

# COMPARISON OF THREE HIGH-END ENDOSCOPIC VISUALIZATION SYSTEMS ON TELESURGICAL PERFORMANCE

David Mintz<sup>1</sup>, Volkmar Falk, MD<sup>2</sup>, and J. Kenneth Salisbury, Jr<sup>3</sup>

<sup>1</sup> DM is a Senior Systems Analyst with Intuitive Surgical Inc,  
1340 W. Middlefield Rd., Mountain View, CA 94043  
david\_mintz@intusurg.com  
<http://www.intusurg.com>

<sup>2</sup> VF is with the Department of Cardiothoracic Surgery Stanford School of Medicine

<sup>3</sup> KS is a Scientific Advisor at Intuitive Surgical Inc. and a member of the faculty in  
Computer Science at Stanford University

**Abstract.** Previous studies and surgeon interviews have shown that most surgeons prefer quality standard definition (SD)TV 2D scopes to first generation 3D endoscopes. The use of a telesurgical system has eased many of the design constraints on traditional endoscopes, enabling the design of a high quality SDTV 3D endoscope and an HDTV endoscopic system with outstanding resolution. The purpose of this study was to examine surgeon performance and preference given the choice between these. The study involved two perceptual tasks and four visual-motor tasks using a telesurgical system using the 2D HDTV endoscope and the SDTV endoscope in both 2D and 3D mode. The use of a telesurgical system enabled recording of all the subjects motions for later analysis. Contrary to experience with early 3D scopes and SDTV 2D scopes, this study showed that despite the superior resolution of the HDTV system surgeons performed better with and preferred the SDTV 3D scope.

## 1 Background

Various studies have examined the effect of 3D versus 2D imaging systems on performance for a variety of laparoscopic tasks or procedures. These studies led to inconclusive results. While some reports demonstrate a clear benefit for 3D vision systems, with a reduction in overall procedure time [1] [2] [3], others could not confirm these results [4] [5] [6].

These studies were performed using first generation 3D systems which have lower resolution than standard 2D systems. The first generation 3D endoscopes used were primarily single channel scopes with minimal stereo separation, resulting in limited stereopsis. These stereo scopes also all used single chip CCD cameras with limited video resolution. Most of the 3D camera systems used frame interlaced displays with shutter glasses to present the 3D images to the surgeons.

Despite these limitations, a number of conclusions can be drawn from the studies referenced above. Depth perception is regarded to be superior with 3D by most users. With first generation systems, however, this did not always result in improved performance, particularly for more complex tasks and for users with high levels of endoscopic experience. This is supported by our interviews of experienced endoscopic surgeons, where we found that they prefer quality 2D scopes with 3 CCD cameras to low resolution 3D endoscopes. The current lack of commercial promotion of these first generation 3D systems by endoscope manufacturers is anecdotal evidence which further supports this view.

The design of the first generation endoscopes used in these studies was severely constrained by the size and weight that a surgeon is willing to support throughout a surgical procedure. In contrast, the *da Vinci*<sup>TM</sup> telesurgical system used for this study employs a robot arm, with a comparatively large payload capability, to hold the endoscope during the procedure. The easing of size and weight constraints afforded by the robot arm allowed the development of an endoscope with superior stereo effect and three-chip CCD cameras resulting in a system with resolution on par with current 2D scopes. These less stringent design specifications also enabled the use of a very high resolution 2D scope attached to a three-chip HDTV camera that would be prohibitively large for use in a standard endoscopic case.

The primary motivation for our study was to look at surgeon performance and preference given the choice between an endoscope with superior 3D effect and three-chip CCD's, and a 2D scope with resolution superior to anything currently available. We seek to determine whether surgeons would continue to value additional resolution above depth perception in a camera system, or if passing a certain threshold in quality for the 3D scope has changed this rule.

## 2 Methods

In this study we compared the perceptions and performance of 15 experienced endoscopic surgeons (> 100 endoscopic cases a year) when using three different endoscopic viewing systems with the *da Vinci*<sup>TM</sup> telesurgical system. The three camera systems were: the 3D system that is standard with the *da Vinci*<sup>TM</sup> system; this same system in 2D mode; and a high quality 2D scope and prototype HDTV camera system. Prior to participating in the study, each of the surgeons went through a standard visual acuity test (eye chart). Only subjects with a corrected acuity of 20/20 or better were included. The study consisted of several perceptual tasks, followed by a group of visual-motor tasks, and finally a questionnaire.

### 2.1 The Telesurgical system

The *da Vinci*<sup>TM</sup> system is a surgical teleoperator developed by Intuitive Surgical Inc. of Mountain View, CA. The *da Vinci*<sup>TM</sup> system creates an immersive operating environment for the surgeon by providing both high quality stereo

visualization and a man-machine interface that directly connects the surgeon's hands to the motion of his surgical tool tips inside the patient's body [7]. The registration, or alignment, of the surgeon's hand motion to the motion of the surgical tool tips is both visual and spatial. The system projects the 3D image of the surgical site atop the surgeon's hands (via two CRT's and mirrored overlay optics), restoring hand-eye coordination and providing a natural correspondence in motions. Furthermore, the controller transforms the spatial motion of the tools into the camera frame of reference, so that the surgeon feels as if his hands are inside the patient's body. Lastly, the *da Vinci*<sup>TM</sup> system restores the degrees of freedom lost in conventional laparoscopy by placing a 3 degree-of-freedom wrist inside the patient and controlling this wrist naturally, bringing a total of seven freedoms of motion to the control of the tool tip (3 orientation, 3 translation and grip). The system also uses its digital controller to filter out surgeon tremor, making the tool tips steadier than the unassisted hand. Additionally, the system allows for variable motion scaling from the masters to the slaves. For example, a 3:1 scale factor maps 3 cm of translation on the masters into 1 cm of translation at the slaves.



**Fig. 1.** *da Vinci*<sup>TM</sup> System Surgeon console

The surgeon's console on the *da Vinci*<sup>TM</sup> system, Fig. 1, transmits the image of the surgical site to the surgeon through a high-resolution stereo display, which uses two medical grade CRT monitors to display one image to each of the sur-

geon's eyes [7]. The virtual image plane of the stereo viewer is placed coincident with the range of motion of the masters using two sets of mirrors. This stereo viewer supports standard definition television (NTSC or PAL) signals. For this study NTSC was used. In order to facilitate the use of the HDTV scope, a custom surgeon's console was built that uses an HDTV monitor and two mirrors to place the high resolution 2D image in the workspace of the masters, thereby replicating the immersive sensation of the production console.

## 2.2 Endoscopic Camera systems

Two vision systems were used for this study. The *InSite*<sup>TM</sup> system was used to provide both the 2D and 3D environments in NTSC format. A second vision system, including a high quality monocular endoscope, an HDTV camera head and a custom built stereo viewer, was used to provide the 2D HDTV view.

***InSite*<sup>TM</sup> Vision system.** The *InSite*<sup>TM</sup> vision system, provided standard with *da Vinci*<sup>TM</sup>, includes a 3D-NTSC endoscope, developed by Precision Optics Corp. This scope has two separate optical channels that provide wide stereo separation while maintaining symmetric optical distortion in each eye. The payload of the *da Vinci*<sup>TM</sup> system camera arm enabled the scope and camera head designers to opt for higher quality whenever there was a design trade-off between quality and size or weight. The result is a scope with a stereo separation of 5.5 mm, giving good stereo effect, and optical resolution of 3.3 mrad/lp for each channel. Two 3-chip CCD cameras are used to capture the full resolution provided by the optics of the endoscope.

The stereo viewer in the surgeon console is specially designed to match the endoscope. The nominal working distance for the endoscope, when used in cardiac surgery is 38 mm. Given the 5.5 mm stereo separation of the optical channels, the angle of convergence with the tissue plane is 8.3 deg. Human interocular spacing is approximately 6.5 cm. The stereo viewer was designed such that the image plane is placed 43 cm away from the eye piece. The angle subtended by the surgeon's eyes is the 8.6 deg, resulting in a realistic stereo effect. This scope was used in 2D mode for a portion of the study by providing the image captured by one of its optical chains to both monitors in the surgeon console. This allows a direct comparison of 2D vs. 3D performance independent of resolution.

**HDTV System.** The prototype HDTV endoscopic system used in this study has exceptional resolution (1.4 mrad/lp) throughout its field of view. A high quality 10 mm monocular endoscope was used attached to a broadcast quality 3-chip CCD HDTV camera to obtain this resolution. Because the stereo viewer included in the production *da Vinci*<sup>TM</sup> surgeon console cannot display HDTV signals, a custom HDTV-capable console was built. The viewer in this console consisted of an HDTV monitor and a pair of mirrors used to project the image into the master manipulator workspace. The design of this console mimicked that of the SDTV stereo viewer described above.

### 2.3 Experiments

All subjects underwent a structured introduction to the system and were allowed to become acquainted with its features. The two scopes were set at the same fixed distance from the tasks.

Six tasks were performed using each visualization system. Subjects completed each of the tasks with a given visualization system before moving on to the next. The order in which the visualization systems were presented to the surgeons was randomized in order to minimize any bias due to learning. The tasks were always presented in the same order.

The first two tasks were purely perceptual in nature while the remaining four involved manipulations with the telesurgical system. This permitted a separation of the subjects' perceptual and sensorimotor skills under each viewing condition. The six tasks are outlined below.

1. **Resolution Estimation.** Subjects were presented standard resolution patterns and asked to determine the level of resolution they could discriminate in both the horizontal and vertical directions.
2. **Relative Distance Estimation** Subjects were presented with a set of six pegs of varying height positioned at varying distances from the scope. This was designed to cause confusion between distance and size. They were told that the pegs were each in one of three rows. The subjects were asked to classify each peg as near, medium, or far.
3. **Switch Touching.** Each subject was given control of a telesurgical manipulator with his dominant hand and told to touch each of four switches placed at varying heights about the viewing field. The state of the switches was recorded, along with the position and orientation of the manipulator tip.
4. **Peg Placement.** Using the same manipulator subjects was directed to pick up cylindrical beads from a starting location and place them on the pegs viewed in task 2. The number of beads dropped as well as the total time to place all six beads was measured. Dropped beads were returned to the starting position for the subjects to retrieve and continue with the task.
5. **Suturing.** Subjects were asked to perform a running suture (4-0 silk) connecting printed dots (distance 4 mm horizontally, 5 mm vertically) on a fixed rubber glove using two telesurgical instruments (needleholders). Time to complete the task and errors (stitch outside a 1mm circle around the marked dot) were recorded.
6. **Knot Tying.** Using the same two instruments, each subject was instructed to tie three instrument knots using a 4-0 Silk suture that was previously stitched into a rubber glove. Time to perform the task was recorded.

### 2.4 Questionnaire

After completion of the tasks, all subjects were interviewed with regard to their performance under the various viewing conditions using a standardized questionnaire. System-related questions included subjective assessment of resolution, depth perception and color perception with the different set-ups.

### 3 Results

The results of the study are summarized below. A full statistical analysis has not been provided, but can be found in Falk et al. [8]

#### 3.1 Experiments

1. **Perceived Resolution.** As expected the HDTV system showed significantly better perceived resolution both horizontally and vertically when compared to both of the 2D and 3D NTSC setups. Interestingly, there was a perceived resolution difference between the 2D and 3D NTSC. Although the difference was small, both the horizontal and vertical resolution was perceived as better in the 3D mode. This may perhaps be explained by the fact that various cerebral summation effects (by cortical vision centers) enhance among other factors acuity and pattern recognition beyond what can be expected from two equal monocular inputs [9]. According to Cagnello and Rabin, acuity is enhanced by roughly 10% when comparing binocular to monocular vision [10]. Given the same resolution, a 3D-system may therefore not only enhance depth perception, but also give better visualization.

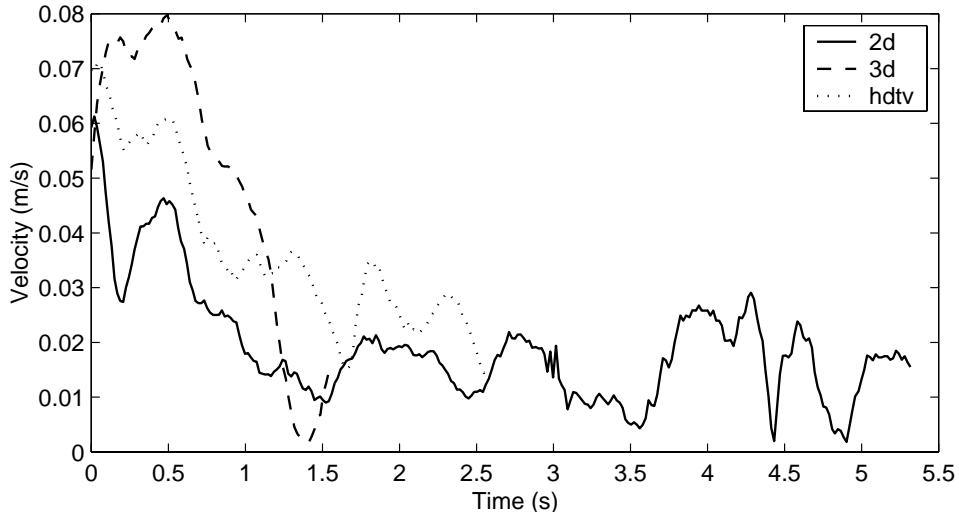
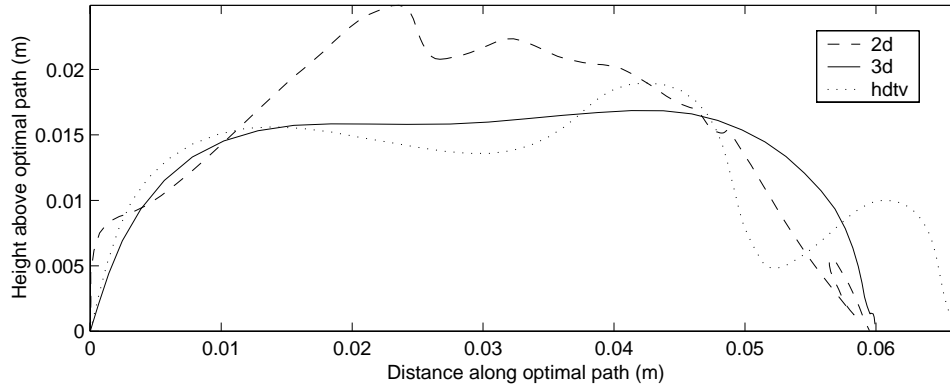


Fig. 2. Example velocity profile. (Subject #3 moving from switch 1 to 2)

2. **Switch Touching Task.** The large amount of raw data taken during this task has allowed a significant amount of post processing and analysis. In the interest of space, not all the metrics studied will be reported here. A more complete analysis of these data is available in Falk et al. [8]. The 3D NTSC system performed as well or better than both of the 2D systems for all metrics

considered. For several of the metrics calculated (total time between switch contacts, mean velocity, and mean acceleration, in particular) performance for the 3D system was significantly better ( $p < 0.03$ ) than for either of the two 2D setups. Additionally the length of the deceleration phase, defined as time from peak velocity to contact with the target switch, was significantly shorter for the 3D system than either 2D system. A typical velocity profile is shown in Fig. 2.



**Fig. 3.** Movement from switch 1 to 2 for subject #3

Inspection of the raw data supplied some other interesting insights, which are difficult to quantify. For instance, Fig. 3 shows the motion of the instrument tip mapped onto a plane defined by the straight line distance from switch to switch, and the maximum deviation from that line. What is interesting about this view of the data is that it supports the claim that motion under 3D visualization was made more confidently and efficiently than with the 2D setups. This can be seen in the smoothness and directness of the path taken. This was typical of many of the subjects. It is also interesting to note that in the case pictured, the motion under HDTV viewing did not finish at the same point as the motions under the other viewing systems. This is because the instrument tip was past the target switch when the switch was contacted, presumably by the wrist or shaft of the tool. This is a further indication of the difficulty encountered in determining depth while using a 2D system.

- 3. Relative Distance Estimation and Peg Placement Tasks.** All subjects were able to assign correctly the pins to their corresponding row under all viewing conditions. Interestingly, using the same pegboard patterns, but now requiring placement of objects on the pins, the results from task 4 were markedly different. With the involvement of motor-skill (acquisition and placement of objects), performance was clearly improved under 3D vision. Although the difference in time to complete the task was only statistically

significant between the 2D-NTSC and the 3D-NTSC systems, the task was performed faster under 3D than under HDTV by 12 of the 14 subjects, as seen in Fig. 4.

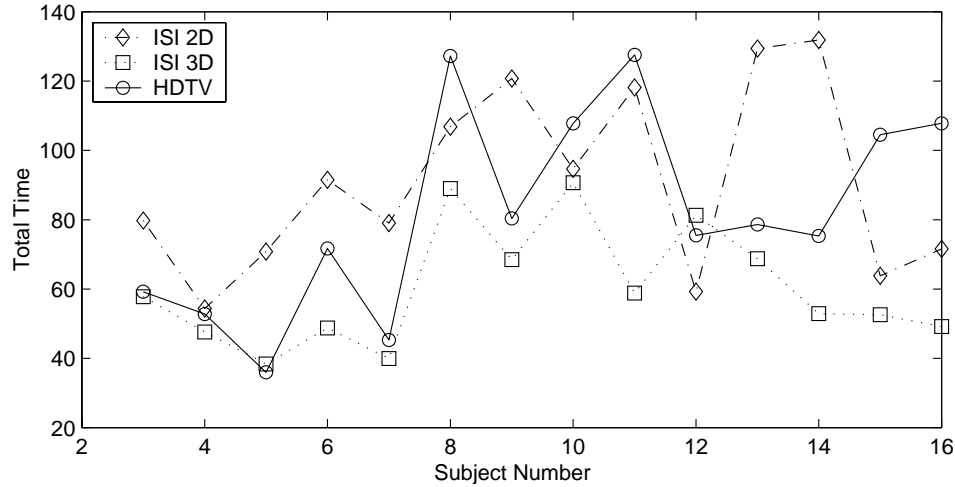


Fig. 4. Total time for peg placement

- 4. Suturing and Knot Tying Tasks.** Results for both the suturing drill and knot tying tasks were again in favor of the 3D-setting [8]. Most subjects experienced difficulties with needle orientation under 2D vision.

### 3.2 Questionnaire

Asked for the best visualization, 53% of the surgeons preferred the HDTV-set-up followed by the 3D-NTSC set-up (33%), while 13% did not see a difference. Eighty percent of the subjects evaluated their own performance under 3D viewing conditions as superior to either 2D setting. As expected, 93% of subjects felt depth perception was best with the 3D-set-up. In terms of resolution the majority favored the HDTV system (HDTV 73%, 3D 13%, undetermined 13%). When asked if problems with depth arose during performance, 87% of the subjects found problems with the 2D-NTSC setting and 47% also for the 2D-HDTV setting (3D: 0%). Asked for the overall effectiveness of working under 2D and 3D conditions, 87% voted for 3D. Eighty percent of the subjects rated the relative importance of depth perception more important than resolution. When asked about specific problems with 2D vision, subjects complained about incorrect needle positioning and inability to grasp suture. On a verbal rating scale (1-10) of subjective importance for endoscopic work, depth perception averaged at a level of 8.4, demonstrating its perceived importance for endoscopic work. This was despite the fact that 93% of all subjects worked with a 2D-scope in their



daily endoscopic practice and stated in the questionnaire that they felt that 3D vision was not an improvement prior to the trial.

## 4 Discussion

Despite a predisposition to 2D systems, the surgeons participating preferred, and performed best when using, the 3D system even when presented with a 2D system with truly superior resolution. This is in contrast with many of the earlier studies and our previous experience with first generation 3D endoscopic systems compared to standard resolution 2D systems. We conjecture that there is a threshold of quality, in either resolution or 3D effect, when the addition of depth cues provided by a 3D system becomes more valuable than better resolution. This assertion is reasonably supported by the results of the tasks and the questionnaire, and warrants further investigation.

The telesurgical platform provides several design advantages that have allowed this 3D endoscope quality threshold to be reached. First the scope and camera are held by a robot rather than by a surgeon, so they can be designed larger and heavier than those used in standard endoscopic surgery. Secondly the surgeon is seated and immersed in the endoscopic view with his head in a fixed position. This relieves the constraint for mobility and peripheral vision that is imposed when the surgeon is working at the operating table. This allows the use of two properly aligned medical grade monitors to present the three-dimensional view, rather than shutter glasses or heads up displays which produce an inferior 3D image.

It is reasonable to expect that surgeons will welcome the better visual resolution that will come with next-generation HDTV-based stereo endoscopes. However, to make cost-effective design choices, it will be important to better understand the tradeoffs between depth perception and spatial resolution touched upon in this study and their effect upon surgical task performance.

## References

1. Dion, Y.M., Gaillard, F.: Visual integration of data and basic motor skills under laparoscopy. Influence of 2-D and 3-D video-camera systems. *Surg Endosc* 1997;11:995-1000
2. Buess, G.F., van Bergen, P., Kunert, W., Schurr, M.O.: Comparative study of various 2D and 3D vision systems in minimally invasive surgery. *Chirurg* 1996;67:1041-6
3. Von Pichler, C., Radermacher, K., Boeckmann, W., Rau, G., Jakse, G.: Stereoscopic Visualization in Endoscopic Surgery: Problems, Benefits and potentials. *Perception* 1997;6:198-217
4. Hanna, G.B., Shimi, S.M., Cuschieri, A.: Randomised study of influence of two dimensional versus three dimensional imaging on performance of laparoscopic procedures. *Lancet* 1998;351:248-51
5. Jones, D.B., Brewer, J.D., Soper, N.J.: The influence of three-dimensional video systems on laparoscopic task performance. *Surg Laparosc Endosc* 1996;6:191-7

6. Chan, A.C., Chung, S.C., Yim, A.P., Lau, J.Y., Ng, E.K., Li, A.K.: Comparison of two-dimensional vs three-dimensional camera systems in laparoscopic surgery. *Surg Endosc* 1997;11:438-40
7. Guthart, G.S., and Salisbury, J.K.: The Intuitive Telesurgery System: Overview and Application, Proc. of the IEEE International Conference on Robotics and Automation (ICRA2000), San Francisco CA, April 2000.
8. Falk, V., Mintz, D., Reitz, B.A.: Influence of 3D-Vision on Telemanipulator Performance. Submitted for publication 2000
9. Frisen, L., Lindblom, B.: Binocular summation in humans: Evidence for a hirschman model. *J Physiol* 1988;402:773-782
10. Rabin, J.: Two eyes are better than one: binocular enhancement in the contrast domain. *Ophthalmic Physiol Opt* 1995;15:45-48