Autostereoscopic
Three-Dimensional Viewer
Evaluation Through Comparison
With Conventional Interfaces in
Laparoscopic Surgery

Michele Silvestri, MSc¹, Massimiliano Simi, MSc¹, Carmela Cavallotti, MSc¹, Monica Vatteroni, PhD¹, Vincenzo Ferrari, PhD², Cinzia Freschi, PhD², Pietro Valdastri, PhD¹, Arianna Menciassi, PhD¹, and Paolo Dario, PhD¹

Abstract
In the near future, it is likely that 3-dimensional (3D) surgical endoscopes will replace current 2D imaging systems given the rapid spreading of stereoscopy in the consumer market. In this evaluation study, an emerging technology, the autostereoscopic monitor, is compared with the visualization systems mainly used in laparoscopic surgery: a binocular visor, technically equivalent from the viewer’s point of view to the da Vinci 3D console, and a standard 2D monitor. A total of 16 physicians with no experience in 3D interfaces performed 5 different tasks, and the execution time and accuracy of the tasks were evaluated. Moreover, subjective preferences were recorded to qualitatively evaluate the different technologies at the end of each trial. This study demonstrated that the autostereoscopic display is equally effective as the binocular visor for both low- and high-complexity tasks and that it guarantees better performance in terms of execution time than the standard 2D monitor. Moreover, an unconventional task, included to provide the same conditions to the surgeons regardless of their experience, was performed 22% faster when using the autostereoscopic monitor than the binocular visor. However, the final questionnaires demonstrated that 60% of participants preferred the user-friendliness of the binocular visor. These results are greatly heartening because autostereoscopic technology is still in its early stages and offers potential improvement. As a consequence, the authors expect that the increasing interest in autostereoscopy could improve its friendliness in the future and allow the technology to be widely accepted in surgery.

Keywords
ergonomics and/or human factors study, SILS, single site surgery, surgical education, biomedical engineering

Introduction
For the past 20 years, laparoscopy has been widely used in surgical operations. The main advantages of laparoscopy are smaller incisions, quicker recovery time, and less post-operative pain than traditional surgery.¹ Laparoscopic procedures, however, can be stressful for the surgeon². Traditional laparoscopic tools have only 4 degrees of freedom, motion direction is reversed by the fulcrum effect of the access point, and tactile feedback is strongly limited.³ These limitations make visual information more and more crucial for performing laparoscopic procedures. Although standard laparoscopic imaging systems provide optimal performance in terms of image quality, detail, and color sharpness, they provide only 2-dimensional (2D) images of the operative field, without any depth cue.⁴,⁵ Consequently, surgeons have to relearn their operative skills and accustom their brain to acquire spatial orientation in the laparoscopic 2D images.⁶

¹Scuola Superiore Sant’Anna, Pontedera, Italy
²University of Pisa, Pisa, Italy

Corresponding Author:
Michele Silvestri, The BioRobotics Institute, Scuola Superiore Sant’Anna, viale R. Piaggio, 34, 56025 Pontedera, Italy
Email: m.silvestri@sssup.it
To overcome these problems, new devices have been developed that guarantee three-dimensional (3D) vision of the operative field. The most famous and commercially available platform is the da Vinci robotic system (Intuitive Surgical, Mountain View, CA) that delivers a high-quality 3D view of the operative field using a binocular visor (BV). There are several studies in literature that demonstrate that the da Vinci stereoscopic display improves surgical performances with respect to traditional 2D viewers because of better distance and motion perception. Therefore, the BV can be considered today as the gold standard for 3D visualization systems in laparoscopy. Another commercially available 3D imaging system is the Triacam (Karl Storz, Tuttingen, Germany) This device uses a standard monitor and polarized glasses to provide the image from the left (right) camera to the left (right) eye of the observer. A research prototype using the same kind of stereoscopic display is reported in literature, consisting in a 3D insertable imaging device for surgical robotics developed by Fowler et al. In connection with this technique, comparative studies already available demonstrate that polarized glasses are more efficient than 2D devices. They mainly emphasize that the advantages of 3D vision increase with the complexity of the task and that depth perception improves the learning curve of surgical skills.

All the stereoscopic displays mentioned above belong to the category of binocular devices since they use additional accessories—such as glasses or helmets, or 2 separate video displays, as in the case of the BV—to show a different image to each eye. The main drawback of these displays is the total immersion in the console, which isolates the surgeon from the surrounding environment and sometimes causes a wrong posture. In the da Vinci platform, the BV is integrated in a console, studied to help the surgeon work in a more comfortable position with a steady position of the head; however, this does not guarantee a better overall posture and increases the surgeon’s estrangement from both the environment and the patient. Nowadays, autostereoscopic devices have been introduced to provide a stereoscopic vision without additional items, such as visors and glasses. Autostereoscopic devices rely on a geometrical solution that keeps the images separate for each eye. Over the past years, consumer electronics, mainly driven by the entertainment market, has greatly invested in autostereoscopic displays. For example, the new Nintendo DS, the Microsoft autostereoscopic prototype, and the Sharp 3D glass-free monitor have improved the handiness and quality of vision thanks to the use of innovative technological solutions. These examples suggest that autostereoscopic technology may become more effective and cheaper in the near future, allowing its use both in the mass market and in specialized fields. Autostereoscopic monitors (AMs) offer 2 main advantages compared with binocular devices: They do not estrange the viewer from the surrounding environment, and they can be viewed by more than one person at the same time. These features can dramatically improve the surgeon’s performance if applied in the operating room.

Only a few studies have been presented in literature, since autostereoscopic devices have not yet found specific medical application. The work of Alpaslan et al. demonstrates the equivalence, in terms of stereoscopic effect, between shutter glasses (a binocular device) and AMs in general manipulation and interaction tasks. In 2003, Mueller-Richter et al. compared surgical performance using a first-generation AM, polarized glasses, and standard 2D devices. This study demonstrated that 2D images are slightly better than 3D cues. Furthermore, polarized glasses showed better performance than the AM.

As summarized in Figure 1, the aim of this work is to evaluate AMs as an emerging technology with significant potential in endoscopic applications, with the 2 most conventional visualization systems currently used in laparoscopic surgery: a BV and a standard 2D system. It is worth mentioning that passive and active 3D displays requiring the use of glasses are widespread and are the cheapest and easiest stereoscopic displays commercially available at the moment. Providing a comparison with these devices, however, is beyond the scope of this paper. In fact, our goal is to verify whether the improvements achieved by AMs since 2003 are able to guarantee performances—in terms of spatial effect and applicability—equivalent to conventional laparoscopic visualization systems, whose surgical effectiveness has been already demonstrated. It is worth mentioning that this study is only a preliminary analysis because, as illustrated above, AMs may have an impressive growth in the near future and could possibly replace 2D and binocular devices currently used in laparoscopic surgery.

This article is divided into the following sections: “Materials and Methods” describes the image acquisition system and the visualization systems used in the comparative study and presents a summary of the tests carried out. “Results” summarizes the results obtained, whereas the final section, “Discussion and Conclusions,” discusses the experimental findings and outlines future work in the field.
Materials and Methods

A dedicated stereoscopic acquisition system was developed to acquire images of the operative field. The system uses 2 cameras, set up in parallel configuration, with 2 separate optical channels. The cameras are commercial (Misumi Electronics Corp, Taiwan) video graphics array (VGA; 640 × 480 pixels) CMOS color imagers, measuring 8 mm × 8 mm × 9 mm, and provided with pinhole lens. Each camera has approximately a 60° horizontal and 52° vertical field of view. Moreover, a fixed distance of 8.7 mm between the camera centers allows comfortable stereoscopic vision ranging between 40 and 150 mm in scene depth. The lighting system is based on 8 high-efficiency white light emitting diodes. Both the acquisition and the illumination systems are embedded in a plastic case measuring 25 mm in diameter and 95 mm in length. Thanks to a set of permanent magnets embedded in the camera body, rough positioning of the point of view can be obtained along the laparoscopic simulator (see Figure 2) by dragging an external magnetic handle.22,23

The features of the 3 different visualization systems compared in this study are illustrated below. The first is the 19-inch (1280 × 1024) AM24 (Pavonine Korea Inc, Korea) that employs the parallax barrier technology for 3D visualization. The parallax barrier is an electro-optic panel with vertical, regularly spaced slits attached to the surface of a liquid crystal display. The slits are used to obscure parts of the 2 images coming from the cameras and to spread 2 separate 2D images ahead of the monitor.25 This architecture guarantees the projection of 2 different images in the observer’s left and right eye, thus allowing a 3D view in specific regions, called “sweet-spots,” as illustrated in Figure 3. The observer has an optimal viewing distance of around 80 cm depending on individual eye separation and on whether he or she is positioned up to ±55° with respect to the Z axis of Figure 3. This device also allows 2D vision: By simply switching off the electro-optical panel, the 2D image stream coming from one camera is displayed on the liquid crystal display.

The second 3D display is a stereoscopic head-mounted display (HMD), eMagin Z800, which offers SVGA resolution (800 × 600 pixels) on organic light emitting diode panels positioned in front of each eye. The panels can be adjusted to the anthropometric user measures.26 An ad hoc software program was developed to acquire camera images by means of a frame grabber (Picolto Altert by Euresys) and to send the images alternatively to the right and left panels at a frequency of 60 frames per second through a single VGA output. The visor receives the images and displays them to the correct eye.

The same acquisition system was used in all experiments so that each visualization system could have equal resolution and image quality. This allowed us to exclude the influence of acquisition system-related features from the evaluation and to avoid the main problem encountered in early comparative studies between 2D and 3D techniques,27,28 where low-resolution 3D displays were compared with high-definition 2D monitors. The acquisition system was set in an optimal visual position at the beginning of the test to prevent biasing when positioning it because of individual experience.

The participants in this study were 16 physicians who had no previous experience in 3D vision and had different surgical skills. They were asked to perform 5 different tasks (see Figure 4) in a laparoscopic simulator using the AM, the BV, and the 2D monitor as control (see Figure 5). Since the same tasks were repeated 3 times, the order of the different viewing conditions was randomly varied to minimize bias because of learning. A 5-minute initial training was carried out for each visualization system so that the users could become acquainted with the viewing conditions. The tasks, summarized in Figure 4, were organized by increasing complexity to avoid motion-memory effects.8

Each task is briefly described below:

1. **Visual task**: The user was asked to count the black and white squares in his or her visual field (Figure 4A). Square pitch was 0.7 mm. The aim of this task was to check (especially with 3D displays) whether the observer correctly perceived the stereoscopic effect without distortions. The number of errors was recorded.

2. **Pick-and-place task**: The user was asked to place 10 rings onto 10 pins. Each pin had a different height, position, and different colored head (see Figure 4B). This task was performed by using
a laparoscopic pincer. The execution time and the number of errors were recorded.

3. **Peg-in-hole task**: The user was asked to stick 10 needles into 10 holes, at different depths and heights, as shown in Figure 4C. The execution time and the number of errors were recorded.

4. **Cutting task**: The user was asked to cut a rubber surface by following a precise pattern, as shown in Figure 4D. This task is common in open surgery, whereas radiofrequency ablation is the standard for cutting a tissue in laparoscopic surgery. For this reason, the laparoscopic cutting task was unfamiliar to all laparoscopic surgeons and eliminated any potential bias due to personal skills and experience. A special laparoscopic grasper, with a no. 20 scalpel blade fixed to the tip, was used for performing the cut. The execution time and the number of errors were recorded.

5. **Suturing task**: The user was asked to perform a single suture on a tissue simulator surface (Figure 4E). This was the most complex task since the user had to use both hands to perform it. The execution time was recorded.

A questionnaire was filled in after the test by all participants to collect their opinions and suggestions.

A statistical analysis was carried out on the recorded data to compare performance in the 3 visual conditions. We used the one-way analysis of variance (ANOVA) to identify
differences between the 3 independent groups, corresponding to the 3 visualizations systems. Differences were considered statistically significant when $P \leq .05$.29

Results

The one-way ANOVA starts from recorded data and from a priori defined number of groups to work out a variance between groups and a variance within groups. The ratio between these parameters is called $P$ value.29 If this value is lower than .05, at least one difference in the average values obtained for the 3 groups is statistically significant, meaning that it is not random. To detect exactly which differences were statistically significant, a multiple comparison was carried out between each pair using the Scheffé post hoc analysis29 and fixing a .05 threshold value for the $P$ value. The following results were obtained:

1. **Visual task:** No errors were recorded with all the visualization systems; therefore, no statistical analysis was carried out.

2. **Pick-and-place task:** The average execution times, reported in Table 1, were as follows: 128 ± 30 seconds for the AM, 136 ± 25 seconds for the BV, and 144 ± 31 seconds using the 2D monitor. Application of the ANOVA to the recorded data resulted in a $P$ value of .004; consequently, at least one difference in the average values was statistically significant. Application of the post hoc analysis (see Table 1) revealed that the pick-and-place task was performed significantly better when using 3D systems than the 2D monitor (AM vs 2D, $P = .0011$ and BV vs 2D, $P = .0008$). Although this task was performed faster when using the AM than when using the BV, this difference was not statistically significant ($P = .8197$). No errors were recorded for all testers when executing this task.

3. **Peg-in-hole task:** The average execution times, showed in Table 1, were as follows: 119 ± 59 seconds for the AM, 107 ± 57 seconds for the BV, and 114 ± 47 seconds when using the 2D monitor. In this case, application of the one-way ANOVA to the average execution times resulted in a $P$ value of .79 (see Table 1). Therefore, the differences obtained in the average values were not statistically significant. The average number of errors (see Table 2) was also analyzed in this task: 0.50 ± 0.85 for the AM, 0.71 ± 0.87 for the BV, and 0.57 ± 0.64 using the 2D monitor. However, the differences obtained were not statistically significant because the overall $P$ value was .7784.

4. **Cutting task:** The average execution times, showed in Table 1, were as follows: 21 ± 2 seconds for the AM, 27 ± 1 seconds for the BV, and 25 ± 3 seconds when using the 2D monitor. Application of the one-way ANOVA resulted in a $P$ value equal to .034; therefore, at least one difference in the average execution times was statistically significant. The post hoc analysis demonstrated that the cutting task was performed significantly better when using the AM than using the BV ($P = .0426$). The same analysis also demonstrated that this task was performed significantly faster when using the 2D monitor than using the BV ($P = .0469$). However, the difference between AM and 2D average execution times was not statistically significant ($P = .7847$). The average numbers of errors, shown in Table 2, were as follows: 0.40 ± 0.50 for the AM, 0.73 ± 1.03 for the BV, and 1.00 ± 0.92 when using the 2D monitor. Application of the one-way ANOVA resulted in a $P$ value equal to .1676; thus, the differences obtained were not statistically significant.

5. **Suturing task:** The average execution times were as follows: 105 ± 30 seconds for the AM, 102 ± 45 seconds for the BV, and 124 ± 40 seconds when using the 2D monitor. The one-way ANOVA demonstrated that these differences were statistically significant ($P = .0099$), whereas post
Table 1. Test Results in Terms of Execution Times Using the AM, the BV and the 2D Displays$^a$

<table>
<thead>
<tr>
<th>Task</th>
<th>Average Execution Time (s)</th>
<th>Overall P</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AM</td>
<td>BV</td>
<td>2D</td>
</tr>
<tr>
<td>Pick-and-place</td>
<td>128 ± 30</td>
<td>136 ± 25</td>
<td>144 ± 31</td>
</tr>
<tr>
<td>Peg-in-hole</td>
<td>119 ± 59</td>
<td>107 ± 57</td>
<td>114 ± 47</td>
</tr>
<tr>
<td>Cutting</td>
<td>21 ± 2</td>
<td>27 ± 1</td>
<td>25 ± 3</td>
</tr>
<tr>
<td>Suturing</td>
<td>105 ± 30</td>
<td>102 ± 45</td>
<td>124 ± 40</td>
</tr>
</tbody>
</table>

Abbreviations: AM, autostereoscopic monitor; BV, binocular visor; 2D, 2-dimensional monitor.
$^a$For each task, the 3 average execution times and the overall P value, carried out applying the one-way analysis of variance on the 3 groups, are presented. In addition, this table shows the P values obtained with the post hoc analysis for each pair of visualization systems.

Table 2. Test Results in Terms of Number of Errors Using the AM, the BV, and the 2D Displays$^a$

<table>
<thead>
<tr>
<th>Task</th>
<th>Average Number of Errors</th>
<th>Overall P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AM</td>
<td>BV</td>
</tr>
<tr>
<td>Peg-in-hole</td>
<td>0.50 ± 0.85</td>
<td>0.71 ± 0.87</td>
</tr>
<tr>
<td>Cutting</td>
<td>0.40 ± 0.50</td>
<td>0.73 ± 1.03</td>
</tr>
</tbody>
</table>

Abbreviations: AM, autostereoscopic monitor; BV, binocular visor; 2D, 2-dimensional monitor.
$^a$For the peg-in-hole and cutting tasks, the 3 average numbers of errors and the overall P value, carried out applying the one-way analysis of variance on the 3 groups, are presented.

The post hoc analysis demonstrated that this task was performed significantly faster when using 3D displays than using the 2D monitor (AM vs 2D, $P = .0202$ and BV vs 2D, $P = .0059$). However, although the BV obtained an average execution time lower than the AM, this difference was not statistically significant ($P = .4986$).

As regards the questionnaire, all participants appreciated the 3D view provided by both stereoscopic displays. Around 65% of participants preferred the BV because of its user-friendliness and 60% of them found the AM tiring in a long-lasting operation.

**Discussion and Conclusions**

An evaluation on surgical performance was performed by comparing emerging autostereoscopic technology with conventional visualization systems used in laparoscopic surgery, that is, a BV equivalent to the da Vinci console and a standard 2D monitor. The acquisition system, based on 2 VGA cameras, was used with all visualization devices to achieve equal resolution and image quality. Participants were asked to perform 5 tasks. Then, execution time and number of errors were recorded and statistically analyzed.

The visual task was used to verify whether the participants were able to correctly perceive and fuse the stereoscopic image stream they were not accustomed to. Results demonstrated that all the participants found it effortless to interface with 3D vision without distortions after the warranted training time.

Both 3D displays allowed better performances in terms of execution speed, with respect to conventional 2D vision, in both low-complexity (pick-and-place) and high-complexity (suturing) tasks. As in previous studies, which used the da Vinci surgical console, the final outcome depends on the improvement of eye–hand coordination, because of the binocular cues and depth perception provided by 3D displays. Above all, no statistically significant difference was found between the BV and the AM, thus proving that they are equally effective. This result disagrees with the findings of Mueller-Richter et al$^{15}$: Interestingly, it demonstrates that improvements in autostereoscopic technology have increased its effectiveness and reveals that autostereoscopy can be competitively compared to conventional BVs. It is worth mentioning that although the AM did not achieve a statistically significant improvement in terms of execution times, it showed equivalent effectiveness when compared with the BV, which is a technologically mature and commercial system. Therefore, since AMs are still in their early stages, they have substantial room for improvement that may drive autostereoscopy to become the leading visualization system over the next few years.

The peg-in-hole task did not show statistically significant differences. This mainly depends on low camera resolution, as pointed out by the participants during the test. This hampers the correct view of both ends of the pin at the edges of the operative field and alters the results.
The cutting task was introduced because the cold cutting of tissue is not a traditional surgical practice, consequently, each participant practically started basically from the same training level. However, no statistically meaningful differences between 3D and 2D conditions were found in this task with respect to both execution time and number of errors. This result could depend on the fact that since the surface to be cut was planar, the participants based their operations mostly on force feedback. As a result, the advantages gained from the 3D perception of the scalpel tip position were lost. However, this study showed a statistically significant difference in execution times between the AM and BV. Since all the participants had never used a 3D display before this study, this result can be considered a preliminary indication that the AM has a shorter learning curve than the BV.

Questionnaires showed how participants were all able to perceive the third dimension using both the BV and AM. They appreciated the depth perception and relative motion perception guaranteed by binocular cues. This result agrees with the study of Alpaslan et al., in which the spatial effect obtained with an AM was found equivalent to that obtained using a binocular device. However, most of the participants (about 65%) preferred the 3D user-friendly features of the BV. This may depend on a required training time for the AM that is longer than 5 minutes during which the observer must find a perfect viewing position to properly fuse the stereoscopic images in the brain. In addition, execution friendliness and speed of activities strongly depend on the observer’s personal skills. A total of 35% of participants who preferred the AM were mostly those who found it effortless and natural to fit into the third dimension of the autostereoscopic screen. This initial adaptation step also led 60% of participants to prefer the binocular visor for a long-lasting task, because they occasionally developed ocular fatigue at the end of the tests performed with the AM. However, as demonstrated by Barkowsky and Le Callet, viewing a 3D content on an AM does not have any measurable influence on the subsequent performance of activities, and the visual fatigue that is sometimes felt is only temporary and fully recoverable. It is worth mentioning that AM technology is still in its early stages and that the specific solutions adopted to achieve autostereoscopy may have a different outcome in terms of user-friendliness. For this reason, similar studies should be carried out on each new autostereoscopic technique, once available. Above all, the short duration of the tests in this study may not have emphasized the stress resulting from the use of the BV in a long surgical task. For example, we noticed that some participants fell into a wrong posture caused by estrangement from the surrounding environment when wearing the BV. The development of new-generation autostereoscopic displays may improve this condition. For example, the new Microsoft prototype uses a camera to track the viewer, so that the system knows where to steer the light. This allows one of the training steps to be deleted and decreases the learning curve. At the same time, the advantage of viewing the images by more than one observer at the same time is not lost because the prototype can deliver a 3D video to 2 viewers by presenting different images to their left and right eyes.

In conclusion, a surgical performance evaluation was carried out between the emerging AM technology and 2 conventional visualization systems in laparoscopy. Similarly to previous works, results distinctly demonstrated that the use of conventional 2D system increases execution time and decreases accuracy compared with 3D interfaces. This suggests that 3D monitors will replace 2D systems in surgery. Moreover, by comparing the mature commercial technology of the BVs with the emerging autostereoscopic technology, this study demonstrated that the latter is already equally effective in both low- and high-complexity tasks. Although the need for a longer training times still limits wide acceptance of the AM by the surgical community, its effectiveness is highly appealing, especially in the view of the fact that AM technology has become fast growing over the past few years thanks to the great interest shown by the entertainment market. This growth may pave the way to wide use of AMs in surgery in the near future.

In the future, we will increase the pool of the statistical analysis to strengthen the results. The involvement of a large number of participants will also allow separate analysis between and within groups of homogenous surgical experience. Above all, tracking of the instrument tip with respect to the target will be introduced in the next study to further investigate the spatial orientation and hand–eye coordination provided by the adopted visualization system. It is worth mentioning that it is easier to embed an eye-tracking system in AM than in binocular devices. In the latter case, it is difficult to integrate the required electronics without reducing comfort of the device itself.

Acknowledgments

The authors wish to thank the resident surgeons who were asked to attend this study. They would also like to thank Mr Nicodemo Funaro for manufacturing the prototypes.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article:
This work was supported in part by the European Commission within the framework of the ARAKNES European Project EU/IST-2008-224565.

References