Phantom-based Comparison of the Accuracy of Point Clouds Extracted from Stereo Cameras and Laser Range Scanner

Ankur N. Kumar*[1], Thomas S. Pheiffer[2], Amber L. Simpson[2], Reid C. Thompson[3], Michael I. Miga[2], and Benoit M. Dawant[1]
[1]Department of Electrical Engineering, Vanderbilt University, Nashville TN, USA
[2]Department of Biomedical Engineering, Vanderbilt University, Nashville TN, USA
[3]Department of Neurological Surgery, Vanderbilt University, Nashville TN, USA

ABSTRACT

Using computational models, images acquired pre-operatively can be updated to account for intraoperative brain shift in image-guided surgical (IGS) systems. An optically tracked textured laser range scanner (tLRS) furnishes the 3D coordinates of cortical surface points (3D point clouds) over the surgical field of view and provides a correspondence between these and the pre-operative MR image. However, integration of the acquired tLRS data into a clinically acceptable system compatible throughout the clinical workflow of tumor resection has been challenging. This is because acquiring the tLRS data requires moving the scanner in and out of the surgical field, thus limiting the number of acquisitions. Large differences between acquisitions caused by tumor resection and tissue manipulation make it difficult to establish correspondence and estimate brain motion. An alternative to the tLRS is to use temporally dense feature-rich stereo surgical video data provided by the operating microscope. This allows for quick digitization of the cortical surface in 3D and can help continuously update the IGS system. In order to understand the tradeoffs between these approaches as input to an IGS system, we compare the accuracy of the 3D point clouds extracted from the stereo video system of the surgical microscope and the tLRS for phantom objects in this paper. We show that the stereovision system of the surgical microscope achieves accuracy in the 0.46-1.5mm range on our phantom objects and is a viable alternative to the tLRS for neurosurgical applications.

Keywords: image-guided neurosurgery, stereovision, laser range scanner, intraoperative imaging, surgical microscope, accuracy, brain shift

1. INTRODUCTION

In brain tumor surgery, brain tissue deformations, commonly referred to as brain shift or cortical displacements/deformations, can produce inaccuracies of 1-2.5cm in the preoperative plan and within image-guided surgical (IGS) systems [1-2]. Furthermore, the changing resection cavity and tumor removal compound the inaccuracies caused by brain shift. These issues hinder the process of establishing accurate correspondence between the physical state of the patient and their images during neurosurgery. This causes intraoperative real-time inaccuracies in IGS systems. Several forms of intraoperative imaging have been used to account for intraoperative inaccuracies to improve IGS systems in neurosurgery. These intraoperative modalities include magnetic resonance imaging, ultrasound, laser range scanner and stereovision systems. In this paper, we primarily focus on the textured laser range scanner (tLRS) and stereo camera systems.

The optically tracked tLRS can provide intraoperative digitized cortical surfaces at various stages (post-duratomy, mid-resection, and post-resection) of the neurosurgery. These digitized surfaces or point clouds have sub-millimetric accuracy. These point clouds have been used in accounting for brain deformations and they have been used for driving a patient-specific biomechanical brain model to update the IGS system [3-7]. Furthermore, the tLRS holds color information for each digitized point originating from the digital image of the field of view. This attribute has been used to develop a methodology for matching vessels preoperatively and postoperatively on the obtained surfaces. The deformation measurements between the two surfaces states have been used to drive our guidance update pipeline for the IGS system [8]. While the tLRS provides valuable intraoperative information, the measurements are temporally too

sparse for computing intermediate updates for all cases. Indeed, establishing correspondences between digitized surfaces extracted at different time points in the surgical procedure can be difficult especially in the presence of large tumor resections. These conditions induce substantial cortical displacement and the consistent visibility of salient features such as vessels may not be available for correspondence. In addition, the positioning and acquisition of intraoperative data using the tLRS takes on the order of 1 minute, which can be disruptive to the surgical workflow. Along with the tLRS, monocular video from the surgical microscope has been probed as a possible solution for acquiring temporally dense data [9, 25]. This approach does not find 3D coordinates of points in the surgical field of view and is effective for short continuous sequences only.

A surgical microscope is routinely used to perform brain tumor surgery and it can be used as a continuous intraoperative data source by augmenting the microscope with two cameras or a stereo camera system. Equipped with stereovision concepts of calibration and reconstruction from the field of computer vision, video streams from the stereo camera system can be used to produce a textured point cloud at every pair of frames of the video stream. As a result, temporally dense 3D measurements of the cortical surface could be available for driving the updates to the IGS system. Early work involved the manual initialization of the seed points in the left and right camera images acquired by the stereo camera system to extract a point cloud through surface growing [10-12]. Two different stereovision systems, both attached externally to the surgical microscope, have been developed in [13-14] and [15].

The stereo camera system designed for brain shift compensation in [13-14] has a few disadvantages. The cameras are attached externally to the microscope optics and the assistant ocular arm is rendered unusable when the cameras are powered on. Often, the assistant ocular arm of the microscope is used by the neurosurgeon as a teaching tool. This limits the acquisition of temporally dense cortical surface measurements, which is the key purpose of using the surgical microscope. An update of the system in [13-14] was made in [16-17], where the external stereo camera system is able to determine 3D coordinates for 1000 uniformly spaced points on the stereo-pair images. This is followed by an interpolation stage, which helps in establishing the 3D surface of the surgical field of view (FOV). This process approximately takes 40 seconds per stereo-pair image frames, and yields an average error of 1.2mm. The external stereo camera system developed as part of [15] used the interesting approach of combining intensity information in a game theoretic approach to perform stereo camera calibration and reconstruction of 3D points from the FOV. However, this nontraditional approach makes it difficult to ascertain its accuracy for the produced 3D point clouds against traditional computer vision methods developed for stereovision. In addition, the system in [15] was tested on patients undergoing neocortical epilepsy surgery. This surgical procedure requires opening of the dura for the placement of intracranial electrodes on the cortical surface but requires no resection. The method in [15] also relies on manually delineated sulcal features on the cortical surface for extracting 3D surfaces or point clouds. This limitation is compounded by the runtime execution of the algorithm, which is on the order of 20 minutes.

In this paper, we seek to eliminate the workflow disadvantages and design issues innate to the developed intraoperative stereovision systems in [13-14, 16-17] and [15]. The proposed stereovision platform will be internal to the operating surgical microscope and will provide 3D digitization of points in the cameras' FOV in near real-time. Since the cameras are internal, the modifications and disruptions to the surgical workflow are minimal. The tLRS has been reliably used for updating the IGS system but lacks dense temporal resolution. For the stereovision intraoperative imaging system to be a viable alternative, an accuracy comparison between the points digitized by both systems is essential. The 3D point clouds extracted on realistic CAD phantoms of known dimensions from the tLRS and stereovision system will form the basis of this comparison. The main objective of this paper is to understand the tradeoffs between these two intraoperative approaches as input to an IGS system. Knowing the accuracy of the stereovision system can aid in the exploration of other real-time surgical guidance approaches to improve the utility of IGS systems. These include real-time surgical instrument tracking, cortical surface tracking, and image-to-physical registration within the surgical microscope environment. To the best of our knowledge, 3D digitized points from an internal stereovision platform native to the surgical microscope has not been compared against a laser range scanner for neurosurgical applications.

2. DATA ACQUISITION

The tLRS (Pathfinder Therapeutics, Inc., Nashville TN, USA) is capable of generating 500,000 points with a resolution of 0.15-0.2mm at the approximate range used during neurosurgery. The tLRS used in carrying out this comparison study is shown in Figure 1. These digitized points are of the form (x, y, z, rgb) and form a colored point cloud. This tLRS has been used at Vanderbilt University Medical Center (VUMC) to acquire intraoperative cortical surfaces to drive the IGS

updates [3-8]. Any commercially available surgical microscope that has an internal stereo camera system can be used with the methodologies presented in this paper. We primarily focus on the accuracy of the 3D point clouds extracted from the cameras of the OPMI® Pentero™ (Carl Zeiss, Inc., Oberkochen, Germany) microscope as it is the current microscope used in neurosurgery cases at VUMC. The internal stereo camera rig of this surgical microscope is comprised of two CCD cameras, Zeiss' MediLive® Trio™, and have NTSC (720x540) resolution with an acquisition video frame rate of 30 frames per second (fps). FireWire® Video cards at the back of the Pentero microscope are connected via cables to a desktop, which acquires video image frames from both cameras using Open Computer Vision Library or OpenCV (Willow Garage, Inc., Menlo Park, California, USA). Figure 2 shows the stereo video acquisition system for the Pentero microscope.

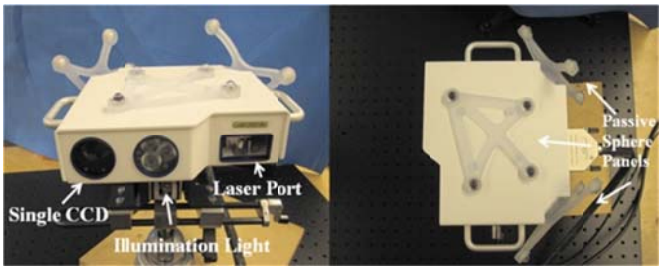

Figure 1: The tLRS made by 3D Digital, Inc. (Bethel CT, USA). This technology uniquely associates RGB color values to corresponding digitized points in its FOV.

Figure 2: The Pentero microscope as a test-bed. (Left) Microscope, (middle) the two FireWire videocards for acquisition and (right) the head of the microscope.

To verify the accuracy of the digitization of points in the FOV for the tLRS and the stereovision system, CAD design was used to create phantom models of known dimensions. The phantom object, shown in Figure 3, was rapid prototyped within a tolerance of 0.1mm vertically (EMS, Inc., Tampa, Florida, USA). The bitmap texture on this phantom object was a cortical surface from a real brain tumor surgery case. This is the kind of RGB texture expected in the FOV of the surgical microscope during neurosurgery. LEGO® (The LEGO Group, Billund, Denmark) bricks are manufactured with a tolerance level of 0.02mm with specific dimensions (LEGO, 2012) and were used to build a Lego-based phantom [18], shown in Figure 4. These two phantom objects were used to compare the accuracy between the point clouds extracted from the stereovision system of the Pentero microscope and the tLRS.

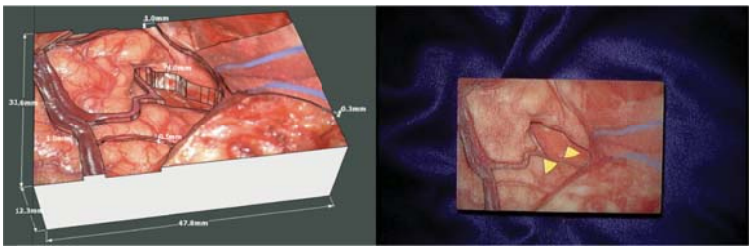

Figure 3: (Left) CAD model of a cortical surface, where the texture is from a real brain tumor surgery case, and (right) phantom in FOV of Pentero microscope.

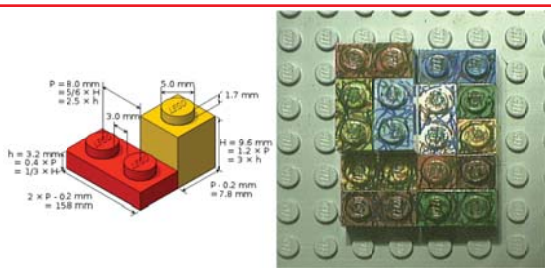

Figure 4: (Left) Lego dimensions used to build a phantom object, and (right) the textured Lego phantom under the FOV of Pentero microscope.

3. METHODOLOGY

Using a pair of cameras to form a stereovision system is a standard computer vision technique to obtain depth information of the FOV [19-20]. This requires calibrating the stereo cameras [22-23], rectifying the left and right images, and reconstructing the 3D surface using any of the techniques described in [24]. This section describes the stereovision process required to go from a 2D pixel to a 3D point. The OpenCV library [21] and Point Cloud Library or PCL (Willow Garage, Inc., Menlo Park, California, USA) are used for implementing the methods in this study.

3.1 Stereo Calibration

Camera calibration is a crucial step to determine the intrinsic and extrinsic parameters of the cameras contributing to the properties of an image in the pinhole camera model. Writing projection equations linking known coordinates of a set of 3D points and their camera image projections to solve for camera lens parameters yields metric information from 2D images. Furthermore, understanding how the two cameras relate to each other in the stereo camera system provides 3D information. For each camera, the intrinsic parameters are defined as the focal length, (f_x, f_y), the location of image-center in pixel coordinates, (c_x, c_y), the radial distortion parameters of the lens, (k_1, k_2, k_3), and tangential distortion parameters (p_1, p_2). The intrinsic parameters characterize the transformation mapping of an image point from a camera to its pixel coordinates. The extrinsic parameters describe the rigid transformation, rotation R and translation T, and this transformation brings the two cameras of the stereovision system into a common reference frame, which is designated as the left camera [19-21].

Known 3D coordinates of points can be determined by employing a calibration pattern. A 3D calibration object whose geometry is known in 3D space with great precision can be used, and such objects are composed of 2-3 orthogonal planes [22]. In addition, the 3D calibration object needs to be optically tracked to provide the method with world coordinates, and relies on intrinsic parameters supplied by camera manufacturer for an acceptable initialization of the estimation method [16-17, 22-23]. In this work, we use the flexible camera calibration technique described in [23], where a planar calibration pattern depicting a rectangular chessboard, shown in Figure 5, is used for calibrating the stereo-pair cameras. The planar calibration pattern is observed in 2 or more different poses. Unlike the technique in [22], the flexible technique of [23] does not require the use of optical trackers, a precisely machined 3D calibration object, and is computationally more efficient. The technique in [23] allows for a printed calibration pattern attached to a hard surface to be moved by hand and the motion can remain unknown. This method can be considered an offline approach to stereo calibration as it depends on 10-15 poses of the chessboard pattern for an acceptable calibration error, and does not depend on the content of the FOV. Each square of the rectangular chessboard is of size 3mm and this makes each interior corner of the chessboard 3mm apart. This provides the scale to the camera calibration stage.

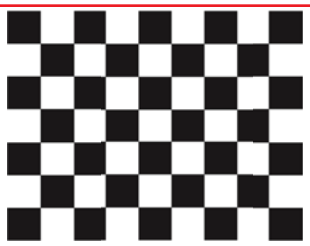

Figure 5: Rectangular planar chessboard is glued to a hard surface to create the calibration pattern seen by the left and right cameras of the stereo camera system. The shown pattern has 8x6 interior chessboard corners and the size of each chessboard square can be adjusted in metric units. We use 3mm as our chessboard square size.

The method by Zhang et al. [23], implemented in OpenCV [21], depends on 2D metric information of the chessboard's interior corners to calibrate each of the cameras in the stereo rig. The basic geometry of a stereo camera rig is known as epipolar geometry and is illustrated in Figure 6. From the pinhole model of each camera, a point P in the physical world is projected as p_l and p_r on each camera's projective planes, π_l and π_r, with centers of projection, O_l and O_r. An epipole, e_l, on the image plane π_l, represents the image of the projection center of the right camera, O_r, and vice versa. The plane identified by the viewed point P, e_l, and e_r, is defined as the epipolar plane and intersects each image plane in lines known as epipolar lines, namely, $p_l e_l$ and $p_r e_r$. The epipolar constraint specifies that a projected point in one image plane and its matching projected point in the other image plane must lie along the corresponding epipolar line of the other image plane. This means that given p_l on π_l, P can lie anywhere on the ray from O_l through p_l, and this line is projected onto π_r as a corresponding epipolar line, $p_r e_r$. The epipolar constraint maps a point in the left image as the line that goes through the corresponding point and the epipolar point in the right image. This mapping is captured in pixel coordinates, $\overline{p_l}$ and $\overline{p_r}$, and encoded as the fundamental matrix F in Equation 1. If F can be estimated using a number of point correspondences in the two image views, one can reconstruct the epipolar geometry with no prior knowledge of intrinsic and extrinsic parameters of the cameras [19-21, 23].

$$\overline{p_r}^T F \overline{p_l} = 0 \qquad (1)$$

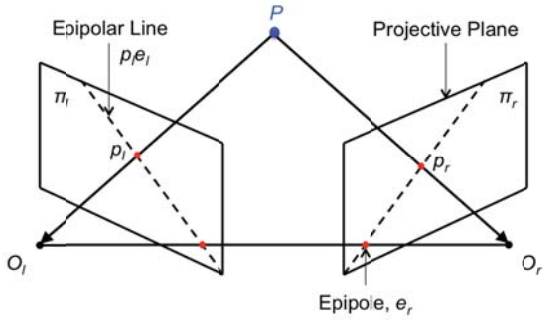

Figure 6: Epipolar geometry relates a point, P, in the FOV and the two image planes of the left and right cameras into a projective equation, shown in Equation 1.

The inner chessboard corners observed by both cameras of the stereo rig provide the needed $\overline{p_l}$ and $\overline{p_r}$ points to determine F. The image containing the chessboard is first adaptively thresholded, the inner chessboard corners are detected, and then refined with subpixel accuracy [21]. Figure 7 shows multiple views of chessboards and their detected corners. The known coordinates of the planar chessboard corners, (x, y, z=0), and the detected chessboard corners in multiple views per camera form constraints that are used in solving Equation 1 for the fundamental matrix, F [19-21, 23].

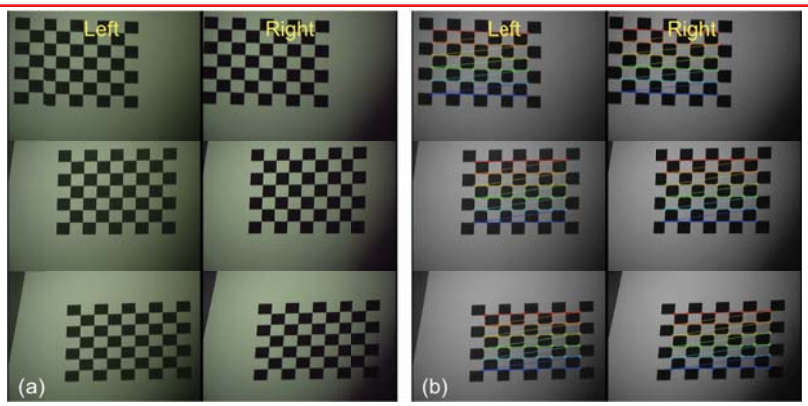

Figure 7: (a) Different poses of the chessboard seen by left and right cameras, and (b) shows the detected interior corners of the chessboard in each camera for stereo calibration stage.

3.2 Stereo Rectification

Once the fundamental matrix, F, is computed, an image transformation can now be determined that makes epipolar lines of the left and right camera images collinear and parallel to the horizontal axis. This makes the stereo-pair images appear as if they had been taken by two cameras with row-aligned image planes or in a frontal parallel configuration. This process is known as stereo rectification and results in the optical axes of the two cameras to be intersecting at infinity or be parallel to one another. In addition, stereo rectification reduces the stereo correspondence problem, which is the identification of homologous points in the left and right camera views, to a reliable and tractable 1D search on horizontal scan-lines, the collinear epipolar lines. We apply Bouguet's method for stereo rectification [21, 26-28]. This method minimizes reprojection distortions while maximizing the view overlap for both cameras. The final results are two remapped row-aligned images. An example of stereo rectification is shown in Figure 8. Furthermore, Bouguet's method yields a reprojection matrix, Q, shown in Equation 2(a). Q is employed for reprojecting a 2D homologous point and its associated disparity d to 3D by applying equation 2(b). In Equation 2(a), the principal point, (c_x, c_y), is the left (reference) image origin, c_x^r is the right image's principal point's x coordinate, f_x is the focal length of the rectified image, and T_x is horizontal translation between left and right cameras. Note that all elements of Q are in pixels.

$$Q = \begin{bmatrix} 1 & 0 & 0 & -c_x \\ 0 & 1 & 0 & -c_y \\ 0 & 0 & 0 & f_x \\ 0 & 0 & \frac{-1}{T_x} & \frac{(c_x - c_x^r)}{T_x} \end{bmatrix} \quad (2a), \qquad Q \begin{bmatrix} x \\ y \\ d \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix} \quad (2b)$$

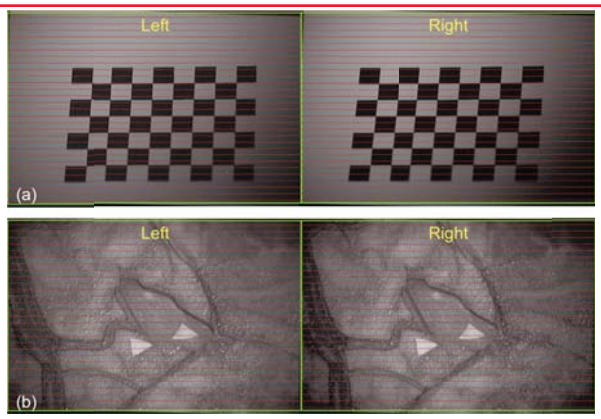

Figure 8: (a)-(b) show examples of the stereo rectification process where the epipolar lines (red) of the left and right cameras are aligned bringing the stereovision system into a frontal parallel configuration. The green rectangles indicate the FOV common to both cameras.

3.3 Stereo Reconstruction

After stereo rectification, Equation 2(b) indicates that observed 2D homologous points in the left and right images, with associated disparity d using the left image as the reference, produces the 3D coordinates $\left(\frac{X}{W}, \frac{Y}{W}, \frac{Z}{W}\right)$, where disparity $d = x^l - x^r - (c_x^l - c_x^r)$. Additionally, stereo rectification constrains the order of the search for these homologous points and, finally, the computed disparity map on the reference image yields a 3D point cloud. Several methods, real-time and non-real-time, can produce disparity maps of varying accuracy based on ground truth datasets. A full taxonomy of these stereo reconstruction methods is presented in [24]. In our work, the point clouds of the FOV generated using two real-time stereo reconstruction algorithms available in OpenCV – Block Matching (BM) and Semi-Global Block Matching (SGBM) are compared. Both techniques are limited by the need for non-repetitive textured FOV that give strong matching points between the two images. They also require that the matching points in both images are fully visible. The 3D point is undetermined if the disparity value for a pixel cannot be resolved. The resulting BM and SGBM point clouds are cleaned for any outliers via a 5x5 Median Filter. The stereo reconstruction parameters used to digitize points in the Pentero microscope's FOV are listed in Table 1 and Table 2 shows the runtimes for the steps involved in the entire stereovision process.

3.3.1 Block Matching (BM)

This real-time algorithm is based on [29] and works by employing sliding Sum of Absolute Differences (SAD) windows to find matching points between left and right stereo rectified images. BM is based on a sliding SAD window of size n_{BM}. First, the images are preprocessed to normalize image brightness and enhance texture. Then, the correspondence for a pixel in the left image is searched with a SAD window on the right image along the epipolar line over a disparity range, δ_{BM}. A uniqueness parameter, ψ_{BM}, based on the SAD ratio of two disparity candidates for a point (x, y) in the left image, is used to filter out false matches. Inherently, SAD windows have issues around boundaries of objects as the sliding window can simultaneously catch foreground and background, which results in local regions of large and small disparities or speckle. To account for speckles, a match is allowed only when the selected disparity for a point (x, y) is within a speckle range, σ_{BM}. The point clouds generated on the Lego phantom model and the cortical surface phantom model using BM are shown in Figure 9. The selection of n_{BM} is discussed in the Results section.

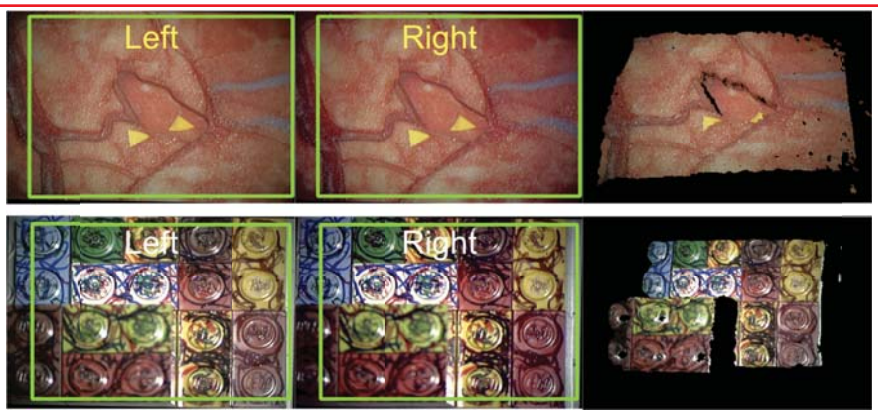

Figure 9: Block Matching (BM) stereo reconstruction results on (a) the cortical surface phantom object with n_{BM}=13, and (b) the Lego phantom object with n_{BM}=27. The respective point clouds are shown in the rightmost column. The green rectangles indicate the FOV used by BM to compute the point cloud.

3.3.2 Semi-Global Block Matching (SGBM)

A simplified variant of Semi-Global Matching with Mutual Information [30] uses Birchfield-Tomasi (B-T) subpixel cost metric [31] instead of mutual information for identifying strong matching points along epipolar lines of the rectified stereo images. B-T cost is computed in a block window size of n_{SGBM} and is aggregated. The cost aggregation step in SGBM is implemented with a multi-path dynamic programming approach, where movement from pixel to pixel, in a disparity range of δ_{SGBM}, can be parameterized as movement along a ray r. This means that a horizontal ray at a point (x, y) is denoted as ($x + r$, y) and a diagonal ray at the same point is denoted as ($x + r$, $y + r$). The cost along a ray r is aggregated in 5 directions by combining the block B-T cost over n_{SGBM} with a surface continuity constraint. The piecewise smooth and continuous nature of objects in a FOV constitutes this surface continuity constraint and is enforced by two parameters – P_1 and P_2. The parameter P_1 penalizes small changes of ±1disparity and P_2 penalizes large changes in disparity > 1. An energy function comprised of the aggregated cost and neighborhood continuity constraints is minimized to find the disparity map [30]. The lower penalty, P_1, helps adapting the energy function to slanted and curved surfaces, whereas, the higher penalty, P_2, preserves discontinuities. There is no penalty for disparities < 1. The SGBM matching is based on subpixel matching of disparities obtained by fitting a local quadratic to each disparity and its neighbors. The preprocessing steps as well as the uniqueness parameter, ψ_{SGBM}, and speckle range, σ_{SGBM}, used for BM are also used for SGBM. In general, SGBM results tend to be less noisy and smoother under slightly varying illumination conditions than BM. SGBM uses a smaller FOV than BM to compute the point cloud and this leads to clipping of points in the point cloud at the edges of the common stereo-pair FOV, indicated by green rectangles in Figure 10. Figure 10 shows the point clouds of the Lego phantom model and the cortical surface phantom model using the SGBM technique.

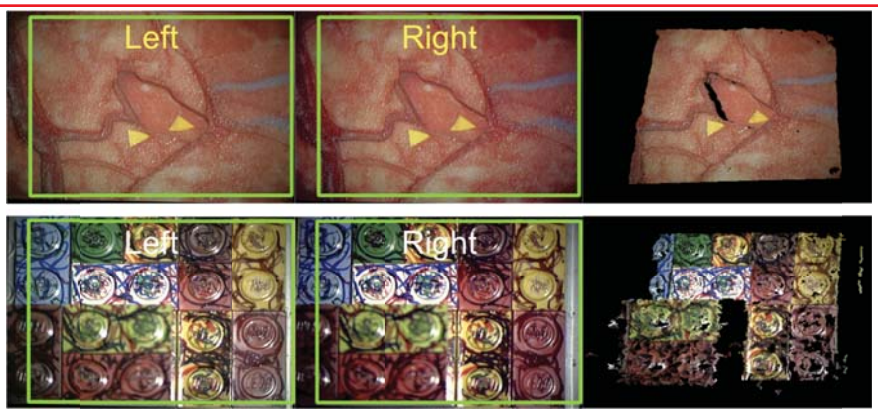

Figure 10: Semi-Global Block Matching (SGBM) stereo reconstruction results on (a) the cortical surface phantom object with n_{SGBM}=17, and (b) the Lego phantom object with n_{SGBM}=9. The respective point clouds are shown in the rightmost column. The green rectangles indicate the FOV used by SGBM to compute the point cloud. The region of interest in SGBM is smaller than BM.

Table 1: Stereo reconstruction parameters

BM		SGBM	
n_{BM}	Varied	n_{SGBM}	Varied
δ_{BM}	32	δ_{SGBM}	32
ψ_{BM}	5	ψ_{SGBM}	5
σ_{BM}	32	σ_{SGBM}	32
		P_1, P_2	8, 32

Table 2: Average Runtimes

Acquisition of Stereo Video	30 fps
Stereo Calibration (done once)	30s
Stereo Rectification (per stereo image pair)	0.06s
Stereo Reconstruction (per stereo image pair)	0.5s (BM), 1.1s (SGBM)

4. RESULTS

Stereo calibration employing multiple poses of a planar chessboard calibration pattern's interior corners needs to be accurately characterized. An acceptable stereo calibration drives the stereo rectification and reconstruction stages. Using the dot product in Equation 1, one can check how close the detected chessboard corners on one image lie on the epipolar lines of the other image. Ideally, these dot products should sum to 0, and thus, the calibration error is computed as the accumulated absolute distance. Fifteen different chessboard views were used to calibrate the stereo system with an error of 0.81 pixel^2. Typically, any subpixel stereo calibration error is regarded acceptable for continuation into the stereo rectification stage of the stereovision framework.

Stereovision and tLRS point clouds are compared to the ground truth relative depths of both phantom objects. The known relative depths, z, of the phantoms are annotated for each pixel (x, y) in the tLRS bitmap image and in the reference left camera image, which is used in stereo reconstruction. Both the stereovision and tLRS point clouds intrinsically keep their mappings from a pixel to its 3D point. Root mean square error (RMS) is computed between all the points in a BM and SGBM point cloud to its respective ground truth z values. The variable parameter for BM and SGBM are n_{BM} and n_{SGBM} respectively. The window block-size for each algorithm, n_{BM} and n_{SGBM}, directly affects the computation of the disparity values. A smaller block-size yields finer disparity values depicting more details of the FOV but tends to be noisier. A larger block-size produces smoother disparity values at a loss of detail and is less noisy. For the Lego phantom model and the cortical surface phantom model, this parameter is swept and the RMS is computed, and compared to the ground truth. The RMS values for each window block-size for both phantoms are graphed in Figure 11 against the RMS values of the tLRS. Table 3 shows the optimal window block-size found for BM and SGBM for each phantom and the respective RMS error values. Figure 12 shows the tLRS bitmap and point cloud for the cortical surface and the Lego phantom objects. Absolute-deviation-based error heat maps of the cortical surface and Lego phantom objects are computed for the LRS, BM and SGBM stereo reconstruction methods. They are presented in Figure 13. The n_{BM} and n_{SGBM} that gave the smallest RMS values were picked to compute these error heat maps. These error heat maps contrast the stereovision system against the tLRS system.

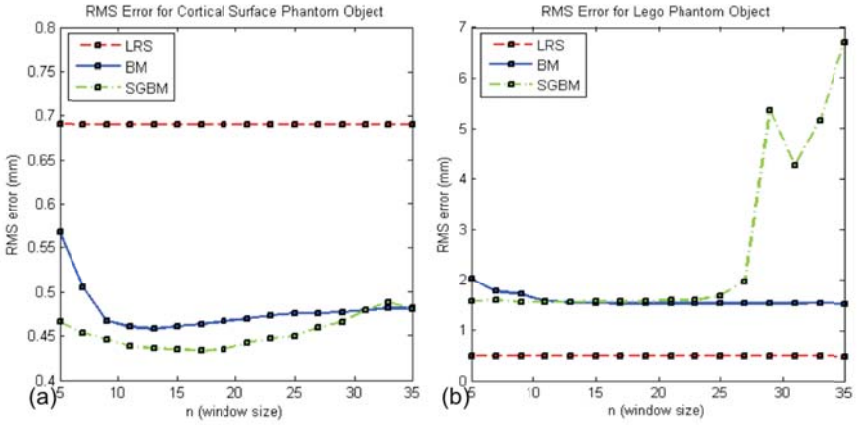

Figure 11: RMS error for varying window block-sizes, n_{BM} and n_{SGBM}, for BM and SGBM stereo reconstruction methods for (a) cortical surface phantom object and (b) Lego phantom object. The optimal values of n_{BM} and n_{SGBM} are shown in Table 3.

	Table 3: RMS Errors (mm)				
Phantom	n_{BM}	BM	n_{SGBM}	SGBM	tLRS
Cortical	13	0.4574	17	0.4328	0.6898
Lego	27	1.5171	9	1.5460	0.4731

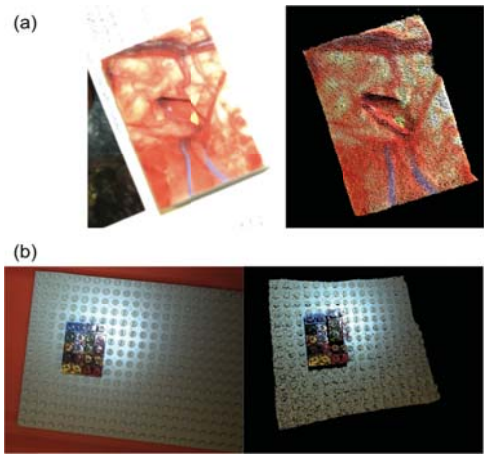

Figure 12: (Left column) the bitmaps captured by the tLRS, and (right column) the point clouds acquired by tLRS for (a) the cortical surface phantom and (b) the Lego phantom.

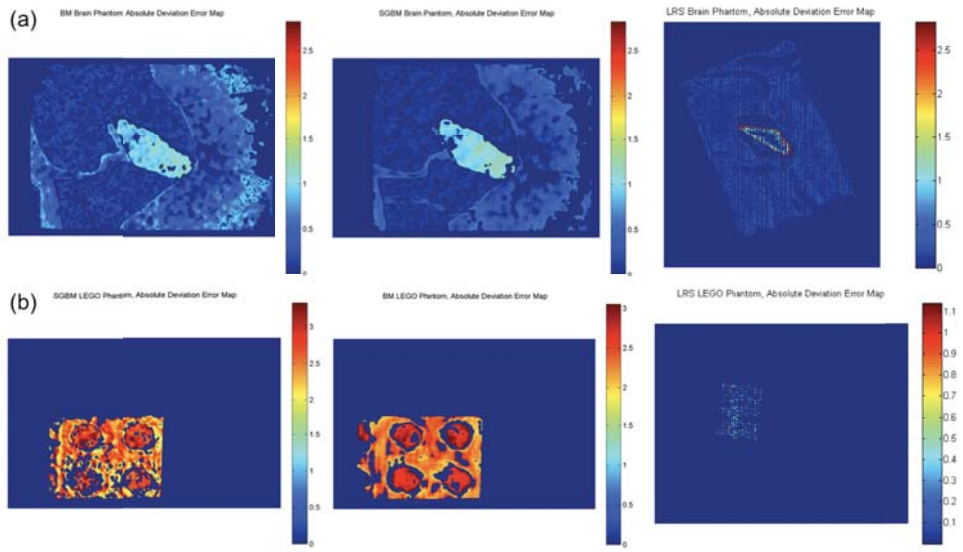

Figure 13: Absolute deviation error heat maps for (a) the cortical surface phantom and (b) the Lego phantom objects. The limitations of the (left column) BM and (middle column) SGBM stereo reconstruction techniques can be seen on their error heat maps. The gold standard for comparison is the error heat map of the (right column) tLRS.

The RMS errors for the BM and SGBM point clouds are similar. For the purpose of developing a real-time stereovision system, BM is the better candidate for stereo reconstruction because its runtime is twice as fast than SGBM's runtime (see Table 2). The stereovision system lacks accuracy in estimating abrupt transitions or boundaries and topology that is far away from the cameras' image planes. Boundaries of objects in the left-camera may be occluded in the right-camera and this causes stereo reconstruction to be inaccurate around boundaries. This is illustrated in Figure 13 at sharp transitions around the resection area of the cortical surface phantom for BM and SGBM techniques. The discontinuity step and the many knobs of the Lego phantom cause occlusions in the stereo reconstruction stage and this results in a higher RMS error. Non-real-time stereo reconstruction approaches address the occluded boundary problem [24]. This issue is not a critical limitation for neurosurgery applications because the cortical surface is relatively smooth when compared to the abrupt edges in the Lego phantom for example. Stereovision point clouds could thus be accurate enough to drive a patient-specific biomechanical model.

5. CONCLUSION

In this paper we proposed an internal stereovision platform, native to the surgical microscope, capable of 3D digitization points in the FOV. Although the stereo calibration is done offline, the digitization of surface points in the FOV, also called the point cloud, is acquired in near real-time. A comparison study was conducted to determine the accuracy with which the 3D coordinates of the cortical surface could be estimated with the stereovision camera system of the surgical microscope. On our phantoms, we show accuracy in the 0.46-1.5mm range. While not as good as what can be achieved with the tLRS, the temporal resolution this approach affords makes it an attractive alternative for driving the patient-specific biomechanical brain model in cases where brain shift is on the order of 2cm, as observed in tumor resection surgeries. The accuracy of the stereovision system within the microscope justifies the exploration of the system to other real-time guidance methodologies in neurosurgery such as surgical instrument tracking.

6. ACKNOWLEDGEMENTS

National Institute for Neurological Disorders and Stroke Grant R01-NS049251 funded this research. We acknowledge the Vanderbilt University Medical Center, Department of Neurosurgery, EMS Prototyping, Carl Zeiss Inc., and The Armamentarium, Inc. (Houston TX, USA) for providing the needed equipment.

REFERENCES

[1] Nimsky, C., Ganslandt, O., Cerny, S., Hastreiter, P., Greiner, G., and Fahlbusch, R. (2000). Quantification of, visualization of, and compensation for brain shift using intraoperative magnetic resonance imaging. *Neurosurgery* 47, 1070-9.

[2] Hartkens, T., Hill, D. L. G., Castellano-Smith, A. D., Hawkes, D. J., Maurer, C. R., Martin, A. J., Hall, W. A., Liu, H., and Truwit, C. L. (2003). Measurement and analysis of brain deformation during neurosurgery. *IEEE Transactions on Medical Imaging* 22, 82-92.

[3] Sinha, T. K., Dawant, B. M., Duay, V., Cash, D. M., Weil, R. J., Thompson, R. C., Weaver, K. D., and Miga, M. I. (2005). A method to track cortical surface deformations using a laser range scanner. *IEEE Transactions on Medical Imaging* 24, 767-781.

[4] Sinha, T. K., Miga, M. I., Cash, D. M., and Weil, R. J. (2006). Intraoperative cortical surface characterization using laser range scanning: preliminary results. *Neurosurgery* 59, 368-376.

[5] Dumpuri, P., Thompson, R. C., Dawant, B. M., Cao, A., and Miga, M. I. (2007). An atlas-based method to compensate for brain shift: preliminary results. *Medical Image Analysis* 11, 128-145.

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

[7] Pheiffer, T.S., Simpson, A.L., Lennon, B., Thompson, R.C., and Miga, M.I. (2012). Design and evaluation of an optically-tracked single-CCD laser range scanner. *Medical Physics* 39, 2, 636-642.

[8] Ding, S., Miga, M. I., Noble, J. H., Cao, A., Dumpuri, P., Thompson, R. C., and Dawant, B. M. (2009). Semiautomatic registration of pre- and postbrain tumor resection laser range data: method and validation. *IEEE Transactions on Biomedical Engineering* 56, 770-780.

[9] Ding, S., Miga, M. I., Pheiffer, T. S., Simpson, A. L., Thompson, R. C., and Dawant, B. M. (2011). Tracking of vessels in intra-operative microscope video sequences for cortical displacement estimation. *IEEE Transactions on Biomedical Engineering* 58, 1985-1993.

[10] Skrinjar, O.M., Tagare, H., and Duncan, J. (2000). Surface growing from stereo images. *IEEE Conference on Computer Vision and Pattern Recognition* (CVPR 2000) 2, 2571-2576.

[11] Skrinjar, O.M., Studholme, C., Nabavi, A., and Duncan, J.S. (2001). Steps toward a stereo-camera-guided biomechanical model for brain shift compensation. In *Proceedings of the Seventeenth International Conference on Information Processing in Medical Imaging* (IPMI 2001), M.F. Insana and R.M. Leahy, eds., 183-189.

[12] Skrinjar, O., and Duncan, J. (2002). Stereo-guided volumetric deformation recovery. *IEEE International Symposium on Biomedical Imaging* (ISBI 2002), 863-866.

[13] Sun, H., Lunn, K. E., Farid, H., Wu, Z., Roberts, D. W., Hartov, A., and Paulsen, K. D. (2005). Stereopsis-guided brain shift compensation. *IEEE Transactions on Medical Imaging* 24, 1039-1052.

[14] Sun, H., Roberts, D. W., Farid, H., Wu, Z., Hartov, A., and Paulsen, K. D. (2005). Cortical surface tracking using a stereoscopic operating microscope. *Neurosurgery* 56, 86-97.

[15] DeLorenzo, C., Papademetris, X., Staib, L. H., Vives, K. P., Spencer, D. D., and Duncan, J. S. (2010). Image-Guided Intraoperative Cortical Deformation Recovery Using Game Theory: Application to Neocortical Epilepsy Surgery. *IEEE Transactions on Medical Imaging* 29, 322-338.

[16] Ji, S., Fan, X., Roberts, D. W., Hartov, A., and Paulsen, K. D. (2010). An integrated model-based neurosurgical guidance system. *Proceedings of SPIE Medical Imaging 2010: Visualization, Image-Guided Procedures, and Modeling*, 7625, 762535.

[17] Fan X., Ji S., Fontaine K., Hartov A., Roberts D.W., and Paulsen K.D. (2011). Simulation of Brain Tumor Resection in Image-guided Neurosurgery. *Proceedings of SPIE Medical Imaging 2011: Visualization, Display and Image-Guided Procedures*, 7964, 79640.

[18] LEGO. (2012). LEGO: A short presentation 2011. The LEGO Group. http://bit.ly/q6fBqX.

[19] Hartley, R., and Zisserman, A. (2004). Multiple View Geometry in Computer Vision. (Cambridge University Press).

[20] Trucco, E., and Verri, A. (1998). Introductory techniques for 3-D computer vision. (Prentice Hall).
[21] Bradski, G., and Kaehler, A. (2008). Learning OpenCV: Computer Vision with the OpenCV Library. M. Loukides and R. Monaghan, eds. (O'Reilly Media).
[22] Tsai, R. (1987). A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal on Robotics and Automation* 3, 323-344.
[23] Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, 1330-1334.
[24] Scharstein, D., and Szeliski, R. (2002). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision* 47, 7-42.
[25] Paul, P., Morandi, X., and Jannin, P. (2009). A surface registration method for quantification of intraoperative brain deformations in image-guided neurosurgery. *IEEE Transactions of Information Technology in Biomedicine*, vol. 13, no. 6, pp. 976-983.
[26] Siciliano, B., and Khatib, O. (2008). Handbook of Robotics. (Springer).
[27] Bouguet, J.-Y. (1999). Visual methods for three-dimensional modeling. *Ph.D. Thesis*, California Institute of Technology.
[28] Bouguet, J.-Y. (2006). Camera calibration toolbox for Matlab. http://www.vision.caltech.edu/bouguetj/calib_doc/.
[29] Konolige, K. (1997). Small vision system: hardware and implementation. *Proceedings of Eighth International Symposium on Robotics Research*, 111-116.
[30] Hirschmuller, H. (2008). Stereo Processing by Semiglobal Matching and Mutual Information. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30, 328-41.
[31] Birchfield, S. and Tomasi, C. (1998). Depth discontinuities by pixel-to-pixel stereo. *International Journal of Computer Vision* 35, 3, 269-293.