Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

A Clinical Procedure for Case-Specific Analytical Validation of Mono-Modality Image Fusion in Image Guided Radiotherapy

Joseph Hanley, Joseph Barbiere, and Yulin Song

Abstract – **Image guided radiotherapy can be performed by fusing the daily treatment and reference planning Computed Tomography scans. Decreased errors in patient setup can lead to smaller target margins that significantly improve treatment efficacy and outcome. The purpose of this work is to present a clinical procedure to analytically compute the accuracy of the registration. Accepted techniques such as Normalized Mutual Information intensity based three-dimensional image registration can be validated using a large automated point sample. Without a user independent metric it is not possible to determine effect of the fusion error on the calculated correction in patient setup.**

The use of in-room Computed Tomography (CT), KV and MV cone beam CT, and Helical Tomotherapy have been introduced to assist in achieving reproducible patient setup [1,2]. Image guided radiotherapy can then be performed by fusing the daily treatment and reference planning CT scans. However, the accuracy of the registration is difficult to quantify and often we resort to a simple qualitative visual assessment as shown in Figure 1. Standard point-based analytical quantitative metrics may require invasive placement of fiducial markers or variable user dependent identification of corresponding anatomical landmarks. Furthermore, the point sample is very small, often only three or four and seldom over twelve.

Numerous previous studies have noted that the visual approach inherently lacks scientific rigor and the landmark approach is both statistically insignificant and user dependent [3]. The purpose of our work is to present an analytical procedure that can be used clinically to validate the fusion accuracy on a case-by-case basis. The procedure includes methodology to automatically extract a relatively large number of sample points for computation of a standard error metric.

There is no gold standard for fusion and competing technologies are available. Errors as large as 6 mm have been reported. If the goal of image-guided radiotherapy is to reduce target margins from 5 mm to 3 mm for example, then fusion error component can be significant and should be evaluated for each patient. The procedure presented in this work is intended for clinical application. The tools required are already available as part of treatment planning systems.

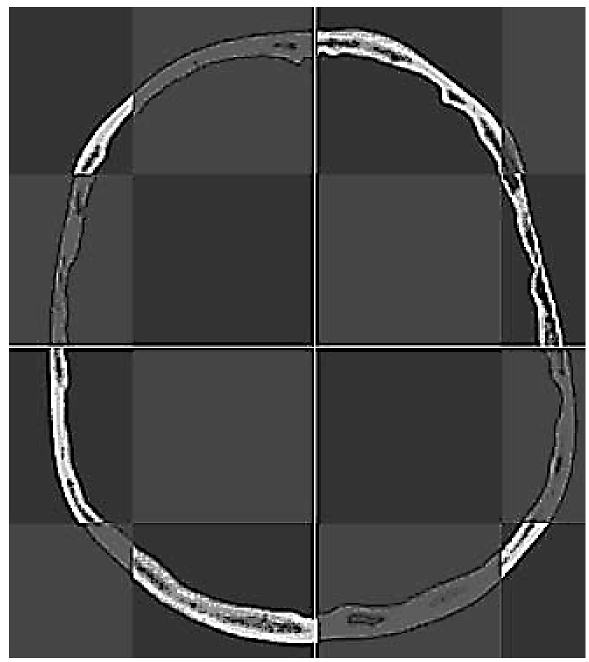

Figure 1. Visual qualitative validation in checkerboard format. There are noticeable misregistrations especially in the upper right quadrant that require quantitative analysis.

The basic premise of CT-CT Image Guided Radiotherapy (IGRT) setup correction is that the required correction in patient position at the time of treatment is equivalent to the rigid body transformation which fuses the plan CT to the setup CT as closely as possible. The procedure described in this work as outlined in Figure 2 is designed to analytically validate the fusion accuracy which depends on variations in the non-rigid anatomy, mechanical factors such as couch sag or digitization accuracy, and the fusion algorithm itself including items such as sample size or the initial parameters selected by the operator.

II. DATA

A. Reference

Data was analyzed using Philips Pinnacle treatment planning system (TPS) that includes Syntegra as a Normalized Mutual Information (NMI) fusion algorithm.

Joseph Hanley is with the Department of Radiation Oncology, Hackensack University Medical Center, Hackensack, NJ 07601 USA (e-mail: JHanley@ humed.com).
Joseph Barbiere is with the Department of Radiation Oncology, Hackensack University Medical Center, Hackensack, NJ 07601 USA (e-mail: JBarbiere@humed.com)
Yulin Song is with the Department of Medical Physics, Memorial Sloan-Kettering Cancer Center, New York, NY 10021 (e-mail: songy@mskcc.org)

With the advent of DICOM RT format standards it is possible to seamlessly transfer images between systems so that our process can be applied regardless of the original image acquisition system. Two diagnostic CT scans on a patient being treated for Head and Neck (H&N) carcinoma are an example of the necessary data. The primary reference CT data set labeled PLAN is used for treatment planning. During the actual treatment the secondary daily CT data set labeled SETUP is acquired and fused with PLAN. The CT data set consists of a series of 512x512 matrices consisting of relative CT numbers in Hounsfield Units (HU) such that air is 0 HU and water is 1000 HU.

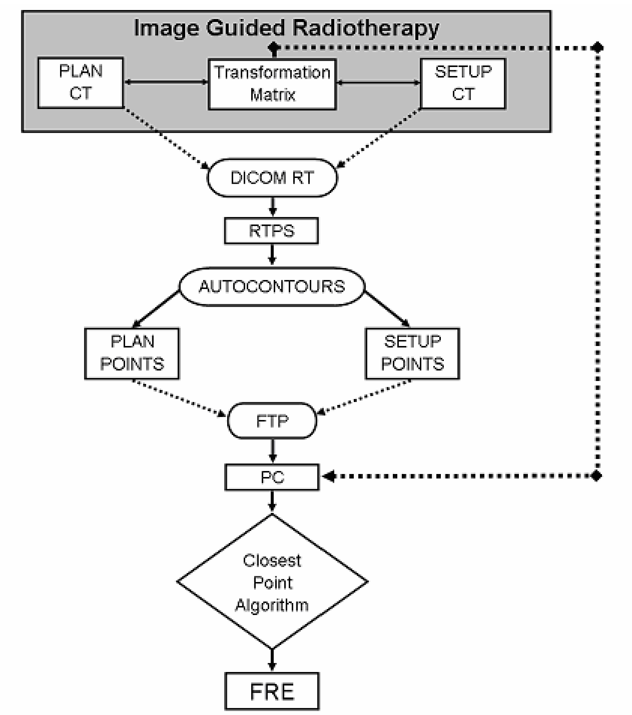

Figure 2. The process utilizes available data, hardware, and software. FRE is a standard analytical validation error metric.

Autocontouring tools that draw a contour on pixels that cross a designated threshold seed value are part of the RTPS. As proof of principle we selected the outer cranial surface and subarachnoid space but the same procedure can be applied to other anatomical structures such as the ventricles, optic globes, or inner parietal bone. In order to obtain comparable primary and secondary data we analyzed the bone-tissue interface as shown in Figure 3 and set the threshold to 1600 HU. Figure 4 shows a primary skull contour consisting of 404 points with a mean spacing of 0.12 mm suitable for a complete primary PLAN data set at 0.30 mm slice spacing. Typically for the skull we obtain 30 slices for a total of approximately 12,000 points.

B. Transformed

The secondary SETUP data set was fused to the primary using NMI [4]. The algorithm computes the homogeneous transformation matrix M such that the error [PLAN – M*SETUP] is minimized. M contains three rotational and three translational components about each axis. It is possible to simultaneously contour on both the reference and transformed data sets.

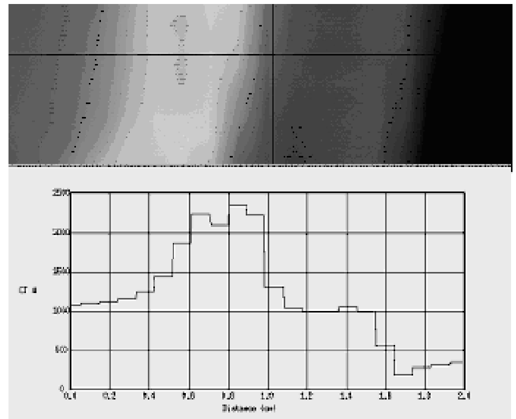

Figure 3. A CT number profile of the outer skull was used to determine a threshold of 1600 HU for all autocontouring of the outer skull.

For analysis we employ a point reduction tool and sample the transformed contour at approximately 5.00 cm intervals. As indicated on Figure 4, there are on the average about 10 data points on the skull per slice that yield about 200 points for 20 slices for sample frequency is 1.7%. It is important to note that for typical manual entries our automated technique increases the sample size by an order of magnitude.

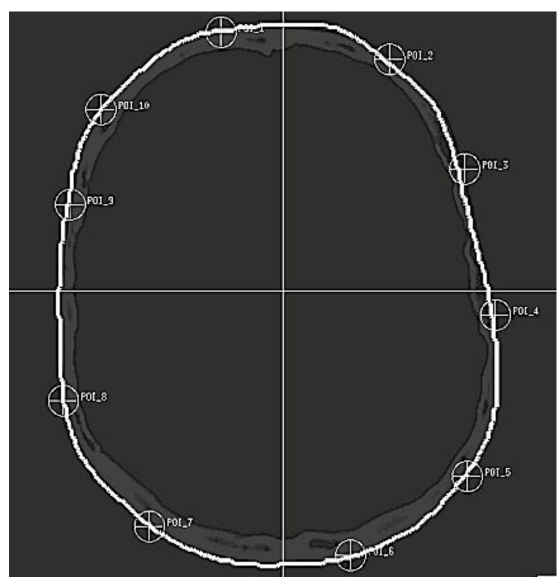

Figure 4. The outer skull reference contour consists of 404 individual points. Ten transformed points are sampled for quantitative analytical validation.

As an added feature, we are also able to similarly contour the spine in a saggital plane. This is a very efficient method to quickly evaluate in-plane transformation for the cord, which is a major critical structure in H&N radiotherapy. The paradox with NMI is that it seeks a global best fit solution for a rigid body homogenous transformation which encompasses the entire data set including areas such as the neck that are flexible. Deformable image registration option is not appropriate for mechanical setup correction which is physically limited to uniform translation and rotation without scaling, shear, or local perturbations. However, using our validation technique it is possible to generate individual contours for each structure and make anatomy specific validation calculations. It may then be possible for the user to compromise the overall fusion accuracy by manually adjusting a transformation parameter to avoid potentially very harmful overdosage to a critical structure.

III. ERROR METRIC

The primary and secondary contours are a series of points stored in a text file by the TPS. The file is exported to an Excel spreadsheet on standard PC. After text to column conversion the coordinate start point and length in the text file for each contour are identified and all contours converted to Euclidean coordinate points. The 4x4 homogenous transformation matrix is also exported so that the primary and secondary contours can be registered.

Let Dij be the distance from a point Pi in the set PLAN to point Sj in the set SETUP as shown in Figure 5. The 12,000x200 matrix D is well within the Excel size limitations. The minimum value of each column is Dj and is the registration error for point Sj. This process is analogous to the search performed by Iterative Closest Point (ICP) algorithms. Fiducial Registration Error (FRE) is a standard error metric [5] and is defined as

$$FRE^2 = 1/n \sum_{j=1,n} D_j^2 .$$

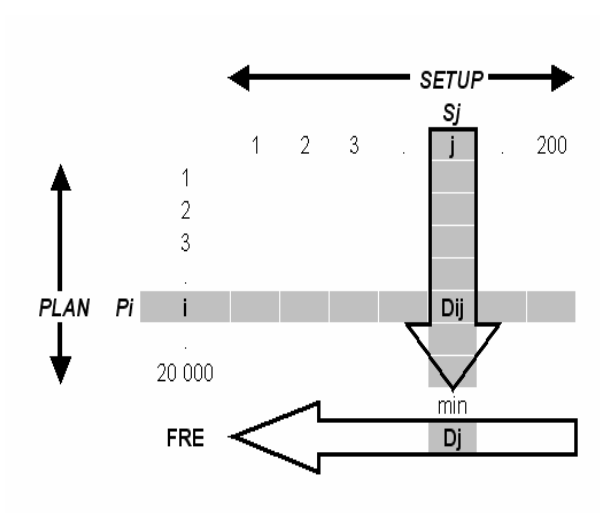

Figure 5. The FRE calculation matrix. Since the 20,000x200 array is in one spreadsheet worksheet, it is possible to dramatically increase the sample size by simply adding more worksheets.

For sensitivity evaluation we varied the translation and rotation parameters for three adjacent slices and summarized the results in Table 1. The technique appears to have sufficient sensitivity to distinguish individual translations and rotations of one millimeter and one degree. Composite transformations about all axes are also clearly indicated.

	cm						
Tx	0.0	0.1	0.0	0.0	0.1	0.2	0.3
Ty	0.0	0.0	0.1	0.0	0.1	0.2	0.3
Tz	0.0	0.0	0.0	0.1	0.1	0.2	0.3
FRE	0.0139	0.0841	0.0834	0.1010	0.1416	0.2253	0.2957
FRE/FRE0	**1**	**6**	**6**	**7**	**10**	**16**	**21**
	deg						
Ax	0.0	1.0	0.0	0.0	1.0	2.0	3.0
Ay	0.0	0.0	1.0	0.0	1.0	2.0	3.0
Az	0.0	0.0	0.0	1.0	1.0	2.0	3.0
FRE	0.0139	0.1027	0.0893	0.0552	0.1258	0.1313	0.2110
FRE/FRE0	**1**	**7**	**6**	**4**	**9**	**9**	**15**

Table 1. Random contour points on adjacent slices were used to generate simulated fused data. FRE0 was then the error associated with no transformation. The ratio of the FRE as the data is translated by Tx, Ty, and Tz or rotated by Ax, Ay, and Az about each axis to FRE0 indicates the sensitivity to small individual and combined transformations.

IV. RESULTS

A. Skull

A representative set of data is shown in Figure 6. The NMI fusion FRE was 0.2387 cm and is in general in excellent agreement with the visual observation. By comparing sample data points from adjacent CT slices to a contour between the two samples we demonstrate the relative insensitivity to the 0.3 cm slice thickness. The analytical evaluation indicates that 0.2 setup error and 0.3 cm margins in the brain area are appropriate in this case.

B. Spine

A sagittal slice through the cord is shown in Figure 7. The NMI fusion FRE was 0.3553 cm and is consistent with the visual observation. The analytical evaluation indicates that 0.2 setup error and 0.3 cm margins in the cord area are inappropriate in this case. Individual analysis of the error due to each transformation parameter indicates that most of the error is due to a translation in the vertical axis. Therefore, the user may wish to alter the NMI global result to obtain a locally clinically acceptable setup in the very critical cord area.

A recent article has also noted differences in relative positional variations within the head and neck area [6]. The authors note that individual structures should be considered both for setup correction and setting target margins.

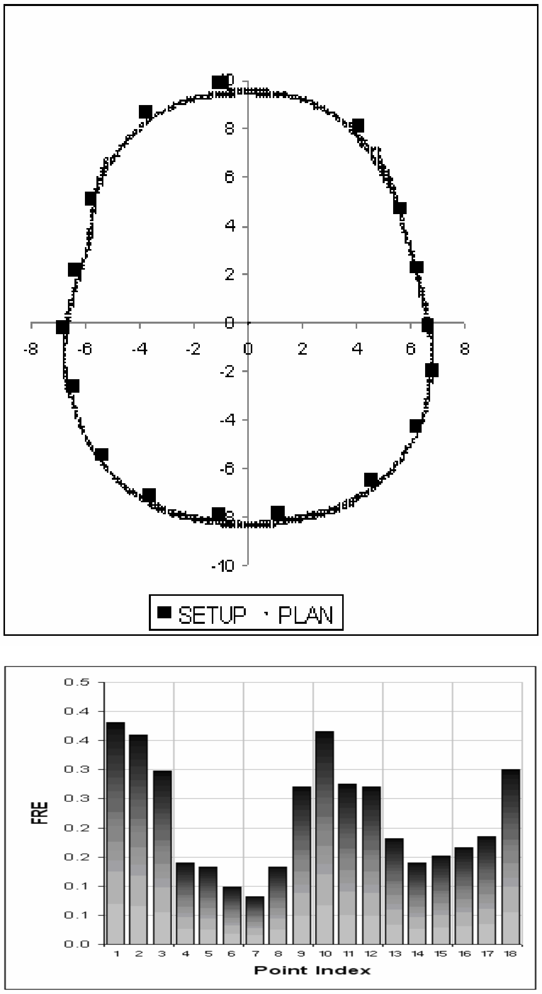

Figure 6. A graphical display of three adjacent PLAN contours and eighteen neighboring transformed SETUP points. The error distribution for the individual points is also displayed. The FRE for all points is 0.2387.

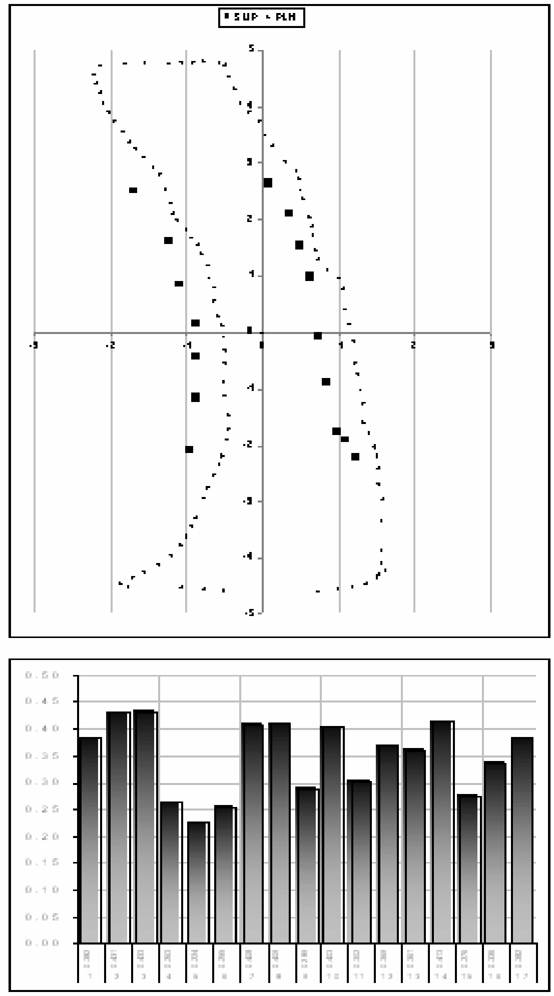

Figure 7. The SETUP sample points for the flexible cord are clearly displaced from the PLAN contour after the NMI fusion. The FRE shown for the individual points is equivalent to 0.3553 for the entire structure.

V. CONCLUSION

We have described a procedure to effectively increase the number of sample points used in quantitative evaluation of a commercial fusion algorithm for image guided radiotherapy. A representative clinical example demonstrates that the required data can be automatically generated and extracted to compute an analytical error metric. For a Head and Neck patient the calculated transformation requires manual adjustment to avoid harmful irradiation of the sensitive flexible spinal cord.

VI. REFERENCES

[1] *MedicalDosimetry,* vol. 31, issue 1, pp. 1-88, 2006.
[2] L. J. Forrest, T. R. Mackie, et. al., "The utility of megavoltage computed tomography images from a helical tomotherapy system for setup verification purposes," *International Journal of Radiation Oncology Biology Physics,* vol. 60, issue 5, pp. 1639-1644, 1998.
3] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis,* vol. 2, no. 1, pp. 1-36, 1998.
[4] F. Maes, A. Collington, D Vandermeullen, G. Marchal, and P. Stuens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging,* vol. 16, no. 2, pp. 187-198, 1997.
[5] J. M. Fitzpatrick, J. B. West, and C. R. Maurer, "Predicting error in rigid-body point-based registration," *IEEE Transactions on Medical Imaging*, vol. 17, pp. 694-702, 1998.
[6] L. Zhang, A. S. Garden, et. al., "Multiple regions-of-interest analysis of setup uncertainties for head-and-neck cancer radiotherapy," *International Journal of Radiation Oncology Biology Physics,* vol. 64, issue 5, pp. 1559-1569, 2006.