35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Automatic Prostate Segmentation in Cone-Beam Computed Tomography Images using Rigid Registration*

Christine Boydev*†, David Pasquier‡§, Foued Derraz*¶, Laurent Peyrodie‖,
Abdelmalik Taleb-Ahmed¶ and Jean-Philippe Thiran†**
*Laboratoire d'Automatique, de Mécanique et d'Informatique Industrielles et Humaines (LAMIH),
Université de Valenciennes et du Hainaut-Cambrésis, France
†Signal Processing Laboratory (LTS5), École Polytechnique Fédérale de Lausanne, Switzerland
‡Centre de Radiothérapie et d'Oncologie Galilée de Lille, France
§Academic Department of Radiation Oncology, Centre Oscar Lambret, Lille, France
¶Unité de Traitements de Signaux Biomédicaux (UTSB), Faculté Libre de Médecine, Lille, France
‖Unité de Traitements de Signaux Biomédicaux (UTSB), Hautes Études d'Ingénieurs, Lille, France
**Department of Radiology, University Hospital Center (CHUV) and University of Lausanne (UNIL), Switzerland
Email contact: christine.boydev@epfl.ch

***Abstract*— We propose to evaluate automatic three-dimensional gray-value rigid registration (RR) methods for prostate localization on cone-beam computed tomography (CBCT) scans. In total, 103 CBCT scans of 9 prostate patients have been analyzed. Each one was registered to the planning CT scan using different methods: (a) global RR, (b) pelvis bone structure RR, (c) bone RR refined by local soft-tissue RR using the CT clinical target volume (CTV) expanded with a 1, 3, 5, 8, 10, 12, 15 or 20-mm margin. To evaluate results, a radiation oncologist was asked to manually delineate the CTV on the CBCT scans. The Dice coefficients between each automatic CBCT segmentation - derived from the transformation of the manual CT segmentation - and the manual CBCT segmentation were calculated. Global or bone CT/CBCT RR has been shown to yield insufficient results in average. Local RR with an 8-mm margin around the CTV after bone RR was found to be the best candidate for systematically significantly improving prostate localization.**

***Index Terms*— Rigid registration, segmentation, cone-beam computed tomography, image-guided radiotherapy, prostate cancer.**

I. Introduction

HIGHLY conformal radiation therapy has provided an opportunity for dose escalation but requires in return greater precision in treatment set-up and delivery [1]. However the prostate is known to be a moving and deformable gland, and its motion nature is well-documented [2]. This limits the effectiveness of skin marks in patient set-up. Daily prostate image guidance helps to correct for inter-fractional set-up errors and to optimize tumor coverage and organs at risk avoidance. Image guidance produces images immediately before each treatment fraction, with the patient in the treatment position. These images are compared, either manually or automatically, with initial images acquired prior to treatment planning at an earlier stage. The spatial differences between the planned and the daily tumor positions are referred to as patient set-up errors.

*This work was financially supported by ELEKTA SAS, Boulogne Billancourt, France

Techniques for daily prostate localization include kilo-voltage cone-beam computed tomography (CBCT) systems attached to the treatment unit. The CBCT image consists of a three-dimensional volumetric image resulting from the reconstruction of mutiple kV projections acquired at different angles [3]. Prostate localization on CBCT scans is challenging due to the relatively poor image quality. It is due to low soft-tissue contrast (more scatter and hence more noise than in a CT image at the same level of radiation dose [4]) and streaking artifacts caused by possible moving gas pockets in the rectum during CBCT image acquisition (i.e. intra-fraction motion) [5]. Moreover Deurloo *et al.* reported that the deformation of prostate and seminal vesicles during the course of radiotherapy is small compared to organ motion and therefore in image-guided radiotherapy (IGRT) of prostate cancer, in first order, only set-up error and organ motion need to be corrected for, whereas prostate and seminal vesicles deformation can be considered as a second-order effect [2]. They concluded that this finding drastically simplifies the task of on-line image guidance for the prostate.

Court *et al.* have developed an automatic monomodal CT/in-room-CT rigid registration of the prostate for IGRT [6]. The present study appears as a preliminary study where we propose to evaluate automatic segmentation of the prostate on CBCT scans using different methods of (multimodal) CT/CBCT rigid registration (RR) for IGRT, namely global RR, bone RR, and local soft-tissue RR focused on the prostate region expanded with several margins. To our knowledge, there are no published data about the effect of the margin size on local CT/CBCT RR quality. We also discuss and quantify the impact of air in the rectum on RR quality.

II. Material and Methods

The process of automatic CBCT prostate segmentation operates by performing RR between the (daily) CBCT and the (initial) planning CT scans. The resulting displacement is then applied to the contours manually delineated on the

978-1-4577-0216-7/13/$26.00 ©2013 IEEE

planning CT scan to create the new CBCT contours. Three types of RR methods are tested :

(a) global RR,
(b) RR of the pelvis bone structures of CT and CBCT images,
(c) bone RR followed by local soft-tissue RR. The latter is conducted using a region of interest (ROI) defined by the CT clinical target volume (CTV) expanded with several margins, namely 1, 3, 5, 8, 10, 12, 15 or 20 mm. The CTV represents the whole prostate gland (i.e. the volume delineated by the physician).

A. Data collection

In total, 103 images of 9 prostate cancer patients have been analyzed. All patients were asked to follow a dietary protocol in order to have a full bladder and an empty air-free rectum at the moment of the treatment. The planning CT data were acquired using a General Electrics Light Speed scanner. The treatment system was an ELEKTA Synergy linear accelerator (LINAC) equipped with CBCT imaging (named x-ray volume imaging (XVI)) whose axis is perpendicular to the treatment axis. During CT (CBCT) acquisition, the peak-voltage, X-ray tube current and exposure time were 120 kVp (120 kVp), 300 mA (40 mA or 64 mA) and 1000 ms (40 ms), respectively. Slice thickness was 3 mm and 1 mm for CT and CBCT scans, respectively. Combined with the beam geometry (fan for CT and cone for CBCT), these parameters account for CBCT's lower image quality compared to CT. The number of slices ranged from 89 to 132 slices in each CT scan, and was 168 in each CBCT scan. Each CT (CBCT) slice had 512 x 512 (410 x 410) pixels, with a typical in-plane resolution of 0.98 (1.00) mm.

For clinical requirements, the prostate CTV was manually delineated on each planning CT scan by a radiation oncologist using clinical software (this step is always required in clinical practice for treatment planning). For the purposes of this study, i.e. for validation, the same radiation oncologist manually delineated the CTV on each CBCT scan, using ITK-SNAP, a free cross-platform open-source software application for manual segmentations [7]. Because it was an extremely tedious and time-consuming task, only every three slices was delineated on each CBCT scan while all slices were delineated on the CT scans. The missing 1 mm-thick CBCT contours were estimated and the manually delineated 3 mm-thick CT contours were upsampled to a slice thickness of 1 mm so that both segmented volumes had the same z-resolution (necessary for the quantitative validation). To do so, we implemented the shape-based interpolation method proposed by Raya and Udupa [8] (code available at http://hdl.handle.net/10380/3390).

The CT contours were used both in the definition of the soft-tissue mask associated to the CT gray-value image input to the local RR and as ground truth in the quantitative RR validation. The CBCT contours were only used as ground truth in quantitative RR validation.

B. Registration algorithm

A typical image registration framework has four basic components: metric, optimizer, transform and interpolator. The metric measures image similarity between the fixed (CT) and the moving (CBCT) images. It is defined as a statistical measure between the gray-value distribution of the two images. Because the context is multimodal image registration, normalized-cross correlation was chosen as an advanced cost function. When masks are associated to the images to be registered, only pixels that belong to the intersection of the masks are considered for the computation of the metric. The similarity between images was intensity-based, allowing the registration method to be fully automatic. Optimization was performed with the regular step gradient method. Transformations were rigid and hence they had only six degrees of freedom (translations and rotations). Linear interpolation was used in all our experiments. Three resolution levels were used.

In prior to CT/CBCT RR, there were some preprocessing steps to carry out (Fig. 1 for steps 1-4):

1) Convert the planning CT Dicom image set and the CBCT DicomRT image set into 3D volumetric meta-image files to get rid off all patient-related information. These are the gray-value images that are passed as inputs to the global RR and the local soft-tissue RR. All further processings use meta-images.
2) Offset the planning CT image so that its isocenter coincides with the treatment isocenter (i.e. the CBCT system's isocenter). To do so, the information about the planning CT isocenter coordinates are retrieved from the Dicom RTStruct file and compared with the LINAC isocenter coordinates (which are set to 0,0,0). This is the starting point for RR.
3) Create the 3D binary masks that identify the patient's body on both gray-value images. These were the masks by default (i.e. if no other masks were used instead) associated to the gray-value images in the RR process. They allow ignoring pixels "outside" the patient (e.g. treatment table or artifacts) and may adversely influence the registration process.
4) *(This step concerns the bone RR only.)* Threshold the CT and CBCT gray-value images to show exclusively the pelvis bone structure. The threshold level used to extract the bone anatomy was 150 HU (HU stands for Hounsfield unit, an XRay attenuation unit used in CT scan interpretation and characterizing the relative density of a substance) for the CT gray-value image, and -140 a.u. for the CBCT gray-value image (the CBCT system is usually not calibrated in its normal clinical use, and hence the units are arbitrary). These thresholded images are registered, each one being associated with its corresponding patient body mask.
5) *(This step concerns the local soft-tissue RR only.)* Expand the physician-drawn CT CTV by the tested margin (1, 3, 5, 8, 10, 12, 15 or 20 mm). This creates the 3D region-of-interest (ROI) masks to be associated to the

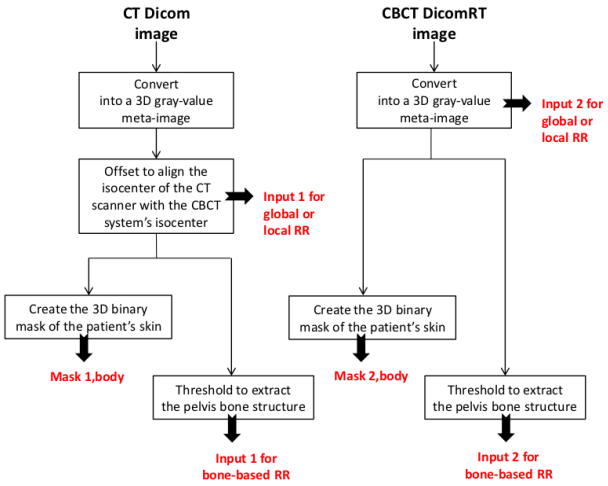

Fig. 1: *Diagram of the data pre-processing carried out in this study*

CT gray-value image in the local soft-tissue RR process, replacing the CT patient body mask.

For this study, all the data processing and visualization were performed on a Linux computer with distribution openSUSE 11.4 x86_64, with an Intel Dual Core i5-560M 2.66 GHz processor, 3MB L2 cache, 4 threads, and 8GB RAM.

For our implementation, the following open-source software, based on C++, was used:

- the Insight Toolkit ITK [9], freely available at www.itk.org,
- the ITK-based Command Line Image Toolkit clitk, freely available at http://www.creatis.insa-lyon.fr/rio/clitk.

C. Validation

To evaluate the RR results, we calculated the Dice similarity coefficient between the automatic CBCT segmentation and the manual CBCT segmentation (referred to as ground truth). Dice's similarity coefficient between a volume, A, and a volume, B, is defined as follows [10] :

$$\textit{Dice's similarity coefficient} = \frac{2 \cdot (A \cap B)}{A + B} \qquad (1)$$

Ideally, when the two volumes perfectly overlap, the Dice coefficient equals 1. A null Dice coefficient corresponds to two disjoint volumes.

D. Statistical analysis

Differences in the Dice results across the multiple RR methods were tested for significance using the inferential non-parametric Friedman statistical test (with α set to 0.05), a version of the parametric repeated-measures ANOVA. The null hypothesis for the Friedman test is that there are no differences between the RR methods, i.e. that the distributions are the same across the multiple RR methods. If the calculated probability is low ($p < 0.05$), the null hypothesis is rejected and it can be concluded that at least two of the RR methods are significantly different from each other. The Wilcoxon-Nemenyi-McDonald-Thompson post-hoc test was conducted to decide which methods are significantly different from each other [11, page 295]. A p-value < 0.05 (< 0.01) was considered (highly) significant. Software R, version 2.12, was used for all statistical analysis [12].

III. Results

The Dice histograms for all data per RR method are illustrated in Fig. 2.

The statistical analysis showed that there was a highly significant difference between the following RR methods: (c)5mm vs (a) ($p = 2.0\,10^{-5}$), (c)5mm vs (b) ($p = 6.6\,10^{-8}$), (c)5mm vs (c)1mm ($p = 1.1\,10^{-6}$), (c)8mm vs (a) ($p = 2.0\,10^{-4}$), (c)8mm vs (b) ($p = 4.0\,10^{-7}$), (c)8mm vs (c)1mm ($p = 2.8\,10^{-5}$), and (c)10mm vs (b) ($p = 6.1\,10^{-3}$). All RR methods were found to yield DICE results significantly different from those obtained without registration. In increasing order, the medians were found to be 0.785, 0.786, 0.787, 0.796, 0.796, 0.799, 0.806, 0.806, 0.819, 0.823 for methods (b), (a), (c)1mm, (c)20mm, (c)12mm, (c)15mm, (c)3mm, (c)10mm, (c)8mm, (c)5mm, respectively. The median without registration was 0.742. Therefore, statistically, RR gives the best agreement when performed locally on soft tissue, with 5-mm and 8-mm CTV expansions.

For a given CBCT image, results are considered successful if all the RR methods give a Dice result higher than that obtained without registration (otherwise they are considered unsuccessful). Fig. 3 (A) shows the boxplots of successful results per RR method (51 CBCT images over 103). Global and bone RRs resulted in the same boxplot aspect. Indeed the brightest pixel intensities in CT and CBCT images are those of the bone structure. The lowest standard deviation (SD) was observed for 1-mm CTV expansion. The bigger the CTV expansion, the more the boxplots looked like those of the global and bone RRs.

IV. Discussion

Fig. 2 shows that the ANOVA could not be applied to compare the RR methods because the distributions do not seem to be normal and the homoscedasticity assumption is not met either. The distributions are all asymetric, the left side tails being longer that the right side one in all cases. The reason for that seems to be dual. Other than the fact that RR is unable to account for the change in prostate shape (a deformable registration would be able to do so) and hence to match the prostate perfectly well on both images, the inter-modality and intra-observer variability of prostate delineation on CT and CBCT scans reduces the expected actual Dice coefficient [13], [14]. In other words, even in the ideal situation of a perfect agreement between the automatic CBCT segmentation (derived from the transformation of the CT manual segmentation) and the manual CBCT segmentation, the Dice coefficient will never reach 1.

Smitsmans *et al.* reported that CT/CBCT RRs mainly failed because of streaks in the CBCT scans caused by moving gas pockets in the rectum during CBCT acquisition [15].

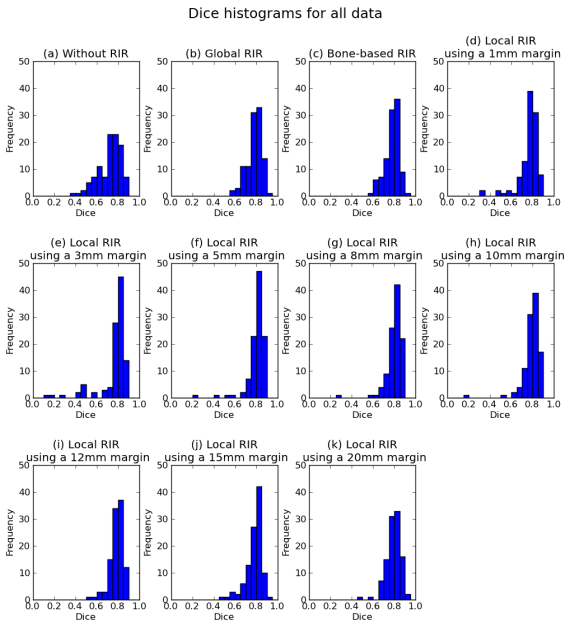

Fig. 2: *Histograms of the Dice results obtained with several methods of CT/CBCT rigid image registration (RIR) methods: (a) without RIR, (b) global RIR, (c) bone RIR, and (d) to (k) a bone RIR refined by a local soft-tissue RIR using the CT clinical target volume expanded with a 1, 3, 5, 8, 10, 12, 15 and 20-mm margin, respectively. There were 20 bins with 0.05 width from 0 and 1.*

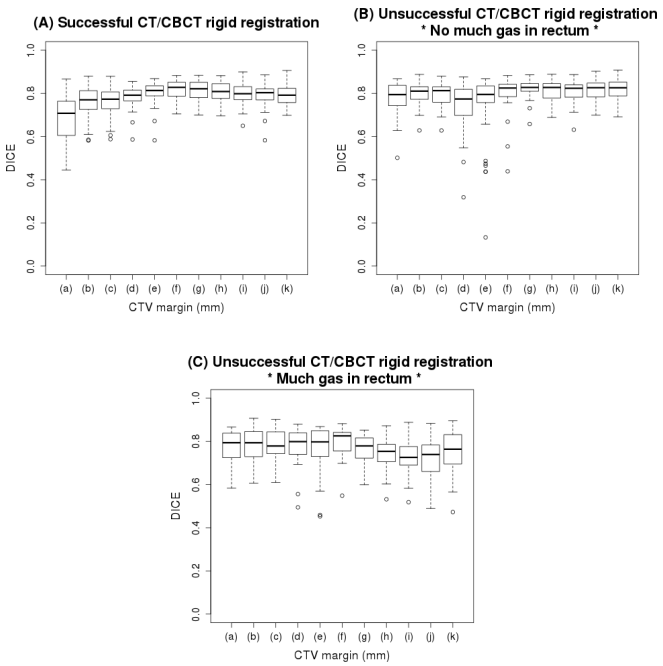

Fig. 3: *Rigid registration (RR) results have been split into two groups: successful and unsuccessful results. RR is said to be successful if it gives a Dice result higher than that obtained without registration. For a given CBCT image, if at least one of the RR methods gives a lower Dice result than that obtained without registration, all RR results for that CBCT image are considered unsuccessful. (A) Boxplots of successful results per RR method, namely (a) without RIR, (b) global RR, (c) bone RR, and (d) to (k) bone RR refined by local soft-tissue RR using the CT clinical target volume expanded with a 1, 3, 5, 8, 10, 12, 15 and 20-mm margin, respectively. Boxplots of unsuccessful results per RR method with (B) the ratio of the mean pixel intensity in rectum to that in prostate, r, lower than 1.3, and (C) r higher than 1.3. The bottom and top of each box are the 25th and 75th percentiles (the lower and upper quartiles, respectively), and the band within the box is the 50th percentile (the median). The ends of the whiskers represent the lowest datum still within 1.5 interquartile range (IQR) of the lower quartile, and the highest datum still within 1.5 IQR of the upper quartile. Anything outside these fences are considered as outliers and marked as dots.*

Furthermore we investigated the influence of air quantity in the rectum on prostate RR quality. Let r be the ratio of the mean pixel intensity in the rectum to that in the prostate. r represents the proportion of air in the rectum (an air-free empty rectum and a prostate, being both soft tissue, have the same range of pixel intensities). The unsuccessful results were split into two groups characterized by lower and higher r values, respectively. We observed that a critical r value of 1.3 resulted in the emergence of two meaningful groups, as shown in Fig. 3 (B) (33 CBCT images over 103) and Fig. 3 (C) (19 CBCT images over 103). For CBCT images that contained relatively little gas in the rectum ($r < 1.3$, Fig. 3 (B)), registration failures concerned global RR, bone RR or local RR with small margins (1, 3 or 5 mm) only. Local RR with an 8-mm margin yields the best results. When global or bone RR fails, the first reason to consider should be the motion/deformation of the prostate gland with respect to the bone pelvis structure. When local RR with small margins fails, it could be caused by the lack of contrast and/or the frequently observed presence of (moving or not) gas pockets situated in the rectum and contiguous with the prostate membrane. For CBCT that contained much gas in the rectum ($r > 1.3$, Fig. 3 (C)), all RRs tended to prove insufficient. We note that all successful results concerned CBCT images with $r < 1.3$, except for three cases ($r = 1.33$, $r = 1.34$ and $r = 1.41$).

V. Conclusion

Automatic RR results strongly depended on the presence of air in the rectum. If relatively little gas was present in the rectum, the best candidate for systematically significantly improving prostate localization was found to be bone RR followed by local soft-tissue RR with an 8-mm margin around the CT CTV. For CBCT that contained much gas in the rectum, all RRs tended to prove insufficient. Future work will investigate the ability of deformable registration to better localize prostate on CBCT images for IGRT.

References

[1] A. L. Zietman, M. L. DeSilvio, J. D. Slater, C. J. Rossi, Jr, D. W. Miller, J. A. Adams, and W. U. Shipley, "Comparison of conventional-dose vs high-dose conformal radiation therapy in clinically localized adenocarcinoma of the prostate: a randomized controlled trial." *JAMA*, vol. 294, no. 10, pp. 1233–1239, Sep 2005. [Online]. Available: http://dx.doi.org/10.1001/jama.294.10.1233

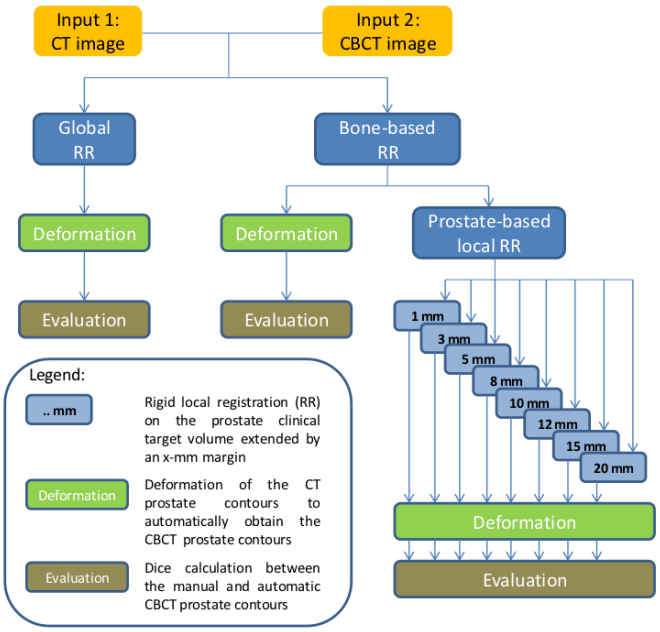

Fig. 4: *The pipeline followed in this study to evaluate different methods of rigid registration (RR) for contour propagation from CT to CBCT images and validation*

[2] K. E. Deurloo, R. J. Steenbakkers, L. J. Zijp, J. A. de Bois, P. J. Nowak, C. R. Rasch, and M. van Herk, "Quantification of shape variation of prostate and seminal vesicles during external beam radiotherapy," *International Journal of Radiation Oncology*Biology*Physics*, vol. 61, no. 1, pp. 228 – 238, 2005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360301604025830

[3] D. A. Jaffray, J. H. Siewerdsen, J. W. Wong, and A. A. Martinez, "Flat-panel cone-beam computed tomography for image-guided radiation therapy," *International Journal of Radiation Oncology*Biology*Physics*, vol. 53, no. 5, pp. 1337 – 1349, 2002.

[4] J. H. Siewerdsen and D. A. Jaffray, "Cone-beam computed tomography with a flat-panel imager: magnitude and effects of x-ray scatter." *Med Phys*, vol. 28, no. 2, pp. 220–231, Feb 2001.

[5] M. H. P. Smitsmans, F. J. Pos, J. de Bois, W. D. Heemsbergen, J.-J. Sonke, J. V. Lebesque, and M. van Herk, "The influence of a dietary protocol on cone beam ct-guided radiotherapy for prostate cancer patients." *Int J Radiat Oncol Biol Phys*, vol. 71, no. 4, pp. 1279–1286, Jul 2008. [Online]. Available: http://dx.doi.org/10.1016/j.ijrobp.2008.03.036

[6] L. E. Court and L. Dong, "Automatic registration of the prostate for computed-tomography-guided radiotherapy." *Med Phys*, vol. 30, no. 10, pp. 2750–2757, Oct 2003.

[7] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3d active contour segmentation of anatomical structures: significantly improved efficiency and reliability." *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, Jul 2006. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2006.01.015

[8] S. P. Raya and J. K. Udupa, "Shape-based interpolation of multidimensional objects." *IEEE Trans Med Imaging*, vol. 9, no. 1, pp. 32–42, 1990. [Online]. Available: http://dx.doi.org/10.1109/42.52980

[9] L. Ibanez, W. Schroeder, L. Ng, and J. Cates, *The ITK Software Guide*, 2nd ed., Kitware, Inc. ISBN 1-930934-15-7, http://www.itk.org/ItkSoftwareGuide.pdf, 2005.

[10] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, July 1945. [Online]. Available: http://www.jstor.org/pss/1932409

[11] M. Hollander and D. A. Wolfe, *Nonparametric statistical methods*. Wiley-Interscience, 1999.

[12] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2012, ISBN 3-900051-07-0. [Online]. Available: http://www.R-project.org

[13] Z. Gao, D. Wilkins, L. Eapen, C. Morash, Y. Wassef, and L. Gerig, "A study of prostate delineation referenced against a gold standard created from the visible human data," *Radiotherapy and Oncology*, vol. 85, no. 2, pp. 239 – 246, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167814007003891

[14] E. Weiss, J. Wu, W. Sleeman, J. Bryant, P. Mitra, M. Myers, T. Ivanova, N. Mukhopadhyay, V. Ramakrishnan, M. Murphy, and J. Williamson, "Clinical evaluation of soft tissue organ boundary visualization on cone-beam computed tomographic imaging." *Int J Radiat Oncol Biol Phys*, vol. 78, no. 3, pp. 929–936, Nov 2010. [Online]. Available: http://dx.doi.org/10.1016/j.ijrobp.2010.02.007

[15] M. H. P. Smitsmans, J. de Bois, J.-J. Sonke, A. Betgen, L. J. Zijp, D. A. Jaffray, J. V. Lebesque, and M. van Herk, "Automatic prostate localization on cone-beam ct scans for high precision image-guided radiotherapy." *Int J Radiat Oncol Biol Phys*, vol. 63, no. 4, pp. 975–984, Nov 2005. [Online]. Available: http://dx.doi.org/10.1016/j.ijrobp.2005.07.973