ICART: Imaging and Computer Assistance in Radiation Therapy

Marta Peroni, Alexander Schlaefer, Wolfgang Birkfellner, Chiara Gianoli, Antje-Christin Knopf, Jamie R. Mcclelland, Marco Riboldi, Simon Rit

MICCAI 2015 MUNICH
18th International Conference 2015 · October 5th to 9th · Munich, Germany
on Medical Image Computing and Computer Assisted Interventions

ICART:

Imaging and Computer Assistance in Radiation Therapy

A workshop held on Friday 9th October at part of MICCAI 2015 in Munich, Germany.

Preface

Radiation Therapy has been a key clinical application scenario for image guidance, image processing and motion compensation. It remains one of the most prominent clinical application fields for new imaging and image processing methods and it is highly computerized. Recent advances in fast and integrated imaging lead to image guided setup and inter-fractional plan adaptation. Current trends point to real-time image guidance and compensation of intra-fractional motion and deformation as well as continuous monitoring of macro- and microscopic changes in the tumor environment.

This workshop will present novel and state-of-the-art approaches for medical image processing in image-guided radiotherapy (IGRT), including segmentation and registration, treatment planning and plan adaptation, image guidance and motion compensation, and imaging and visualisation. It will also provide an exciting venue for exchanging experiences and ideas on clinical and technical challenges, and for discussing recent trends in research and clinical application. The 20 papers accepted for presentation at this workshop are included in these proceedings.

It is a privilege to hold this workshop in the context of MICCAI, and we appreciate that the organizers provide this exciting venue. We would like to thank all authors, program committee, and reviewers for helping to compile a set of highly interesting papers and we would also like to acknowledge the use of EasyChair (www.easychair.org) when preparing the proceedings.

August 5, 2015
London

Marta Peroni
Alexander Schlaefer
Wolfgang Birkfellner
Chiara Gianoli
Antje-Christin Knopf
Jamie McClelland
Marco Riboldi
Simon Rit

Table of Contents

Program Committee

Oscar Acosta	INSERM, U19099, University of Rennes 1, LTSI
Christoph Bert	Department of Radiation Oncology, University Clinic Erlangen , Friedrich-Alexander-University Erlangen-Nürnberg and GSI Helmholtz Centre for Heavy Ion Research, Darmstadt, Germany
Wolfgang Birkenfellner	Center for Medical Physics and Biomedical Engineering, Medical University Vienna
Kristy Brock-Leatherman	University of Michigan
Philippe Cattin	Medical Image Analysis Center (MIAC), University of Basel
Floris Ernst	Institute for Robotics and Cognitive Systems, University of Lübeck
Christoph Fürweger	European Cyberknife Center, Munich
Chiara Gianoli	Experimental Physics and Medical Physics, Ludwig-Maximilians Universität München
Antje Christin Knopf	Joint Department of Physics at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust
Daniel Low	Department of Radiation Oncology, UCLA
Jamie McClelland	Centre for Medical Image Computing, Dept. Medical Physics and Bioengineering, University College London
Katia Parodi	Experimental Physics and Medical Physics, Ludwig-Maximilians Universität München
Marta Peroni	Paul Scherrer Institut
Per Rugaard Poulsen	Department of Oncology, Aarhus University Hospital and Institute of Clinical Medicine, Aarhus University
Bas Raaymakers	Department of Radiotherapy, University Medical Center Utrecht
Marco Riboldi	Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano
Simon Rit	Université de Lyon - CREATIS
Dan Ruan	Department of Radiation Oncology, Department of Bioengineering, UCLA
Sairos Safai	Center for Proton Therapy, Paul Scherrer Institut
Alexander Schlaefer	Institute of Medical Technology, Hamburg University of Technology
Gregory C. Sharp	Massachusetts General Hospital and Harvard Medical School
Jan-Jakob Sonke	Department of Radiation Oncology, The Netherlands Cancer Institute-Antoni van Leeuwenhoek Hospital

Philipp Steininger	Institute for Research and Development on Advanced Radiation Technologies (radART), Paracelsus Medical University Salzburg
Marcel Van Herk	Institute of Cancer Sciences, The University of Manchester
Jef Vandemeulebroucke	Department of Electronics and Informatics and iMinds Department of Medical IT, Vrije Universiteit Brussel
René Werner	Department of Computational Neuroscience, Center of Experimental Medicine, University Medical Center Hamburg-Eppendorf
Lei Xing	Department of Radiation Oncology, Stanford University
Paolo Zaffino	ImagEngLab - Experimental and Clinical Medicine Department, Magna Graecia University

Additional Reviewers

Simon, Antoine	INSERM, U19099, University of Rennes 1, LTSI

Machine-learning based image segmentation using Manifold Learning and Random Patch Forests

K. Fritscher[1], P. Raudaschl [1], P. Zaffino[2], G.C. Sharp[3], M.F. Spadea[2], R. Schubert[1]

[1] Department for Biomedical Image Analysis, UMIT, Austria
[2] Department of Experimental and Clinical Medicine, Magna Graecia University, Italy
[3] Massachusetts General Hospital, Harvard Medical School, USA

Abstract. Accurate segmentation of organs at risk is an essential precondition for successful treatment planning in radiotherapy. In this paper we introduce a novel approach for the fully automatic segmentation of organs at risk in CT images.

The method is based on the usage of Manifold Learning in order to extract features of multi-scale image patches. In combination with spatial locality information, the resulting embedding coordinates are used as features for a multi-step classification approach based on using multiple Random Patch Forests.

Experiments using 17 CT images of the head-neck region demonstrate the high potential of the presented method. Quantitative evaluation shows that the approach provides accurate segmentation results even with a small number of training subjects.

1 Introduction

Efficient and exact delineation of anatomical structures is one of the key areas in medical imaging. This is especially true for medical fields like radiation oncology, in which the segmentation of anatomical structures (e.g. organs at risk (OAR)) directly acts as an input for treatment planning. Hence, the need for methods that provide accurate segmentation with minimal user input is obvious. Most frequently used segmentation approaches in radiotherapy are (multi) atlas based segmentation (MABS) strategies and methods based on statistical models of shape and/or appearance (SSM, SAM). MABS offers a high level of robustness and usually does not need any user interaction. Model based approaches on the other hand provide closed surface and anatomically plausible shapes, which is a desirable property for most applications.

Despite their advantageous properties, both approaches suffer from shortcomings when the anatomical structure of interest shows high morphological variability. Most methods for deformable registration have problems to correctly deform structures showing high shape variability. On the other hand the creation of statistical shape models for structures with highly varying shape or even topology is extremely challenging and often requires a large number of labeled training images.

In this paper a novel approach for the fully automatic segmentation of organs at risk in CT images is introduced. Using the proposed method, image patches with different sizes and shapes are extracted for voxel neighborhoods in predefined regions of interest. Using these patches, manifold learning is performed by using Laplacian Eigenmaps in order to obtain the embedding coordinates of each voxel/patch in the patch manifold space. The low-dimensional embedding is then used in combination with locality information to train a Random Forest Classifier in a two-step approach. During the training phase, multiple Random Forests (RF) are trained in order to assess the correct labeling for 9 different locations within a voxel neighborhood. The final class label for each voxel is obtained from the probability maps coming from all 9 random forest classifiers.

The presented approach is evaluated by segmenting the parotid glands in 17 head neck CT scans of cancer patients. Being the major salivary gland, parotid glands are highly critical OARs for radiotherapy of head and neck cancer patients. Unintended irradiation of the parotid glands can result in difficulties for mastication and deglutition. Apart from this, the automatic segmentation of parotid glands is highly challenging. This is due to high anatomical shape variability of the parotid gland on the one hand and low soft-tissue contrast as well as image noise (caused by dental implants) on the other hand.

1.1 Related work

The usage of random forests [1] is highly popular in numerous machine learning applications - not only, but also in medical imaging. Criminisi proposed random forests for the localization of anatomical organs [2], [3]. In addition, random forests were used for segmentation, e.g. for the segmentation of brain tumors [4] or prostates in MRI [5]

Patch based segmentation has also been applied frequently for different anatomical structures [6]–[8]. In [9] the usage of geodesic distances in combination with patch based segmentation and spatial-context information was proposed. However, the usage of this spatial information is dependent on user defined landmarks or additional labeled structures. In [10] the benefit of different features for patch based segmentation has been evaluated for the segmentation of the parotid glands leading to very accurate segmentation results. In addition, patch based segmentation was used in combination with atlas based segmentation [7], [11], [12] as an additional refinement step.

In [13] class probabilities have been used to train a random forest classifier for infant brain segmentation. In [14] iterative, voxel-wise classification was applied for the segmentation of the prostate. However, using this approach segmentation is dependent on previously segmented prostates of the same patient.

A combination of Laplacian Eigenmaps [15] and Random Forests for segmentation has been proposed in [16]. However, the approach is based on computing Laplacian Eigenmaps using whole images instead of patches as input. This does not only result in high computational demands, but also requires a large set of labeled training images.

Summing up, both Random Forest Classification/Regression and Manifold Learning are very active fields of research. However, to the best of our knowledge, the presented combination of 1) Manifold Learning based on multi-scale image patches, 2) Iterative Random Forest Training based on embedding coordinates resulting from Laplacian Eigenmaps and 3) patch based label fusion using multiple random forests 4) without the need for prior elastic registration has not been presented before. The evaluation will show that the proposed method provides accurate segmentation results and performs equally well or better than existing methods for the segmentation of parotid glands in CT images.

2 Methods

2.1 Preprocessing and patch extraction

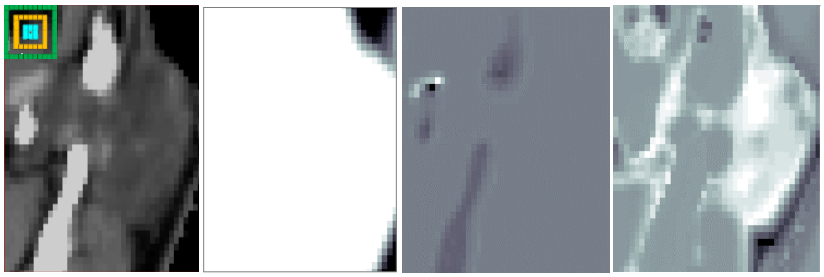
In a first preprocessing step, all images of the training set are registered onto a reference dataset using a rigid transform. In order to account for image noise Mutual Information (MI) is used as a metric. Registration is performed separately for left and right parotid gland using different sub-regions (defined in the reference dataset) for MI computation.

After registration, patches are extracted for each voxel within a predefined region of interest (ROI) containing the respective structure(s) of interest (in this project the parotid gland). The ROI is defined in the reference image and must have the same size for all training and test images. Patches with a 3x3x3 (type A), 7x7x3 (type B) and 11x11x5 (type C) neighborhood are extracted for each voxel. In order to avoid redundancies in overlapping regions and also to reduce the computational burden for the following steps, only patch type A contains all voxels within the patch. Patch types B and C only contain the border voxels of the respective patches (see also fig. 1 (left)). Anisotropic size of the patches is chosen to account for the anisotropic resolution of the images used in this project (1.2x1.2x2.5mm voxel size). Using these patches independently for the following manifold learning step shall on the one hand provide information about small structures (type A) and on the other add knowledge about the appearance of neighboring structures as well as robustness towards local noise (types B and C)

2.2 Manifold learning

It has been shown in various research projects, that non-linear manifold learning approaches like Laplacian Eigenmaps (LapE) [15] can successfully be used for unravelling the inherent structure of the input data [17].
LapE compute a low-dimensional representation of the data in which the distances between a datapoint and its k nearest neighbors are minimized. First a neighborhood graph G is computed, in which every datapoint is connected to its k nearest neighbours. Based on using the Gaussian kernel function the weights for each edge in G are computed and stored in adjacency matrix W. For the computation of the low dimen-

sional representations y_i the cost function $\phi(Y) = \sum_{ij} \|y_i - y_j\|^2 \omega_{ij}$ is used. Using spectral graph theory, the minimization of the cost function can be defined as an eigenproblem so that minimizing $\phi(Y)$ is proportional to minimizing $Y^T L Y$, where the graph Laplacian $L = M - W$. Degree matrix M is a diagonal matrix, whose entries are the row sums of W. By solving $Lv = Mv\lambda$, d eigenvectors $v_i$corresponding to the d smallest eigenvalues can be calculated. As a result, each patch p can be represented by its respective position in patch space. Fig. 1 is presenting the result of manifold learning based on LapE using patch type A as input. Looking at fig. 1, one can see that the resulting eigenvectors have the potential to separate anatomically meaningful structures like skin, bone or different types of soft tissue.



sample *EV1* *EV2* *EV3*

Fig. 1. Illustrating the output of LapE by projecting the embedding coordinates back into image space (sample: a sample from an original image with an illustration of the applied patch configuration in the upper left corner;EV1-EV3: respective embedding coordinates along the coordinate axis formed by 3 exemplary eigenvectors for patch type A)

In this project, LapE are calculated independently for the 3 patch types A, B, C. The resulting embedding coordinates for each patch type are used as feature vector f for the following classification step based on the usage of Random Forests. In this project a value of $d = 20$ was chosen, which results in 60 image features per patch.

In addition to the image features, the offset of each voxel relative to the corner voxel of the predefined ROI is also added to the feature set. By this means, two patches which have similar appearance, but come from very different locations, can be more easily distinguished. Apart from this, the location information can be regarded as a "weak" shape model, which assists in inferring the location (and vague shape) of the parotid gland. Usage of this additional locality information is also the reason why rigid pre-registration was performed during preprocessing.

2.3 Voxel Classification using Random Patch Forests

Random Forest Classification is applied in order to determine a class label $c \in C$ for a given test voxel $x \in \Omega$.

Training. Using a feature representation f for each voxel of the training set, each tree t learns a weak predictor $p(c|f)$. In each node of a tree, the set of training voxels

is split based on a binary test using $s(x; f_k) < t$, where t is a threshold value for each split node and f_k represents the k^{th} feature of f. Based on the test result, the respective voxel is sent to one of its child nodes. During training f_k and t are optimized in a way that the information gain among class distributions is maximized [16]. For improved generalization ability, randomness is induced by only using a randomly sampled subset of f to optimize the parameters for each split node. Optimization is stopped when the maximum information gain falls under a predefined threshold. Growth of a tree stops when a predefined depth is reached or when a node contains less than a predefined number of training samples.

Testing. During testing each new voxel x of an unseen image is pushed through each trained tree by applying the learnt split parameters f_k and t. For each tree the tested voxel will read different leaf nodes, which are associated with different class probabilities. The final probability for a test voxel based on a single random forest is calculated using $p_F(c|x) = \frac{1}{T}\sum_{t=1}^{T} p_t(c|f)$, where T is the number of trees that are used.

Multiple Forests. In the presented approach, classification for each voxel is obtained by using multiple forests. Instead of training only one forest to obtain the label for the center voxel of an image patch, 8 additional RF are calculated in order to assess the class label for the corner voxels of a 3x3x3 neighborhood. Using highly overlapping patches, the probability for a given voxel can then be obtained by $P_m(c|x) = \frac{1}{9}\sum_{t=1}^{9} p_F(c|f)$. The usage of multiple forests is supposed to increase the robustness of the final labeling especially in the presence of image noise compared to using only one random forest.

Two Step Random Forest training and testing. Similar to the approach presented in [13], a two-step approach is used for RF training and testing: In a first training iteration a feature vector $f = \{p_1^A, \dots, p_{20}^A, p_1^B, \dots, p_{20}^B, p_1^C, \dots, p_{20}^C, i_x, i_y, i_z\}$ is used. p_d^L refers to the features obtained for each patch type using Laplacian Eigenmaps and $\{i_x, i_y, i_z\}$ refer to the offset of each voxel relative to the center (voxel) of the predefined ROI (see also section 2.2).Using the Multiple Random Forest approach described above 9 probability maps $P_{f\ 1..9}^1$ for each training dataset can be obtained.

For the second training iteration the probability maps $P_{f\ 1..9}^1$ are also fed into the classifier in addition to the features used in the first iteration. Based on the fact that similar shaped structures should also result in similar probability maps, the inclusion of the probability maps in the second RF training adds additional locality and shape information to the feature vector, which is expected to be useful for improved RF classification.

During testing, the same two step approach that was used for RF training is applied in order to obtain a final probability map P_m and a final class label C for each x in an unseen image. C is obtained by using

$$C\big(p(x,f)\big) = \begin{cases} 1, p(x,f) > 0.5 \\ 0, p(x,f) \leq 0.5 \end{cases} \quad (1)$$

3 Results

Evaluation scenarios: On the whole, 4 different scenarios have been evaluated in order to test the influence of the suggested contributions:

- **I)** Only LapE embedding coordinates for patch type A in combination with 1 training iteration and 1 RF (for center voxel) are used
- **II)** Equal to I, but using LapE embedding for all patch types
- **III)** Equal to II, but using 9 RF
- **IV)** Equal to III, but using 2 RF iterations.

During evaluation only one LapE had to be computed due to the usage of a leave-one-out strategy. In other settings out-of-sample extension for LapE could be applied for improved efficiency.

Data and parameter settings. The presented approach has been evaluated using 17 CT images of the head neck region with manually labeled parotid glands that acted as ground truth. Image resolution is 1.25x1.25x2.5mm. Accuracy of the segmentation is measured using the DICE coefficient (=$2|A \cap B|/(|A| + |B|)$, where A represents the ground truth and B the results using the proposed approach. In addition, maximum Hausdorff Distance is used to quantify the largest segmentation error.

Pre-registration of the test image to a randomly chosen training image has been performed using rigid transform in combination with Mutual Information metric (to account for the significant noise level in some of the test images). The size for the ROI of which the patches have been extracted was 50x50x87.5mm left Parotid) and 68x75x87.5mm (right parotid) respectively. Based on observations in preliminary test runs the size of the neighborhood for the computation of the Laplacian Eigenmaps was set to $n = 50$. As already mentioned, 20 embedding coordinates have been calculated for each patch type. A leave-one-out cross validation approach was used for the evaluation of all 4 scenarios.

Figure 2 presents results for all 4 testing scenarios. More specifically, the average Dice coefficient and average values for max. Hausdorff distance for left and right parotid glands are depicted. Error bars indicate 25th and 75th percentile.

Figure 3 shows the final labels as well as the final probability maps (for one sample region of an exemplary dataset) resulting from RF classification using scenarios II-IV.

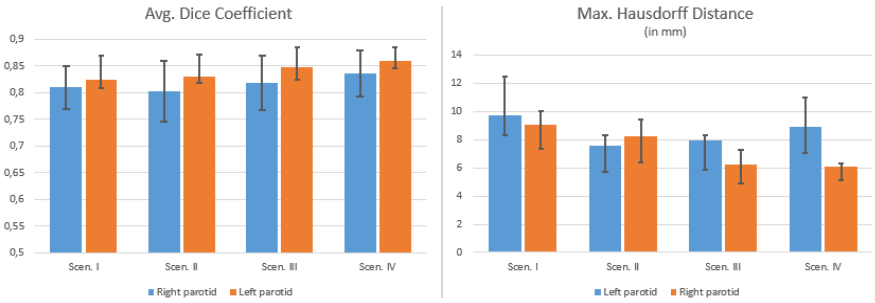


Fig. 2. Comparing average Dice scores and max. Hausdorff distances obtained for left and right parotid glands for all 4 scenarios explained in section 3.

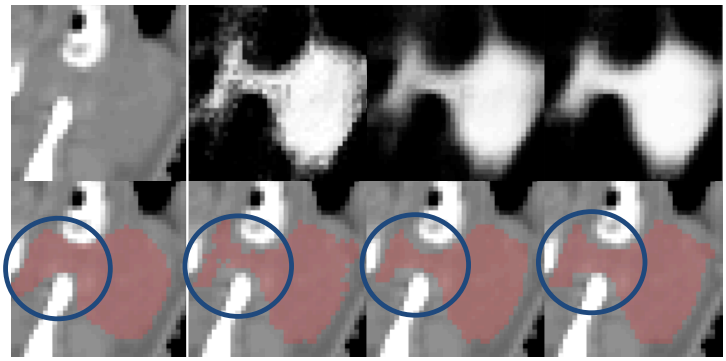


Fig. 3. First **col.**: Sample of original image (top row) and manual segmentation (bottom row); **second to fourth col.**: final probability maps (top row) and corresponding segmentation results (bottom row) obtained with scenarios II-IV, blue circles mark an exemplary area in which gradual improvements of the segmentation accuracy from scen. II – IV can be observed best

4 Discussion

A novel approach based on the usage of manifold learning in combination with a multi-step learning approach based on random patch forests has been presented. The evaluation has shown that the presented method is highly suitable for the segmentation of the parotid glands in CT scans. It could be shown, that the different extensions based on using of multi scale patches (scenario II), an ensemble of random forests (scenario III), as well as the usage of a two-step Random Forest classification scheme (scenario IV) leads to an increase of the respective Dice scores. It can also be observed, however, that max. Hausdorff Distances are lowest using Scenario II for the right parotid gland. This might be due to a smoothing effect that occurs in scenarios III and IV. This smoothing effect (which is also visible in fig. 3) potentially caused that single voxels which are labeled as foreground in scenario II are labeled as back-

ground in scenario III. That has very little or no effect on real segmentation quality, but a positive effect on max. Hausdorff values. This is also supported by the fact that Dice scores are higher using scenarios III and IV. Looking at fig. 3, it can also be observed that the two-step RF classification used in scen. IV results in more distinct probability maps leading to higher segmentation accuracy compared to the ground truth.

Compared to other approaches for parotid gland segmentation which used the same dataset [18] or a (sub)-set of the dataset used in this work [7] for evaluation the presented approach provides equal or higher Dice scores and smaller Hausdorff distances. Also compared to the Dice scores presented in other publications [18]–[20] the obtained results are highly competitive. However, it is also evident that several aspects of the presented approach can potentially be improved: e.g. alternative patch configurations will be tested in order to further increase the positive impact of using multi-scale patches. Moreover, the usage of additional locality information and more sophisticated voting schemes for classification based on multiple RFs as well as the application of different sampling approaches for increased efficiency will be evaluated.

References

1. L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, pp. 5–32, 2001.
2. O. Pauly, B. Glocker, A. Criminisi, D. Mateus, A. M. Möller, S. Nekolla, and N. Navab, "Fast multiple organ detection and localization in whole-body MR dixon sequences," in *LNCS,* 2011, vol. 6893 LNCS, pp. 239–247.
3. A. Criminisi, D. Robertson, E. Konukoglu, J. Shotton, S. Pathak, S. White, and K. Siddiqui, "Regression forests for efficient anatomy detection and localization in computed tomography scans," *Med. Image Anal.*, vol. 17, pp. 1293–1303, 2013.
4. D. Zikic, B. Glocker, and A. Criminisi, "Atlas encoding by randomized forests for efficient label propagation," in *LNCS,* 2013, vol. 8151 LNCS, pp. 66–73.
5. S. Ghose, J. Mitra, A. Oliver, R. Marti, X. Llad, and I. Segmentation, "A random forest based classification approach to prostate segmentation in MRI," *MICCIA Grand Challenge*, 2012.
6. Z. Wang, C. Donoghue, and D. Rueckert, "Patch-based segmentation without registration: Application to knee MRI," in *LNCS*, 2013, vol. 8184 LNCS, pp. 98–105.
7. C. Wachinger, G. C. Sharp, and P. Golland, "Contour-driven regression for label inference in atlas-based segmentation," in *LNCS*, 2013, vol. 8151 LNCS, pp. 211–218.
8. P. Coupé, J. V. Manjón, V. Fonov, J. Pruessner, M. Robles, and D. L. Collins, "Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation," *Neuroimage*, vol. 54, pp. 940–954, 2011.
9. Z. Wang, K. Bhatia, B. Glocker, and A. Marvao, "Geodesic patch-based segmentation," MICCAI 2014 (pp. 666-673) ,2014.
10. C. Wachinger, M. Brennan, G. Sharp, and P. Golland, "On the Importance of Location and Features for the Patch-Based Segmentation of Parotid Glands," *people.csail.mit.edu*.
11. X. Han, M. Learning, and in M. Imaging, "Learning-boosted label fusion for multi-atlas auto-segmentation," *Mach. Learn. Med. Imaging*, 2013.
12. W. Bai, W. Shi, D. P. O'Regan, T. Tong, H. Wang, S. Jamil-Copley, N. S. Peters, and D. Rueckert, "A probabilistic patch-based label fusion model for multi-atlas segmentation with

registration refinement: Application to cardiac MR images," *IEEE Trans. Med. Imaging*, vol. 32, pp. 1302–1315, 2013.
13. L. Wang, Y. Gao, F. Shi, G. Li, J. Gilmore, W. Lin, D. Shen, "LINKS:learning-based multi-source integration framework for segmentation of infant brain images," *Neuroimage*, 2015.
14. Y. Gao, S. Liao, D. Shen, and M. physics, "Prostate segmentation by sparse representation based classification," Medical physics, 39(10), 6372-6387.
15. M. Belkin and P. Niyogi, "Laplacian Eigenmaps for Dimensionality Reduction and Data Representation," *Neural Computation*, vol. 15. pp. 1373–1396, 2003.
16. H. Lombaert, D. Zikic, A. Criminisi, and N. Ayache, "Laplacian Forests: Semantic Image Segmentation by Guided Bagging," MICCAI 2014 (pp. 496-504) 2014.
17. P. Aljabar, R. Wolz, and D. Rueckert, "Manifold learning for medical image registration, segmentation, and classification," Medical Imaging Intelligence and Analysis, 351.
18. K. Fritscher, M. Peroni, P. Zaffino, M. Spadea, G. C. Sharp, R. Schubert, "Automatic segmentation of head and neck CT images for radiotherapy treatment planning using multiple atlases, statistical appearance models, and geodesic active contours ," *Medical Physics, 41,* 2014.
19. V. Fortunati, R.F. Verhaart, F. van der Lijn, W.J. Niessen, J.F. Veenland, M. M. Paulides, T. van Walsum, "Tissue segmentation of head and neck CT images for treatment planning: A multiatlas approach combined with intensity modeling," *Medical physics* 40.7 (2013)
20. A. A. Qazi, V. Pekar, J. Kim, J. Xie, S. L. Breen, and D. A. Jaffray, "Auto-segmentation of normal and target structures in head and neck CT images: a feature-driven model-based approach.," *Medical Physics*, vol. 38, no. 11, pp. 6160–70, Nov. 2011.

A novel atlas-selection approach for multiple atlas segmentation based on Manifold Learning and Random Forests using Multi-Scale Image Patches

P. Raudaschl[1], K. Fritscher[1], P. Zaffino[2], G.C. Sharp[3], M.F. Spadea[2], R. Schubert[1]

[1] Department for Biomedical Image Analysis, UMIT, Austria
[2] Department of Experimental and Clinical Medicine, Magna Graecia University, Italy
[3] Massachusetts General Hospital, Harvard Medical School, USA

Abstract. Atlas-based segmentation is a frequently used approach in medical imaging and multi atlas-based segmentation (MABS) has achieved great success for various applications. In order to simultaneously exploit the capabilities of MABS, limit execution time and maintain robustness, it is preferable to select a (preferably small) subset of atlases to be used for segmentation.
In this work, an atlas selection strategy using Manifold Learning and Random Forest Regression is presented. The approach aims at learning the relationship between the pairwise appearance of structures and the Dice coefficients of their respective labeling. For this purpose, multi-scale patches in relevant regions of interest are extracted. Local patch models based on linear and non-linear dimensionality reduction are created. Resulting coordinates of the patch embeddings are used to train a regression model for Dice score prediction using Random Forests. Predicted Dice scores are used to rank/select atlases for MABS.
The newly developed approach is applied for segmentation of the left and right parotid glands in CT images of cancer patients. Quantitative evaluation shows that the presented atlas selection approach performs distinctly better than other commonly used selection strategies, especially for a small number of atlases.

Keywords: Atlas-based segmentation, Atlas selection, PCA, Machine Learning

1 Introduction

Image segmentation is a frequently applied task in medical imaging, especially in radiotherapy. Automated segmentation is a substantial component of image-guided adaptive radiotherapy (e.g. for segmentation of organs at risk). The utilization of a priori knowledge of structures that should be segmented is the basis to accomplish a reliable and robust automated segmentation result. In case of atlas-based segmentation this knowledge is available through already segmented atlas images. The segmentation of structures in new images is performed by registering these new images to an already segmented atlas image. Especially for subjects with high inter-structure variability, multi atlas-based segmentation (MABS) approaches have shown to be more accurate than single atlas-based segmentation attempts [1, 2]. Apart from the selected

approach for elastic registration, the selection of an appropriate voting strategy and the selection of atlases have an important influence on the segmentation accuracy and reliability. Moreover, restricting registration to an appropriate subset of atlases can significantly improve segmentation speed. The presented approach is used for segmenting parotid glands in head and neck CT scans of cancer patients. Being the largest salivary glands, the parotid glands are a highly critical OAR. Moreover, due to the high anatomical shape variability of parotid glands, low contrast and image noise caused by dental artifacts, automated segmentation of the parotids is very challenging.

1.1 Related Work

Common atlas selection approaches use image similarity metrics for atlas selection. Typical image similarity metrics are sum of squared distances (SSD), cross correlation [3], and normalized mutual information (NMI) [4]. Based on similarity values, a ranking can be built and used for atlas selection. In a recent work, an atlas selection strategy based on the correlation of inter-atlas similarities was proposed [5]. This approach outperformed common NMI-based atlas selection, especially if only a smaller number of atlases was used. Wolz et al. developed an atlas selection approach based on learning an embedding for atlas selection [6]. Their results have shown an increasing gain in accuracy with increasing distance between the new image that should be segmented and the atlas images compared to common multi-atlas segmentation approaches. In a work by Cao et al. images are projected into a low-dimensional manifold [7]. They used Euclidian distance in this low-dimensional space to determine intrinsic image similarity that is used as an atlas selection criterion. Sanroma et al. [2] used a learning based approach to rank atlases. They tried to learn the relationship between the pairwise appearance of analyzed images and the final labeling performance expressed by Dice values.

1.2 Contribution

In this work an atlas selection strategy using Manifold Learning and Random Forest Regression is presented. The Dice score has shown to be a good (virtual) parameter for atlas selection [2, 5]. Hence, in this work the Dice score was chosen as target parameter for prediction through Random Forest Regression. In contrast to the inspiring approach of Sanroma et al. [2], where HOG features are extracted out of the images and used for model learning with Support Vector Machines, in this work features are generated by using Manifold Learning in combination with multi-scale image patches and Random Forest Regression. Multi-scale image patches are used to assess (semi-) local image characteristics. The patches form the input for a Manifold Learning step. In contrast to [2], the signed distances between the embedding coordinates of the ensemble of all patches extracted from a pair of images will be used to learn (training) and predict (testing) their respective Dice scores. The predicted Dice scores are used to rank and select atlases for MABS. To the best of our knowledge there is no atlas selection approach that uses Manifold Learning with multi-scale image patches in combination with PCA or Laplacian Eigenmaps and Random Forests.

2 Methods

In Fig. 1 the principle overview of the developed atlas selection approach can be seen.

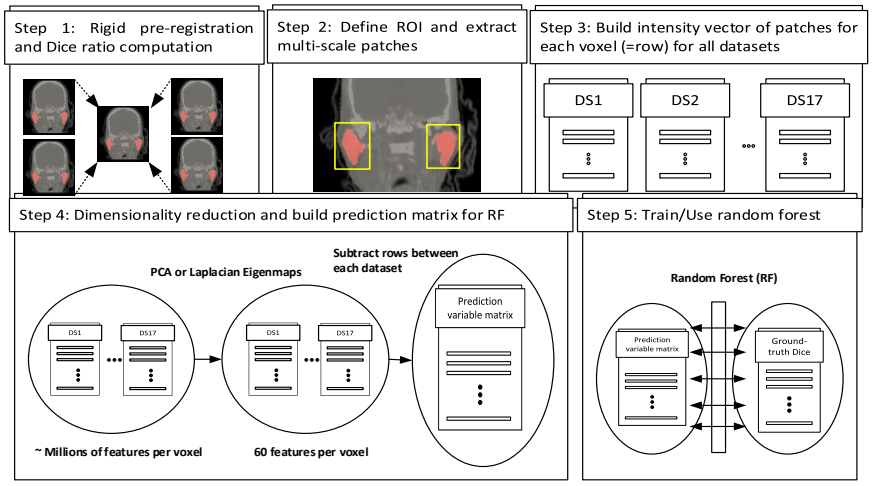


Fig. 1. Principal overview of the proposed atlas selection approach for MABS.

2.1 Data Preprocessing

In the first step all atlas images are rigidly registered on a randomly selected reference atlas image using MI. In order to improve image registration quality, metric computation was restricted to a masked region around left and right parotid gland. Subsequently, regions of interest (ROI) containing all voxels for which patches shall be extracted are defined. The size of the ROIs was constrained to be equal for atlases and target images. Dice ratio, which will be used for atlas ranking (section 2.3) can be calculated as

$$Dice = \frac{2|A \cap B|}{|A| + |B|}. \quad (1)$$

A and B are labeled regions that are compared.

2.2 Multi-scale Patch Extraction

The second step is the extraction of three different image patches for each voxel within the predefined ROI (see Fig. 2). The smallest patch has a cubic structure (patch 1: 3x3x3 voxels). The two remaining patches have a cuboid-like structure: However, only margin voxels that form a hull around the cuboid are used for the larger patch types. These patches are surrounding patch type 1 and have a size of 7x7x3 (patch2)

and 11x11x5 voxels (patch 3). In order to take account of anisotropic image resolution, patch sizes are also anisotropic.

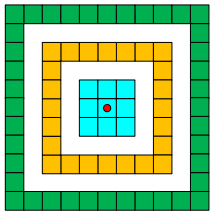


Fig. 2. Principle 2D structure of multi-scale image patches. Red dot: current investigated voxel. Cyan: patch 1. Orange: patch 2. Green: patch 3. Thickness of patches 2 and 3 is one voxel.

By using the border voxels of the cuboid-like patches instead of cubic patches a multi-scale approach can be applied without introducing redundancy compared to using only cubic patches (due to overlapping regions). Furthermore, the number of variables for subsequent dimensionality reduction and Random Forest learning is distinctly lower. This has a significant positive effect on computation time. For each voxel within the predefined ROI the intensity values of its respective patches are arranged in a feature vector $\boldsymbol{t_i}$ for all p atlas datasets for each patch type i. The length of these vectors $\boldsymbol{t_i}$ is m, where m = number of patches per training image x number of voxels per patch type.

2.3 Patch Model Creation

In order to create a patch model, two different approaches have been evaluated. On the one hand, Principal Component Analysis (PCA) [8] for performing linear dimensionality reduction and on the other hand Laplacian Eigenmaps (LapE) [9, 10] providing non-linear embeddings of the input patches. For both approaches, all p feature vectors $\boldsymbol{t_i}$ of patch type i are arranged in three matrices M_i, which are used as input for PCA and LapE.

Principal Component Analysis. PCA aims at identifying the linear combinations of the original variables which maximize their variance. The distribution of data is estimated by the mean value $\boldsymbol{x} = \varepsilon(\boldsymbol{x}) = (\varepsilon(x_1)\varepsilon(x_2) \ldots \varepsilon(x_m))^T \in \mathbb{R}$ and the covariance of the data between the vectors $c_{i,j} = \varepsilon((x_i - m_i)(x_j - m_j))$. Covariances can be represented by the covariance matrix $\boldsymbol{C} = \varepsilon((\boldsymbol{x} - \boldsymbol{m})(\boldsymbol{x} - \boldsymbol{m})^T) \in \mathbb{R}^{mxm}$. From the covariance matrix the eigenvectors and eigenvalues can be calculated. Finally, only the d eigenvectors are chosen that cover a certain (high) degree of all variations. In this project a value of d=20 was chosen.

Manifold Learning. Laplacian Eigenmaps (LapE) are used for non-linear Manifold Learning [9, 10]. A low-dimensional representation of the data is calculated in which the distances between a data point (=patch) and its k nearest neighbors are minimized. LapE generate a graph G based on the neighborhood information. Each data point is represented by a node. The connectivity is determined by a k-nearest neighbor search. Connectivity is weighted by a Gaussian kernel and is stored in the adjacency

matrix $\boldsymbol{W}$. For dimensionality reduction the solution of the minimization problem of the cost function $arg\ min_Y \sum_i \sum_j ||\boldsymbol{y}_i - \boldsymbol{y}_j||^2 w_{i,j}$ can be computed by the eigensystem $\boldsymbol{Lf} = \boldsymbol{\lambda f}$. $\boldsymbol{L}$ is the non-normalized graph Laplacian defined by $\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{W}$. $\boldsymbol{D}$ is the diagonal matrix with $d_{ii} = \sum_i w_{j,i}$. By solving the generalized eigenvalue problem $\boldsymbol{Lf} = \boldsymbol{\lambda DF}$ d eigenvectors f_i corresponding to the d smallest eigenvalues can be calculated. Consequently, each patch can be represented in the respective position in the patch space. In this project, the number of neighbors used for LapE computation was set to 50.

Random Forest Regression. The first d=20 principal components (PCs) or eigenvectors (EVs), respectively, are used. Using the first 20 PCs/EVs for each patch type, 60 features (= coordinates in PCA space and embeddings resulting from LapE respectively) per voxel are used for Random Forest Regression. For this purpose they are again arranged in vector form. Since Dice ratios are based on pairwise comparison of voxel quantities, also pairwise differences of features between all pairs of images of the training set are calculated and stored in the prediction matrix $\boldsymbol{P}$.

Random Forests (RF) are an ensemble learning approach that can be used for classification and regression [11]. RFs are composed of a multitude of decision trees, where each tree t consists of nodes. During training, the prediction matrix $\boldsymbol{P}$ (containing the embedding coordinates for each patch resulting from PCA and LapE) is trained for predicting the Dice value of each image pair within the training set. The set of embedded coordinates of the training sets gets binary split into two parts at each node of a tree and the respective input is then assigned to one of its child nodes. The target variable of RF regression is the Dice ratio dr after rigid pre-alignment. j is a split variable out of the feature space and s is a splitting point. A binary splitting plane can then be defined: $R_1(j,s) = \{X | X_j \leq s\}$ and $R_2(j,s) = \{X | X_j > s\}$. The splitting variable j and splitting point s are determined by

$$\min_{j,s}\{\min_{c_1}(\sum_{p_i \in R_1(j,s)}(dr_i - c_1)^2) + \min_{c_2}(\sum_{p_i \in R_2(j,s)}(dr_i - c_2)^2)\}. \qquad (2)$$

In each step an optimal pair (j,s) is searched.

For new target images the features have to be computed using PCA or LapE. During testing, new embedding coordinates for the patches of an unseen image are evaluated by the trained trees. By this means a prediction for the "virtual" Dice ratio between (not yet existing) parotid glands label in the new image and every image in the training set can be computed.

2.4 Atlas Selection

Based on the predicted Dice ratios resulting from RF regression an atlas ranking is built for atlas-based segmentation. Higher predicted Dice ratios are ranked first.

3 Results

3.1 Evaluation Strategy

In order to evaluate the segmentation results of the proposed method the results were compared with the following atlas selection strategies:

- **Atlas selection based on original Dice values:** Dice values after deformable registration and segmentation were used to build an atlas ranking. Indeed, this is an unrealistic approach, however, for testing purposes this is often used as a reference. In the following this selection strategy will be referred to as "Oracle" selection.
- **Atlas selection based on common image similarity:** Atlas ranking was generated according to NMI values between image regions. NMI-based selection was chosen as a reference since this is one of the most frequently used selection strategies.

Based on the determined rankings, segmentation performance was tested using different numbers of ranked atlases. Previous results of works with the same datasets have shown that using more than 11 atlases deteriorates segmentation results [5]. Hence, the maximum number of atlases that were tested was set to 11.

A leave-one-out strategy was applied in order to evaluate different atlas selection approaches. The Dice ratio and the 95% Hausdorff distance were used to quantitatively evaluate segmentation accuracy for a varying numbers of atlases

In order to test the significance of the improvement of the respective atlas selection strategy with respect to other atlas selection strategies and different numbers of atlases used for selection, a paired t-test with $p \leq 0.05$ was used [2].

3.2 Data

The atlas database consists of 17 CTs of the head and neck. All images have a voxel size of 1.25x1.25 x 2.5 mm and a dimension of 512 x 512 x 89 slices. Manual delineations of left and right parotid glands were used as gold standard to evaluate segmentation accuracy.

The dimension of the bounding box of the left parotid is 40 x 40 x 35 (56,000 voxels) and of right parotid it is 45 x 50 x 35 (78,750 voxels).

3.3 Experiments

In Fig. 3 (first row) the average Dice ratios for the segmentation of the left and right parotid with respect to different numbers of atlases are visualized. Colored bars represent average Dice ratio for different atlas selection strategies. In Fig. 3 (second row) the average 95% Hausdorff distances of the left and right parotid gland with respect to different numbers of atlases are depicted. Black lines indicate 25^{th} and 75^{th} percentile.

Significance tests performed in order to reveal significant accuracy improvement (in terms of higher Dice scores) for every newly added atlas showed, that no significant improvement could be obtained when using more than 5 datasets. The underlying

atlas ranking for this significance tests is based on Oracle selection, which is considered as "virtual" (i.e. practically not available) gold standard for atlas ranking.

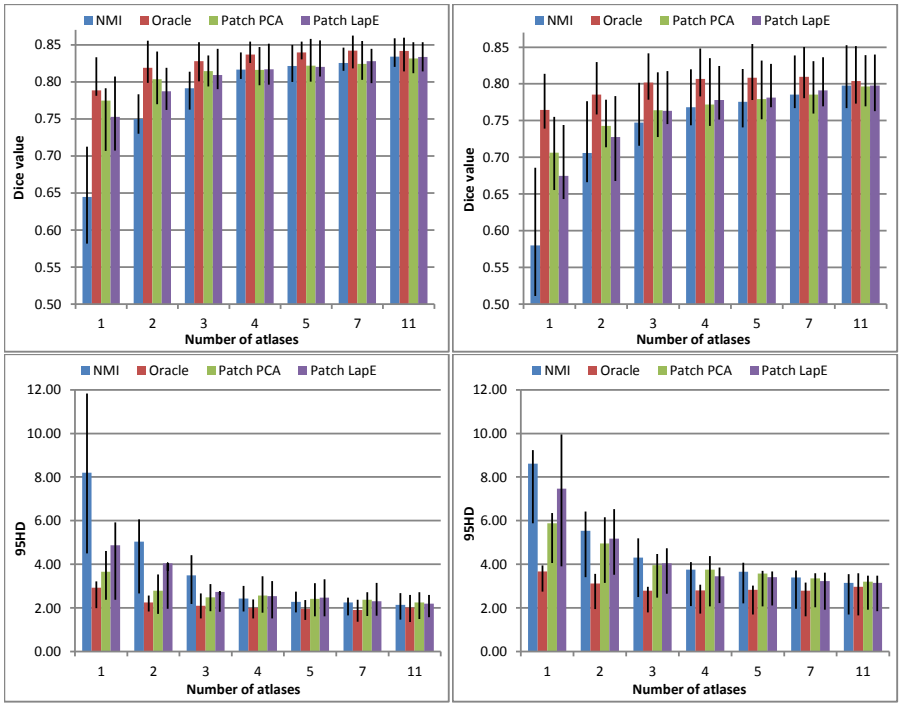


Fig. 3. Evaluation of segmentation performance: First row: Colored bars represent average Dice ratio of the left (first column) and right (second column) parotid gland. Second row: Colored bars represent average 95% Hausdorff distance in mm of the left (first column) and right (second column) parotid. Black lines indicate 25th and 75th percentile.

4 Discussion

Significance tests revealed that using more than 5 atlases does not lead to a significant improvement of the segmentation accuracy even when using Oracle selection. Moreover, evaluation has shown that when using less than 4 atlases the segmentation accuracy is significantly higher when using PCA- and LapE-based atlas selection compared to NMI-based selection (first row in Fig. 3). Furthermore, for the left parotid there was no significant accuracy improvement using 3 atlases based on PCA-based ranking and 5 atlases using NMI-based ranking. The same is true for the right parotid gland when using 4 atlases based on PCA-based ranking. This confirms the positive property of the newly developed approach that equally accurate segmentation results can be achieved with a comparably lower number of atlases.

Comparing PCA vs. LapE of the presented approach, it can be said that PCA-based ranking performs better than the LapE-based ranking when using less than 4 atlases (first row in Fig. 3). For higher number of atlases there are no significant differences concerning segmentation accuracy.

5 Conclusion

It could be shown, that the developed atlas selection technique performs clearly better than NMI-based selection if a low number of atlases is used. In addition, equal segmentation accuracy can be obtained with a lower number of atlases compared to NMI-based selection. Overall, PCA-based atlas selection performed slightly better than LapE-based selection.

In further analyses, the presented approach will be compared to additional selection strategies in addition to NMI. Moreover, more organs in the head neck area will be included in future evaluations. It has to be said, however, that the parotid gland is a very suitable and highly non-trivial test structure, which is frequently used for the evaluation of segmentation approaches for the head and neck region.

6 References

1. van Rikxoort EM, Isgum I, Arzhaeva Y, Staring M, Klein S, Viergever, MA, Pluim, JPW and van Ginneken, B: *Adaptive local multi-atlas segmentation: Application to the heart and the caudate nucleus.* Medical Image Analysis, Vol. 14, No. 1, pp. 39-49, (2010).
2. Sanroma G, Wu G, Gao Y and Shen D: *Learning to rank atlases for multiple-atlas segmentation.* IEEE Trans Med Imaging, Vol. 33, No. 11, pp. 1939-1953, (2014).
3. Aljabar, P., Heckemann, R.A., Hammers, A., Hajnal, J.V. and Rueckert, D.: *Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy.* NeuroImage. Vol. 46, No. 3, pp. 726-738, (2009).
4. Studholme C, Hill DLG and Hawkes DJ. *An overlap invariant entropy measure of 3D medical image alignment.* Pattern Recognition. Vol. 32, pp. 71-86, (1999).
5. Raudaschl P, Fritscher K, Zaffino P, Sharp GC, Spadea MF and Schubert R. *A novel atlas-selection approach for multi-atlas based segmentation using the correlation of inter-atlas similarities.* MICCAI - Image-guided adaptive radiotherapy (IGART) workshop, MICCAI 2014, Boston, (2014).
6. Wolz R, Aljabar P, Hajnal JV, Hammers A and Rueckert D. *LEAP: Learning embeddings for atlas propagation.* NeuroImage, Vol. 49, pp. 1316-1325, (2010).
7. Cao Y, Yuan Y Li X, Turkbey, Choyke PL and Yan P. *Segmenting Images by Combining Atlases on Manifold.* MICCAI 2011, pp.272-279, (2011).
8. Rencher AC. Methods of Multivariate Analysis – Principal Component Analysis. John Wiley & Sons, ISBN: 9780471418894, chapter. 12, (2003).
9. Belkin M, Niyogo P. *Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering.* Advances in Neural Information Processing Systems 14, pp. 585-591, (2001).
10. Gerber S, Tasdizen T and Whitaker Ross. *Robust Non-linear Dimensionality Reduction using Successive 1-Dimensional Laplacian Eigenmaps.* Proc. of 24th International Conference on Machine Learning, Corvallis, OR, (2007).
11. Breiman L. *Random Forests.* Machine Learning, Vol. 45, No. 1, pp. 5-32, (2001).

Groupwise Registration for Robust Motion Field Estimation in Artifact-Affected 4D CT Images

Alexander Tack[1], Yuske Kobayashi[1], Tobias Gauer[2], Alexander Schlaefer[3], and René Werner[1]

[1] Department of Computational Neuroscience, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany
[2] Department of Radiotherapy and Radio-Oncology, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany
[3] Institute of Medical Technology, Hamburg University of Technology, 21073 Hamburg, Germany
{a.tack,r.werner}@uke.de

Abstract. Precise voxel trajectory estimation in 4D CT images is a prerequisite for reliable dose accumulation during 4D treatment planning. 4D CT image data is, however, often affected by motion artifacts and applying standard pairwise registration to such data sets bears the risk of aligning anatomical structures to artifacts – with physiologically unrealistic trajectories being the consequence. In this work, the potential of a novel non-linear hybrid intensity- and feature-based groupwise registration method for robust motion field estimation in artifact-affected 4D CT image data is investigated. The overall registration performance is evaluated on the DIR-lab datasets; Its robustness if applied to artifact-affected data sets is analyzed using clinically acquired data sets with and without artifacts. The proposed registration approach achieves an accuracy comparable to the state-of-the-art (subvoxel accuracy), but smoother voxel trajectories compared to pairwise registration. Even more important: it maintained accuracy and trajectory smoothness in the presence of image artifacts – in contrast to standard pairwise registration, which yields higher landmark-based registration errors and a loss of trajectory smoothness when applied to artifact-affected data sets.

1 Introduction

During 4D radiotherapy treatment planning of moving tumors, 4D CT- and registration-based motion estimation is commonly used for 4D dose calculation or dose accumulation purposes, i.e. to estimate the dose applied to tumor and normal tissue under consideration of respiratory motion. The standard approach to estimate motion fields and voxel trajectories in 4D CT images is to sequentially apply non-linear pairwise registration between two of the 4D CT image volumes; concatenation of the estimated motion fields forms the sought voxel trajectories. 4D CT image data is, however, often affected by motion artifacts [11] and – since information from other volumes is neglected – pairwise registration tends to align anatomical structures to artifacts in artifact-affected image frames and areas [8].

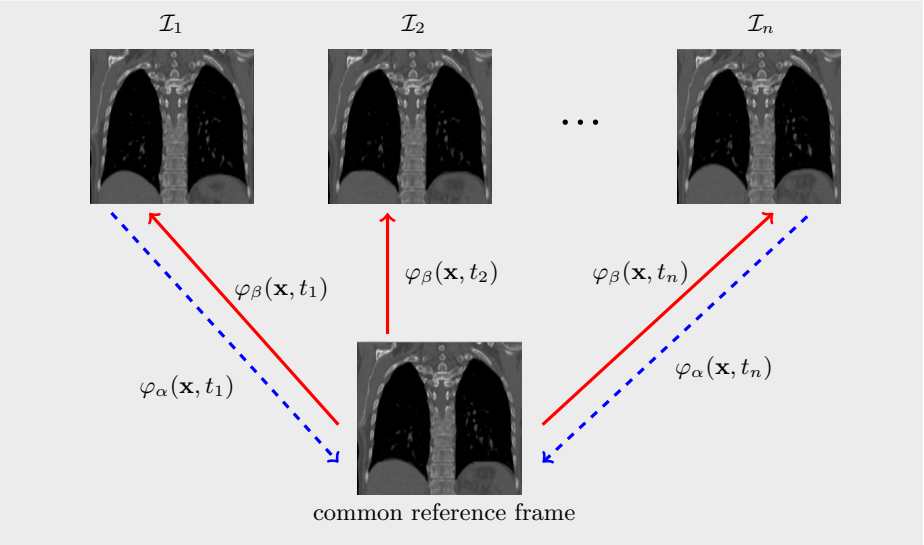


Fig. 1. Illustation of the 4D transformations φ_α and φ_β between a set of images $\mathcal{I}_i$, $i = 1, \ldots, n$, and the common reference frame.

Temporal regularization and groupwise registration should have the potential to mitigate this problem and to compute more realistic voxel trajectories in artifact-affected data sets.
In this paper, we present a novel hybrid intensity- and feature-based spatially and temporally regularized groupwise registration method, and the aforementioned hypothesis is analyzed both by means of motion phantom data and clinical 4D CT images.

2 Methods

The proposed non-linear groupwise registration approach is based on Metz *et al.* [4] and additionally inspired by essential components of the registration method of Xu and Li [10]. The resulting algorithm is implemented within the open-source registration framework elastix (http://elastix.isi.uu.nl).
Let the four-dimensional input image $\mathcal{I}(\mathbf{y})$ consist of voxels with spatio-temporal coordinates $\mathbf{y} = (\mathbf{x}^T, t)^T \in \mathbb{R}^3 \times \mathbb{R}$. The sought transformations from image $\mathcal{I}_r$ to image $\mathcal{I}_t$ (here: $r, t \in \mathcal{T}$ as breathing phases) is obtained by concatenating 4D B-spline transformations $\varphi_\alpha(\mathbf{y})$ and $\varphi_\beta(\mathbf{y})$ as illustrated in Fig. 1. Following [3], 4D B-spline transformations can be parameterized as

$$\varphi_\mu(\mathbf{y}) = \mathbf{y} + \sum_{\mathbf{y}_i \in \mathcal{N}_\mathbf{y}} \boldsymbol{c}_i \beta^k\left(\frac{\mathbf{y} - \mathbf{y}_i}{\delta_i}\right), \tag{1}$$

with $\mathbf{y}_i \in \mathcal{N}_\mathbf{y}$ as control points within the compact support of the B-spline at y, δ_i as control point spacing, with the k-th order (here: cubic) multidimensional

B-spline polynomial $\beta^k(\mathbf{y})$ and the B-spline coefficient vectors $\boldsymbol{c}_i$. The parameter vector μ is optimized during registration and is associated to the control point coefficients [4, 6].
To obtain φ_α and φ_β of Fig. 1, two steps are performed: First, the transformation φ_α is calculated by minimizing the cost function $\mathcal{D}$. $\mathcal{D}$ is in this work given by

$$\mathcal{D} = \lambda\mathcal{D}_M + \rho\mathcal{D}_F + \zeta\mathcal{D}_S, \tag{2}$$

with the weighting coefficients λ, ρ, $\zeta \in \mathbb{R}_+$ and the distance or regularization terms $\mathcal{D}_M$, $\mathcal{D}_F$, and $\mathcal{D}_S$.
The first term in (2), the distance measure $\mathcal{D}_M$, is defined as

$$\mathcal{D}_M(\alpha) = \frac{1}{|\mathcal{S}||\mathcal{T}|}\sum_{\mathbf{x}\in\mathcal{S}}\sum_{t\in\mathcal{T}}(\mathcal{I}(\varphi_\alpha(\mathbf{x},t)) - \bar{\mathcal{I}}_\alpha(\mathbf{x}))^2, \tag{3}$$

with

$$\bar{\mathcal{I}}_\alpha(\mathbf{x}) = \frac{1}{|\mathcal{T}|}\sum_{t\in\mathcal{T}}\mathcal{I}(\varphi_\alpha(\mathbf{x},t)) \tag{4}$$

as the common reference frame of Fig. 1; $\mathcal{S}$ and $\mathcal{T}$ denote the sets of spatial and temporal voxel coordinates, respectively.
Similar to [10], a feature term $\mathcal{D}_F$ is integrated into $\mathcal{D}$ and defined as

$$\mathcal{D}_F = \frac{1}{n|\mathcal{T}|}\sum_{t\in\mathcal{T}}\sum_{i=0}^{n}||\varphi_\alpha(p_{it},t) - \bar{\varphi}_\alpha(p_i)||_2\,. \tag{5}$$

p_{it} denotes the spatial position of the i-th feature point or landmark at breathing phase t. The mean temporal coordinate of the i-th feature point is calculated by

$$\bar{\varphi}_\alpha(p_i) = \frac{1}{\mathcal{T}}\sum_{t\in\mathcal{T}}\varphi_\alpha(p_{ij},t). \tag{6}$$

Analogous to [5], the landmarks p_{it} are identified using a two-step algorithm: First, feature points are detected in end-inspiration CT image based on the answers of the Förstner operator applied to the individual voxels. Then, the feature points are transferred to the other 4D CT image frames by intensity-based block matching and subsequent plausibility checks (see [7] for details).
The final curvature regularization term $\mathcal{D}_S$ is defined by means of the Hessian $\mathcal{H}_{\varphi_\alpha}$ of φ_α and given by

$$\mathcal{D}_S = \frac{1}{|\mathcal{S}||\mathcal{T}|}\sum_{\mathbf{x}\in\mathcal{S}}\sum_{t\in\mathcal{T}}||\mathcal{H}_{\varphi_\alpha}(\mathbf{y})||_F^2. \tag{7}$$

In the second registration step and for a given parameter vector α, the inverse transformation of φ_α, φ_β, is approximated by minimizing the inverse consistency term

$$\mathcal{D}_{IC} = \frac{1}{|\mathcal{S}||\mathcal{T}|}\sum_{\mathbf{y}\in\mathcal{S}\times\mathcal{T}}||\varphi_\beta(\varphi_\alpha(\mathbf{y})) - \mathbf{y}||_2^2. \tag{8}$$

The transformations φ_α and φ_β can be combined to a transformation $\varphi_{\bar{\alpha}}^{rt}$ from a reference image I_r to a template image I_t.

Table 1. Published target registration errors for groupwise registration algorithms and the DIR-lab datasets, compared to the pairwise and the proposed groupwise registration approaches PW and GW applied in this work

	Target Registration Error (TRE) [mm]						
Dataset	**Castillo [1]**	**Delmon [2]**	**Wu [9]**	**Metz [4]**	**Xu&Li [10]**	***PW***	***GW***
DIR-lab 01	0.97	1.3	0.64	1.02	1.28	0.88	0.80
DIR-lab 02	0.86	1.0	0.56	1.06	0.56	0.92	0.76
DIR-lab 03	1.01	1.7	0.70	1.19	0.59	1.06	1.02
DIR-lab 04	1.40	1.6	0.91	1.57	0.69	1.49	1.39
DIR-lab 05	1.67	1.9	1.10	1.70	1.10	1.63	1.57
DIR-lab 06	1.58	1.6	3.28	—	—	1.47	1.48
DIR-lab 07	1.46	1.7	1.68	—	—	1.51	1.59
DIR-lab 08	1.77	1.8	1.70	—	—	1.63	1.73
DIR-lab 09	1.19	1.5	1.72	—	—	1.31	1.27
DIR-lab 10	1.59	1.6	1.48	—	—	1.34	1.42
ØTRE	**1.35**	**1.57**	**1.38**	**1.31**	**0.84**	**1.32**	**1.30**

3 Experiments and Results

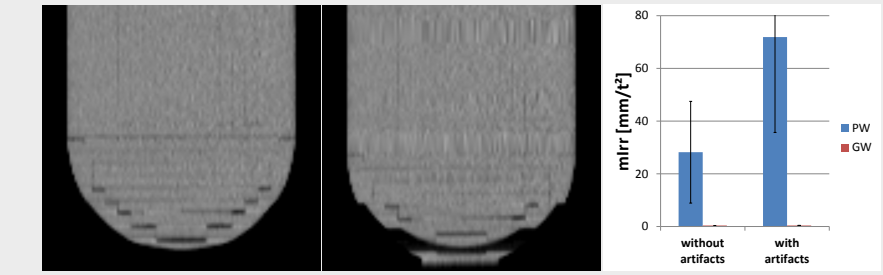
The defined functionals were minimized using adaptive stochastic gradient descent within a multi-resolution setting (four levels). The control point spacing was chosen to 13 and 10 mm for registration steps one and two. The applied weighting coefficients were $\lambda = 0.2$, $\rho = 0.5$ and $\zeta = 0.2$.
For evaluation purposes, we used the freely available DIR-lab 4D CT data sets, acquired motion phantom images, and worked with 4D CT image data acquired in our facility (only expiration phases included into the experiments).

3.1 DIR-lab datasets

The DIR-lab 4D CT data sets (http://www.dir-lab.com) were applied for registration accuracy evaluation of the proposed groupwise registration scheme (GW). The results are summarized in Table 1 and refer to target registration errors (TRE) between end inspiration and expiration, evaluated by means of the 300 landmarks provided by the DIR-lab. Similar to the following experiments, we compared our GW results to a publicly available B-spline-based pairwise registration approach PW that is conceptually close to GW [4]. For both approaches, the TRE is in the order of 1.3 mm (differences not significant).

3.2 Motion phantom datasets

The motion phantom consists of a QA phantom on a motion stage. Images were acquired during regular phantom motion (Siemens SOMATOM AS Open, phase-based reconstruction). To evaluate the potential of GW for robust registration in artifact-affected data, we reconstructed two data sets: (1) almost artifact-free by standard reconstruction, and (2) a data set containing a single



(a) w/o artifacts (b) with artifacts (c) Trajectory irregularity

Fig. 2. Coronal view of the phantom for (a) standard reconstruction and (b) with enforced motion artifacts. (c): irregularity mIrr of the voxel trajectories, evaluated for pairwise (PW) and groupwise registration (GW) and phantom datasets with and without motion artifacts.

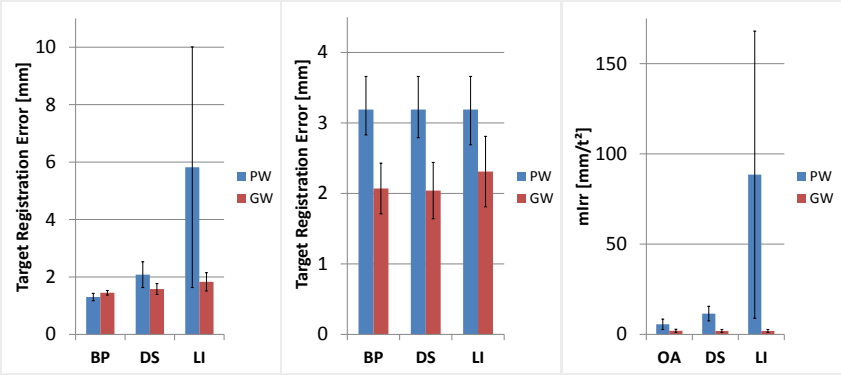
artifact-affected image frame (here: 40% expiration phase T20; artifacts enforced by manipulating the phase-sorting process, as shown in Fig. 2). GW and PW were applied for voxel trajectory estimation in the artifact-free and -affected data sets. Figure 2 (c) illustrates that the irregularity mIrr of the trajectories ($\text{mIrr} = \frac{1}{T|P_1|}\sum_t^T \sum_{\mathbf{x}\in\Omega} \left|\left|\frac{\partial^2 \varphi_\mu^{rt}(\mathbf{x})}{\partial t^2}\right|\right|_2^2$, cf. [4]) is far higher for PW than for GW – especially for the artifact-affected data set.

3.3 Clinical data

Similar to the phantom setting, five patient 4D CT data sets without significant artifacts after standard reconstruction (BP = best possible reconstruction) were selected for evaluation purposes. By specific manipulation of the phase sorting process we deliberately induced double structure (DS) and missing data/linear interpolation (LI) artifacts in the diaphragm area during reconstruction. We ended up with three 4D CT images for each patient: the original one (BP), an LI dataset containing LI artifacts in the 40% expiration phase T20, and an DS dataset with DS artifacts in T20.
Landmark-based TRE values for PW and GW registration of the end inspiration phase T00 to T20 and T00 to end expiration (T50) as well as mIrr values are summarized in Fig. 3. Corresponding motion fields (Fig. 4) and difference images before and after registration (Fig. 5) visually confirm the result that GW improves robustness of motion estimation in artifact-affected 4D CT images.

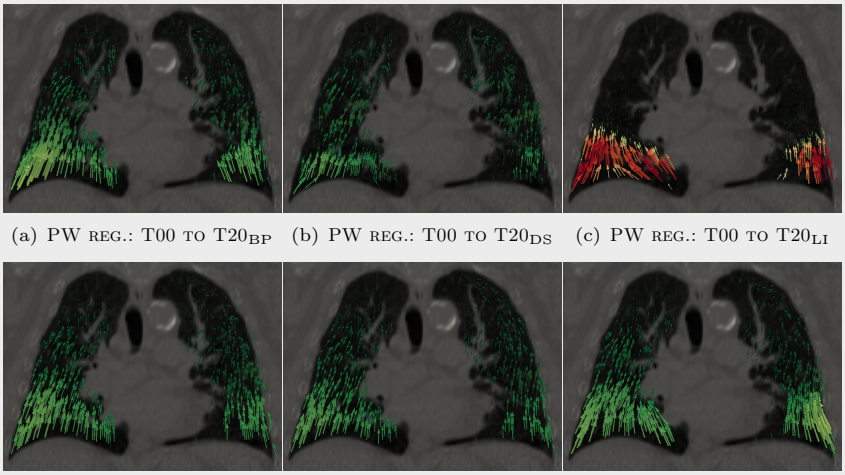
4 Discussion and Conclusions

In this work, a hybrid intensity- and feature-based spatially and temporally regularized groupwise registration approach has been proposed and its potential for



(a) TRE T00 to T20 (b) TRE T00 to T50 (c) Trajectory irregularity

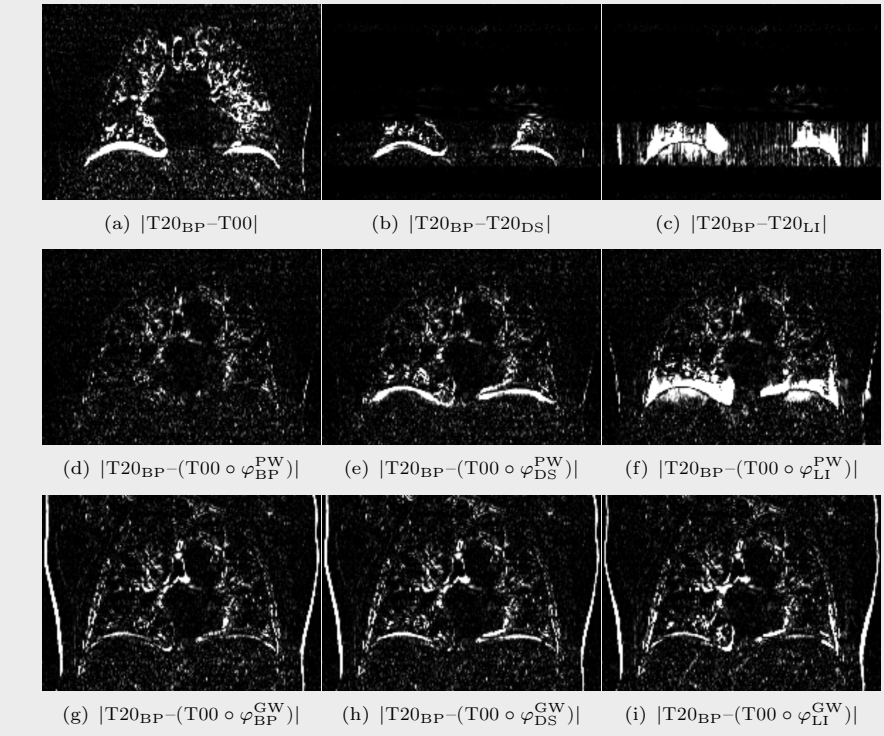
Fig. 3. TRE and mIrr data with std. dev. for pairwise (PW) and groupwise (GW) registration in patient 4D CT images (BP: standard reconstruction; LI: BP data set with enforced linear interpolation artifacts in T20; DS: double structure artifacts in T20).



(a) PW reg.: T00 to $T20_{BP}$ (b) PW reg.: T00 to $T20_{DS}$ (c) PW reg.: T00 to $T20_{LI}$

(d) GW reg.: T00 to $T20_{BP}$ (e) GW reg.: T00 to $T20_{DS}$ (f) GW reg.: T00 to $T20_{LI}$

Fig. 4. Motion fields from end inspiration phase T00 to T20. The amount of motion is visualized in a spectrum from green (approx. 10 mm) to red (approx. 40 mm). Top: PW registration of T00 to T20 in the (a) BP, (b) DS, and (c) LI data set. Note that only the T20 phases are manipulated, i.e. T00 is similar for all data sets. Especially the influence of the LI artifacts on the motion fields becomes obvious. Bottom: similar motion fields for GW registration; the influence of the artifacts visibly decreases compared to PW.



(a) $|\mathrm{T20_{BP}}–\mathrm{T00}|$ (b) $|\mathrm{T20_{BP}}–\mathrm{T20_{DS}}|$ (c) $|\mathrm{T20_{BP}}–\mathrm{T20_{LI}}|$

(d) $|\mathrm{T20_{BP}}–(\mathrm{T00}\circ\varphi^{\mathrm{PW}}_{\mathrm{BP}})|$ (e) $|\mathrm{T20_{BP}}–(\mathrm{T00}\circ\varphi^{\mathrm{PW}}_{\mathrm{DS}})|$ (f) $|\mathrm{T20_{BP}}–(\mathrm{T00}\circ\varphi^{\mathrm{PW}}_{\mathrm{LI}})|$

(g) $|\mathrm{T20_{BP}}–(\mathrm{T00}\circ\varphi^{\mathrm{GW}}_{\mathrm{BP}})|$ (h) $|\mathrm{T20_{BP}}–(\mathrm{T00}\circ\varphi^{\mathrm{GW}}_{\mathrm{DS}})|$ (i) $|\mathrm{T20_{BP}}–(\mathrm{T00}\circ\varphi^{\mathrm{GW}}_{\mathrm{LI}})|$

Fig. 5. Top row: (a) difference image between phase T00 and $\mathrm{T20_{BP}}$ before registration; (b) difference between the original phase $\mathrm{T20_{BP}}$ and the corresponding image $\mathrm{T20_{DS}}$ with enforced DS artifacts; (c) difference between $\mathrm{T20_{BP}}$ and $\mathrm{T20_{LI}}$. Middle row: difference between the original $\mathrm{T20_{BP}}$ data set and the T00 data set warped by the transformation φ^{PW} obtained by pairwise registration in the (d) original 4D CT data set, (e) the DS 4D CT data set, and (f) the LI data set. Bottom: similar to the middle row, but transformations φ^{GW} computed by groupwise registration GW.

robust voxel trajectory estimation in artifact-affected 4D CT data sets investigated.

The landmark-based evaluation on the 4D CT DIR-lab datasets demonstrated accuracy in the order of the voxel spacing and comparable performance to a related pairwise registration method. Similar results were obtained for a landmark-based evaluation in in-house acquired clinical data sets (TRE pairwise vs. groupwise registration of T00 to T20: 1.30 vs. 1.45 mm). However, TRE values for registration in artifact-affected data sets increased considerably for pairwise registration (2.08 mm for double structure artifacts and 5.82 mm for linear interpolation artifacts), but maintained almost stable for groupwise registration (1.45 mm vs. 1.58 mm and 1.83 mm, respectively). In addition and for both phantom and patient data sets, the groupwise registration approach has been

shown to result in considerably smoother (in terms of mIrr) and most probably more realistic voxel trajectories than pairwise registration – again especially in artifact-affected data sets.
Both findings – maintaining a registration accuracy in the order of the voxel spacing and providing a high degree of voxel trajectory smoothness even in the presence of motion artifacts – strongly confirms the hypothesis underlying this paper that groupwise registration improves robustness of motion field and voxel trajectory estimation in artifact-affected 4D CT image data.

Acknowledgments Funded by the Forschungsförderungsfond of the University of Hamburg, Medical faculty (FFM).

References

1. Castillo, E., Castillo, R., Martinez, J., Shenoy, M., Guerrero, T.: Four-dimensional deformable image registration using trajectory modeling. Phys. Med. Biol. 55, 305–327 (2010)
2. Delmon, V., Rit, S., Pinho, R., Sarrut, D.: Registration of sliding objects using direction dependent B-splines decomposition. Phys. Med. Biol. 58, 1303–1314 (2013)
3. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.W.: Elastix: A toolbox for intensity-based medical image registration. IEEE Trans. Med. Imag. 18, 196–205 (2010)
4. Metz, C.T., Klein, S., Schaap, M., van Walsum, T., Niessen, W.J.: Nonrigid registration of dynamic medical imaging data using nD+t B-Splines and a groupwise optimization approach. Med. Image Anal. 15, 238–249 (2011)
5. Polzin, T., Rühaak, J., Werner, R., Handels, H., Modersitzki, J. et al.: Lung registration using automatically detected landmarks. Methods Inf. Med. 53, 250–256 (2014)
6. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Non-rigid registration using Free-Form Deformations: Application to breast MR images. IEEE Trans. Med. Imag. 18, 712–721 (1999)
7. Werner, R., Wolf, J.C., Ehrhardt, J., Schmidt-Richberg, A., Handels, H.: Automatische Landmarkendetektion und -übertragung zur Evaluation der Registrierung von thorakalen CT-Daten. In: Bildverarbeitung für die Medizin, pp. 31–35. Springer, Heidelberg (2010)
8. Werner, R. and Gauer, T.: Reference geometry-based detection of (4D-)CT motion artifacts: a feasibility study. In: Proc. SPIE 9413, Medical Imaging 2015: Image Processing, pp. 94130S (2015).
9. Wu, G., Wang, Q., Lian, J., Shen, D.: Estimating the 4D respiratory lung motion by spatiotemporal registration and super-resolution image reconstruction. Med. Phys. 40, 031710 (2013)
10. Xu, H. and Li, X.: A symmetric 4D registration algorithm for respiratory motion modeling. In: MICCAI 2013. LNCS, vol. 8150, pp. 149156. Springer, Heidelberg (2013)
11. Yamamoto, T., Langner, U., Loo Jr, B.W., Shen, J., Keall, P.J.: Retrospective analysis of artifacts in four-dimensional CT images of 50 abdominal and thoracic radiotherapy patients. Int. J. Radiat. Oncol. 72, 1250–1258 (2008)

Pythagorean Mean Images for Efficient Groupwise Registration

Mathias Polfliet[1,2], Wyke Huizinga[3], Stefan Klein[3], Johan de Mey[4], and Jef Vandemeulebroucke[1,2]

[1] Vrije Universiteit Brussel, Dept. of Electronics and Informatics (ETRO), Pleinlaan 2, B-1050 Brussels, Belgium
[2] iMinds, Dept. of Medical IT, Gaston Crommenlaan 8 (box 102), B-9050 Ghent, Belgium
[3] Biomedical Imaging Group Rotterdam, Depts. of Radiology & Medical Informatics, Erasmus MC, Rotterdam, the Netherlands
[4] Department of Radiology, Universitair Ziekenhuis Brussel (UZ Brussel), Laarbeeklaan 101, 1090 Brussels, Belgium

Abstract. Groupwise registration is a powerful technique allowing to simultaneously align multiple images using an unbiased approach. Its need arises from population studies or motion estimation across dynamic sequences. An often used class of efficient groupwise metrics measures similarity as the sum of the pairwise similarities between the images and a template image, commonly chosen to be the arithmetic mean image in the current iteration. However, arithmetic averaging in intensity space limits the applications to closely related modalities, and may produce fuzzy images compromising the performance of the metric. Geometric and harmonic averaging is capable of handling range and scale differences without adding computational complexity. Groupwise similarity metrics based on mutual information and the three Pythagorean means were investigated. Experiments performed on monomodal and multimodal data demonstrated superior performance of geometric and harmonic over arithmetic averaging and the corresponding pairwise registration.

1 Introduction

Groupwise registration has shown to be of interest in several areas of research, such as the unbiased construction of atlases [1] and motion estimation across temporal sequences for radiotherapy planning [2]. Scalability of the approach is an important property for groupwise registration, given the fact that such acquisitions can contain over a hundred images to be registered.

The direct application of entropy-based metrics using joint probability density functions (PDFs) such as mutual information [3] (MI), results in an exponentially increasing probability space. Such metrics suffer from the *curse of dimensionality* with the sparsity in the joint PDF limiting their applicability. Several methods have been proposed to work around this problem. Spiclin *et*

al. [4] proposed a method based on hierarchically subdividing the joint intensity space similar to a tree code, and reported good results for up to ten images, after which sparsity limits the robustness of the method.

Instead of focusing on estimating the probability density functions, others have focused on its parent measure, entropy. Hero *et al.* [5] proposed a method based on entropic graphs, where each node in the graph represents an intensity pair in the joint intensity space and the minimal length needed to span the graph is related to the entropy of the system. The main drawback of this method stems from the lengthy optimization required to find this minimal length.

An interesting class of metrics is based on constructing a template image and measuring similarity as the sum of the pairwise similarities between the images and the template [6, 7]. This approach yields an algorithm with linearly increasing computational complexity with respect to the number of images to be aligned, making it suitable for large groupwise registration problems. As a template image, the arithmetic mean image in the current iteration is commonly used, though several authors have demonstrated the importance of the use of a sharp and more representative template to improve robustness and accuracy of the registration [8, 9].

Indeed, arithmetic averaging over the intensity space tends to lead to fuzzy mean images. In addition, it is less suited in the presence of scale and range differences, often present in multimodal data. In this work we propose two novel multimodal metrics for groupwise registration based on internally computing the geometric and harmonic mean images. The performance of these metrics is evaluated for mono- and multimodal groupwise registration, and compared to results when using the arithmetic mean image.

2 Methods and Materials

2.1 Registration Metrics

Following Bhatia *et al.* [7], pairwise mutual information S_{MI} [3] can be extended to arithmetic average mutual information[5] (AAMI)

$$S_{AAMI}\left(I_i(\boldsymbol{x}), \ldots, I_n(\boldsymbol{x})\right) = \sum_{i=1}^{n} S_{MI}(I_i(\mathcal{T}_i(\boldsymbol{x})), \overline{I}_A(\boldsymbol{x})) \tag{1}$$

where $\boldsymbol{x}$ is the spatial coordinate, $\mathcal{T}_i$ the spatial transformation and I_i the intensity function associated with the i^{th} image, for which we assumed an interpolation scheme. $\overline{I}_A(\boldsymbol{x})$ is the voxel-wise arithmetic mean intensity image defined as

$$\overline{I}_A(\boldsymbol{x}) = \frac{1}{n} \sum_{i=1}^{n} I_i(\mathcal{T}_i(\boldsymbol{x})) \ . \tag{2}$$

[5] Note that we preferred the use of MI instead of normalised MI, as proposed in [7].

The arithmetic mean tends to lead to fuzzy mean images. In addition, it is not suited for handling differences in intensity ranges and scales as found in multimodal data. We propose the use of the geometric and harmonic mean images, given by

$$\overline{I}_G(\boldsymbol{x}) = \sqrt[n]{\prod_{i=1}^{n} I_i'(\mathcal{T}_i(\boldsymbol{x}))} \ , \tag{3}$$

$$\overline{I}_H(\boldsymbol{x}) = \frac{n}{\sum_{i=1}^{n} \frac{1}{I_i'(\mathcal{T}_i(\boldsymbol{x}))}} = \frac{n \prod_{i=1}^{n} I_i'(\mathcal{T}_i(\boldsymbol{x}))}{\sum_{i=1}^{n} \prod_{j=1, j\neq i}^{n} I_j'(\mathcal{T}_j(\boldsymbol{x}))} \ . \tag{4}$$

In equations (3) and (4), I_i' is the modified intensity function to ensure non-negative intensities. The three Pythagorean means become equal if and only if, the samples over which are averaged are identical. In all other cases, the arithmetic mean will be higher than the geometric mean, which in turn will be higher than the harmonic mean. As such, intensities tend to fall-off faster at edges, limiting the blurred region, while overlapping structures are retained. We hypothesize that this higher specificity could lead to better registration accuracy.

Two novel groupwise similarity metrics, using MI as a submetric and based on the geometric average image (GAMI) and harmonic average image (HAMI) were implemented as an extension to the software package `elastix` [10] together with AAMI and will be made available in the future. Partial derivatives with respect to the transformation parameters were computed analytically and determined following the approach of Thévenaz *et al.* [11].

To illustrate the behavior of these three metrics, a simulation was developed in which nine squares were simultaneously rotated with a different speed and one was kept stationary. The rotation speed of each square was a multiple of the rotation speed of the first square, such that after a rotation of 90° for the first square, the initial situation is recovered where all squares overlap and another global minimum in the metric space is found. Two simulations were performed, mimicing monomodal and multimodal data. For the monomodal experiment, we used squares with identical intensity. A single image consisted of an outer square, with an intensity of 0.5, an inner square, with an intensity of 1.0, and the background, with an intensity of 0.0. For the multimodal experiment, we set the intensities of the inner and outer squares randomly between 0.5 and 1.0 for all images separately. Figure 1 shows the metric values for a rotation from 0° to 90° for the slowest rotating image.

It can be seen that the proposed metrics have less local minima and those that are common between all metrics are far less pronounced compared to AAMI.

2.2 Monomodal Experiments and Validation

Monomodal 4DCT data was taken from the POPI and DIR-LAB databases [12, 13], consisting of, respectively, 6 and 10 CT images of the thorax for a total

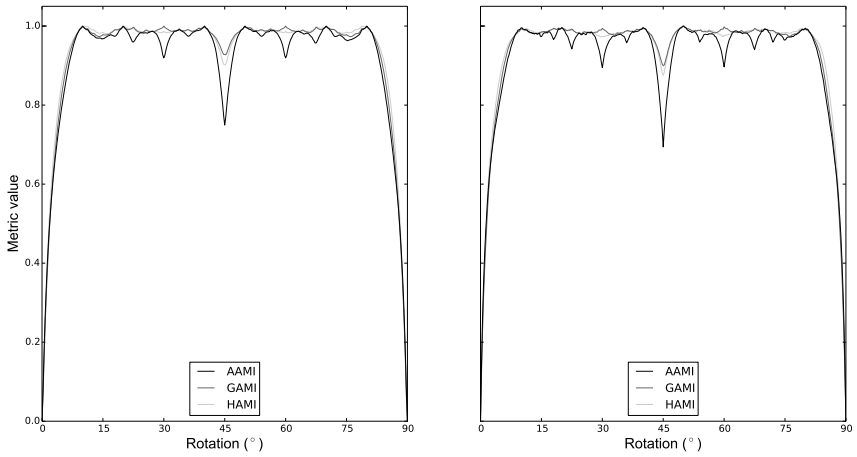


Fig. 1. Behaviour of the three negated similarity metrics under discussion, arithmetic average mutual information (AAMI), geometric average mutual information (GAMI) and harmonic average mutual information (HAMI), when 10 squares are rotated simultaneously in the monomodal case (a) and the multimodal case (b). A lower metric value corresponds to a better match. The angle of the slowest rotating square is given on the x-axis.

of 16 patients. For the POPI database, three patients had 100 manually identified landmarks in the lungs for every breathing phase while the remaining three patients had 100 landmarks in inspiration and expiration phases. The images in the DIR-LAB dataset had 300 landmarks in the lungs for the inspiration and expiration phases and 75 in the four phases in between. Registration accuracy was determined with respect to the inspiration landmarks for all phases in which landmarks were available. We define the groupwise target registration error (gTRE) over a group of images for a single patient as a measure for the accuracy of the registration:

$$gTRE(r) = \frac{1}{n}\sum_{i\neq r}^{n}\frac{1}{|P_i|}\sum_{\boldsymbol{p}_{i,j}\in P_i}||\mathcal{T}_{i,r}(\boldsymbol{p}_{i,j}) - \boldsymbol{p}_{r,j}|| \ . \tag{5}$$

Herein, r is the reference time point, P_i the collection of landmarks in time point i, $\mathcal{T}_{i,r}$ the transformation that maps the coordinates from the i^{th} timepoint to the reference timepoint and $\boldsymbol{p}_{i,j}$ the j^{th} landmark from the i^{th} timepoint.

We performed a deformable registration using cubic B-splines and a final control point spacing of 12 mm. For the optimization four resolutions were used together with an adaptive stochastic gradient descent. Registrations were performed with AAMI as a groupwise metric and MI as a pairwise metric and the results were compared to those obtained from the proposed groupwise metrics GAMI and HAMI. All other registration parameters were kept constant. Pair-

wise registrations were only performed for the phases in which landmarks were available with respect to the inspiratory phase whilst groupwise registrations always included all breathing phases.

2.3 Multimodal Experiments and Validation

Multimodal brain images obtained from the RIRE database [14] were used for the multimodal experiments. Data from 18 patients was included for which at least three of the following modalities were available: CT, PET, MR-T1, MR-T2, MR-proton density (PD). Ground truth transformations were determined through the use of fiducial markers and a stereotactic frame for CT to MR and PET to MR. Four to ten landmarks were available for each patient as a ground truth for the registrations, allowing to compute the gTRE.

Given the large range differences present between CT and PET, it is trivial to see that AAMI will fail to obtain a good and unbiased template image. Prior to the registration, the images were normalized between 0 and 1 to eliminate this effect. The registrations were performed with three translational degrees of freedom prior to rigid registrations which had six degrees of freedom. This allowed for a more robust optimization. The performance of groupwise registration using AAMI, GAMI and HAMI, was compared to pairwise registration using MI. Once again all other registration parameters were kept equal.

3 Results and discussion

3.1 Monomodal Data

The results for the deformable registrations of the 4DCT images of the thorax are pooled for all 16 patients and summarized in Table 1. Significance testing was performed using a two-tailed Wilcoxon signed-rank test given the non-normality of the data. Pairwise registrations were outperformed by groupwise registrations using AAMI ($p = 0.11$), GAMI ($p = 0.016$) and HAMI ($p = 0.0019$), illustrating the added value of a groupwise approach. We obtained better accuracies using GAMI with respect to AAMI in all 16 registrations and the difference was significant ($p = 4.4 \times 10^{-4}$). The HAMI metric gave better results than GAMI in all 16 registrations and the difference was significant ($p = 4.4 \times 10^{-4}$).

The obtained accuracies for pairwise registration using MI corresponded well to other results reported for this data and when performing registration without masks. Delmon *et al.* [15] reported 3.82 ± 4.15 mm for a pairwise registration using MI. The improvement in results is most likely due to lower control point spacing used here (32 mm versus 12 mm) which allowed for finer deformations to be modeled.

The use of geometric and harmonic mean images in the groupwise registration framework has little to no computational overhead compared to the AAMI, and constitutes an elegant alternative to approaches which require separate optimization of the template image [8] and minimal spanning tree [9].

Table 1. Results for the registration of the 4DCT of the lungs for the four compared methods: mutual information (MI), arithmetic average MI (AAMI), geometric average MI (GAMI) and harmonic average MI (HAMI).The values, expressed in mm, correspond to the mean, standard deviation, median and maximal groupwise target registration error.

Method	Mean gTRE	StDev gTRE.	Median gTRE	Max gTRE
MI	2.15	1.03	1.83	4.80
AAMI	2.17	1.10	1.83	5.42
GAMI	2.12	1.06	1.79	5.22
HAMI	2.06	1.01	1.77	5.01

3.2 Multimodal Data

The results for the rigid registrations on the RIRE dataset are shown in Table 2. In accordance with Tomaževič *et al.* [16], we considered the registration successful if the gTRE of a subject was less than the largest voxel spacing of the images under study (8 mm). This lead us to exclude a total of two subjects from the statistical analysis and Table 2 for all metrics to allow for a fair and honest comparison. Two misregistrations were obtained for AAMI (patient 008 and patient 105) and one for GAMI (patient 008).

Overall, groupwise registration using HAMI gave the best results. No misregistrations were recorded and accuracies were significantly better than AAMI ($p = 0.0013$, using a two-tailed Wilcoxon signed-rank test) whilst GAMI also outperfomed AAMI significantly ($p = 0.0037$). GAMI and HAMI did not produce significantly different results compared to pairwise MI.

Results per modality pairing are illustrated in Figure 2. Given the limited data available for some modality pairings no significance testing was performed. It can be seen from the figure that a pairwise approach for the PET-MR registrations is better. It is possible that images with low signal-to-noise ratio, such as PET, disturb the groupwise approach by clouding the mean image and wiping away some of the details present therein.

Table 2. Results for the registration of the multimodal brain images for the four compared methods: mutual information (MI), arithmetic average MI (AAMI), geometric average MI (GAMI) and harmonic average MI (HAMI). The values, expressed in mm, correspond to the mean, standard deviation, median and maximal groupwise target registration error.

Method	Mean gTRE	StDev gTRE.	Median gTRE	Max gTRE
MI	2.31	0.80	2.05	4.50
AAMI	4.19	1.91	3.46	7.62
GAMI	2.60	1.00	2.45	5.43
HAMI	2.35	0.66	2.39	3.94

Our results for pairwise registration were worse compared to those of Thévenaz *et al.* [11] (reported as the median TRE per modality pair), even though a similar implementation is followed for MI. This discrepancy might be explained by differences in multiresolution strategy and optimization method, or other detailed settings such as the number of histogram bins, interpolation method, etc.

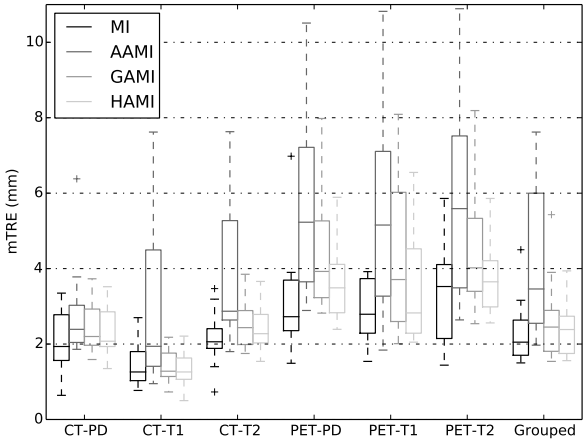


Fig. 2. Boxplots (using John Tukey's definition) of the gTREs of the 16 registrations per modality pair and grouped over all modalities. The compared metrics were mutual information (MI), arithmetic average mutual information (AAMI), geometric average mutual information (GAMI) and harmonic average mutual information (HAMI).

4 Conclusion

In this work we presented an extension to the average image mutual information through the use of the different Pythagorean mean images. The proposed metrics handle scale and range differences better and have an improved optimization behavior without adding computational complexity.

The geometric and harmonic average mutual information proved to be superior to the commonly used arithmetic average mutual information and the difference was significant for mono- and multimodal experiments.

Acknowledgements

The research of S. Klein and W. Huizinga is supported by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 601055, VPH-DARE@IT.

References

1. Bhatia, K.K., Hajnal, J.V., Puri, B.K., Edwards, A.D., Rueckert, D.: Consistent groupwise non-rigid registration for atlas construction. In: Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on, IEEE (2004) 908–911
2. Schreibmann, E., Thorndyke, B., Li, T., Wang, J., Xing, L.: Four-dimensional image registration for image-guided radiotherapy. International Journal of Radiation Oncology* Biology* Physics **71**(2) (2008) 578–586
3. Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., Marchal, G.: Automated multi-modality image registration based on information theory. In: Information processing in medical imaging. Volume 3. (1995) 263–274
4. Spiclin, Z., Likar, B., Pernus, F.: Groupwise registration of multimodal images by an efficient joint entropy minimization scheme. Image Processing, IEEE Transactions on **21**(5) (2012) 2546–2558
5. Hero, A., Ma, B., Michel, O.J., Gorman, J.: Applications of entropic spanning graphs. Signal Processing Magazine, IEEE **19**(5) (2002) 85–95
6. Joshi, S., Davis, B., Jomier, M., Gerig, G.: Unbiased diffeomorphic atlas construction for computational anatomy. NeuroImage **23** (2004) S151–S160
7. Bhatia, K.K., Hajnal, J., Hammers, A., Rueckert, D.: Similarity metrics for groupwise non-rigid registration. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007. Springer (2007) 544–552
8. Fletcher, P.T., Venkatasubramanian, S., Joshi, S.: The geometric median on riemannian manifolds with application to robust atlas estimation. NeuroImage **45**(1) (2009) S143–S152
9. Wu, G., Jia, H., Wang, Q., Shen, D.: Sharpmean: Groupwise registration guided by sharp mean image and tree-based registration. NeuroImage **56**(4) (2011) 1968–1981
10. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.: Elastix: a toolbox for intensity-based medical image registration. Medical Imaging, IEEE Transactions on **29**(1) (2010) 196–205
11. Thévenaz, P., Unser, M.: Optimization of mutual information for multiresolution image registration. Image Processing, IEEE Transactions on **9**(12) (2000) 2083–2099
12. Vandemeulebroucke, J., Rit, S., Kybic, J., Clarysse, P., Sarrut, D.: Spatiotemporal motion estimation for respiratory-correlated imaging of the lungs. Medical physics **38**(1) (2011) 166–178
13. Castillo, R., Castillo, E., Guerra, R., Johnson, V.E., McPhail, T., Garg, A.K., Guerrero, T.: A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. Physics in medicine and biology **54**(7) (2009) 1849
14. West, J., Fitzpatrick, J.M., Wang, M.Y., Dawant, B.M., Maurer Jr, C.R., Kessler, R.M., Maciunas, R.J., Barillot, C., Lemoine, D., Collignon, A., et al.: Comparison and evaluation of retrospective intermodality brain image registration techniques. Journal of computer assisted tomography **21**(4) (1997) 554–568
15. Delmon, V., Rit, S., Pinho, R., Sarrut, D., Delmon, V.: Direction dependent b-splines decomposition for the registration of sliding objects. In: Proceedings of the Fourth International Workshop on Pulmonary Image Analysis. (2011) 45–55
16. Tomaževič, D., Likar, B., Pernuš, F.: Multi-feature mutual information image registration. Image Analysis & Stereology **31**(1) (2012) 43–53

Feature Based Atlas Selection Strategy for Segmentation of Organs at Risk in Head and Neck District

Paolo Zaffino[1], Davide Limardi[1], Salvatore Scaramuzzino[1], Daniela Alterio[2], Federico Javier Diaz[3], Sabrina Vigorito[2], Delia Ciardo[2], Rosalinda Ricotti[2], Barbara Alicja Jereczek-Fossa[2,4] Patrik Raudaschl[5], Karl Fritscher[5], Gregory C. Sharp[6], and Maria Francesca Spadea[1]

[1] Department of Experimental and Clinical Medicine, ImagEngLab, Magna Graecia University of Catanzaro, Italy
[2] Department of Radiation Oncology, European Institute of Oncology, Milano, Italy
[3] Mevaterapia, Medical Radiation Oncology, Buenos Aires, Argentina
[4] Department of Health Sciences, Universitá degli Studi di Milano, Milano, Italy
[5] Department for Biomedical Image Analysis, UMIT, Austria
[6] Department of Radiation Oncology, Massachusetts General Hospital, Harvard Medical School, USA

Abstract. The aim of this work is to test a novel methodology for atlas selection for Multi Atlas Based Segmentation. A feature-based method, relying on histogram of oriented gradients and support vector machine, was used to predict best atlas selection. This was compared to *i*) optimal selection, named oracle, identified by a-priori knowledge and to *ii*) Normalized Mutual information (NMI) ranking, currently used by most methodologies. Final contours were computed by using gaussian- and STAPLE-based voting algorithms. The algorithm was trained and tested on 55 and 20 CT images respectively, where manual contours were drawn for mandible, brainstem and parotid glands. Feature based ranking resulted closer to oracle than NMI, while segmentation accuracy was comparable with NMI strategy. Inaccuracies were mainly caused by label fusion, thus highlighting the voting importance. In conclusion, results proved the effectiveness of the methodology and suggested to include the voting effect into the prediction workflow.

Keywords: Multi altas based segmentation, atlas selection, machine learning, head and neck radiotherapy

1 Introduction

The automatic identification of anatomical structures on medical images is one of the most challenging task in image guided radiotherapy. The most important

advantages of automatic contouring include planning time and the inter- and intra-rater variability reduction. Multi Atlas Based Segmentation (MABS) [1] is a powerful strategy to segment, in a completely unsupervised manner, a query subject. MABS algorithm accomplishes this task by taking advantage of image registration technique and of a database of formerly contoured images that are called *atlases*. To reach the aim, each atlas is non-rigidly registered onto the subject of interest and the contours are remapped accordingly. For each structure, n possible contours will be available and the final label will be computed by means of a statistical algorithm (this last step is generally called *voting* or *label fusion*). Recent studies have shown that the composition of the atlas database plays an important role for the accuracy and the reliability of the entire process [2]. On one side, it is important to have enough cases to represent patient variability; on the other side, atlases that are anatomically very different from the query subject might generate noise and decrease the accuracy. The best approach is to identify, in a large database, an optimal subset of atlases for the subject of interest, thus excluding cases that can introduce redundant and/or misleading information. Besides improving the output accuracy, a further advantage is the decrease of the total segmentation time. Several selection methodologies have been proposed, ranging from very basic (selection criteria based on metadata as age, sex, pathological condition) to the most elaborate ones (based on image similarity computation, as Normalized Mutual Information, NMI [3]). In this work we investigated a new approach for atlas selection. It relies on image feature detection and machine learning techniques in order to identify the best atlases to segment the query subject. Our implementation was inspired by the Sanroma's et al. work [4], who proposed their method in the field of brain Magnetic Resonance Imaging (MRI). We adapted their approach in the context of Head and Neck (HN) radiotherapy. The rational behind the method is to predict, before deformable registration, the contribution that each single atlas will give to the labeling. This contribution is measured in terms of Dice Similarity Coefficient (DSC) [5] between the atlas and the query subject contours and it is learned by the algorithm during the training phase. When the algorithm is run on line, the selection of the optimal atlas subset will be performed on the basis of the best predicted DSC.

2 Materials and Methods

2.1 Data

In this study, 75 simulation CT volumes of HN cancer patients were retrospectively used. All patients were treated at the Department of Radiation Oncology of the European Institute of Oncology, Milan, Italy. The following research was performed within the general notification of the head and neck cancer studies to the Ethics Committee of the European Institute of Oncology, Milan, Italy (IEO

N94/11). All patients gave consent for treatment and the use of their anonymized clinical/imaging data for research. Mandible, brainstem and parotid glands of each subject were manually contoured by an expert physician. The dataset was randomly split in two parts: 55 patients were used to build the atlas database, while the remaining 20 were used as testing subjects. All the cases were rigidly aligned on a reference atlas, which was randomly selected.

2.2 Methods

The method workflow can be divided into two sections: feature extraction and atlas selection (based on machine learning algorithm).

Feature Extraction Histogram of Oriented Gradient (HOG) features [6] was used to describe the patients morphology in each image. HOG identifies image characteristics by counting the occurrences of image intensity gradients. In this work, each image was firstly preprocessed in order to enhance soft tissue according to the procedure described by [7]. HOGs were then computed inside a Region Of Interest (ROI), by setting 9 gradient orientations and a cell size equal to 8, 12 and 4 voxels for parotid glands, mandible and brainstem respectively. For a given structure, ROI extent was defined by considering a bounding box including the 55 co-registered atlas contours. VLFeat library [8] in Matlab environment (MathWorks, Natick, MA) was used to calculate HOGs. Finally, each couple of images to be compared was represented by a vector containing the squared differences between HOGs of each image. This vector will be referred as vector of features from now on. For each pair of images, more than 100000 features were computed. In order to reduce the computational time and to improve the algorithms performace, only 500 features were selected. The selection was performed by means of the univariate feature selection algorithm embedded in scikit-learn library [9] (Python environment [10]). Finally, features values were normalized by means of the standard scaler algorithm available in scikit-learn.

Atlas selection A Support Vector Machine (SVM) [11] algorithm, with a linear kernel, was employed as learning model to analyze the data. In this work, the scikit-learn implementation was used. Training was run on each possible image volume couple extracted from the 55 atlases. So far, a total of 2970 (55×54) training examples were generated. The SVM input was the feature vector. The output was the DSC computed between each pair of corresponding structures in each image couple. The DSC was computed after deforming the atlas on the target image by means of non rigid registration. During the testing phase, the input vectors were used to predict the DSC value. DSC was finally used to rank the atlases, going from the best (up to 1) to the worst (down to 0).

Image Registration and Segmentation MABS algorithm, available in Plastimatch toolkit [12, 13] was used to execute image registration and segmentation.

Plastimatch MABS is able to execute atlas selection based on NMI or using an external ranking, run image registration (both rigid and deformable) and fuse the contours using gaussian [14] and/or STAPLE [15] voting strategies. In this work, a deformable registration, consisting of 2 B-Spline stages, was used to deform the images. Root mean square error between image intensities was used as fitness function. The final labels were generated by using both gaussian and STAPLE voting algorithms. Registrations and segmentations were computed using the first 10, 15, 20 and 25 atlases of each ranking.

2.3 Data Analysis

For validation purpose, the ideal ranking, called oracle, was extracted by registering each atlas on the query subject and calculating the real DSC between the deformed structure and the Ground Truth (GT). In addition, also NMI was used to select the atlas and compare with feature based strategy. Data analysis was performed in terms of Ranking Similarity (RS) and segmentation accuracy (computing DSC between contours). RS was quantified by computing the distance between corresponding atlases included in the two different rankings. Both oracle-NMI and oracle-feature were considered. Moreover, the number of atlases comprised in the first 15 positions both for oracle-NMI selection and oracle-feature selection was computed. Finally the DSC between automatically computed contours and GT was calculated.

2.4 Hardware

The tests were executed on a machine equipped with CPU 32 cores @ 2.80GHz, 60 GB of RAM and powered by GNU/Linux. The machine was provided by Amazon as part of an educational/research grant.

3 Results

The training phase required about 1 week for 2970 example and 4 structures. Depending on the number of atlases included in the subset, the testing phase took from 1 (10 atlases) to 2 hours (25 atlases). In table 1 the medians and the quartiles of the distance between oracle-feature and oracle-NMI rankings are reported. Ranking performed by feature method was closer to oracle than NMI one, with an average distance of 11.

This result is confirmed by looking at the number of atlases included in the first 15 positions for oracle-feature and oracle-NMI rankings respectively, as it is shown in table 2. This number is always larger, or at least equal (such as for right parotid) in the feature based ranking than NMI.

Results regarding segmentation accuracy are reported in figure 1 and in table 3, were DSCs median $\pm$ quartiles are shown both for gaussian and STAPLE

	Median features	25° perc features	75° perc features	Median NMI	25° perc NMI	75° perc NMI
Left parotid	11.25	5	19.25	15	7.12	24
Right parotid	11.75	5	20.75	13.75	7	24.75
Mandible	10.5	5	19.25	12	5	21
Brainstem	11	5	21.75	14.5	6.5	24

Table 1. *Distance between feature and NMI ranking compared to oracle one.*

	Features	NMI
Left parotid	6 ± 0.75	4.5 ± 0.625
Right parotid	5 ± 1.25	5 ± 0.625
Mandible	7 ± 1.125	6 ± 1.125
Brainstem	6 ± 1.125	4.5 ± 1.0

Table 2. *Number of atlases (median ± quartiles) present in the first 15 positions of both oracle-feature and oracle-NMI rankings.*

label fusions. It is possible to observe that not strong conclusions can be made about the performance of one method compared to the other. Although previous results suggested a better behavior of feature based ranking, the accuracy test showed high variability and dependence on the voting algorithm, number of atlases included in the dataset and analyzed structure. Overall, for a single structure the average variability was in included in the range of ± 0.05.

	Left parotid	Right parotid	Mandible	Brainstem
NMI	0.799 ± 0.039	0.769 ± 0.026	0.856 ± 0.024	0.813 ± 0.025
Features	0.797 ± 0.036	0.762 ± 0.040	0.856 ± 0.025	0.815 ± 0.025
Oracle	0.814 ± 0.019	0.788 ± 0.029	0.869 ± 0.024	0.829 ± 0.024

Table 3. *Best DSC values (median ± quartiles) for each structures and for each selection strategy.*

4 Conclusions

In this work we tested a novel atlas selection strategy based on feature detection and machine learning algorithms. The aim is to predict, for each atlas and for each structure, the DSC prior to deformable registration. This is a very important novelty since the possibility to estimate the contour accuracy before registration can substantially decrease the computational time and potentially improve the results. The original idea was proposed by Sanroma et al. and it was tested on neurological data (MRI). Here, we applied the method on CT images

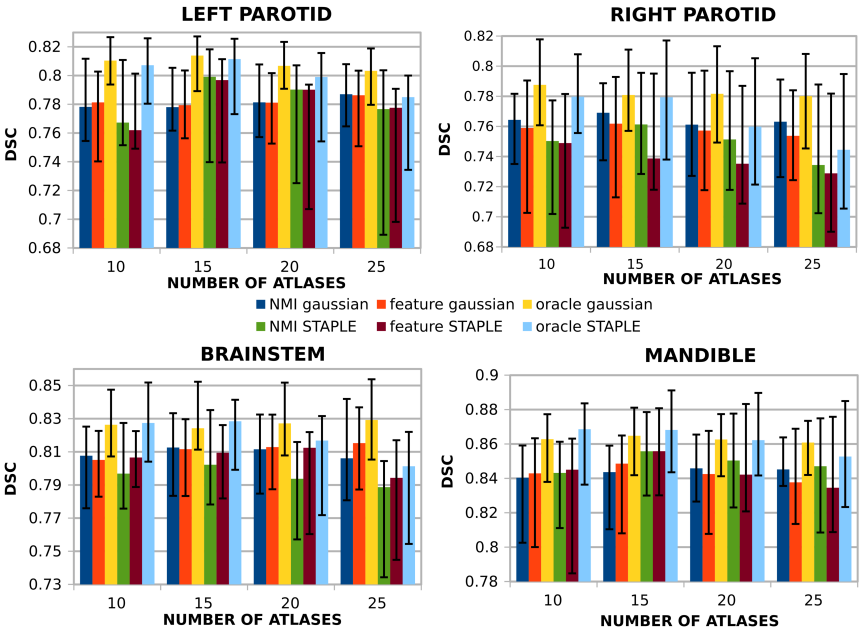


Fig. 1. *DSC values (median ± quartiles) obtained for each selection strategy, for each structure and for each atlas cluster size.*

of the HN district. The high computational time does not represent an issue for clinical applications, since atlas HOG extraction and SVM training phase can be executed offline. Only the query subject HOG computation and the SVM prediction are left to the clinical routine. Results showed that feature based selection was closer to the ideal one than the NMI based, demonstrating the effectiveness of the method. However, DSC values did not reflect such a trend. Looking at the results of RS analysis we were expecting to get an average DSC higher in feature based selection than NMI based selection. Nevertheless, the two methods gave similar results. One possible explanation is that, although a single atlas might be closer to the GT, when it is combined with other atlases the final contour is affected by different weight. This is confirmed also by looking at results coming from the two different voting methods, gaussian and STAPLE. In fact, it is possible to see how, using the same contours as input, the final segmentations generated by the two voting algorithms are rather different, proving the weight of the label fusion phase. Finally, as it is shown in figure 1, it was not possible to identify a specific number of atlases to use to run the segmentation regardless the structure under analysis. Our conclusion is that, although the method is very promising and potentially better than NMI based selection, a deeper investigation is required to understand how the final segmentation can be effected by the voting step and by the number of selected atlases. Our idea is that it is not important to select the best single atlases, but the atlases that when used

together provide the best segmentation. The future goal is then to include the voting effect in the training step, in order to select the optimal subset of atlases that combined togheter maximize the labeling accuracy.

Acknowledgement We would like to thank Amazon for providing us, for free, the server that was used for the experiments. The sponsorship is part of the Amazon Web Services grants for educational/research purpose. This work was partially supported by Associazione Italiana per la Ricerca sul Cancro (AIRC, grant number IG-13218) and by "Life Sciences and Technologies" School, PhD Programme "Molecular and translational oncology and innovative medical-surgical technologies" (XXIX cycle), "Magna Graecia" University of Catanzaro, Italy.

Bibliography

1. R. A. Heckemann, J. V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers, "Automatic anatomical brain mri segmentation combining label propagation and decision fusion," *NeuroImage*, vol. 33, no. 1, pp. 115–126, 2006.
2. P. Aljabar, R. A. Heckemann, A. Hammers, J. V. Hajnal, and D. Rueckert, "Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy," *Neuroimage*, vol. 46, no. 3, pp. 726–738, 2009.
3. C. Studholme, D. L. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3d medical image alignment," *Pattern recognition*, vol. 32, no. 1, pp. 71–86, 1999.
4. G. Sanroma, G. Wu, Y. Gao, and D. Shen, "Learning to rank atlases for multiple-atlas segmentation." *IEEE transactions on medical imaging*, 2014.
5. L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
6. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1. IEEE, 2005, pp. 886–893.
7. P. Zaffino, D. Ciardo, G. Piperno, L. Travaini, S. Comi, A. Ferrari, D. Alterio, B. Jereczek-Fossa, R. Orecchia, G. Baroni *et al.*, "Radiotherapy of hodgkin and non-hodgkin lymphoma a nonrigid image-based registration method for automatic localization of prechemotherapy gross tumor volume," *Technology in cancer research & treatment*, 2015.
8. A. Vedaldi and B. Fulkerson, "VLFeat: An open and portable library of computer vision algorithms," 2008.
9. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in python," *The Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

10. G. van Rossum and J. de Boer, "Interactively testing remote servers using the python programming language," *CWI Quarterly*, vol. 4, no. 4, pp. 283–303, 1991.
11. C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.
12. G. Sharp, R. Li, J. Wolfgang, G. Chen, M. Peroni, M. Spadea, S. Mori, J. Zhang, J. Shackleford, and N. Kandasamy, "Plastimatch-an open source software suite for radiotherapy image processing," in *Proceedings of the XVIth International Conference on the use of Computers in Radiotherapy (ICCR), Amsterdam, Netherlands*, vol. 7, 2010.
13. J. Shackleford, N. Shusharina, J. Verberg, G. Warmerdam, B. Winey, M. Neuner, P. Steininger, A. Arbisser, P. Golland, Y. Lou, C. Paganelli, M. Peroni, M. Riboldi, G. Baroni, P. Zaffino, M. F. S. Spadea, A. Aptea, Z. Saleh, J. Deasy, S. Mori, N. Kandasamy, and G. C. Sharp, "Plastimatch 1.6: current capabilities and future directions," in *Proceedings of MICCAI 2012 Image-Guidance and Multimodal Dose Planning in Radiation Therapy Workshop.*, 2012.
14. M. R. Sabuncu, B. T. Yeo, K. Van Leemput, B. Fischl, and P. Golland, "A generative model for image segmentation based on label fusion," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 10, pp. 1714–1729, 2010.
15. S. K. Warfield, K. H. Zou, and W. M. Wells, "Validation of image segmentation and expert quality with an expectation-maximization algorithm," in *Medical Image Computing and Computer-Assisted InterventionMICCAI 2002*. Springer, 2002, pp. 298–306.

Optimization of Multimodal and Multitemporal Deformable Image Registration for Head and Neck Cancer

Catarina Veiga[1], Ruheena Mendes[2], Dhanasekaran Kittappa[2], Swee-Ling Wong[2], Rachel Bodey[3], Marc Modat[4], Sebastien Ourselin[4], Gary Royle[1] and Jamie McClelland[4]

[1] Radiation Physics Group, Department of Medical Physics and Biomedical Engineering, University College London, London, UK
[2] Radiotherapy Department, University College London Hospital, London, UK
[3] Radiotherapy Physics Group, University College London Hospital, London, UK
[4] Centre for Medical Imaging Computing, Department of Medical Physics and Biomedical Engineering, University College London, London, UK

Abstract. With the growing interest in translating multimodal and functional imaging into the patient pathway it becomes important to accurately co-register the information from different imaging modalities at different timepoints. In this work we optimize and evaluate an open-source deformable image registration algorithm (NiftyReg) for CT and MR registrations in the head and neck region. The accuracy of the registrations was assessed using similarity between manual and deformed landmarks and structures. The properties of the deformation fields were also quantified.

1 Motivation

Over the last years there has been a growing interest in further introducing multimodal and functional magnetic resonance (MR) imaging into the patient treatment workflow. The additional and complementary information can potentially be used to improve the outcome of cancer radiotherapy treatments. To improve treatment planning multimodal and multiparametric imaging can be used to aid target delineation [1], for dose painting applications [2] and to decide on treatment strategies and beam arrangements considering the biological characteristics of the patient. In adaptive radiotherapy (ART) applications, it can be used as an early biomarker of the patient response and the additional anatomical and functional information can be fed into the treatment adaptation process. Early recognition of failure may allow alternative treatments to be explored, avoiding unnecessary radiation exposure and associated side effects [3]. Functional imaging also provides alternative methods to assess treatment outcome. Additionally, the treatment follow-up information may be correlated with the physical dose and biological properties of the tissue pre-therapy to develop predictive models of treatment outcome [4].

The ability to fully use clinically the biological information obtained with anatomical and functional imaging relies on the accuracy in co-registering the multiple sources of information. Therefore, accurate deformable image registration (DIR) is a key part in the different applications of multimodal and multiparametric imaging. In the head and neck (HN) region MR images are more challenging to register than other popular imaging modalities (such as CT or CBCT) for several reasons: (i) the image resolution is generally smaller and within certain volumes, such as the shoulders and thoracic area, the image quality is considerably poorer (ii) imaging patients in treatment positioning is not always possible due to coil placement and patient comfort, (iii) image-specific artifacts, such as those caused by the inhomogeneities of the magnetic field (bias), and field-of-view (FoV) clipping are common issues of current clinical acquisition protocols.

Validation of automatic DIR for multimodal and multitemporal data in the HN region in the context of radiotherapy is a rather unexplored topic in the literature. Leibfarth *et al.* use DIR between planning PET/MR and CT images for HN patients, comparing three different optimization metrics of a B-Spline DIR, for dose painting applications [5]. Slagmolen *et al.* present a small feasibility study on CT-MR and MR-MR DIR for radiotherapy treatment planning [6]. On a more technical side, some groups having been developing specialized DIR algorithms [7,8]. Other authors looked at various applications in different anatomical sites [9,10]. In this work the use of DIR for multimodal and multitemporal registrations in the HN region was investigated and optimized. An in-house DIR tool is used to co-register images from different modalities at similar time points (CT and MR) and from the same modality at different time points (MR), and the quality of the registrations was assessed using manually annotated structures. The properties of the deformation vector fields (DVFs) were also assessed.

2 Methods and Materials

2.1 Patient data acquisition

A total of three head and neck datasets were used in this study. Each patient received a routine radiotherapy planning CT (pCT), a pre-treatment MR booked as close as possible in time to the pCT (MR_1), and a follow-up MR 6 months after treatment (MR_2). The MR study consisted of T_2-weighted sequences, but functional sequences were available for future studies. MR_1 was acquired in treatment position. The inclusion criteria of this study was solely based on minimizing acquisition issues characteristic of routine MR (such as clipping of the FoV), and not to select patients with smaller anatomical changes.

The imaging protocol consisted of a planning CT (GE Widebore 16 slice system, GE Healthcare, Little Chalfont, UK) with contrast injection and reconstructed with a resolution of $0.977 \times 0.977 \times 2.5$ mm. The MR images were acquired using the MAGNETOM Avanto (Siemens Healthcare, Erlangen, Germany) MRI scanner (1.5T). In T_2-weighted images, TE varied between 90 and

110 ms, TR between 2400 and 8100 ms, slice thickness between 3 or 5 mm with a gap of 0.5 mm, and number of slices between 29 and 61. In addition, image resolution was 0.703×0.703 mm^2 or 0.859×0.859 mm^2.

2.2 Multimodal and multiparametric imaging in an ART workflow

For integration of MR data into the radiotherapy pathway, image registration is necessary between CT at the planning stage and repeat MR at different time points. To co-register multitemporal MRs with CT, two registration pathways can be followed:

1. the pos-RT MR is registered with the pre-RT MR, which is independently registered to the CT;
2. the pos-RT MR is registered directly with the CT.

In this work only the results from the first pathway were assessed quantitatively. Due to the 6 months gap between pCT and MR_2 we found that it was very challenging to tune the DIR parameters to be universally good, and in general the results were poor and physically implausible. Therefore, in our opinion it was a better approach to independently register similar anatomical information from different modalities at similar time points (pCT-MR_1) and anatomical deformations from the same modalities at different time points (MR_1-MR_2). This allows to decouple the difference in image intensity between modalities from the anatomical deformations that occur over time. Therefore, two registration methods were investigated:

- CT-MR_1: if the two images were acquired close in time and with same immobilization, a rigid registrations is the easier and natural approach. However, DIR can be used to compensate for residual setup errors. This may introduce additional issues, which will be investigated here.
- MR_1-MR_2: monomodal DIR was investigated to track over time changes in anatomy. The ability to map anatomy between timepoints also allows to propagate co-registered functional information (using the same DVFs as in the anatomical sequences).

2.3 Image registration settings

All the registrations were performed using the open-source DIR software NiftyReg (http://cmic.cs.ucl.ac.uk/home/software/). It includes a Block Matching based affine registration [11], and several B-spline Free Form Deformation based algorithms [12]. NiftyReg's stationary velocity fields implementation was the algorithm chosen for the registrations [13]. This is a symmetric and inverse-consistent algorithm that explicitly generates the transformations in the forward and inverse direction.

A total of six registration parameters with variable weight of the bending-energy penalty terms (BE) and control point spacing (CPS) were investigated per registration type. The values of BE varied between 0.01% and 1%, while the CPS

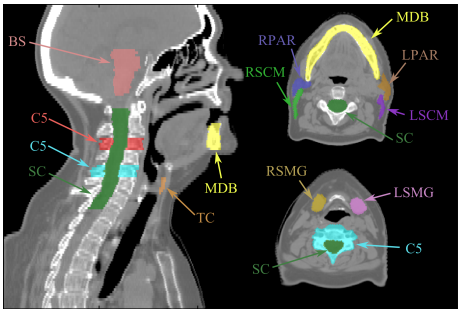


Fig. 1. Structure set manually delineated on the CT and MRs of each patient.

values tested ranged from 8 to 12 mm. NMI was chosen as similarity measure for multimodal registrations, and LNCC for monomodal registrations. LNCC was preferred over other monomodal similarity measures since it handles better the nonuniform biases that cause artifacts in MR images. To minimize the impact of artifacts due to field inhomogeneities, the MR images were corrected for bias using the N4ITK algorithm [14]. To avoid the optimization of the transformation in regions were there is no anatomical matching, the tumour was masked out in MR_1-MR_2 registrations for patients where the gross tumour disappeared between MR_1 and MR_2 (as a result of the treatment). This avoids unrealistic deformations in these regions of no real one-to-one matching. The resulting deformation is a smooth interpolation between the mapping outside the mask, guided by the regularisation of the registration.

2.4 Quantitative analysis

The registrations were compared qualitatively, by visual inspection, and quantitatively by similarity of deformed points and structures with the manually delineated gold-standard. The registration quality was assessed in both directions.

A total of 12 structures were manually delineated on the CT, MR_1 and MR_2 by expert radiation oncologists. The structures chosen provided an indication of how well the registration accounted for anatomical differences and positioning errors (Fig. 1). It consisted of vertebrae C3 and C5, mandible, thyroid cartilage (bony anatomy), spinal canal, brainstem, parotids (organs at risk, OAR), submandibular gland and sternocleidomastoid muscles (soft tissues).

We calculated the dice similarity coefficient (DSC), which provides information about the similarity between volumes, and the distance transform (DT), which is the signed Euclidean distances between the manual and deformed surfaces and infers about the closeness between contours. We computed mean, standard deviation and 95% percentile of the DT distribution (DT_{mean}, DT_{std} and $DT_{95\%}$).

Additionally, the properties of the DVFs were assessed for all the registrations. The smoothness of the transformations was analysed using the harmonic

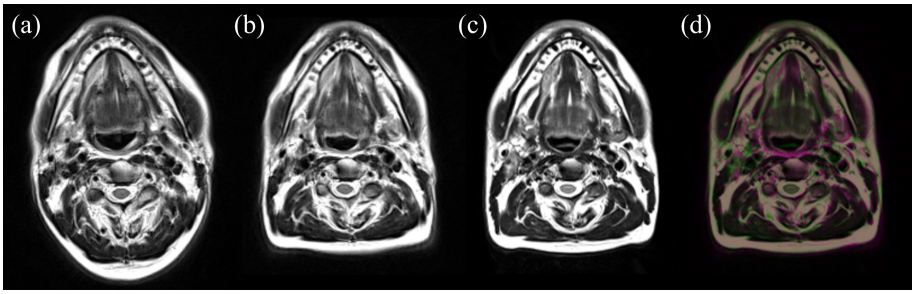


Fig. 2. Example of registrations: (a) MR_2, (b) MR_1, (c) MR_2 deformed to MR_1, and (d) overlay between MR_1 (magenta) and deformed MR_2 (green).

energy (HE) and the properties of the determinant of the Jacobian of the transformation [det(Jac)]. The HE is inversely proportional to the smoothness of the deformation, while det(Jac) indicates the level of expansion/contraction at each voxel with negative values indicating noninvertible and unrealistic deformations. Additionally, the inverse-consistency error (ICE) was calculated to investigate if the generated transforms were true inverses.

3 Results

Analysing the quantitative results obtained for all the registrations performed, we concluded that the combination of parameters that worked the best in our datasets were BE=1% and CPS=12 mm for CT-MR_1, and BE=0.1% and CPS=12 mm for MR_1-MR_2. In CT-MR_1 registrations, we found it was preferable to use a higher weight of the BE than for MR_1-MR_2. This reduced the risk of the registration causing additional uncertainties in comparison with rigid-only alignment (such as deformation of bones). However, if the immobilization is not present, those constrains should be relaxed to give the algorithm enough freedom to recover larger deformations. Since multimodal registrations had to capture larger anatomical changes the constrains had to be relaxed (Fig. 2), and the properties of the DVFs reflect also this. A higher CPS spacing in general resulted in DVFs with more desirable properties, which did not compromise the similarity between structures. For this combination of registration parameters, the mean and standard deviation obtained for the quantitative evaluation of the DIR can be found in Table 1.

Additionally to the global results provided in Table 1, we also looked at the results grouped by structure type. In MR_1-MR_2 registrations, the DSC values were 0.62±0.12, 0.77±0.08 and 0.84±0.07 for bony anatomy, soft tissues and OAR in DIR cases, and 0.4±0.3, 0.63±0.18 and 0.65±0.19 when using a rigid-only registration. The use of a rigid-only transform in MR_1-MR_2 registrations was not adequate, as we found that for some anatomical structures the overlap could be close to zero due to the large anatomical changes and differences in positioning between pre and pos-RT scans. For CT-MR_1 registrations the DSC

Table 1. Qualitative assessment of the registrations (mean± standard deviation). The results are averaged for all patients, structures or DVFs, and registration directions.

	Similarity of structures			
	DSC	DT_{mean} (mm)	DT_{std} (mm)	$DT_{95\%}$ (mm)
CT-MR_1	0.72±0.16	-0.1±1.8	2.2±1.2	6±4
MR_1-MR_2	0.74±0.13	0.1±1.7	2.3±1.6	6±5
	Properties of the DVFs			
	HE	$det(Jac)_{1\%}$	$det(Jac)_{99\%}$	ICE_{mean} (mm)
CT-MR_1	0.15±0.01	0.72±0.10	1.32±0.09	0.08±0.10
MR_1-MR_2	0.39±0.06	0.5±0.3	1.7±0.3	0.8±0.7

values were 0.57±0.18, 0.79±0.06 and 0.81±0.04 for bony anatomy, soft tissues and OAR when using DIR, and 0.57±0.18, 0.74±0.09 and 0.79±0.04 when using rigid-only registrations. The results obtained with DIR and rigid-only registrations were very similar, with DIR performing marginally better in the soft tissue regions.

4 Discussion

Promising results were found for multimodal and multitemporal registrations. For the soft tissues and OARs, values found were in agreement with results from other multimodal studies [5,6], and comparable to monomodal or quasi-monomodal (CT-CT/CT-CBCT) studies [15,16]. In spite of the large deformations between pre and pos-RT images, it was possible to achieve similar registration accuracy as for CT and MR in treatment position.

The registration of bony anatomy was poorer than for other types of structures. On one side, the reduced contrast between soft tissue and bone in MR difficulties the delineation of bones, particularly for complexly shaped structures such as the vertebrae, resulting in a non-ideal gold-standard. This low contrast also affects the quality of the registrations. However, the main interest in using MR is not to provide additional information on the bone anatomy (where CT is more relevant), but rather on the soft tissues. Thus misregistrations of the bones is of reduced importance when considering clinical applications and, in fact, in regions of higher clinical relevance, such as OAR and soft tissues, DIR performed in a higher level of accuracy. Nevertheless, the poor registration of the bones may affect nearby soft tissues so it is of importance to develop DIR strategies that account for the rigid behavior of bony anatomy.

We found that CT-MR_1 DIR slightly improved the anatomical matching in comparison to a rigid registration; however, the difference was not clinically significant. Additionally, we found that one must carefully tune its DIR registration to avoid introducing errors in this process. Further studies with a larger patient dataset are necessary to fully understand this additional uncertainty, and also to

validate CT-MR registrations for patients that can not acquire MR_1 in treatment position (i.e., with considerable setup variation between scans).

The FoV clipping was found to limit of the quality of the registrations. A clipped body contour reduces the ability to capture global deformations and generates unrealistic deformations within the patient near the edges of the FoV. For example, this was found to interfere in the registration of the mandible, which was very often clipped in the MR scans.

One of our main interests is to use multimodal and multiparametric MR imaging in the context of ART applications, which is becoming an increasingly relevant topic with the advent of the MR-Linac [17]. We acknowledge that our study is not ideal to validate the use of DIR for such applications, as the multiple MR were not acquired throughout the treatment. However we believe that using a pos-RT MR results in more challenging registrations, and therefore MR_2 can be considered a surrogate for MR acquired during treatment. This is however only true when considering the mapping of healthy tissues. Tracking of tumour volumes has to still be properly validated when the MRs are acquired throughout the course of radiotherapy. Additionally, the work here presented focused on anatomical information only, so future work will also focus on tracking functional information associated with the anatomical scans.

5 Conclusions

In this work we optimized an open-source DIR algorithm for the registration of CT and MR datasets of the head and neck. This is a first step toward incorporating additional imaging into the radiotherapy pathway.

6 Acknowledgments

The authors would like to thank Pankaj Daga, Gergely Zombori, and Matt Clarkson for all the help and support with NifTK, Shonit Punwani for all the useful discussions on multiparametric MR, and Maria Kilkenny, Heather Fitzke, Ana Mónica Lourenço and Paul Doolan for data retrieval.

C.V. is funded by Fundação para a Ciência e a Tecnologia (FCT) grant SFRH/BD/76169/2011, co-financed by ESF, POPH/QREN and EU.

References

1. X. Geets, M. Tomsej, J. Lee, *et al.* "Adaptive biological image-guided IMRT with anatomic and functional imaging in pharyngo-laryngeal tumors: Impact on target volume delineation and dose distribution using helical tomotherapy," *Radiother Oncol* 85(1):105-115 (2007)
2. P. Dirix, V. Vandecaveye, F. De Keyzer, *et al.* "Dose painting in radiotherapy for head and neck squamous cell carcinoma: value of repeated functional imaging with (18)F-FDG PET, (18)F-fluoromisonidazole PET, diffusion-weighted MRI, and dynamic contrast-enhanced MRI," *J Nucl Med* 50:1020-1027 (2009)

3. D. C. Colvin, M. E. Loveless, M. D. Does, *et al.* "Earlier detection of tumor treatment response using magnetic resonance diffusion imaging with oscillating gradients," *Magn Reson Imaging* 29:315-323 (2011)
4. V. Vandecaveye, F. De Keyzer, S. Nuyts, *et al.* "Detection of head and neck squamous cell carcinoma with diffusion weighted MRI after (chemo)radiotherapy: correlation between radiologic and histopathologic findings," *Int J Radiat Oncol Biol Phys* 67:960-971 (2007)
5. S. Leibfarth, D. Mönnich, S. Welz, *et al.* "A strategy for multimodal deformable image registration to integrate PET/MR into radiotherapy treatment planning," *Acta Oncol* 52:1353-1359 (2013)
6. P. Slagmolen, D. Loeckx, S. Roels, *et al.* "Nonrigid Registration of Multitemporal CT and MR Images for Radiotherapy Treatment Planning," in *Biomedical Image Registration* 297-305 (2006)
7. A. du Bois d'Aische, M. De Craene, X. Geets, *et al.* "Estimation of the deformations induced by articulated bodies: Registration of the spinal column," *Biomedical Signal Processing and Control* 2:16-24 (2007)
8. M. Söhn, M. Birkner, Y. Chi, *et al.* "Model-independent, multimodality deformable image registration by local matching of anatomical features and minimization of elastic energy," *Med Phys* 35:866-878 (2008)
9. K. Brock and Deformable Registration Accuracy Consortium, "Results of a multi-institution deformable registration accuracy study (MIDRAS)," *Int J Radiat Oncol Biol Phys* 76:583-596 (2010)
10. A. Akbarzadeh, D. Gutierrez, A. Baskin, *et al.* "Evaluation of whole-body MR to CT deformable image registration," *J Appl Clin Med Phys* 14(4):4163 (2013)
11. S. Ourselin, A. Roche, G. Subsol, *et al.*, "Reconstructing a 3d structure from serial histological sections," *Image Vis Comput* 19:25-31 (2001)
12. D. Rueckert, L. I. Sonoda, C. Hayes, *et al.*, "Nonrigid registration using free-form deformations: application to breast MR images," *IEEE Trans Med Imaging* 18:712–721 (1999)
13. M. Modat, P. Daga, J. Cardoso, *et al.*, "Parametric non-rigid registration using a stationary velocity field," *IEEE Workshop MMBIA* 145-150 (2012)
14. N. Tustison, B. Avants, P. Cook, *et al.*, "N4itk: Improved N3 Bias Correction," *IEEE Trans Med Imaging* 29:1310-1320 (2010)
15. P. Castador, J. Lee, A. Parraga, *et al.*, 'Comparison of 12 deformable registration stratefies in adaptive radiation therapy for the treatment of head and neck tumors," *Radiother Oncol* 89(1):1-12 (2008)
16. C. Veiga, J. McClelland, S. Moinuddin, *et al.*, "Toward adaptive radiotherapy for head and neckpatients: Feasibility study on using CT-to-CBCT deformable registration "dose of the day" calculations," *Med Phys* 41(3):031703 (2014)
17. B. W. Raaymakers, J. J. W. Lagendijk, J. Overweg, *et al.*, "Integrating a 1.5T MRI scanner with a 6 MV accelerator: proof of concept," *Phys Med Biol* 54:N299–N237 (2009).

4D Lung CT Segmentation for Radiation Therapy Applications

S.G. Yeary[1], G.E. Christensen[1], J.E. Bayouth[2], S. Bodduluri[1], Y. Pan[1], J. Guo[1], K. Du[2], J.H. Song[1], B. Zhao[1], I. Oguz[1], and J.M. Reinhardt[1]

[1] University of Iowa
[2] University of Wisconsin-Madison

Abstract. Radiation therapy protocols for lung cancer treatment planning commonly use a 3D CT image to delineate critical regions of interest and a 4D respiratory-gated CT to track tumor respiratory motion. 4D CT data sets contain a great deal of information and can be analyzed to obtain quantitative data on respiratory system dynamics, mechanics, and function. We describe a method that transfers segmentation contours from the 3D planning CT to the entire 4D data set, even though the lung changes size and shape during image acquisition due to respiration. Our proposed method uses deformable image registration to align all phases of the 4D CT into the coordinate system of the 3D planning CT. A graph optimization algorithm is used to segment the lung boundaries of the aligned phases of the 4D data set. The algorithm is initialized with the 3D segmentation of the planning CT. The segmentations of the individual phases in the planning CT coordinate system are then transformed back into each of the original phase coordinate systems providing 3D segmented volumes at each phase of the 4D CT. We tested the method on six data sets from subjects about to undergo radiation therapy for lung cancer. For each subject a 3D planning CT was segmented as part of the treatment planning process. This planning CT was used to segment the 4D CT at ten phase points across the respiratory cycle. The results of our method were compared against manual segmentations of four phases giving an average surface distance of 0.417 mm and average Dice coefficient of 0.983. The 4D segmentations appear to be more consistent across phases than the manual segmentations, especially near the mediastinum. The results show that this method can provide accurate 4D segmentations from a single segmented 3D CT image.

Keywords: segmentation, 4D image analysis, radiation therapy

1 Introduction

Respiratory-gated 4D CT imaging is used in radiation therapy planning for lung cancer treatment to measure tumor position at different points in the respiratory cycle. These data sets can be further analyzed to obtain quantitative measurements of lung dynamics [1], lung mechanics [2], and regional lung ventilation [3].

Accurate lung segmentation of the constituent 3D volumes of a 4D data set is required before quantitative analysis can be performed.

An initial 3D CT scan (referred to here as the "planning" CT) is conducted on all patients that undergo a 4D CT scan. The planning scan is performed in breath-hold and typically provides higher signal-to-noise and superior spatial resolution than a 4D CT acquisition. As a result, the planning CT image provides a high quality image of the lung anatomy and — due to acquisition during a single breath hold — has fewer motion artifacts than a 4D CT data set. The planning CT is commonly used for radiation therapy planning and includes lung and tumor segmentations generated by an image analyst.

Several different approaches for lung tissue segmentation in 3D CT images have been proposed, including optimal thresholding [4], adaptive border tracking [5], graph search algorithms [6], and segmentation-by-registration with the use of a lung atlas [7]. These algorithms vary in segmentation accuracy, computational complexity, and robustness to abnormalities and disease, but it is unlikely any 3D algorithm could accurately and consistently segment all volumes of a 4D data set independently. It seems likely that a direct 4D segmentation may give better results than 3D volume-by-volume segmentation as all of the available image information is used and temporal coherence across respiratory phases can be ensured. Additionally, since 4D CT data sets typically have lower image quality and more motion artifacts than the typical planning CT, conventional segmentation algorithms may fail or perform poorly when applied to the individual 3D volumes of the 4D CT.

Lung segmentation in a 4D CT data set can be viewed as a multiple surface segmentation problem, where prior known interactions between surfaces can be used to improve segmentation robustness. Optimal surface finding is a graph search method that has been used in many applications to segment multiple surfaces in n-D medical images [8]. In [9], optimal surface finding is used to segment multiple objects and surfaces in the brain. In [10] and [11], an active shape model is combined with optimal surface finding to simultaneously segment multiple lung volumes. In this work, we propose a similar method for lung segmentation in 4D CT data sets consisting of ten volumes imaged at different phases during the breathing cycle.

Our 4D segmentation algorithm is based on a combination of deformable image registration and 4D optimal surface finding. First, image registration is used to align all volumes of the 4D data set to the coordinates of the planning CT so that the planning CT segmentation can be used as a shape prior for all 4D CT lung volumes. A 4D geometric graph is constructed using shape priors and the optimal surface is found using an efficient maximum flow algorithm.

2 Methods

Our proposed 4D segmentation method consists of two stages: image registration and optimal surface finding. The processing pipeline is shown in Figure 1. We assume that the 4D CT consists of N individual 3D phase volumes sampled

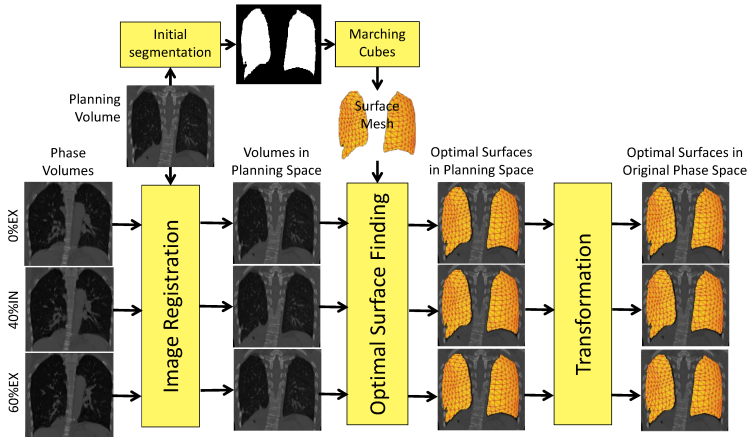


Fig. 1. Processing flowchart. The process starts with the planning CT segmentation and produces N segmented 3D phase volumes.

across the respiratory cycle. We use the notation 20%IN to refer to the 3D image reconstructed at 20% of tidal lung volume during inspiration, and 80%EX to refer to the 3D image reconstructed at 80% tidal lung volume during expiration.

2.1 Registration

In the first stage, each 3D volume of the 4D data set is roughly aligned to the planning CT volume using deformable image registration. Our method does not require an accurate registration, so a low resolution registration framework can be used to reduce computation time. In this work, we used the Elastix image registration software [12] to perform the alignment of the planning CT to each of the 3D volumes. A b-spline transform was used with a grid spacing of 20 mm. Normalized correlation was used for the similarity metric and a gradient descent optimization algorithm was used with a maximum of 1000 iterations. Note that it is possible to perform all registrations simultaneously since each planning CT to 3D phase volume registration is performed independently. Thus, performing the registrations in parallel can reduce the overall computation time.

After performing the image registration, we obtain a set of deformed 3D images in the same coordinate space as the planning CT image, allowing the existing planning CT segmentation to be used as a subject specific shape prior for the next stage.

2.2 Optimal Surface Finding

Optimal surface finding is a graph search framework used for simultaneous segmentation of multiple interacting surfaces in images [9]. The lungs at each phase image are treated as surfaces and the interactions between different phases represent temporal constraints. A shape prior similar to the true surface is used to

initialize the graph search space. Graph $G(V, E)$ consisting of node set V and edge set E is constructed in a region around the shape prior. In our work, the planning CT segmentation is utilized as the shape prior for all phases. Since the same shape prior was used for all phases, all sub-graphs G^t had the same geometric structure in space. This greatly simplified the 4D graph construction.

For multiple surface segmentation the following graph construction is done for each surface, using similar graph construction parameters to those used in [11]. First, the marching cubes algorithm is used to transform the planning CT segmentation into a mesh representation with approximately 10,000 vertices. A search profile, or graph column V_i, is constructed at each mesh vertex and nodes $v_{i,j}$ are sampled at equal distances along the profile. For the experiments presented in this paper, the graph column length was 60 nodes sampled at a distance of 0.35 mm. The search profiles were defined using the electric field lines generated by treating each mesh vertex as a charged point source as described in [13].

Graph edges are introduced to define feasible surfaces as those which maintain the topology of the shape prior and have a certain degree of surface smoothness. This is achieved by introducing intracolumn edges that force the optimal surface to intersect each graph column V_i at exactly one node $v_{i,j}$ and intercolumn edges that restrict the surface height from changing more than Δ nodes between neighboring columns. The intracolumn edges $< v_{i,j}, v_{i,j-1} >$ were introduced within each column of the graph and the intercolumn edges $< v_{i,j}, v_{k,j-\Delta} >$ were introduced between all neighboring columns V_i and V_k. For experiments used in this paper, we used $\Delta = 12$ nodes. Each node $v \epsilon V$ was assigned a cost inversely proportional to the likelihood the surface contains $v_{i,j}$. The gradient magnitude of the deformed phase images was used as the cost function in this work. The cost of a feasible surface is the summation of all node costs on the surface. The optimal surface corresponds to the surface with the minimum cost among all feasible surfaces.

Interactions between surfaces were enforced by introducing temporal edges $< v^t_{i,j}, v^{t+1}_{i,j-\delta} >$ between corresponding columns of the surfaces. In our 4D lung segmentation, the deformed lung of each image phase $t = 1 \ldots N$ was treated as a surface and the temporal constraint δ enforced surface consistency between phases. Here we used $\delta = 20$ nodes.

A maximum flow algorithm was used to find the globally optimal solution of the cost function as described in [14]. This resulted in N surfaces in the deformed image space. The surfaces were transformed back to the coordinates of each original 3D phase volume using the transformations obtained from the registration.

2.3 Data Sets and Experimental Setup

4D CT data sets and a planning CT image from six lung cancer subjects about to undergo radiation therapy were used for this study. All data were gathered under a protocol approved by the University of Iowa Institutional Review Board (IRB 200905703). Each 4D CT contained ten 3D volumes retrospectively reconstructed in 20% steps from inspiration to exhalation. All images were resampled

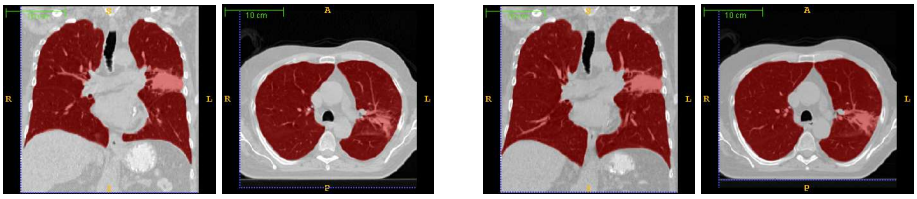


Fig. 2. Automatic 4D segmentation results for subject A. From left to right: 0%EX coronal plane, 0%EX transverse plane, 100%IN coronal plane, 100%IN transverse plane.

to obtain $1 \times 1 \times 1$ mm^3 voxels. Ground truth segmentations were generated for four respiratory phases (0%EX, 40%IN, 100%IN, 60%EX) by a trained image analyst (Y.P.) using the Analyze^{TM} software package. The same image analyst segmented the planning CT image for each subject to provide initial shape priors.

To test the sensitivity of the method to the segmentation shape prior, we ran an additional experiment for one data set where different 3D phase volumes were used in place of the planning CT image. This was done by using the phase volumes at 20%IN, 60%IN, 40%EX, and 20%EX to establish the initial shape priors for the optimal surface finding.

3 Results

3.1 4D Segmentation with planning CT as Initial Segmentation

Tables 1 and 2 show the Dice coefficients and mean absolute surface distance for each of the four evaluation phases. Figure 2 shows the automatic 4D segmentation, displayed at two of the individual phase reconstructions, for one subject.

Subject	0%EX		40%IN		100%IN		60%EX		Mean
ID	Left	Right	Left	Right	Left	Right	Left	Right	
A	0.985	0.986	0.988	0.988	0.985	0.986	0.985	0.986	0.986
B	0.976	0.983	0.976	0.986	0.973	0.984	0.974	0.984	0.980
C	0.978	0.978	0.977	0.985	0.984	0.982	0.984	0.977	0.981
D	0.988	0.988	0.981	0.982	0.983	0.978	0.982	0.979	0.983
E	0.981	0.983	0.982	0.981	0.982	0.983	0.983	0.981	0.982
F	0.985	0.983	0.984	0.984	0.984	0.985	0.985	0.985	0.985

Table 1. Dice coefficients for the 4D segmentation compared to the ground truth segmentations for each of the four 3D phase volumes used for evaluation. Dice coefficients are calculated separately for left and right lungs. Mean is calculated across four phases and both lungs for each subject.

Subject ID	0%EX		40%IN	
	Left	Right	Left	Right
A	0.385 ± 0.807	0.437 ± 1.203	0.315 ± 0.860	0.353 ± 0.790
B	0.428 ± 0.890	0.422 ± 0.766	0.410 ± 0.715	0.394 ± 0.743
C	0.431 ± 0.755	0.536 ± 1.375	0.501 ± 1.126	0.320 ± 0.682
D	0.242 ± 0.573	0.297 ± 0.784	0.435 ± 0.959	0.491 ± 1.194
E	0.499 ± 1.315	0.536 ± 1.555	0.455 ± 1.086	0.498 ± 0.959
F	0.307 ± 0.566	0.443 ± 1.135	0.342 ± 0.655	0.376 ± 0.836

Subject ID	100%IN		60%EX		Mean
	Left	Right	Left	Right	
A	0.407 ± 0.841	0.459 ± 0.976	0.417 ± 0.854	0.429 ± 0.853	0.400
B	0.464 ± 0.950	0.434 ± 0.773	0.441 ± 0.778	0.407 ± 0.717	0.425
C	0.365 ± 0.896	0.423 ± 0.909	0.349 ± 0.839	0.520 ± 0.871	0.431
D	0.380 ± 0.694	0.542 ± 0.911	0.399 ± 0.795	0.547 ± 1.197	0.417
E	0.422 ± 0.747	0.468 ± 0.986	0.424 ± 1.009	0.495 ± 0.983	0.475
F	0.347 ± 0.662	0.338 ± 0.625	0.340 ± 0.716	0.353 ± 0.716	0.356

Table 2. Average symmetric absolute surface distance ± standard deviation (mm) for the 4D segmentation compared to the ground truth segmentations for each of the four 3D phase volumes used for evaluation. Mean is calculated across four phases and both lungs for each subject.

3.2 Sensitivity to Initial Segmentation

The sensitivity to the choice of initial segmentation was tested by segmenting one 4D data set (subject A) with four different phase volumes used as the initial segmentation. Table 3 gives the Dice coefficients for the four evaluation volumes with four different initializations.

Initialization	0%EX		40%IN		100%IN		60%EX		Mean
Phase Volume	Left	Right	Left	Right	Left	Right	Left	Right	
20%IN	0.982	0.985	0.985	0.986	0.984	0.987	0.982	0.985	0.984
60%IN	0.981	0.984	0.985	0.987	0.983	0.987	0.981	0.986	0.984
40%EX	0.982	0.985	0.985	0.987	0.982	0.986	0.982	0.986	0.984
20%EX	0.981	0.985	0.985	0.986	0.983	0.987	0.981	0.985	0.984

Table 3. Dice coefficients for the 4D segmentation of subject A using four different initial segmentations as the shape prior. Mean is calculated across four phases and both lungs for each subject.

4 Discussion

The results show that image registration and optimal surface finding can be used to produce an accurate 4D segmentation from one initial 3D segmentation. In our experiments an average symmetric absolute surface distance of 0.417 mm and a mean Dice coefficient of 0.983 were obtained. This mean surface error is less than half of a voxel. We observe that most errors were where the airways enter the lungs at the mediastinum. This is a subjective area to manually segment, as different boundaries may be chosen by different analysts. For comparison, a 3D thresholding based algorithm similar to [4] was applied to each volume and an average symmetric absolute surface distance of 0.5644 mm and a mean Dice coefficient of 0.9790 were obtained. Although the simple thresholding gave good results, tumors were not included in the segmentation thus it is not suitable for this application. The same mean Dice coefficient of 0.984 was obtained for all initial segmentations showing the algorithm it robust to the initial segmentation. Since all phase volumes were segmented simultaneously by the optimal surface finding, our final segmentation contours are consistent across all phases. The segmentations appear to be more consistent across phases than the manual segmentations, especially near the mediastinum. Additionally, the method was able to include large chest wall tumors in the segmentation, which is a difficult problem due to the similarity in intensity with the surrounding tissue.

We were able to use a computationally inexpensive registration by using the optimal surface finding to refine the initial segmentation result. In many cases it is even possible to use an affine registration when the lung deformation across respiration is not too large. The optimal surface finding incorporates shape prior information, surface smoothness constraints, and temporal surface constraints making it ideal for 4D lung segmentation. The optimal surface finding guarantees a globally optimal minimization of the selected cost function.

The experiments were run on a Linux machine with an Intel Xeon 2.27 GHz CPU and 48 GB of RAM. The deformable image registration takes 2 minutes per phase, which can be done in parallel. The optimal surface finding takes approximately 4.5 minutes with our parameters. This yields a total time of 6.5 minutes to produce segmentations of all 10 volumes of a 4D data set. The manual segmentation used for evaluation took approximately 30 minutes per 3D volume.

The proposed framework is extensible to segmentation of other surfaces. For example, sublobar segmentations can be used to study the mechanical properties and sliding of the lung lobes. Our framework can be easily extended to handle lobar segmentations.

5 Summary

We proposed a method for 4D segmentation of lung tissue in respiratory-gated data sets. The method utilizes the planning CT segmentation to obtain accurate segmentations for all phase volumes of a 4D data set. An average surface distance of 0.417 mm and an average Dice coefficient of 0.983 were achieved.

Based on preliminary results we showed that the method is robust to the initial planning CT image segmentation. The 4D segmentation is valuable for further quantitative analysis of the data sets.

6 Acknowledgments

This work was supported in part by NIH grant CA166703.

References

1. Wu, G., Wang, Q., Lian, J., Shen, D.: Estimating the 4D respiratory lung motion by spatiotemporal registration and super-resolution image reconstruction. Medical Physics **40** (2013)
2. Amelon, R.E., Cao, K., Ding, K., Christensen, G.E., Reinhardt, J.M., Raghavan, M.L.: Three-dimensional characterization of regional lung deformation. J Biomechanics **44** (Sep 2011) 2489–2495
3. Du, K., Bayouth, J.E., Cao, K., Christensen, G.E., Ding, K., Reinhardt, J.M.: Reproducibility of registration-based measures of lung tissue expansion. Medical Physics **39** (2012)
4. Hu, S., Hoffman, E.A., Reinhardt, J.M.: Automatic lung segmentation for accurate quantification of volumetric X-ray CT images. IEEE Transactions on Medical Imaging **20** (2001)
5. Pu, J., Roos, J., Yi, C.A., Napel, S., Rubin, G.D., Paik, D.S.: Adaptive border marching algorithm: Automatic lung segmentation on chest CT images. Comput Med Imaging Graph **32**(6) (Sep 2008) 452–462
6. Hua, P., Song, Q., Sonka, M., Hoffman, E.A., Reinhardt, J.M.: Segmentation of pathological and diseased lung tissue in CT images using a graph-search algorithm. IEEE International Symposium on Biomedical Imaging (2011)
7. van Rikxoort, E.M., de Hoop, B., Viergever, M.A., Prokop, M., van Ginneken, B.: Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. Medical Physics **36** (2009)
8. Li, K., Wu, X., Chen, D.Z., Sonka, M.: Optimal surface segmentation in volumetric images—A graph-theoretic approach. IEEE Transactions on Pattern Analysis and Machine Intelligence **28** (2006)
9. Oguz, I., Sonka, M.: LOGISMOS_B: Layered optimal graph image segmentation of multiple objects and surfaces for the brain. IEEE TMI **33** (2014) 1220–1235
10. Wilms, M., Ehrhardt, J., Handels, H.: A 4D statistical shape model for automated segmentation of lungs with large tumors. Proceedings of MICCAI (2012) 347–354
11. Sun, S., Sonka, M., Beichel, R.R.: Graph-based 4D lung segmentation in CT images with expert-guided computer-aided refinement. IEEE 10th International Symposium on Biomedical Imaging (2013)
12. Klein, S., Staring, M., Murphy, K., Viergever, M., Pluim, J.: Elastix: A toolbox for intensity-based medical image registration. IEEE Transactions on Medical Imaging **29** (Jan 2010) 196–205
13. Yin, Y., Zhang, X., Williams, R., Wu, X., Anderson, D.D., Sonka, M.: LOGISMOS-layered optimal graph image segmentation of multiple objects and surfaces: Cartilage segmentation in the knee joint. IEEE Transactions on Medical Imaging **29** (2010)
14. Boykov, Y., Jolly, M.: An experimental comparison of min-cut/max-flow algorithms for energy minimizations in vision. IEEE Transactions on Pattern Analysis and Machine Intelligence **26** (2004) 1124–1137

Application and Evaluation of Wavelet-based Surface Reconstruction for Contour Propagation in Radiotherapy

Stefano Moriconi[1], Elisa Scalco[1], Tiziana Rancati[2], Antonella Messina[3], Tommaso Giandini[4,5], Riccardo Valdagni[2,6], and Giovanna Rizzo[1]

[1]Institute of Molecular Bioimaging and Physiology (IBFM-CNR), Segrate (MI), Italy
{stefano.moriconi, elisa.scalco, giovanna.rizzo}@ibfm.cnr.it
[2]Prostate Cancer Program, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy
{tiziana.rancati, riccardo.valdagni}@istitutotumori.mi.it
[3]Radiology Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy
antonella.messina@istitutotumori.mi.it
[4]Medical Physics Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy
[5]Radiotherapy Division, European Institute of Oncology, Milan, Italy
tommaso.giandini@istitutotumori.mi.it
[6]Dept. of Radiation Oncology1, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

Abstract. A wavelet-based open-source approach for three-dimensional (3D) surface reconstruction of anatomical structures, not yet used for radiotherapy (RT) applications, was presented. This was obtained from manual cross-sectional contours by combining both image voxel segmentation processing and implicit surface streaming methods using wavelets. 3D meshes reconstructed with the proposed approach were compared to those obtained from other three traditional triangulation algorithms. Evaluation was performed in terms of mesh quality metrics and accuracy of contour propagation in the pelvic district. Results have shown a smoothness and regularity of the reconstructed surface, comparable, or even better, than the other methods, and an accuracy of contour propagation in line with the state-of-art literature. This demonstrated the efficacy of the proposed approach for the 3D surface reconstruction in RT.

Keywords: surface reconstruction; wavelet; contour propagation; mesh quality

1 Introduction

An accurate and reliable organ representation is of great interest in Radiotherapy (RT), since dose planning, treatment evaluation and toxicity model are based on organ contours or organ volume. The availability of tomographic images acquired during the course of RT has opened the possibility to daily or weekly monitoring organ changes, thus improving the efficacy of the treatment taking into account these modifications. Deformable image registration combined with contour propagation methods are able to recover these spatial changes and to automatically recontour organs of interest. This approach was proposed by many authors [1-4]; however, not every work has described in detail the whole procedure. In general, contour propagation can be faced by deforming binary masks extracted from 2D contours [3], [4] or by deforming 3D surface meshes generated from 2D contours [1], [3]. In this last case, the choice of the method for the surface reconstruction has an impact on the final result: in fact, it is desirable to have regular and smoothed meshes for a good rep-

resentation of the organs of interest, in order to make some evaluation on the organs surface or volume (e.g. Dose Volume Histogram and Dose Surface Histogram) or to obtain an accurate contour propagation. However, the algorithm used to generate surface representation is not always described. Several studies in the past decades proposed different methods to obtain 3D surfaces from generic cross-sectional data, as, for example, marching cubes [5] and power crust [6]. Furthermore, licensed commercial software provide with automatic built-in tools for 3D surface reconstruction (e.g. Amira®, Mercury Systems, MA, USA). By using these approaches, acceptable results in terms of smoothness, regularization and mesh quality are not always obtained. Moreover, they can show a complex and time-consuming processing, which sometimes results in failure, especially with large and complex datasets, and, regarding commercial software, they are black-box or little is customizable about the surface reconstruction from contours.

In this work we proposed a different approach for 3D surface reconstruction of organs of interest from slice contours, based on wavelet [7], which theoretically provides smooth, accurate and interpolated 3D geometrical model of the organ, and which has not yet been proposed in RT applications. We compared it with other classical methods already used in this context [1], [8], [9], by evaluating them in terms of their intrinsic mesh quality and of accuracy of contour propagation. In this work we considered patients treated for prostate cancer and we focused on relevant structures of the pelvic district: the prostate itself, the bladder, the rectum and the penile bulb.

2 Materials and methods

2.1 Wavelet-based Surface Reconstruction (WSR).

Bi-dimensional (2D) manual contours are first mapped into binary voxel image segmentation. Since serial slice data is usually anisotropic, an optional inter-slice thickening step is used to obtain isotropic-like voxel segmentation of the structure. The inter-slice thickening step generates new slices between two adjacent ones by interpolating the binary segmentation using logical operators. The binary segmentation is then converted into a set of 3D oriented points in the physical-space coordinates. Each 3D oriented point, i.e. $\bar{p} = [x_p\ y_p\ z_p]$, is associated with its outward-pointing normal vector $\bar{n}$. The estimation of the 3D points' normals is determined by the cross product between the vectors joining the two closest points of $\bar{p}$ on the same slice and those on adjacent slices respectively. The set of normals is then normalized by the Euclidean norm. The outward-pointing direction is verified by evaluating the sign of the dot product between $\bar{n}$ and the vector joining $\bar{p}$ and the closest inner reference of the structure (i.e., the binary skeletonization), otherwise the direction of $\bar{n}$ is inverted accordingly. This step is performed in order to obtain a dense and homogeneous cluster of 3D oriented points which represents the discrete sampling of the surface to reconstruct. We refer to [10] for a more detailed description of the binary segmentation pre-processing step.

WSR algorithm is employed then to extract a 3D smooth, interpolated mesh of the structure from the cluster of oriented points. The algorithm, which is accurately described in [7], estimates an approximation of the solid structure by means of an indicator function (χ) using the input points. The algorithm constructs then an approximation of the original surface as the zero level-set of the indicator function itself. WSR algorithm can be summarized

in four sub-steps, namely, octree configuration, estimation of χ, smoothing χ and polygon generation. By configuring the octree, the multi-resolution spatial grid for estimating the indicator function is defined. It represents a hierarchical structure of maximal depth d_m that encodes subtrees to cells, corresponding to higher-resolution detail. The greater d_m, the finer the surface reconstruction. Limitations in d_m can be introduced by the finite allocable memory and by the intrinsic spatial resolution and extension of the structure's input points. The estimation of χ at each cell of the octree takes advantage of the multi-scale and hierarchical structure of the compactly supported wavelet basis. This allows an efficient representation of χ with a relatively small number of coefficients. By applying the divergence theorem, the wavelet coefficients are approximated over the input oriented points. The resulting indicator function χ is 0 outside the approximated solid structure, whereas it is 1 on the inside. The smoothness of χ depends on the choice of the wavelet basis. We adopted the Daubechies wavelet (D4) for the surface reconstruction. In general, as the support of the wavelet decreases, so it also does the smoothness of the reconstructed surface. To this purpose a further smoothing step is performed on the indicator function before extracting the surface mesh. It consists in pruning the isolated cells of the octree, and in filtering the octree with a convolution mask on adjacent cells. The convolution mask adopted is the tensor product of the mask (¼ ½ ¼) in $\mathbb{R}^3$. Lastly the polygonal (triangle) surface extraction is obtained by applying an octree contouring method [11] and eventually by running Marching Cubes [5], which is guaranteed to produce a water-tight topological and geometrical manifold. Adopted implementation: *http://josiahmanson.com/research/wavelet_reconstruct*

2.2 Dataset

T2-weighted MRI images acquired before and three months after RT were considered for five patients (MRI1 and MRI2, respectively). A set of 2D contours of prostate, bladder, rectum and bulb were manually drawn on each MRI study, consisting thus in a total set of 40 contours (i.e., 5 patients x 2 time-points x 4 structures). The choice of these four organs was justified by the need of testing methods on different conditions of size and amount of deformation. Surfaces were then generated from each set of 2D contours using the WSR method previously described and other three classical approaches already used in RT context, namely Marching Cubes (MCB) [5] combined with a Taubin non-shrinking filter [12], Power Crust (PWC) [6] and Amira surface generation (AMR) using the built-in *unconstrained* smoothing modality [13].

2.3 Image registration and contour propagation

MRI2 were mapped on the MRI1 using a non-rigid image registration method implemented in the open-source software Elastix [14]. The chosen registration method is the classical Free-Form Deformation based on B-splines [15], with the parameters optimization available in the Elastix implementation. In particular, Normalized Mutual Information was chosen as the similarity metric, the adaptive stochastic gradient descent was the adopted optimization algorithm, and by setting a sufficiently high number of iterations the registration is guaranteed to reach convergence. A multi-resolution approach in 5 steps was adopted, using a uniform control-points grid at each step; the final grid was made up by isometric cubes of 10 mm.

To automatically deform organs' contours of MRI1 on MRI2, a contour propagation method was adopted. The propagation procedure was integrated in the Elastix toolbox by applying the estimated deformation field on the vertices of the mesh generated from the 2D contours.

2.4 Mesh evaluation

The evaluation of the WSR method was carried out by comparing it with MCB, PWC and AMR algorithms. This comparison was done in terms of indices of mesh quality and accuracy of the contour propagation.

Mesh quality. The evaluation of WSR, MCB, PWC and AMR surfaces, before and after deformation, is here performed in terms of aspect ratio (AR) and curvature. AR quantifies the mesh quality by evaluating the regularity of each triangular element of the reconstructed surface relative to an optimal reference, i.e. the equilateral triangle [16]: the greater the AR, the worse the mesh quality. Curvature is a measure of the surface departure from planarity: meshes with smooth curvature have a more regular and realistic look. This was done by calculating the mean curvature of each vertex of the mesh [17] using the Kitware's open-source program ParaView (*www.paraview.org*).

Since in RT it is also of interest to have accurate and smoothed contours on 2D slices of the images, the reconstructed surfaces were cut on the plane of the correspondent MRI and contours were qualitatively evaluated.

Accuracy of contour propagation. In order to estimate the influence of the choice of the surface reconstruction method on contour propagation, accuracy was evaluated in terms of distances between the deformed mesh and the reconstructed surface from the originally delineated contours on MRI2. For this calculation, each deformed mesh was compared with the original one generated with the same method: i.e. WSR deformed surfaces were compared with surfaces reconstructed with WSR, and so on. Distances between surfaces were calculated as the Euclidean distance between every vertex point in the deformed mesh and the closest vertex point in the manual mesh. The same procedure was repeated inverting the reference surface and the mean (D_mean) and maximum (D_max) distances were stored.

3 Results

An exemplificative result of the four surface reconstruction methods and their effects on contour propagation is reported in Fig. 1, where the prostate is represented before and after deformation. WSR, MCB and AMR successfully generated all the surfaces, whereas PWC failed in 9 of 40 cases due to a non-reached convergence of the algorithm.

Mean percentage AR values for the prostate meshes and curvature of the four organs obtained using the considered methods are reported in Table 1 and Fig. 2, respectively. In the remaining three organs, values of AR had the same trend of the prostate and were not further reported.

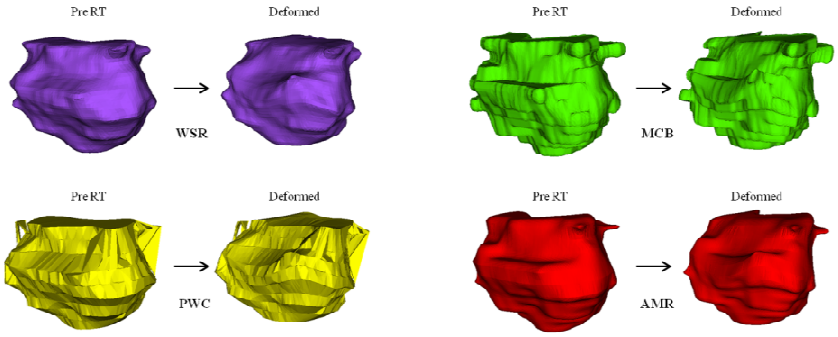


Fig. 1. Prostate surface reconstruction using the four described method starting from 2D contours (PreRT) and after deformation (Deformed). WSR: purple; MCB: green; PWC: yellow; AMR: red.

3.1 Mesh quality evaluation

Table 1. Values percentage's distribution of aspect ratio for the four methods in the prostate.

	Prostate AR percentage								
	1-10	**10-10^2**	**10^2-10^3**	**>10^3**		**1-10**	**10-10^2**	**10^2-10^3**	**>10^3**
WSR	0.96	0.04	0.00	0.00	**PWC**	0.63	0.17	0.08	0.13
MCB	1.00	0.00	0.00	0.00	**AMR**	0.99	0.01	0.00	0.00

As a general remark, the distribution of mean percentage AR for meshes obtained with WSR, AMR and MCB was limited to values in range 1-100, with more than 95% of triangles having an AR index lower than 10, thus meaning overall regular shape of the triangular elements. In particular, MCB showed the best performance, having the 100% of the triangles with AR lower than 10. Conversely, the AR percentage distribution of meshes obtained with PWC spans the whole range of values with a relatively greater amount of degenerate triangles (>35%). This was observed in original meshes as well as in propagated meshes, after the deformation process.

Even if MCB had the best AR, when surface curvature was observed, it showed a staircase effect, with anomalous curvature on the edges correspondent to slice changes; this did not happen with AMR and WSR, which had smooth and regular surfaces, with a planar mean curvature. PWC, which was already the worse method when AR was considered, deviated from planarity with a lot of vertices with anomalous curvature. This was verified on meshes generated from 2D contours as well as on deformed surfaces (Fig. 2).

2D contours obtained from the deformed prostate meshes were represented in Fig. 3, confirming what found in curvature. In fact it is possible to note that both AMR and WSR have a regular and smoothed shape, whereas PWC and MCB contours have an irregular look.

3.2 Accuracy of contour propagation

Results of accuracy of contour propagation are reported in Table 2.

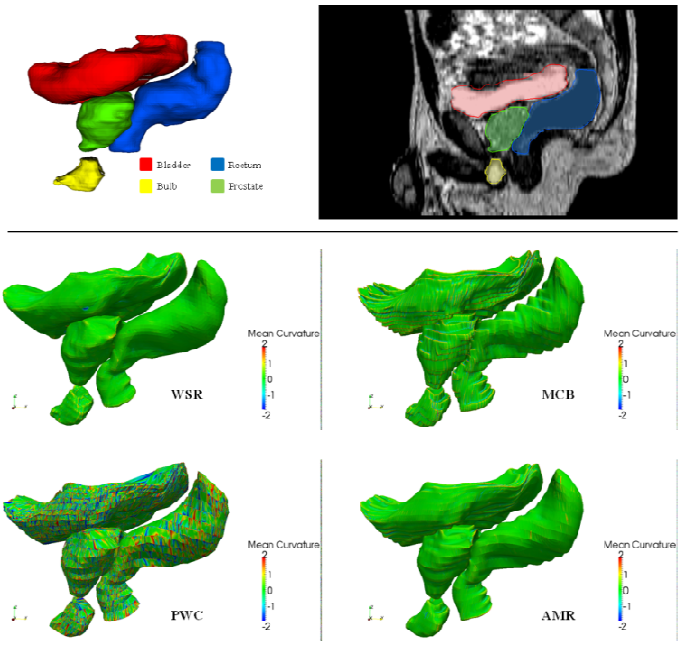


Fig. 2. Representation of the mean curvature of the deformed surfaces of the four organs, reconstructed with WSR, MCB, PWC and AMR. First row: model of the four organs and their area on a 2D sagittal plane.

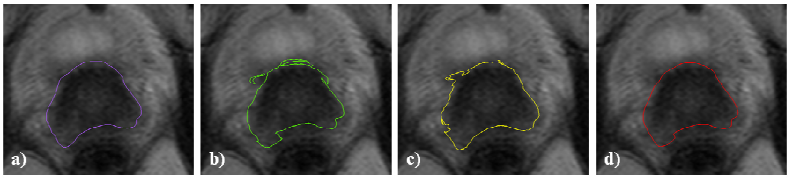


Fig. 3. A 2D representation of the deformed contours of the prostate on the correspondent T2w-MR image. a) WSR; b) MCB; c) PWC; d) AMR.

Among the four presented methods, PWC was the less robust, since it failed in numerous cases, thus preventing a successful contour propagation. For this reason, results of accuracy of PWC cannot be directly compared with the other methods and they were removed from Table 2. Regarding WSR, MCB and AMR, accuracy was quite different in bladder and bulb, whereas was practically the same in prostate and rectum. In general, WSR reported lower values of D_mean and D_max with respect to MCB and AMR.

4 Discussion and Conclusion

In this work a method based on wavelet was proposed to generate surfaces useful for contour propagation in RT; this method was compared with other three approaches already used in this context. Results have shown that globally the mesh quality and the accuracy of

Table 2. Accuracy of contour propagation in bladder, bulb, prostate and rectum using different surface reconstruction methods. Data are presented as mean ± st. dev. over the five subjects.

	D_mean [mm]	D_max [mm]	D_mean [mm]	D_max [mm]
	Bladder		**Prostate**	
WSR	1.65 ± 0.46	11.10 ± 3.33	1.59 ± 0.45	7.45 ± 2.56
MCB	1.55 ± 0.48	14.78 ± 6.65	1.57 ± 0.41	8.16 ± 2.47
AMR	1.80 ± 0.57	11.24 ± 3.02	1.62 ± 0.47	7.71 ± 2.44
	Bulb		**Rectum**	
WSR	1.18 ± 0.24	6.22 ± 2.35	1.95 ± 0.45	15.82 ± 5.24
MCB	1.27 ± 0.24	5.45 ± 1.77	2.19 ± 0.54	16.87 ± 5.04
AMR	1.42 ± 0.31	5.35 ± 1.67	2.19 ± 0.57	16.10 ± 6.19

contour propagation obtained with the proposed method are comparable, or even better, than the others.

Although all the considered approaches make use of intrinsically different smoothing paradigms, distinctive features and critical aspects of the resulting meshes are highlighted prior to deformation in terms of mesh quality. The analysis of mesh quality metrics has reported a very high shape regularity of the triangles of WSR and AMR surfaces; MCB, which presented the best values of AR percentages, revealed a staircase effect in curvature, not observable by the AR. It is likely that the staircase effect observed on MCB meshes can be reduced by introducing an anisotropic gaussian-smoothing kernel, however this would require a specific and ad-hoc tuning process of the pre-filtering step. PWC, presenting the worse yield both in AR and curvature, resulted in inadequate triangles. This suggests that surfaces reconstructed with WSR and AMR have a more regular lattice of the triangular mesh, resulting in more reliable organ surfaces.

The accuracy of contour propagation is comparable with other state-of-art works in the pelvic district [18], [19]. These results highlighted that the contribution of the registration error is larger than the effect of the surface reconstruction method, as can be noted from the quantitative analysis, where distances were similar over the considered methods. However, this second aspect has an impact on the final yield of the 2D deformed contours. In fact, as can be seen in Fig. 3, the MCB and PWC methods resulted in irregular contours, while WSR and AMR showed a smoothed shape. This is particularly true for bladder and prostate, whereas smoothed contours were found in rectum and bulb, due to a lower variability along the z-axis, which minimizes the staircase effect of the MCB.

Lastly, runtime performances showed similar values among the considered methods. Overall, a mean processing time of <1.0 s is observed for surface generation and mesh propagation together in each evaluated structure using the described configuration.

In conclusion, the WSR proposed method was able to generate qualitatively optimal surfaces, which can be used in the contour propagation framework with an accuracy in line with the state-of-art works. No limiting assumption, nor any ad-hoc tuning of the parameters is performed with the proposed approach. These results are comparable with those obtained with AMR; nonetheless, WSR has also the advantage of being open-source and can be easily integrated in Elastix.

Acknowledgements. The study was partially funded by "Ministero degli Affari Esteri e della Cooperazione Internazionale, Direzione Generale per la Promozione del Sistema Paese (MAECI - DGSP)". Fondazione Italo Monzino is acknowledged for supporting clinical research in prostate cancer. AIRC is aknowledged for supporting one of the authors (TG), grant AIRC N-14300.

References

1. Faggiano, E., Fiorino, C., Scalco, E. et al.: An Automatic Contour Propagation Method to Follow Parotid Gland Deformation during Head-and-Neck Cancer Tomotherapy. Phys. Med. Biol., **56** (2011) 775-791
2. Thor, M., Petersen, J. B., Bentzen, L. et al.: Deformable Image Registration for Contour Propagation from CT to Cone-Beam CT Scans in Radiotherapy of Prostate Cancer. Acta Oncol., **50** (2011) 918-925
3. Hardcastle, N., Tomé, W. A., Cannon, D. M. et al.: A Multi-Institution Evaluation of Deformable Image Registration Algorithms for Automatic Organ Delineation in Adaptive Head and Neck Radiotherapy. Radiation Oncology, **7** (2012) 90
4. Gu, X., Dong, B., Wang, J. et al.: A Contour-Guided Deformable Image Registration Algorithm for Adaptive Radiotherapy. Phys. Med. Biol., **58** (2013) 1889
5. Lorensen, W. E., & Cline, H. E.: Marching Cubes: A High Resolution 3D Surface Construction Algorithm. **21** (1987) 163-169
6. Amenta, N., Choi, S., Kolluri, R. K.: The Power Crust. (2001) 249-266
7. Manson, J., Petrova, G., Schaefer, S.: Streaming Surface Reconstruction using Wavelets. **27** (2008) 1411-1420
8. Deurloo, K. E., Steenbakkers, R. J., Zijp, L. J. et al.: Quantification of Shape Variation of Prostate and Seminal Vesicles during External Beam Radiotherapy. International Journal of Radiation Oncology* Biology* Physics, **61** (2005) 228-238
9. Wang, H., & Fei, B.: Nonrigid Point Registration for 2D Curves and 3D Surfaces and its various Applications. Phys. Med. Biol., **58** (2013) 4315
10. Moriconi, S., Scalco, E., Broggi, S. et al.: High Quality Surface Reconstruction in Radiotherapy: Cross-Sectional Contours to 3D Mesh using Wavelets. **In press** (2015)
11. Schaefer, S., & Warren, J.: Dual Marching Cubes: Primal Contouring of Dual Grids. **24** (2005) 195-201
12. Taubin, G., Zhang, T., Golub, G.: Optimal surface smoothing as filter design. Springer (1996)
13. Stalling, D., Zöckler, M., Sander, O. et al.: Weighted labels for 3D image segmentation. Konrad-Zuse-Zentrum für Informationstechnik Berlin (1998)
14. Klein, S., Staring, M., Murphy, K. et al.: Elastix: A Toolbox for Intensity-Based Medical Image Registration. Medical Imaging, IEEE Transactions on, **29** (2010) 196-205
15. Rueckert, D., Sonoda, L. I., Hayes, C. et al.: Nonrigid Registration using Free-Form Deformations: Application to Breast MR Images. Medical Imaging, IEEE Transactions on, **18** (1999) 712-721
16. Knupp, P. M.: Algebraic Mesh Quality Metrics for Unstructured Initial Meshes. Finite Elements Anal. Des., **39** (2003) 217-241
17. Dyn, N., Hormann, K., Kim, S. et al.: Optimizing 3D Triangulations using Discrete Curvature Analysis. Mathematical methods for curves and surfaces, **28** (2001) 135-146
18. Gardner, S. J., Wen, N., Kim, J. et al.: Contouring Variability of Human-and Deformable-Generated Contours in Radiotherapy for Prostate Cancer. Phys. Med. Biol., **60** (2015) 4429
19. Varadhan, R., Karangelis, G., Krishnan, K. et al.: A Framework for Deformable Image Registration Validation in Radiotherapy Clinical Applications. J. Appl. Clin. Med. Phys., **14** (2013) 4066

Organ Contour Adaptor to create new structures to use for adaptive radiotherapy of cervix cancer using Matlab Bridge and 3DSlicer / SlicerRT

Y. Seppenwoolde[1,2], M. Daniel[2], H. Furtado[2,3], D. Georg[1,2]

[1] Christian Doppler Laboratory for Medical Radiation Research for Radiation Oncology, Medical University Vienna, Austria
[2] Department of Radiation Oncology, Medical University Vienna, Austria
[3] Center for Medical Physics and Biomedical Engineering, Medical University Vienna, Austria

Abstract. In commercially available clinical radiotherapy software there is no turn-key solution for adaptive radiotherapy. For example the applications that are designed for organ contouring and treatment plan optimization do not provide solutions to generate new contours based on a (non-rigid) organ motion model. To interpolate linearly between two extreme positions and/or filling states of patient's organ contours, a MATLAB-based tool was integrated into 3DSlicer / SlicerRT using the Matlab Bridge, allowing generation of intermediate uterus as well as bladder contours that are used for treatment plan optimization in adaptive radiotherapy of cervix cancer.

Keywords: Adaptive radiotherapy, workflow automation, Slicer, Matlab.

1 Introduction

For patients with cervical cancer, despite the increased accuracy of irradiation techniques, sparing of bladder, rectum and small bowel is still challenging because all organs at risk (OAR) in the pelvic area change shape and position on a daily basis due to variations in filling (Fig. 1). With the introduction of cone-beam CT scanners that are mounted directly at the linear accelerator, it became possible to observe these changes of internal organ configurations of patients during each treatment fraction. Theoretically, this enables re-adaptation of plans according to tumour shrinkage and changes in OAR morphology, resulting in reduction of toxicity [1,2]. Full online plan adaptation requires that re-delineation, re-optimizing of dose distributions and repetition of all legally required quality assurance steps should be performed in a few minutes. These workload intensive procedures would require a high degree of automation and workflow-integration that is currently absent in off-the-shelf products.

Nonetheless, by finding a well-balanced compromise between full automation and degree of plan adaptation, it is possible to apply a simplified scheme of adaptation that provides improved treatment.
Based on our own experience and that of other research groups [3], patients can be divided into two groups: the first group consists of patients who show uterus motion as a function of bladder filling (called “Movers”) and the second group are those patients whose uterus position stays relatively stable regardless of bladder volume (“Non-Movers”). With a model for the uterus position, a pre-determined set of plans for can be constructed for the “Movers” and the most appropriate treatment plan can be selected on a daily basis, while for the “Non-Movers” a single plan will suffice (See Fig. 2).
Although an empty bladder might be more reproducible and easy to obtain for the patients, it is favourable for the patient to have a full bladder because it pushes the bowel out of the high dose area. In addition, a full bladder will increase the distance of the ventral bladder wall to the high dose area. Patients are routinely instructed to have a full bladder at the time of treatment by means of a standardized drinking protocol. This in order to ensure that we are able to irradiate the largest portion of fractions with the half full to full library plan before it starts to get difficult for the patient to maintain a full bladder due to the early bladder radiation response (irritation, onset of inflammation).

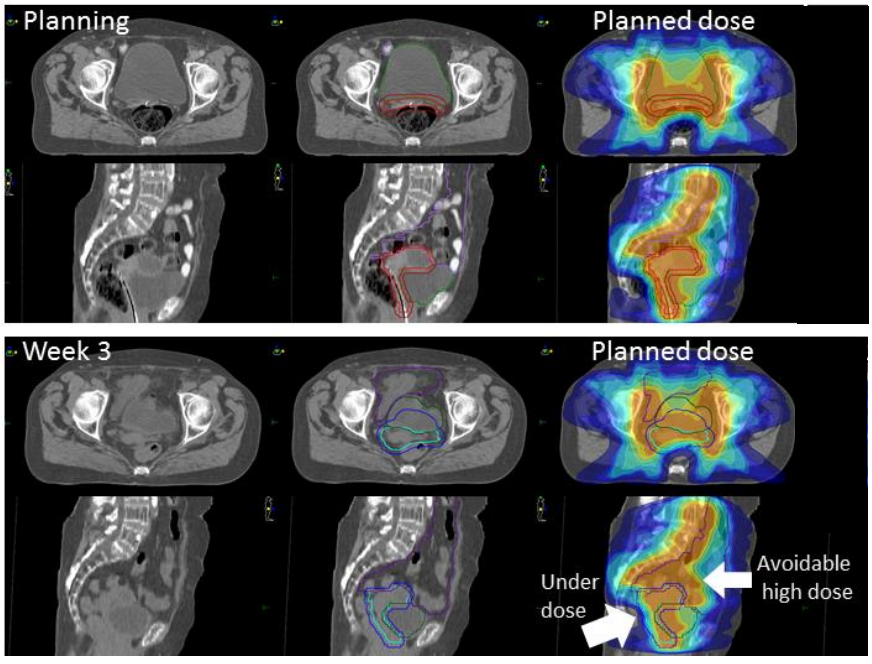


Fig. 1. Example of changing organ positions between planning and 3rd week of treatment.

The patients are asked to empty their bladder after treatment and the weight of the patient is assessed before and after this micturition to provide immediate feedback about their bladder volume and the effectiveness of the drinking protocol. However, despite these measures, bladder filling is hard to manage on an individual basis.
The patient specific relation between bladder filling and the position of the uterus can be assessed by making a set of CT scans with full and empty bladder. A two stage approach, consisting of two treatment plans, one for an empty to half full and one for half full to full bladder, has been shown to give a good level of plan adaptiveness [2], ensuring both a good tumor coverage as sparing of the surrounding healthy tissue.
Commercially available clinical software that is designed for organ contouring and treatment plan optimization does not provide solutions to generate new contours based on a motion model that interpolates between two extreme (filling) positions of an organ.
We developed a MATLAB-based tool that allows generating intermediate contours of uterus as well as bladder, according to the available bladder volumes. Its main purpose was to interpolate linearly between two extreme positions and/or filling states of patient's organ contours. Non-rigid deformation between one organ position and the other is made by matching the outer contour of both structures. To facilitate data handling and DICOM import/export options, the Matlab code is integrated to 3DSlicer / SlicerRT [4,5] by using the MatlabBridge.

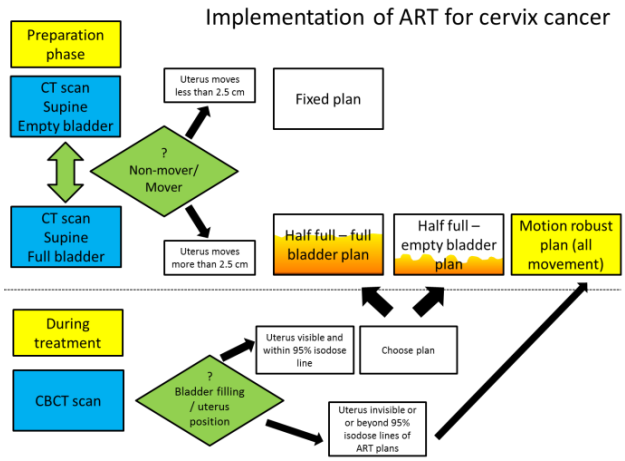


Fig. 2. Clinical workflow.

2 Materials and Methods

MatlabBridge is an extension of 3DSlicer / SlicerRT to allow running Matlab functions directly in 3D Slicer: it takes the input from the data loaded into 3DSlicer / SlicerRT, processes the Matlab routine inclusive user interaction and transfers the pro-

cessed data back to Slicer allowing visualization of the results in Slicer right after the execution is completed.
The Matlab code to interpolate contours has been developed in-house, but the generated contours could not directly be read by the treatment planning system. Loading them into 3DSlicer / SlicerRT and saving the contours solved this problem. To avoid unnecessary and time-consuming data transfer between the two programs, MatlabBridge was used to connect them. Furthermore, all 3DSlicer / SlicerRT's built-in image processing features could be used, without having to program them into Matlab, and on the other hand also all Matlab functionality can be used unlimited.
The registration of empty and full organ structures was based on the continuous point drift code package (CPD2) [6]. CPD2 simultaneously finds both the non-rigid transformation and the correspondence between two point sets without making any prior assumption of the transformation model except that of motion coherence. This method can estimate complex non-linear non-rigid transformations, and is shown to be accurate on 2D and 3D examples and robust in the presence of outliers and missing points. For our specific purpose, an interpolated half full contour was created by using an interpolation factor (Fig. 2) that could be automatically calculated based on detected bladder volumes, or could be user defined. The interpolation factor depends on the bladder filling as measured in the available full and empty bladder CT scans and does not necessarily have to be 50% because sometimes a considerable remaining volume is present in the empty bladder scan, or the bladder in the full bladder scan is not completely full (as often can be estimated based on the patient anatomy as appearing in the scans). In case one of the two CT scans of the patient has a too small full bladder contour, or a too large empty bladder contour, extrapolated contours can be created as well (to some extent and with sanity check of the results). This reduces the need for the patient to be rescanned and reduces patient radiation exposure. The (extra- and) interpolated contours were added to an existing DICOM file and transferred to the commercial treatment planning system (Monaco, Elekta AB, Stockholm, Sweden). Connectivity to other planning systems can be easily added.

3 Results

For adaptive radiotherapy for patients with Cervix cancer, a two stage approach, consisting of two treatment plans, one for an empty to half full and one for half full to full bladder was found to increase target coverage (Table 1). For the creation of an intermediately filled bladder contour and the corresponding intermediate uterus contours out of two CT scans with full and empty bladder, the new software was used.
Some results and the integration of the in-house developed Matlab code into 3DSlicer / SlicerRT is shown in Fig. 3. Newly generated contours can be checked directly after creation using the corresponding CT images in the 3DSlicer / SlicerRT user interface before they are exported to the clinical treatment planning system. The whole procedure takes between 30 and 45 minutes. With the newly created structures we were able to create two different treatment plans (Fig. 4). With daily CBCT scans that are made just before the treatment starts with the patients at the treatment table of the

linear accelerator, the current position of bladder and uterus can be determined and the most appropriate dose distribution can be chosen. The procedure was tested for 3 patients and a "sanity" check on the produced contours was performed by comparing the generated intermediate structures by CBCT scans that had similar bladder volumes. All intermediate structures created by the system were considered to be anatomically logical with a clinically relevant level of accuracy. The effect of the new structures on the dose distributions is shown in Table 1. Especially for the Mover patient (P3, Fig. 4), the target coverage increased a lot, by the cost of higher dose to the bladder. For the non-Mover patients the differences are smaller, although for P1 the dose to the bowelbag decreased significantly due to ART.

		Target V<42.75(%)		Bladder V30 (%)		Rectum V30 (cc)		Bowelbag V30 (cc)	
		average	SD	average	SD	average	SD	average	SD
P1	Clinical	8	5	52	9	85	21	601	180
	ART combi	4	4	53	9	71	20	393	183
P2	Clinical	0	0	75	4	30	10	392	75
	ART combi	0	1	79	5	27	9	346	75
P3	Clinical	22	15	49	4	53	24	447	148
	ART combi	1	3	86	15	53	25	447	139

Table 1. Dose-volume parameters for radiotherapy treatment plans that were made using the newly created structures, compared to the clinical pan that was based only on the planning-CT scan. For the target, the relative volume that received less than the prescribed dose of 42.75 Gy was calculated, for the OARs, the relative (bladder) or absolute (rectum and bowelbag) volume that received more than 30 Gy was calculated. P1 and P2 were non-movers, P3 was a mover. The average values were determined by evaluating the planned dose distributions for contours obtained from daily CBCT scans.

4 Conclusions

For adaptive radiotherapy for patients with Cervix cancer, a two stage approach, consisting of two treatment plans, one for an empty to half full and one for half full to full bladder will be implemented soon in our clinic. 3DSlicer's MatlabBridge enabled us to integrate in-house developed Matlab code in a way so that the image viewing features and DICOM in- and export routines of SlicerRT could be used, combined with the more flexible programming options regarding contour deformation of Matlab. Especially the reviewing option for the newly created contours provided us with an extra safety step before importing the generated contours directly into the clinical planning system.

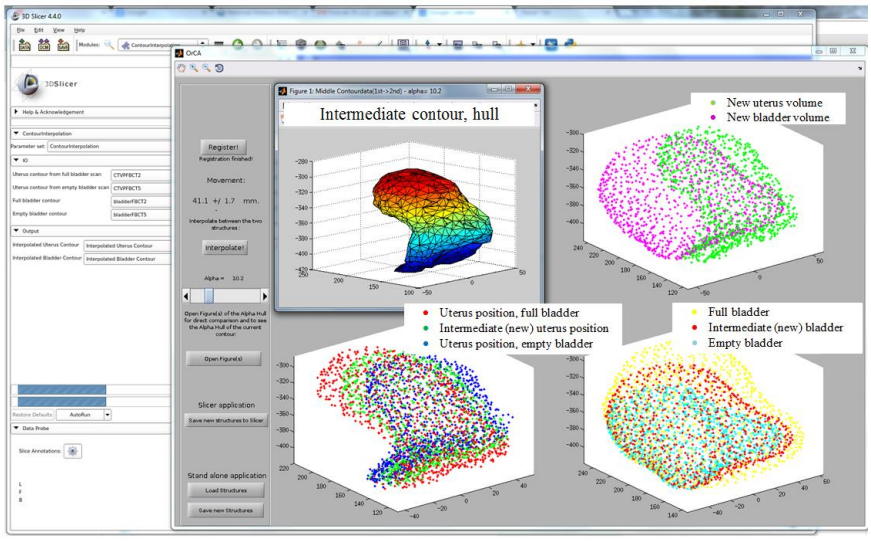


Fig. 3. Screenshot of the user-interface of 3DSlicer / SlicerRT and Matlab screens. Point clouds of input uterus and bladder contours and interpolated middle uterus and bladder contour are shown. New slices are determined by the triangulated surfaces. DICOM export is done using 3DSlicer / SlicerRT modules and user interface.

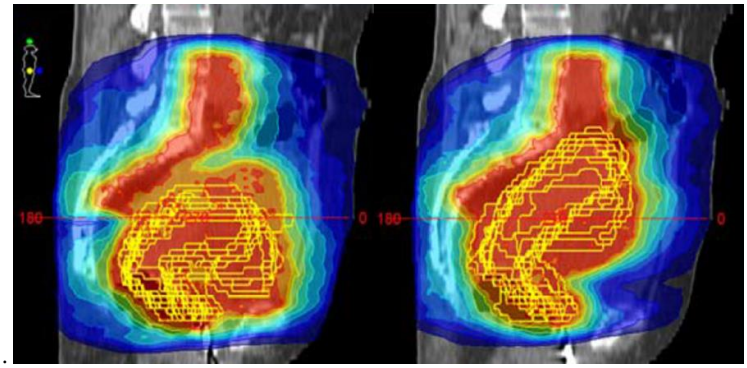


Fig. 4. Example of two optimized dose distributions (color wash) for the motion model of the uterus for empty to half-full (left panel) and half-full to full (right panel) bladder. The corresponding uterus positions as measured by the daily CBCT (yellow) for all treatment fractions of one patient are overlaid.

Acknowledgments.

The financial support by the Federal Ministry of Science, Research and Economy and the National Foundation for Research, Technology and Development is gratefully acknowledged.

References

[1] Bondar L, Hoogeman M, Mens JW, Dhawtal G, De Pree I, Ahmad R, et al. Toward an individualized target motion management for IMRT of cervical cancer based on model-predicted cervix-uterus shape and position. Radiother Oncol 2011;99:240–5.

[2] Heijkoop S, Langerak T, Quint S. Clinical Implementation of an Online Adaptive Plan-of-the-Day Protocol for Nonrigid Motion Management in Locally Advanced Cervical Cancer IMRT. Int J … 2014;90:673–9.

[3] Ahmad R, Hoogeman MS, Bondar M, Dhawtal V, Quint S, De Pree I, et al. Increasing treatment accuracy for cervical cancer patients using correlations between bladder-filling change and cervix-uterus displacements: Proof of principle. Radiother Oncol 2011;98:340–6.

[4] Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging 2012;30:1323–41.

[5] Pinter C, Lasso A, Wang A, Jaffray D, Fichtinger G. SlicerRT: Radiation therapy research toolkit for 3D Slicer. Med Phys 2012;39:6332.

[6] Myronenko A, Song X. Point set registration: Coherent point drifts. IEEE Trans Pattern Anal Mach Intell 2010;32:2262–75.

Prediction of Geometric Variations of the Bladder in Image Guided Radiotherapy of Prostate Cancer

Richard Rios[1,2*], Juan. D. Ospina[1], Frederic Commandeur[1], Caroline Lafond[3], Oscar Acosta[1], Jairo J. Espinosa[2], and Renaud de Crevoisier[3]

[1] Université de Rennes 1, LTSI-INSERM U1099, Rennes, F-35000, France
[2] Universidad Nacional de Colombia, Facultad de Minas, Medellín, Colombia
[3] Département de Radiothérapie, Centre Eugène Marquis, Rennes, F-3500, France
[*] rriospa@unal.edu.co

Abstract. In prostate cancer radiotherapy, inter-fraction organ variations introduce geometrical uncertainties that cause that the delivered dose to healthy organs to considerably differ from the planning dose distribution. The aim of this work is to propose a new model based on principal component analysis and mixed-effects model to describe and predict the bladder motion and deformation during treatment by separating intra and inter-patient variability. Two population databases are used for training and validating the motion/deformation model. Using the first database, a set of directions of maximum organ variability is determined to characterize the organ shape and variation. Later on, a mixed-effect model is fit to predict geometrical variability along each direction. Two validation phases are carried out to assess the performance of the model: in the first phase, the set of directions of geometrical variability are evaluated using a leave-one-out cross validation over the first database; in the second phase, the prediction performance of the model is assessed by comparing the estimated motion/deformation region with the region obtained with the available images of the second database.

1 Introduction

In prostate cancer radiotherapy during the treatment planning a single computed tomography (CT) is usually used to design the planned dose distribution and to determine the toxicity effects in the healthy organs called organ at risk (OAR). However, the OARs exhibit considerably internal organ motion and deformation during the treatment, which may lead to considerably difference between the actual and planned delivered dose ([7]). Thus, geometrical uncertainties can lead to dosimetric consequences that may impact on the tumor control and normal tissue complication probabilities ([8]). Therefore, quantify and characterize the inter-fraction organ variability during the treatment is important to design robust treatment plan with less side effects in the OARs.

Several approaches have been developed to quantify and characterize geometrical uncertainties produced by inter-fraction organ variations. Some of these approaches are: serial imaging measurements ([10]), fiducial markers ([3]), OAR margins ([9]), rigid-body motion ([12]), parametrization of the organ structure ([6]), biomechanical models ([13]), and statistical models ([11], [1]). However, a promising approach is given by using weighted scenarios of fundamental directions of patient's geometrical variability

([11], [1]). These directions of variability called *motion/deformation modes* are obtained by applying *principal component analysis* (PCA) to a data set of (pre-)treatment organ geometries, which are parametrized by sets of *corresponding surface points* ([2]).

To model anatomical organ variation, PCA based model was initially applied by Shön et al. (2005) to the rectum, bladder and prostate, where intra-subject variability was just considered to developed a model. However, this approach was limited by the required number of CT scans per patient for developing the model. Therefore, Budiarto et al. (2011) proposed a method that considered a population database of prostate and seminal vesicles, where the aim was to describe the geometrical uncertainties for a new patient by using a small number of CT scans. The developed method is based on the assumption that organs move and deform on a similar way despite the variation of the shape of a specific organ over a population. However, the prediction of inter-fraction organ variations is still not properly approached. Therefore, in this paper we propose a model that predicts the intra-fraction geometrical variations by separating and quantifying the inter and intra-patient variability. It allows us to predict the patient's motion/deformation region during the treatment by means of an image that describes the occupation probability of the organ and is called as *probability map*. Basically, given a new patient with a planned delineation, the clinical perspective behind this work is to quantify the geometrical uncertainties produced by inter-fraction organ variations, and to analyze its effects in the relation between dose distribution and morbidity prediction in the OARs.

2 Materials

For this study, the method is applied to two population databases treated for prostate cancer with external beam RT. The first database consists of 20 patients, where for each patient a planning CT scan and several (ranging from 5 to 8) on-treatment CT scans are available. Similarly, the second database consists of 3 patients, where for each patient a planning CT scan and several (ranging from 35 to 39) on-treatment CBCT scans are available. For each patient, the bladder, rectum, prostate and seminal vesicles (SV) are manually contoured following the same protocol. For each patient, the prescribed dose is computed in a standard Treatment Planning System (TPS) step and then is re-sampled into the CT native space. Finally, the size of the images is $135 \times 215 \times 55$ with 1 mm pixels and 2 mm thick slices.

In this paper, all the segmented bladders are brought into the same spatial referential system by selecting a reference patient. Thus, all the patients with his organ delineations are rigid registered on this reference patient by aligning the barycentre of the prostates. Then, the segmented bladders are propagated using the same transformations. Finally, each segmented bladder is reconstructed with a triangulated surface points. The triangulation process is initially done by finding the point with maximum distance to the bladder surface; and later on, a regular meshing process is carried out with constant sampling that ensure the corresponding points. Therefore, all the segmented bladder are discretized with 64442 points over its surface.

3 Methods

The proposed method consists in two phases of dimensionality reduction and a fitting process. An initial phase of dimensionality reduction is made by coding each segmented bladder into a vector of spherical harmonics coefficients. Then, a second phase is done by applying PCA to a population of bladders represented in the spherical harmonic space. Finally, a mixed-effects model is fit to predict the motion/deformation patterns of a training set.

3.1 Spherical Harmonics

In mathematics, spherical harmonics (SP) are solutions of Laplace's equation expressed in spherical coordinate system and define an orthonormal base of the solution space. In several applications based on medical images, they are used to model anatomical shapes with high accuracy and dimensionality reduction ([4]). Then, any organ shape $f \in \mathbb{R}^3$ in spherical coordinates can be approximated as a linear combination of spherical harmonics:

$$f(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-L}^{L} c_l^m Y_l^m(\theta,\varphi) \approx \sum_{l=0}^{L}\sum_{m=-l}^{l} c_l^m Y_l^m(\theta,\varphi), \quad 0 < L < \infty \tag{1}$$

where l denotes the harmonic degree, $Y_l^m(\theta,\phi)$ is the spherical harmonic function of degree l and azimuthal degree m, and the coefficients c_l^m are the coordinates in the space spanned by spherical harmonics basis. The coefficients c_l^m could be obtained as the projection of the sampled points function onto the spherical harmonics basis, *i.e*, $c_l^m =< f(\theta,\phi, Y_l^m(\theta,\varphi) >$. Therefore, fixing a harmonic degree to L, the bladder's surface at the *j-th* CT scan of the *i-th* patient denoted by $x_{i,j}$ can be represented in spherical harmonics coordinates as follows:

$$C_{ij} = \left[c^0_{ij,0}\ c^0_{ij,1}\ \cdots\ c^L_{ij,L}\ c^{-1}_{ij,1}\ \cdots\ c^{-L}_{ij,L}\right]^T \in \mathbb{R}^p \tag{2}$$

where $p = L^2 - (L^2 - L)/2$.

3.2 Principal Component Analysis

In this study, our motion/deformation model is based over some assumptions that are already stated in the methodology developed by Budiarto et al. (2011) and Shön et al. (2005), which are described as follows: first, it is assumed that the organs move and deform according to a limited number of variability directions that are imposed by the body anatomy; second, it is assumed that the inter-fraction geometrical organ variation is random along the set of variability directions; third, although the organs may be different in sizes and shapes from one patient to other, it is assumed that the directions of variability are similar for all the patients.

In order to find the directions of maximum variability, a data matrix of bladders represented in spherical harmonics is build up in order to obtain a *singular value decomposition*(SVD), where the aim is to have a sample that describe all the possible

shapes and sizes of the *bladder space*. Thus, consider the bladder structure of the *i-th* patient at the *j-th* CT parametrized as C_{ij}, the data matrix would be defined as follows:

$$\mathbf{C}_{snapshot} = \left[C_{10}\ C_{11}\ \cdots\ C_{1j_1}\ C_{21}\ \cdots\ C_{n1}\ \cdots\ C_{nj_n}\right] = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{3}$$

where j_i is the number of observations available for the *i-th* patient, $\mathbf{\Sigma}$ is the diagonal singular value matrix, and $\mathbf{U}$ and $\mathbf{V}$ are the left and right singular vectors, respectively. The column vectors of the matrix $\mathbf{U}$ are the directions of variability of the bladder population and they are called *modes* of variability. As a result, each bladder in the spherical harmonics space could be approximately expressed as an linear combination of the modes, as it is described as follows:

$$C_{ij} \approx z_{ij,1}\,\varphi_1 + \cdots + z_{ij,k}\,\varphi_q, \quad \text{with } q << p \tag{4}$$

where the vectors φ_i are the singular vectors of matrix $\mathbf{U}$, $z_{ij,k} = C_{ij}^T \varphi_k$. Mathematically, the modes φ_i define a new coordinate system, where each bladder C_{ij} can be represented by L new coordinates $z_{ij,1}, \ldots, z_{ij,q}$.

3.3 Linear Mixed-effects model

Mixed-effects models provide a powerful tool for analyzing longitudinal data, where there are two sources of variability: the intra and inter-subject variability. Linear mixed-effects models allows to model separately both source of data variability ([5]).

Considering that $z_{ij,k}$ represents the variability measurement along the *k-th* mode of the *i-th* patient at the *j-th* delineated bladder. Then, the population database is represented as a serial of population measurements of geometrical variability as follows:

$$\{z_{11,k}, \ldots, z_{1j_1,k}\}, \{z_{21,k}, \ldots, z_{2j_2,k}\}, \cdots, \{z_{n1,k}, \ldots, z_{nj_n,k}\} \tag{5}$$

where $k = 1, \ldots, q$. Then, a mixed-effects model is proposed to describe the intra and inter-variability of the mode's measurements, as it is depicted as follows:

$$\begin{aligned} z_{ij,k} &= \mu_k + b_{i,k} + \epsilon_{ij,k}, \\ b_{i,k} \sim \mathcal{N}(0, \sigma_b^2),\ \epsilon_{ijk} &\sim \mathcal{N}(0, \sigma^2), \end{aligned} \tag{6}$$

where $i = 1, \ldots, n$ and $j = 1, \ldots, j_i$. The term μ_k is the mean geometrical variation measurement along the *k-th* mode across the population, $b_{i,k}$ is a random variable representing the deviation from the population mean of the mean for the *i-th* patient, the $\epsilon_{ij,k}$ is a random variable representing the deviation for observation j on patient i from the mean for the *i-th* patient. Finally, the variances σ_b^2 and σ^2 denote the inter and intra-patient variability, respectively.

Given a new patient with his planned delineated bladder, the representation of the organ structure in the reduced space $\hat{z} = [\hat{z}_1, \ldots, \hat{z}_q]$ is obtained, see (4). Then, the inference of the most probable structures of the organ along the treatment fractions is given by the following predictive model:

$$C = Wz, \tag{7}$$

$$C = \left[c_0^0 \; c_1^0 \; \cdots \; c_L^{-L}\right], \; W = \left[\varphi_1 \; \varphi_2 \; \cdots \; \varphi_q\right], \; z = \left[z_1 \; z_2 \; \cdots \; z_q\right]$$

$$z_k \sim \mathcal{N}(\hat{\mu}_k, \hat{\sigma}_k)$$

$$\hat{\mu}_k = \mu_k + \zeta_k \, (\hat{z}_k - \mu_k), \;\; \zeta = \frac{1}{1 + \sigma^2/\sigma_b^2}$$

$$\hat{\sigma}_k^2 = \left(1 - \frac{\sigma_b^2}{\sigma^2}\right) \frac{n}{\sigma^2 + \sigma_b^2} \left(1 - \frac{\sigma_b^2}{\sigma^2}\right) - \frac{(\sigma_b^2)^2}{\sigma^2} + \sigma^2 + \sigma_b^2$$

Sampling each normal distribution of z_k, a 3D region of motion/deformation of the patient's bladder can be predicted, where each voxel provides the probability of being occupied by the bladder during the treatment. Such 3D region is called *probability map* (PM).

3.4 Validation

The validation of the model is split in two phases in order to determine whether the result obtained can be applied to other patient outside of the population database. First, in a first phase a leave-one-out cross validation procedure is used to evaluate the population modes obtained with PCA. The aim is to determine if the modes accomplish to describe the motion/deformation patterns of a typical patient not included in the database. Thus, consider a delineated bladder C_i and its approximation $\hat{C}_i$ given by the Eq. 4, a reconstruction error e_i and a metric $\mathcal{M}$ are defined as follows:

$$e_i = 1 - \frac{\| C_i - \hat{C}_i \|}{\| C_i \|}, \;\; \mathcal{M} = \frac{1}{n} \sum_{i=1}^{n} e_i, \tag{8}$$

where the metric $\mathcal{M}$ is used to provide a quantitative measure of the approximation error in the training and validation set. Finally, in a second phase, the prediction performance of the model is assessed by using the second database. For each patient two PMs are calculated: a PM estimated by means of the bladder structures generated by the model, which is called *estimated PM*; and a second PM obtained by means of the bladder structures available from the patient, which is called *empirical PM*.

4 Results

Fig. 1 illustrates an example of the rigid registration process that is carried out over the bladder population in order put them into the same framework. The segmented bladder and prostate of an unregistered patient are illustrated before and after the registration process. The impact of the number of the spherical harmonic coefficients on the bladder reconstruction is evaluated on the bladder population. The evaluation is done in terms of the Dice Score (DS) and Hausdorff distance (HD) as a function of the level of the

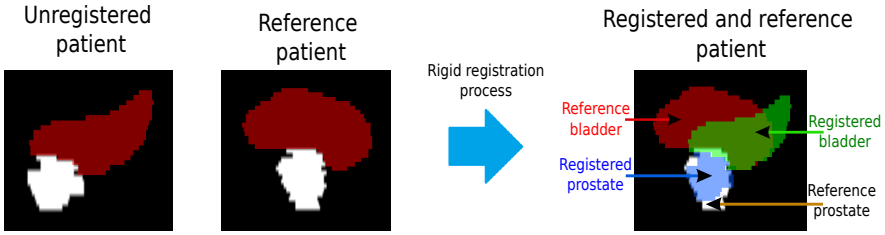


Fig. 1. Visualization of the rigid registration process that is used to put all delineated bladders in the same spatial referential system.

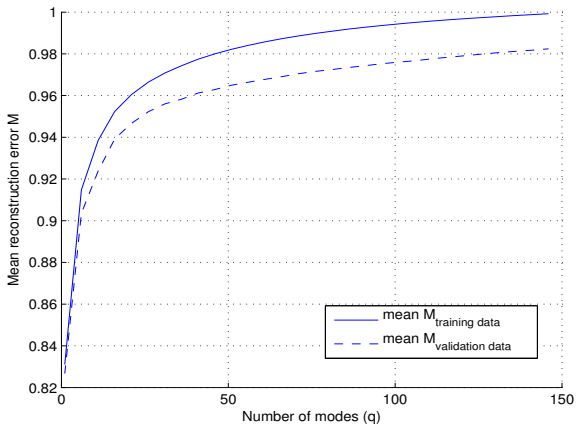


Fig. 2. Mean metric error $\mathcal{M}$ as a function of the number of modes in both training and validation database.

basis functions. A level of $L = 15$ is selected as a good bladder shape representation with a mean DS equals to 0.954 ± 0.0522 and a mean HD equals to $6.112\,mm \pm 4.997$.

Following a leave-one-out procedure over the first database, motion/deformation modes are obtained from the training population of 19 patients and evaluated with the remaining patient. Thus, SVD is applied to a data matrix $\mathbf{C}_{snapshot} \in \mathbb{R}^{p \times n}$ with n ranging between $151 - 154$ images and $p = 512$. In average, from all the 20 combinations of cutting the first database, the first mode has the biggest variance contribution with a mean value of 34.7%, and the first 40 modes covers in average 90% of the accumulated variability. Also, Fig. 2 shows that the first 40 modes, the mean reconstruction error $\mathcal{M}$ in both training and validation sets are 0.98 and 0.96, respectively.

Figure 3 illustrates the estimated and empirical probability maps that are obtained from a patient from the second database. It depicts the deformation/motion region that is estimated by means of the model after sampling the normal distribution of the vector z. Similarly, an empirical deformation/motion region is obtained for the same patient by means of the 36 segmented bladders available.

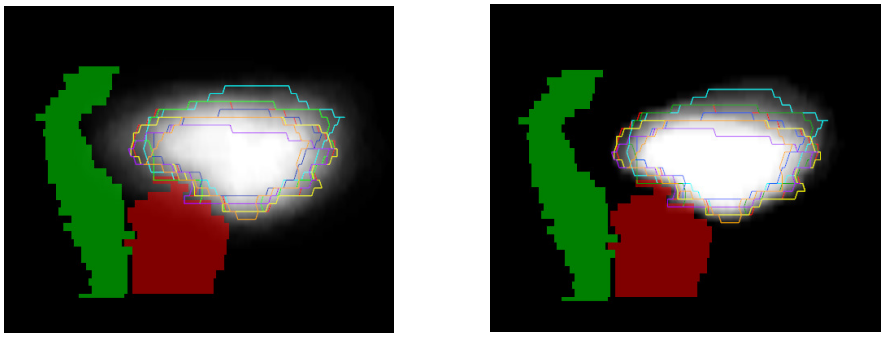


Fig. 3. The left images shows the estimated probability map using the deformation/motion model. The right images shows the empirical probability map using the available images of the patient.

5 Discussions and Conclusions

We have proposed a methodology for predicting inter-fraction geometrical variations of the bladder in prostate cancer radiotherapy. A rigid registration process was done for mapping the population of bladders into the same space. The proposed methodology consisted in two phases of dimensionality reduction and a fitting process. Our proposed model has been validated in two phases by using two databases. In the first phase, leave-one-out cross validation was applied to validate the direction of geometrical variability by using the first database. In the second phase, the estimated motion/deformation region was compared with the region obtained with the available images of the second database.

A rigid registration process was proposed to define a spatial referential system by using the prostate barycentre of a reference patient. Although the prostate presents rigid motions during the treatment, its volume is nearly identical to the volume in the planned CT ([10]). Therefore, we have chosen the prostate barycentre because its overall structure remains stable during the course of the treatment. At difference to previous works, an initial phase of dimensionality reduction with SP was proposed, which allows a pre-reduction of the number of variables needed to represent each bladder shape. Therefore, the combination of SP and PCA achieved a significant dimensionality reduction of the number of variables needed to explain the internal structure in the population data, where a delineated bladder represented with of millions of voxels was reduced to a few number of scores.

Comparing our results with previous works ([1]), we found that a higher number of modes are needed to properly describe the variability of the data. However, there is significant difference between the organ considered in both studies, where geometrical variations by motion and deformation during treatment in prostates and seminal vesicles are less significant than those in the bladder. As a result, a lower number of modes may be required to describe the geometrical variability in a composite of prostate/seminal vesicles than in the bladder.

To summarize, a method was proposed for predicting motion/deformation region for a typical patient by using just his planned segmented bladder. Therefore, as significant

contribution of this study is to propose a population based model that accomplishes to better personalize the prediction of the inter-fraction bladder variations. In addition, after learning geometrical variations in a training population, the population based model allows to quantify *a priori* the treatment uncertainties due to internal motion and deformation of the organ. Thus, possible applications of such information could be margin evaluations, estimation of delivered actual dose, toxicity prediction and robust design of planned dose. In addition, when more CT scans are available inside an adaptive RT scheme, the model could be used to adapt the dose during the treatment by using such information to improve the personalization of the model patient-specific parameters.

References

1. Budiarto, E., Keijzer, M., Storchi, P.R., Hoogeman, M.S., Bondar, L., Mutanga, T.F., de Boer, H.C.J., Heemink, A.W.: A population-based model to describe geometrical uncertainties in radiotherapy: applied to prostate cases. Phys. Med. Biol. 56(4), 1045–1061 (Feb 2011)
2. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active Shape Models-Their Training and Application. Computer Vision and Image Understanding 61(1), 38–59 (Jan 1995)
3. Crook, J.M., Raymond, Y., Salhani, D., Yang, H., Esche, B.: Prostate motion during standard radiotherapy as assessed by fiducial markers. Radiotherapy and Oncology 37(1), 35–42 (Oct 1995)
4. Dillenseger, J.L., Guillaume, H., Patard, J.J.: Spherical harmonics based intrasubject 3-D kidney modeling/registration technique applied on partial information. IEEE Trans Biomed Eng 53(11), 2185–2193 (Nov 2006)
5. Frees, E.W.: Longitudinal and Panel Data: Analysis and Applications in the Social Sciences. Cambridge University Press, 1 edition edn.
6. Hoogeman, M.S., van Herk, M., Yan, D., Boersma, L.J., Koper, P.C.M., Lebesque, J.V.: A model to simulate day-to-day variations in rectum shape. International Journal of Radiation Oncology Biology Physics 54(2), 615–625 (Oct 2002)
7. Jaffray, D.A., Lindsay, P.E., Brock, K.K., Deasy, J.O., Tom, W.A.: Accurate accumulation of dose for improved understanding of radiation effects in normal tissue. Int. J. Radiat. Oncol. Biol. Phys. 76(3 Suppl), S135–139 (Mar 2010)
8. Kvinnsland, Y., Muren, L.P.: The impact of organ motion on intestine doses and complication probabilities in radiotherapy of bladder cancer. Radiotherapy and Oncology: Journal of the European Society for Therapeutic Radiology and Oncology 76(1), 43–47 (Jul 2005)
9. Mageras, G.S., Fuks, Z., Leibel, S.A., Ling, C.C., Zelefsky, M.J., Kooy, H.M., van Herk, M., Kutcher, G.J.: Computerized design of target margins for treatment uncertainties in conformal radiotherapy. Int. J. Radiat. Oncol. Biol. Phys. 43(2), 437–445 (Jan 1999)
10. Roeske, J.C., Forman, J.D., Mesina, C.F., He, T., Pelizzari, C.A., Fontenla, E., Vijayakumar, S., Chen, G.T.Y.: Evaluation of changes in the size and location of the prostate, seminal vesicles, bladder, and rectum during a course of external beam radiation therapy. International Journal of Radiation Oncology*Biology*Physics 33(5), 1321–1329 (Dec 1995)
11. Söhn, M., Birkner, M., Yan, D., Alber, M.: Modelling individual geometric variation based on dominant eigenmodes of organ deformation: implementation and evaluation. Phys. Med. Biol. 50(24), 5893–5908 (Dec 2005)
12. Thor, M., Apte, A., Deasy, J.O., Muren, L.P.: Statistical simulations to estimate motion-inclusive dose-volume histograms for prediction of rectal morbidity following radiotherapy. Acta Oncol. 52(3), 666–675 (Apr 2013)
13. Yan, D., Jaffray, D., Wong, J.: A model to accumulate fractionated dose in a deforming organ. International Journal of Radiation Oncology Biology Physics 44(3), 665–675 (Jun 1999)

Population-based Correspondence Models for Respiratory Motion Estimation in the Presence of Inter-fraction Motion Variations

Matthias Wilms[1], René Werner[2], Tokihiro Yamamoto[3], Heinz Handels[1], and Jan Ehrhardt[1]

[1]Institute of Medical Informatics, University of Lübeck, Lübeck, Germany
[2] Department of Computational Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[3] Department of Radiation Oncology, University of California Davis, Sacramento, USA
wilms@imi.uni-luebeck.de

Abstract. Many respiratory motion compensation approaches in radiation therapy of thoracic and abdominal tumors are guided by external breathing signals. Patient-specific correspondence models based on planning 4D data are used to relate signal measurements to internal motion. The motion estimation accuracy of these models during a treatment fraction depends on the degree of inter-fraction motion variations. Here, we investigate whether motion estimation accuracy in the presence of inter-fraction motion variations can be improved by (sub)population models, which incorporate patient-specific motion information and motion data from selected additional patients. A sparse manifold clustering approach is integrated into a regression-based correspondence modeling framework for automated identification of subpopulations of patients with similar motion characteristics. In an evaluation with repeated 4D CT scans of 13 patients, subpopulation models, on average, outperform patient-specific correspondence models in the presence of inter-fraction motion variations.

1 Introduction

Respiratory motion is a key problem in external beam radiation therapy (RT) of thoracic and abdominal tumors. In clinical practice, this problem is typically tackled by the use of technical motion compensation approaches (e.g., gating or tumor tracking) [1]. Most compensation approaches are guided by (external) breathing/surrogate signals (e.g., spirometry, skin surface displacements). Given a signal measurement, a trained correspondence model, which relates signal measurements to internal motion patterns, can be used to estimate the complex 3D motion of internal structures during the treatment [2].

Several approaches for correspondence modeling and model-based respiratory motion estimation in RT in general have been proposed over the past decade (see [2] and [3] for overviews). From a methodological point of view, most approaches

are either patient-specific or population-based. Patient-specific correspondence models [4–6] are built prior to the treatment based on planning 4D CT/MRI data of the specific patient. In contrast, (inter-patient) population-based models are build from motion data of several patients and typically model the mean motion (+ deviations from the mean) [7, 8]. Population models can, e.g., be adapted to unseen patients to allow for model-based motion estimation in the absence of patient-specific 4D planning data.

In most population-based approaches, the complete population is used to compute the model. Hence, the possible existence of subpopulations that optimally resemble the unseen patient's true motion is ignored. Recently, approaches for selecting/determining subpopulations of patients with motion patterns most similar to those of the unseen patient have been proposed [9–12].

Despite these developments, most of the time patient-specific models still outperform population models. However, the accuracy of a patient-specific correspondence model built on planning data and applied during a treatment fraction highly depends on the degree of inter-fraction motion variations [6]. We hypothesize that a patient's inter-fraction motion variability resembles the inter-patient motion variations observed in a subpopulation of patients with similar breathing characteristics. In this work, we, therefore, investigate whether the motion estimation accuracy in the presence of inter-fraction motion variations can be improved by correspondence models, which incorporate both, patient-specific motion information obtained from a planning data set as well as motion information from selected additional patients. To automatically identify subpopulations of patients with similar motion characteristics, we propose using a sparse manifold clustering approach, which is integrated into a regression-based correspondence modeling framework. The proposed approach is evaluated by means of repeated 4D CT scans of 13 lung cancer patients.

2 Methods

The goal of our approach is to build a correspondence model for a patient $p = 0$ with available planning data acquired prior to the treatment. We assume the planning data of this patient to consist of a 4D CT data set $(I_{0,j})_{j=1}^{n_{ph}}$ with n_{ph} 3D images $I_{0,j} : \Omega_0 \to \mathbb{R}$ ($\Omega_0 \subset \mathbb{R}^3$) capturing the patient's anatomy at breathing phases j and corresponding n_{sur}-dimensional surrogate signal measurements $(\zeta_{0,j})_{j=1}^{n_{ph}}$ with $\zeta_{0,j} \in \mathbb{R}^{n_{sur}}$. We furthermore assume a population of n_{pat} other patients to be given. The data available for each of these population patients $p \in \{1, \ldots, n_{pat}\}$ also consists of a 4D CT image sequence $(I_{p,j})_{j=1}^{n_{ph}}$ with $I_{p,j} : \Omega_p \to \mathbb{R}$ and n_{ph} corresponding surrogate signal measurements $\zeta_{p,j} \in \mathbb{R}^{n_{sur}}$. For the sake of simplicity, we assume temporal correspondence between the phases j across all patients.

After explaining the preprocessing of the data (Sec. 2.1), we will briefly review our patient-specific correspondence modeling approach (Sec. 2.2), before a general population-based extension (Sec. 2.3) and the new subpopulation-based approach (Sec. 2.4) are presented.

2.1 Preprocessing

As a first step, the 4D images sequences of all patients are mapped to a common atlas space Ω_A to establish anatomical correspondence between all data sets. The atlas space is generated following the approach presented in [13] by using the population patients. After this preprocessing, the reference breathing phases $j = 1$ (here: end-inspiration) of all patients are anatomically aligned, while breathing-related anatomical differences between patients in other phases are preserved. Please note, that the surrogate signals need to be transformed accordingly.

2.2 Patient-specific correspondence modeling

Our patient-specific correspondence modeling approach [4] is based on the assumption that a linear relation between a surrogate signal measurement $\hat{\zeta}_0$ and the corresponding internal motion $\hat{\varphi}_0$ exists. The non-linear transformation $\hat{\varphi}_0 = id + \hat{u}_0 : \Omega_A \rightarrow \Omega_A$ represents the deformation of the internal structures between the reference phase $I_{0,1}$ and the breathing state represented by $\hat{\varphi}_0$. Here, $\hat{u}_0$ denotes a displacement field. Mathematically, we define the relationship between $\hat{\zeta}_0$ and $\hat{\varphi}_0/\hat{u}_0$ as

$$\hat{\mathbf{u}}_0 = \overline{\mathbf{u}} + \mathbf{B}\hat{\zeta}_0 \ , \tag{1}$$

where $\hat{\mathbf{u}}_0 \in \mathbb{R}^{3m}$ (m: number of image voxels) denotes a vector containing all elements of the displacement field $\hat{u}_0$, $\mathbf{B} \in \mathbb{R}^{3m \times n_{sur}}$ is a coefficient matrix, and $\overline{\mathbf{u}} \in \mathbb{R}^{3m}$ is the mean displacement vector.

Model training consists of learning the coefficient matrix $\mathbf{B}$ (and calculating $\overline{\mathbf{u}}$) in Eq. (1). First, internal motion data is derived from the planning data by estimating non-linear transformations $(\varphi_{0,j})_{j=1}^{n_{ph}}$ between the reference phase $I_{0,1}$ and each $I_{0,j}$ via image registration [14]. Subsequently, least squares regression is performed based on the vectorized displacements fields $(\mathbf{u}_{0,j})_{j=1}^{n_{ph}}$ of the estimated transformations and the corresponding surrogate signal measurements $(\zeta_{0,j})_{j=1}^{n_{ph}}$ to obtain $\mathbf{B}$. See [4] for further details. This model represents the breathing characteristics captured by the patient-specific planning data.

2.3 Population-based correspondence modeling

The patient-specific correspondence model built in Sec. 2.2 is extended to a general population-based model by adding internal motion information and corresponding surrogate signal measurements from all n_{pat} population patients to the training data used to estimate coeffcient matrix $\mathbf{B}$ in Eq. 1. Therefore, non-linear transformations $(\varphi_{p,j})_{j=1}^{n_{ph}}$ between the reference phase $I_{p,1}$ and all other images $I_{p,j}$ of each patient p have to be computed. Finally, the training data consists of displacement fields $(\mathbf{u}_{p,j})_{p=0,j=1}^{n_{pat},n_{ph}}$ and corresponding surrogate signal measurements $(\zeta_{p,j})_{p=0,j=1}^{n_{pat},n_{ph}}$.

As all patients are represented by the same number of samples in the training data and no weighting is applied, least-squares regression averages out differences between patients. Hence, the influence of the patient-specific planning data of patient $p = 0$ depends on the population size.

2.4 Subpopulation-based correspondence modeling

As stated before, we hypothesize that a patient's inter-fraction motion variability resembles the inter-patient motion variations observed in a (sub)population of patients with similar breathing characteristics. The general population model in Sec. 2.3 is built based on data from all population patients, which might lead to an improved estimation accuracy. However, we assume that most of the time the heterogeneity of the population/large inter-patient differences will negatively impact the accuracy. We are therefore interested in identifying a subset $S \subseteq \{1, \ldots, n_{pat}\}$ of population patients with breathing characteristics most similar to that of the unseen patient $p = 0$ to build a more accurate model.

Our idea is to determine subset S by clustering the motion of all patients with respect to their similarity. In the end, the cluster the new patient $p = 0$ belongs to is chosen as S. Motion clustering has also been used by Peressutti *et al.* [11] to identify patients with similar breathing characteristics to personalize affine population models for cardiac respiratory motion compensation. However, in our application we have to deal with non-linear deformations and, furthermore, our clustering approach based on sparse coding is more general than the one used in [11] as it is able to effectivly handle high-dimensional data with complex structure.

Motion representation We start by representing the motion of each patient $p \in \{0, \ldots, n_{pat}\}$ by a vector $\mathbf{m}_p = [\mathbf{u}_{p,EE}^T, \mathbf{u}_{p,MI}^T, \mathbf{u}_{p,ME}^T]^T \in \mathrm{R}^{9m}$. Each vector $\mathbf{m}_p$ consists of a concatenation of three selected displacement field vectors $\mathbf{u}_{p,j}$, which encode the motion between the reference phase EI (end-inspiration) and the phases of end-expiration ($j = EE$), mid-inspiration ($j = MI$), and mid-expiration ($j = ME$) extracted from the 4D CT data sets (cf. Sec. 2.3). While the motion between EI and EE serves as a general representative for the motion of patient p, hysteresis-related patterns are captured by integration of the motion between EI and MI/ME.

Motion clustering We now assume that these vectors $\mathbf{m}_p$ lie in or close to multiple low-dimensional manifolds embedded in the high-dimenional ambient space. A (spectral) clustering solely based on pairwise distances computed in the ambient space might ignore this (possible) structure and assign motion vectors belonging to different manifolds to the same cluster. In order to avoid this, we employ the Sparse Manifold Clustering and Embedding (SMCE) method [15]. The approach aims to find a small number of neighbors $\{\mathbf{m}_i\}_{i \neq p}$ of each data point/motion vector $\mathbf{m}_p$ that belong to the same manifold and, therefore, approximately span a low-dimensional affine subspace passing near $\mathbf{m}_p$. An advantage of SMCE over other approaches is that these neighbors are automatically found, without the manual selection of a neighborhood radius or a fix neighborhood size.

Next, for each motion vector $\mathbf{m}_p$ a sparse weighting vector $\mathbf{w}_p \in \mathbb{R}^{n_{pat}+1}$ is computed whose non-zero elements indicate the (inverse) distances of the selected neighbors of $\mathbf{m}_p$ to $\mathbf{m}_p$. From these vectors $\{\mathbf{w}_p\}$, a similarity matrix

$$\mathbf{W} = \left[|\mathbf{w}_0| \ \ldots \ |\mathbf{w}_{n_{pat}}| \right] \tag{2}$$

is build, which is subsequently used for k-means-based spectral clustering [16]. Here, k denotes the number of clusters to be obtained. This parameter should be chosen with respect to the low-dimensional structure of the high-dimensional data ($k \geq \#$ of manifolds).

Finally, the subpopulation of patients (including $p = 0$) specified by motion vectors belonging to the same cluster as $\mathbf{m}_0$ are used to generate a subpopulation model following the steps outlined in Sec. 2.3.

2.5 Experiments

An evaluation of the different models presented in Sec. 2.2–2.4 in the presence of inter-fraction variations of respiratory lung motion is carried out on repeated 4D CT data sets (10 breathing phases; resampled to a spatial resolution of $2.5 \times 2.5 \times 2.5\,\text{mm}$) of 13 lung cancer patients (see [17] for details). For each of these 13 patients, 2 4D CT data sets acquired at different days (Day1 and Day2) are available. 20 4D CT data sets of different patients are additionally used as population patients. In total, 46 4D CT data sets are used for the experiments.

During preprocessing (cf. Sec. 2.1), intra-patient/intra-fraction registrations restricted to the lungs are performed to estimate the respiratory lung motion of all patients in atlas space. The resulting transformations are used to build different models and serve as ground-truth motion data for evaluation. Model-based estimation accuracy is quantitatively evaluated by computing mean vector differences between a displacement field $\hat{u}_j$ estimated by a correspondence model and the ground-truth field u_j computed via registration for all inner-lung voxels. Due to the lack of available real surrogate data for the data sets used, a $n_{sur} = 2$-dimensional spirometry signal (signal value + time derivative) is simulated by an image-based analysis of the air content inside the lungs (see [4] for details).

For each of the 26 4D CT data sets of the different day cohort (13 patients with 2 repeated scans), 4 different correspondence models are built: (1) a patient-specific intra-fraction model, (2) a patient-specific inter-fraction model, (3) a population model, and (4) a subpopulation model. Each model is used to estimate the lung motion between the reference phase EI and the phases at MI, ME, and EE in the experiment-specific test data set.

Patient-specific intra-fraction model To establish a reference for performance comparison, a patient-specific intra-fraction model (cf. Sec. 2.2, [4]) is built for each data set by using a leave-out strategy. Phases at MI, ME, and EE are left out during training, respectively, and the motion between the reference phase EI and the left-out phases is estimated by the trained correspondence model.

Patient-specific inter-fraction models Patient-specific inter-fraction models (cf. Sec. 2.2) are built based on the internal motion data and the surrogate signal of the Day1 (Day2) data set to perform a model-based estimation of the motion in the Day2 (Day1) data set. Results of these models will give an impression on how well a patient-specific correspondence model build on pre-treatment planning data is able to estimate motion during a treatment fraction at a different day.

Table 1. Mean estimation errors obtained for the surrogate-based estimation of inner lung motion for the different correspondence models, given as mean±standard deviation per phase for the 26 data sets considered. Last column: mean results over all 3 phases.

	Mean estimation error [mm]			
Motion estimation	**EI → EE**	**EI → MI**	**EI → ME**	**Mean**
No motion estimation	8.59 ± 3.92	4.80 ± 2.27	7.09 ± 3.36	6.83 ± 3.58
Inter-fraction motion difference	4.06 ± 1.48	3.17 ± 0.85	3.62 ± 1.36	3.62 ± 1.30
Models:				
Patient-specific intra-fraction	1.45 ± 0.70	2.19 ± 1.02	1.47 ± 0.73	1.71 ± 0.89
Patient-specific inter-fraction	3.75 ± 1.31	2.91 ± 1.04	3.34 ± 1.14	3.34 ± 1.20
Population model	4.52 ± 1.56	2.87 ± 0.91	4.06 ± 2.08	3.82 ± 1.72
Subpopulation model	3.33 ± 0.99	2.59 ± 0.85	3.01 ± 0.93	2.97 ± 0.96

Population model For each data set, a population model (cf. Sec. 2.3) is built using a population of $n_{pat} = 44$ data sets (both data sets of the remaining 12 patients with repeated data + 20 patients with single session data) and the Day1 (Day2) data set of the specific patient. Each model is used to estimate the lung motion in the corresponding Day2 (Day1) data set.

Subpopulation model For each data set, a subpopulation model (cf. Sec. 2.4) is built using the same population as for the population models. However, a spectral motion clustering based on the Day1 (Day2) data set is carried out as described in Sec. 2.4 to identify a suitable subpopulation of patients. The $k-$means algorithm is run with $k = \{2, \ldots, 44\}$ clusters. In this work, k is retrospectively optimized for each patient with respect to the estimation error. Resulting models are used for motion estimation in the Day2 (Day1) data sets.

3 Results

Quantitative results of our evaluation are listed in Tab. 1. As expected, patient-specific intra-fraction models are giving (by far) the best results in terms of estimation accuracy. For the experiments where the motion in the different day data sets has to be estimated, the patient-specific inter-fraction models have on average a significantly higher estimation accuracy than the general population models (paired t-test; $p < 0.05$). This result supports our hypothesis that the heterogeneity of the population negatively impacts the accuracy of general population models. In contrast, the results of the subpopulation models show, on average, a significant improvement in terms of estimation accuracy compared to the results of the patient-specific inter-fraction models ($p < 0.01$). For some patients with large inter-fraction motion differences, mean improvements $>$ 1mm are achieved by using the subpopulation models instead of the patient-specific inter-fraction models (see Fig. 1). On average, 7.77 ± 10.18 patients were used

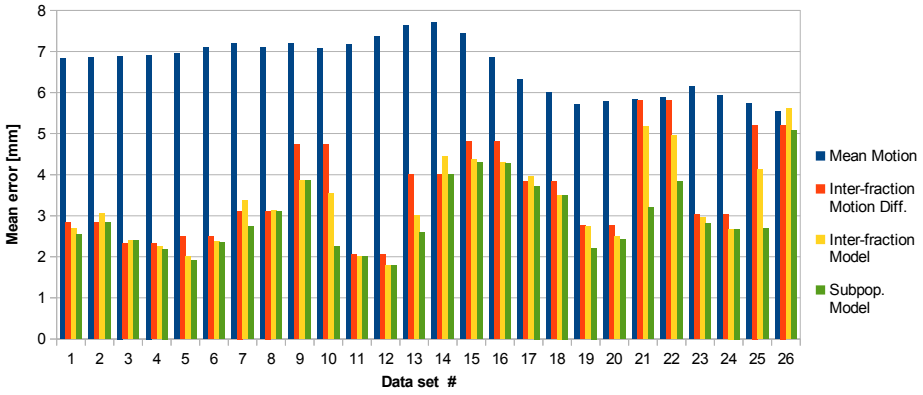


Fig. 1. Mean estimation errors of inter-fraction and subpopulation models for all 26 data sets (mean results over all 3 phases). In addition mean motion and inter-fraction motion differences are given.

for building the subpopulation models. A strong correlation (correlation coefficient of 0.85) between inter-fraction differences and mean estimation accuracy improvements achieved by the subpopulation models compared to the patient-specific inter-fraction models is observed.

4 Conclusion

Current respiratory motion compensation approaches in radiation therapy usually employ patient-specific correspondence models to relate surrogate signal measurements to internal motion patterns. In this work, an investigation on whether the motion estimation accuracy in the presence of inter-fraction motion variations can be improved by correspondence models that incorporate both patient-specific motion information obtained from a planning data set as well as motion information from selected additional patients with similar breathing motion was carried out. A sparse manifold clustering approach was employed to automatically identify subpopulations of patients with similar motion characteristics. The evaluation based on 13 patients with repeated 4D CT scans showed that these so-called subpopulation models, on average, outperform patient-specific correspondence models in the presence of inter-fraction motion variations. However, for most patients, only small differences between both models exist. Furthermore, it has to be noted that the parameter k in our clustering approach, which controls the number of patients in the chosen subpopulation, was retrospectively optimized with respect to the estimation error. Future work will therefore focus on finding ways to choose this parameter automatically. Moreover, all experiments in this work were carried out in a common atlas space. Hence, an efficient integration of this approach into a clinical workflow will be another challenge.

References

1. Keall, P.J., et al.: The management of respiratory motion in radiation oncology report of AAPM Task Group 76. Medical Physics **33**(10) (2006) 3874–3900
2. McClelland, J.: Estimating internal respiratory motion from respiratory surrogate signals using correspondence models. In: 4D Modeling and Estimation of Respiratory Motion for Radiation Therapy. Springer (2013) 187–213
3. Ehrhardt, J., Klinder, T., Lorenz, C.: Computational motion phantoms and statistical models of respiratory motion. In: 4D Modeling and Estimation of Respiratory Motion for Radiation Therapy. Springer (2013) 215 – 247
4. Wilms, M., Werner, R., Ehrhardt, J., Schmidt-Richberg, A., Schlemmer, H.P., Handels, H.: Multivariate regression approaches for surrogate-based diffeomorphic estimation of respiratory motion in radiation therapy. Physics in Medicine and Biology **59** (2014) 1147–1164
5. King, A., Buerger, C., Tsoumpas, C., Marsden, P., Schaeffter, T.: Thoracic respiratory motion estimation from MRI using a statistical model and a 2-D image navigator. Medical Image Analysis **16**(1) (2012) 252 – 264
6. McClelland, J.R., Hughes, S., Modat, M., Qureshi, A., Ahmad, S., Landau, D.B., Ourselin, S., Hawkes, D.J.: Inter-fraction variations in respiratory motion models. Physics in Medicine and Biology **56**(1) (2011) 251–272
7. Ehrhardt, J., Werner, R., Schmidt-Richberg, A., Handels, H.: Statistical Modeling of 4D Respiratory Lung Motion Using Diffeomorphic Image Registration. IEEE Trans Med Imag **30**(2) (Sep 2011) 251–65
8. Klinder, T., Lorenz, C., Ostermann, J.: Prediction framework for statistical respiratory motion modeling. Proc. MICCAI 2010 **6363** (2010) 327–334
9. Samei, G., Tanner, C., Szkely, G.: Predicting liver motion using exemplar models. In: Abdominal Imaging. Computational and Clinical Applications – 4th International Workshop at MICCAI 2012. Volume 7601 of LNCS., Springer (2012) 147–157
10. Tanner, C., Samei, G., Szekely, G.: Robust exemplar model of respiratory liver motion and individualization using an additional breath-hold image. In: Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on. (2015) 1576 – 1579
11. Peressutti, D., Penney, G.P., Kolbitsch, C., King, A.P.: Personalising cross-population respiratory motion models using anatomical features. In: Proc. of the 17th Conference on Medical Image Understanding and Analysis. (2013) 45 – 50
12. Peressutti, D., Penney, G.P., Kolbitsch, C., King, A.P.: Personalising population-based respiratory motion models of the heart using neighbourhood approximation based on learnt anatomical features. Med Image Anal **18**(7) (2014) 1015–1025
13. Ehrhardt, J., Handels, H.: Motion Estimation in Artifact-affected 4D CT Images Using Temporal Consistent Model-based Registration with Local Adaptive Weights. In: MICCAI IGART Workshop - MICCAI 2014, Boston (2014) 5–12
14. Werner, R., Schmidt-Richberg, A., Handels, H., Ehrhardt, J.: Estimation of lung motion fields in 4D CT data by variational non-linear intensity-based registration: A comparison and evaluation study. Phys Med Biol **59** (2014) 4247–4260
15. Elhamifar, E., Vidal, R.: Sparse manifold clustering and embedding. In: Proc. NIPS. (2011) 55–63
16. Ng, A.Y., Jordan, M.I., Weiss, Y., et al.: On spectral clustering: Analysis and an algorithm. Proc. NIPS (2002) 849–856
17. Yamamoto, T., Kabus, S., von Berg, J., Lorenz, C., Chung, M.P., Hong, J.C., Loo, B.W., Keall, P.J.: Reproducibility of four-dimensional computed tomography-based lung ventilation imaging. Academic Radiology **19**(12) (2012) 1554–1565

Fast automated non-linear contour propagation for adaptive head and neck radiotherapy.

Florian Weiler[1], Christoph Brachmann[1], Nadine Traulsen[2], Reinoud Nijhuis[3], Grzegorz Chlebus[1], Mark Schenk[2], Dörte Corr[1], Stefan Wirtz[1], Ute Ganswindt[3], Christian Thieke[3], Claus Belka[3], and Horst K. Hahn[1]

[1] Fraunhofer MEVIS, Universitätsallee 29, 28359 Bremen, Germany.
[2] Fraunhofer MEVIS, Maria-Goeppert-Straße 3, 23562 Lübeck, Germany.
[3] Ludwig-Maximilians-University, Department of Radiation Oncology, Munich, Germany.
florian.weiler@mevis.fraunhofer.de

Abstract. A crucial and time-consuming task in adaptive radiotherapy is the propagation of contours from an initial planning CT image to a control image taken during the course of treatment. Precise adaptation of contours for organs at risk, as well as target volumes is necessary in order to calculate an adapted treatment plan. Although several commercially available solutions exist that aim at solving this task, manual editing and correction of such automated mappings is still an inevitable requirement making the overall process tedious and time-consuming in clinical routine. We present a processing pipeline aiming at fast and fully automated propagation of contours between different datasets of an ongoing therapy. The method is based on a non-linear image registration combined with GPU accelerated contour generation. We evaluate our method by calculating Dice similarity coefficients and 3D Hausdorff distances between our results, and manually generated contours which serve as a ground truth. Additionally, we compare our results against contours mapped using a state-of-the-art commercially established contouring software.

Keywords: Adaptive radiotherapy, contour propagation, non-linear registration

1 Introduction

Adaptive radiotherapy attempts to improve the outcome of radio-therapeutic tumor treatment by adapting an original treatment-plan during the course of therapy to physiological changes occurring in the patients body. To facilitate this, CT-imaging is performed at certain stages during therapy, building the foundation for adaptation of the treatment plan. In order to calculate an adapted plan, contours of both target structures as well as organs at risk need to be transferred to the control CT image, taking into account the current anatomical situation.

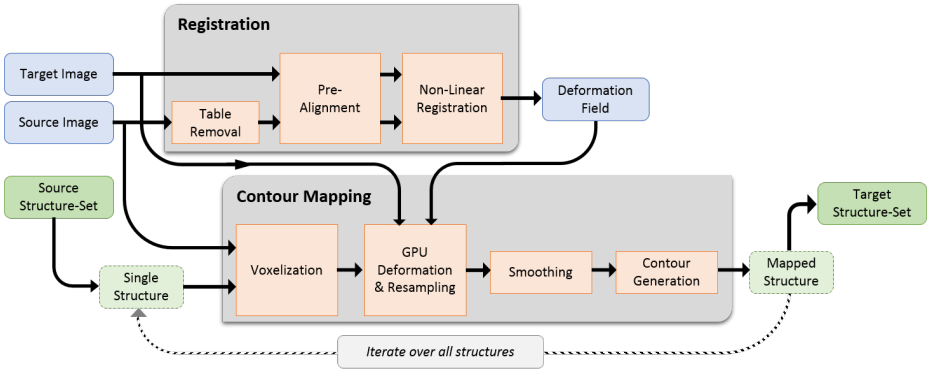


Fig. 1. Overview of the complete processing pipeline.

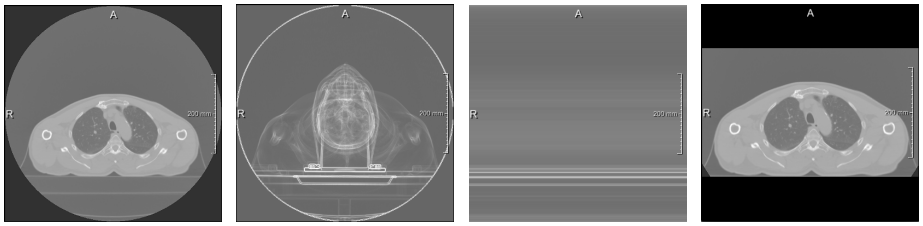
This goal can theoretically be achieved in two ways. First, the control CT could be completely re-contoured, resulting in a new set of contours, independent of the initial structures. Alternatively, the original structures can be transferred to the target image by means of non-linear image registration with subsequent transformation of all structures. The latter approach offers two significant advantages: First, the time requirement can be reduced given the fact that a well fitting set of contours builds the basis for typically required corrections. Second, the correspondence between structures from the planning image to the control image remains intact, facilitating precise calculations of volumetric changes of individual structures over time.

Both approaches are supported by current commercially available software systems such as ABAS [1]. Complete re-contouring is supported by means of multi-atlas based pre-segmentation of organs at risk, followed by inevitable manual adaptations of the resulting contours. Alternatively, the original contours can be used for a single-atlas based segmentation.

In this paper, we present a novel approach for fully automated transfer of contours from different CT images relying on a non-linear image registration with automatic patient table removal, followed by re-generation of contours of the original structures on the target image. We evaluate our method by comparing the results of the mapping process to those obtained using ABAS.

2 Methods

Our method is based on a processing-pipeline consisting of two main stages: First, the non-linear image registration and second the image based contour mapping. The registration stage consists of a pre-processing step aiming at automatic table removal, followed by a coarse linear volume alignment with subsequent non-linear image registration as presented in [4]. The mapping stage consists of structure voxelization, GPU accelerated deformation and resampling, and finally contour



(a) (b) (c) (d)

Fig. 2. A representative slice of the input image volume is displayed in (a). The projection along the z-axis results in (b), and a further projection along the x-axis in (c). (d) displays the resulting cropped output image after automatic table removal.

generation using the marching squares approach. Figure 1 gives a schematic overview of the complete pipeline.

2.1 Automatic table removal

To eliminate the influence of the CT scanning table on the results of the image registration, we have implemented a simple technique for table removal. As an exact registration of the patient table would have no impact on the addressed problem, a VOI including the patient volume only is automatically calculated. This eliminates the influence of structures outside the patient on the distance measure and also reduces computation costs effectively speeding up the registration process as the considered image domain Ω is reduced to the VOI only.

The proposed method is derived from [9], but has been adapted to images used in radiotherapy. We exploit the circumstance that the patient table is typically aligned in parallel to the direction of image acquisition. This allows to easily identify the table by performing an orthogonal projection of locally maximal image intensities along the z-axis of the input image. This sums up all table voxels, forming a very dominant line in the projected image. A second projection along the x-axis generates a one-dimensional image profile, where the maximum image intensity depicts the surface of the table. Figure 2 gives a short overview of the performed steps.

2.2 Non-linear image registration

One of our general expectation regarding the non-linear, intra-patient image registration here is, that the deformations will be considered as smooth. The future-oriented trend of standardized patient positioning in advanced head and neck radiotherapy supports this assumption. Therefore, we chose the Normalized Gradient Fields (NGF) [3] as distance measure, focusing on edges in the CT images to be registered, and a curvature regularizer, aiming for smooth deformations.

Consider $\mathcal{R} : \mathbb{R}^3 \to \mathbb{R}$ as the fixed reference image and $\mathcal{T} : \mathbb{R}^3 \to \mathbb{R}$ as the moving template image with compact support in domain $\Omega \subseteq \mathbb{R}^3$. Image registration aims to find a transformation $y : \Omega \to \mathbb{R}^3$ such that $\mathcal{T}(y)$ is similar to $\mathcal{R}$. In our specific application of contour propagation we aim to be able to propagate contours as images with the transformation y.

The variational approach used for the non-linear image registration step here, models the image registration process by the objective function $\mathcal{J}$

$$\mathcal{J} = \mathcal{D}(\mathcal{R}, \mathcal{T}(y)) + \alpha \cdot \mathcal{S}(y), \tag{1}$$

where $\mathcal{D}$ is the distance measure, $\mathcal{S}$ is the regularizer, and α is the regularization parameter, a weighting factor affecting data fit and regularity.
We used the Normalized Gradient Fields (NGF) [3] as distance measure. For $x \in \Omega$ it is given by

$$\mathcal{D}(\mathcal{R}, \mathcal{T}(y)) := \int_{\Omega} 1 - \left(\frac{\langle \nabla \mathcal{T}(y(x)), \nabla \mathcal{R}(x) \rangle_\eta}{\| \nabla \mathcal{T}(y(x)) \|_\eta \| \nabla \mathcal{R}(x) \|_\eta} \right)^2 \mathrm{d}x \tag{2}$$

with

$$\langle f, g \rangle_\eta := \sum_{j=1}^{3} f_j g_j + \eta^2 \quad \text{and} \quad \|f\|_\eta := \sqrt{\langle f, f \rangle_\eta} \tag{3}$$

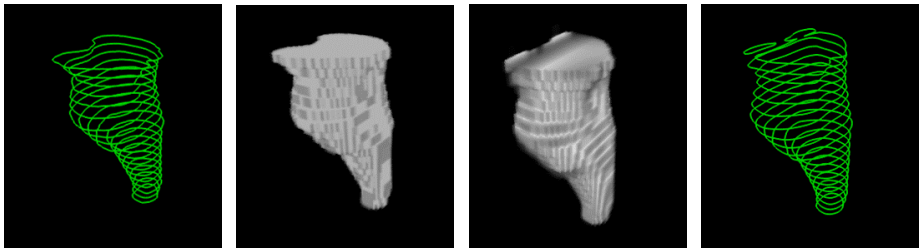
to describe image similarity. The NGF distance measure considers the angle between image gradient vectors in the reference and the template image at each point. The edge parameter η is used to define a threshold, that specifies which gradients are counted among the noise level.

For regularization purpose we used the curvature regularizer as proposed in [2], which is based on second order derivatives, penalizing the Laplacian of the displacement components. The curvature regularizer $\mathcal{S}$ is given by

$$\mathcal{S}(y) := \frac{1}{2} \sum_{l=1}^{3} \int_{\Omega} \|\Delta u\|^2 \mathrm{d}x \tag{4}$$

with the decomposition $y(x) = x + u(x)$, where u is the displacement. One beneficial characteristic of the curvature regularizer is, that it results in very smooth deformations, meeting our expectations.

A discretize-then-optimize scheme [5] is performed to optimize $\mathcal{J}$ using a quasi-Newton L-BFGS optimizer [7]. To avoid local minima, the iteration scheme is embedded into a multi-level approach [5]. Therefore, the optimization problem is solved consecutively on coarse to fine image resolution levels. In this implementation, the objective function is evaluated on 5 levels. The proposed results were obtained with an edge parameter of $\eta = 0.1$ and a regularization parameter of $\alpha = 1$.



(a) (b) (c) (d)

Fig. 3. Illustration of the mapping process: Axial source contours (a). 3D image representation (b). Deformed image representation (c). Resulting contours (d).

2.3 Contour propagation

Contour propagation is performed individually for each ROI loaded from a DICOM RTSTRUCT file by following a sequence of standard image-processing operations. The mapping is always performed from the source image to the target image, such that no deformation is applied to the target image and the resulting contours are placed in the correct coordinate system.

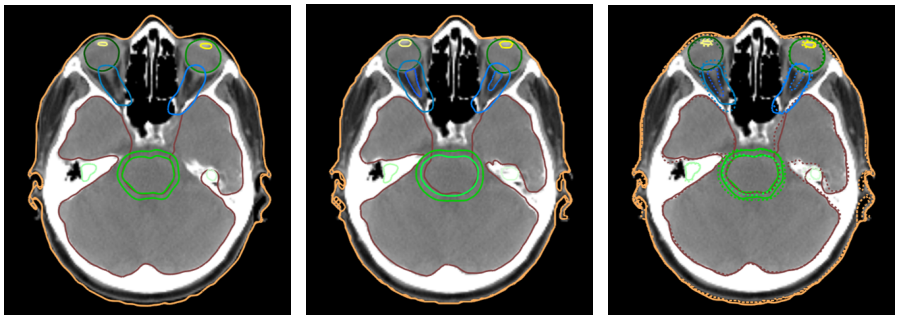
First, the axially aligned contours defining an anatomical structure are converted into a 3D image object, by rasterizing them in to the coordinate system of the corresponding reference image. Anti-aliasing is applied in order to correctly account for partial volume effects. This image mask is then deformed non-linearly using an OpenCL accelerated GPU implementation. During the deformation, resampling to the target coordinate system with tri-linear filtering is also performed on the GPU .

This results in a deformed 3D image representation of the given structure living in the coordinate system of the target image. An additional smoothing step using a Gaussian filter kernel is applied to reduce aliasing artifacts.

Finally, result contours for each structure are generated by processing the deformed image representation slice by slice. A marching-squares algorithm with interpolation is utilized for this step. The combination of anti-aliasing during rasterization, tri-linear filtering during resampling, and interpolation in the marching-squares step assures to minimize undesired artifacts in the resulting contours. Figure 3 illustrated this process.

3 Evaluation

We have implemented our pipeline in a prototypical software assistant called *CUTE*. An initial quantitative assessment of our method was presented in [6]. Here, we present results of an in depth comparison of the performance of our method over a set of 15 structures. Eight replanning CTs of five randomly selected head and neck cancer patients have been retrospectively auto-contoured using CUTE. These contours were compared to manually created contours from



(a) (b) (c)

Fig. 4. Comparison of automatically propagated contours (a), and manually delineated contours (b). Image (c) shows both contour sets simultaneously, manual contours drawn with dashed lines.

an experienced radiotherapist. In addition, the commercially available ABAS software (Atlas-based Autosegmentation, Elekta AB, Stockholm) [1] was used to automatically create another set of contours. Then, the two automatically propagated contour sets were individually compared to the manual re-contourings. Similarly to [8] Dice similarity coefficient (DSC) and 3D Hausdorff distance (HD) were chosen as evaluation metrics. Results are shown for a subset of all contours that includes the main target volumes with nearby organs-at-risk (OARs). Target volumes have been classified to low or high depending on the patient's prescriptions dose.

4 Results

Figures 5 and 6 show box plots illustrating distribution of DSC and HD metrics of the two methods for all structures of the 8 replanning CTs. The box plot for all structures shows that both ABAS and CUTE perform well with no significant differences. Our method delivers more robust results as the quantiles are closer to the median. The best DSC-values of ABAS outperform the best ones of CUTE, yet the variation and median DSC is better in contours propagated by CUTE. Comparison of box plots created for OARs and target volumes shows that ABAS performs better for latter structures, while CUTE for the former ones. Taking a look at box plots for individual structures provides an insight into which method is superior for a given structure. We found comparable results for certain structures (e.g., larynx, PTV high) as well as significant differences for others (e.g., brainstem, mandible, both parotids).

5 Discussion and Conclusion

Our evaluation results show, that the overall quality of contouring in sense of DSC and HD is quite similar for both methods. The atlas-based method performs

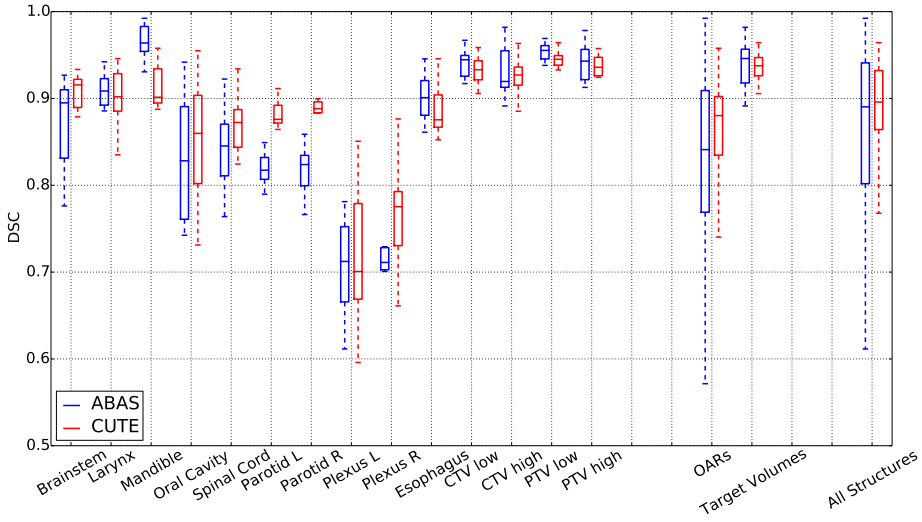


Fig. 5. Dice similarity coefficients of ABAS/CUTE contours compared to manual recontourings.

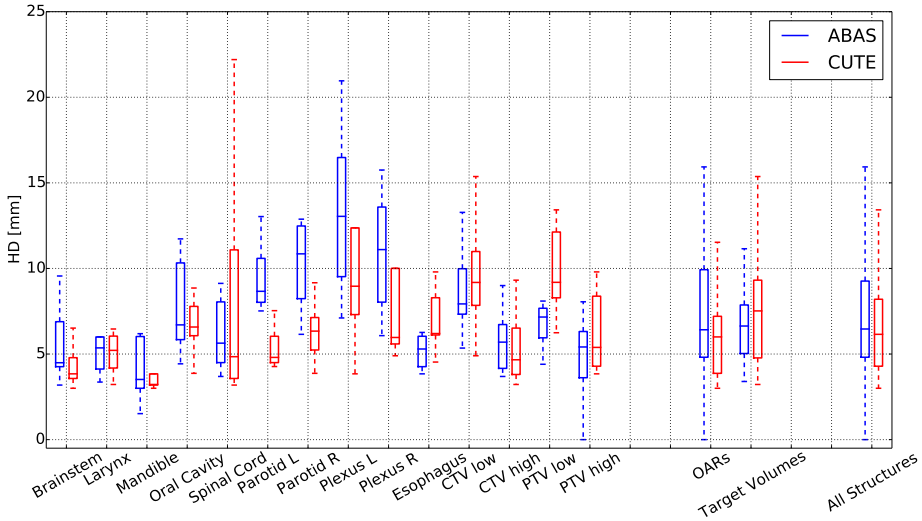


Fig. 6. Hausdorff distances of ABAS/CUTE contours compared to manual recontourings.

better than our method for target volumes resulting in higher DSC and lower HD. This is quite reasonable, considering that ABAS takes knowledge about nature of PTV into account. PTV contains large safety margins which do not correlate with the intensity values. Our registration based method cares about true image similarities and smooth transformations and not about safety margins. For OARs it is vice versa. Here only the intensity contrast and the surrounding structures rule and give an advantage to the individual registration.

Summarizing, atlas-based methods are good for bone structures and target volumes, for structures where the surroundings determine the concrete alterations, our approach performed significantly better in terms of both quality and robustness. We expect, that a combination of ABAS and CUTE could potentially optimize the overall workflow of re-planning situations in daily routine. Furthermore the fully automatic and robust approach can be performed in the background supporting automatic quantifications and dose accumulations. This will be a focus of future research.

In conclusion, contour propagation using automated mapping can be considered a reliable way to reduce the manual effort of re-contouring structures in adaptive radiation therapy.

References

1. Elekta: Atlas-based autosegmentation, http://www.elekta.com/healthcare-professionals/products/elekta-software/treatment-planning-software/contouring-software.html
2. Fischer, B., Modersitzki, J.: Curvature based image registration. Journal of Mathematical Imaging and Vision 18, 81–85 (2003)
3. Haber, E., Modersitzki, J.: Intensity gradient based registration and fusion of multimodal images. Methods of Information in Medicine 46(3), 292–299 (2007)
4. König, L., Kipshagen, T., Rühaak, J.: A non-linear image registration scheme for real-time liver ultrasound tracking using normalized gradient fields. In: Proc. MICCAI Challenge on Liver Ultrasound Tracking (CLUST 2014). Boston, USA (September 2014)
5. Modersitzki, J.: FAIR: Flexible Algorithms for Image Registration. SIAM, Philadelphia (2009)
6. Nijhuis, R., Brachmann, C., Weiler, F., Traulsen, N., Corr, D., Schenk, M., Ganswindt, U., Thieke, C., Wirtz, S., Belka, C.: A novel, efficient contour mapping method facilitates adaptive radiotherapy in head and neck patients. In: 3rd Estro Forum (2015)
7. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer (1999)
8. Pekar, V., Allaire, S., Qazi, A.A., Kim, J.J., Jaffray, D.A.: Head and neck autosegmentation challenge: Segmentation of the parotid glands (2010)
9. Zhu, Y.M., Cochoff, S.M., Sukalac, R.: Automatic Patient Table Removal in CT Images. J. Digital Imaging 25(4), 480–485 (2012)

Respiratory Motion Compensation with Topology Independent Surrogates

Christoph Jud[1], Frank Preiswerk[2] and Philippe C. Cattin[1]

[1]Department of Biomedical Engineering, University of Basel
[2]Brigham and Women's Hospital, Harvard Medical School
{christoph.jud,philippe.cattin}@unibas.ch

Abstract. We present a method for organ motion compensation based on a statistical motion model. The novelty of our method is, that the surrogates for prediction can be independent of the model's topology i.e. the surrogate signal does not have to correspond to a point of the motion model. By non-linear regression, we grasp the correlation between a captured signal during free breathing and the motion model parameters. Such a signal could be a spirometer, a breathing belt or an abdominal ultrasound signal. Our method performs on par with state-of-the-art methods, where model points have to be tracked. However, we solely incorporate one-dimensional and even topology independent surrogates. In our experiments on the right liver lobe we have achieved an average motion prediction accuracy of 2-3mm using population models and below 1mm with patient specific models.

1 Introduction

Respiratory organ motion compensation is central in image-guided thoracical and abdominal interventions. Especially in dynamic dose delivery methods as in radiotherapy or in high intensity focused ultrasound, tumor localization is crucial. Typically, a motion model is used to predict the organ motion given some external respiratory signal (surrogates). Thus, the treatment of healthy tissue can be reduced.

A prominent class of methods [3,7] is based on statistical motion models of organ shape deformations. During treatment, the shape deformation is inferred based on detected points in ultrasound (US) images which serve as surrogates. However, obtaining such surrogates is difficult, since they have to correspond to specific points in the model.

In this paper, we present a statistical motion model in which surrogates are incorporated that originate from arbitrary signal sources provided that they are correlated to the organ motion. For breathing, these might be a spirometer, a breathing belt or a 1D US where the sensor is placed on the abdominal skin [10]. The key idea is to predict the statistical motion model parameters given these surrogate signals using non-linear regression. This greatly simplifies the treatment setup since the absolute position of the surrogate sensor is no longer needed.

Beside the motion prediction, our method enables the synthesis of a respiratory cycle with a high temporal resolution. This can be used to investigate the patient specific motion pattern for planning. Further, in contrast to standard statistical motion model approaches, we construct our motion model with a different number of temporal samples per individual, to make full use of the training data. Finally, the dense underlying shape model opens the possibility to automatically fit the model into the 3D volume. However, this is not in the focus of the present paper.

In our experiments, we study the respiratory motion of the right liver lobe. To predict the organ motion, as surrogates, we simulate a correlated signal to the liver's motion. As such, we have reached an average accuracy between 2 and 3mm using a population based model. Using patient specific motion models, we have reached an accuracy of less than 1mm. The influence of organ drift [11] or rotations caused by change in posture during treatment has not been investigated.

Previously, [7] presented a statistical motion model, where physical points on the diaphragm serve as surrogates which correspond to points in the model. They detected and tracked such points using lateral US images. In [10], a low-cost 1D US signal has been proposed. As surrogate data, they map the US signal to positions on the organ. In our experiments, the depth of the diaphragm is simulated in the view to capture the signal from such a 1D US sensor which is placed on the abdominal skin. A comprehensive review on respiratory motion models can be found in [6].

The estimation of relationships among random variables has priorly been shown in [2] with the application to attribute manipulation in 3D face models. There, regression analysis of facial attributes and face models have been studied. However, to our best knowledge, such an approach has previously *not* been established in clinical applications.

2 Materials and Methods

2.1 Shape Modeling

For each volunteer v, we have a 4DMRI sequence of τ^v time steps, stacked by the method of [12], resulting in τ^v times an MR image $I_t^v : \Omega \to \mathbb{R}$ where $\Omega \subset \mathbb{R}^3$ is the image domain and I_t^v denotes the image at a time point t. For each volunteer, an exhalation master image $I_{t=m}^v$ is selected which serves as reference to determine relative displacements to each other time point. By non-rigid free form registration [9] of the exhalation master to each other image of this volunteer $I_{t\neq m}^v$ the displacement fields $D_t^v : \Omega \to \mathbb{R}^3$ are derived. To deal with across organ boundaries the surrounding of the liver structure is masked out during registration.

The liver structure within each exhalation master image has been manually segmented yielding a label map $L : \Omega \to \{0, 1\}$ which indicates liver structure when $L(x) = 1, x \in \Omega$. Using margin cubes, for each label map a shape $S \subset \mathbb{R}^3$

is obtained. We perform an iterative group-wise registration of the shapes to reduce a bias of the mean shape $\bar{S}$ to a specific exhalation master shape S^v.

$$\bar{S}_1 = S^{v=m}, \quad \bar{S}_{i+1} = \frac{1}{V}\sum_v \bar{S}_i + \Delta S_i^v, \tag{1}$$

where m is randomly chosen and ΔS_i^v is obtained by the Gaussian process registration method of [5] such that $\bar{S}_i + \Delta S_i^v \approx S^v$. In the following, by S^v we always mean the registered shapes $\bar{S} + \Delta S^v$, if nothing else is mentioned.

We equidistantly sample inside of $\bar{S}$ and with thin-plate-spline interpolation we add several interior points to each S^v. To recap, for each volunteer we have now an exhalation master shape S^v, which is in correspondence with the population mean shape $\bar{S}$. By the displacement fields $F_t^v := D_t^v(S^v)$ induced by motion, for each time point we derive a shape $S_t^v := S^v + F_t^v \subset \mathbb{R}^3$ with 2571 surface and 368 interior points.

2.2 Statistical Shape Model

We distinguish between the modeling of shape and the modeling of shape motion. In the shape model, the variation among a population of shapes originating from different individuals is considered. Whereas in the motion model, the shape deformation over time relative to a reference shape is investigated.

For each volunteer v, we have a segmented exhalation master shape S^v of the right liver lobe. The exhalation master shapes, which were brought into correspondence (Equation 1), are assumed to be Gaussian distributed $p(S^v|S_\mu, \Sigma_S) \sim \mathcal{N}(S\mu, \Sigma_S)$ where

$$S_\mu = \frac{1}{V}\sum_v S^v \mathrel{\widehat{=}} \bar{S}, \quad \Sigma_S = \frac{1}{V-1}\sum_v (S^v - S_\mu) \otimes (S^v - S_\mu) \tag{2}$$

are the maximum likelihood estimates of $p(S^v|S_\mu, \Sigma_S)$, V is the number of volunteers and $\otimes$ is the outer-product. Thus, a shape can be parametrized by $S^\alpha = S_\mu + \sum_{i=1}^M \alpha_i \psi_i$, where ψ_i are orthogonal basis vectors of Σ_S weighted by the model parameters α_i and M denotes the number of basis vectors.

2.3 Statistical Motion Model

In addition to the shape variation among a population we model the relative shape deformation over time. Since each volunteer has been observed for a different amount of time, we assume that the displacements are a mixture of Gaussian distributions

$$p(F) = \sum_v^V p(v)p(F|v) = \sum_v^V \pi^v p(F|v), \tag{3}$$

where $\sum_v^V \pi^v = 1, \pi^v \in (0,1), \forall v = 1, ..., V$. Each component distribution is assumed to be Gaussian $p(F^v) \sim \mathcal{N}(F_\mu^v, \Sigma_{F^v})$ with

$$F_\mu^v = \frac{1}{\tau^v}\sum_t^{\tau^v} F_t^v, \quad \Sigma_{F^v} = \frac{1}{\tau^v - 1}\sum_t^{\tau^v} (F_t^v - F_\mu^v) \otimes (F_t^v - F_\mu^v). \tag{4}$$

The first two moments of the mixture $p(F)$ are estimated by

$$F_\mu = \sum_v^V \pi^v F_\mu^v, \quad \Sigma_F = \sum_v^V \pi^v \left(\Sigma_{F^v} + \left(F_\mu^v - F_\mu\right) \otimes \left(F_\mu^v - F_\mu\right)\right), \tag{5}$$

where $\pi^v = \tau^v / \sum_v^V \tau^v$ is the weighting of the component distribution with respect to the number of temporal samples per volunteer (see more details about moments of Gaussian mixtures in [4])

The variation of the shape displacements is finally parametrized by $F = F_\mu + \sum_{i=1}^N \beta_i \phi_i$, where ϕ_i are N orthogonal basis vectors of Σ_F.

With the combination of the shape and the motion model a shape to a particular time point can be synthesized by

$$S_\beta^\alpha = \underbrace{S_\mu + \sum_{i=1}^M \alpha_i \psi_i}_{\text{shape model}} + \underbrace{F_\mu + \sum_{i=1}^N \beta_i \phi_i}_{\text{motion model}}, \tag{6}$$

where α_i and β_i are coefficients of the shape and the motion model respectively. Typically, the shape model is priorly fitted to an exhalation master shape [5]. Subsequently, the shape motion is additively imposed to the derived shape.

2.4 Attributes and Regression

Let $a \in \mathbb{R}^d$ be a d-dimensional attribute vector which corresponds to a particular time point. Consider an observed finite set $A = \{(a_0, \beta_0), ..., (a_n, \beta_n)\} \subset \mathbb{R}^d \times \mathbb{R}^N$ of n pairs of i.i.d. attribute vectors a_i and motion coefficient vectors β_i. Let further assume that there exists a function $f : \mathbb{R}^d \to \mathbb{R}^N$ which maps the attribute vectors to the coefficient vectors, while we only observe noisy instances of β such that $\beta \sim \mathcal{N}(f(a), \sigma_\epsilon \mathbf{I})$.

Gaussian Process Regression Let $f \in \mathcal{GP}(0, k)$ be a Gaussian process with the covariance function $k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$. Assuming a Gaussian likelihood, the posterior distribution $p(f|A)$ is given in closed form [8] and is again a Gaussian process $\mathcal{GP}(\mu_A, k_A)$ with

$$\mu_A(a) = \mathbf{K}_{a,A}^T (\mathbf{K}_{A,A} + \sigma_\epsilon \mathbf{I})^{-1} \mathbf{B} \tag{7}$$

$$k_A(a, a') = k(a, a') - \mathbf{K}_{a,A}^T (\mathbf{K}_{A,A} + \sigma_\epsilon \mathbf{I})^{-1} \mathbf{K}_{a',A}, \tag{8}$$

where $\mathbf{K}_{a,A} = (k(a_i, a))_{i=1}^n \in \mathbb{R}^n$, $\mathbf{K}_{A,A} = (k(a_i, a_j))_{i,j=1}^n \in \mathbb{R}^{n \times n}$ and $\mathbf{B} = (\beta_0, \ldots, \beta_n)^T \in \mathbb{R}^{n \times N}$. The expectation of an unseen output β^* given an attribute a^* yields Equation 7.

In our application, we use a straight forward Gaussian process model, where we apply a Gaussian kernel as covariance function

$$k_g(x, x') = \theta_0^2 \exp\left(-\frac{\|x - x'\|^2}{2\theta_1^2}\right), \tag{9}$$

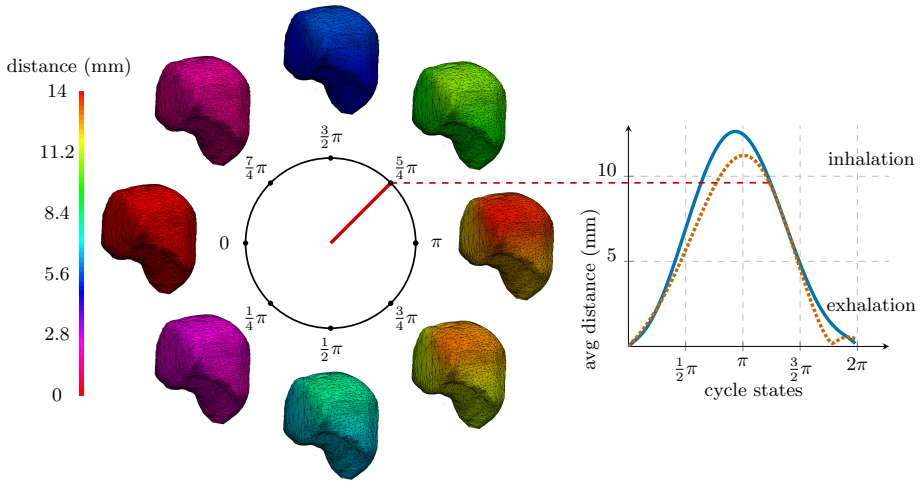


Fig. 1: Example liver shape S^v deformed by an average respiratory cycle $c = [0, 2\pi]$. The shape is colored with respect to the absolute value of the displacement of a point $|\Delta p| \in [0, 14]$ mm. In the right plot, the average distance to the exhalation master is plotted (*blue* - population mean, *orange dotted* example of a patient specific cycle).

where θ_0 is a scaling parameter and θ_1 is a length scale or smoothness parameter.

Given an attribute signal which is correlated with the organ motion, we have defined now the tools to predict the motion model parameters β of the current respiratory state with Equation 7. Based on that, a shape deformation can be synthesized using Equation 6.

In the following, we first synthesize a high-temporal resolution respiratory cycle. This is followed by the evaluation of the prediction performance of our method.

3 Results

3.1 Average Breathing Cycle

In this first study, we analyze the respiratory motion of the liver in general. We built a motion model out of the motion samples among all the $V = 9$ volunteers, while we have kept 99.9% of the variance. For each sample shape S_t, an attribute $c \in [0, 2\pi]$ has been considered which indicates the cycle state of t within a respiratory cycle[1]. This rather abstract attribute is applied to synthesize an average respiratory cycle of the liver shape. For the regression, 8 000 pairs

[1] This cycle attribute was computed using a greedy cycle detection algorithm which is based on the average vertical coordinates of the displacement fields F_t.

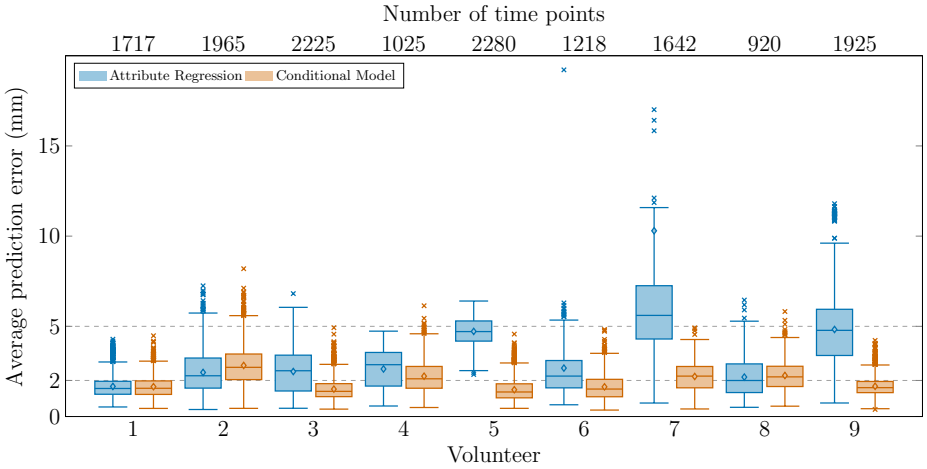


Fig. 2: For each L1O experiment, the average corresponding point difference between the ground truth and the predicted shape is visualized. We compare our method *Attribute Regression* with the *Conditional Model* [1,7]. The upper x-axis indicates for how many time points the motion has been predicted.

of cycle attributes resp. motion model coefficients have been randomly picked among all volunteers.

In Figure 1, an example right liver lobe and its displacements within this average respiratory cycle is visualized. Note that here, we synthesize a semantic and non-linearly captured respiratory cycle of a shape, where we do *not* simply vary the most dominant principal component of the motion model. Thus, for each patient we can generate an average respiratory cycle e.g. for planning. While the source 4DMRI has a framerate of 2.8Hz the temporal resolution can be arbitrary high. In this example, 100 samples have been synthesized which corresponds to approximately 25Hz.

3.2 Motion Model Prediction

In the motion prediction experiment, we simulate a surrogate signal which indicates the depth of the diaphragm measured from the abdominal skin for example by a 1D US sensor. We define a 1D signal which is generated by a ground truth model point in the region of the diaphragm. Let $s : [0, \tau] \to \mathbb{R}^3$ be the 3D signal of absolute coordinates of this point at time point $t \in [0, \tau]$. To get invariant to the absolute positioning of the patient lets project the signal into its dominant mode of variation

$$\mathcal{F}[s] = (s - \mu_s)\psi_0 + \epsilon, \tag{10}$$

where $\mu_s = \frac{1}{\tau}\int_0^\tau s(t)\mathrm{d}t$ is the signal mean value, $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon)$ is additive noise, which has been set to $\sigma_\epsilon = 2$mm and ψ_0 is the orthonormal eigenfunction corre-

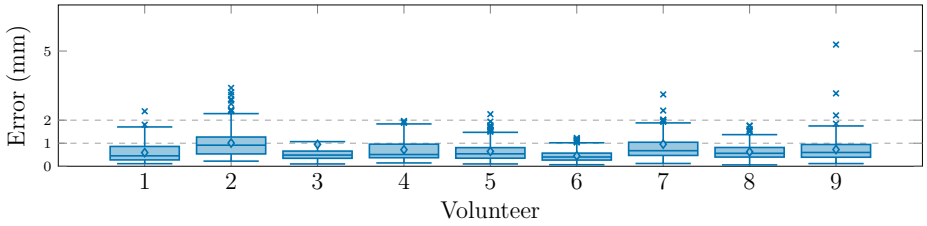


Fig. 3: Evaluation of the average prediction error for each patient specific experiments, where $\theta_1 = 5$, 700 training samples have been picked and 210 tests have been performed.

sponding to the largest eigenvalue λ_0 of the equation $\int_0^\tau \mathrm{cov}(s_i, s_j)\psi_0(s_i)\mathrm{d}s_i = \lambda_0\psi_0(s_j)$. Here, cov is the covariance function of the signal s.

In this evaluation, we show the motion prediction performance of our method given the simulated signal $\mathcal{F}[s]$. For each volunteer, we generated a leave-one-out (L1O) motion model, where only motion samples from the other volunteers have been considered. 99.9% of the variance has been kept yielding L1O-models of 20 to 22 principal modes. Note that for each sample, we additionally computed $\mathcal{F}[s]$ for the later usage as an attribute (Equation 10).

For the Gaussian process regression, we randomly picked 8 000 $\mathcal{F}[s]$-attributes resp. ground truth coefficient vector pairs, again only from the other volunteers. We manually optimized the parameters and used $\theta_1 = 3$, while the exact value of $\theta_0 = 5\,000$, had only minor influence to the prediction performance. In Figure 2, for each volunteer the average prediction error is plotted. The prediction error is robustly kept below 5mm, whereas the median stays around 2 to 3mm. For radiotherapy these are reasonable error bounds.

We compare our method to [7] where the simulated 3D point signal s serves as surrogate data. The prediction is performed by estimating the mean of a statistical motion model which is conditioned on s [1]. For a fair comparison, we added to s an isotropic Gaussian noise $\mathcal{N}(0, \sigma_\epsilon/\sqrt{3})$. The conditional model performs equally well, while it better generalizes in the experiment with volunteer 7 and 9. Certainly, our model with only 9 volunteers is not capable to generalize to the respiratory motion of these two subjects. This can be confirmed when comparing to the results with patient specific models (Figure 3). Here, for each volunteer, we built a motion model where only samples from the volunteer of interest are considered. For the regression, 700 attribute/coefficient pairs have been randomly picked and we adjusted $\theta_1 = 5$. For all volunteers including for volunteer 7 and 9, the average error has been considerably improved to less than 1mm. In Figure 1 on the right, a patient specific average respiratory cycle is plotted for a comparison to the population mean cycle.

4 Conclusion

We have presented a new method for organ motion compensation based on a statistical motion model, in which the surrogate data can be independent of the model's topology. The major novelty of our method is the non-linear regression of the surrogate data and the model parameters. Although already a simple Gaussian process model yields reasonable results the potential of our method is far from being exhausted. The regression is *not* limited to one attribute and will gain robustness and precision with additional surrogate sources as e.g. 2D US data. Further, we will investigate more complex and combined covariance functions and a full Gaussian process inference to obviate parameter selection. In the experiments, we have shown a reasonable prediction performance using population based models. The generalization of the respiratory motion was further improved by a patient specific regression as shown in the last experiment.

References

1. Albrecht, T., Lüthi, M., Gerig, T., Vetter, T.: Posterior shape models. Medical image analysis 17(8), 959–973 (2013)
2. Amberg, B., Paysan, P., Vetter, T.: Weight, sex, and facial expressions: On the manipulation of attributes in generative 3d face models. In: Advances in Visual Computing, pp. 875–885. Springer (2009)
3. Boye, D., Samei, G., Schmidt, J., Székely, G., Tanner, C.: Population based modeling of respiratory lung motion and prediction from partial information. In: SPIE Medical Imaging. pp. 86690U–86690U. International Society for Optics and Photonics (2013)
4. Carreira-Perpiñán, M.Á.: Mode-finding for mixtures of gaussian distributions. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(11), 1318–1323 (2000)
5. Lüthi, M., Jud, C., Vetter, T.: A unified approach to shape model fitting and non-rigid registration. In: Machine learning in medical imaging, pp. 66–73. Springer (2013)
6. McClelland, J.R., Hawkes, D.J., Schaeffter, T., King, A.P.: Respiratory motion models: A review. Medical image analysis 17(1), 19–42 (2013)
7. Preiswerk, F., Arnold, P., Fasel, B., Cattin, P.C.: A bayesian framework for estimating respiratory liver motion from sparse measurements. In: Abdominal Imaging. Computational and Clinical Applications, pp. 207–214. Springer (2012)
8. Rasmussen, C.E.: Gaussian processes for machine learning (2006)
9. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J.: Non-rigid registration using free-form deformations: application to breast mr images. Medical Imaging, IEEE Transactions on 18(8), 712–721 (1999)
10. Schwartz, B.M., McDannold, N.J.: Ultrasound echoes as biometric navigators. Magnetic Resonance in Medicine 69(4), 1023–1033 (2013)
11. von Siebenthal, M., Székely, G., Lomax, A.J., Cattin, P.C.: Systematic errors in respiratory gating due to intrafraction deformations of the liver. Medical physics 34(9), 3620–3629 (2007)
12. Von Siebenthal, M., Székely, G., Gamper, U., Boesiger, P., Lomax, A., Cattin, P.: 4d mr imaging of respiratory organ motion and its variability. Physics in medicine and biology 52(6), 1547 (2007)

Patient localization for robotized ultrasound-guided radiation therapy

Ivo Kuhlemann[1,2], Philipp Jauer[1], Achim Schweikard[1], and Floris Ernst[1]

[1] Institute for Robotics and Cognitive Systems, University of Lübeck, Ratzeburger Allee 160, 23562 Lübeck, Germany
{kuhlemann, jauer, schweikard, ernst}@rob.uni-luebeck.de
http://www.rob.uni-luebeck.de
[2] Graduate School for Computing in Medicine and Life Sciences, University of Lübeck, Ratzeburger Allee 160, 23562 Lübeck, Germany

Abstract. Accurate localization and tracking of moving targets is one of the major challenges faced today during high-precision radiotherapy. Typically, the position of the treatment target is either determined using infrequent X-ray images or cone-beam CT scans. A totally different approach currently under active development makes use of ultrasound imaging to continuously track the target region. We have evaluated a robotized setup where Microsoft's Kinect v2 sensor is used to localize the patient and specific ultrasonic view ports previously defined in the planning CT. The setup is validated using an anthropomorphic torso phantom and four predefined view ports (apical and parasternal echocardiography, liver sonography, suprapubic prostate sonography). The Kinect sensor and an optical tracking system (used to determine the position of the torso phantom) were calibrated to the robot using the QR24 hand-eye calibration algorithm. Then each view port was approached fifteen times from different directions, showing that the accuracy achievable is, on average, approximately 2.1 *cm*. This number can mostly be attributed to the difficulty of obtaining accurate calibration of the geometric relationship between the robot and the Kinect sensor. It was observed that the Kinect sensor system suffers from substantial distortion in the centimeter range, severely compromising the accuracy of the whole setup.

1 Introduction

To compensate for tumor movements and thus to minimize the injury of healthy tissue is the aim of today's robotized radiation therapy systems like the CyberKnife. These systems would greatly benefit from developing a method for non-invasive live tracking of tumors. Given that the only currently existing such method relies on live monitoring using fluoroscopy [12], the need to come up with an approach not relying on ionizing radiation is evident. In recent years, there have been multiple studies on using ultrasound for target localization during radiotherapy; see for example [3,11]. These approaches, however, all lack one important feature: it is implicitly assumed that the robot can find the target or that it is guided to the target manually. In reality, this approach is not feasible: it

might be necessary to automatically position or re-position the robot during delivery because the arm may obstruct the treatment beam path for certain beam directions. Consequently, a fully autonomous setting for transducer placement would be the ultimate goal. In this work, we present an approach to dynamically locate the patient during treatment using a spatial sensor (Microsoft Kinect v2) mounted to the tool flange of an industrial robot (Adept viper s850). This mounting is envisioned to be in parallel to an ultrasonic transducer also attached to the tool flange. Clearly, this setup now requires knowledge about spatial transforms describing the geometric relationship between the individual components, being the robot, the ultrasonic transducer and the spatial sensor. More specifically, these transforms need to be known with varying degrees of accuracy: The transform from the patient's body surface to the robot coordinate system needs to be known with medium (centimeter range) accuracy only, since it is expected that the optimal transducer position needs to be found adaptively by machine vision and force feedback. In reality, it is not possible to determine this transformation with high (millimeter range) accuracy because it will inevitably be based on some kind of pre-treatment volumetric scan (CT, MRI or CBCT) and cannot take body deformation or physiological changes into account. The other transforms, however, which describe the location of the spatial sensor and the transducer with respect to the tool flange must be determined with the highest possible accuracy (ideally with an accuracy of better than one millimeter). The accuracy of these transformations will directly affect the capabilities of the system to reach a defined target position.

2 Material and Methods

To be able to more easily asses the accuracy of patient localization, the ultrasonic probe has been replaced by an optically trackable pointing device. This pointing device is factory calibrated with respect to the tracking system, thus allowing us to accurately determine the position and orientation of the pointer's tip with respect to the tracking system. In a real scenario, it would of course be required to determine the spatial relationship between the ultrasonic probe's transducer surface and the carrying system. This can be done with high accuracy by following the procedure described in [5]. It was shown that sub-millimeter accuracy can be achieved when the calibration procedure is carried out with the appropriate care.

2.1 Robot-to-Ultrasound (Pointer) calibration

In a first step, the transformation between the robot's effector and the tip of the pointing device ${}^{E}\mathbf{T}_{M}$ has to be determined. Here, a matrix of type ${}^{A}\mathbf{T}_{B}$ describes the homogeneous transform from the coordinate system A to the coordinate system B. In our setup, E stands for *effector*, i.e. the coordinate system at the robot's tool flange. This transformation matrix can be readily obtained by using any of the many algorithms available for hand-eye calibration [6,10,13].

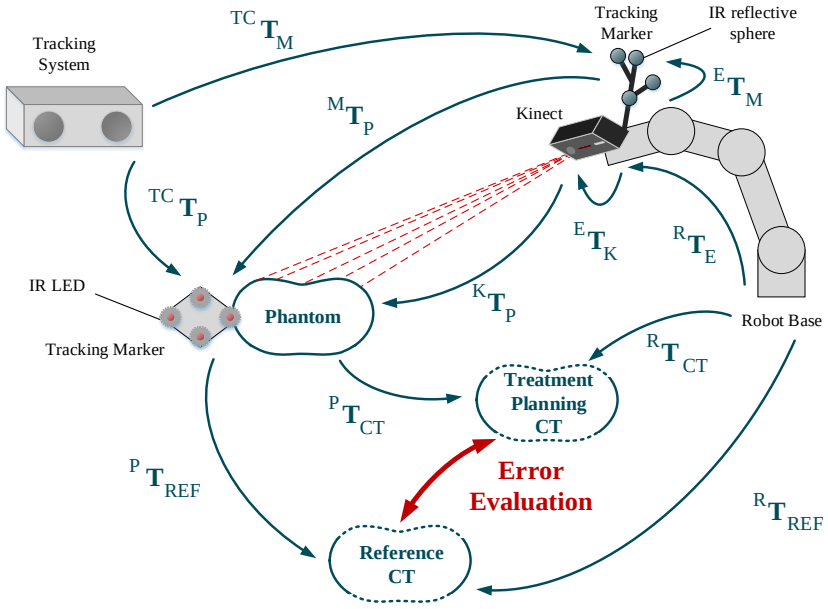
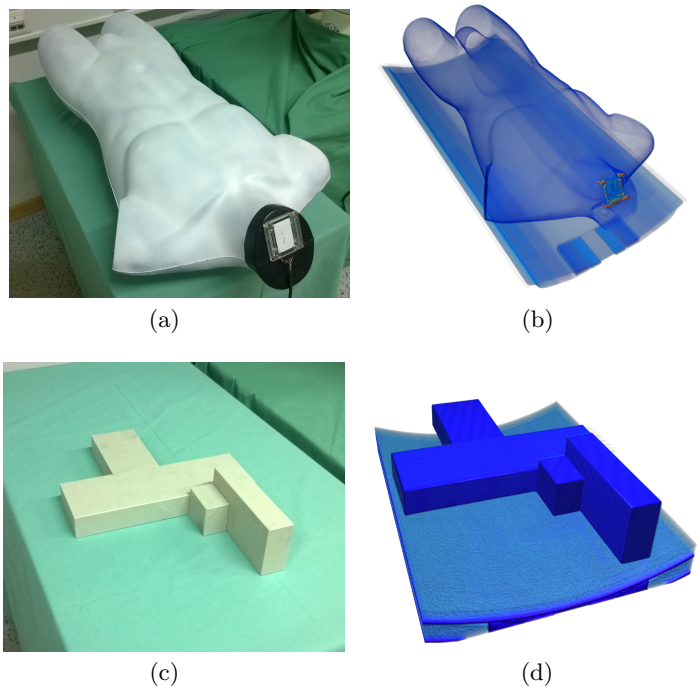


Fig. 1: Calibration and evaluation setup of our experiment

We use the *QR24 algorithm* for hand-eye-calibration, see [7], since it also takes imperfections of the tracking system into account. This algorithm not only determines the transformation ${}^{E}\mathbf{T}_{M}$ but also the transformation ${}^{R}\mathbf{T}_{TC}$, i.e. the transformation from the robot's base coordinate system R to the tracking camera's base coordinate system TC. This calibration is carried out by moving the robot to multiple random poses ($n \geq 30$) within a sphere while the position of both the robot's effector M_i and the tracked marker N_i is recorded. The resulting equation systems of the form $M_i \cdot {}^{E}\mathbf{T}_{M} = {}^{R}\mathbf{T}_{TC} \cdot N_i$, $i = 1, \ldots, n$, can then be solved for the unknown transformation matrices. These transformations are conceptually shown in Figure 1.

2.2 Robot-to-Kinect calibration

To calibrate the Kinect v2 sensor's internal coordinate system with respect to the robot's effector, the same algorithm is used. Note, however, that in this case the calibration process is an eye-in-hand calibration since the Kinect sensor is mounted to the robot's flange. A custom-built tracking phantom (see Figure 2, (c)) is used as the marker object required for calibration. A ground truth for the geometry of this object was obtained from a CT scan (Siemens Somatom Balance; 1 mm slice thickness, 0.59 mm in-plane resolution). Then the surface of the phantom was extracted from the CT data (visualized in Figure 2, (d)) and was consequently used as a template for ICP-based registration [2] to the point



(a) (b)

(c) (d)

Fig. 2: A photo of the anthropomorphic phantom (a) and its CT data (b). The calibration tool for the Kinect sensor (c) and the corresponding CT data (d).

clouds delivered by the Kinect sensor. In the same way as for the tracking camera, the QR24 calibration algorithm was performed on multiple measurements of the static phantom from different robot poses.

2.3 Coordinate transformation CT-to-Tracking camera

To be able to accurately determine the overall accuracy of the system, we used an anthropomorphic torso phantom, see Figure 2, (a). An optically trackable active marker device (IR LEDs) was rigidly attached to the torso phantom. This marker was also detectable by the same tracking camera as the pointing device used to mimic the ultrasound probe. A surface point cloud of the torso phantom was again extracted from a CT scan (same protocol as before). A ray casted rendering of this scan — also showing the attached marker device — is given in Figure 2, (b). The CT data was also used to determine the relationship between the LEDs of the marker device and the phantom's surface.

2.4 Definition of ultrasonic view ports

To determine the accuracy achievable with this setup, four patches representing typically used ultrasonic view ports (two echocardiographic ports, a liver port

and a suprapubic prostate port) were selected as target templates from the torso phantom's CT point cloud. These patches are shown in Figure 3. Using the approach described in [4], these view ports can be determined from the planning CT data. An example is given in Figure 4.

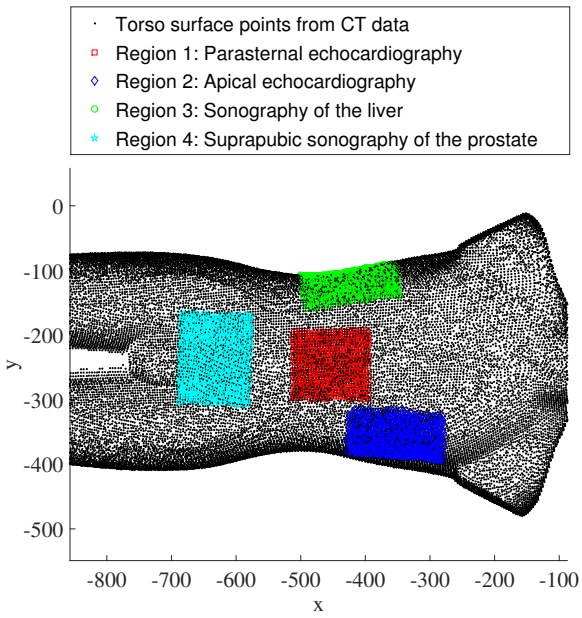


Fig. 3: The four patches used to test the registration and calibration accuracy

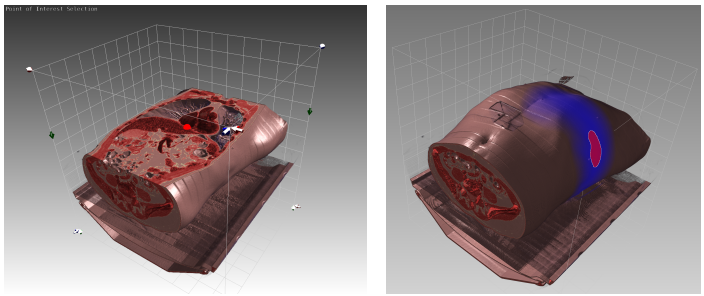


Fig. 4: View port determination based on the planning CT. Left: manual target definition (red sphere). Right: Quality of ultrasonic view port (blue). The darker the color, the better the visibility of the target becomes. Optimal view port is shown in red. Figures from [1].

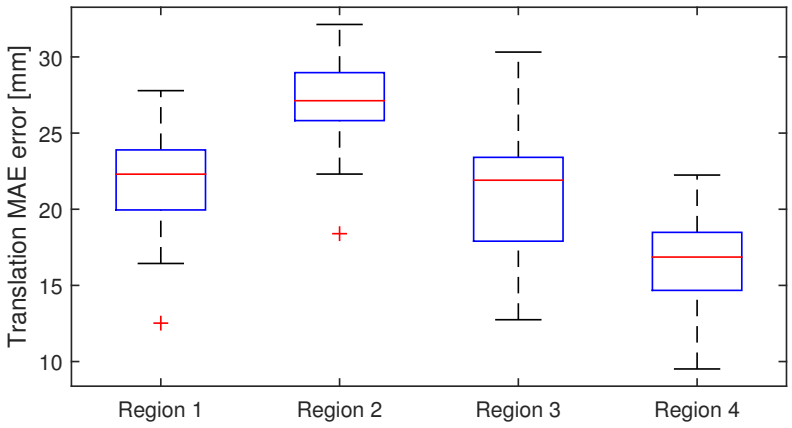


Fig. 5: The translational point-to-point Mean Absolute Error in [mm] based on fifteen localization approaches and four view port patches.

3 Results and Discussion

To evaluate the overall accuracy, we compared all points in the preselected view port patches inside the reference CT (used as ground truth), to the position of the corresponding patches after being registered by ICP to the point cloud coming from the Kinect sensor (see Figure 1). In fact, all points in a single patch in the reference CT have their corresponding points in the registered planning CT. Hence, in case of perfect calibration and tracking of the Kinect, these two patches would be exactly superimposed, but are practically not. We calculated the point-to-point translation and rotation error based on fifteen initial patient localizations for four different view port regions. To determine the translation error, we used the *Mean Absolute Error* (MAE) with a point-to-point correspondence. The rotation error was calculated by using the *Horn* algorithm [8] for determining the transformation between two corresponding point clouds (in this case, the two patches). The calculated rotation matrix was then converted into the axis-angle representation to obtain the rotation error in one axis. In Figure 5, we can see that the MAE point-to-point translation error varies around a mean of 21.6 mm with a maximum of 32.1 mm (Region 2) and a minimum of 9.5 mm (Region 4). Figure 6 shows the results for the rotation error with a mean error of 21.8°, a maximum of 23.9° in Region 3 and a minimum of 21.0° in Region 1.

4 Conclusions

Clearly, the results obtained are not perfect: even under the initial assumption of the Kinect sensor being a consumer product – and thus probably not perfectly

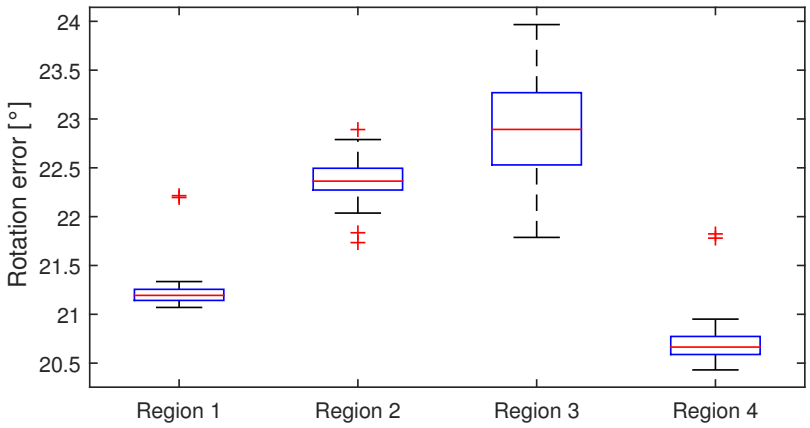


Fig. 6: The Mean Absolute rotation error in [°] based on fifteen localization approaches and four view port patches.

calibrated – and its limitations (noise, resolution, etc.), we had hoped for better results. Nevertheless, the results obtained are acceptable and useable in so far that exact positioning of the ultrasonic probe is not really needed. As stated before, the scanned patches will always be registered to a surface extracted from a planning CT scan taken days (or even weeks) before treatment. Additionally, the patient will not necessarily stay in the exact same position, resulting in deformation of the skin surface. Furthermore, as can be seen in Figure 4, a typical view port is not exactly small and it is not required to perfectly place the transducer. Fully automatic systems will also incorporate machine vision methods and force adapting strategies to improve the image quality (see, for example, [9]). Of course, there is a potential risk of unwanted collisions between probe and patient, but since robotized ultrasound is partially based on force-servoing, there exist several features to avoid this complication [9].

During our experiments, it was observed that the Kinect's reconstruction of a planar surface (lab floor) deviated substantially from the expected. When the scanned plane (about 1.2 by $0.5\,m$) was aligned to the coordinate axes using PCA, the z coordinate exhibited a strong distortion, where the top part of the plane was displaced by as much as two centimeters relative to the bottom part of the plane. Considering these findings, the following points will have to be analyzed in future works:

- What is the quality of the intrinsic calibration of the Kinect v2 sensor?
- Is there a nonlinear behavior on the depth values reported by the sensor?
- Can calibration of the spatial relationship between the Kinect sensor and the robot be improved by additionally incorporating information from the – currently not used – color image?

Assuming that these questions can, at least in part, be solved, it is expected that the registration accuracy of the setup presented will fall considerably.

References

1. Ammann, N.: Robotized 4D Ultrasound for cardiac Image- Guided Radiation Therapy. Master's thesis, University of Lübeck (2012)
2. Besl, P.J., McKay, H.D.: A method for registration of 3-D shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence 14(2), 239–256 (Feb 1992)
3. Blanck, O., Jauer, P., Ernst, F., Bruder, R., Schweikard, A.: Pilot-Phantomtest zur ultraschall-geführten robotergestützten Radiochirurgie. In: Treuer, H. (ed.) 44. Jahrestagung der DGMP. vol. 44, pp. 122–123. DGMP, Cologne, Germany (Sep 2013)
4. Bruder, R., Ernst, F., Schweikard, A.: SU-D-220-02: Optimal transducer positions for 4D ultrasound guidance in cardiac IGRT. In: 53rd Annual Meeting of the AAPM. Medical Physics, vol. 38, p. 3390. Vancouver, BC, Canada (Jul 2011)
5. Bruder, R., Griese, F., Ernst, F., Schweikard, A.: High-accuracy ultrasound target localization for hand-eye calibration between optical tracking systems and three-dimensional ultrasound. In: Handels, H., Ehrhardt, J., Deserno, T.M., Meinzer, H.P., Tolxdorff, T. (eds.) Bildverarbeitung für die Medizin 2011. pp. 179–183. Informatik aktuell, Springer, Berlin, Heidelberg (2011)
6. Daniilidis, K.: Hand-eye calibration using dual quaternions. International Journal of Robotics Research 18(3), 286–298 (1999)
7. Ernst, F., Richter, L., Matthäus, L., Martens, V., Bruder, R., Schlaefer, A., Schweikard, A.: Non-orthogonal tool/flange and robot/world calibration for realistic tracking scenarios. International Journal of Medical Robotics and Computer Assisted Surgery 8(4), 407–420 (Dec 2012)
8. Horn, B.K.P.: Closed-form solution of absolute orientation using unit quaternions. Journal of the Optical Society of America A 4(4), 629–642 (1987)
9. Kuhlemann, I., Bruder, R., Ernst, F., Schweikard, A.: WE-G-BRF-09: Force- and image-adaptive strategies for robotised placement of 4D ultrasound probes. In: 56th Annual Meeting of the AAPM. Medical Physics, vol. 41, p. 523. Austin, TX, USA (Aug 2014)
10. Park, F.C., Martin, B.J.: Robot sensor calibration: solving AX=XB on the Euclidean group. IEEE Transactions on Robotics and Automation 10(5), 717–721 (1994)
11. Schlosser, J., Salisbury, K., Hristov, D.: Telerobotic system concept for real-time soft-tissue imaging during radiotherapy beam delivery. Medical Physics 37(12), 6357–6367 (2010)
12. Shirato, H., Shimizu, S., Kitamura, K., Nishioka, T., Kagei, K., Hashimoto, S., Aoyama, H., Kunieda, T., Shinohara, N., Dosaka-Akita, H., Miyasaka, K.: Four-dimensional treatment planning and fluoroscopic real-time tumor tracking radiotherapy for moving tumor. International Journal of Radiation Oncology, Biology, Physics 48(2), 435–442 (2000)
13. Tsai, R.Y., Lenz, R.K.: A new technique for fully autonomous and efficient 3D robotics hand/eye calibration. IEEE Transactions on Robotics and Automation 5(3), 345–358 (June 1989)

Real-time tumor tracking during VMAT radiotherapy treatments based on 2D/3D registration using CBCT projections

Hugo Furtado[1,3], Yvette Seppenwoolde[2,3], Dietmar Georg[2,3], and Wolfgang Birkfellner[1,3]

[1] Center for Medical Physics and Biomedical Engineering, Medical University Vienna, Austria
[2] Department of Radiation Oncology, Medical University Vienna, Austria
[3] Christian Doppler Laboratory for Medical Radiation Research for Radiation Oncology, Medical University Vienna, Austria

Abstract. Rotational radiotherapy treatments such as volumetric modulated arc therapy (VMAT) can have superior overall quality while having shorter delivery time with respect to conventional intensity modulated radiotherapy (IMRT). As with conventional treatments, intra-fractional tumor motion is a major source of uncertainty in dose application. To ensure full tumor coverage the planning target volume (PTV) margins are enlarged. Tumor tracking using intensity based 2D/3D registration can reduce uncertainty and would enable margin reduction. The challenge for full arc rotational treatments is the poor tumor visibility at certain irradiation angles. In this work we investigate the feasibility of tumor tracking in VMAT treatments with partial arcs where the tumor is well visible. For three patients, we determined for which subset of angles it was possible to track the tumor using raw projections from CBCT acquisition. We then generated VMAT plans with the obtained partial arcs. Full arc VMAT plans were used as benchmark. For all cases is was possible to track the tumor in two arcs of about 90 degrees, typically with imaging around anterior-posterior (AP) or posterior-anterior (PA) projections. The plans with the partial arcs were clinically comparable to the plans with full arc in terms of target coverage and OAR sparing. The results indicate the feasibility of partial arc VMAT treatments with tumor tracking. Based on this we will investigate generation of plans with reduced PTV margins. These treatments could be delivered by gating the LINAC beam if the tumor motion exceeds a certain margin.

Keywords: 2D/3D registration, motion tracking, radiation therapy, VMAT

1 Introduction

In radiotherapy, rotational treatments such as intensity modulated arc therapy (IMAT) or volumetric modulated arc therapy (VMAT) have shown superior overall quality while treatment times are shorter [3]. As with conventional treatments, intra-fractional tumor motion remains as one of the main challenges to

solve. The uncertainty due to tumor motion imposes an enlargement on the planned target volume (PTV) in order to assure sufficient dose delivery to the tumor to achieve local control [9]. The PTV enlargement typically leads to increased dose delivery to organs at risk (OAR). In the case of lung tumors, the target site of this work, the main cause for motion is breathing. But, depending on the tumor location heartbeat can also contribute.

Continuous efforts to manage tumor motion have been made especially in the last decade. Management can be done by tracking the tumor position during treatment in order to reduce the uncertainty. Among the different approaches for tumor tracking, purely intensity based 2D/3D registration [11] using x-ray images acquired intra-fractionally is a feasible approach. This technique can deal well with both periodical and non-periodical motion patterns in conventional SBRT treatments [8][7]. However, results for this technique are for the time being limited to specific imaging angles. Investigation at tumor tracking for different gantry angles is underway [13] but early results suggest that for many cases there are imaging angles where tumor tracking will not be possible due to poor visibility. This would limit the applicability of 2D/3D registration for rotational treatments as typically they consist of a full arc rotations around the patient. Tracking approaches using implanted fiducial markers [5] are very successful and have therefore been investigated for VMAT treatments as well [1]. But, implanting markers in the lung has associated complications [4].

In this work we investigate the feasibility of performing partial arc VMAT treatments for lung cases using markerless real-time 2D/3D registration for tumor tracking. The partial arcs consist of the angle intervals where the tumor is well visible and therefore tracking is possible. To assess the feasibility of tumor tracking we used x-ray projections acquired for CBCT reconstruction of the patients under treatment. The projections form a basis for assessing the angle intervals where tracking is possible. Based on individual patient findings, partial arc treatment plans were generated and their clinical relevance compared to a regular full arc VMAT plan. Our aim is to create partial arc treatment plans with reduced PTV margins and to deliver treatment with the tumor under constant monitoring.

2 Materials and methods

2.1 Image datasets and Image preprocessing

For this investigation we used data from three patients undergoing hypo-fractionated SBRT treatment at our hospital. Each patient was treated with 3 fractions. As a routine at each fraction, a cone beam CT (CBCT) is acquired for the purpose of daily patient positioning. This setup step is well established and guarantees that during treatment, uncertainties in tumor location are typically only due to breathing or heartbeat.

Each of the CBCTs is reconstructed from 657 planar kV x-rays. These x-rays are the ones used in our study. The diaphragm position was manually annotated

in each of the projections and is the basis for deciding when the tumor was correctly tracked or not.

Tumor visibility depends on the x-ray path through the body. Therefore, for optimal tumor tracking the image intensity settings have to be adjusted as a funtion of each projection angle. An automatic intensity adjustment method based on contrast maximization in the PTV region was used [6]. As an example, figure 1 shows CBCT projections for 12 different gantry angles which are automatically intensity adjusted based on this method.

For creating the new VMAT plans, we used the conventional planning CT which has all the target volumes and organs at risk (OARs) delineated. Both the full and partial arc VMAT plans were generated using these CT datasets.

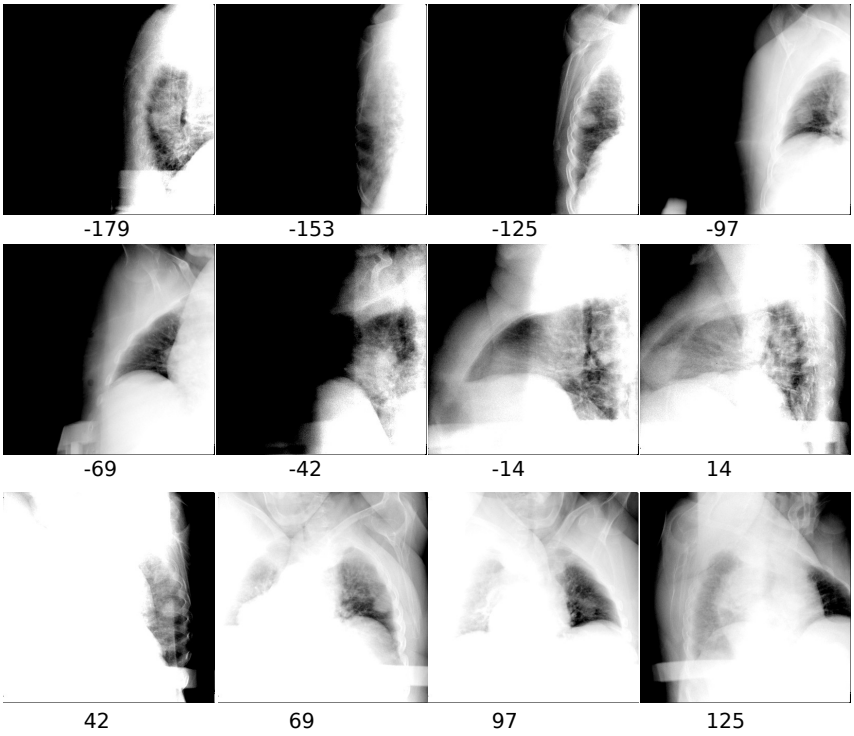


Fig. 1. Illustrative set of x-ray projections obtained during CBCT acquisition for patient 2. The figure shows 12 projections with the angle indicated. All images were automatically intensity adjusted to maximize contrast around the tumor region. The tumor is well visible in projections around the AP or PA directions (angles -153, -125, -97, -69, 42, 69, 97) and not well distinguishable in the other angles.

2.2 2D/3D registration

Intensity based 2D/3D registration is an optimization process which aims at finding the spatial transform for a volume dataset of the patient that generates a digitally reconstructed radiograph (DRR) that best matches a real x-ray image acquired during treatment.

The first step of the registration consists of generating an initial DRR from the planning CT. The planning CT is initially aligned by 3D/3D registration to the CBCT volume. The initial transform for the first DRR is defined automatically by using the imaging angle for each of the CBCT projections. The DRR is then compared to the CBCT projection by means of a merit function. An optimizer iterates through the previous steps (new DRR and merit calculation) searching for the rigid spatial transformation T with the highest similarity between the DRR and the x-ray. The final translational and rotational parameters $(t_x, t_y, t_z, \omega_x, \omega_y, \omega_z)$ represent the tumor displacement.

The registration was limited to a region-of-interest (ROI) defined as the projection of the PTV on the x-ray images. We selected this ROI mainly because this is the area where we want to track motion and where this motion is assumed to be rigid as our method does not account for deformation. One advantage of using a small ROI is also that rendering a smaller region is much faster.

DRR generation is the most time consuming step therefore we used ray-casting implemented on a general purpose graphics processing unit (GPGPU). We used normalized mutual information [10] or stochastic rank correlation [2] as merit functions. We used the NEUWOA algorithm proposed by Powell [12] for the optimization.

2.3 Evaluation methodology

The evaluation consists of two steps. The first step is to determine the angle interval for which the tracking is feasible. This was done by applying the 2D/3D registration for each of the CBCT images and comparing the obtained results with the annotated diaphragm motion. The range of angles where tracking is possible was determined to be the range where the correlation of tracking and annotated motion is good. This was determined by visual inspection of the plots shown in Figure 2 and by cross checking with inspection of the CBCT images.

The second step is the creation of a VMAT plan for the partial angle arcs where tracking is possible. Since the kV images are acquired with a fixed 90 degree angle in relation to the treatment beam (gantry), the angle arcs used for planning are perpendicular to the imaging angles. The generated plan is then compared with a clinically "conventional" full arc VMAT plan. If the the partial arc plan fulfills the clinical requirements for the given treatment it is then considered accepted.

3 Results

Table 1 summarizes the tracking results for all three patients. For all cases, two arcs were defined with the angle intervals where the tumor could be tracked.

The angles were always with imaging roughly around anterior-posterior (AP) and posterior-anterior (PA) directions. The table shows for each patient, the angle interval for imaging around AP, the interval for imaging around PA and the total angle range for which imaging is available. All but one of the partial arcs are over 90 degrees and the total range of angles where tracking is possible exceeds 180 degrees for all three patients. Figure 2 shows for each patient plots of the full tracking sequence from the CBCT projections which was used to determine the arc ranges. Finally, figure 3 shows the obtained plans for full arc (left) the partial arcs (middle) and the corresponding DVHs (right). As it can be seen, the DVHs of the full and partial arc plans are comparable in terms of both target coverage and OAR sparing. The dose distributions for the partial plans are in general acceptable though for patient 1 there is an increase of low dose to the heart which should be better investigated.

	Angles (deg)				Total (deg)
Patient	AP		PA		
1	-136	-25	32	155	234
2	-160	-65	20	108	183
3	-141	-48	52	149	190

Table 1. Summary of the angle intervals for which tumor tracking was possible. For each of the patients two arcs were obtained, one around the AP and the other around the PA directions. For these two partial arcs, the minimum and maximum angle where tracking was possible is indicated. The last column shows the total angle range.

4 Discussion and Conclusions

The results obtained in our study suggest that it is clinically feasible to perform VMAT treatments with partial arcs where tumor imaging is possible.

In terms of tumor tracking feasibility, in all cases the total range of angles where imaging is possible exceeded 180 degrees that is, equivalent to over a half of a full arc. The imaging angles are always around the AP or PA directions which is to be expected as this is the direction where the heart or other structures within the mediastinum are not occluding the tumor.

During actual treatment, scatter from the treatment beam will have a negative effect on the kV images so they will actually contain more noise than these CBCT projections. Nevertheless, in our previous work, the same scatter was present when using simultaneous kV-MV acquisition but the additional noise had a small impact on tracking accuracy [7].

The treatment plans that were generated based on the tumor visibility show DVHs which are clinically acceptable. The dose distributions however need to be closely analyzed. In the case of patient 1, there is an increased low dose

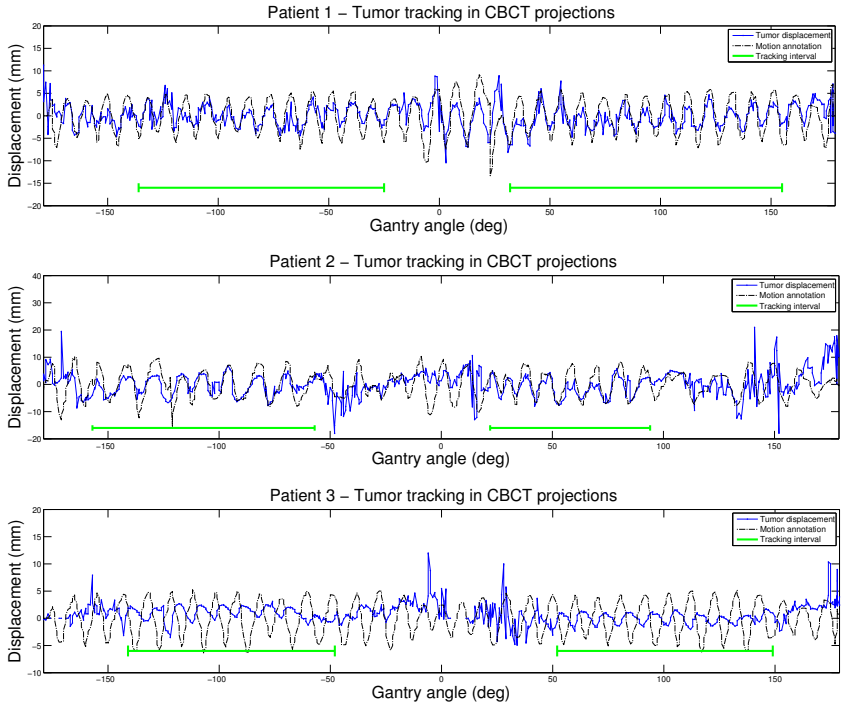


Fig. 2. Tumor tracking in the full range of 360 degrees (657 CBCT x-ray projections). The tracking results (blue line) are compared to the annotated diaphragm motion (dashed black line). The angles where the tumor could be tracked is shown in green.

delivered to the heart. Since there is no contour of the heart on this dataset (as it was not considered an OAR for the SBRT planning), is was for the time being not possible to evaluate the amount of delivered dose. With a contour it would be possible to optimize treatment planning to minimize the dose. For this and subsequent cases, relevant organs will be additionally delineated and used as constraints in treatment planning. In all three cases higher doses were observed close to the spinal cord. Though the study is promising, these plans have to be considered as preliminary and need careful plan optimization and clinical validation by a radiation oncology expert.

In terms of clinical implementation, tracking the tumor position in the CBCT projections was done retrospectively and can therefore be seen as a planning step. The tracking method used was the same which is to be used intra-fractionally which means that as long as there are no important anatomical changes, the tumor can be tracked during treatment for the same angles. In terms of workflow this means that for each patient, a partial arc and a full arc plans would be generated and the decision of which one to use would be based on the CBCT acquired immediately before treatment.

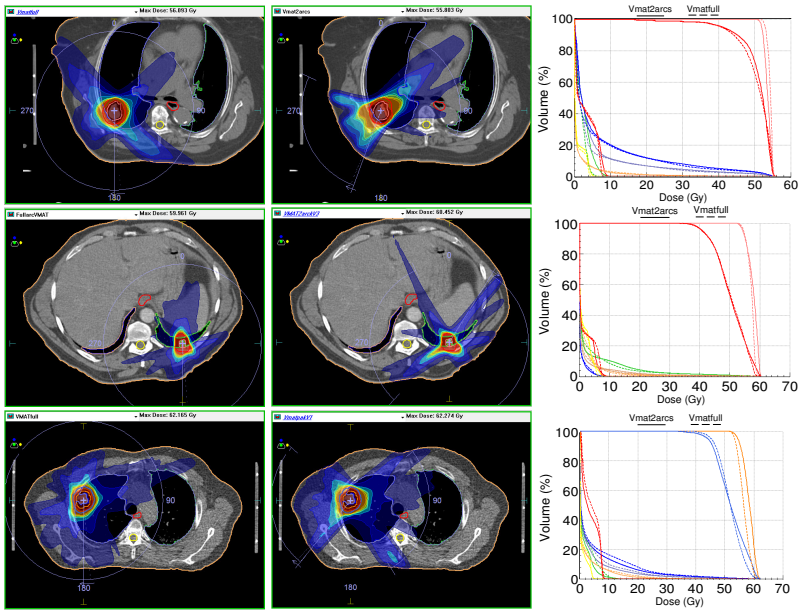


Fig. 3. Comparison of full arc VMAT (left) partial arc VMAT (middle) dose distributions and corresponding DVHs (right) for the three patients in the study.

Also, in a clinical scenario, imaging and registration latency will have to be taken into account. The imaging latency could be up to a few hundred milliseconds while one registration takes slightly less than 100 ms. This means that a motion prediction scheme will have to be used to estimate the true tumor position at a given time.

The method shown here is dependent on the availability of a CBCT prior to irradiation. CBCT is commonly available as it is used for patient setup. In cases where it might not be available, an alternative approach could be used. The planning CT can be used to generate DRRs in the full 360 degree arc. These DRRs could then be the base to asses on which angle arcs tumor tracking would be possible.

The final aim will be to create treatment plans with reduced PTV margins. As the tumor position will be known during treatment, margins can be reduced and the plan delivered as long as the tumor moves within the expected limits. If this is not the case, the beam can be turned off until the tumor position is again within limits. In the future, the dosimetric benefits of margin reduction have to be weighted against the additional dose burden imposed by kV imaging and the additional planning time required to implement this approach.

Despite the preliminary nature of these results, they represent a very positive indication that tracking during VMAT treatments is feasible, bringing the benefit of shorter delivery times, high conformal dose distributions and reduced margins.

Acknowledgments. The financial support by the Federal Ministry of Science, Research and Economy and the National Foundation for Research, Technology and Development is gratefully acknowledged.

References

1. Azcona, J., Li, R., Mok, E., Hancock, S., Xing, L.: Development and clinical evaluation of automatic fiducial detection for tumor tracking in cine megavoltage images during volumetric modulated arc therapy. Medical Physics 40(3) (2013)
2. Birkfellner, W., Stock, M., Figl, M., Gendrin, C., Hummel, J., Dong, S., Kettenbach, J., Georg, D., Bergmann, H.: Stochastic rank correlation: A robust merit function for 2D/3D registration of image data obtained at different energies. Medical Physics 36(8), 3420–3428 (Aug 2009)
3. Cozzi, L., Dinshaw, K., Shrivastava, S., Mahantshetty, U., Engineer, R., Deshpande, D., Jamema, S., Vanetti, E., Clivio, A., Nicolini, G., Fogliata, A.: A treatment planning study comparing volumetric arc modulation with rapidarc and fixed field imrt for cervix uteri radiotherapy. Radiotherapy and Oncology 89(2), 180–191 (2008)
4. De Mey, J., Van De Steene, J., Vandenbroucke, F., Verellen, D., Trappeniers, L., Meysman, M., Everaert, H., Noppen, M., Storme, G., Bossuyt, A.: Percutaneous placement of marking coils before stereotactic radiation therapy of malignant lung lesions. J. Vasc. Interv. Radiol. 16(1), 51–56 (2005)
5. Fledelius, W., Worm, E., Elstrøm, U., Petersen, J., Grau, C., Høyer, M., Poulsen, P.: Robust automatic segmentation of multiple implanted cylindrical gold fiducial markers in cone-beam ct projections. Medical Physics 38(12), 6351–6361 (2011)
6. Furtado, H., Steiner, E., Stock, M., Georg, D., Birkfellner, W.: Ep-1665: 2d/3d registration for pre-treatment lung tumor motion analysis using cbct for intrafractional tracking. Radiotherapy and Oncology 111, S232
7. Furtado, H., Steiner, E., Stock, M., Georg, D., Birkfellner, W.: Real-time 2d/3d registration using kv-mv image pairs for tumor motion tracking in image guided radiotherapy. Acta Oncologica 52(7), 1464–1471 (2013)
8. Gendrin, C., Furtado, H., Weber, C., Bloch, C., Figl, M., Pawiro, S., Bergmann, H., Stock, M., Fichtinger, G., Georg, D., Birkfellner, W.: Monitoring tumor motion by real time 2d/3d registration during radiotherapy. Radiotherapy and Oncology 102(2), 274–280 (2012)
9. van Herk, M.: Errors and margins in radiotherapy. Semin Radiat Oncol 14(1), 52–64 (2004)
10. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P.: Multimodality image registration by maximization of mutual information. IEEE Trans. Med. Imag. 2, 187–198 (1997)
11. Markelj, P., Tomaževič, D., Likar, B., Pernuš, F.: A review of 3d/2d registration methods for image-guided interventions. Medical Image Analysis 16(3), 642–661 (2012)
12. Powell, M.: Nonconvex Optimization and Its Applications, chap. The NEWUOA sorftware for unconstrained optimization without derivatives, pp. 255–297. Springer US (2006)
13. Steiner, E., Furtado, H., Viehböck, C., Birkfellner, W., Georg, D., Stock, M.: Ep-1660: 2d/3d image registration for various gantry angles for online tumor motion tracking. Radiotherapy and Oncology 111, S229–S230

A Generalized Strategy for 3D Dose Verification of IMRT/VMAT Using EPID-measured Transit Images

Aiping Ding, Bin Han, Lei Wang, Lei Xing

Department of Radiation Oncology, Stanford University School of Medicine, Stanford, CA 94305

Abstract
The aim in this study is to develop a generalized strategy for 3D dose verification of IMRT and VMAT planes using EPID transit images in combination with Monte Carlo (MC) simulations. An EPID-based dosimetric verification procedure was developed to convert EPID-measured transit images into 2D exit photon fluence by de-convoluting with the MC-simulated EPID response kernels. The present scatter from the phantom to the EPID was iteratively corrected by using a series of pencil beam scatter kernels derived from MC simulations. The primary fluence is therefore yielded by subtracting the corrected scatter from the total reconstructed exit fluence and used to reconstruct the dose distribution in multiple 2D planes parallel to the EPID by convoluting with the pencil beam deposition kernels. After summing up all the reconstructed 2D dose planes, the 3D dose distribution is obtained. The EPID-based dosimetric system was validated using 6 MV photon beam available from Varian TrueBeam STXTM. The results show that the EPID-based dosimetry system developed in this study is an accurate and robust tool for dose verification of IMRT/VMAT plans.

Keywords:

EPID, transit image, Monte Carlo, dose verification

1 Introduction

With more than 1.6 million new diagnoses of cancer and over one-half million projected to die of cancer in 2013, in the U.S. alone, almost 1,600 people died per day and cancer has been becoming a major health problem [1]. Radiation therapy (RT) is a major modality in cancer treatment and approximately 60% of all cancer patients in the U.S. receive RT as therapy or for palliation as an adjunct to surgery or chemotherapy. Over past two decades, intensity-modulated radiation therapy (IMRT) and volumetric-modulated arc therapy (VMAT) have become the mainstay for treating different types of cancers because of its capability of producing highly conformal dose distribution [2-9]. Practically, IMRT is composed of 5 to 10 radiation beams, with the intensity of each

adfa, p. 1, 2011.

beam modulated by a computer-controlled dynamic movement of a multi-leaf collimator (MLC). In a VMAT delivery [10-12], the position and speed of MLCs, and the gantry rotation vary dynamically [13]. While IMRT or VMAT offers a valuable tool for enhancing the therapeutic ratio and shows significant potential for improved survival and treatment outcome, their treatment planning and delivery are more technically involved and prone of errors due to dramatically increased complexity of the process. It is widely recognized that the efficacy of IMRT/VMAT can only be fully exploited with an effective quality assurance (QA) procedure to ensure the safe and efficient delivery of the exquisite dose distributions. However, current IMRT/VMAT QA procedure is labor intensive and inefficient. Moreover, it is unsafe and may lead to wrong conclusion as the measurement data are collected on one or two 2D planes, instead of 3D volume. The objective of this study is to develop a clinically practical 3D pretreatment dose verification for rapid IMRT and VMAT QA using a high spatial-resolution and high frame rate a-Si EPID and transit images in combination with Monte Carlo (MC) simulations.

2 Materials and methods

2.1 The portal electronic portal imaging device

The EPID used in this study is a standalone portable PerkinElmer XRD-0822 AP20 a-Si flat panel detector (PerkinElmer, Sunnyvale, CA). The size of detector was 20.48 × 20.48 cm^2, with a matrix of 1024 ×1024 pixels and a minimum pixel size of 0.2 mm. Its maximum frame rate is 50 frames per second (fps). The images were acquired in a "cine-mode" and a PerkinElmer image acquisition software XIS (version 3.0, PerkinElmer, Waltham, MA) was used to acquire and process all the EPID images. Measurements were performed on a Varian TrueBeam Stx Linac (Varian Medical Systems, Palo Alto, CA) for 6 MV photon beam. A source to detector distance (SDD) of 130cm and source to axis distance of 100cm were used. The thickness of the water-equivalent slab phantom was set to 20 cm and was positioned at the iso-center of the linear accelerator.

2.2 Corrections of the EPID raw images

Before images acquisition, a dark field (DF) image and a flood field (FF) image were acquired for offset and gain corrections. The offset correction took into account the dark current of each pixel and acquired with photon beam off. In order to create the offset correction image, an averaged image ($\overline{EPID}_{DF}$) of 300 frames of DF images had to be acquired and would be subtracted from the incoming pixel data during acquisition time. To homogenize differences in pixel sensitivities, an FF gain correction was carried out by irradiating the EPID with the incident photon beam fully covering the entire detector sensitive field (20×20 cm^2). To create the FF image, an averaged image ($\overline{EPID}_{FF}$) of 300 frames of offset-corrected images has to be acquired. Each EPID-measured raw image is corrected by using the following equation:

$$EPID_{raw}|_{corrected} = \frac{EPID_{raw} - \overline{EPID}_{DF}}{\overline{EPID}_{FF} - \overline{EPID}_{DF}} \quad (1)$$

2.3 EPID images to incident fluence conversion kernel

To determine the incident photon beam fluence, it was necessary to simulate and calibrate the EPID device to establish a relationship between EPID pixel values and radiation dose. Detailed structure and composition of the EPID were provided by the manufacturer and were modeled using the GATE (Geant4 Application for Tomographic Emission) [14] to generate a deconvolution kernel $K_{de}(x, y)$. The incident photon fluence $\Phi_M(x, y)$ on the EPID can thereafter be reconstructed from the corrected EPID raw image using the flowing equation:

$$\Phi_M(x, y) = (\overline{EPID}_{raw}|_{corrected}(x, y)) \otimes^{-1} (K_{de}(x, y)) \quad (2)$$

2.4 Scatter prediction kernels

When the EPID was placed close to the phantom, a large amount of scattered radiation is incident on the EPID surface. Therefore, the EPID-measured transit fluence (Φ_M) behind a slab phantom comprises primary Φ_p (un-scattered) and scattered contribution (Φ_S), that is

$$\Phi_M = \Phi_p + \Phi_S \quad (3)$$

In order to reconstruct the dose distribution with the phantom using the transmit images Φ_M, those scatter must be removed firstly.

MCNPX [15] was used to produce a series of MC scatter kernels, which allows scatter fluence predications from uniform water slab phantoms exposed to a divergent beam. It simulates a pencil beam impinging upon a slab phantom from divergent angels separately. The phantom is water-equivalent with a thickness of 20 cm. The scattered radiation present in the EPID surface was estimated point-by-point with the MC-generated pencil beam scatter kernel and was iteratively scatter-corrected using the MC simulated scatter kernels which give the primary fluence Φ_p at the plane of the EPID.

2.5 Dose deposition kernels

The extracted primary fluence at the EPID plane is scatter-corrected and converted to 2D dose distributions within the phantom in multiple planes parallel to the EPID. By summing up the 2D dose planes, the 3D dose distribution is obtained. A series of pencil-beam dose kernel $K_{pb}(x, y)$ was simulated using the MCNPX code at different depths of the slab phantoms for the dose reconstruction.

3 Results

A de-convolving dose kernel $K_{dp}(x, y)$ was generated from GATE MC simulations with the consideration of the MV photon dose deposition in the EPID screen, the optical photon creating and scattering process, as shown in Fig.1. The incident photon fluence on the EPID was therefore reconstructed from the corrected EPID raw images.

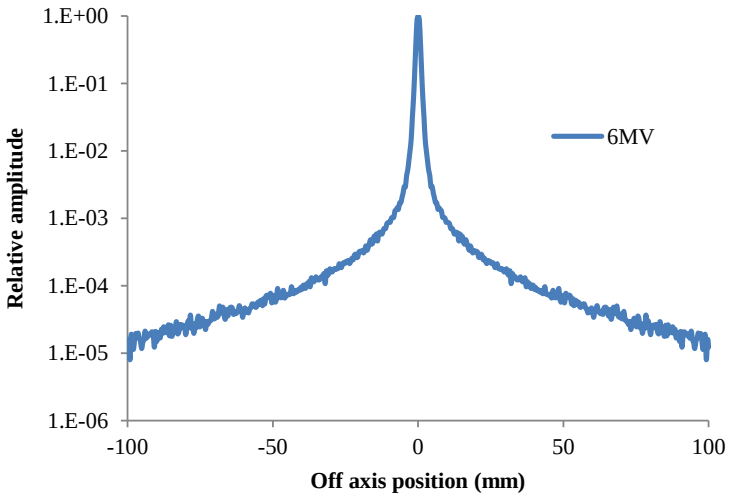


Fig. 1. The GATE generated kernel used in de-convolution of EPID-measured raw images to incident photon fluence.

Fig.2 shows an example of MCNPX simulated scatter prediction kernel. During the MC simulation, two separated tallies were simultaneously used to record the fluence from scatter only and primary plus scatter on the EPID plane.

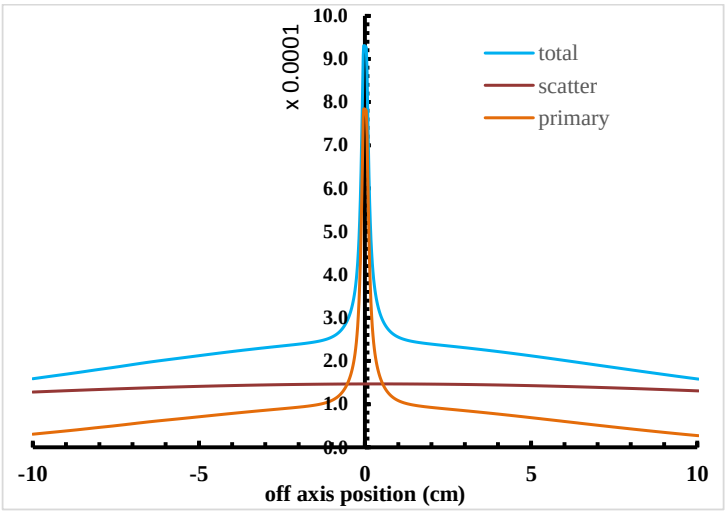


Fig. 2. An example of MCNPX simulated scatter prediction kernel used to remove the scatter from the reconstructed transit fluence.

Fig.3. shows the simulated $K_{pb}(x, y)$ of pencil beam used in the dose reconstruction.

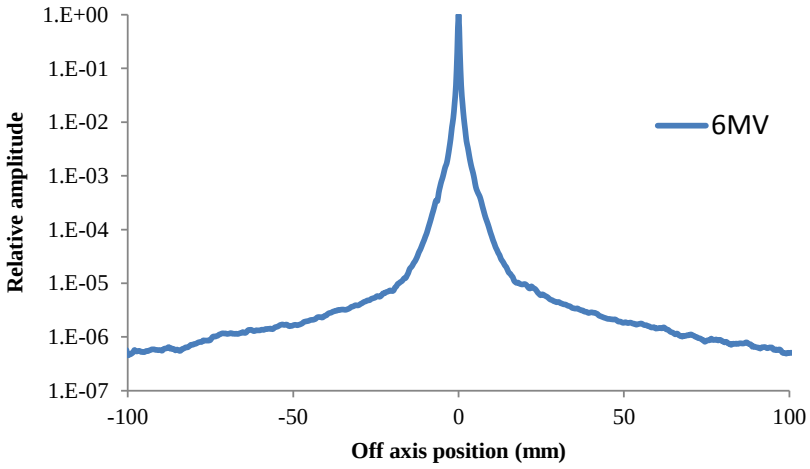


Fig. 3. The MCNPX generated dose kernel used in the dose reconstruction

To evaluate the performance EPID dosimetry system on photon beam applications, the absolute 2D dose distribution of square fields of 4×4 to 15×15 cm^2 fields were tested firstly against water scan results and PTW729 ion chamber array measurements, as shown in Fig.4.

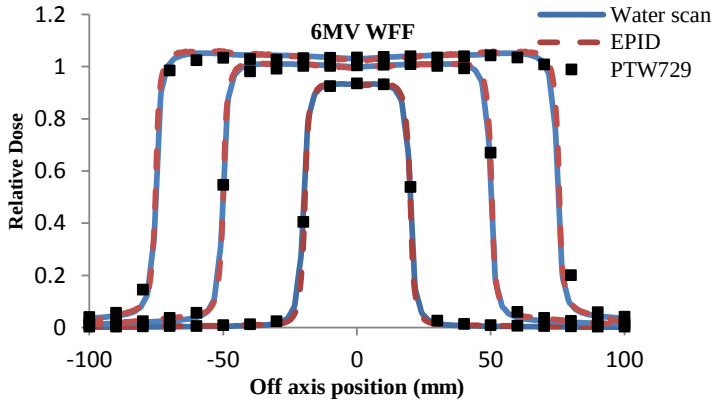


Fig. 4. Water-scan, PTW729, and EPID-measured dose profiles

To further validate the EPID measurement, one typical patient-specific case was delivered and compared with TPS calculation. Overall, as shown in Fig.5., the comparison data showed good agreement for both cases. EPID measurements vs TPS calculation, the γ-index pass rates were greater than 99% for criterion of 3%/3mm in the selected dose plane.

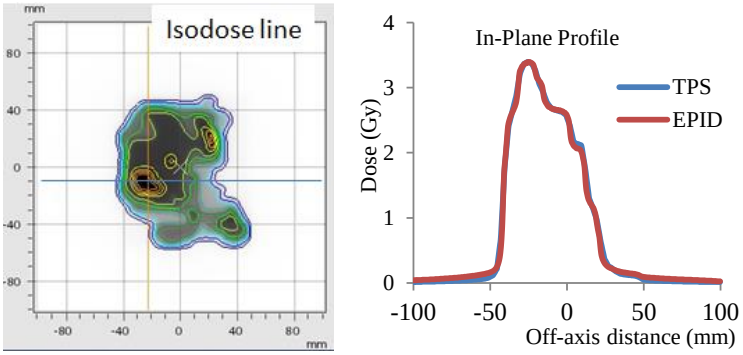


Fig. 5. Isodose line overlay of EPID measurements with TPS calculation and in-plane profiles of EPID and TPS calculation for the tested patient case.

4 Conclusions

We have developed a generalized procedure for dose verification of IMRT and VMAT using EPID transit images in combination with MC simulations. It provides a viable solution to the unmet need for a 3D dosimetric tool for IMRT/VMAT plan validation and for a number of intractable dosimetry problems, such as small fields and fields with high dose rates.

5 References

1. American Cancer Society Annual Report. http://www.cancer.org/research/cancerfactsstatistics/cancerfactsfigures2013/index (Access date: 01/13/2014).
2. Gregoire, V. and T.R. Mackie, State of the art on dose prescription, reporting and recording in Intensity-Modulated Radiation Therapy (ICRU report No. 83). Cancer Radiotherapie, 2011. 15(6-7): p. 555-559.
3. Gaspar, L.E. and M.S. Ding, A Review of Intensity-Modulated Radiation Therapy. Current Oncology Reports, 2008. 10(4): p. 294-299.
4. Sanghani, M. and J. Mignano, Intensity modulated radiation therapy: A review of current practice and future directions. Technology in Cancer Research & Treatment, 2006. 5(5): p. 447-450.
5. Nutting, C., D.P. Dearnaley, and S. Webb, Intensity modulated radiation therapy: a clinical review. British Journal of Radiology, 2000. 73(869): p. 459-469.
6. Mell, L.K., A.K. Mehrotra, and A.J. Mundt, Intensity-modulated radiation therapy use in the U.S, 2004. Cancer, 2005. 104(6): p. 1296-1303.
7. Sheets, N.C., et al., Intensity-Modulated Radiation Therapy, Proton Therapy, or Conformal Radiation Therapy and Morbidity and Disease Control in Localized Prostate Cancer. Jama-Journal of the American Medical Association, 2012. 307(15): p. 1611-1620.
8. Mell, L.K., A.K. Mehrotra, and A.J. Mundt, Intensity-modulated radiation therapy use in the U.S., 2004. Cancer, 2005. 104(6): p. 1296-1303.
9. Sheets, N.C., et al., INtensity-modulated radiation therapy, proton therapy, or conformal radiation therapy and morbidity and disease control in localized prostate cancer. JAMA, 2012. 307(15): p. 1611-1620.
10. Mackie, T.R., et al., Tomotherapy - a New Concept for the Delivery of Dynamic Conformal Radiotherapy. Medical Physics, 1993. 20(6): p. 1709-1719.
11. Yu, C.X., Intensity-Modulated Arc Therapy with Dynamic Multileaf Collimation - an Alternative to Tomotherapy. Physics in Medicine and Biology, 1995. 40(9): p. 1435-1449.
12. Otto, K., Volumetric modulated arc therapy: IMRT in a single gantry arc. Medical Physics, 2008. 35(1): p. 310-317.
13. Wolff, D., et al., Volumetric modulated arc therapy (VMAT) vs. serial tomotherapy, step-and-shoot IMRT and 3D-conformal RT for treatment of prostate cancer. Radiotherapy and Oncology, 2009. 93(2): p. 226-233.
14. Jan S et al 2010 GATE V6: a major enhancement of the GATE simulation platform enabling modelling of CT and radiotherapy Phys. Med. Biol. 56 881-901
15. D.B. Pelowitz, MCNPX USER'S MANUAL. (Los Alamos National Laboratory, Los Alamos, NM, 2005).

Simulation of abdominal MRI sequences in a computational 4D phantom for MRI-guided radiotherapy[A]

Paganelli C.[1], Summers P.[2], Bellomi M.[2,3], Baroni G.[1,4] and Riboldi M.[1]

[1]Dipartimento di Elettronica Informazione e Bioingegneria, Politecnico di Milano, Milano, Italy.
[2]Division of Radiology, Istituto Europeo di Oncologia, Milano, Italy.
[3]Department of Health Sciences, Università di Milano, Milano, Italy.
[4]Bioengineering Unit, Centro Nazionale di Adroterapia Oncologica, Pavia, Italy.

Abstract. A preliminary simulation of MRI sequences to derive a 4D MRI phantom for abdominal organs is presented, with the aim of providing a framework for the evaluation of MRI-guided methods in external beam radiotherapy. Specifically, we propose an extension of the 4D NCAT phantom including tissue parameters obtained via T1 and T2 DESPOT sequences, the simulation of dedicated abdominal MR sequences such as VIBE and TrueFISP, the modeling of radiofrequency coil response and noise followed by k-space sampling and image reconstruction. Analysis of tissue parameters and reconstructed images were performed to show the robustness of the implemented phantom.

Keywords. 4D MRI phantom, abdominal MRI, MRI reconstruction

1 Introduction

Over the last few years, there has been growing interest in the use of Magnetic Resonance Imaging (MRI) in image guided radiotherapy [1,2]. Due to the absence of ionizing radiation and increased soft tissue contrast relative to Computer Tomography (CT), MRI is an attractive technology for target definition [2]. Moreover, it has sufficient temporal resolution to study organ motion due to respiration [3]. Early studies investigating the use of MRI in treatment planning for organ motion quantification mainly followed cineMRI approaches [4,5], in which few slices are acquired allowing the description of a significant amount of breathing cycles. New four-dimensional MRI (4D MRI) retrospective sorting methods were also proposed to provide volumetric information of respiratory motion [6,7]. These features of MRI motivated also several recent technological developments towards the integration of MRI with radiation therapy treatment units, raising the prospect of fully MRI-guided treatments [8].

A standard approach to validation is the use of acquired in-vivo images as ground truth [9], which are not always available. In 4D MRI sorting, for example, a full resolu-

[A] **Acknowledgment:** The author would like to thank the Associazione Italiana per la Ricerca contro il Cancro (AIRC) for the support and Pifferi A. for the help during acquisitions.

tion 4D MRI cannot be obtained due to limited spatial and temporal trade-offs. Other solutions involve scanning MRI-compatible moving-structure phantoms [7,10], which do not often reflect the real internal anatomy. An MRI-compatible anthropomorphic moving phantom is under development [11], but the use of physical phantoms is generally limited due to their cost and the lack of flexibility to mimic the range of motions encountered in-vivo. Numerical digital phantoms offer a practical approach to evaluate and determine optimized methods. Examples have already been proposed for CT with the 4D cardiac-torso (NCAT) phantom, which incorporates natural beat-to-beat heart rate and respiratory motion variations, and then the extended (XCAT) version [12]. A number of MRI simulators have been developed, mainly for application to brain imaging, based on Bloch equations [13,14], but largely neglecting organ motion effects. The first implementations of numerical phantoms that incorporate motion in MRI for body radiotherapy were based on the extension of the 4D NCAT (and its extended XCAT version) by assigning MR properties to each tissue mask. Sharif et al. [15] proposed a physiologically improved NCAT (PINCAT) phantom in which the signal intensities were modified for MRI application to validate a dynamic MR imaging scheme in real-time cardiac MRI. In [16], a moving phantom was generated by segmenting cardiac images from an in-vivo acquisition by using a random affine transformation to simulate motion. However, in these cases a gray level to each tissue signal was also assigned, i.e. neglecting specific tissue properties. Wissmann et al. [17] designed a more realistic numerical phantom for cardiovascular MRI (MRXCAT) by extending the 4D XCAT phantom to MRI. They applied different tissue signal, multiple receiver coils and noise models, and selected arbitrary trajectories and undersampled acquisition of the k-space for accelerated cine and myocardial perfusion imaging, with specific cardiac MRI sequences.

In this work, we propose the basis for the construction of a 4D MRI phantom based on the 4D NCAT with an approach similar to the one proposed by Wissmann et al. [17] but extending the phantom to the abdominal site (abdoMRCAT) to account for organ motion due to respiration.

2 Materials and Methods

As proposed by Wissmann et al. [17], the phantom $P(\vec{k},t)$ is described in k-space through a combination of several weighting functions:

$$P(\vec{k},t) = R \cdot F \cdot [S(N_{coil}) \cdot T(T1,T2,\rho) \cdot C(TE,TR,\alpha) \cdot O(\vec{x},t) + n(\vec{x},t)]$$

where $O(\vec{x},t)$ represents the 4D NCAT phantom defined in space and time and $n(\vec{x},t)$ is the noise model. T is the tissue contribution as a function of relaxation times T1 and T2 and proton density ρ. The MR sequence is described by the operator C, which expresses the acquired signal as a function of repetition time TR, echo time TE and flip angle α. S describes the sensitivity of N_{coil} coils. These physical space functions undergo Fourier transformation F and the sampling of k-space R is applied to produce the raw k-space phantom $P(\vec{k},t)$. Once a complete k-space representation has been obtained, image reconstruction of the phantom can then follow via inverse Fourier transformation, as in the case of data sampled during image acquisition.

2.1 Tissue parameters T and MR sequences C

Two pulse sequences that are typically performed in abdominal MRI imaging were of interest for this study: (1) a T1-weighted spoiled volumetric interpolated breath-hold sequence (VIBE) [18] to acquire a 3D volumetric image in breath-hold and (2) a T2-weighted balanced steady-state free precession sequence (TrueFISP) [19] used during free-breathing to repeatedly acquire fast 2D images able to describe respiratory motion. Specifically at steady-state, the signal equations of the two different sequences C applied to each organ mask according to the specific relaxation times T1, T2 and proton density ρ values were as follows:

(1) VIBE: $C = \frac{\rho \sin\alpha(1-e^{\frac{-TR}{T1}})}{(1-\cos\alpha e^{\frac{-TR}{T1}})} e^{\frac{-TE}{T2}}$

(2) TrueFISP: $C = \rho \sin\alpha \frac{1-e^{\frac{-TR}{T1}}}{1-(e^{\frac{-TR}{T1}}-e^{\frac{-TR}{T2}})\cos\alpha - e^{\frac{-TR}{T1}} e^{\frac{-TR}{T2}}} e^{\frac{-TE}{T2}}$

The target sequence for in-vivo T1 and T2-weighted acquisitions (Fig. 1) used the following parameters (1.5T Siemens Magnetom Avanto):
- axial VIBE: TR/TE: 4.8msec/1.75msec; α: 68°; bandwidth: 601Hz per pixel; scan matrix: 256x224 pixels with spacing of 1.28x1.28mm; slice thickness of 5mm; acquisition time: 180msec/slice.
- sagittal TrueFISP: TR/TE: 2.9msec/1.26msec; α: 10°; bandwidth: 252Hz per pixel; percentage sampling: 70%; scan matrix: 240x320 pixels with spacing of 1.25x1.25mm; slice thickness of 4mm.

The imaging parameters of the acquired in-vivo images (TR, TE and α) were used in equations (1) and (2) for the generation of the abdoMRCAT images.
The dominant properties determining tissue appearance in MR imaging are T1 and T2 relaxation times and the proton density of the tissue. Proton density values ρ were not available for all abdominal organs in the literature [20], whereas relaxation times T1 and T2 were well defined in different works [21]. Therefore, in a first approach, an estimation of ρ (i.e. estimated ρ) was derived from in-vivo T1-weighted (i.e. VIBE) and T2-weighted (i.e. TrueFISP) acquisition, by segmenting a region of interest in each organ, and deriving ρ from the sequence equations (1) and (2) and performing a mean between the two ρ values. However, in order to avoid dependency from other factors (such as noise and potential artifacts), specific MRI sequences were acquired to estimate both T1, T2 and ρ values. Driven equilibrium single pulse observation of T1 (DESPOT1,11 flip angles in the range 3°-45°) and T2 (DESPOT2, 8 flip angles in the range 8°-64°) as described in [22] were acquired on the abdomen of one healthy volunteer. These acquisitions were used to derive T1-map, T2-map and ρ-map.
T1, T2 values were compared with the literature and ρ values derived from DESPOT acquisition were compared to the estimated values from equation (1) and (2) of in-vivo VIBE and TrueFISP (Wilcoxon test, alpha=5%). Differently from Wissmann et al. [17], no contrast agent concentration was considered.

2.2 Coil operator S and noise n

The effect of the coil sensitivity operator was implemented as proposed by [16], in which the phantom signal $(T(T1,T2,\rho)\cdot C(TE,TR,\alpha)\cdot O(\vec{x},t))$ was combined with a

simulated sensitivity map of the coil. The maps were designed as a linear fall-off with the sum of squares was equal to 1 and a circular arrangement of the coils around the abdomen (N_{coils}=8).
The presence of noise in MRI images was simulated by adding a Gaussian noise in order to satisfy the signal to noise ratio of in-vivo acquisitions (SNR=20).

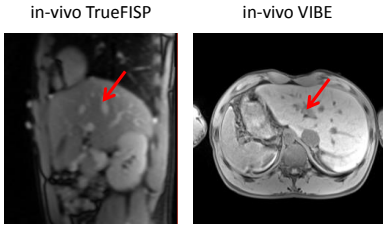


Fig. 1. In-vivo acquisitions. T2-weighted TrueFISP (left) and T1-weighted VIBE (right). Red arrows indicate blood.

2.3 k-space generation *F* and k-space sampling *R*

The image-domain model was transformed to the k-space domain via the discrete Fourier transform. Then, the k-space was sampled according to the sequence sampling approach. Specifically:
- VIBE [18]: the k-space was filled with a 70% sampling factor in the phase encoding (PE) directions. The remaining k-space was filled with zeros to provide smaller voxel size and maintain short imaging time.
- TrueFISP [19]: the k-space sampling was based on a generalized auto-calibrating partially parallel acquisition (GRAPPA) [23], with the aim to speed up the MRI pulse sequence in order to acquire fast slices able to describe the respiratory cycle(i.e. acquisition time ranged from 180msec to 300msec, as proposed in the literature [5]).
In the GRAPPA algorithm, unsampled k-space lines (i.e. sampling factor = 2, i.e. acquisition of interleaved lines) are synthesized by a linear combination of acquired neighboring k-space data (N_{FE}=256 and N_{PE}=256) using spatial information contained in the coil elements. The acquisition of additional lines in the k-space center is a form of self-calibration (i.e. auto-calibration (AC) lines, N_{ac}=16).

2.4 Image reconstruction

Once all k-space (3D for VIBE and 2D for TrueFISP) samples were obtained for a particular coil, an inverse Fourier transform was used to generate the uncombined image for that coil. The full set of N_{coils} uncombined images can then be combined using a normal sum of squares reconstruction [23].
Reconstructed images of all respiratory phases were qualitatively compared to in-vivo T1-weighted (VIBE) and T2-weighted (TrueFISP) acquisitions. Errors were quantified by computing the difference between the reference phantom and the reconstructed images.

3 Results

Table I shows the T1 and T2 values reported in the literature and ρ values derived from VIBE and TrueFISP equation (i.e. estimated ρ). In addition, T1, T2 and ρ values quantified via DESPOT acquisitions are also reported. Background and air were set to 0. Blood was not measured via DESPOT because of flow effects that cannot be compensated with the acquisition.T1 and T2 values derived from DESPOT differed from the literature [21] of 28% and 38% on average across tissues respectively, and the two populations (i.e. T1 and T2 literature vs. T1 and T2 DESPOT) did not result significantly different. The difference between ρ values derived from DESPOT and the ones estimated from VIBE and TrueFISP equations was as 37%. No significant difference was also observed between these two groups (i.e. estimated ρ vs. DESPOT ρ). Bowel, spleen and heart yielded higher estimated ρ values than the ones obtained with DESPOT. Values for ρ obtained with DESPOT for liver, spleen and bone were closer to literature ones [20] with respect to the estimated ρ.

Fig. 2 shows the abdoMRCAT images obtained by changing the proton density value of blood ρ_{blood} by using the estimated blood value from TrueFISP (12766) and VIBE (1397) separately and by using the mean of the two acquisitions (7081) for both sequences. A comparison with an in-vivo acquisition (Fig.1) shows that in the abdoMRCAT with a ρ_{blood} specific for each sequence, blood in the heart for the phantom resulted more similar to blood in vessels shown in in-vivo acquisitions (i.e. dark for VIBE and bright for TrueFISP) with respect to using the mean ρ_{blood} value (i.e. 7081) for both TrueFISP and VIBE simulations.

A quantitative analysis was also performed by computing the error as difference between the original phantom and the reconstructed one for all the respiratory phases (Fig. 3).The error maps showed pronounced edge effects for VIBE, whereas a more homogeneous distribution was maintained for TrueFISP. The errors obtained computing the difference between the original phantom and the reconstructed one were 10% and 7% (mean among all the phases) for VIBE and TrueFISP respectively.

Table 1. Tissue parameters. First and second columns: T1 and T2 values [msec] in the literature. Third column: estimated proton density ρ values [a.u.] from VIBE and TrueFISP acquisitions (literature values in brackets). Last columns: T1, T2 and ρ obtained from DESPOT maps.

	T1 [msec] literature	**T2 [msec] literature**	**estimated ρ (literature value) [a.u.]**	**T1 [msec] DESPOT**	**T2 [msec] DESPOT**	**ρ [a.u.] DESPOT**
background	0	0	0	0	0	0
air lung	0	0	0	0	0	0
body	240	85	1198	376	30	1336
bowel	100	10	1325	122	8	117
muscle	900	50	3342	825	28	3195
kidney	650	70	3832	921	40	1972
heart	1000	20	6150	1032	20	1346
liver	420	45	2797 (2182)	506	30	2023
blood	1500	20	7081	-	-	-
spleen	1514	65	4968 (2088)	1466	52	1428
cartilage	1060	35	977	588	16	1100
bone	732	30	916 (1343)	753	36	1041

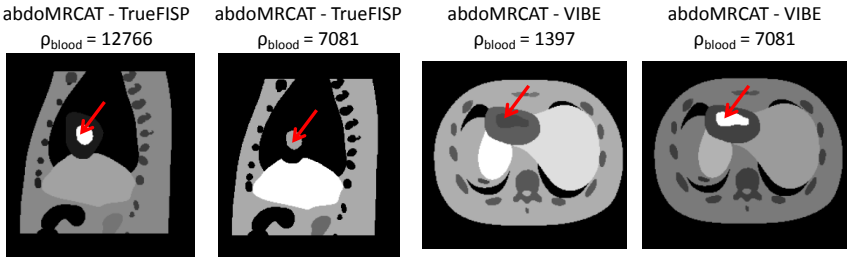


Fig. 2. Blood proton density. abdoMRCAT with a ρ_{blood} value selected in a sequence-specific fashion (i.e. 12766 for TrueFISP and 1397 for VIBE) and a mean value approach (7081) for both sequences. Red arrows indicate blood.

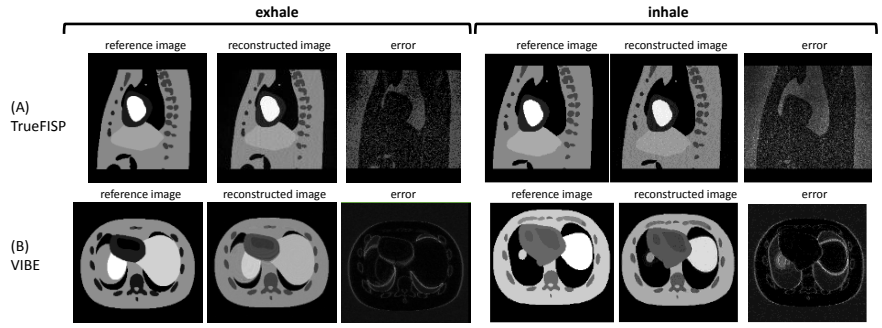


Fig. 3. Image reconstruction. (A) TrueFISP and (B) VIBE with the original reference MRI phantom, the reconstructed image and the error for both exhale and inhale.

4 Discussion and Conclusion

In this work, we describe the basis for the implementation of an abdominal 4D MRI phantom by extending the 4D NCAT phantom designed by Segars et al. [12]. A similar work has previously been reported in the literature by Wissmann et al. [17] for cardiac acquisitions. Our aim was to provide a phantom able to describe respiratory motion of abdominal organs in MRI that can be used for validation purposes.

The tissue parameters T1 and T2 obtained using specific MR sequences (i.e. DESPOT1 and DESPOT2) [22] were consistent with the literature [21], supporting their use in the phantom. Our acquisitions did not however take into account inhomogeneities which are recognized to influence T1 and T2 measurement. Subtle corrections in the T1 and T2 values may therefore be warranted in the future. Proton density values for abdominal organs appear to be largely absent from the literature. The values obtained from ρ-maps derived from the DESPOT data were compared with estimated values extracted from in-vivo T1-weighted and T2-weighted acquisitions (i.e. deriving ρ values from equation (1) and (2)): a mean discrepancy of 37% was found, with comparable values in all organs, without a significant difference between estimated ρ and DESPOT ρ. Spleen, bowel and heart ρ values were over-estimated by in-vivo T1/T2

acquisitions (4968). However for spleen, a lower value was observed in DESPOT (1428) resulting more similar to the literature (2088). Also bowel and heart presented lower values in DESPOT estimation, however these two organs were difficult to analyze due to artifacts associated with filling and contraction. Furthermore, we did not derive the proton density value for blood as flow effects [24] could not be adequately compensated for with the available DESPOT sequences, and in the presence of in/out-flowing blood, the signal equations (i.e. equation (1) and (2)) are not appropriate for the VIBE and TrueFISP signals. Moreover, the magnetization described by the signal equations does not include signal alteration due to motion during sampling [17]. In addition, steady-state free precession MRI as TrueFISP is a type of gradient echo MRI pulse sequence in which a steady, residual transverse magnetization is maintained between adjacent breathing cycles. Conversely, spoiled gradient echo MRI, such as VIBE, is an MRI technique which destroys residual transverse magnetization at the end of each excitation cycle. Blood and other fluids may therefore exhibit spoiled contrast behavior, even though stationary tissue remains in a steady-state free precession. For this purpose, in order to obtain simulations similar to in-vivo acquisitions, we applied to blood the specific estimated ρ value obtained via in-vivo T1 and T2-weighted acquisitions, allowing to have dark blood and bright blood in VIBE and TrueFISP respectively, even if blood flow is not adequately accounted for by their signal equations. In addition, the reconstruction errors were 10% and 7% for VIBE and TrueFISP respectively, with more errors along the edges being present in the VIBE reconstruction (Fig. 3) due to a central sampling of the k-space and to a volumetric interpolation in contrast to the calibrated and bi-dimensional reconstruction of the TrueFISP acquisition.

In conclusion, we presented the preliminary steps on the simulation of MRI sequences for the construction of an abdominal 4D MRI phantom that can be considered as a framework for the validation of MR image reconstruction and quantitative post-processing approaches to improve organ motion quantification and compensation. Furthermore, our work provided also a preliminary quantification of tissue parameters including proton density values of abdominal organs. Future works on tissue parameters are needed as MR-based measurements are dependent on sequence used and sources of error such as field inhomogeneity. Because a spin's history of radio-frequency pulses and relaxation intervals determines the appearance in MRI, a more extensive consideration of the object model is required in MR simulations. In addition, a deep analysis of the reconstruction effects on the 4D motion will be taken into consideration. Further extensions to the phantom include the incorporation of spin history and additional effects such as susceptibility and magnetization transfer.

5 References

1. Hugo G D, Rosu M. Advances in 4D radiation therapy for managing respiration: Part I – 4D imaging. Med Phys; 22:258–271 (2012)
2. Metcalfe P, Liney GP, Holloway L et al. The Potential for an Enhanced Role For MRI in Radiation-therapy Treatment Planning Technology in Cancer Research & Treatment; 1-18 (2013)
3. Biederer J, Hintze C, Fabel M, Dinkel J. Magnetic Resonance Imaging and Computed Tomography of Respiratory Mechanics. Journal of Magnetic Resonance Imaging; 32:1388–1397 (2010)

4. Sawant A, Keall P, Pauly KB, et al. Investigating the feasibility of rapid MRI for image-guided motion management in lung cancer radiotherapy. BioMed Res Int;2014:48506 (2014)
5. Paganelli C, Seregni M., Fattori G. et al. Magnetic Resonance Imaging Guided versus Surrogate-Based Motion Tracking in Liver Radiation Therapy: A Prospective Comparative Study, Int J Radiation Oncol Biol Phys; 91: 840-848 (2015)
6. vonSiebenthal M, Székely G, Gamper U et al. 4D MR imaging of respiratory organ motion and its variability. Phys Med Biol;52:1547–1564 (2007)
7. Cai J, Chang Z, Wang ZH et al. Four-dimensional magnetic resonance imaging (4D-MRI) using image-based respiratory surrogate: A feasibility study. Med Phys;38:63-84 (2011)
8. Keall PJ, Barton M, Crozier S. The Australian magnetic resonance imaging-linac program. Semin Radiat Oncol;24:203-206 (2014)
9. Weon C, Nam WH, Lee D, et al. Position tracking of moving liver lesion based on real-time registration between 2D ultrasound and 3D preoperative images, Medical physics, 42: 335-345(2015)
10. http://modusmed.com/qa-phantoms/mri-respiratory-motion
11. Perrin R, Peroni M, Bernatowicz K et al. A Realistic Breathing Phantom of the Thorax for Testing New Motion Mitigation Techniques with Scanning Proton Therapy. Med. Phys. 41, 111 (2014)
12. Segars PW, Sturgeon G et al. 4D XCAT phantom for multimodality imaging research. Med. Phys; 37: 4902-4915 (2010)
13. http://mrilab.sourceforge.net/
14. Benoit-Cattin H, Collewet G et al. The SIMRI project: a versatile and interactive MRI simulator. Journal of Magnetic Resonance 173: 97–115 (2005)
15. Sharif B and Bresler Y. Adaptive real-time cardiac MRI using PARADISE: validation by the physiologically improved NCAT phantom. Proc IEEE Int Symp Biomed Imaging (2014)
16. Aja-Fernandez S, Cordero-Grande L et al. A MRI phantom for cardiac perfusion simulation. IEEE ISBI (2012)
17. Wissmann L, Santelli C, Segars WP, Kozerke S. MRXCAT: Realistic Numerical Phantoms for Cardiovascular Magnetic Resonance. J Cardiovasc Magn Reson;16:63 (2014)
18. Rofsky NM, Lee VS et al. Abdominal MR Imaging with a Volumetric Interpolated Breath-hold Examination. Radiology, (1999)
19. Fuchs F, Laub G et al. TrueFISP: technical considerations and cardiovascular applications. European Journal of Radiology 46: 23-32 (2003)
20. Nyman R, Ericsson A, et al. T1, T2, and relative proton density at 0.35 T for spleen, liver, adipose tissue, and vertebral body: normal values. Magn Reson Med.;3(6):901-10 (1986)
21. Bernstein MA, King KF, Zhou XJ: Basic pulse sequences. Handbook of MRI Pulse Sequences (2004)
22. Deoni SCL, Rutt BK et al. Rapid Combined T1 and T2 Mapping Using Gradient Recalled Acquisition in the Steady State. Magnetic Resonance in Medicine 49:515–526 (2003)
23. Griswold M A, Jakob P M, et al. Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA). Magnetic Resonance in Medicine 47:1202–1210 (2002)
24. Axel L. Blood flow effects in magnetic resonance imaging. Magn Reson Annu. 237-44 (1986)

Flat-Field Correction Pipeline for a Cone-Beam Computed Tomography Imaging Device with Independently Movable Source and Detector

Peter Keuschnigg[1], Philipp Steininger[1,2], Horst Schoedl[2], Katja Presich[1], Daniel Kellner[1,2], Philipp Huber[1], Ulrich Mayer[1,2], Markus Mehrwald[1,2], and Heinz Deutschmann[1,2]

[1] Paracelsus Medical University, Institute for Research and Development on Advanced Radiation Technologies (radART), Strubergasse 16, 5020 Salzburg, Austria
[2] medPhoton GmbH, Strubergasse 16, 5020 Salzburg, Austria

Abstract. The ImagingRing, a novel cone-beam computed tomography (CBCT) imaging device with independently rotatable x-ray source and flat-panel detector arms, allows for planar and CBCT acquisitions of arbitrary regions inside the device's cylindrical imaging volume. Due to its special geometry, an adjustment of standard methods for generating flat-field planar images is required in order to fit the particular geometric conditions. This work gives an overview of the current implementation of the flat-field correction pipeline for the ImagingRing as well as an outlook to further planned improvements for CBCT enhancing quality.

1 Introduction

The ImagingRing [1,2], a cone-beam computed tomography (CBCT) imaging device, has recently been developed by medPhoton GmbH which is a spin-off company of the Paracelsus Medical University. This novel device is a couch-mounted scanner with independently rotatable x-ray source and flat-panel detector arms, which allow for special use cases in conventional and particle radiotherapy. The scanner is moveable along the couch to allow for covering all relevant anatomical regions and the x-ray beam is shaped with a source-mounted 2D collimator. Supporting independent rotations, the ImagingRing is capable of performing planar and CBCT x-ray acquisitions focusing on arbitrary regions inside the cylindrical imaging volume. Due to the ImagingRing's special geometry, an adjustment of standard methods for generating flat-field planar images is required in order to fit the particular geometric conditions.

This work gives an overview of the current implementation of the flat-field correction pipeline integrated in the open-radART ion software suite [4] for the ImagingRing as well as an outlook to further planned improvements.

2 Methods

The pipeline for flat-field correction of planar images acquired with the ImagingRing consists of several customized standard methods. In a first step the multi-level-

gain (MLG) correction is responsible for creating dose-equivalent images by accounting for the non-linear dose response of the flat-panel detector. A subsequent bad-pixel correction handles outlier intensities resulting from dysfunctional pixels or sectors of the detector. Afterwards, the images are oriented in the device's imaging coordinate system (ICS) using a flexmap calibration to account for the gravity-induced flex of the source and detector arms. After successful image orientation, the projections are corrected for the beam fluence of the x-ray source. Finally, beam attenuation caused by varying incidence angles and source-to-detector distance (SDD) is accounted for.

2.1 Geometry of the ImagingRing

The basic geometry of the ImagingRing is defined as shown in Figure 1. S denotes the position of the x-ray source focal spot and R denotes the flat panel detector image origin. The unit vectors v_1 and v_2 depict the detector row and column orientation vectors, respectively. The mechanical isocenter (ISO) is different from the virtual axis of rotation N. The SDD measured from S to the detector center C usually cannot be considered constant, due to the possibility of independent rotations for imaging a region around an arbitrary point N. Even in an opposite position, where the isocenter is on the line between S and C, v_1 is not perpendicular to this line to enable for large-field-of-view CBCT acquisitions. Hence, for simplification of post-processing (e.g. reconstruction) the images are projected on a consensus plane which is perpendicular to $\overline{SISO}$.

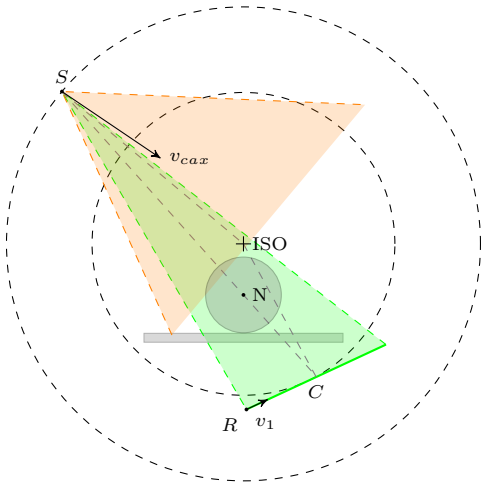
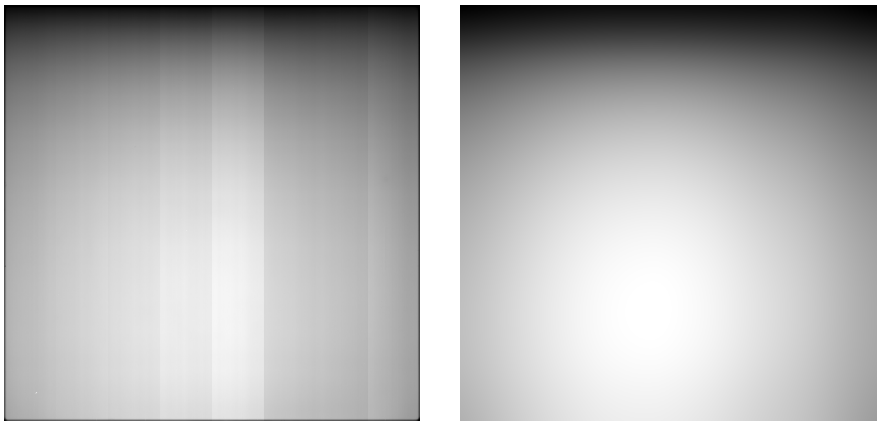


Fig. 1: ImagingRing geometry for an acquisition of a region around the effective imaging isocenter N, showing the focal spot position S, the detector origin R, the detector center C and orientation vector v_1 with respect to the table and a cylindrical phantom object.

2.2 Multi-Level-Gain Correction

As described in Huber et al. [3], MLG correction is used for correcting the nonlinear detector response, especially at low dose rates. For this method, an offset image without irradiation and multiple flood-field gain images at different dose rates are acquired. In contrast to the application for other imaging devices, the flood-field MLG calibration for the ImagingRing is more complex since a short SDD results in varying incidence angles of the x-rays at the detector. To overcome this limitation a slit collimation is utilized, allowing for irradiation of just a few columns of the detector while rotating the detector through the slit. Each column is irradiated with the same slit beam profile, avoiding the requirement of correcting for the inhomogeneous fluence of the x-ray tube in rotation direction. Since the source-mounted 2D collimator allows for arbitrary jaw positions, it can be used to form a slit with adjustable size. The mean value of the intensities for which a specific pixel on the detector has been irradiated then defines its intensity on the MLG calibration image as shown in Figure 2a. The beam profile at calibration time normalized in the range $[0.0, 1.0]$, which is visualized in Figure 2b, can be extracted from the MLG calibration image.



(a) MLG calibration image (b) calibration beam profile

Fig. 2: MLG calibration image (a) which has been generated from slit collimation acquisitions of the ImagingRing, showing different intensity levels for sub-panels. The beam-profile (b) extracted from the MLG image is used for correcting the beam-dependent intensity variations due to, e.g., the anode heel effect.

The correction method itself has been implemented similar to what is described in Huber et al. [3]. Using the MLG supporting points measured at multiple known dose-rates, a nonlinear mapping between a pixel intensity and a specific dose rate value can be performed, resulting in a dose-equivalent inten-

sity. For a given intensity at the detector I_d, the dose-equivalent intensity D_d is computed by linear interpolation between the closest MLG calibration images (high, low) followed by a correction for the calibration beam profile D_{calib}.

$$D_d = \frac{D_{low} + (I_d - I_{low})\frac{D_{high} - D_{low}}{I_{high} - I_{low}}}{D_{calib}} \quad : I_{low} \leq I_d \leq I_{high} \tag{1}$$

2.3 BadPixel Correction

A pixel on the detector p_d is considered bad if its intensity deviation from the surrounding pixel intensities U is greater than a given intensity threshold I_t. Static bad pixels are present in all acquisitions, whereas dynamic bad pixels appear and vanish at certain energy levels. Therefore, the calibration of this method consists of acquisition of homogeneous frames at multiple beam energies and subsequent evaluation of the outliers in each image. Hence, the resulting binary bad-pixel map in which non-zero pixels identify bad pixels is computed by:

$$I_b = \begin{cases} 0, & \text{if } |I_i - I_d| \leq I_t \quad \forall I_i \in U \\ 1, & \text{otherwise} \end{cases} \tag{2}$$

The correction method utilizes a spiral interpolation approach of n intensities of the original image for replacing the intensities of marked bad pixels defined in (2). The intensity sampling is started at a random angle and continues with golden angle increments and progressive distance p_d. For each interpolated intensity I_i, the Euclidean distance to p_d is used as a weighting factor w_i. The replacement intensity is then computed by the weighted mean value of the interpolated intensities given in (3).

$$I_c = \sum_{i=1}^{n} w_i I_i \quad \text{where } \sum_{i=1}^{n} w_i = 1 \tag{3}$$

2.4 Image Orientation and Flexmap Correction

For further corrections, the acquired image has to be oriented in the ICS. In order to account for dynamic movement of the source- and panel-arms due to gravity based effects, a 9-degrees-of-freedom (9-DOF) flexmap is applied. The 9-DOF flexmap specifies for combinations of source and panel arm positions a vector $[tS_x, tS_y, tS_z]$ for translations of S, a vector $[tC_x, tC_y, tC_z]$ for translations of C and a set of 3 Euler angles for transformation of v_1 and v_2. This approach is similar to what has been described in Warmerdam et al. [5] for investigating geometric uncertainties in 2D/3D image registration of cranial images.

2.5 Incidence Angle Correction

For a single ray emerging from the focal spot position S and passing through a specific detector pixel, the acquired intensity at the detector is attenuated

depending on the incidence angle α and the beam energy E. The correction of the former effect is referred to as geometric incidence angle correction whereas the latter is referred to as physical incidence angle correction. Measurements for different energies $E = \{60kV, 120kV\}$ and various angles $\alpha \in [-40\,\text{deg}, 40\,\text{deg}]$ to account for the maximum field-of-view have shown the separability of those effects.

The geometric incidence angle correction as shown in (6) can be seen as a relative correction from a pixel position at a given detector plane p_d to another plane p_c (e.g. the consensus plane) using the incidence angles α_d and α_c respectively. Specifically, α depends on S, its projection onto the specific plane S_d, S_c and the point on the plane p_d, p_c.

$$\alpha_d = \arctan\left(\frac{||S - S_d||}{||p_d - S_d||}\right) \tag{4}$$

$$\alpha_c = \arctan\left(\frac{||S - S_c||}{||p_c - S_c||}\right) \tag{5}$$

$$I_c = I_d \frac{\cos(\alpha_c)}{\cos(\alpha_d)} \tag{6}$$

The physical incidence angle correction can be seen as the deviation of the measured angle dependency from the cosine function, depending on the used energy E and the incidence angle α. Having specified the measurements in a polynomial of degree four for a low energy $Y_{low}(x) = 1 + c_{2low}x^2 + c_{4low}x^4$ and a high energy $Y_{high}(x) = 1 + c_{2high}x^2 + c_{4high}x^4$, a linear interpolation for arbitrary energies is used for the correction:

$$I_c = \frac{I_d}{Y_{low}(\alpha) + (Y_{high}(\alpha) - Y_{low}(\alpha))\frac{E - E_{low}}{E_{high} - E_{low}}} \tag{7}$$

2.6 Distance Correction

The intensity produced on the detector surface is dependent on the distance of the focal spot to the specific detector pixel. Projecting the intensities from a pixel position at the detector plane p_d onto another plane p_c (e.g. consensus plane) therefore requires a relational correction according to the inverse square law. Using the Euclidean distance $d(p, q) = ||p - q||$ from the focal spot S, the equation can be stated as:

$$I_c = I_d \frac{||p_d - S||^2}{||p_c - S||^2} \tag{8}$$

2.7 Fluence Correction

For the calibration of the fluence correction, multiple projections are taken such that the whole cone has been imaged at least once. A continuous detector movement as described for MLG calibration can be used again for this acquisition.

The acquired images are subsequently pre-processed by bad-pixel correction and MLG correction. The fluence plane is defined perpendicular to v_{cax} at the isocenter. A stitched projection of all pre-processed images onto the fluence plane considering flex calibration, incidence angle correction and distance correction followed by a normalization to the range $[0.0, 1.0]$ creates the fluence map. A fluence map as shown in Figure 3 is created for multiple energies E. For a given position of a pixel at the detector p_d, the most relevant fluence map, i.e. closest matching E, for the given energy is evaluated at a point p_f which is the intersection of a ray between S and p_d with the fluence plane. The evaluated intensity is subsequently corrected for the geometric incidence angle and the distance of the detector plane before using it for the correction, as stated in (9).

$$I_c = \frac{I_d}{I_f \frac{\cos(\alpha_d)}{\cos(\alpha_f)} \frac{||p_f - S||^2}{||p_d - S||^2}} \tag{9}$$

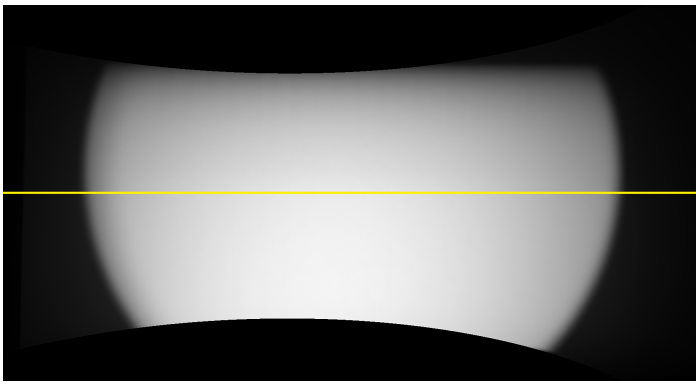


Fig. 3: Fluence map for 120kV with a source cone angle of 50 deg which is used for fluence correction. The anode heel effect is visible at the top of the cone. The yellow line represents the horizontal intensity profile extracted in Figure 4.

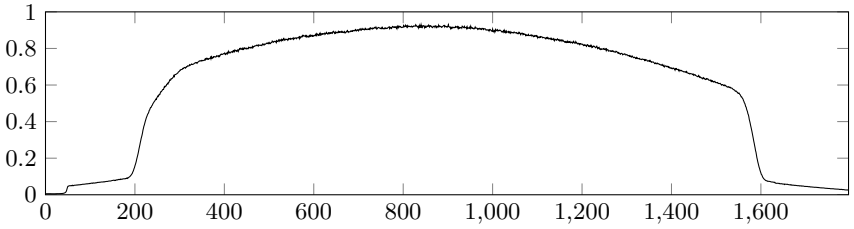
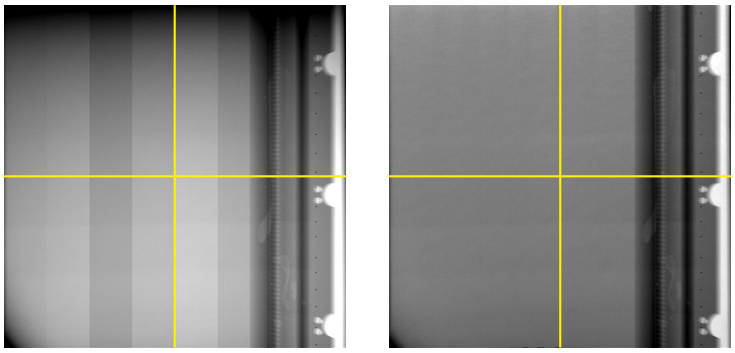


Fig. 4: Horizontal intensity profile of the fluence map for 120kV with a source cone angle of 50 deg.

3 Results

An acquisition of images with the ImagingRing at various positions of the source and detector arm show clearly an effect of the specific geometry to the intensity profile, as can be seen in Figure 5a, Figure 6 and Figure 7. After correcting the images with the proposed pipeline, the bad-pixels and differences in sub-panel intensity levels have been removed as shown in Figure 5b, Figure 6 and Figure 7. Furthermore, the impact of the geometric effects and the heel effect as part of the beam fluence have been minimized. Especially in the homogeneous table region, the flattening of the intensity profiles is obvious. However, even though the vertical intensity profile of the corrected table image shows a straightened curve, there is a remaining slope visible. In the left lower corner of the images, the cone border is visible since the 9DOF-flexmap correction does not account for a flex induced tilt of v_{cax}, effectively leading to an insufficient fluence correction.



(a) original (b) corrected

Fig. 5: A planar x-ray acquisition of the couch carrying the ImagingRing before (a) and after (b) the application of the correction pipeline. The yellow lines represent the horizontal (Figure 6) and vertical (Figure 7) intensity profile positions.

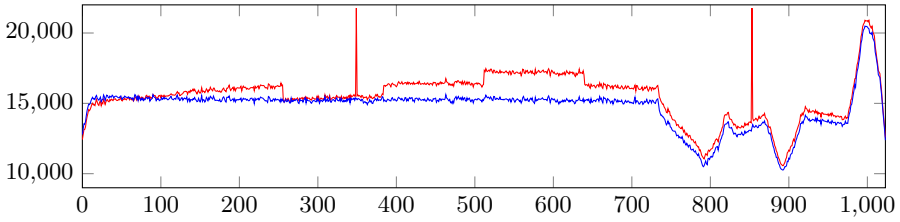


Fig. 6: Comparison of the horizontal beam profiles from the images of Figure 5 before (red) and after (blue) application of the correction pipeline.

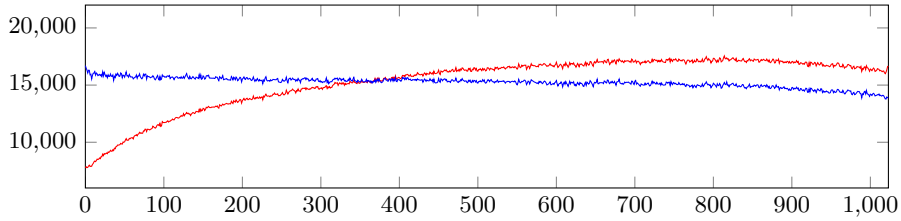


Fig. 7: Comparison of the vertical beam profiles from the images of Figure 5 before (red) and after (blue) application of the correction pipeline.

4 Discussion

As successfully shown in the results section, the presented correction pipeline is able to handle x-ray acquisitions created by arbitrary source and panel arm positions. However, the offset response of the flat-panel detector used for MLG-correction may change over time due to reasons of varying detector temperature and age factors. Since the detector response is non-linear especially in a range close to the offset value, a dynamic replacement of the offset values is considered critical for long-term operation of the device and currently under investigation. For the case of CBCT acquisitions, a basic lag correction may also be added to the correction pipeline. However, the dynamic offset replacement can be seen as a prerequisite for successful ghosting correction. The use case of CBCT reconstruction with correct hounsfield units in addition requires for a scatter correction considering head scatter and patient scatter as key factors. A replacement of the 9DOF-flexmap correction with an 11DOF equivalent, additionally accounting for tilts of v_{cax}, is currently examined. Future developments may also include advanced temperature correction algorithms as have been described in Huber et. al [3] for electronic portal imaging.

5 Acknowledgments

This work was supported by grant $I1383 - B23$ (DEXTER project) of the Austrian Science Fund (FWF).

References

1. Deutschmann, H.: Patient positioning and imaging system. European Patent EP 2687159 A1 (2013)
2. Deutschmann, H., Neuner, M., Steininger, P., Pinzger, M., Buck, M., Sedlmayer, F.: Robotic positioning and imaging. Strahlenther Onkol. 2013 189(2), 185–185 (2013)
3. Huber, S., Mooslechner, M., Mitterlechner, B., Weichenberger, H., Serpa, M., Sedlmayer, F., Deutschmann, H.: Image quality improvements of electronic portal imaging devices by multi-level gain calibration and temperature correction. Physics in Medicine and Biology 58(18), 6429–6446 (2013)

4. Steininger, P., Mehrwald, M., Neuner, M., Kellner, D., Memelink, M., Pinzger, M., Böhler, A., Mitterlechner, B., Keuschnigg, P., Weichenberger, H., Sedlmayer, F., Deutschmann, H., et al.: The open-radart ion (oraion) software suite. In: National Institue of Radiological Sciences & EBG MedAustron GmbH (Hrsg.) editors(s). Joint Symposium on Carbon Ion Radiotherapy. pp. 157–167. Wr. Neustadt (2014)
5. Warmerdam, G., Steininger, P., Neuner, M., Sharp, G., Winey, B.: Influence of imaging source and panel position uncertainties on the accuracy of 2D3d image registration of cranial images. Med. Phys. 39(9), 5547–5556 (2012)

Real-Time Beam Visualization for Monitoring External Beam Radiotherapy: A Feasibility Study

Cesare Jenkins[1,2], Dominik Naczynski[1], Shu-Jung Yu[1], and Lei Xing[1]

[1] Department of Radiation Oncology, Stanford University School of Medicine, Stanford, CA, USA,
[2] Department of Mechanical Engineering, Stanford University, Stanford, CA, USA

Abstract. As external beam radiation therapy has advanced in complexity and precision there have been few new tools developed to monitor the accuracy of treatment delivery. To this end, a system for real-time beam visualization (RT-BV) was developed. The system consists of a flexible scintillating film, common optical detectors and image processing algorithms. The scintillating films were formed by mixing Gd_2O_2S:Tb (GOS) with silicone. The films were placed in the path of therapeutic beams generated by a medical linear accelerator (LINAC). The emitted light was captured by cameras and processed. The RT-BV system was able to achieve sufficient contrast-to-noise at 20 frames/second with normal ambient room lighting. Pixel intensities were found to be linear with dose-rate and the system demonstrated sub-millimeter resolution in identifying collimator movements.

Keywords: Image Guided Radiation Therapy · Quality Assurance · Cameras

1 Introduction

Each year, thousands of radiation therapy (RT) treatments are performed accurately. However, reports indicate that between 0.6 to 4.7 incidents per 100 radiation therapy visits have been reported to occur even in advanced oncology centers operating with modern equipment and trained staff [1,2,3,4]. A significant number of these incidents occur during the delivery of RT and can result in underdosing or overdosing patients, irradiating healthy tissue or, at worst, patient death [5]. As EBRT techniques continue to increase in complexity and dose delivered per fraction, there is a clear and growing need for a monitoring and validation methodology that enables clinicians to have greater confidence that they are delivering the planned dose where and how it was intended [6].

In current practice, on-board MV electronic portal imaging (EPID) is utilized for monitoring the RT beam after it exits the patient [7]. However, due to a lack of soft tissue contrast and multileaf collimator (MLC) blockage of the field of view, the available anatomic references provided by the approach are limited

to bony anatomy or implanted fiducial markers that lie within the beam's eye view making the interpretation of the data non-trivial [8,9]. More recently Jarvis et al. have presented work demonstrating the potential of Cherenkov emission as a means of visualizing therapy in real time [10]. Fahimian et al. also proposed a method of utilizing air scintillation to visualize the radiation beam [11]. While both seek to visualize therapeutic beam delivery, the signals generated are at least three orders of magnitude smaller than typical room lights [12]. The technical complexity involved in utilizing such techniques has limited their implementation. Hence, the ability to visualize the position, shape and intensity of the radiation beam as it passes through the patient in real-time is not yet available in clinical practice.

As a first step toward developing a system capable of realizing this ability, we developed a simple imaging approach utilizing a flexible scintillating film and an inexpensive digital camera to image the radiation beam as it enters or exits the patient's anatomy.

2 Materials and Methods

The real-time beam visualization (RT-BV) system consists of a flexible scintillating film and a digital camera. The film is placed on the patient's skin such that it will emit an optical signal when the beam passes through it. The camera is arranged such that it can image the emitted signal as well as the surrounding patient surface anatomy.

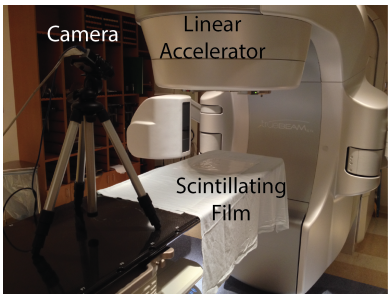


Fig. 1. Photograph of the imaging setup.

The desirable features of a scintillation film to be placed on the patient include ample signal generation, minimal dose attenuation and sufficient flexibility to conform to anatomy. In our study, scintillating films were prepared by mixing GOS powder and a silicone elastomer in a 1:1 mass ratio. The resulting mixture was cast on a level surface at a thickness of 0.8 mm and allowed to cure at room temperature.

The imaging setup, shown in Figure 1, consisted of a 2048 x 2048 CMOS digital camera equipped with a 50 mm lens and a band-pass filter. The camera

was placed 100 cm inferior and 50 cm anterior of iso-center. The halogen lights in the treatment room were set to maximum intensity for all image acquisition. The camera was selected primarily based on its resolution, bandwidth, and cost with secondary consideration given to noise characteristics. All images were acquired with a 42 ms integration time with the gain of the camera set to its minimum value. At this integration time a frame rate of 23 frames per second (fps) is achievable. In general, a set of data for a static field was created by sampling 10 frames over a 2 second interval. Images were analyzed using ImageJ 1.47v (National Institute of Health, Bethesda, MD), Matlab 7.8.0 (The MathWorks Inc. Natick, MA) and OpenCV.

Radiation was delivered using Varian EX and TrueBeam medical linear accelerators (LINAC) (Varian Medical Systems, Palo Alto, CA). Beam energies between 6 and 15 MV were investigated with dose rates ranging between 200 and 600 MU/min. Square fields were delivered while imaging. Bright outliers, caused by x-ray interactions with the imaging sensor were median filtered and a background image was subtracted from each frame. Average pixel values and contrast-to-noise ratios (CNR) where evaluated. See Figure 2 for a graphical overview of the image processing approach.

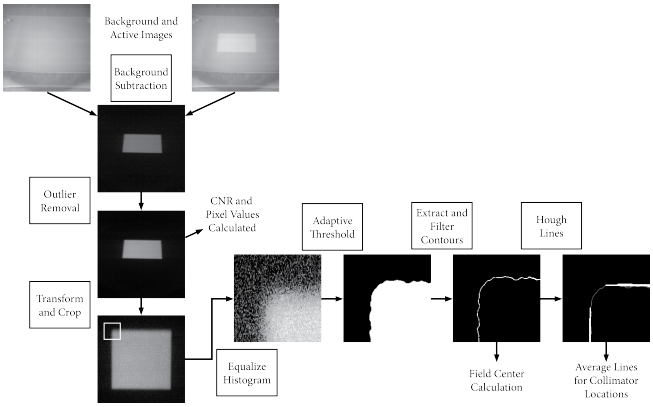


Fig. 2. An overview of the image processing algorithm used to evaluate images and locate field edges. The left column indicates processing that was applied to all images. The small square region indicates the image area magnified at right for demonstrating the localization portion of the algorithm.

In order to verify the resolution and beam shape fidelity of the system, complex MLC shaped fields with gaps ranging from 0.1 cm to 8 cm were delivered and imaged simultaneously with the RT-BV system and a Varian EPID portal imaging system (see Figure 4a). Finally, a full VMAT plan was imaged as it was delivered to an anthropomorphic phantom.

An image processing algorithm was designed to robustly identify and locate the beam within each image in a beam-on dataset. Refer to Figure 2 for a graphical overview of the image-processing algorithm utilized.

A perspective transformation was determined from extracting the corners of a square field in a test image and used to correct for the distortion introduced due to the angled positioning of the camera. The Otsu method was used to identify a threshold value that would best segment the resulting images [13]. A binarization based on this value was applied to the images. Connected components were extracted from the binary images using the algorithm presented by Suzuki et al. [14]. The contours were then filtered by size to identify one that corresponds to the beam profile. The center of the field was calculated from the mass center of the pixels along the identified contour. The outline of the contour was then fit for straight lines using a probabilistic Hough Transform [15]. Lines were grouped according to their angle and location within the image to identify each of the four collimators. An average location for each group was calculated from

$$c_d = \frac{\sum \bar{d} l e^{-|\delta|}}{\sum l e^{-|\delta|}} \tag{1}$$

where c_d is the reported collimator location, $\bar{d}$ is the average pixel coordinate for each line (in the direction of interest), l is the length of the line and δ is the angular deviation of the line from the ideal collimator direction (i.e. 90° for the x-collimators).

3 Results and Discussion

The fabricated films were confirmed to have a thickness of 0.8 mm ± 0.1 mm. They were flexible and molded easily to the contours of an anthropomorphic phantom (Figure 5).

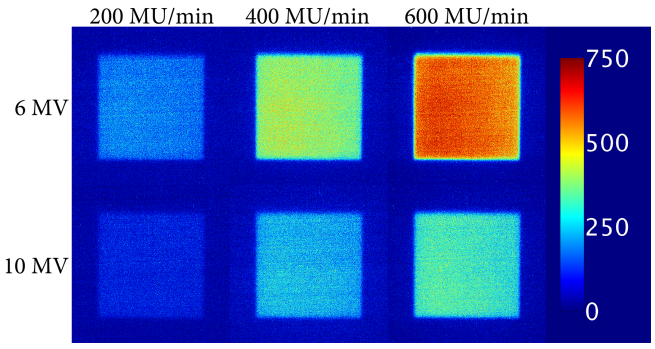


Fig. 3. Images of 10 cm x 10 cm fields at various beam settings. Pixel intensities are reported in arbitrary units.

Processed images of 10 x 10 cm fields for 6 and 10 MV beams with dose rates between 200 and 600 MU/min are shown in Figure 3. Images show clear field boundaries and reasonable uniformity across the field. For beams delivered at a higher dose rate, the penumbra of the beam can also be observed. Pixel intensity values increased with dose rate in a highly linear (R2¿0.97) fashion (data not shown). Increasing beam energy reduces the slope of this relationship. This is consistent with the reduction, at higher energies, in the rate at which dose is deposited at the surface. The average contrast-to-noise (CNR) for sets of 10 images at each beam setting is reported in Table 1. The standard deviation is also presented offering insight into the variation in measurements. These results indicate that the system is capable of monitoring beams across the clinical range in the presence of full room lighting.

Table 1. CNR (SD) for 10 x 10 cm fields at various beam settings.

	Dose rate (MU/min)		
Beam energy (MV)	300	400	600
6	4.0 (0.11)	9.4 (0.38)	18.7 (0.82)
10	1.9 (0.12)	4.9 (0.16)	8.6 (0.32)

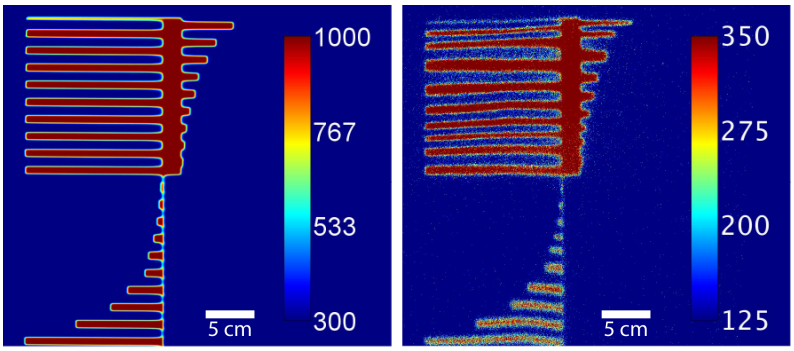


Fig. 4. Images of a test field imaged simultaneously with an EPID portal imaging system (a) and the RT-BV system (b).

Figure 4 shows the result of a simultaneous acquisition with an EPID portal imager and the presented system for a complex test field. All features are discernible, including those created by 1 mm leaf movements near the center of the field. A discernible difference is also observable between features created by leaf

locations that vary by 1 mm. The distortions in the RT-BV system image are due to the fact that the film was not perfectly flat on the couch at the time of imaging. While these distortions are undesirable on a flat surface, they demonstrate the ability of the film to represent the field shape as it appears on the contours of a patients anatomy. Several frames from the imaged VMAT plan are presented in Figure 5. The field, film and phantom are all visible. This test also confirmed that the system was also capable of imaging beam exit profile as well as entry. The algorithm was generally able to properly locate the collimators well within the 1 mm machine tolerance (see Figure 6). The average deviation between machine and estimated locations was 0.5 mm. The maximum deviation was 2mm.

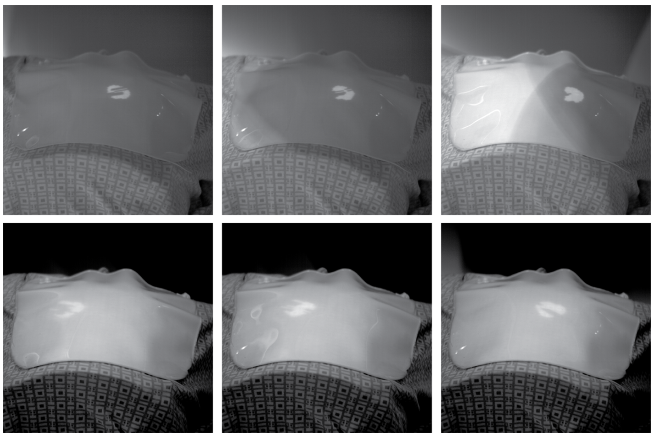


Fig. 5. Sequential images of a VMAT plan being delivered to an anthropomorphic phantom. The top three frames show an entry beam profile as the gantry moves from approximately 0° to 30°. The bottom three frames show a beam exit profile as the gantry moves from approximately 70° to 90°. Full video is available at `http://youtu.be/X6n-aDPBDGw`.

Imaging approaches based on air scintillation [11] or Cherenkov radiation [10] also offer the ability to visualize the beam. These techniques enjoy the advantage of not perturbing the treatment beam but generally require much more advanced imaging setups and image processing. The data presented here was collected using cameras costing approximately $1,500 as compared with an minimum cost of $10,000 for the cameras needed to perform air scintillation or Cherenkov imaging. EPID based techniques measure the beam aperture exiting the patient, but interpreting the data relative to patient anatomy is non-trivial. The proposed system enables high-resolution images of the beam in a context that is immediately interpretable. It also offers the potential for gathering beam dosimetry data at beam entry based on the intensity of emitted optical signal from the film. A combined use of EPID and the RT-BV system may be valuable

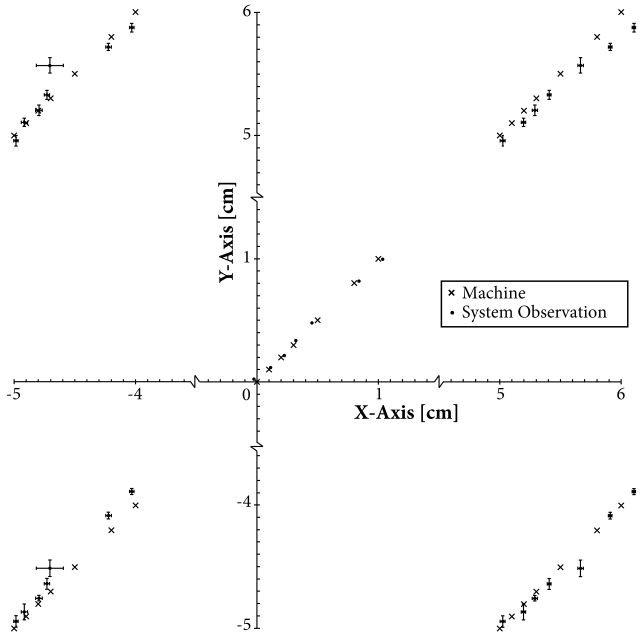


Fig. 6. The results of the collimator location algorithm.

and allow us to harness the strengths of both systems to provide a complete view of therapy.

In summary, the RT-BV system has demonstrated the capability of obtaining high CNR images of therapeutic beams and treatment context in real time. Processing of these images resulted in accurate, sub-millimeter identification of collimator locations. It is envisioned that the proposed system could be used to provide a real-time view/recording of the patient treatment delivery process. While further developments leading to a quantitative, automated system for assessing treatment validity would be ideal and are currently being explored, the availability of qualitative information on beam intensity, shape and location in the context of the patient's anatomical contours offers a valuable, high level view of treatment.

References

1. Morganti, A.G., Deodato, F., Zizzari, S., Cilla, S., Digesu', C., Macchia G., Panuni, S., Gaetano, A.D., Piermattei, A., Cellini, N., Valentini, V.: Complexity index (COMIX) and not type of treatment predicts undetected errors in radiotherapy planning and delivery. Radiother. Oncol. 89, 320-329 (2008)
2. Yeung, T.K., Bortolotto, K., Cosby, S., Hoar, M., Lederer, E.: Quality assurance in radiotherapy: evaluation of errors and incidents recorded over a 10 year period. Radiother. Oncol. 74, 283-291 (2005)

3. Huang, G., Medlam, G., Lee, J., Billingsley, S., Bissonnette, J.P., Ringash, J., Kane, G., Hodgson, D.C.: Error in the delivery of radiation therapy: results of a quality assurance review. Int. J. Radiat. Oncol. Biol. Phys. 61, 1590-1595 (2005)
4. Bissonnette, J.P., Medlam, G.: Trend analysis of radiation therapy incidents over seven years. Radiother. Oncol. 96, 139-144 (2010)
5. Shafiq, J., Barton, M., Noble, D., Lemer, C., Donaldson, L.J.: An international review of patient safety measures in radiotherapy practice. Radiother. Oncol. 92, 15-21 (2009)
6. Yorke, E., Gelblum, D., Ford, E.: Patient safety in external beam radiation therapy. AJR Am. J. Roentgenol. 196, 768-772 (2011)
7. Wiersma, R.D., Mao, W., Xing, L.: Combined kV and MV imaging for real-time tracking of implanted fiducial markers. Med. Phys. 35, 1191-1198 (2008)
8. Timmerman, R.D., Xing, L.: Image-guided and Adaptive Radiation Therapy. Lippincott Williams & Wilkins (2009)
9. Berbeco, R.I., Hacker, F., Ionascu, D., Mamon, H.J.: Clinical Feasibility of Using an EPID in cine Mode for Image-Guided Verification of Stereotactic Body Radiotherapy. Int. J. Radiat. Oncol. Biol. Phys. 69, 258-266 (2007)
10. Jarvis, L.A., Zhang, R., Gladstone, D.J., Jiang, S., Hitchcock, W., Friedman, O.D., Glaser, A.K., Jermyn, M., Pogue, B.W.: Cherenkov video imaging allows for the first visualization of radiation therapy in real time. Int. J. Radiat. Oncol. Biol. Phys. 89, 615-622 (2014)
11. Fahimian, B., Ceballos, A., Trkcan, S., Kapp, D.S., Pratx, G.: Seeing the invisible: Direct visualization of therapeutic radiation beams using air scintillation. Med. Phys. 41, 010702 (2014)
12. Glaser, A.K., Zhang, R., Davis, S.C., Gladstone, D.J., Pogue, B.W.: Time-gated Cherenkov emission spectroscopy from linear accelerator irradiation of tissue phantoms. Opt. Lett. 37, 1193-1195 (2012)
13. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. 9, 62-66 (1979)
14. Suzuki, S., be K.: Topological structural analysis of digitized binary images by border following. Computer Vision, Graphics, and Image Processing 30, 32-46 (1985)
15. Matas, J., Galambos, C., Kittler, J.: Robust Detection of Lines Using the Progressive Probabilistic Hough Transform. Comput. Vis. Image Und. 78, 119-137 (2000)

Fast processing of CBCT to improve delivered dose assessment

Kiran Joshi[1] and Tom Marchant[1]

Christie Medical Physics and Engineering, The Christie NHS Foundation Trust, Manchester, M20 4BX, UK kiran.joshi@physics.cr.man.ac.uk

Abstract. In a previous work we developed an algorithm to improve the Hounsfield Unit accuracy of CBCT images based on prior information from a CT scan. However, the processing time required to run the algorithm may be a barrier to clinical implementation. Here we describe work to speed up two key processing steps: 3D binary morphological operations and image interpolation.

Efficient binary morphological operators have been implemented in three dimensions using `C++`, extending the open source `leptonica` image processing library. Processing time comparisons have been made to implementations of three dimensional binary morphology available in ITK, MATLAB and IDL. The efficient implementations presented in this report have been found to require processing times up to three orders of magnitude shorter than the alternatives.

Image downsampling has also been investigated as a method to enable faster processing. Downsampling images by a factor of 2.0 (4.0) can produce a speedup of 1.8x (3.4x) in a processing step that interpolates into masked regions of an image. Preliminary studies of the effect of downsampling on the quality of final processed images suggest that downsampling by a factor of 2.0 produces a negligible decrease in final processed image quality. When used together these two developments can allow significantly faster processing of CBCT images.

1 Introduction

In recent years Cone Beam CT (CBCT) imaging systems have been integrated in the majority of linear accelerators used for radiation therapy. The kilovoltage (kV) x-ray source results in images with better soft tissue detail than those produce by megavoltage (MV) portal imaging or cone beam imaging systems. The kV-CBCT images are important for ensuring that patients are set up accurately prior to treatment. Since the images are acquired multiple times over the course of a treatment they also allow the possibility for adaptive radiotherapy. Delivered dose could be computed using the CBCT images, to assess whether the planned treatment remained optimal as tumour and patient reacted. It is often necessary, however, to apply image processing techniques to the CBCT image volumes, to restore accurate calibration before use in radiation therapy dose assessment. However given their large size – up to tens of millions of voxels per

image volume – even conceptually simple image processing operations such as binary erosion and dilation can be slow. Slightly more complex operations such as smoothing with a Gaussian kernel or interpolating into masked regions of an image can increase processing time considerably. The aim of the work presented here was (i) to produce an efficient implementation of three dimensional binary morphological operators and (ii) investigate the use of image downsampling, to allow CBCT image volumes to be processed as quickly as possible.

2 CBCT and Adaptive Radiotherapy

If significant changes in patient anatomy are observed in the CBCT images, current clinical practice often requires the acquisition of an additional fan-beam CT image. This enables the patient's treatment to be re-planned to take into account the altered anatomy, but is an expensive and time-consuming process and results in the patient receiving a larger dose of radiation.

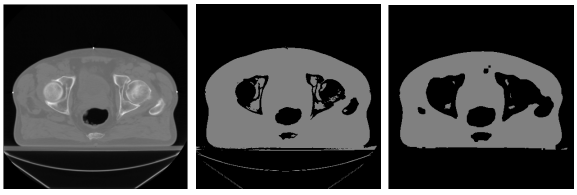
A preferable solution would be to re-plan the treatment directly using the CBCT image. However, the wide cone-beam of x-rays results in larger amounts of scatter than in conventional fan-beam CT scans, and produces artefacts that can make identification of anatomical structures difficult. The CBCT image voxels must also be corrected so that their values accurately represent x-ray attenuation. Dose calculations based on uncorrected CBCT images are prone to error, and can result in large dose inaccuracies when compared to doses calculated using fan-beam CT scans.

Several techniques have been proposed that aim to correct the CBCT images. One such correction procedure, proposed in [1], uses the information contained in the CT scan acquired prior to treatment to correct the overall normalization of the voxel values in the CBCT image, producing an image calibrated in Hounsfield units. The uniformity of the voxel values is also greatly improved by the correction procedure[1]. In order to get maximum value out of the information contained in the CT images and to avoid degrading the quality of the CBCT images, the correction algorithm is applied to the large three-dimensional (3D) image volumes wherever possible. The resulting high quality CBCT image aides clinicians with tissue visualisation, and allows for the assessment of the delivered dose. In cases where large anatomical changes occur between CT and CBCT scans, this procedure is expected to be more robust than techniques such as deformable image registration. It is important that such a correction procedure can be performed quickly. If it is to have a large impact on clinical workflow the aim should be to correct the image volumes in as close to real-time as possible.

At various stages of the correction procedure 3D binary masks are created to select anatomical regions of the image (i.e. bone, soft tissue, gas). Fig. 1(a) shows an example of one slice of the image produced in a CT scan of a male pelvis. In Fig. 1(b) a threshold has been applied to separate the soft tissue (grey areas of

[1] An artefact present in CBCT images results in the voxels representing, for example, soft tissue such as muscle or fat having different values depending on their location in the image.

the mask) from the regions of bone and gas (black areas). Column (c) shows the binary mask after it has been eroded in three dimensions using a structuring element of size 5x5x3 voxels. Binary erosion can be performed to prevent artefacts arising at the boundaries between the soft tissue and bone or gas when performing subsequent image processing techniques. The implementations of three



(a) (b) (c)

Fig. 1: (a): One slice of the image volume produced in a CT scan of a male pelvis. (b): A binary mask created after applying a threshold to separate soft tissue. (c): The result of eroding the binary mask in (b) using a 3D structuring element of size 5x5x3.

dimensional binary morphology currently available in the Insight Toolkit (ITK) have been found to require processing times upwards of 1s (cf. Sect. 3.2). Given that binary morphological operations are performed many times over the course of the image correction procedure, the total time spent on binary morphology can become several seconds. Therefore, given the nature of the algorithm and the aim that images be processed in as little time as possible, work was undertaken on producing a more efficient implementation of three-dimensional binary morphology and is discussed in Sect. 3.

In a subsequent step of the correction algorithm, the masked regions of the image are interpolated into based on data from surrounding regions (i.e. image inpainting). This process is computationally intensive and must be performed for multiple masks. The interpolation of one image volume can take several seconds, depending on the size of the image. In an attempt to reduce the time taken to perform the interpolation downsampling was used. The amount of downsampling and the effect on the final corrected images is described in Sect. 4.

3 Efficient 3D binary morphology

3.1 Implementation

The three-dimensional binary morphological operations are implemented as an extension to the open-source image processing library `leptonica` [2], which provides efficient implementations of binary morphology in two dimensions. Full details of the two-dimensional implementations can be found in [3], though a

brief summary will be repeated here to aid in the discussion of the extension to three dimensions.

The basic binary morphological operations of dilation and erosion can be expressed in several ways. A standard definition given here lends itself well to demonstrating the low-level implementation. Let A represent an N-dimensional binary image, and B represent an N-dimensional binary structuring element (`Sel`). The dilation, $\oplus$, and erosion, $\ominus$, of an image A by `Sel` B are given, respectively, by

$$A \oplus B = \bigcup_{\boldsymbol{b} \in B} A_{\boldsymbol{b}} \tag{1}$$

$$A \ominus B = \bigcap_{\boldsymbol{b} \in \bar{B}} A_{\boldsymbol{b}} \tag{2}$$

where $A_{\boldsymbol{b}}$ is the image A after translating it by pixel vector $\boldsymbol{b}$. If the `Sel` is considered as a set of pixels with locations $b = (i, j)$ relative to an origin, typically at its centre, then the pixel vector $\boldsymbol{b}$ is the vector pointing from the origin to pixel b. $\bar{B}$ represents the inversion of B about its origin, i.e. $\bar{B} = \{-b \mid b \in B\}$.

In order to increase CPU and memory efficiency the binary images are packed such that each horizontal 32-pixel line segment is represented by a single 32-bit integer. I.e., each bit of the integer represents a single on/off pixel of the binary image. In this way the union and intersection operations can be implemented using bitwise OR and AND operators, and the required image translations can be implemented using the bit-shift operators available in `C++`. After packing the binary image each application of the bit-shift and bitwise logical operators can perform the necessary translate and logical OR/AND operations on 32 pixels at once.

Within the `leptonica` framework a packed two-dimensional image is represented by a `Pix` object. In the extension to three dimensions, each slice of the image is packed individually and the image volume is stored as an array of `Pix`. The object representing the structuring element must also be extended to take into account an additional dimension. Shifts in the z-direction are performed by selecting the corresponding slices of the volume from the array of `Pix`. The 3D morphological operations are then performed by looping over the required shifts defined by the `Sel`, selecting the necessary slices from the array of `Pix`, and taking the logical OR/AND of all corresponding pairs of 32-bit integers.

Boundary conditions are handled by padding the input image by an appropriate amount in each dimension prior to the application of the morphological operations. This padding must be performed by the user and is not implemented internally in the algorithm presented here[2].

The morphological operations presented here have been implemented in such a way that arbitrary-shaped structuring elements can be used. When the `Sel` is separable the morphological operations can be performed in the z-direction in-

[2] Padding must also be performed by the user when using the ITK, MATLAB and IDL binary morphology functions.

dependently, before applying two-dimensional morphological operations to each slice of the volume.

3.2 Performance

The performance of the algorithm implemented here, referred to as "Fast" in the following, is compared to the implementations of three dimensional binary morphology available in ITK[3] (itkBinaryDilateImageFilter), IDL[4] (DILATE) and MATLAB[5] (imdilate). A binary mask is created from a CT scan of a male pelvis and dilated using separable, cube-shaped structuring elements with a range of sizes. The CPU time taken to perform the dilation is isolated from that taken to create the test image volume. For a given structuring element size the test is repeated multiple times to obtain an average dilation time. All tests are performed using a single core of a 3.33 GHz Intel Core i7 CPU in a machine with 16 GB of RAM.

Figure 2 shows the average times taken to dilate the image volume using cube-shaped structuring elements of various sizes. CPU times taken to perform a binary erosion show the same trends as those shown in Figure 2.

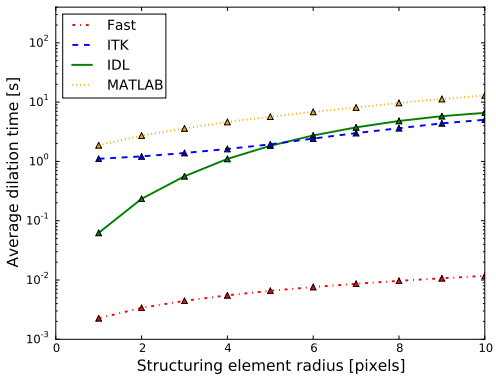


Fig. 2: Average CPU times spent performing binary dilation on a 512x512x31 voxel binary mask created from a CT scan, using structuring elements of various sizes. The structuring element radius is the distance, in voxels, from the centre of the structuring element to the outer faces. A `Sel` of radius 1 is therefore a 3x3x3 cube. The fast implementation presented in this report (red, dot-dashed) is compared to the implementations available in ITK (blue, dashed), IDL (green, solid) and MATLAB (orange, dotted). The fast implementation is typically two to three orders of magnitude faster than the alternatives.

[3] ITK v.4.7.2, The Insight Software Consortium
[4] IDL v7.1, Exelis Visual Information Studio, Boulder, Colorado, USA
[5] MATLAB v6.5 Release 13, The Mathworks, Inc., Natick, Massachusetts, USA

Despite the ITK implementation making use of the techniques presented in [4] it is only able to perform binary morphological operations in times longer than 1s.

The IDL implementation runs in times an order of magnitude shorter than ITK when using very small structuring elements. However the processing time increases rapidly as the `Sel` radius increases.

Binary dilation implemented in MATLAB has approximately the same performance as ITK, with average processing times varying between 1.8s and 14.7s.

The fast binary morphological operators implemented in this report typically run several hundreds of times faster than the functions available in ITK, IDL and MATLAB. When dilating the CT mask using very small `Sel`s the processing time required is just 0.002s, compared to 0.06s with IDL and 1.1s with ITK.

The resulting dilated masks produced by each of the algorithms were compared by subtraction, and were found to be identical.

4 Image downsampling

Interpolation into masked regions of the image volumes is performed on each slice of the image volume separately, often referred to as "2.5D", and takes place prior to the creation of a correction map that is subsequently applied to the CBCT image. The temporary images used to create the correction map are downsampled using the ITK ResampleImageFilter, with linear interpolation used to compute pixel values at non-grid locations of the full-sized images. The original CBCT and CT images themselves are not downsampled at any point.

Since the interpolation is performed in 2.5D, the downsampling is only performed in the x- and y- directions; the number of slices in the volume is kept fixed. For example, a downsample factor of 2.0 implies that an image slice of size 410x410 pixels becomes 205x205 pixels. The computation time taken to perform the interpolation after different amounts of downsampling is shown in Table 1. Downsampling the images by a factor of 2.0 results in a speed increase of 1.8x,

Downsample factor	Mean time/image [s]
1.0	1.29
2.0	0.70
4.0	0.38

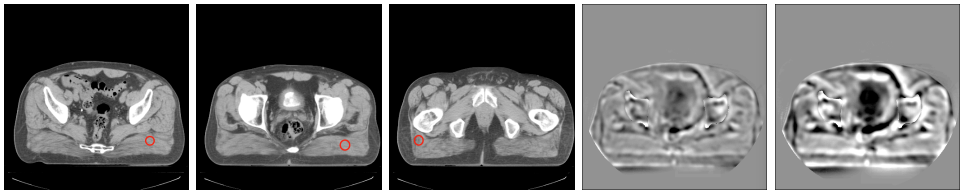
Table 1: Variation of 2.5D interpolation speed with downsampling factor.

and downsampling by a factor of 4.0 results in a 3.4x speedup.

The effect of the downsampling, and subsequent upsampling, on the final corrected CBCT images was estimated using two methods: by assessing the uniformity of the various CBCT images corrected after downsampling by different

amounts, and by subtracting the CBCT images corrected after downsampling from the CBCT image corrected without downsampling.

The image uniformity was assessed by defining small regions of muscle tissue which were expected to have similar densities. Examples of some of the regions are shown in Fig. 3(a)-(c). The same set of regions were used when analysing all images. The mean pixel value in each of the regions was calculated and the average and standard deviation of these means was used to quantify the uniformity. As well as the corrected CBCT images, the uniformity was calculated using a CT image of the same patient. A comparison of the uniformity in each of the images is show in Table 2.



(a) (b) (c) (d) (e)

Fig. 3: (a)-(c): Small regions containing similar types of muscle tissue, in different parts of the image, were used to estimate image uniformity. (d), (e) CBCT images corrected using downsampling factors of 2.0 and 4.0, respectively subtracted from that corrected without downsampling.

Image	Average $\pm$ std. dev.
CT	1011.8 ± 16.3
Uncorrected CBCT	1044.2 ± 47.8
Corrected CBCT	1010.0 ± 16.6
Corrected CBCT, Downsample 2.0	1008.6 ± 17.7
Corrected CBCT, Downsample 4.0	1006.9 ± 18.5

Table 2: Comparisons of image uniformity estimates.

Little difference is observed between the uniformities of the corrected CBCT images. The average of the mean pixel values in the muscle tissue regions remains approximately constant, while the standard deviation (non-uniformity) increases slightly as the amount of downsampling increases.

The result of subtracting the CBCT image corrected using a downsampling factor of 2.0 (4.0) from the CBCT image corrected without downsampling is

shown in Fig. 3d (e). When downsampling by a factor of 2.0 the differences are typically small. Differences in pixel values of around 10 are observed across the majority of the image. In regions near the edges of bones, and around the outer edge of the patient, the differences are slightly larger with maximum differences of approximately 70 appearing in isolated areas. When downsampling by a factor of 4.0 the differences are increased. Across the majority of the image differences in pixel value are approximately 15-20, with maximum differences of around 150.

Whether differences of this size can be tolerated is unclear. The ultimate test of the effect of downsampling will be to use the various corrected CBCT images to calculate the dose delivered by a radiotherapy plan.

5 Summary

Two methods have been investigated to speed up processing of CBCT images for delivered dose assessment. The efficient implementation of 3D binary morphological operators has been presented and their performance compared with the implementations available in ITK, IDL and MATLAB. The algorithms are implemented as an extension to the `leptonica` image processing library and have been found to run up to three orders of magnitude more quickly than commonly available alternatives. It is likely that these fast 3D operators could be of use in a wide range of applications.

Image downsampling has also been investigated as a method for enabling faster image processing. Downsampling by a factor of 2.0 or 4.0 can significantly reduce overall processing times. The effect on the final corrected images has been estimated by directly comparing corrected images and assessing image uniformity. Preliminary results suggest that downsampling by a factor of 2.0 has a negligible impact on the quality of the corrected image, though further work will be performed by using corrected CBCT images to calculate dose delivered by a radiotherapy plan.

The combination of these techniques produces an overall speedup of 3x and allows CBCT images to be corrected in under 10s.

This research is funded by UK MRC grant MR/L023059/1.

References

[1] T E Marchant et al. “Shading correction algorithm for improvement of cone-beam CT images in radiotherapy”. In: *Physics in Medicine and Biology* 53.20 (2008), p. 5719.

[2] D S Bloomberg. URL: `http://www.leptonica.com`.

[3] D S Bloomberg. “Implementation Efficiency of Binary Morphology”. In: *International Symposium for Mathematical Morphology VI* (2002). URL: `http://www.leptonica.com/papers/binmorph.pdf`.

[4] N. Nikopoulos and I. Pitas. “A fast implementation of 3-D binary morphological transformations”. In: *Image Processing, IEEE Transactions on,* 9.2 (Feb. 2000), pp. 283–286. ISSN: 1057-7149. DOI: `10.1109/83.821743`.

Author Index

Z