A Registration Framework for Preoperative CT to Intraoperative White Light Images

M.Gooroochurn[a 1], M.Ovinis[a †], D.Kerr[a ‡], K.Bouazza-Marouf[a §], M.Vloeberghs[b **]

[a] Loughborough University, Loughborough, UK, [b] Queen's Medical Centre, Nottingham, UK

Abstract. A registration framework for image-guided robotic surgery is proposed for three major neurosurgical procedures. The registration uses preoperative CT images and intraoperative white light camera images as modalities. A point-based rigid body registration based on the manual identification of homologous anatomical features in CT and white light images is used. A clinical accuracy of 5 mm is considered sufficient for the three procedures. Tests with CT data and projection views from a surface rendered CT model gave a maximum RMS registration error of 0.68mm when tested over fifteen datasets. This paper discusses the proposed framework and the results obtained from preliminary tests on simulated data.

1 Introduction

Point-based rigid body registration is a common registration technique in image-guided neurosurgery, with skull implanted fiducial markers considered the gold standard. However, because the need for surgical intervention can usually only be established after a scan, a retrospective registration basis such as anatomical features is appealing. Intraoperatively, these landmarks may be found through relatively inexpensive white light imaging. One of the earliest implementations of a CT image-to-patient registration system using white-light imaging was by Colchester et al. [1]. A surface model of a patient, reconstructed intraoperatively using a stereo video system, was matched to a surface derived from CT.

Clarkson et al. [2] used an alignment by mutual information approach described by Viola and Wells [3] to register CT images to multiple video images. They also developed a technique to register two or more video images of the human face to a 3D surface model using a similarity measure based on photo consistency [4]. The technique had an error between 1.45 and 1.59 mm when the initial mis-registration was up to 16 mm/degrees. As these methods are based on intensity rather than features, feature extraction or segmentation is not necessary and are therefore suited in applications where features cannot be reliably extracted. However, these techniques require that the surfaces to be matched be roughly aligned, preventing its use for gross misalignments.

The three targeted neurosurgical procedures are Intracranial Pressure (ICP) Monitoring, External Ventricular Drainage (EVD) and evacuation of a Chronic Subdural Haematoma (CSDH). They are routinely performed using a freehand technique. Image-guided solutions are sought for these procedures for a robotic surgery system operating in conventional medical set-ups. Since the target anatomy is the head, a rigid body registration is performed, using salient features in CT and corresponding features found on the patient's head. Under the proposed method, craniofacial features are selected based on their availability in the two modalities used, their saliency and efficacy of detection as well as their effect on the accuracy of the point-based registration algorithm. The craniofacial landmarks chosen as the registration basis should be visible and reproducible in the CT model and corresponding landmarks need to be found in white light images as well. Employing only facial features for the registration does not guarantee good accuracy as they cover only a small volume of the head. Moreover, the landmarks should be spaced out evenly and located as far away from each other as possible [5]. Furthermore, head CT scans are routinely specified from the base to the vertex of the skull and the nose and mouth features are not normally available in CT images.. In view of these considerations, the ear tragus and the outer eye corner on each side of the head are chosen in this work as natural landmarks for the registration. The selected landmarks are intuitive and straightforward for a non-expert operator to pick manually.

2 Registration Protocol

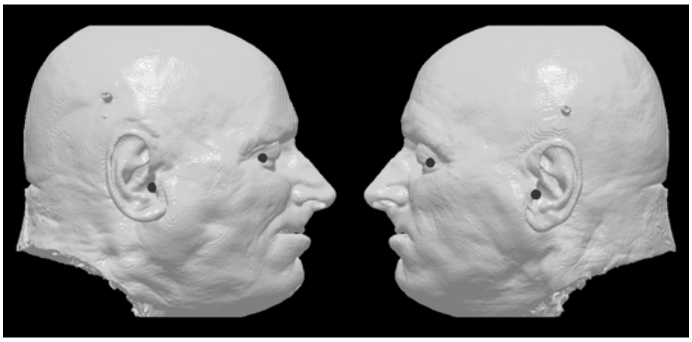
2.1 Simulation of Target Registration Error

Fiducial registration errors (FREs) of 1.0 to 3.0 mm are common for systems that use between eight and sixteen anatomical landmarks. Systems that use skin-affixed markers have typical FREs of less than 2.0 mm and employ six to ten markers. FREs for systems using bone-implanted markers are generally less than 1.0 mm, and use three to five markers [6]. However, in medical applications, target registration error (TRE), the error between the perceived and

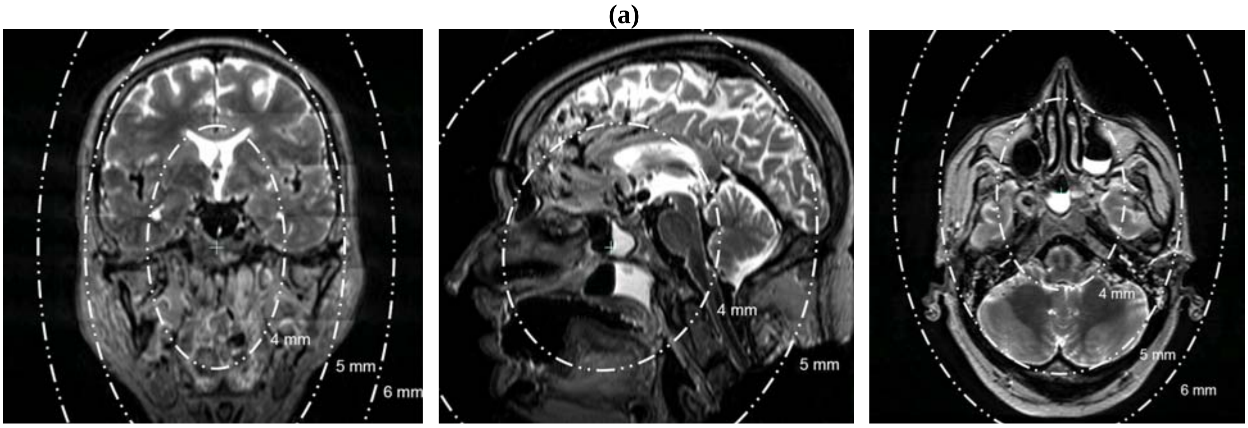
[1] M.Gooroochurn@lboro.ac.uk † M.Ovinis2@lboro.ac.uk ‡ d.kerr@lboro.ac.uk § k.bouazza-marouf@lboro.ac.uk
** Michael.Vloeberghs@Nottingham.ac.uk

true position of the desired location, is a more significant measure of accuracy. A TRE of 5mm is considered sufficient for these three procedures. The required accuracy for the targeted procedures is low as the ‘targets’ in an ICP monitoring bolt placement, EVD and evacuation of a CSDH are in general quite large, providing a large margin of error for positioning. For example, for EVD, the average volume of the ventricles is 30.9 ± 5.7 ml [7] while for the evacuation of a CSDH, the average thickness of a haematoma requiring surgical evacuation is more than 10 mm [8]. In addition, the brain may be held relatively stationary by immobilising the head, reducing potential error due to movement.

The possible TREs using the proposed registration basis has been simulated for landmark localisation errors (FLEs) of 5 mm (based on the uncertainties obtained in practice) in both modalities. The analytical expression of estimated TREs for point-based registration [6] has been used for the simulation. The resultant loci of TREs are shown in Figure 1 on the coronal, sagittal and axial views. The estimated TRE are within the desired 5mm clinical accuracy as the 5mm TRE isocontour envelopes the targeted anatomy. These errors are better in some instances than that of a neurosurgeon performing the procedures using a free-hand technique. In a study by O’Leary [9] which involved 24 patients who underwent an EVD, the accuracy of catheter placement was 9.7 ± 6.3 mm using a free-hand technique. Another more recent study of 97 patients by Huyette et al. [10] found an accuracy of 16 ± 9.6 mm and an average of two passes using the free-hand technique. A high percentage (22.4%) of catheter tips was actually placed in non-ventricular spaces. The proposed registration approach is therefore sufficiently accurate for the targeted procedures.



(a)



(b)

Figure 1: (a) Anatomical landmarks used, the lateral canthus and tragus, and (b) the expected TREs isocontours (in coronal, sagittal and axial views respectively) for FLEs of 5 mm in both modalities.[2]

2.2 Equipment and System Placement

In [11], Ansari et al. uses a camera set-up with two views (frontal and profile) to reconstruct the 3D coordinates of facial features. However, the proposed registration technique uses the ear tragus as a feature. It should thus appear in at least two views to enable 3D reconstruction. An intermediate view, midway between the frontal and profile view would contain both the ear tragus and the eye corner on one side of the face.

Using a technique similar to [11] with these three camera views would enable the reconstruction of the two outer eye corners and one ear tragus fully, and with these three points, point-based registration would be feasible. A schematic of the three camera system is shown in Figure 2(a). However, for the preliminary investigation presented, the 5 camera set-up shown in Figure 2(b) has been used.

[2] MRI data from US National Library of Medicine's Visible Human Project®

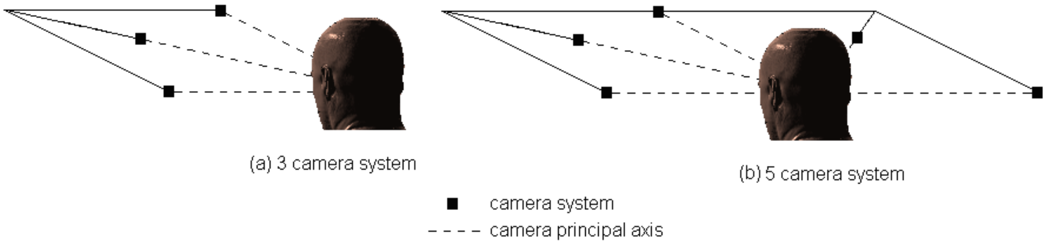


Figure 2: Schematic of Camera System

Figure 3 shows the projected frontal, profile and intermediate views of the head obtained from a CT surface rendered model based on this set-up. A simple Direct Linear Transformation (DLT) method has been used to calibrate these views and reconstruct 3D coordinates. The method described in Section 3 uses such simulated views from CT models to generate a frontal, two intermediate and two profile views. These are then used for simulating the errors in reconstructing the selected features and the registration error for mapping a preoperative plan to the reconstructed coordinate system. Performing reconstruction from stereo views requires a calibrated camera system. The workspace should be calibrated such that the object to be reconstructed lies within the calibrated volume as extrapolation outside that space can lead to large errors [12]. A calibration object encompassing the volume of the human head should thus be used for calibration.

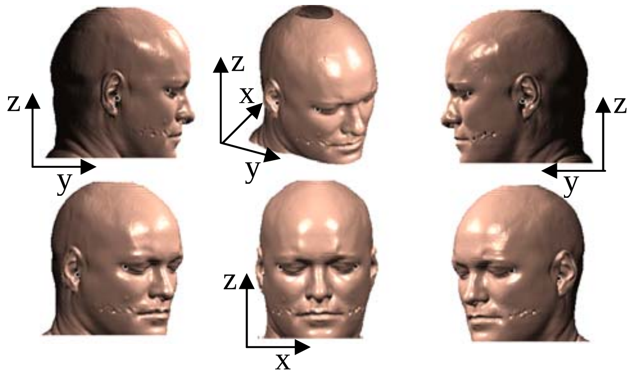


Figure 3: Projected Head Views[3]

Any image captured should be delimited to fall within the field of view in each camera corresponding to that occupied during calibration. This ensures accurate reconstruction of the landmarks. Figure 4 illustrates three possible scenarios of coverage for the field of view of the frontal camera with respect to the calibrated space. Case (b) has the correct coverage. Similar coverage in the profile view would make it absolutely certain that the head is correctly placed. Figure 4(a) and (c) show cases for incorrect placement.

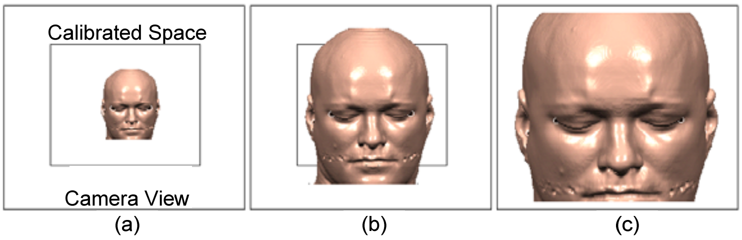


Figure 4: Fitting Image in Calibrated Region[3]

Therefore, for a field of view of the camera sensor corresponding to a particular working distance and camera focal length, vertical and horizontal limits can be chosen for the location of the head during initial camera placement. The user interface for initial positioning of the camera system with respect to the patient has demarcations similar to Figure 5. To help the operator further in obtaining a proper frontal view, the image showing the demarcated area for the frontal view can be marked with a horizontal line on which both eyes should lie. Furthermore, having the patient's ears (viewed frontally) horizontally adjacent to each other and within the region spanned vertically by the eyes and nose ensures proper initial set-up of the camera system.

[3] CT slices have been taken from the patient contributed image repository at http://www.pcir.org

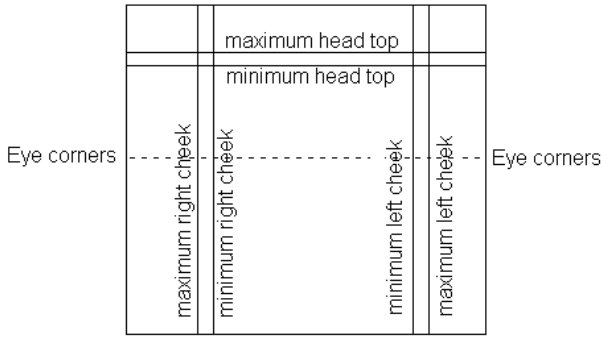


Figure 5: Demarcations for camera set-up

3 Methodology for Preliminary Validation

3.1 Projection View Generation and Processing

Due to the lengthiness of the procedure for undertaking clinical trials and the associated high costs, a common dataset in the two modalities could not be obtained at this point. This has precluded testing of the proposed registration framework on real data. As a preliminary study, a simulation of the registration is presented next for white light images generated from a surface rendered CT model. It is believed that the error obtained in extracting the features replicates the error incurred in practice with real white light images. The 3D surface rendered models were created from CT datasets from various sources such as the US National Library of Medicine (Visible Human Project), the Patient Contributed Image Repository[4], the Association of Electrical and Medical Imaging Equipment Manufacturers[5], and numerous databases available in the public domain containing DICOM compliant CT images[6] as well as from anonymised CT images of patients. The images used are 512 × 512 × 1 voxel(s) with slice thickness ranging from 0.4 to 1.25 mm. A surface rendered model of the DICOM compliant scans was created by constructing an isosurface.

The five views are projected from the CT model at azimuth angles of 0, 45, 90, -45 and -90 degrees corresponding to the frontal, left intermediate, left profile, right intermediate and right profile images respectively. These projected views are marked with control points whose 3D coordinates are known in the CT model and appear on the projected views. This enables calibration of these views. The craniofacial landmarks are not used as part of the calibration. The recovered DLT parameters are used for reconstructing the selected craniofacial landmarks extracted in stereo views. Clearly, only the intermediate and profile views suffice for this purpose and have been used for tests on 15 head models. The maximum RMS error obtained for these simulated tests was 1.04 mm for the left ear corresponding to model 1. The maximum error in the three dimensions was 1.28mm in the x-direction corresponding to model 2 for the left ear.

Due to the same pose of the model in both datasets, the rigid-body transformation between them has a rotation matrix equal to identity and zero for the translation vector. The next calculation performed was to register the two datasets using landmarks from the CT model and the reconstructed space. This served to validate how well the estimated features corresponded and contributed to the overall registration error. Four points were picked over the face and mapped from the CT space onto the reconstructed coordinate system. For the fifteen head models tested, the maximum RMS error incurred from the mapping was 0.68mm with a maximum error of -0.88mm in the z-direction.

3.2 Simulating the Procedure

This section aims at replicating the protocol followed in a normal surgery scenario with the selection of entry and target points by the neurosurgeon (Figure 6). These are done on given CT scans which can then be located on the patient. The registration transformation obtained earlier is used to map these points onto the reconstructed coordinate system. A spherical representation of the trajectory vector provides the length of the trajectory and its orientation in space. These metrics are computed in each coordinate system as a means to assess accuracy with φ being the angle from the z-axis and θ the angle in the x-y plane from the x-axis. For the fifteen head models, the maximum error for φ was 0.25° for which θ equals 0.23° while the maximum value for θ was 0.44° with an associated value of 0.07° for φ.

---

[4] http://pcir.org

[5] ftp://medical.nema.org/MEDICAL/Dicom/Multiframe

[6] http://apps.sourceforge.net/mediawiki/gdcm

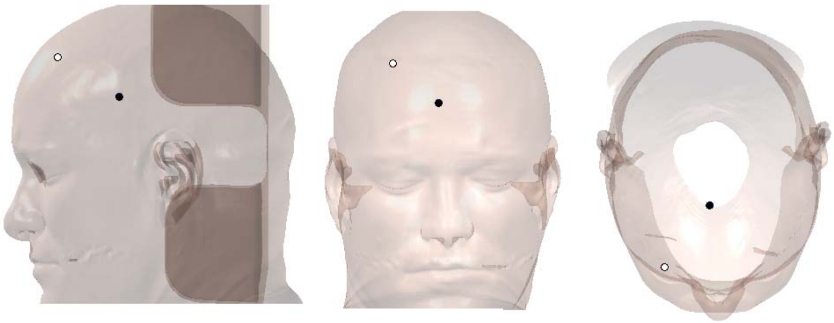


Figure 6: Entry and Target Points[7]

4 Discussion

The methodology employed for simulating the registration is based on projected views of 3D surface rendered CT models. Due to the non-availability of a common dataset in the CT and white light modalities, a simulation approach was adopted to illustrate the registration procedure while providing error estimates. Nevertheless, the use of the selected landmarks as a basis for registration shows that rigid registration can be used to map the chosen entry and target points to the desired accuracy. Registration RMS errors less than 1 mm were obtained for points selected around the face while case studies involving the selection of entry and target points on the different head models showed acceptable angular deviation from the original set trajectory.

5 Conclusion

A registration protocol for preoperative CT to intraoperative white light images has been described. Specifically, the proposed registration has been devised in view of supporting three neurosurgical procedures that are emergency in nature. Simulation of the registration framework shows a maximum RMS error of 0.68mm while the maximum error along the Cartesian axes was -0.88mm for the z-axis. Hence the maximum simulated error is within the 5mm accuracy needed to undertake these procedures. The next step will be to implement such a protocol on real dataset in both modalities. Machine vision tools will also be introduced into the registration framework to help in localising and extracting craniofacial features in the CT and white light modalities thereby reducing subjectivity in the process. The proposed registration method can also be used as part of other registration methods where gross alignment is first needed.

References

1. A.C. Colchester, J. Zhao, K.S. Holton-Tainter et al. , "Development and preliminary evaluation of VISLAN, a surgical planning and guidance system using intra-operative video imaging," *Med. Image Anal.*, vol. 1, pp. 73-90, Mar. 1996.
2. M.J. Clarkson, D. Rueckert, D.L. Hill et al., "Registration of multiple video images to preoperative CT for image-guided surgery," *Proceedings of SPIE*, vol. 3661, pp. 14, 2003.
3. P. Viola and W.M. Wells III, "Alignment by Maximization of Mutual Information," *International Journal of Computer Vision*, vol. 24, pp. 137-154, 1997.
4. M.J. Clarkson, D. Rueckert, D.L.G.Hill et al., "Using photo-consistency to register 2D optical images of the human face to a 3D surface model," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, pp. 1266-1280, 2001.
5. J.B. West, J.M. Fitzpatrick, S.A. Toms et al., "Fiducial point placement and the accuracy of point-based, rigid body registration", *Neurosurgery*, vol. 48, pp. 810-6; discussion 816-7, Apr. 2001.
6. J.M. Fitzpatrick, J.B. West and C.R. Maurer Jr, "Predicting error in rigid-body point-based registration," *IEEE Transactions on Medical Imaging*, vol. 17, pp. 694-702, 1998.
7. F. Brassow and K. Baumann, "Volume of brain ventricles in man determined by computer tomography," *Neuroradiology*, vol. 16, pp. 187-189, 1978.
8. M.R. Bullock, R. Chesnut, J. Ghajar, et al., "Guidelines for the Surgical Management of Traumatic Brain Injury," *Neurosurgery*, vol. 58, pp. S2, 2006.
9. S.T. O'Leary, M.K. Kole, D.A. Hoover et al., "Efficacy of the Ghajar Guide revisited: a prospective study," *J. Neurosurg*., vol. 92, pp. 801-803, May. 2000.
10. D.R. Huyette, B.J. Turnbow, C. Kaufman et al., "Accuracy of the freehand pass technique for ventriculostomy catheter placement: retrospective assessment using computed tomography scans," *J. Neurosurg.*, vol. 108, pp. 88-91, Jan. 2008.
11. A-Nasser Ansari, Mohamed Abdel-Mottaleb, “Automatic facial feature extraction and 3D face modelling using two orthogonal views with application to 3D face recognition”, *Pattern Recognition* 38, 2549 – 2563, 2005
12. LIANG CHEN, C.W. ARMSTRONG, D.D. RAFTOPOULOS, “An investigation on the accuracy of three-dimensional space reconstruction using the direct linear transformation technique”, *Journal of biomechanics*, Vol. 27, No. 4, pp. 493-500, 1994

[7] CT slices have been taken from the patient contributed image repository at http://www.pcir.org