Ian J. Gerard, Jeffery A. Hall, Kelvin Mok, and D. Louis Collins

New Protocol for Skin Landmark Registration in Image-Guided Neurosurgery: Technical Note

Neurosurgery Volume 11 Supplement 3, September 2015, p 376–381

doi: 10.1227/NEU.0000000000000868

New Protocol for Skin Landmark Registration in Image Guided Neurosurgery: Technical Note

ABSTRACT

Background: Newer versions of the commercial Medtronic Stealth Station© allow the use of only 8 landmark pairs for patient-to-image registration as opposed to 9 landmarks in older systems. The choice of which landmark pair to drop in these newer systems can have an effect on the quality of the patient-to-image registration.

Objective: The objective of this work was to investigate four landmark registration protocols based on 8 landmark pairs and compare the resulting registration accuracy to a 9 landmark protocol.

Methods: Four different protocols were tested on both phantoms and patients. Two of the protocols involved using 4 ear landmarks and four facial landmarks and the other two involved using 3 ear landmarks and 5 facial landmarks. Both the fiducial registration error and target registration error was evaluated for each of the different protocols to determine any difference between them and the 9 landmark protocol.

Results: No difference in fiducial registration error was found between any of the 8 landmark protocols and the 9 landmark protocol. A significant decrease ($p < 0.05$) in target registration error was found when using a protocol based on four ear landmarks and four facial landmarks compared to the other protocols based on three ear landmarks.

Conclusion: When using 8 landmarks to perform the patient-to-image registration, the protocol using 4 ear landmarks and 4 facial landmarks greatly outperformed the other 8-landmark protocols and 9-landmark protocol, resulting in the lowest target registration error.

Keywords: image guided neurosurgery, landmark registration, neuronavigation, anatomical landmarks, registration error

INTRODUCTION

Since the introduction of frameless stereotactic neurosurgery[1], neuronavigation is used in many different neurosurgical procedures, notably in tumour resections, as these systems allow for visualization of preoperative images, segmented surfaces and planning information and help reduce trauma through precise location of target surgical areas. During surgery, these image guided neurosurgery (IGNS) systems provide guidance through an environment that tracks both the patient and the surgical tools with a patient-to-image mapping. This allows a surgeon to point to a specific location on a patient and see the corresponding anatomy in the preoperative images. One of the most widespread technique for patient-to-image mapping involves choosing corresponding landmarks on both preoperative images as well as on the patient in the operating room (OR). While the accuracy for this technique has been reported with variable success[2-5], the main advantage is minimal invasiveness to the patient, as opposed to bone implanted markers or stereotactic frames, and the short amount of time needed to create the mapping. Since this is the only registration procedure generally done on a neuronavigation system it is important to minimize registration errors to maintain the highest level of accuracy as long as possible throughout the intervention. The current landmark registration protocol in frameless stereotactic IGNS procedures using the Medtronic Stealth

Station© involves choosing 9 corresponding landmark pairs on both a patient's preoperative images and their anatomy in the OR[2]. The landmarks include: i) bridge of the nose (BN), ii) right medial canthus (rMC), iii) right lateral canthus (rLC), iv) right tragus valley (rTV), v) right tragus (rT), vi) left medial canthus (lMC), vii) left lateral canthus (lLC), viii) left tragus valley (lTV), ix) left tragus (lT) (see Fig. 1). The use of anatomical landmarks instead of skin implanted markers was put into practice in order to improve patient comfort and to save imaging time so that diagnostic magnetic resonance imaging (MRI) could be used for preoperative registration. However, due to different factors related to MRI image acquisition and the technician or surgeon choosing the landmarks, the accuracy of this registration technique has been reported to vary between 1 mm and 7 mm[2-5] depending on the neuronavigation system used and the metrics used to measure misregistration. In newer versions of the commercial Medtronic software, the point-based landmark registration protocol now permits the use of only eight landmarks. With this constraint, one landmark pair must be removed from the current protocol for clinical cases. From initial visual inspection, the result of this protocol change appears to have a higher rate of misresgistration along the anterior-posterior (A-P) axis for different 8-landmark landmark protocols when compared to the 9-landmark protocol.

Figure 1 goes here

In this study we evaluate the effect of fiducial registration error (FRE) and target registration error (TRE) on patient-to-image mapping in IGNS through the use of 4 different 8-landmark pair matching protocols as well as the effect of repeated landmark use on improving patient-to-image registration quality in a 9-landmark protocol.

METHODS

We investigate the effect on registration quality of 4 different protocols (see Table 1). The protocols are separated based on which landmark, either the tragus or medial canthus, is dropped from which side, left or right.
Protocol T-L and T-R exclude the left tragus and right tragus respectively. Protocol C-L and C-R exclude the left medial canthus and right medial canthus respectively.

Phantom experiments: Phantom experiments were initially performed on a plastic anthropomorphic head phantom and a corresponding CT scan to determine if testing should be done with real patient data. Each of the four protocols was tested five times. Each transformation was evaluated by measuring the Euclidian distance between two well defined, easily located target point pairs on the phantom and the corresponding CT scans. The target point pairs corresponded to points on the phantom where there was a clear indentation on both the left and right sides (Figure 1 iii)). As shown below, this experiment yielded significant results following a Wilcoxon sign test, and we proceeded with patient experiments.

Patient experiments: Registration data was collected on 10 different patients undergoing tumour resection surgery at the Montreal Neurological Hospital (MNH). For 6 of the patients in this series, landmarks were chosen five times by the same technician on the preoperative images. Each time, corresponding landmarks were chosen on the patient by the neurosurgeon in the operating room. For the other four patients, the landmarks were recorded only once. This resulted in a total of 34 patient-image landmark sets for evaluation. All data were collected and analyzed on a prototype neuronavigation system, IBIS Neuronav[3]. Each set of patient-image landmarks was used to compute a patient-to-image transformation that was used to estimate FRE and TRE.

Registration Errors

Recent literature suggests three main error metrics when analyzing the accuracy of point-based registration methods: i) fiducial localization error (FLE) – the error in locating fiducial points, ii) fiducial registration error (FRE) – the distance between corresponding fiducial points after registration, and iii) target registration error (TRE) – the distance between corresponding points other than the fiducial points after registration[6-8] . For our analysis we used the FRE and the TRE as registration quality metrics.

Any nonzero misalignment between a transformed fiducial point and its corresponding fiducial point in the target space is an FRE. It is expressed mathematically as:

$$FRE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\|T^{W-I}(f_i^I) - f_i^W\right\|^2} \tag{1}$$

where N is the number of points used, $\boldsymbol{f_i^W}$ is the fiducial point in world coordinates $\boldsymbol{f_i^I}$ is the fiducial in image coordinates and $\boldsymbol{T^{W-I}}$ is a rigid transformation matrix calculated to minimize the RMS distance between landmark pairs. The term $\boldsymbol{T^{W-I}}(\boldsymbol{f_i^I}) - \boldsymbol{f_i^W}$ can be interpreted as the FRE for an individual landmark pair. Previous work[9] has shown that with sufficiently well chosen landmark pairs this registration metric will plateau to a steady value, motivating the identification of the same landmarks multiple times on the same patient in order to reduce FREs. Any nonzero misalignment between the location of a landmark not used as a fiducial and its corresponding transformed point on the preoperative images is a TRE. It is expressed mathematically as:

$$TRE(t^I) = \|T^{W-I}(t^I) - t^W\| \tag{2}$$

Where t^I and t^W correspond to the target point positions in the image coordinates and world coordinates respectively.

To assess the TRE for each estimated transformation, the single landmark point not used as a fiducial in the registration procedure was used as a target point to obtain an estimation of the TRE. The TRE for the protocol using all 9 landmark pairs was estimated with the following relationship[10]:

$$\langle TRE^2(t_i^I)\rangle \approx \langle FLE^2\rangle\left(\frac{1}{N} + \frac{1}{3}\sum_{i=1}^{3}\frac{d_i^2}{f_i^2}\right) \tag{3}$$

where the angled brackets represent the expected value of the functions, d_i^2 is the distance of the target from principal axis *i* and f_i^2 is the RMS distance of the fiducials from the same axis . The $\langle FLE^2\rangle$ was determined by adding in quadrature the error contributions from tracking as previously determined by Gerard 2014[9] as well as the RMS of the pointer after calibration during the procedure done on the day of the operation.

Analysis of patient data was performed in two steps. Initially, a two-tailed student's t-test was performed between results of protocols with the same landmark dropped on opposite sides. If no significant difference was found between them then their results were treated as a single group for a single factor ANOVA. The FRE for the 6 patients with landmark pairs chosen 5 times were evaluated with a one tailed student's t-test.

To evaluate whether an improved registration quality based on TRE could be attributed to the modified protocol rather than the choice of target (landmark not being used) variability the bias between landmarks was investigated. First, the standard deviation of the landmarks chosen by the surgeon was calculated to observe if there was any difference in the surgeon's ability to consistently choose different landmarks. In addition, the FRE of the 4 landmark points to be used as target points in the modified protocols will be compared when used in the 9 landmark protocols to ensure that no difference can be attributed to the quality of the chosen point.

Ethics

The research ethics board of the Montreal Neurological Institute and Hospital approved this study and all patients gave informed consent.

RESULTS

The results of the phantom test demonstrated a significantly smaller TRE for the C-L and C-R protocols ($p<0.05$, Wilcoxon sign test) relative to the T-L and T-R protocols.

For the patient data, analysis of the FRE between protocols showed there was no significant difference ($p = 0.989$) between the different protocols (Table 2, left side). Analysis of the TRE between protocols showed there was a significant difference ($p = 5.83$ e-8) between different protocols with the protocols involving dropping a medial canthus landmark (i.e., C-L and C-R) having a significantly lower TRE (Table 2, right side).

The results of using repeated recordings of the same landmark pairs are summarized in Table 3. It can be seen that in 5 of the 6 cases, this strategy significantly decreases the mean FRE, while there is no significant change in the remaining case. When evaluating the TRE, there was no improvement on multiple landmarks when using the current protocol. In protocols T-(L/R) and C-(L/R) some significant improvement on TRE was observed for several cases.

The standard deviation of the surgeon's ability to consistently pick the same anatomical landmark for each patient is summarized in Table 4. There is no specific landmark that varies significantly between the six patients, however in some cases there is higher variability on some anatomical landmarks for a certain patient. Comparing the FRE between landmarks used as target points in the modified protocols when used as fiducial points in the 9 landmark protocol is summarized in Table 5. The comparison yielded no significant difference (ANOVA $p<0.05$) between the quality of different chosen landmarks.

DISCUSSION AND CONCLUSION

In this report we investigate four new protocols for landmark selection for patient-to-image registration in IGNS when constrained to 8 points for the landmark registration protocol. Analysis was performed in terms of both the FRE and TRE for the new protocols in comparison to the current protocol in practice. No significant difference was found in the FRE between the current protocol in practice and any of the new protocols. A significant difference was observed between the TRE for the current protocol and the observed protocols. A Tukey post-hoc test revealed that the protocols involving dropping a medial canthus landmark were different than the current protocol and the protocol involving removing a tragus landmark. This is confirmed in part with phantom tests. Reasons for these protocols showing improvement may stem from the increased number of points on the lateral aspects of the patient's head and also the fact that less points near the front of the face creates less bias towards a stronger fit along this area with greater potential for superior-inferior error in the bridge of the nose, as compared to the back of the head near the ears. This creates less rotational bias in the anterior-posterior direction resulting in smaller registration errors. While not investigated here, previous work[9] characterizing registration errors based on landmark choice has suggested that fiducial points should be selected further away from areas of interest to avoid lever effects that increase the TRE closer to these fiducial points. This would suggest that the side on which the landmark was removed should be on the opposite side of the tumour to be operated.

When comparing the mean FRE and TRE for the six patient registrations using the single measurement strategy and the repeated measurement strategy, there was a significant decrease in the mean FRE for five of the six cases. The reasons that there was no improvement in one case could be related to the quality of the preoperative images. For instance, in some scans the patients wear ear plugs due to the noise of the MRI machine which can cause the anatomy of the ears (tragus and tragus valley) to be bent inwards and not be accurately represented on the preoperative images compared to the true anatomy. When evaluating the TRE, there was no improvement when choosing landmarks multiple times for the current protocol. This is probably related to the fact that the TRE is estimated using the fiducial configuration. Since the configuration does not change between single and multiple landmark strategies the estimated TRE does not change in a significant manner. Significant changes were observed in the TRE in some instances for Protocol T-(L/R) however it never resulted in a TRE that was better than the results of protocol C-(L/R). The TRE in Protocol C-(L/R) improved in only one case. This may be related to the fact that the TRE was already quite low using this protocol so the additional landmarks added no benefit to reducing the registration error.

To ensure improvement in registration quality could be attributed to the new protocol rather than choice of target point the quality of the points used as targets was evaluated for bias. The neurosurgeon's ability to consistently pick the same anatomical landmark was evaluated in order to observe if there were any specific landmarks that showed high variability. There was no particular landmark of the 9 tested that varied consistently between each of the patients, however, in some patients certain landmarks varied more than others. In addition to this the FRE of the target points in the modified protocols were compared when used as fiducials in the 9 landmark protocol to ensure that they had similar individual FREs thus strengthening the improved registration being attributed to the protocol rather than the choice of target point.

In conclusion, the work shown here suggests an improved protocol for determining the patient-to-image mapping for IGNS interventions using 8 landmarks consisting of the BN, rLC, rT, rTV, lLC, lT, lTV, and one of either the *rMC or the lMC*.

REFERENCES

1 Roberts DW, Strohbehn JW, Hatch JF, Murray W, Kettenberger H. A frameless stereotaxic integration of computerized tomographic imaging and the operatin microscope. *J Neurosurg* (1986) 65:545-549

2 Wolfsberger S, Rössler K, Regatschnig R Ungersböck K. Anatomical landmarks for image registration in frameless stereotactic neuronavigation. *Neurosurg Rev* (2002) 25:68-72.

3 Mercier L, Del Maestro RF, Petrecca K, Kochanowska A, Drouin S, Yan CXB Janke AL, Chen SJS Collins DL. New prototype neuronavigation system based on preoperative imaging and intraoperative freehand ultrasound: System description and validation. *Int J CARS* (2011) 6(4):507-522.

4 Sipos EP, Tebo SA, Zinreich SJ, Long DM, Brem H. In vivo accuracy testing and clinical experience with the ISG Viewing Wand. *Neurosurgery* (1996) 39(1):194-202.

5 Pfisterer WK, Papadopoulus WKS, Drumm DA, Smith K, Preul MC. Fiducial verus nonfiducial neuronavigation registration assessment and considerations of accuracy. *Neurosurgery* (2008) 62(3):201-207.

6 Moghari MH, Abolmaesumi P. Understanding the Effect of Bias in Fiducial Localization Error on Point-Based Rigid-Body Registration. *IEEE T Med Imaging* (2010) 29(10): 1730 - 1738.

7 Labadie RF, Davis BM, Fitzpatrick JM. Image-guided surgery: what is the accuracy. *Current opinion in otolaryngology & Head and Neck Surgery* (2005) 13(1): 27 - 31.

8 Liu W, Ding H, Han H, Xue Q, Sun Z, Wang G. The Study of Fiducial Localization Error of Image in Point-based Registration. *31st Annual International Conference of the IEEE EMBS* (2005). pp:5088 - 5091.

9 Gerard IJ, Collins DL. An analysis of tracking error in image-guided neurosurgery. *Int J CARS* (2014). DOI: 10.1007/s11548-014-1145-2

10 Fitzpatrick JM, West JB. The Distribution of Target Registration Error in Rigid-Body Point-Based Registration. *IEEE Transactions on Medical Imaging* (2001) 20(9):917 – 927.

FIGURE LEGEND

Figure 1: A) Location of skin landmarks on anthropomorphic head phantom. B) Registration target points.