CHAPTER 5

Discussion

5.1 DIR accuracy on MVCT images

Regarding the DIR accuracy evaluation, the known and unknown offset investigation were assessed on the phantoms and NPC patients with rigid and non-rigid volume changes. The agreement between the automatically deformed contour using 8 DIR methods and the reference structure was shown to be dependent on the areas of interest. This study demonstrated that DIR methods in MVCT images were not significantly different from kVCT images, based on intensity, volume, and physical characteristics of the deformation field. The results demonstrated that the use of different deformation models and different ROIs affects the accuracy of DIR.

Regarding the rigid volume changes for both phantoms and clinical cases, there were concordance of the results with MSD, CC and NMI. The Horn and Schunck optical flow showed the better performance than the Demons in both transformations and mapping directions as showed in the AsyHS$_{BW}$, AsyHS$_{FW}$, SymHS$_{BW}$ and SymHS$_{FW}$ methods. As regards the non-rigid volume changes, there were concordance of the results with MSD and CC. The Horn and Schunck optical flow still showed the better performance than the Demons in the AsyHS$_{BW}$, AsyHS$_{FW}$, SymHS$_{BW}$ and SymHS$_{FW}$ methods. Yeo *et al* (2013) assessed the accuracy of 12 DIR algorithms from DIRART software and quantitatively examine low contrast regions by developing the deformable gel (DEFGEL). The greatest accuracy was exhibited by the original Horn and Schunck optical flow algorithm and the modified demons algorithm exhibited the greatest error.

As regards the volume-based analysis, there were concordant between the phantom and NPC cases evaluation. The DSC values were also concordant with the intensity-based analysis. The DIR with AsyHS_{BW}, AsyHS_{FW}, SymHS_{BW} and SymHS_{FW} methods also showed better performance in both rigid and non-rigid organ. However, for the AsyDM_{FW}, the mean value of DSC was found to be significantly increasing in the non-rigid changes and reached the best performance with DSC = 0.812 ± 0.07 for NPC cases. Rigaud *et al.* (2015) compared the performance of ten DIR approaches using different registration methods (Demons or B-spline free-form deformation (FFD)), preprocessing, and similarity metrics. The most effective DIR methods were the demons with both the mutual information metric and the filtered CTs. The mean value of DSC for Demons with original CTs with mean square error metric were 0.75 for parotid gland and showed that the choice of the metric or of the image preprocessing was at least as important as the registration method. Moreover, Varadhan *et al*. (2013) described a framework to test the accuracy of DIR using both B-spline and diffeomorphic demons algorithms with the forward and inverse direction. For head and neck study sets, the mean value of DSC for diffeomorphic demons were 0.74.

Regarding the individual investigation in ROIs, Figure 5.1 demonstrates the overlapping volume analysis of the target and OAR with the DSC value. Mostly, the best DIR methods for both kVCT and MVCT images were the AsyDM_{FW} methods. Regarding the MVCT images, normal organs were more accurately deformed than the target structure, with a mean value of DSC for the target were 0.770 ± 0.09 (GTV), 0.842 ± 0.066 (CTV), 0.483 ± 0.11 (nodal-GTV), and 0.660 ± 0.17 (nodal-CTV) with AsyDM_{FW} methods. The mean value of DSC for the OARs were 0.810 ± 0.06 (left parotid gland, AsyHS_{FW}), 0.824 ± 0.06 (right parotid gland, AsyDM_{FW}) and 0.806 ± 0.06 (spinal cord, AsyHS_{BW}). The average of DSC value was less than 0.7 for nodal-GTV and nodal-CTV, the results showed unsatisfactory volume matching for adaptive radiotherapy application. Hardcastle *et al.* (2012) showed good agreement between the DIR-propagated ROIs and the expert-drawn ROIs. About 94% of all ROIs generated using DIR were scored as clinically useful, requiring minimal or no edits. However, 27% (12/44) of the GTVs required major edits. Hardcastle *et al.* (2012) also showed that the agreement was not as good for the nodal-GTVs due to poor soft-tissue contrast surrounding these structures.

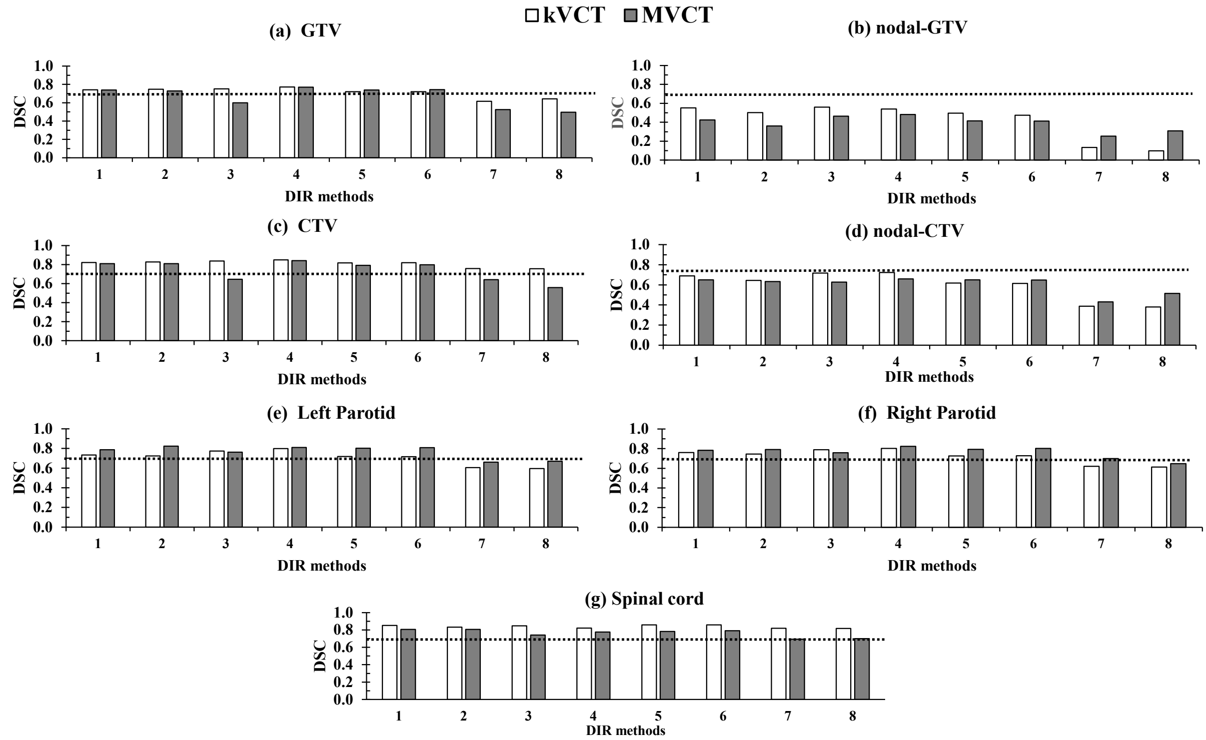

Figure 5.1 The dice similarity coefficient (DSC) comparison of kVCT and MVCT images for all targets (a) GTV, (b) nodal-GTV, (c) CTV, (d) nodal-CTV and all organs (e) left parotid gland, (f) right parotid gland and (g) spinal cord in eight DIR methods (1=AsyHS$_{BW}$, 2=AsyHS$_{FW}$, 3=AsyDM$_{BW}$, 4= AsyDM$_{FW}$, 5=SymHS$_{BW}$, 6=SymHS$_{W}$, 7=SymDM$_{BW}$ and 8=SymDM$_{FW}$)

For the deformation field analysis in terms of ICE, mostly the asymmetric transformation showed better performance than the symmetric transformation. The average of ICE for all DIR methods in asymmetric and symmetric transformation were 0.136 and 0.788 mm, respectively. Monica *et al*. (2013) assessed and evaluated different methods of computing the inverse of a B-Spline transformation for five head and neck patients. The results showed the average of ICE were 1.755 mm (asymmetric) and 0.462 mm. (symmetric) transformation.

The accuracy of DIR for kVCT and MVCT images showed not significantly difference with validation techniques, intensity-based, volume-based and deformation field. Although, the NMI on kVCT images for rigid volume changes in phantom investigation showed the significantly better than MVCT images with the mean NMI were 1.261 (kVCT) and 1.198 (MVCT) but the NMI value of MVCT images can yield the acceptable value. Whereas, Monica (2013) which assessed and evaluated different methods of computing the inverse of a B-Spline transformation, showed the mean of NMI for DIR on kVCT images were 1.138 and 1.147 for asymmetric and symmetric transformation, respectively.

Regarding the accuracy of DIR, normally, kVCT images have more accuracy in the case of DIR generated by high contrast. However, as the kVCT images in Figure 5.1, left (e) and right (f), demonstrate, the DSC value of the parotid glands reveals the poor performance when compared with the DIR in the MVCT images. The inferior results are due to the fact that the two NPC patients had dental filling as in Figure 5.2. As for the kVCT images, the metal (much higher atomic number, Z) increased the photoelectric attenuation and caused streak artifacts, but for the MVCT images, streak artifacts were found to be minimum (Ruchala *et al*., 1999). Artifact images decrease the accuracy of DIR. The means of the DSC values of both the parotid glands in patients with dental filling were 0.651±0.12 (left) and 0.666±0.11 (right) on the kVCT images.

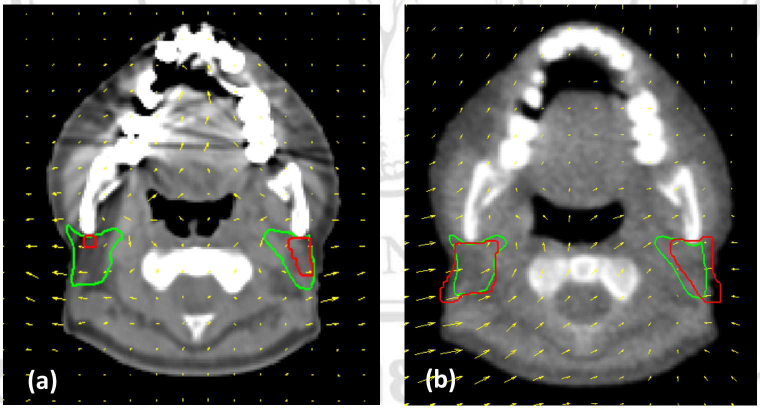

Figure 5.2 The automatic deformed contour (red-line) and reference contour (green-line) in transverse images of (a) kVCT and (b) MVCT of NPC patient with dental filling.

This study describes the application of known deformations on any image data set to evaluate the accuracy and limitations of a DIR algorithm used in radiation oncology. These two techniques using phantom and clinical MVCT images provided possibilities to verify a variety of the deformation methods for the quality assurance of DIR. The results for the volume matching by the dice similarity coefficient can yield the acceptable level for application in adaptive radiotherapy (Zimring *et al*., 2005).

5.2 Accumulated dose evaluation

The validation of DIR on weekly MVCT images were consistent in terms of the volume-based criterion, DSC, and the deformation field analysis, ICE. Accuracy in terms of DSC analysis tended to decrease as the treatment progressed as a result of organs with large-scale deformation causing reduction in the DIR accuracy (Fabri *et al*., 2013). The AsyDM_{FW} method showed the best performance with the highest average DSC value of all ROIs for entire the treatment, the mean value of DSC = 0.804 ± 0.07, and also enabled the minimization of the inverse consistency error by the lowest mean value of ICE = 0.006 ± 0.002 mm.

As regards the weekly dose variation from the initial plan, the results demonstrated that the median, the near-minimum dose, and the near-maximum dose of the target slightly varied, by less than 0.5% of the initial plan. However, with regard to organ doses like the bilateral parotid gland, the discrepancy between the planned and the delivered mean doses was 6.8% (right) and 15.8% (left) higher than the initial plan. The accumulated mean parotid dose increased to be in the range of -0.6 to 12.4 Gy. Lee *et al*. (2008) analyzed the changes in the parotid gland dose with reference to the anatomic changes throughout the course of radiotherapy. The daily parotid mean dose of the 10 patients differed from the planned dose by an average of 15%. At the end of the treatment, 3 of the 10 patients were estimated to have received greater than 10% higher mean parotid dose than in the original plan (range: 13 to 42%), whereas the remaining 7 patients received doses that differed by less than 10% (range: 6 to 8%).

Regarding the correlation between dose accumulation and DIR accuracy, there was concordant between the accuracy of ROI deformation and discrepancy of dose accumulation. The DIR method that yielded the highest DSC value was considered the best method for dose accumulation with the lowest variation from the reference dose. The AsyDM_{FW} method demonstrated the best performance for target deformation at the end of treatment with the highest mean DSC value of 0.836 ± 0.03, as presented in Figure 4.19, and showed the best agreement for target dose accumulation with the lowest average variation of 0.05% with the reference dose, as presented in Figure 4.22. This method also gave the highest mean DSC value of 0.766 ± 0.05 for the right parotid gland and the lowest mean parotid dose variation with 3.2% of the reference deformed dose, as demonstrated in Figure 4.23 (a). Whereas, the $SymHS_{FW}$ method gave the best performance for left parotid and spinal cord with the highest mean value of DSC = 0.808 ± 0.06 and 0.790 ± 0.05, respectively. Therefore these methods showed the lowest dose variation with 2.9% and 0.9% of the reference deformed dose, as presented in Figure 4.23.

Regarding the DIR accuracy on weekly MVCT images, the three DIR methods demonstrated satisfactory volume matching for accumulated dose application with DSC values more than 0.7 for all the methods. Moreover, the one-way ANOVA analysis demonstrated that there was no significant difference between the three DIR methods as regards ROI deformation and dose accumulation. However, when uncertainty was considered (difference between the maximum dose and the minimum dose) in the estimation of the accumulated dose for all the DIR methods, the average of the DSC value by using the three DIR methods at the end of treatment for targets were 0.816 ± 0.02. The mean uncertainty for estimating the target dose were 0.21 ± 0.11 Gy (range: 0.06 to 0.32 Gy). Regarding the uncertainty of the parotid dose, the averages of the DSC values by the three DIR methods were 0.816 ± 0.02 (right) and 0.799 ± 0.01 (left). This shows that the mean uncertainty values for estimating the parotid dose were 1.99 ± 0.76 Gy (range: 0 to 3.14 Gy) for the right parotid and 1.19 ± 0.24 Gy (range: 0 to 1.58 Gy) for the left parotid. For the spinal cord, the average of the DSC values was 0.762 ± 0.03, while the mean uncertainty value for the estimated dose was 0.407 ± 0.04 Gy (range: 0.04 to 0.57 Gy). Rigaud *et al*. (2015) compared the performance of ten DIR approaches using different registration methods (demon and B-spline free form deformation), preprocessing, and similarity metrics, then quantify their impact for dose accumulation, in health structure. The results showed that the mean uncertainty (difference between the maximum dose and the minimum dose, considering all the 10 DIR methods) to estimate the cumulated mean dose for the parotid gland was 4.03 Gy (SD = 2.27 Gy, range: 1.06 to 8.91 Gy).

Further investigation would involve applying this methodology to a larger population and other treatment areas to identify patients who may benefit from adaptive treatment. DIR on megavoltage computed tomography imaging makes it possible to calculate daily and accumulated doses. Significant dose variations were observed as a result of inter-fractional anatomic changes, which is information that would benefit adaptive treatment strategies.