Using an Android Application to Assess Registration Strategies in Open Hepatic Procedures: A Planning and Simulation Tool

Derek J. Doss[1], Jon S. Heiselman[1], Jarrod A. Collins[1], Jared A. Weis[1], Logan W. Clements[1], Sunil K. Geevarghese[5], and Michael I. Miga[1,2,3,4]

[1]Vanderbilt University, Department of Biomedical Engineering, Nashville, TN USA
[2]Vanderbilt University Medical Center, Department of Radiology and Radiological Sciences, Nashville, TN USA
[3]Vanderbilt University Medical Center, Department of Neurological Surgery, Nashville, TN USA
[4]Vanderbilt Institute in Surgery and Engineering, Nashville, TN USA
[5]Vanderbilt University Medical Center, Nashville, TN USA

Abstract

Sparse surface digitization with an optically tracked stylus for use in an organ surface-based image-to-physical registration is an established approach for image-guided open liver surgery procedures. However, variability in sparse data collections during open hepatic procedures can produce disparity in registration alignments. In part, this variability arises from inconsistencies with the patterns and fidelity of collected intraoperative data. The liver lacks distinct landmarks and experiences considerable soft tissue deformation. Furthermore, data coverage of the organ is often incomplete or unevenly distributed. While more robust feature-based registration methodologies have been developed for image-guided liver surgery, it is still unclear how variation in sparse intraoperative data affects registration. In this work, we have developed an application to allow surgeons to study the performance of surface digitization patterns on registration. Given the intrinsic nature of soft-tissue, we incorporate realistic organ deformation when assessing fidelity of a rigid registration methodology. We report the construction of our application and preliminary registration results using four participants. Our preliminary results indicate that registration quality improves as users acquire more experience selecting patterns of sparse intraoperative surface data.

Keywords: hepatic, simulation, training, surgery, android application

1. Introduction

Liver cancer incidence rates have increased for both men and women from 2003 to 2012 [1]. Surgical resection represents a potentially curative method of treatment and an increased frequency in hepatic resections has improved the five-year survival status for patients with metastatic colorectal carcinoma [2]. However, open hepatic procedures are complicated operations due to the complexity of the organ vasculature. The biliary vessels, portal vein, and hepatic artery branches require careful localization to guide resection planes and reduce the risk of complication. The resection plane ultimately influences the future liver remnant, which should be preserved for adequate liver function. To combat these complexities, image-guided liver surgery (IGLS) methods register preoperative image volumes or segmented structures to intraoperative physical space, providing information for surgical navigation. However, the process of registration is compromised by

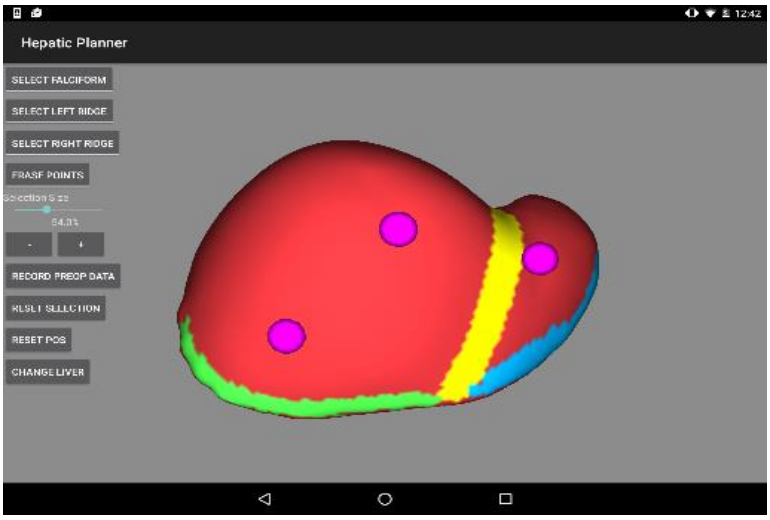


Figure 1: Preoperative model with the falciform ligament (yellow), right inferior ridge (blue), and left inferior ridge (green) selected.

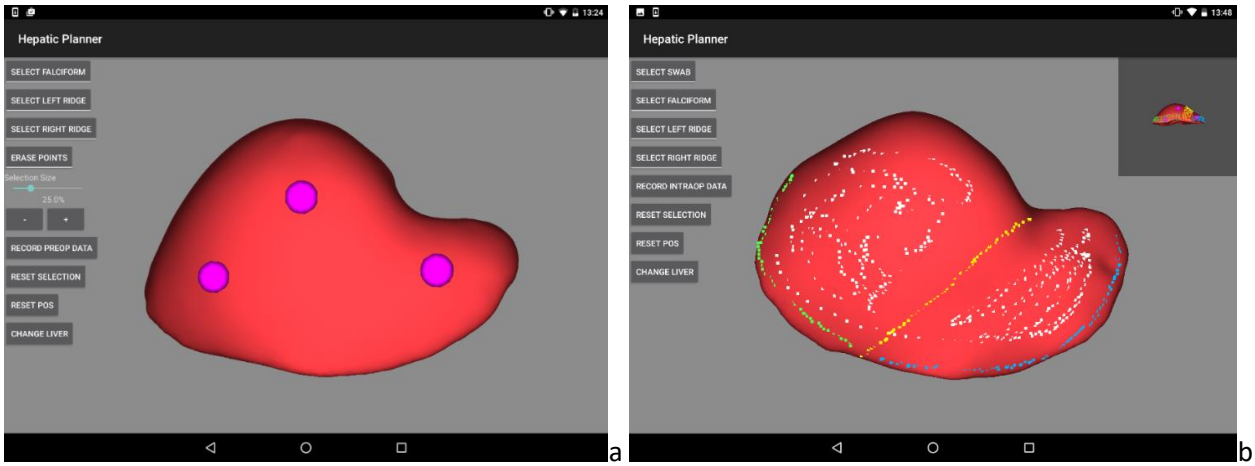
intraoperative organ deformation and the lack of established collection patterns for intraoperative data used to drive the registration.

In previous work, rigid registration methods complemented by deformation-correcting model-based methods have been developed within the context of open hepatic procedures [3]. The alignment methodology begins with a preoperative marking of salient features on the segmented liver surface (i.e. falciform ligament, round ligament, left inferior ridge, and right inferior ridge), as seen in Figure 1. Homologous features in addition to more widely distributed sparse surface data are marked intraoperatively using an optically tracked stylus. Once complete, a salient feature weighted surface registration is performed [4, 5]. Following this initial alignment, an iterative method is employed, minimizing the error between the collected surface data and the preoperative organ model surface. While conceptually easy to understand, in our experience, the marking of salient features both in the planning and intraoperative phases can vary. When patient anatomy can be confounding, technicians may mark preoperative salient features differently. In addition, surgeons may mark salient features to different extents and swab the organ with the tracked stylus in different patterns. Ultimately, these variations can influence the alignment of registered surgical targets.

In order to reduce image-to-physical registration error, an Android tablet application called *Hepatic Planner* was developed. *Hepatic Planner* simulates the process of using a clinically approved, rigid registration IGLS system. In addition, it also allows visualization of each registration. *Hepatic Planner* provides a straightforward method of training surgeons in a rigid registration system before entering the OR. This training has the potential to shorten IGLS surgery times and may eventually lead to more useful navigation systems.

Methods

1.1 Development of Hepatic Planner



a b

Figure 2: (a) The user interface of Hepatic Planner in the preoperative selection stage with targets displayed (purple) and (b) in the intraoperative selection stage with salient features and intraoperative surface data shown.

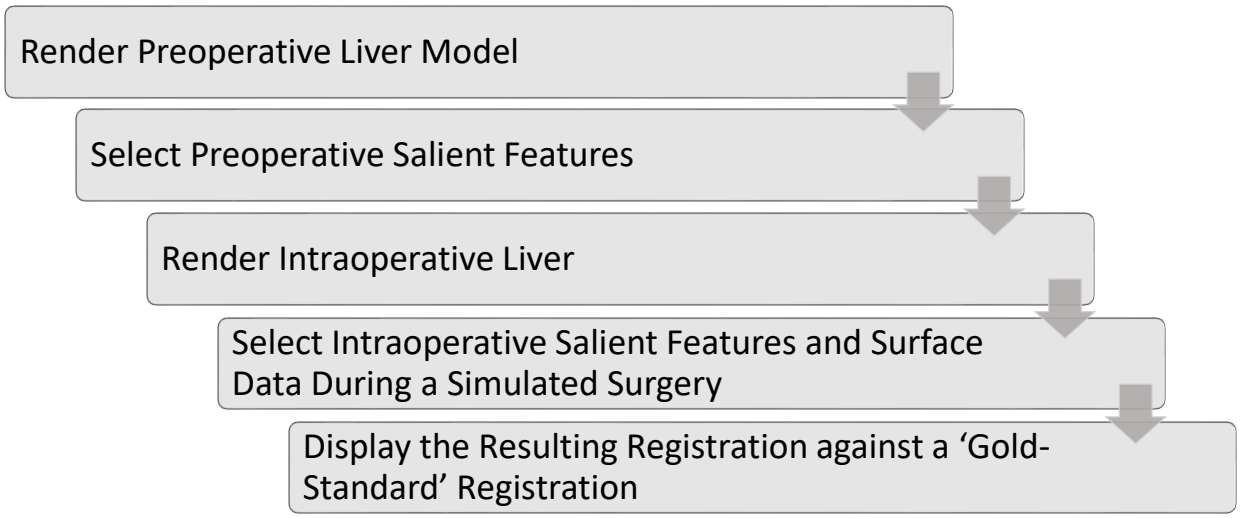


Figure 3: The workflow of *Hepatic Planner*.

Figure 2 shows the *Hepatic Planner* tablet application and Figure 3 shows its overarching workflow. The application was written in Java (JDK 1.8.91) using Android Studio 2.1. The application was built to target Android 6.0 API Level 23. The rendering of 3D models in the application was based on the openGL-ES 2.0 API. The liver files were accurately segmented before the 3D rendering was created and consisted of vertex coordinates, vertex unit normal, and triangular faces. The application was developed and tested on a Nexus 9 tablet with a Dual-core 2.3 GHz processor and 2GB of memory.

The typical segmented liver model in the collected data from our clinical collaborator consisted of 30,000 vertices. The 3D geometry was rendered onto the 2D screen of the Android device using a projection matrix. Due to this, selection could not be completed simply through touch events. To select points through direct

interaction with the 2D projected liver, a k-d tree was used based on the screen position that the touch event was detected at. A tolerance was employed to allow for larger patch selection for preoperative salient feature region designation.

Simulation of intraoperative data collection was accomplished through the k-d tree; however, only one point was selected per touch event and was limited to the sections of the organ typically accessible by optically tracked styluses used in IGLS. A periodic sinusoidal noise pattern (D) was added to simulate the period depression into and elevation off of the surface as has been seen clinically and is reported in Eq. 1. The difference between the previous integer of the start of intraoperative data collection and the time since data collection started is represented by t. A random floating point number ranging from 0.0-1.0 is represented by α. The unit vector pointing away from the center of the model for a point n is represented by P(n). The maximum displacement of noise was set to 2.0 mm.

$$D = (\sin(2\pi t) + \sin(2\pi\alpha))\vec{P}(n) \quad (1)$$

Three targets were rendered along with the preoperative liver model to provide a more realistic application for *Hepatic Planner*. The targets were displayed during preoperative salient feature designation to provide a reference for the user when collecting preoperative data. The preoperative model along with the preoperative data collected and the three targets were presented during intraoperative selection as can be seen in the top right of Figure 2a.

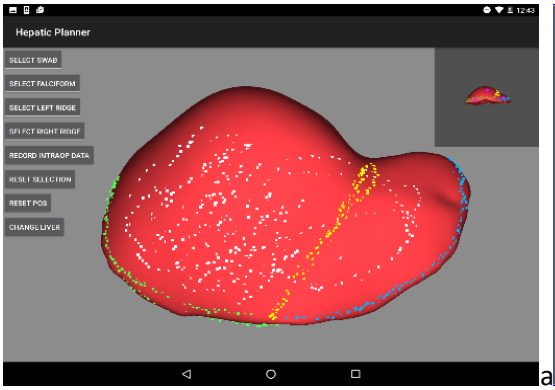
2.2 Data Collection and Analysis

Computed tomography images from a mock liver phantom were used for this investigation. The phantom was created previously using water, 7% by volume Polyvinyl Alcohol, and 10% by volume glycerin. Stiffness was developed though a 12-hour freeze thaw cycle.

The images of the phantom liver consisted of an undeformed preoperative state and a deformed intraoperative state. These images were used to create preoperative and intraoperative 3D models that could be rendered into the application. The preoperative model was used to preoperatively select three salient feature regions (i.e. falciform ligament, left inferior ridge, and right inferior ridge). The intraoperative model was used to simulate intraoperative sparse surface data collection. Each participant first designated corresponding salient features on the intraoperative model and then collected a comprehensive representation of the anterior organ surface.

Clinically collected intraoperative organ surface data was used to determine that 1000 sparse data points are typically acquired during a registration. To limit the amount of data collected during the intraoperative swabbing simulation and standardize the amount of points collected across trials, a notification was added to alert the user when the number of collected points reached 1000.

Accuracy of each registration was determined by using a “gold standard” registration. This gold standard represents a full-organ iterative closest point surface based registration between preoperative and intraoperative states (registering full CT-segmented intraoperative and preoperative surface models). Each sparse data image-to-physical registration determined from the planner was then compared to the gold standard with regard to surface error, calculated as the distance from each point on the trial registration to



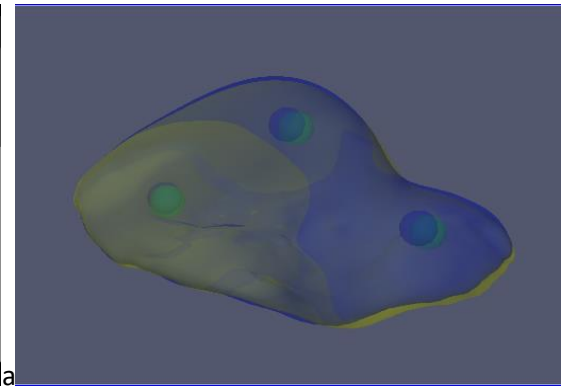
a

b

Figure 4: (a) Intraoperative swabbed data selection. (b) Results of the registration that were displayed to the user.

the gold standard on a closest point basis. Each trial was also compared to the preoperative model via the true surface error calculated between the trial registration of the intraoperative model and the preoperative model. This represented the true surface registration error.

Three targets were created to measure the registration error between each trial and the gold standard. The targets were three spheres that were placed in the gold standard in three different sections of the liver: the left lobe (LL), the right lobe (RL), and centrally (Cent). The tumors were then placed in the intraoperative model by inverting the transformation from the gold standard registration. Therefore, they were only subjected to rigid misalignment and not deformation. The distance between the centroid of each target between the gold standard and each trial registration was computed as target error.

To investigate the viability of *Hepatic Planner* as a training tool, participants were given a goal, instructed to designate preoperative salient features, then were tasked with performing five registration trials. The goals consisted of minimizing error for the target in the right lobe, then minimizing error for the target located in the left lobe, and finally minimizing error for the target located centrally. Between each trial, the registration results were displayed to the participant with the gold standard as a reference. Each participant designated preoperative salient feature regions separately for each goal in the app then proceeded to select the intraoperative salient features and swab data in the app for five trials (Figure 4a). After each trial, the participants were asked to analyze and rate the registration results (e.g. Figure 4b) on a scale from 1-10, with 10 being a perfect registration and 1 being the worst registration possible and determine whether they could be clinically usable. After the five trials for each goal, the participants selected the preoperative salient feature regions for the next goal.

Four participants were used with different backgrounds. Participant 1 was a lower-level graduate student with some understanding of the weighted Iterative Closest Point (wICP) rigid registration system. Participant 2 was a high-level graduate student with knowledge of the inner workings of the rigid registration system. Participant 3 was a research assistant professor with knowledge of the inner workings of the wICP rigid registration system and years of experience working with it. Participant 4 was a surgeon with clinical experience using the wICP rigid registration system.

2. Results

Table 1 shows the mean surface error for each goal given to each participant. The mean surface error is given in terms of both the preoperative liver model and the ideal, gold standard registration. “Average Ideal Surface Error” is the average closest point distance of the produced registrations between all the trials in the provided goal and the gold standard registration. “Average Preop Surface Error” is the average closest point distance magnitudes between the preoperative organ model and each trial registration for all registration trials within the specified goal.

Table 2 shows the mean target error between the trial registration and the Gold-Standard targets. “Avg LL Target Error”, “Avg RL Target Error”, and “Avg Cent Target Error” are the mean errors for the target located in the left lobe, right lobe, and centrally, respectively.

Figures 5 and 6 show the mean ideal surface error (comparison to the gold-standard registration) and the mean preoperative surface error (comparison of preoperative and intraoperative surfaces after registration), respectively, averaged for all trials in a goal for each participant.

Figures 7 and 8 show the TRE and the surface error, respectively, for the first trial in each given goal and the best trial in each given goal.

Figure 9 is a colormap representing the magnitudes of the closest point distance between each trial registration and the gold standard registration. Figure 10 is a colormap of the magnitudes of the closest point distance between each trial registration and the preoperative liver model. The mean of all trials for each goal was taken and used to create the colormaps. The scale for the colormap ranges from 0–12 mm.

Table 1: Average surface errors for each goal and participant. All values are measured in mm.

	Participant 1		**Participant 2**		**Participant 3**		**Participant 4**	
Goal	**Average Ideal Surface Error**	**Average Preop Surface Error**	**Average Ideal Surface Error**	**Average Preop Sufrace Error**	**Average Ideal Surface Error**	**Average Preop Sufrace Error**	**Average Ideal Surface Error**	**Average Preop Sufrace Error**
RL	4.50±3.51	4.25±2.98	5.28±3.80	4.13±3.15	4.32±2.84	4.52±2.92	5.41±4.47	8.70±6.45
LL	5.74±4.32	4.33±3.63	5.56±3.64	4.28±3.12	6.12±4.43	4.72±3.58	3.84±3.26	6.26±4.73
Cent	3.46±2.22	3.68±2.44	5.23±3.19	4.04±2.79	4.94±3.57	4.10±2.98	3.05±2.11	4.85±3.41

Table 2: Average target errors for each goal and participant. All values are measured in mm.

	Participant 1			**Participant 2**			**Participant 3**			**Participant 4**		
Goal	**Avg LL Target Error**	**Avg RL Target Error**	**Avg Cent Target Error**	**Avg LL Target Error**	**Avg RL Target Error**	**Avg Cent Target Error**	**Avg LL Target Error**	**Avg RL Target Error**	**Avg Cent Target Error**	**Avg LL Target Error**	**Avg RL Target Error**	**Avg Cent Target Error**
RL	9.07 ±3.33	7.97 ±3.19	8.91 ±4.82	9.98 ±2.54	10.81 ±2.70	9.86 ±3.12	16.34 ±4.23	10.45 ±3.76	12.02 ±3.97	6.16 ±1.70	11.99 ±5.11	11.06 ±4.77
LL	10.83 ±2.14	10.91 ±3.64	11.48 ±4.24	11.69 ±2.19	13.32 ±2.40	9.94 ±1.52	8.05 ±2.50	10.67 ±1.47	12.62 ±2.74	11.91 ±1.65	12.13 ±2.72	10.28 ±3.45
Cent	5.51 ±2.52	4.97 ±3.03	7.95 ±1.89	13.37 ±2.29	9.35 ±2.98	11.58 ±1.20	10.21 ±2.81	7.68 ±2.10	10.71 ±4.04	9.69 ±1.21	9.89 ±1.68	7.20 ±1.57

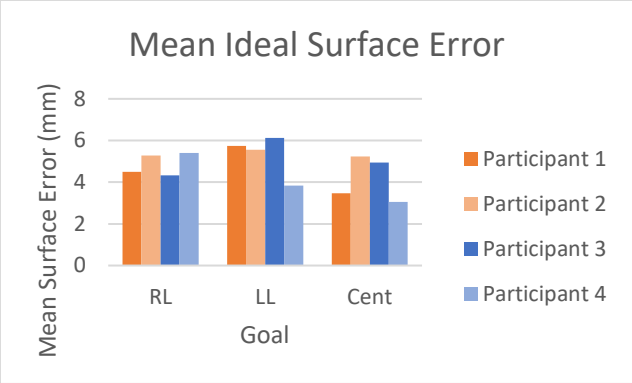


Figure 5: Mean ideal surface error of all trials in each goal for each participant. Compares user registration result to gold-standard result.

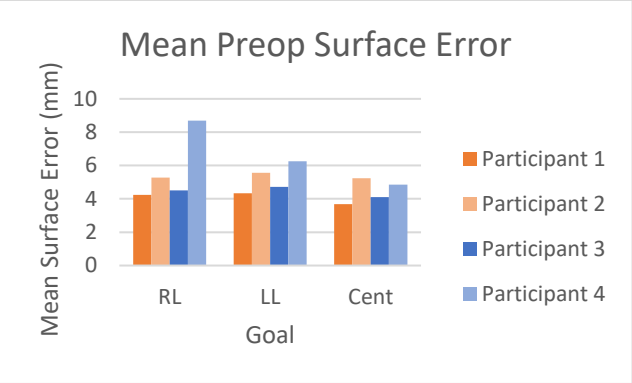


Figure 6: Mean preoperative surface error of all trials in each goal for each participant. Reports error between preop and intraop surfaces after registration.

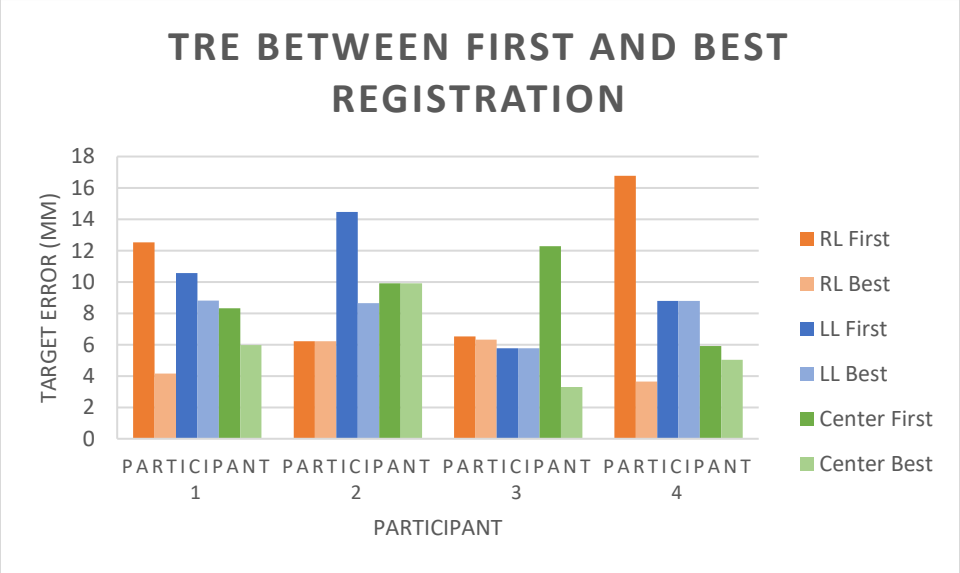


Figure 7: Comparison between the target error for the first trial for each goal and the best target error produced by each participant.

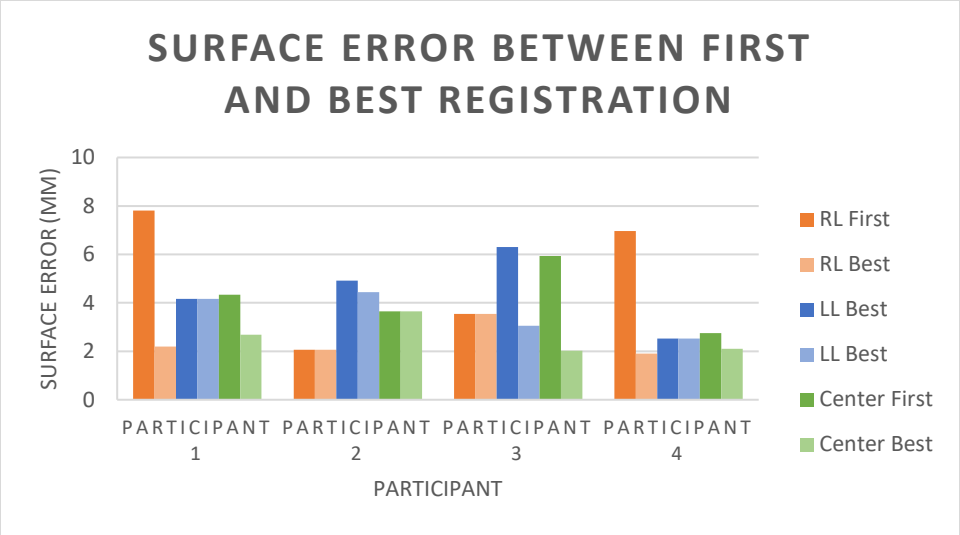


Figure 8: Comparison between the surface error for the first trial for each goal and the best surface error produced by each participant.

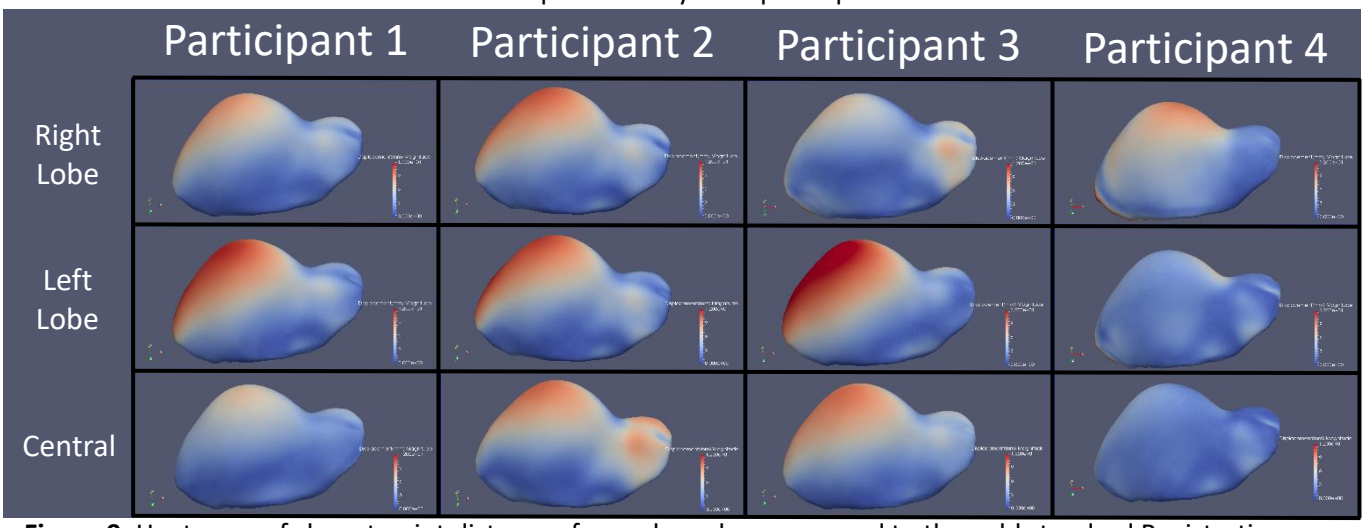


Figure 9: Heatmaps of closest point distances for each goal as compared to the gold standard Registration between each participant. The magnitude of the color scale ranges from 0–12 mm.

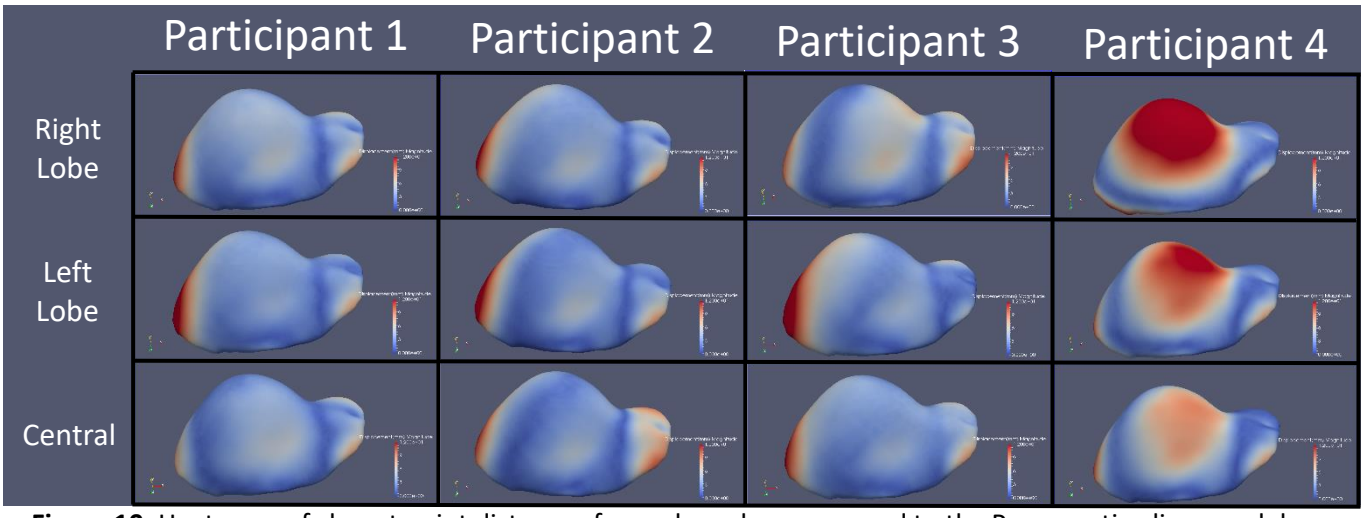


Figure 10: Heatmaps of closest point distances for each goal as compared to the Preoperative liver model between each participant. The magnitude of the color scale ranges from 0–12 mm.

3. Discussion

As seen in Figures 9 and 10, the results suggest an overall improvement in surface error for all users as they gain more experience performing registrations in the application and progress through the goals of registering the right lobe target, left lobe target, then central target. While the summary data in Table 1 seemingly fails to support this trend, the data in Table 1 is based on the averages of the given registration goal for each participant, rather than the participant's progress throughout the trials in each registration task. Registration improvements as participants progressed through specific registration tasks can be further seen in Figure 5 and Figure 6. Although it would be expected that the surface error would decrease with the participant gaining experience, there are other factors at play, mainly the nature of the tasks given. Given a target on the left lobe as a goal, the overall surface error will increase since the left lobe is smaller than the right lobe and less surface data will be distributed on the liver surface. The only exception to this trend is participant 4, the surgeon. The surgeon had dramatic, consistent improvement between each registration task. This could be attributed to participant 4 quickly learning how to use *Hepatic Planner* and applying their previous experience.

One error that was seen when analyzing the results was the large variability in surface error and target error from one trial to another. However, the participants were encouraged to experiment with different methods of data collection to get the best registration they could, so the variability was expected. Furthermore, no other instructions were given outside of a tutorial on how to use *Hepatic Planner* and the basics of how the rigid registration system functioned. Because of this, it seemed the participants were choosing methods based on their own intuition, experience, and feedback from the app. So, it was expected to get the least surface and target error from participants 3 and 4, since these participants were the most experienced with the nuances of registration.

Despite the variability, there was a clear trend of improvement for participants in each goal. All participants' best trials had approximately the same surface and target error, despite the surface and target error from the initial trial being different between participants. Participants 2 and 3, who were more familiar with the inner workings of the rigid registration system and how it used the collected data, typically achieved their best registration on their first try. This could be attributed to their previous knowledge, which provided good intuition for their first trial. The lack of improvement after their initial trial could be attributed to trying new, less optimal theories towards data patterns for registering the organ more accurately. Participants 1 and 4, who were less familiar with the inner workings of the rigid registration system and how it used the collected data, seemed to have improvement from their first attempt. Participant 4 was expected to have the best accuracy in the whole study, since participant 4 has experience using the rigid registration system clinically, but there seemed to be a learning curve with *Hepatic Planner*. The consistent and large improvement may be due to participant 4 becoming more adept with *Hepatic Planner* rather than becoming more adept with the rigid registration system. Participant 1 had a similar jump in accuracy, but it was not expected that the same learning curve that may have impacted participant 1. Participant 4 was not accustomed to the presentation that *Hepatic Planner* used for the intraoperative organ. Participant 1 was much more familiar with models that didn't contain extra anatomy. Therefore, it is expected that the improvement of registration accuracy for participant 1 was due to increased understanding of the rigid registration system through the use of *Hepatic Planner*.

With regard to Table 2, the lowest target error was expected to be in the goal associated with that target. For example, the best registration for the left lobe target would be expected to occur when the participants were

given the left lobe goal. However, this was not the case in all goals. The best registration for the left lobe occurred during the right lobe target goal during several goals between participants. However, this is most likely due to the rigid registration system and the phantom used, rather than *Hepatic Planner*. The rigid registration system does not account for deformation that occurs between the preoperative and intraoperative states of the organ. This deformation will result in surface error and target error even in the gold-standard registration. In addition, the phantom used has substantial deformations, which result in large surface error and target error. To quantify the error more accurately, a liver model that has a greater number of targets could more fully represent the error present due to each trial registration. Furthermore, a liver model with less substantial deformations would allow for trends in registration data to be more easily seen since the less variability in the error would be present.

Another potential reason for the variability in the data collected was the anatomy of the model. The model does not contain obvious anatomical features that surgeons use for orientation of the organ. This was evident from the data collected by participant 4, who immediately searched for the anatomy to determine where to place the falciform on the model. In contrast, participants 1, 2, and 3 seemed to identify the location of the falciform from previous familiarity with similar liver models. This was most likely the reason for the large jump between the first registration and the best registration that participant 4 produced, both in terms of target error and surface error. In order for *Hepatic Planner* to be a viable training option for surgeons, a liver model with the relevant anatomy present should be used.

4. Conclusion

In summary, we find that Hepatic Planner can be a useful tool for improving the accuracy of rigid registration results. Between registration trials, the participants who had yet to fully understand the mechanics of the rigid registration method had regular improvement in the surface error and target error. However, the mean surface error did not significantly change with each goal given for all participants except for the surgeon, participant 4. With further revision and the implementation of a phantom that has known ground truth locations, clinically-derived relevant anatomy, and more realistic deformation, it would be possible to quantify the error in each lobe more accurately. We believe that a longer study with inexperienced trainees would allow more extensive tracking of the registration learning curve using *Hepatic Planner* and would further assess the usefulness of *Hepatic Planner* as a surgical education tool for the rigid registration system.

5. Acknowledgements

This work was supported by the National Institutes of Health with award R01CA162477 from the National Cancer Institute and the training award T32-EB021937 from the National Institute of Biomedical Imaging and Bioengineering.

7. References

[1] Siegel, R. L., Miller, K. D. and Jemal, A., “Cancer statistics, 2016.” CA Cancer J Clin., 66(1), 7–30 (2016).

[2] Kopetz, S., Chang, G. J., Overman, M. J., Eng, C., Sargent, D. J., Larson, D. W., Grothey, A., Vauthey, J. N., Nagorney, D. M. and McWilliams, R. R., "Improved Survival in Metastatic Colorectal Cancer Is Associated With Adoption of Hepatic Resection and Improved Chemotherapy," J Clin Oncol., 27(22), 3677-3683 (2009).

[3] Rucker, D. C., Wu, Y., Clements, L. W., Ondrake, J. E., Pheiffer, T. S., Simpson, A. L., Jarnagin, W. R. and Miga, M. I., "A Mechanics-Based Nonrigid Registration Method for Liver Surgery Using Sparse Intraoperative Data," IEEE Trans. Med. Imaging, 33(1), 147-158 (2014).

[4] Clements, L. W., Chapman, W. C., Dawant, B. M., Galloway, R. L. and Miga, M. I., "Robust surface registration using salient anatomical features for image-guided liver surgery: Algorithm and validation," Med Phys., 35(6), 2528-2540 (2008).

[5] Clements, L. W., Cash, D. M., Chapman, W. C., Galloway, R. L. and Miga, M. I., "Robust surface registration using salient anatomical features in image-guided liver surgery," Medical Imaging 2006: Visualization, Image-guided Procedures, and Display: Proc. of the SPIE, (in press), 2006