http://dx.doi.org/10.1117/12.2217515

# Cost-Effective Surgical Registration Using Consumer Depth Cameras

Michael Potter[1,2], Ziv Yaniv[1,3]

[1]Office of High Performance Computing and Communications, National Library of Medicine, National Institutes of Health, Bethesda, MD, 20894, USA
[2]Deptartment of Biomedical Engineering, Rochester Institute of Technology, Rochester, NY, USA 14623
[3] TAJ Technologies Inc., Mendota Heights, MN 55120, USA

## ABSTRACT

The high costs associated with technological innovation have been previously identified as both a major contributor to the rise of health care expenses, and as a limitation for widespread adoption of new technologies. In this work we evaluate the use of two consumer grade depth cameras, the Microsoft Kinect v1 and 3DSystems Sense, as a means for acquiring point clouds for registration. These devices have the potential to replace professional grade laser range scanning devices in medical interventions that do not require sub-millimetric registration accuracy, and may do so at a significantly reduced cost. To facilitate the use of these devices we have developed a near real-time (1-4 sec/frame) rigid registration framework combining several alignment heuristics with the Iterative Closest Point (ICP) algorithm. Using nearest neighbor registration error as our evaluation criterion we found the optimal scanning distances for the Sense and Kinect to be 50-60cm and 70-80cm respectively. When imaging a skull phantom at these distances, RMS error values of 1.35mm and 1.14mm were obtained. The registration framework was then evaluated using cranial MR scans of two subjects. For the first subject, the RMS error using the Sense was $1.28 \pm 0.01$ mm. Using the Kinect this error was $1.24 \pm 0.03$ mm. For the second subject, whose MR scan was significantly corrupted by metal implants, the errors increased to $1.44 \pm 0.03$ mm and $1.74 \pm 0.06$ mm but the system nonetheless performed within acceptable bounds.

**Keywords:** image-guided procedures, registration, surface acquisition, Microsoft Kinect, 3DSystems Sense

## 1. INTRODUCTION

Intra-operative registration is one of the key technologies facilitating image-guided navigation.[1] Most often, registration is performed using fiducials that are visible in a preoperative volumetric scan, typically CT or MR, and then intra-operatively localized using a tracked calibrated pointer. Alternatively, intra-operative surface scans can be utilized forgoing the need for fiducials in the preoperative scan. The three most common options for acquiring an intra-operative surface scan include touching the surface using a pointer,[2] employing visible light stereo cameras,[3,4] or by laser range scanning[2,3,5,6] .

The use of a tracked calibrated pointer is clinically less desirable, as it requires user interaction and slows the clinical workflow. This is not an issue for the two other surface acquisition methods that are fully automatic. Unfortunately, both stereo-vision and laser range scanning introduce additional equipment into the operating room. These devices, particularly the high-end laser range scanners used today, add non-negligible financial costs to the operation. The cost of innovation has been previously identified as both a major contributor to the rise of health care expenses, and as a limitation for widespread adoption of new technologies.[7] Moreover, it is often unclear whether new technologies actually provide benefits to surgical procedures which outweigh their costs.[8] By designing navigation guidance systems using cost effective consumer grade hardware we potentially improve their chances for widespread clinical adoption, minimizing the financial burden on the healthcare system associated with adopting technological innovation.

We therefore investigate the use of two consumer grade depth cameras, the Microsoft Kinect v1 and 3DSystems Sense, as cost effective alternatives to stereo-vision and laser range scanning technologies. Both devices provide

**E-mail:** mjp2010@rit.edu, zivyaniv@nih.gov

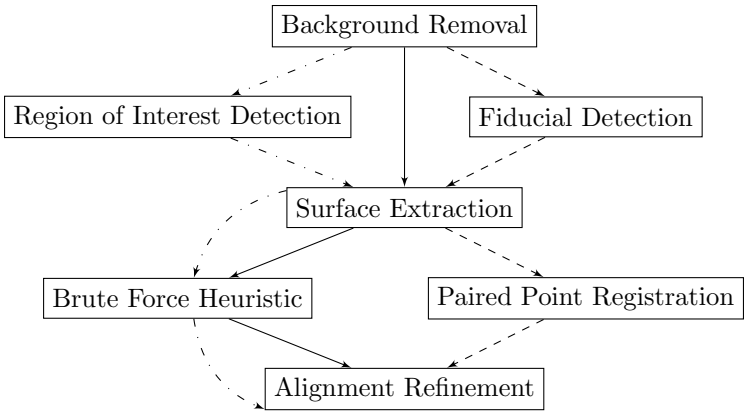



Figure 1: Overview of the surface registration framework. The three registration approaches are identified by their dataflows: no additional information (solid path), region of interest identification (dash-dotted path), and anatomical landmark localization which are used as fiducials (dashed path).

depth information and corresponding color data and are referred to, from here on, as RGBD cameras. In this work we describe a registration framework based on the open-source Point Cloud Library (PCL)[9] that readily fits within the clinical workflow and performs in near-real time. We then empirically identify optimal scanning distances for data acquisition and qualitatively evaluate the influence of lighting conditions on each device's performance. Finally, we evaluate our registration approach using CT data acquired from a skull phantom and MR data acquired from two subjects. It should be noted that from an implementation standpoint, the use of the PCL library enables us to use our implementation in conjunction with any RGBD camera which has OpenNI2 drivers (e.g. Xtion Pro Live).

## 2. METHODS

The proposed intraoperative registration framework is based on the use of a surface model derived from a preoperative diagnostic quality volumetric image. The complexity of obtaining the external body surface may vary based on the modality and region of the body. In our case, we are interested in neurosurgical procedures and thus deal with cranial MR scans. Segmenting the cranial surface in MR is easily done via thresholding, which in our case was performed using the open source program 3DSlicer.[10] In other cases, a simple thresholding may not be sufficient, requiring the use of more complex segmentation algorithms. It is worth mentioning that in CT automatic removal of the bed surface[11] will greatly facilitate segmentation of the anatomical surface.

The framework consists of three approaches towards surface based registration that utilize varying levels of information to increase the convergence success rate of the iterative closest point (ICP) algorithm.[12] The method which requires the most information uses anatomical landmarks manually localized in the preoperative scan. If this approach is the desired one, then we require the operator to identify these distinct landmarks preoperatively, after performing surface segmentation.

We next describe our registration framework in detail.

### 2.1 Registration Framework

The developed registration framework is capable of leveraging three different levels of information to perform the initial alignment for the ICP algorithm. The first approach requires no additional information, directly using the preoperative and intraoperative point data. The second approach requires the user to identify the locations of regions of interest in the intraoperative data by specifying points. The third approach relies on preoperative anatomical landmarks that are also identified intraoperatively. Figure 1 summarizes the dataflows defining each of these approaches within our framework.

The alignment process takes between one and four seconds per frame, but is separated on the processor from the live image feed which allows for alignment to occur in tandem with the live-streaming of data. Note that

this performance speed was measured running the registration framework on a 2011 MacBookPro laptop with a 1GB AMD Radeon HD 6750M GPU. The program uses less than 800MB of RAM, and leverages the GPU only for visualization purposes.

As indicated in Figure 1, the first step for all of the registration approaches is to remove the background data obtained by the RGBD camera. The background was defined to be all points within the intraoperative point cloud having a z-depth, distance from camera, greater than 2 meters. Based on empirical tests, data captured beyond this distance would not lead to effective alignment and we thus expect the anatomical surface to be closer to the depth camera. It should be noted that removing these points improves the computation speed as we are dealing with a smaller point cloud.

#### 2.1.1 Approach I: no additional information

In this approach we begin by identifying the subset of points in the point cloud that potentially correspond to the anatomical structure of interest using a heuristic approach. We start by clustering points based on their Euclidean distance from each other. Points that are more than 1cm away from each other are considered to belong in different clusters. This clustering approach is sped up by using a kD-tree for fast nearest neighbor queries.[13]

To identify the cluster corresponding to the anatomy of interest the following heuristics are used: it is assumed that the anatomy of interest is centered in the image, that it is the closest object to the camera and that its bounding box is smaller than that of the preoperative point cloud but contains at least 40 points, an empirically determined lower bound.

Once the point cluster corresponding to the anatomy of interest has been identified we use a brute force approach to obtain an initial alignment for the ICP algorithm. This is done by generating a set of possible poses and evaluating each of them using the root mean squared error (RMSE) between the intraoperative cloud and the preoperative cloud. For each of the initial poses we run the ICP algorithm with a tight bound on the number of iterations (50). The resulting poses are compared and the one that corresponds to the minimal RMSE serves as the initial pose for the ICP algorithm with a looser bound on the number of iterations (500).

The set of initial poses is generated as follows. The center of mass of the preoperative point cloud and the intraoperative anatomy cloud are first aligned. Then all 24 unique rotation matrices that correspond to all combinations of $90^o$ Euler ZYX angle choices in 3-space are evaluated. It bears mentioning that, while there are 64 different Euler angle permutations of $90^o$ angles, they do not correspond to 64 unique rotation matrices. For example, rotations (xyz) of ($0^o$,$0^o$,$90^o$) and ($180^o$,$180^o$,$270^o$) are parameterizations of the same rotation matrix:

$$r(0,0,90) = r(180,180,270) = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Finally, for each rotation, we perform a translation to match the scanned surface bounding box against the target model bounding box along each axis in turn.

#### 2.1.2 Approach II: region of interest identification

This approach is a slight modification of the previous one. The difference between the two is in the identification of the set of points which correspond to the anatomy of interest in the intraoperative point cloud. In the previous approach we used a heuristic to identify this subset of points while here the user identifies the region of interest directly. This is done by specifying locations on the intraoperative point cloud. Points that are farther than a threshold from the specified locations are removed from the data. The empirical threshold which we use is derived from the axis aligned bounding box of the preoperative point cloud. We use half of the length of the box's diagonal as the threshold.

Specifying locations on the intraoperative point cloud is done in one of two ways, either using a mouse to click on the visualized point cloud data or by placing skin adhesive color markers on the anatomy. In our case these markers are circular color stickers with a diameter of 12.7mm. Each of these methods has its advantages and disadvantages. On the one hand, using a mouse allows the operator to unambiguously specify the locations

(a) low illumination (b) medium illumination (c) high illumination

Figure 2: Skull phantom data acquisition under different lighting conditions, shown here with the Sense RGBD camera.

while using markers requires robust detection and is potentially susceptible to failed and erroneous marker detections. On the other hand, in the clinical setting using a mouse is likely less desirable due to space and sterility considerations, while using markers is straightforward and easily fits within the clinical workflow.

To automatically identify the locations of the color markers in the intraoperative point cloud we analyze the color information associated with each of the points. Analysis is performed using the HSV color space as it is more stable with respect to illumination changes than the RGB color space.[14] As we are using green markers we use the following intervals in HSV space to identify points that potentially belong to the markers, H=[$70^o$,$120^o$], S=[0.25,1.0], and V=[0.25,1.0]. It is worth noting that there was some variability between the color levels reported by the two cameras we investigated. If a different RGBD camera is to be used, these values may need to be modified.

After identifying all points which correspond to the expected color, we search for clusters of between 3 and 30 green points with less than 1cm between them. This step allows us to ignore false positives which are within the HSV thresholding limits but have a size that differs from the expected marker size range. The centroid, mean location, of each cluster defines a region of interest as described above.

#### 2.1.3 Approach III: anatomical landmark fiducials

This approach extends the previous one in the following manner. Preoperatively the operator identifies anatomical landmarks in the preoperative point cloud. Intraoperatively, the operator places the skin adhesive markers on the corresponding physical locations. The markers are detected as described above. Our approach makes two assumption, the marker configuration is visible in a single RGBD frame, and the distances between markers are unique. Once all marker centroids have been detected we automatically pair them with the corresponding landmarks identified in the preoperative point cloud based on their unique distances from each other.

When the pairing is obtained we use a closed-form paired point registration algorithm based on singular value decomposition.[15] The output of this alignment serves as input to the ICP algorithm.

This approach of point landmark localization in one dataset and paired region localization in the other dataset is similar to the spotlight regions of registration approach described in,[16] with the difference being that in that work the spotlight region was defined in the preoperative space whereas in this work it is defined in the intraoperative one.

## 3. EXPERIMENTAL EVALUATION

We first investigated the effect of lighting conditions and distance between the RGBD camera and object of interest on registration accuracy. This evaluation was performed using an anatomical phantom for which we acquired a CT scan. We then evaluated the registration framework performance using MR data from two subjects and the effect of three simulated patient setups on accuracy.

Figure 3: Experimental setup for evaluating the effect of camera distance on registration accuracy. Skull phantom positioned at various distances from the camera.

### 3.1 Lighting and Distance Evaluation

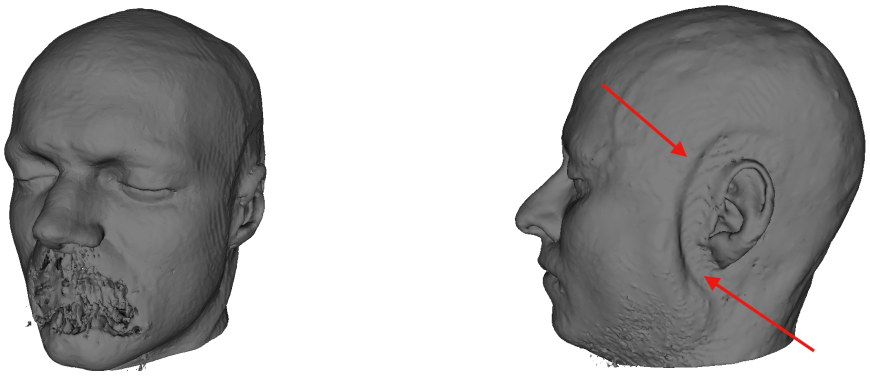
We began by identifying the optimal lighting conditions for each of the cameras. To achieve this, the cameras were placed on a tripod and pointed at a plastic skull phantom, which was also fixed to a tripod. Three different levels of lighting were investigated, as shown in Figure 2.

Once acceptable lighting conditions were determined, we evaluated the effect of the distance between the camera and phantom on registration accuracy. To do this, the phantom was moved from immediately in front of the camera to several meters back and the final alignment accuracies were quantified. Figure 3 illustrates this experimental setup. The preoperative phantom surface was obtained from a CT scan with the following dimensions 512x512x416 and a spacing of 0.48x0.48x0.62mm.

At each location a point cloud of the skull phantom was acquired. The average z-depth, distance from camera, of the points belonging to the object of interest as reported by the camera was also recorded. We then ran the registration framework using the point cloud data to quantify the registration accuracy for each camera distance (z-depth).

(a) (b) (c)

Figure 4: Simulated patient setups used for registration framework evaluation (a) supine (b) supine with headrest (c) prone.



(a) (b)

Figure 5: Illustration of potential issues associated with differences between the skin surface obtained from MR and from intraoperative surface scanning (a) presence of implants corrupting the MR data and consequentially the derived surface and (b) headphones used during MR scan physically deforming the surface.

### 3.2 Simulated Patient Setup Evaluation

Once the ideal imaging ranges were identified, the registration framework was evaluated using surface data derived from MR images and point cloud data acquired from two subjects in a simulated patient setup. The MR images were 256x256x176 with an isotropic 1mm spacing. For each camera we evaluated three patient positions. First, simulating the patient in the supine position, similar to the setup in which the MR scan was acquired. We then had each subject lie with their head suspended over the ground, simulating the use of a headrest. Finally, we acquired scans of the subjects lying face down, in the prone position, with the camera on the ground beneath them. This pose allowed us to evaluate the effects of gravity on registration accuracy, as the MR scan and point cloud were acquired with different patient setups, supine versus prone. Figure 4 shows these simulated setups.

It should be noted that our registration framework assumes that the preoperative and intraoperative surfaces can be accurately aligned using a rigid transformation. There are situations where this assumption is invalidated, potentially degrading alignment accuracy. One such situation, already alluded to, is associated with the effects of gravity. When the preoperative and intraoperative patient poses are different, gravity will introduce a nonrigid shift of the surface. Additional discrepancies between the preoperative and intraoperative surfaces can arise from the presence of implants which may have significant effects on the MR scan and derived surface, as illustrated in Figure 5(a). In such cases the use of anatomical landmark fiducials, as described in section 2.1.3, allows our framework to provide accurate registration results. Finally, issues specific to cranial procedures include the introduction of deformation due to the usage of headphones during the MR scan, as illustrated in Figure 5(b), and the lack of hair in the MR scan. The use of headphones aims to make the MR acquisition process more pleasant for the patient but if need be the scan can be acquired without them. The lack of hair and eyebrows in MR versus their presence in the intraoperative point cloud may also degrade the quality of alignment.

## 4. RESULTS

### 4.1 Lighting and Distance

The lighting evaluation showed very little impact of illumination levels on system performance. For the Sense, the average RMS errors for low, medium and high illumination levels were $1.139 \pm 0.003$mm, $1.132 \pm 0.008$mm, and $1.144 \pm 0.009$mm. These values are identical to within their 95% confidence intervals, meaning that the change in lighting had no detectable effect on performance. For the Kinect in the low, medium, and high conditions, errors of $1.22 \pm 0.01$mm, $1.27 \pm 0.01$mm, and $1.29 \pm 0.01$mm were observed. While these values indicate a slight increase in error as lighting increases, the apparent effect size of 0.07mm is on the same order as the system variability between individual trials (Table 1). For all practical purposes, then, it appears that both scanners function consistently over a reasonable spectrum of ambient lighting intensities. The lack of variation is not

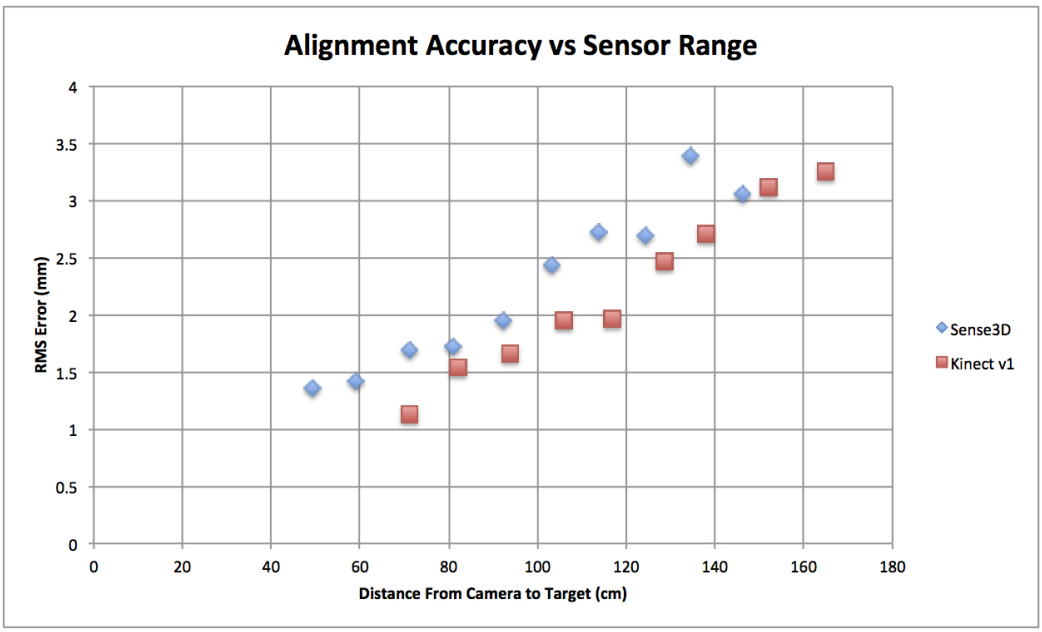
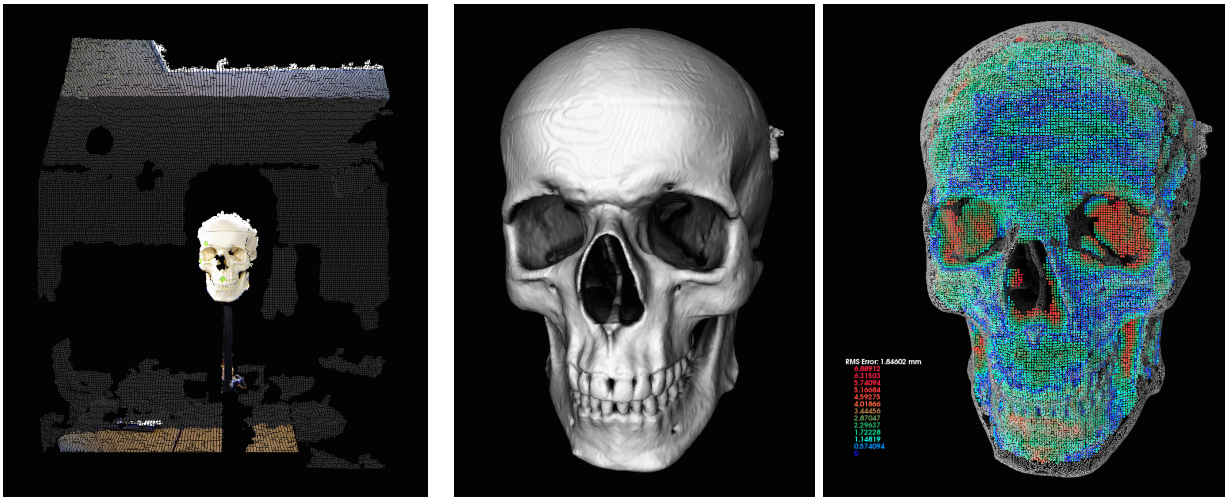



Figure 6: Alignment RMS as a function of distance from the camera.

surprising, as both cameras were designed to automatically correct their exposure levels in order to compensate for these sorts of changes in ambient lighting.

It was determined that the optimal viewing distances for the Sense and Kinect were between 50-60cm and 70-80cm respectively. These distances correspond to the minimum possible distance away from the cameras before portions of the anatomy of interest were no longer visible by the camera. Accuracy continuously degraded as the object was moved further and further away from the camera, as shown in Figure 6.

At their optimal distances, the Sense had an RMS error of 1.35mm compared to 1.14mm for the Kinect, when imaging the skull model. Figure 7c illustrates the spatial distribution of errors, overlaying the aligned points from the Sense point cloud onto the skull model, with the point color indicating its distance to the nearest neighbor in the skull model. Most of the error accrued comes from the eye-socket cavity, which both scanners had difficulty imaging properly.



(a) raw input cloud (b) CT model reconstruction (c) nearest neighbor errors

Figure 7: The skull model throughout the alignment procedure.

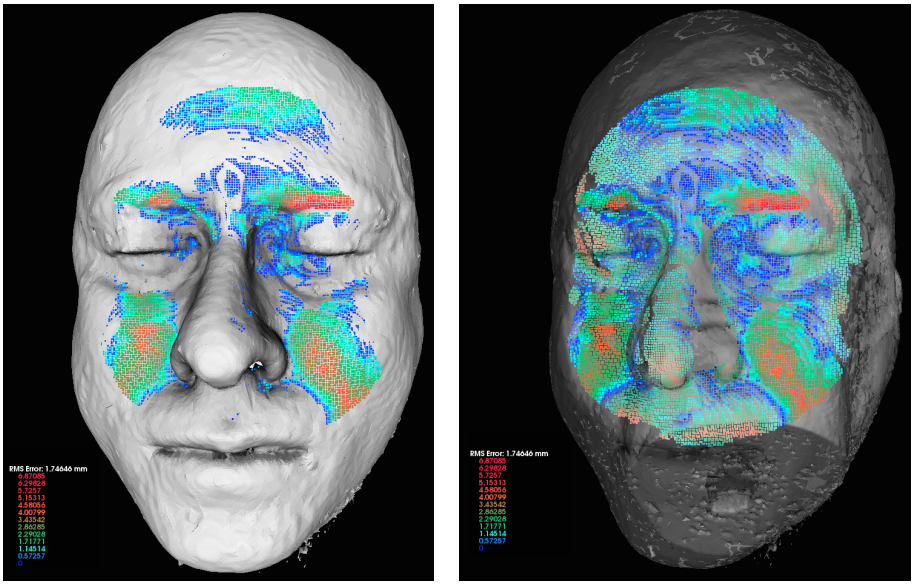
| Subject | Sensor | Setup | RMS Error (95% CI) |
|---|---|---|---|
| 1 | Sense | supine | $1.28 \pm 0.01$ mm |
| | | supine with headrest | $1.30 \pm 0.03$ mm |
| | | prone | $1.73 \pm 0.01$ mm |
| | Kinect | supine | $1.24 \pm 0.03$ mm |
| | | supine with headrest | $1.34 \pm 0.03$ mm |
| | | prone | $1.77 \pm 0.05$ mm |
| 2 | Sense | supine | $1.44 \pm 0.03$ mm |
| | | supine with headrest | $1.90 \pm 0.09$ mm |
| | | prone | $1.80 \pm 0.03$ mm |
| | Kinect | supine | $1.74 \pm 0.06$ mm |
| | | supine with headrest | $1.58 \pm 0.06$ mm |
| | | prone | $2.24 \pm 0.04$ mm |

Table 1: Average alignment errors according to simulated patient setup.

### 4.2 Simulated Patient Setup

The results of the three simulated patient setup evaluations are summarized in Table 1. For the first subject the RMS error using the Sense was $1.28 \pm 0.01$ mm (95% t-test confidence interval). Using the Kinect this error was $1.24 \pm 0.03$ mm. On the second subject, whose MR scan was significantly corrupted by metal implants, the errors increased to $1.44 \pm 0.03$ mm and $1.74 \pm 0.06$ mm, requiring the use of fiducial stickers to ensure a correct alignment, but the registration framework nonetheless remained functional. As the values in Table 1 show, there was no statistically significant difference (at p=0.05) between the performance of the Sense and Kinect cameras for subject 1. The data from the second subject exhibited larger differences between camera performance, but with contradictory results as to which camera was better for different poses. This increased variability is likely a result of the corrupted model rather than indicative of a true difference in hardware performance, as will be explored more thoroughly throughout the remainder of this analysis.

Errors increased in the prone position for several reasons. Of chief concern, when the subjects were prone their noses and some other facial features moved within the minimum viewing distance of the cameras. This was especially problematic for the Kinect, and is certainly an important consideration if one needs to acquire the



(a) error heat map over model Surface (b) error heat map (translucent model)

Figure 8: Spatial distribution of alignment errors for alignment of subject 1 in the prone position.

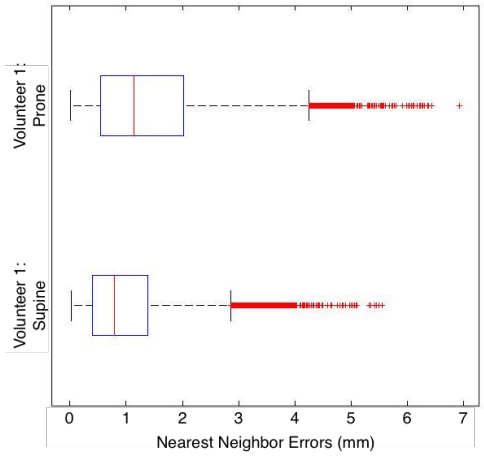


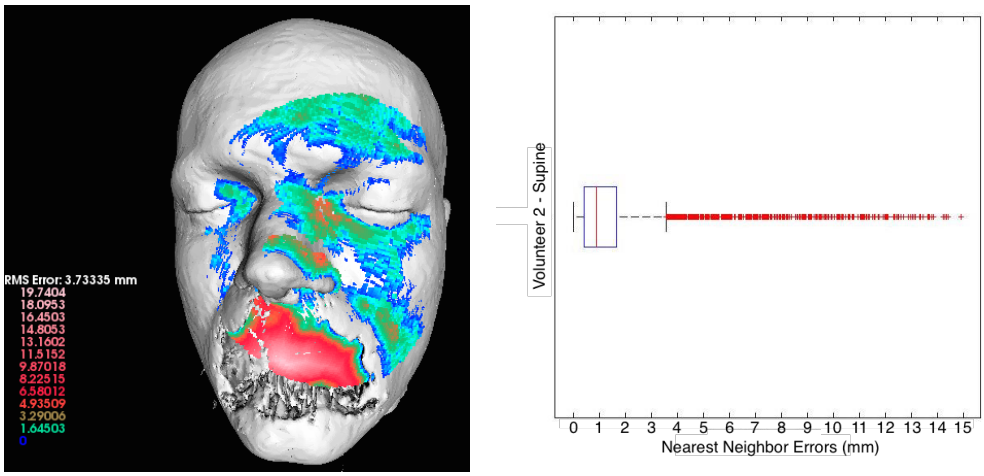

Figure 9: Error distributions as a function of subject setup.

data with the patient in the prone position. In a clinical setting this can possibly be mitigated by increasing the height of the surgical table, something we could not do in our simulated setup. The prone setup also suffered from the effects of gravity, which caused the tip of the nose as well as portions of the cheeks to be displaced by several millimeters compared to their locations in the MR scans which were acquired in the supine position. This phenomena is illustrated in Figure 8.

As Figure 8 also demonstrates, hair, in this case eyebrows, can be a significant contributor to RMS error. While deformation due to gravity resulted in increased errors, this issue is not specific to the use of RGBD cameras for surface acquisition. When the MR data is acquired in the supine position and the surface data is acquired with the patient in the prone position the assumption underlying our registration framework, rigid relationship between the datasets, is invalidated.

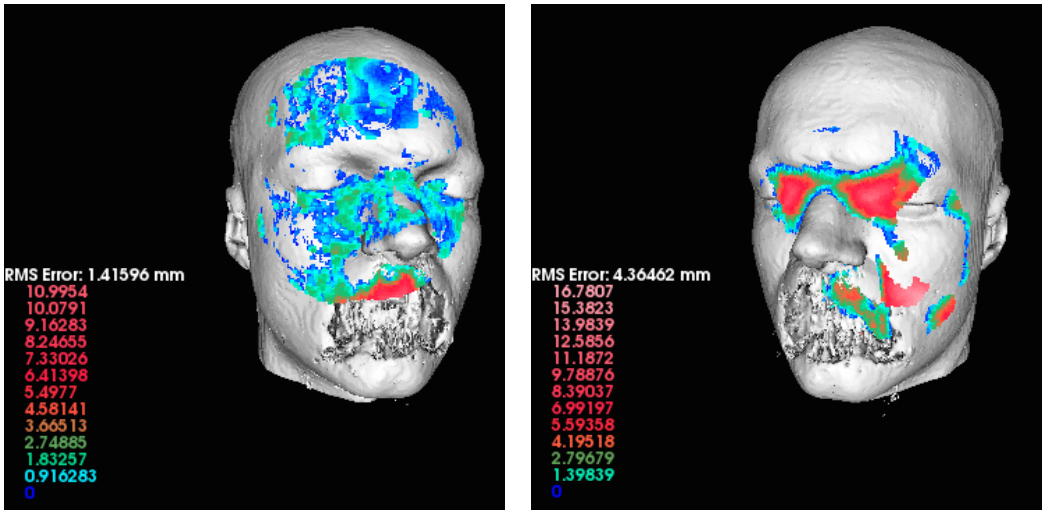
The nearest-neighbor errors over two cropped point clouds are given in Figure 9. These results show that, while the error distribution becomes more spread out when the subject's orientation with respect to gravity changes, the alignment procedure is still able to maintain a high fraction of low-error correspondences.

For subject 2 we observed errors significantly in excess of those observed for subject 1. This increase in error is largely attributable to errors in the preoperative surface reconstruction which arise from imaging artifacts in

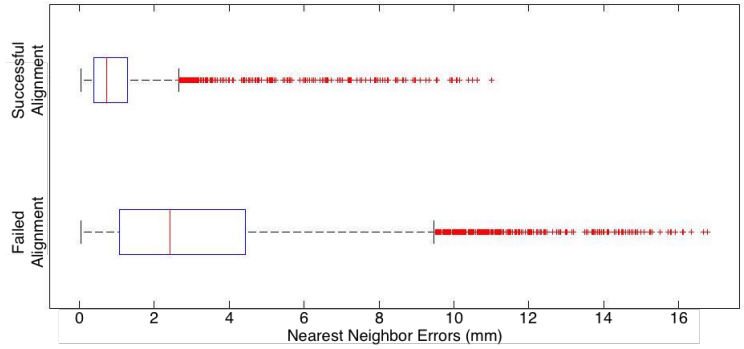



(a) Alignment RMS: corrupted MR model (b) Alignment error distribution

Figure 10: Visualization of alignment errors due to a corrupted preoperative model.




(a) Successful alignment (b) Failed alignment




(c) Alignment error distributions

Figure 11: Identifying alignment failures.

the MR scan due to the presence of an implant, readily seen in Figure 10a. The impact of the preoperative surface errors are easily seen in Figure 10b where the error tail extends far beyond the impacts observed by gravity in Figure 9.

The large error values in the mouth region of Subject 2 made the automatic alignment algorithm unreliable, with an example of a failed alignment given in Figure 11b. In spite of the occasional failures of purely automatic alignment on the corrupted point set, anatomical landmark fiducial based registration was still possible and demonstrates a potential advantage of consumer depth RGBD cameras which can use color information for non-intrusive fiducial markings. In spite of the large errors present at the mouth, the algorithm was still able to fit the rest of the input scan to the model with a reasonable degree of accuracy.

As demonstrated in Figure 11c, not only do incorrect alignments exhibit a higher total RMS error than correct alignments, but their individual nearest-neighbor errors are also significantly more spread than in properly aligned cases. This makes it possible to automatically detect failed alignments and ignore them.

### 4.3 Registration Framework Robustness

As introduced in Figure 1, our registration framework is comprised of three execution paths, each requiring varying degrees of user interaction. One point of failure for the fully automated procedure was in the detection and subsequent extraction of the head from background points before alignment. When performing analysis on Subject 1, this component of the automated procedure had a success rate of 66% (8 out of 12 trials). If the user identified the region of interest locating the head, a single click operation, the success rate increased to 100%.

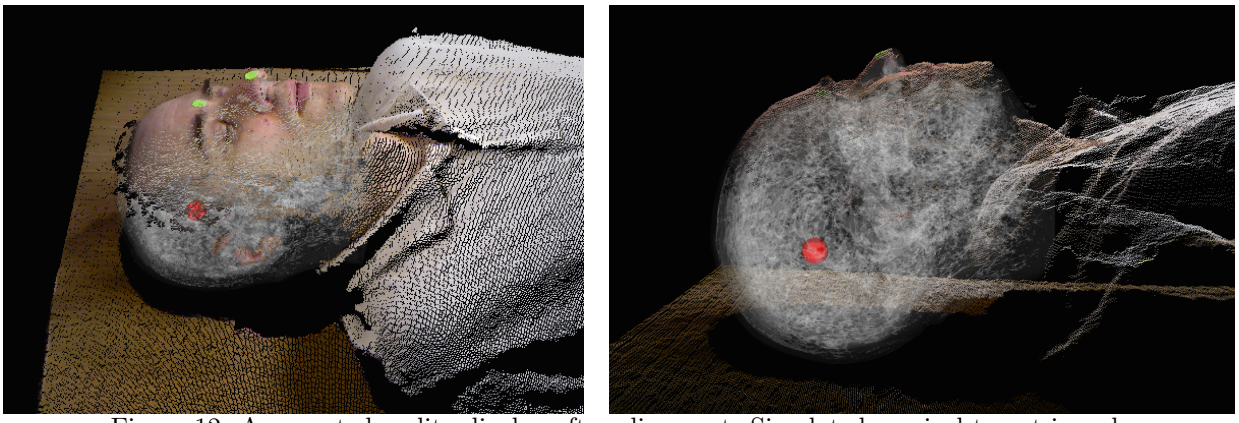


Figure 12: Augmented reality display after alignment. Simulated surgical target in red.

The anatomical landmark fiducial based approach, which required placing three or more green markers on the subject before using acquiring the intraoperative point cloud, likewise saw a success rate of 100%.

Another potential avenue for failure comes in the form of alignment breakdown despite accurate extraction of the head from background points. This sort of error can be seen in Figure 11b and is typically accompanied by increased RMS error values as well as by visually obvious overlay failure. While not a problem in the case of subject 1, the corrupted preoperative model of subject 2 made this error significant. The success rate of both the fully automated approach and the region of interest identification approached dropped to 25%, as incorrect alignments were often favored. When fiducial markers were used the success rate returned to 100%. This demonstrates the usefulness of user input in increasing the robustness of alignment, but also shows that even heavily corrupted datasets can be utilized for registration purposes.

## 5. DISCUSSION AND CONCLUSIONS

The use of RGBD cameras in the clinical setting has been investigated for a wide variety of tasks as described in the survey by Bauer et al.[17] Amongst others, these include registration for various clinical applications: positioning and management of patients in radiation therapy, registration for respiratory motion compensation in tomographic reconstruction and registration for guidance of open and percutaneous procedures. Our work falls into this last category, registration for guidance of neurosurgical procedures.

Similar approaches to that presented in this paper were recently described in[18] and.[19] In[18] the Kinect is used for registration in the context of augmented reality for forensic medicine. Registration is performed in two steps, initial alignment by manually selecting three or more non colinear points in both data sets, this is followed by the use of the ICP algorithm. Our approach differs from this registration approach in that our framework uses color markers to automatically identify the anatomical landmarks for the initial alignment, and we also provide two additional registration approaches that require less information. The reported accuracy, in that work referred to as FRE, was between 1-2mm in a phantom study and between 3-4mm in a cadaver study. Runtime of the registration is 1sec with 10sec required for initialization. In our framework the complete registration takes 1-4sec. In[19] the Kinect is used for registration in the context of augmented reality guidance for percutaneous needle insertions. Registration is performed in the same manner as in.[18] The reported accuracy, fiducial, was about 2mm in phantom studies and 2-3mm in an animal study. It should be noted that in both studies it appears that the cadaver and animal were in the same pose during the preoperative image acquisition and during the intervention. In our work we have shown that the surface deformation in the facial region associated with change in pose is non-negligible and will effect registration accuracy. Depending on the clinical requirements this reduction in accuracy may be acceptable.

The ultimate objective of this registration process is not simply to move points close together, but rather to enable augmented reality visualization to enhance surgical procedures. To this end our system allows for users to input the location of surgical targets which have been determined pre-operatively. These targets will then be overlaid onto the scene, as shown in Figure 12, to help guide the clinician during the procedure.

We have presented an empirical evaluation of two consumer grade RGBD cameras as a means for surface based registration. We first empirically evaluated the effect of ambient lighting and distance from the object of interest on the quality of the data. We found that both sensors are robust with respect to a broad range of lighting conditions, with measurement differences of less than 0.1mm when varying the ambient lighting. As both sensors use a triangulation based approach to point reconstruction the point density on the object surface decreases with increasing distance.[20] Thus, the optimal input for registration, densest point cloud, is acquired when the distance between the object of interest and the RGBD camera is minimized. As we are dealing with a triangulation based reconstruction approach there is also a minimal distance below which the object of interest is too close, causing self occlusions. These occur when the infrared camera cannot see regions that are illuminated by the laser.[20] In our case these distances were empirically shown to be 50-60cm for the Sense and 70-80cm for the Kinect with a skull phantom as the object of interest, which is appropriate as we are targeting neurosurgical procedures.

We have shown that both RGBD cameras yielded similar registration accuracies, RMS below 2mm, which are within the error bounds appropriate for neurosurgical procedures as compared to current registration accuracy obtained in these procedures.[21] Differences between the two RGBD cameras that can potentially influence clinical adoption include the ergonomic design of the Sense which is better suited for handheld scanning and the Kinect's need for a power source other than the USB connection, requiring two cables versus the single USB cable required by the Sense.

## ACKNOWLEDGMENTS

We thank Kevin Cleary and Kendall O'Brien for their help with data acquisition.

## REFERENCES

[1] Yaniv, Z. and Cleary, K., "Image-guided procedures: A review," Tech. Rep. CAIMR TR-2006-3, Image Science and Information Systems Center, Georgetown University (April 2006).

[2] Simpson, A. L., Burgner, J., Glisson, C. L., Herrell, S. D., Ma, B., Pheiffer, T. S., III, R. J. W., and Miga, M. I., "Comparison study of intraoperative surface acquisition methods for surgical navigation," *IEEE Trans. Biomed. Eng.* **60**(4), 1090–1099 (2013).

[3] Kau, C. H., Richmond, S., Incrapera, A., English, J., and Xia, J. J., "Three-dimensional surface acquisition systems for the study of facial morphology and their application to maxillofacial surgery," *International Journal of Medical Robotics and Computer Assisted Surgery* **3**(2), 97–110 (2007).

[4] Kumar, A. N. et al., "Phantom-based comparison of the accuracy of point clouds extracted from stereo cameras and laser range scanner," in [*SPIE Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling*], (2013).

[5] Shamir, R. R., Freiman, M., Joskowicz, L., Spektor, S., and Shoshan, Y., "Surface-based facial scan registration in neuronavigation procedures: a clinical study," *J Neurosurg.* **111**(6), 1201–1206 (2009).

[6] Cao, A., Thompson, R. C., Dumpuri, P., Dawant, B. M., Galloway, R. L., Ding, S., and Miga, M. I., "Laser range scanning for image-guided neurosurgery: Investigation of image-to-physical space registrations," *Med. Phys.* **35**(4), 1593–1605 (2008).

[7] Bodenheimer, T., "High and rising health care costs. part 2: Technologic innovation," *Ann Intern Med.* **142**(11), 932–937 (2005).

[8] Barbash, G. I. and Glied, S. A., "New technology and health care costs - the case of robot-assisted surgery," *N Engl J Med.* **363**(8), 701–704 (2010).

[9] Rusu, R. B. and Cousins, S., "3D is here: Point cloud library (PCL)," in [*IEEE International Conference on Robotics and Automation (ICRA)*], 1–4 (2011).

[10] Fedorov, A. et al., "3D slicer as an image computing platform for the quantitative imaging network," *Magn Reson Imaging.* **30**(9), 1323–1341 (2012).

[11] Zhu, Y.-M., Cochoff, S. M., and Sukalac, R., "Automatic patient table removal in CT images," *J Digit Imaging* **25**(4), 480–485 (2012).

[12] Besl, P. J. and McKay, N. D., "A method for registration of 3D shapes," *IEEE Trans. Pattern Anal. Machine Intell.* **14**(2), 239–255 (1992).

[13] Rusu, R. B., *Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments*, PhD thesis, Technical University Munich (2009).

[14] Gevers, T. and Smeulders, A. W. M., "Color based object recognition," *Pattern Recognition* **32**(3), 453–464 (1999).

[15] Arun, K. S., Huang, T. S., and Blostein, S. D., "Least-squares fitting of two 3-D point sets," *IEEE Trans. Pattern Anal. Machine Intell.* **9**(5), 698–700 (1987).

[16] Ma, B. and Ellis, R. E., "Robust registration for computer-integrated orthopedic surgery: Laboratory validation and clinical experience," *Medical Image Analysis* **7**(3), 237–250 (2003).

[17] Bauer, S. et al., "Real-time range imaging in health care: A survey," in [*Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications*], *Lecture Notes in Computer Science* **8200**, 228–254, Springer Berlin Heidelberg (2013).

[18] Kilgus, T. et al., "Mobile markerless augmented reality and its application in forensic medicine," *International Journal of Computer Assisted Radiology and Surgery* **10**(5), 573–586 (2015).

[19] Seitel, A. et al., "Towards markerless navigation for percutaneous needle insertions," *International Journal of Computer Assisted Radiology and Surgery* **in press** (2015).

[20] Khoshelham, K. and Elberink, S. O., "Accuracy and resolution of kinect depth data for indoor mapping applications," *Sensors* **12**(2), 1437–1454 (2012).

[21] Shamir, R. R., Joskowicz, L., Spektor, S., and Shoshan, Y., "Localization and registration accuracy in image guided neurosurgery: a clinical study," *International Journal of Computer Assisted Radiology and Surgery* **4**(1), 45–52 (2009).