies were acquired for all cases and all had fiducial markers implanted in the liver. The end-exhale phase has been used as the reference CT data for 4DCT-MRI generation.

Motion data. Motions have been extracted from two sources: (1) original 4DCTs from the patients, used as a reference to test our methodology and the accuracy of the deformation algorithms, and (2) from a library of liver motions, acquired either from 4D-MRI of 13 healthy volunteers (XXX) (for the population-based approach) or from the same patient (for the subject-specific approach). The motion library is a collection of XYZ displacements of corresponding liver points (see ***MRI-CT liver correspondence***) extracted from multiple breathing cycles of each imaged volunteer/patient (mean: 504; range: [221,851]). In total more than 6000 motion cycles are available in this library.

In order to make a fair comparison to the original (reference) 4DCT data sets, a 'most similar' motion to that of the 4DCT must be chosen from the library. For the subject-specific approach, the motion extracted from the patient's 4DCT was compared to the 'most-similar' motion found in the 464 breathing cycles available in the 4DMRI data of the same patient, whereas for the population-based approach, 4DCTs of five patients were compared with the 'most-similar' motions found from the volunteer library.

Motion extraction

4D-MRI was generated by retrospective sorting of the MRI sequences, based on the liver state determined from simultaneously acquired navigator slices (XXX). Motion was then extracted using a combination of a multi-resolution affine registration and a B-spline non-rigid registration. The correlation coefficient was used as a similarity measure without additional smoothing of the deformation fields, leading to registration accuracies of the order of one voxel or smaller (XXX).

Finally, this motion data was converted to match the number of phases in the corresponding reference 4DCT.

The reference motion fields were obtained by registering patient 4DCT images using an open source software (www.plastimatch.com). Five stages of multi-resolution optimization with gradual refinement of the grid spacing between control points in the B-spline grid have been used with a mean square error cost function and regularization lambda of 0.05 (Shackleford et al 2010, 2012). Registration was validated using marker position errors (MPE), calculated as the absolute difference between estimated motion (i.e. registration result) and the 'ground truth' motion (i.e. displacement of implanted fiducial markers obtained directly from the 4DCT images).

MRI-CT liver correspondence

Mapping motion information from 4D-MRI to CT images was done based on a pre-defined correspondence between livers (XXX). This assumes that certain corresponding points will move similarly across different livers, e.g. points on the anterior surface will show an SI-AP motion, as they slide along the abdominal wall. As such, a fine mesh (>500 points) is automatically created by gradual refinement of a coarse liver prototype to the segmented liver surface of each patient (see XXX).

In order to determine the 'most-similar' motion, a direct comparison of motion vectors at corresponding points was performed. Once found, a continuous vector field was generated by a B-spline based interpolation using the Insight Toolkit implementation (XXX).

Comparing 4DCT-MRI and 4DCT

Spatial evaluation. 4DCT and 4DCT-MRI images were compared visually (see fig. 1) and then using Summed-Square Differences (SSD), calculated for displacements over all breathing phases for points inside and on the liver surface. SSD was used as the metric for assessing 'most similar' motion from the motion library. Additionally, for the population-based approach, motion data was evaluated at the fiducial markers positions. Marker position errors (MPE) have been calculated as

an absolute difference between the motions extracted from 4D-MRI and the 'ground truth' motion extracted from the originating 4DCT data (see fig. 3(b)).

Density information. Deforming a static CT with externally derived motion vectors may result in somewhat different density information than for 4DCT acquisition (see fig. 1). Therefore, density changes for different incident field directions were quantified using water-equivalent range (WER) calculations through both the reference and 4DCT-MRI data sets. WERs were calculated by converting HU values to proton stopping power using a clinically validated calibration curve and were evaluated in the CTV region only.

Dosimetric comparison. Finally, 4D dose calculations (4DDC) for scanned proton beam have been performed for both 4DCT and 4DCT-MRI, using the deforming dose-grid approach (XXX, XXX) and ray-casting model (XXX), as used by the clinical treatment planning system at XX (XXX). In this algorithm, the positions of calculation points are adjusted considering time-dependent displacements in directions orthogonal to the central beam axis, with lateral displacements obtained from extracted motion fields, and water-equivalent depths from pre-calculated density maps at each motion phase (XXX). For all cases, a breathing period of 5s was assumed and a planning target volume (PTV) was defined using an isotropic expansion of 10mm around the CTV. 4DDC's have been calculated for four plans, three single field (P1 anterior-posterior, P2 right lateral and P3 anterior-superior oblique) and one three field plan (P4, combining P1-P3), similar to those described by XXX. For all 4DDC's, an energy switching time of 80 ms has been assumed (XXX, XXX).

Results

4DCT-MRI subject-specific

Spatial evaluation. 4DCT and the ‘most similar’ 4DCT-MRI for the subject-specific approach are compared in fig. 1(a): motion excursion of 4DCT-MRI resembles that of the original 4DCT. When comparing corresponding points in the liver between the two, SSD differences as low as 0.8mm were found.

Fig. 1: Comparison of 4DCT images with generated 4DCT-MRI for (a) patient-specific and (b) population-based approach. Note that biggest differences occur outside the liver volume (marked with red contour).

Density evaluation. On comparison of HU differences, a good agreement was observed within the liver volume (fig. 1(a)). However, larger differences appeared near ribs, in particular at the anterior site subjected to respiratory motion. Calculated WER’s between 4DCT and 4DCT-MRI were found to be <2mm.

Dosimetric evaluation. Fig. 2 compares dose distributions calculated for 4DCT and subject-specific 4DCT-MRI. Mean dose difference are <1%, and 93($\pm$ 8)% of CTV points pass the gamma evaluation (3%/3mm).

In summary, for the subject-specific approach, 4DCT-MRI has been found to be an excellent surrogate for modeling motion.

Fig. 2: 4D dose calculation results for 4DCT and patient-specific 4DCT-MRI for a single-field plan (P3) and three-field plan (P4); insert shows 3%/3mm gamma analysis – agreement is marked in blue and solid white line defines CTV.

4DCT-MRI – population-based

Spatial evaluation. Examples of 4DCT and 4DCT-MRI for the population-based approach are presented in fig. 1(b). Accuracy of matching of the ‘most similar’ motion from the 4D-MRI motion

library (compared to 4DCT extracted motion) varied both with respect to breathing phase (fig. 3(a)) and from patient to patient, with SSD's within the liver ranging from 1 to 12mm (mean 5.5mm).

For the inserted fiducials, mean MPE's were less than 4mm in 4 of 5 cases, fig. 3(b), which is slightly larger than the voxel size of the CT (voxel diagonal dimension ~3mm). For one case however (patient 4) MPE of 6 mm and a maximum of 9mm were found.

Fig. 3: (a) Reference CT (end-exhale; shown in green) overlaid with the moving phases deformed according to motion extracted from 4D-MRI (magenta). Matching regions are shown in gray scale. (b) Box plot of marker position errors calculated for motion extracted from the generated 4DCT-MRI images of 5 liver patients. Boxes correspond to the 75th percentiles of the MPE at different breathing phases.

Density evaluation. Example of HU differences are shown in fig. 1(b), and WER analysis for five patients and all phases of their 4D data in fig. 4. Mean absolute WER deviations were 2mm with a maximum difference of 25mm (34% of WER_{EE}). For maximum ranges, mean differences were 2mm ranging to a maximum of 24mm, whereas for minimum range, mean differences were 3mm, but differences of up to 63mm were also observed. All worst case differences were found for the same patient (patient 4).

Fig. 4: (a) Mean, (b) maximum, and (c) minimum water-equivalent range (WER) differences between 4DCT and 4DCT-MRI images, calculated in the CTV of 5 liver patients, at 10 respiratory phases, and for 3 different field directions (from anterior, from the right and a non-coplanar field in between). (d) An example of original and warped image liver data (patient 2).

Dosimetric evaluation. Fig. 5 shows a comparison of 4DDC results for 4DCT and population-based 4DCT-MRI. Mean dose differences were about 1%, except for case 4, where they were up to ~4%, see fig. 6(a). On average 79(±14%) of points within the CTV met the gamma criteria of 3%/3mm, fig. 6(b). However, for 3/5 cases with 'most similar' SSD's <2mm, mean dose differences were lower than 1% and more than 80% of points met the gamma criteria of 3%/3mm for single field

plans.

Fig. 5: Example of the 4D dose calculation results for 4DCT and 4DCT-MRI for case 2; insert shows 3%/3mm gamma analysis – agreement is marked in blue.

Fig. 6: Comparison of 4DCT and 4DCT-MRI: (a) mean dose differences and (b) CTV volume meeting the 3%/3mm gamma criteria. C1-C5 refer to the five different cases (patients) studied and P1-P4 to treatment plans.

Discussion

We envisage three main clinical applications for 4DCT-MRI: (1) to evaluate plan robustness under different breathing, for instance using an appropriate model to select different percentiles of motion vectors from the library, (2) to improve volume delineation, employing the superior MRI contrast in liver and realistic motion excursions, particularly needed for ITV definition, and (3) to reconstruct the '4D dose of the day', based on the time-resolved delivery log-files (Tyagi et al 2012, XXX) and motion reconstructed from a surrogate combined with a motion model (XXX). Detected dose errors could then be compensated by plan adaptations, and applied in subsequent treatment fractions. Additionally, online plan QA is an interesting idea, where tools similar to those mentioned above (but very fast) would allow treatment intervention in case of erroneous dose delivery. Although, there are many challenges to be addressed before implementing such sophisticated methods, similar concepts are currently being developed for the photon therapy applications (Ravkilde et al 2014).

We have shown that the subject-specific 4DCT-MRI provides an excellent surrogate to 4DCT with the big advantage being 4DCT-MRI data can be generated for multiple and different breathing cycles. Although results for the population-based approach are not as good, such an approach could nevertheless still be of use for a-priori robustness tests of the sensitivity of plans to a range of different, realistic breathing scenarios.

However, all results are dependent on the quality of the reference CT image, as seen for example in case 4 of the population-based approach, where the quality of the 4DCT images was poor to start with, with high-intensity streaking artifacts, limiting the validity deformable registration results. Moreover, such data should not be used for realistic planning and dose calculations. In general, the end-exhale phase is shown to have the least motion artifacts (Pan et al 2007) and is preferable for the 4DCT-MRI generation. However, if image quality is acceptable, other breathing phases can be used (as demonstrated on end-inhale with the subject-specific approach). In clinical practice, 4DCT-MRI images could be also generated based on a reference CT acquired with a fast CT scan or a gated-CT protocol, which has a clear advantage in terms of imaging dose compared to 4DCT

acquisition.

The current model and analysis is however specific to the liver. The method could potentially be extended to other sites/organs if tissue properties are preserved. This could be a challenge in lung tissue, where WER changes may be larger: Mori et al (2009) reported WER changes of 1.5cm in lung between different CT scans.

4DCT-MRI images were matched and evaluated based on the current clinical 'gold standard' - 4DCT images. 4DCTs represent phases of a single, averaged breathing cycle, which may underestimate the target excursion and exhibit imaging artifacts that present challenges to deformable registration algorithms (Persson et al 2010). However, it is difficult to know whether motion extracted from 4DCT or 4D-MRI is the best representation of the ground truth motion. This is one of the reasons why validation of 4D medical image registration is very challenging. The use of sophisticated, anthropomorphic moving phantoms (XXX) could be a first-step towards development and evaluation of advanced motion imaging techniques.

Conclusion

The 4DCT-MRI method extends the capabilities of 4DCT imaging by generating multiple-breathing cycle image sets of the patient. It represents the variability of deformable liver motion on high-quality images with few discontinuities and relevant density information. The potential clinical applications of 4DCT-MRI include, but are not restricted to, plan robustness evaluations under variable motion conditions and patient-specific QA for moving targets. 4DCT-MRI can also be directly employed in clinical research studies on advanced 4D treatment planning and delivery techniques taking into account patient movement.

References

XXX

XXX

XXX

XXX

XXX

Catana C, van der Kouwe A, Benner T, Michel C J, Hamm M, Fenchel M, Fischl B, Rosen B, Schmand M and Sorensen A G 2010 Toward implementing an MRI-based PET attenuation-correction method for neurologic studies on the MR-PET brain prototype J. Nucl. Med. 51(9) 1431-8

Damkjær S M, Aznar M C, Pedersen A N, Vogelius IR, Bangsgaard J P and Josipovic M 2013 Reduced lung dose and improved inspiration level reproducibility in visually guided DIBH compared to audio coached EIG radiotherapy for breast cancer patients Acta. Oncol. 52(7) 1458-63

Dinkel J, Hintze C, Tetzlaff R, Huber P E, Herfarth K, Debus J, Kauczor H U and Thieke C 2009 4D-MRI analysis of lung tumor motion in patients with hemidiaphragmatic paralysis Radiother. Oncol. 91(3) 449-54

Dowling J A, Lambert J, Parker J, Salvado O, Fripp J, Capp A, Wratten C, Denham J W and Greer P B 2012. An atlas-based electron density mapping method for magnetic resonance imaging (MRI)-alone treatment planning and adaptive MRI-based prostate radiation therapy Int. J. Rad. Oncol. Biol. Phys. 83(1) e5–11

Ehrhardt J, Werner R, Schmidt-Richberg A and Handels H 2011 Statistical modeling of 4D respiratory lung motion using diffeomorphic image registration IEEE Trans. Med. Imaging. 30(2) 251-65

Eilertsen K, Vestad L N, Geier O and Skretting A 2008 A simulation of MRI based dose calculations

on the basis of radiotherapy planning CT images *Acta. Oncol.* 47(7) 1294-302

Eley J G, Newhauser W D, Lüchtenborg R, Graeff C and Bert C 2014 4D optimization of scanned ion beam tracking therapy for moving tumors *Phys. Med. Biol.* 59(13) 3431-52

Langner U W and Keall P J 2008 Prospective displacement and velocity-based cine 4D CT *Med. Phys.* 35(10) 4501-12

Marx M, Ehrhardt J, Werner R, Schlemmer H P and Handels H 2014 Simulation of spatiotemporal CT data sets using a 4D MRI-based lung motion model *Int. J. CARS.* 9(3) 401-409

Mittauer K E, Deraniyagala R, Li J G, Lu B, Liu C, Samant S S, Lightsey J L and Yan G 2015 Monitoring ABC-assisted deep inspiration breath hold for left-sided breast radiotherapy with an optical tracking system *Med. Phys.* 42(1) 134

Mori S, Lu H M, Wolfgang J A, Choi N C and. Chen G T 2009 Effects of interfractional anatomical changes on water-equivalent pathlength in charged-particle radiotherapy of lung cancer *J. Radiat. Res.* 50 513-519

Paganelli C, Summers P, Bellomi M, Baroni G and Riboldi M 2015 Liver 4DMRI: A retrospective image-based sorting method. *Med Phys.* 42(8) 4814

Pan T and Sun X 2007 Improvement of the cine-CT based 4D CT imaging *Med. Phys.* 34(11) 4499-4503

Persson G, Nygaard D, Brink C, Jahn J, Per Munck R and Specht L 2010 Deviations in delineated GTV caused by artefacts in 4DCR Radiother. *Oncol.* 96 61-66

Ravkilde T, Keall PJ, Grau C, Høyer M, Poulsen PR 2014 Fast motion-including dose error reconstruction for VMAT with and without MLC tracking. *Phys Med Biol* 59(23), 7279-96.

Shackleford J A, Kandasamy N and Sharp G C 2010 On developing B-spline registration algorithms for multi-core processors *Phys. Med. Biol.* 55(21) 6329-51

Shackleford J A, Yang Q, Lourenço A M, Shusharina N, Kandasamy N and Sharp G C 2012

Analytic regularization of uniform cubic B-spline deformation fields Med. Image. Comput. Comput. Assist. Interv. 15(2) 122-9

Tyagi N, Yang K, Gersten D and Yan D 2012 A real time dose monitoring and dose reconstruction tool for patient specific VMAT QA and delivery. Med Phys 39, 7149.

Yamamoto T, Langner U, Loo B, Shen J and Keall P J 2008 Retrospective analysis of artifacts in four-dimensional CT images of 50 abdominal and thoracic radiotherapy patients Int. J. Radiat. Oncol. Biol. Phys. 72(4) 1250-8

(a) 4DCT vs patient-specific 4DCT-MRI

(b) 4DCT vs population-based 4DCT-MRI

max: 129.1 %

4D-CT : P3

4D-CT(MRI) : P3

120.0 %

110.0 %

100.0 %

90.0 %

80.0 %

70.0 %

60.0 %

50.0 %

40.0 %

30.0 %

20.0 %

10.0 %

0.0 %

4D-CT : P4

4D-CT(MRI) : P4

(a) Registration error: Overlay of reference and moving 4DCT-MRI images

(b) Registration error: Marker position errors from 4DCT-MRI images

max: 135.4 %

4D-CT : P1

4D-CT(MRI) : P1

4D-CT : P4

4D-CT(MRI) : P4