2020 IEEE International Conference on Robotics and Automation (ICRA)
31 May - 31 August, 2020. Paris, France

Evaluation of Increasing Camera Baseline on Depth Perception in Surgical Robotics

Apeksha Avinash[1], Alaa Eldin Abdelaal[1] and Septimiu E. Salcudean[1]

Abstract— In this paper, we evaluate the effect of increasing camera baselines on depth perception in robot-assisted surgery. Restricted by the diameter of the surgical trocar through which they are inserted, current clinical stereo endoscopes have a fixed baseline of 5.5 mm. To overcome this restriction, we propose using a stereoscopic "pickup" camera with a side-firing design that allows for larger baselines. We conducted a user study with baselines of 10 mm, 15 mm, 20 mm, and 30 mm to evaluate the effect of increasing baseline on depth perception when used with the da Vinci surgical system. Subjects (N=28) were recruited and asked to rank differently sized poles, mounted at a distance of 200 mm from the cameras, according to their increasing order of height when viewed under different baseline conditions. The results showed that subjects performed better as the baseline was increased with the best performance at a 20 mm baseline. This preliminary proof-of-concept study shows that there is opportunity to improve depth perception in robot-assisted surgical systems with a change in endoscope design philosophy. In this paper, we present this change with our side-firing "pickup" camera and its flexible baseline design. Ultimately, this serves as the first step towards an adaptive baseline camera design that maximizes depth perception in surgery.

I. Introduction

In robot-assisted surgery, incisions are made into a patient's body through which robotic instruments and an endoscopic camera are inserted. Guided by the images from this endoscope, the surgeon teleoperates these instruments to perform surgical procedures. For improved perception of the surgical field, modern robot-assisted surgical platforms use stereoscopic cameras to provide 3-dimensional (3D) images to the surgeon which is a vast improvement over the 2-dimensional (2D) viewing conditions typically found in laparoscopic surgery [1].

The da Vinci surgical system™ (Intuitive Surgical, CA) provides this 3D vision by presenting the left image of a stereoscopic pair to the left eye and the right image to the right eye, and relies on the user's brain to fuse the two to obtain depth information. This system attempts to model the actual phenomenon of human vision in which each eye sees a slightly different perspective of the same scene [2]. On average, the left and right eyes are separated by an inter-pupillary distance (IPD) of 63.5 mm. The brain uses the disparity in the two images received, i.e., the horizontal shift between the projection of the same object in both images, to infer the depth of that object from the viewer. Similarly, stereoscopic systems use stereoscopic cameras set at a fixed distance apart to capture these two views and present them separately to each eye.

To facilitate correct and comfortable fusion for 3D vision, one must consider various parameters: the inter-camera separation or camera baseline, the camera-to-object distance, the viewer-to-display screen distance, the camera convergence angle, the camera field of view, and the display screen width, to name a few [3].

In robot-assisted surgery, the endoscopic camera must fit through standard surgical trocar sizes; the largest commonly used is 12 mm. For safety, it is advantageous for the endoscope to guide its own insertion, and hence the stereo cameras are typically placed at the tip of the endoscope [4]. Thus, the largest separation between the cameras, or the camera baseline, is constrained to be less than 12 mm [4]–[6]. Current clinical endoscopes, including the one used with the da Vinci system, typically have a fixed baseline of 5.5 mm [7].

In our previous work, we introduced a stereoscopic "pickup" camera with six degrees of freedom, two more than current rigid endoscopes, that provides a larger number of positions and angles with which the surgical site can be viewed [8]. With a low-resolution camera at its tip to guide insertion into the body, the "pickup" camera is then grasped and manipulated through a da Vinci instrument. A significant advantage of our design is the side-firing position of the main viewing stereo cameras, with an increased and possibly changeable baseline between the left and right cameras, irrespective of the trocar diameter. In this paper, we explore the advantage of this increased baseline.

The main contributions of this paper are as follows:

- This is the first study to evaluate the effect of increasing camera baselines on depth perception in near-field applications such as robot-assisted minimally invasive surgery.
- We show how the design of the "pickup" camera allows having multiple baselines unlike the original endoscopic design used in current robot-assisted surgical platforms.
- We propose the use of a real validated surgical training task for the evaluation of depth perception in stereoscopic systems in contrast to past evaluations

---

Financial support for this work is gratefully acknowledged and is provided by Professor Salcudean's C.A. Laszlo Chair, by the Natural Sciences and Engineering Research Council of Canada, and by the Vanier Canada Graduate Scholarship held by A.E. Abdelaal.

[1] A. Avinash, A.E. Abdelaal and S.E. Salcudean are with the Electrical and Computer Engineering Department, University of British Columbia, 2332 Main Mall, Vancouver, BC Canada apeksha@ece.ubc.ca

978-1-7281-7395-5/20/$31.00 ©2020 IEEE

of depth perception where virtual reality is employed for the same purpose.

II. Related Work

In the broader field of teleoperation, studies show that an increase in the camera baseline leads to an improvement in performance [9]–[11]. [9] studies the effect of an increased baseline on the performance of virtual tasks. Subjects are asked to align a virtual control peg with a virtual target peg placed at a certain distance from the display screen (along the camera depth or z-axis). The authors report a rapid improvement in performance when increasing the baseline from 0 (2D) to 20 mm and an asymptotic increase to a maximum at 30 mm. No measurable improvement is found on further increases of the baseline up to 80 mm. [10] explores the effect of various camera baselines for head mounted displays (HMDs). In this study, subjects are asked to perform a task with their hands but using the HMD to view the scene. Since the working distance is comparable with that of direct viewing working distances, the range of baselines tested is larger. They report that depth perception is reasonable for baselines between 40-130 mm, but is still ideally 63.5 mm which is the IPD. [11] records the time taken to recognize objects at various camera baselines, and the accuracy of the recognition. The images are displayed via two polarized monitors and viewed via a hood fitted with a polarizing filter. Here too, larger baselines lead to shorter recognition times and higher recognition accuracy.

While general rules of thumb have been presented in the literature for baseline estimation, most research has resorted to the trial-and-error method of finding the most suitable baseline for their application. To achieve comfortable stereo, the application and its setup parameters must be considered [3]. The viewer-to-display screen distance in the da Vinci surgical system is approximately 45 cm [12] which differs from the distances used previously (80 cm in [9] while [10] uses an HMD). Although this distance is similar to that reported in [11], the distances to the target object used in that study are much larger: 200-500 m from the cameras (i.e., far-field working distance). The working distance in minimally invasive surgery ranges from 10-35 cm, which is smaller than any of the working distances in the studies mentioned above. Clearly, the camera setups and working environments differ across all the above mentioned experiments and hence we cannot directly translate their results to robot-assisted surgery.

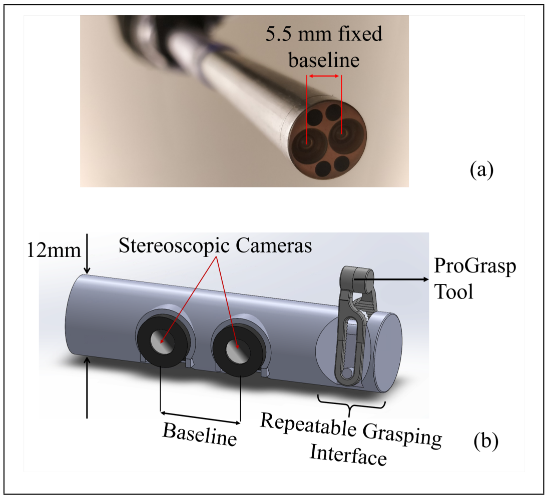


Fig. 1: (a) Current clinical endoscope with a fixed camera baseline used with the da Vinci surgical system. (b) Conceptual stereoscopic “pickup” camera.

For applications such as surgery, depth information in the near-field is of high importance [13]. The major design constraint for stereo endoscopes in surgery is the port diameter. The current standard is 10-12 mm which helps keep the wound sizes small [14]. However, this negatively impacts the possible camera baselines. This motivated [14] to evaluate surgical performance when the camera baseline is reduced. The authors report a clear decline in performance with camera baselines of less than 5.5 mm, but argue that there is still stereoscopic information provided with reduced baselines. This is a strong claim to make, however, since it is also mentioned that subjects likely used other monocular depth cues (like shadows and visual feedback of contact with tools) to successfully complete the experimental task.

It is important to note that the work in [14] is driven by a significantly different goal: an evaluation of miniaturization driven by the constraint of the trocar diameter. Rather than settle for smaller baselines that only slightly compromise performance, we instead strive to find the optimum baseline at which maximum performance can be extracted. The end-firing design of the endoscope has so far remained constant throughout the last few decades [5] and serves as the main obstruction to an increase in the camera baseline. In this work, we do not face the same constraint with our custom stereoscopic camera. Designed to place the stereoscopic cameras along the side of a cylindrical body, the overall diameter of the camera still remains less than 12 mm for easy insertion into the trocar, as seen in Fig. 1b. With this design, we have the freedom to increase the baseline as desired.

III. Methods

A. Creating Stereo Images

In stereoscopy, depth information is generated from the disparity between corresponding projections of the same object in the left and right rectified images of a stereo pair. This disparity can be generated with a camera system by displacing the left and right cameras horizontally by a distance known as the baseline. It has become the standard to keep the axes of these two cameras parallel for such stereoscopic systems [15]. Converged or “toe-in” cameras are not recommended as they introduce keystoning in the images when displayed and these vertical conflicts cause discomfort and strain to the user [15]. With parallel axes, points at infinity will have zero disparity and will appear on the plane of the display. To correctly view points at all other distances from the camera, these left and right images must be shifted horizontally and this is termed as Horizontal Image Translation (HIT) [16].

The images are translated horizontally until the desired target object is at the same position in both images. The images are then said to be "aligned", and the target object will now appear at the plane of the display. Any object in front of the target will appear to be in front of the display and any object behind the target will appear to be behind the display. Depending on the distance of the target object from the cameras as well as the baseline of the cameras, the amount of translation or shift required to produce aligned images will differ.

The da Vinci S surgical system has a built-in 'scope alignment' function that allows the user to manually implement HIT [17]. Since current stereo endoscopes have a fixed baseline, the only variable parameter to consider for HIT is the distance to the target object. A calibration object is used to determine the camera-to-target distance before surgery and this alignment is set for the duration of the procedure.

For our work, we project images from various baselines onto the da Vinci S surgeon console. To project stereo images correctly, we had to: firstly, nullify the effect of the da Vinci's built-in alignment which is calibrated to the endoscope's baseline, and, secondly, implement HIT for our own images. We nullified the da Vinci's alignment by pushing 2D images to the console (the same image to the left and right displays) and used the 'scope alignment' feature to manually adjust the two images so that they correspond exactly. This essentially reduces any previously set horizontal translation to zero. Next, we use our own reference target object to horizontally shift our images. This target is placed around 200 mm away from the cameras. The amount of shift required for each baseline differs and is recorded. These recorded shifts are then used to correctly display image pairs from each baseline so that the target object appears to be on the plane of the display irrespective of the baseline used.

B. Hardware

With an outer diameter of 12 mm, our stereoscopic "pickup" camera can easily be inserted through a surgical trocar and then picked up and grasped by a da Vinci instrument [8]. The side-firing design places the cameras on the side of the cylindrical body (Fig. 1b), as opposed to the end-on design currently used by commercially available endoscopes (Fig. 1a). This allows us to keep the overall diameter of 12 mm for easy insertion inside the insufflated patient. For this proof-of-concept study, we used a commercial stereo camera LI-OV580-OV9782ST from Leopard Imaging. Due to the physical size of the camera heads, the smallest baseline achievable is 10 mm.

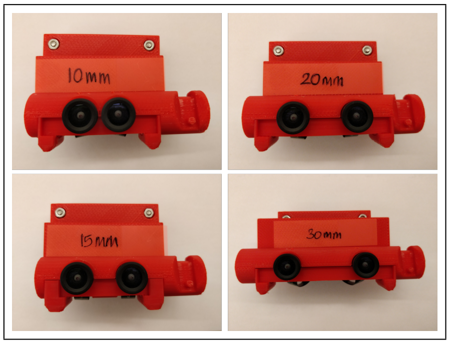


Fig. 2: 3D printed camera probes with various baselines as indicated above.

We used four different baseline values: 10 mm, 15 mm, 20 mm, and 30 mm. Probes with different baselines were 3D printed for the experimental study as shown in Fig. 2. The prototypes shown are bulkier to accommodate the electronics of the camera but the final design will have the electronics embedded within the cylindrical body.

C. Experimental Setup and Task

We conducted our experiments using the da Vinci S surgical system. Held by a ProGrasp tool, the "pickup" camera was positioned at a height of 200 mm, providing a top-down view of the task as shown in Fig. 3. Images from the camera were sent directly to the surgeon console through a DeckLink card (BlackMagic Design [18]), whose system is briefly described in [19].

The task was a modified version of the "Pea on a Peg" task which is part of a validated training curriculum for laparoscopic surgery [20]. The original task's objective is to pick up small-sized beads and place them on top of poles of varying heights. We noticed, however, that contact made between the beads and the poles served as a strong depth cue. This sort of "touch" provides the user with a sense of where the top of the pole is with respect to their tool, thus contributing to depth perception.

For our work, we are interested in the effect an increased baseline has on depth perception. To avoid the effect of any other depth cues, we modified the "Pea on a Peg" task's objective to be a purely perceptual one. Depth judgements along the axis of the camera's focal point (z-axis) are the most challenging [21] and we targeted this direction with our setup. Eight differently sized poles were placed in eight different positions on the board. The occupied positions on the board were labelled from A to H (starting from left to right, and continuing from top to bottom). The objective was to write down the sequence of letters representing the shortest pole to the tallest pole. The heights of the poles varied from 10 mm to 25 mm with increments of 2.5 mm as seen in Fig. 4a. The depth resolution required to resolve such small changes in height would have to be high for subjects to estimate the correct sequence. Subjects did not have to interact with the board or the poles thereby

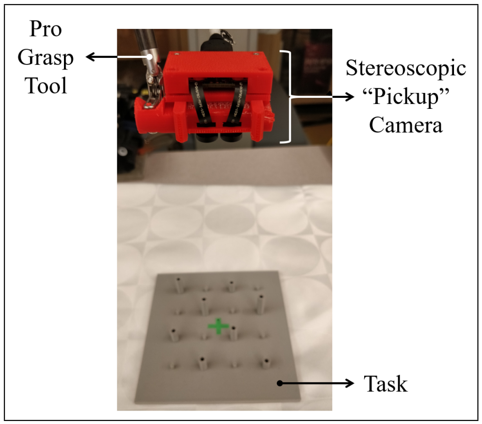
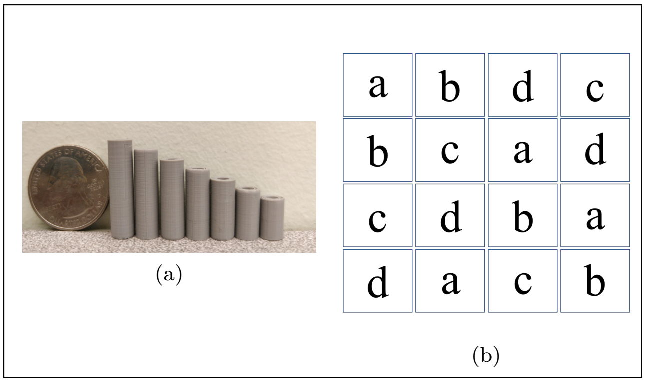
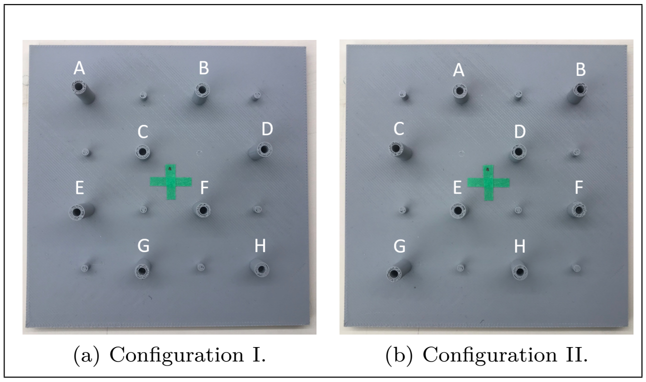


Fig. 3: Experimental Setup.



a	b	d	c
b	c	a	d
c	d	b	a
d	a	c	b

Fig. 4: (a) Poles of heights varying from 10 mm to 25 mm with increments of 2.5 mm were used. Coin shown for scale. (b) Balanced Latin Square for four baseline cases. Each letter represents a baseline: a: 10 mm, b: 15 mm, c: 20 mm, d: 30 mm. Each row represents the order in which an individual subject experiences the four cases. Consecutive subjects will follow consecutive rows and the entire block is repeated once the last row is reached.



(a) Configuration I. (b) Configuration II.

Fig. 5: Board Configurations.

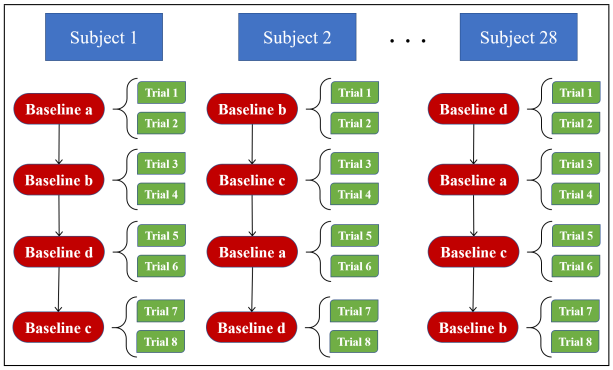


Fig. 6: Experimental Design.

eliminating any cues gained from making contact. Any other monocular cues such as shadows would remain constant across all the baselines as the lighting was maintained constant.

D. Experimental Design

Since depth perception is subjective and varies across people, a within-subject study was chosen to minimize the effect of an individual's intrinsic perception. The four cases (10 mm, 15 mm, 20 mm, 30 mm) were counterbalanced across subjects using the balanced Latin square [22] depicted in Fig. 4b. This dictated the order in which each subject faced the different cases.

The perceived height of an individual pole can vary based on its position in the image captured by the "pickup" camera. To cover all 16 positions on the board, the poles were arranged in one of two different configurations as shown in Fig. 5. The letters A-H represent each position on the board. Out of all the possible combinations available when the poles are placed on the board, eight different combinations were randomly selected. Four of these combinations were arranged in Configuration I, and four in Configuration II.

A single subject's experimental session consisted of eight trials—he/she was shown the task twice under each camera baseline condition. Since the same combination of poles cannot be displayed every time, a different combination was randomly chosen from the set of eight combinations prepared in advance. Fig. 6 illustrates the study's experimental design. All subjects were thus presented with each combination exactly once during their session.

Each subject was scored based on their answer. The score was computed by measuring the Hamming distance of the subject's answer from the correct sequence. Hamming distance measures the number of individual changes required to transform one word into another [23]. So for example, if the correct sequence is 'ADCHBFG' and the subject's answer is 'ACDHFGB', the computed score is 5 (the changes are A$\underline{\text{CD}}$H$\underline{\text{FGB}}$). The lower the score, the closer the answer is to the right sequence.

To smoothly conduct the study, the left and right images of each combination of the board were captured using each baseline prior to testing. A virtually drawn outline of the board and its center helped position the "pickup" camera directly overheard, with its body parallel to the board and its line of sight at approximately 90°

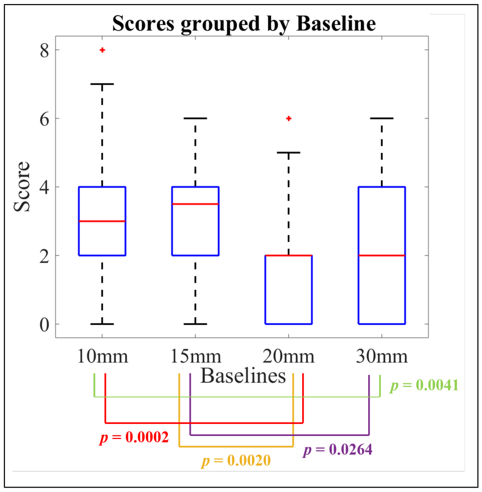


Fig. 7: Box plot of scores per baseline condition. Statistical significance is achieved between the following baselines: (1) 10 mm and 20 mm, (2) 10 mm and 30 mm, (3) 15 mm and 20 mm, and (4) 15 mm and 30 mm. The respective p-values are marked in the figure.

to the board. Board combinations were randomly chosen from the set of eight for each participant during the study, and the corresponding images were loaded onto the console after aligning with HIT. This also ensured that the same images were seen across all subjects. The randomization of the board combinations was to ensure that no particular baseline was biased with an "easy" or "hard" board.

IV. Results

A pilot study was conducted with 12 subjects to determine the required sample size for statistical significance. Fully informed consent was obtained from all participating subjects with approval from the University of British Columbia (UBC) Research Ethics Board (Study Number: H18-01845). The G*Power application [24] was used to compute the desired sample size. With an error probability α of 0.05 and a desired power of $(1 - \beta = 0.90)$, the desired sample size was determined to be 25.

To minimize the effects of any learning bias, we recruited 28 subjects to maintain counterbalancing across the entire set of subjects. As can be seen in Fig. 7, lower scores are achieved with an increase in the baseline with the lowest score on average attained with 20 mm. Two-sample t-tests between each of the groups were conducted and statistical significance was achieved $(p < 0.05)$ between the following baseline conditions: (1) 10 mm and 20 mm, (2) 10 mm and 30 mm, (3) 15 mm and 20 mm, and (4) 15 mm and 30 mm.

Fig. 8 shows the distribution of average scores across all subjects for each combination of poles. The scores are more or less consistent, which implies that there is no bias with respect to any combination, i.e., no combination of poles used is a particularly "hard" or "easy" combination

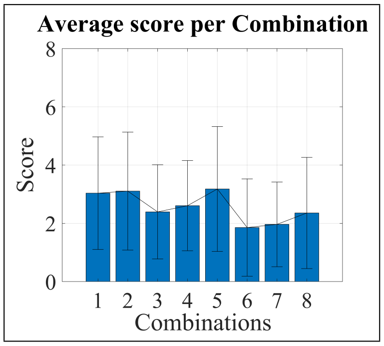
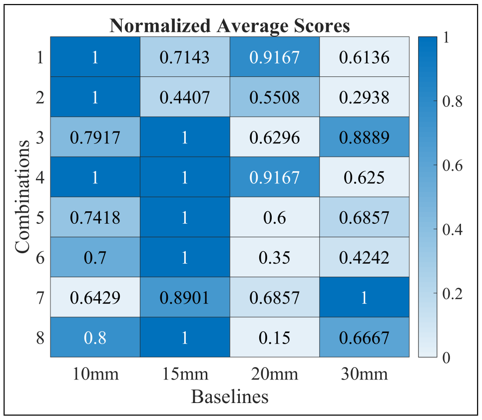


Fig. 8: Average scores across subjects for each combination of poles. There was no statistically significant difference between any of the combinations.



	10mm	15mm	20mm	30mm
1	1	0.7143	0.9167	0.6136
2	1	0.4407	0.5508	0.2938
3	0.7917	1	0.6296	0.8889
4	1	1	0.9167	0.625
5	0.7418	1	0.6	0.6857
6	0.7	1	0.35	0.4242
7	0.6429	0.8901	0.6857	1
8	0.8	1	0.15	0.6667

Fig. 9: Heatmap of normalized average scores across all baselines and combinations of poles. The lighter shades of blue indicate lower scores and better performances. Normalized along each row, the lowest score in the row maps to the lightest shade and the highest score to the darkest shade. The domination of light blue boxes on the right half shows that the higher baselines have almost always performed better.

regardless of baseline. This was important to ensure that the only variable to affect performance is the baseline.

Fig. 9 displays a heatmap of average scores achieved per baseline per combination of poles. Each row represents a single combination of poles while each column represents the baseline condition. Lower scores indicate better performance and this is represented by lighter shades of blue. The normalized values of the average scores are written within the individual boxes. The right half of the map is dominated by lighter blues which shows that better scores were achieved at higher baselines.

V. Discussion

As the literature suggests, it was not surprising to see that depth perception improved with larger baselines. There is a sharp decline from 15 mm to 20 mm as can be seen in Fig. 7 but the scores with 30 mm are similar; the 20 mm and 30 mm groups were not statistically different. The important contribution of this paper, however, is that our study confirms that there is an improvement with larger baselines and that this can easily be achieved with our "pickup" camera design. Current endoscopes are restricted by their design philosophy and hence the focus of current research has not been on evaluating larger baselines. A shift in the design philosophy, which may offer an end-firing camera for insertion and two side-firing cameras for endoscopic view opens doors to better depth perception in robot-assisted surgery.

Treating this as a preliminary study, we chose only four baseline conditions. It would be interesting to see the trend on increasing the baseline further than 30 mm—would performance continue to plateau or would there be a gradual decline as the baseline becomes too large to achieve comfortable stereo for such small working distances? We also note that a further increase in baseline would require using a longer camera which would be limited by the available abdominal space and may vary across surgical procedures and across patient anatomy.

Another opportunity to further expand on the results of this study is to evaluate performance at different camera-to-target distances. For this study, this distance was fixed to 200 mm, which the authors believed served as an approximate average of surgical working distances. It is entirely possible that the same baseline does not perform the best at a smaller or larger distance. Evaluating the performance at varying distances would be the first step towards building an adaptive baseline camera system that would dynamically adjust the camera baseline based on the camera-to-target object distance. Such a system would also serve useful for tracking and/or performing any 3D reconstruction of the surgical field [25].

In our work, we placed the task at the center of the image with the intention that the surgeon's focus of attention will likely be concentrated at the center. It would be interesting to see how peripheral depth perception is affected with an increase in the baseline. While the camera is usually centered around the object of interest, any critical event that occurs in the surgeon's peripheral vision must be brought to their attention and must not go unnoticed. Similarly, the size of the task may affect the depth perceived through the various baselines. A follow-up study to analyze a change in task size may be another possible direction.

A shortcoming of this study is the lack of comparison to the current gold standard: the endoscopic baseline. There were two reasons for this: firstly, the commercial stereo cameras used in this study were physically restricted to baselines larger than 10 mm only. Secondly, we could not use the endoscope itself to compare against the performance with these cameras as the resolution and image quality varied vastly between the two. Depth perception is impacted by image quality and this would not have been a fair comparison. We also note that the effect of magnification in this setup was not studied. The clinical da Vinci system has been reported to have a digital magnification of 10-15× [26], [27]. In the future, it will be necessary to validate the results of this study with a direct comparison to the endoscopic baseline by either utilizing smaller cameras or better quality cameras that can merit a direct comparison.

Lastly, the next logical step for this study is to evaluate the effect of an increased baseline on surgical performance. For this study, we isolated the environment from other cues that could contribute to depth perception. But this is not the actual case with real surgery. Manipulation of tools introduces other depth cues such as "touch" or contact, motion parallax, and moving shadows to name a few. It will be interesting to see the impact a change in baseline has with the addition of these cues and in an environment more closely simulating that of surgery.

VI. Conclusion and Future Work

In this paper, we evaluated the effect an increased camera baseline would have on depth perception in surgical robotics. Current endoscopes are restricted by the diameter of the incision made on the patient's body and thus have a small fixed baseline of 5.5 mm. We propose the use of a custom stereoscopic "pickup" camera with a larger baseline to circumvent this limitation. With our side-firing design, the cameras are placed laterally along the axis of the cylindrical body, thereby maintaining the overall diameter of 12 mm to fit through the surgical trocar. We conducted an experimental user study with 28 subjects to evaluate the effect of larger baselines: 10 mm, 15 mm, 20 mm, and 30 mm. Subjects were shown a board with poles of different heights and asked to record the perceived sequence of shortest to tallest based on their visual perception only. Each subject was scored based on how far their sequence was from the correct one. The results show that subjects fared better when larger baselines were used to view the task, with 20 mm performing the best. This preliminary study shows the value of an increased baseline and also presents a camera design that can achieve this baseline without facing any of the restrictions the current endoscopic design faces.

As mentioned previously, there are many possible future directions this work can take. Firstly, it is important to evaluate these results in the context of surgical robotics by including a task that involves manipulation and interaction with the task. After conducting such a study, we would also be able to evaluate the effect pure depth perception has on surgical performance. Secondly, it is necessary to extend the results of this study by testing with baselines larger than 30 mm and with varying working distances. Together, we would be one step closer to building an adaptive camera baseline design which can maximize surgical performance.

Acknowledgment

The authors would like to thank Dr. Peter Black, Dr. Chris Nguan, Keith Tsang, and Megha Kalia for their valuable input and discussions.

References

[1] V Falk, D Mintz, J Grunenfelder, J. Fann, and T. Burdon, “Influence of three-dimensional vision on surgical telemanipulator performance,” Surgical endoscopy, vol. 15, no. 11, pp. 1282–1288, 2001.

[2] C. Wheatstone, “Xviii. contributions to the physiology of vision.—part the first. on some remarkable, and hitherto unobserved, phenomena of binocular vision,” Philosophical transactions of the Royal Society of London, no. 128, pp. 371–394, 1838.

[3] R. Pepper and J. Hightower, “Research issues in teleoperator systems,” in Proceedings of the Human Factors Society Annual Meeting, SAGE Publications Sage CA: Los Angeles, CA, vol. 28, 1984, pp. 803–807.

[4] P. C. De Groen, “History of the endoscope [scanning our past],” Proceedings of the IEEE, vol. 105, no. 10, pp. 1987–1995, 2017.

[5] R. Satava, “3d vision technology applied to advanced minimally invasive surgery systems,” Surgical endoscopy, vol. 7, no. 5, pp. 429–431, 1993.

[6] A. F. Durrani and G. M. Preminger, “Three-dimensional video imaging for endoscopic surgery,” Computers in biology and medicine, vol. 25, no. 2, pp. 237–247, 1995.

[7] D. Mintz, V. Falk, and J. K. Salisbury, “Comparison of three high-end endoscopic visualization systems on telesurgical performance,” in International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2000, pp. 385–394.

[8] A. Avinash, A. E. Abdelaal, P. Mathur, and S. E. Salcudean, “A “pickup” stereoscopic camera with visual-motor aligned control for the da vinci surgical system: A preliminary study,” International journal of computer assisted radiology and surgery, vol. 14, no. 7, pp. 1197–1206, 2019.

[9] L. B. Rosenberg, “The effect of interocular distance upon operator performance using stereoscopic displays to perform virtual depth tasks,” in Proceedings of IEEE Virtual Reality Annual International Symposium, IEEE, 1993, pp. 27–32.

[10] F. Camposeco, C. Avilés, B. Careaga, L. Spindola, and R. Velázquez, “Constraints on human stereo vision for tele-operation,” in IX Latin American Robotics Symposium and IEEE Colombian Conference on Automatic Control, 2011 IEEE, IEEE, 2011, pp. 1–6.

[11] E. Spain, “Effects of extended camera baseline and image magnification on target detection time and target recognition with a stereoscopic tv system.,” Naval Ocean Systems Center, San Diego, CA, Tech. Rep., 1986.

[12] J. Hallett, 3-d imaging guides surgical operations, 2011. [Online]. Available: https://www.vision-systems.com/factory/robotics/article/16739083/3d-imaging-guides-surgical-operations (visited on 09/11/2019).

[13] J. McIntire, E. Geiselman, E. Heft, and P. Havig, “How much camera separation should be used for the capture and presentation of 3d stereoscopic imagery on binocular hmds?” In Head-and Helmet-Mounted Displays XVI: Design and Applications, International Society for Optics and Photonics, vol. 8041, 2011, p. 804 104.

[14] J. M. Fishman, S. R. Ellis, C. J. Hasser, and J. D. Stern, “Effect of reduced stereoscopic camera separation on ring placement with a surgical telerobot,” Surgical endoscopy, vol. 22, no. 11, pp. 2396–2400, 2008.

[15] M. S. Banks, J. C. Read, R. S. Allison, and S. J. Watt, “Stereoscopy and the human visual system,” SMPTE motion imaging journal, vol. 121, no. 4, pp. 24–43, 2012.

[16] L. Lipton, “Stereographics, developers handbook,” StereoGraphics Corporation, vol. 2, no. 2.2, pp. 2–2, 1997.

[17] D. Gere, C. R. Burns, J. D. Stern, and M. J. Tierney, Stereo imaging system and method for use in telerobotic systems, US Patent 6,720,988, 2004.

[18] Blackmagic design decklink. [Online]. Available: https://www.blackmagicdesign.com/ca/products/decklink (visited on 07/30/2019).

[19] G. Samei, K. Tsang, J. Lobo, C Kesch, S Chang, P Black, and S. Salcudean, “Fused mri-ultrasound augmented-reality guidance system for robot-assisted laparoscopic radical prostatectomy,” Hamlyn symposium on medical robotics, pp. 79–80, 2018.

[20] H. Schreuder, C. Van Den Berg, E. Hazebroek, R. Verheijen, and M. Schijven, “Laparoscopic skills training using inexpensive box trainers: Which exercises to choose when constructing a validated training course,” BJOG: An International Journal of Obstetrics & Gynaecology, vol. 118, no. 13, pp. 1576–1584, 2011.

[21] A. M. Liu, G. K. Tharp, and L. W. Stark, “Depth cue interaction in telepresence and simulated telemanipulation,” in Human Vision, Visual Processing, and Digital Display III, International Society for Optics and Photonics, vol. 1666, 1992, pp. 541–547.

[22] I. S. MacKenzie, “Within-subjects vs. between-subjects designs: Which to use?” Human-Computer Interaction: An Empirical Research Perspective, vol. 7, p. 2005, 2002.

[23] R. W. Hamming, “Error detecting and error correcting codes,” The Bell system technical journal, vol. 29, no. 2, pp. 147–160, 1950.

[24] F. Faul, E. Erdfelder, A.-G. Lang, and A. Buchner, “G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences,” Behavior research methods, vol. 39, no. 2, pp. 175–191, 2007.

[25] Y. Nakabo, T. Mukai, Y. Hattori, Y. Takeuchi, and N. Ohnishi, “Variable baseline stereo tracking vision system using high-speed linear slider,” in Proceedings of the 2005 IEEE international conference on robotics and automation, IEEE, 2005, pp. 1567–1572.

[26] A. Gudeloglu, J. V. Brahmbhatt, and S. J. Parekattil, “Robotic-assisted microsurgery for an elective microsurgical practice,” in Seminars in plastic surgery, Thieme Medical Publishers, vol. 28, 2014, pp. 011–019.

[27] L.-M. Su, Early Diagnosis and Treatment of Cancer Series: Prostate Cancer E-Book. Elsevier Health Sciences, 2015.