

A Framework for Determining Component and Overall Accuracy for Computer Assisted Surgery Systems

A.B. Mor¹, J.E. Moody¹, D. Davidson¹, R.S. Labarca², B. Jaramaz¹, and A.M. Digioia III¹

¹Institute for Computer Assisted Orthopaedic Surgery, The Western Pennsylvania Hospital

²Center for Medical Robotics and Computer Assisted Surgery, Carnegie Mellon University

Pittsburgh, PA, USA

abm@icaos.org

Abstract. Computer assisted surgery (CAS) systems have generated a great deal of interest in the research and medical communities due to their promise of improved results with less invasive access to the surgical site. While some systems have been checked against a minimum acceptable standard set by the system's developers, no overall methodology is in place to insure that a system is thoroughly characterized. A framework to assess the accuracy of CAS systems will provide developers and researchers with a common starting point with which to test their systems, and will provide surgeons with a consistent basis for comparing different systems. We describe this framework and its application to the HipNav system.

1 Background

Multiple groups have explicitly modeled outcome errors for their CAS systems, looking at the final surgical or laboratory result in comparison to a predetermined ideal result. Other groups have examined individual subsystems for their component accuracies. [2] described accuracy validation for image guided orthopaedic surgery, identified sources of error for prototypical CT-based systems, and suggested techniques to minimize the effects of some sources. [3] described a graph-based method to model error propagation through a system.

2 Accuracy Assessment Framework Design

The end-to-end accuracy of a CAS system determines the distribution of surgical results that a surgeon might expect to achieve, and clearly needs to be thoroughly tested. But, while the end-to-end accuracy is of direct importance to a system's overall usefulness and safety, it is vital to understand the contributions from individual system components. After these different subsystem components are identified, a flow chart of the CAS procedure should be generated, and possible contributions to the system error at each step should be identified. Every effort should be made to quantify the error terms at each step, in addition to the end-to-end error. A rigorous statistical analysis is then performed on the data, and the end-to-end and component accuracies are analyzed and compared to a minimum level of accuracy relevant to the surgical application.

To help in the characterization of the CAS system's accuracy, it is valuable to follow a typical case profile with a model of the operative anatomy. A well-constructed and characterized phantom model will aid in determining application-specific error measurements, which can be different than general error terms using an artificial construct. The satisfactory phantom will contain fiducials, whose positions are then accurately determined with a coordinate measuring machine (CMM), that are easily locatable at all stages of a practice procedure. By generating this accurately measured ground truth model, we can then characterize all error terms that are based on patient anatomy.

3 Results

The described framework was applied to the HipNav system [1]. The accuracy threshold for HipNav was set to be 2° of rotation and 1.5mm of translation. The rotational component is defined by the error between the actual placed orientation of the acetabular cup and the HipNav reported orientation. The translational component is set by the desired registration accuracy necessary to achieve the rotational goal. Two plastic phantom models of the pelvis, with spherical fiducials and acetabular cup implants installed on both sides, were constructed to facilitate the accuracy tests. The locations of the fiducials and the orientations of the mounted cups were then measured with a CMM.

We tested the accuracy of the following subsystems: CT scanning, tracking system and tracking bodies, probe noise, probe calibration, registration errors, and end-to-end accuracy. The CT scan had an RMS error of 0.34mm, a standard deviation of 0.16mm, and a maximum point-to-point error of 0.76mm. The tracking system, a passive Polaris system from NDI, was examined with respect to positional noise and accuracy with four different tracker bodies, in two different orientations, varied from 0.02mm to 0.08mm standard deviation (max 1.72mm), while the angular noise varied from 0.008° to 0.12° standard deviation (max 0.60°). The trackers were also characterized for noise when measuring a probe tip location. The positional noise varied from 0.076mm to 0.41mm (max 1.40mm). Accuracy of the trackers, when performing a point-to-point registration across a grid 20" by 40" in size, was approximately 0.2mm mean RMS (0.08mm to 0.12mm standard deviation) for all 4 trackers tested. Maximum errors varied from 0.45mm to 0.64mm. To determine the accuracy of point probe calibration, we performed repeated pivot calibrations. The RMS error in the reported length of the probe was 0.59mm, with a standard deviation of 0.60mm and a maximum of 1.18mm. The registration algorithm was tested on surface models from 33 patients, with an average rotational error of 1.23° and an average translational error of 1.07mm.

In addition to the above tests, we performed end-to-end tests with the four pelvic models. Average orientation errors ranged from 0.78° to 1.2° . Standard deviation varied from 0.16° to 0.36° . The maximum error measured was 1.83° . After examining the data, though, the error for each of the four pelvic models was clustered around a different mean, suggesting the presence of offsets between the CMM reference frame and the coordinate frame used by HipNav. After removing this bias, the maximum error was 0.94° .

4 Conclusions

We have shown a method for examining and characterizing CAS systems for their potential error, showing both error distributions and maximum errors. Both individual component errors and end-to-end errors were examined. Both types of analysis are important for the surgeon, as some surgeons may only use a portion of a CAS system, while others may use the complete package. This analysis, although not yet complete, demonstrates the framework necessary for completing a total characterization of a CAS system.

- [1] DiGioia AM, Jaramaz B, Blackwell M, Simon DA, Morgan F, Moody JE, Nikou C, Colgan BD, Aston CA, Labarca RS, Kischell E, Kanade T. "The Otto Aufranc Award. Image guided navigation system to measure intraoperatively acetabular implant alignment." *Clin Orthop*. 1998 Oct;(355):8–22.
- [2] Simon D, O'Toole RV, Blackwell M, Morgan F, DiGioia AM 3rd, and Kanade T. "Accuracy Validation In Image-Guided Orthopaedic Surgery," *Proceedings of the MRCAS*, 1995, pp. 185–192.
- [3] Viant WJ. "The development of an evaluation framework for the quantitative assessment of computer-assisted surgery and augmented reality accuracy performance." *Stud Health Technol Inform*. 2001, pp. 81:534–40.