

Improved depth perception with three-dimensional auxiliary display and computer generated three- dimensional panoramic overviews in robot-assisted laparoscopy

Fokko P. Wieringa
Henri Bouma
Pieter T. Eendebak
Jean-Paul A. van Basten
Harrie P. Beerlage
Geert A. H. J. Smits
Jelte E. Bos

Improved depth perception with three-dimensional auxiliary display and computer generated three-dimensional panoramic overviews in robot-assisted laparoscopy

Fokko P. Wieringa,^{a,*} Henri Bouma,^a Pieter T. Eendebak,^a Jean-Paul A. van Basten,^b Harrie P. Beerlage,^c Geert A. H. J. Smits,^d and Jelte E. Bos^{a,e}

^aNetherlands Org. for Appl. Scientific Research TNO, P.O. Box 6235, 5600 HE Eindhoven, The Netherlands

^bCanisius Wilhelmina Hospital, P.O. Box 9015, 6500 GS Nijmegen, The Netherlands

^cJeroen Bosch Hospital, P.O. Box 90153, 5200 ME 's-Hertogenbosch, The Netherlands

^dRijnstate Hospital, P.O. Box 9555, 6800 TA Arnhem, The Netherlands

^eVU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

Abstract. In comparison to open surgery, endoscopic surgery offers impaired depth perception and narrower field-of-view. To improve depth perception, the Da Vinci robot offers three-dimensional (3-D) video on the console for the surgeon but not for assistants, although both must collaborate. We improved the shared perception of the whole surgical team by connecting live 3-D monitors to all three available Da Vinci generations, probed user experience after two years by questionnaire, and compared time measurements of a predefined complex interaction task performed with a 3-D monitor versus two-dimensional. Additionally, we investigated whether the complex mental task of reconstructing a 3-D overview from an endoscopic video can be performed by a computer and shared among users. During the study, 925 robot-assisted laparoscopic procedures were performed in three hospitals, including prostatectomies, cystectomies, and nephrectomies. Thirty-one users participated in our questionnaire. Eighty-four percent preferred 3-D monitors and 100% reported spatial-perception improvement. All participating urologists indicated quicker performance of tasks requiring delicate collaboration (e.g., clip placement) when assistants used 3-D monitors. Eighteen users participated in a timing experiment during a delicate cooperation task *in vitro*. Teamwork was significantly (40%) faster with the 3-D monitor. Computer-generated 3-D reconstructions from recordings offered very wide interactive panoramas with educational value, although the present embodiment is vulnerable to movement artifacts. © 2014 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.1.1.015001](https://doi.org/10.1117/1.JMI.1.1.015001)]

Keywords: endoscopic surgery; robot guided interventions; 3-D visualization and 3-D reconstruction; image perception and observer performance; image guided procedures.

Paper 13009RR received Oct. 22, 2013; revised manuscript received Apr. 15, 2014; accepted for publication Apr. 21, 2014; published online May 16, 2014.

1 Introduction

During surgical procedures, team members must focus on a variety of individual tasks, while also constantly maintaining a shared, common situational awareness with the entire team. One of the major ingredients of situational awareness in surgery is spatial perception.

Spatial perception determines greatly how people interact with their environment. It involves a combination of vision, proprioception, sense of equilibrium, and complex memory tasks [to permanently update a mental three-dimensional (3-D) map of the ambient world]. The more delicate a visually guided motor skill becomes, the more demands are placed on these perceptions. This holds especially true for extremely refined teamwork like a surgeon and assistant(s) that must perform complex eye-hand coordination in a restricted space. In such cases, optimizing shared spatial perception is crucial.

During open surgery, a major bottleneck for a shared spatial perception is getting sufficient direct visual access to the surgical area for all team members. In endoscopic surgery, direct

visual access is replaced by the indirect mechanism of the video chain.¹ Thus, adding additional video monitors helps assistants to gain visual access, which is a benefit compared to open surgery, although visualization during endoscopic surgery also has specific challenges which can be split into two main issues.

The first issue is related to impaired depth perception. Stereopsis can significantly contribute to visual perception and cognitive performance during complex tasks in a spatially complex environment.² Indeed introduction of the Da Vinci surgical robot series has considerably reduced the previous mentioned bottlenecks by offering stereopsis within the surgical console (in combination with improved dexterity of instruments and visual augmentation of the surgical field). Various reports in the existing literature have already shown that stereoscopic vision decreases surgery time and increases accuracy for the surgeon behind the console of the Da Vinci robot.^{3–10} However, to the best of our knowledge to date, there is little or no published evidence illustrating the impact of using a 3-D auxiliary monitor for the assistant(s) at the operating table during robot-assisted laparoscopy (RAL). For the assistant(s), these facilities are not

*Address all correspondence to: Fokko P. Wieringa, E-mail: f.wieringa@tno.nl

currently widely available, although it seems some 3-D auxiliary monitor systems are beginning to emerge on the market. For our study, we connected a live 3-D monitor plus 3-D record/playback system to Da Vinci robots of all three generations presently available on the market to improve shared spatial perception for assistants. We quantitatively analyzed the perceived effects for both urologists and assistants at the operating table with a questionnaire. Furthermore, we report console times for partial tumor nephrectomy, and we performed an *in vitro* experiment wherein we measured how fast a defined complex interaction task could be performed by an assistant with a two-dimensional (2-D) monitor or with a 3-D monitor (while teaming up with the surgeon using the console).

The second issue is related to impaired spatial orientation and lack of overview. The live endoscope image is limited by the optical field-of-view (FOV). In order to view the entire surgical volume inside the patient, the endoscope is therefore moved around over time. Apart from having stereoscopic video or not, both endoscopic surgeons and their assistants must constantly reconstruct and update a mental 3-D panoramic model of the surgical volume from endoscopic views in various positions. The methods to create a dense reconstruction from the input of video captured by a free moving handheld camera are already described in the literature,^{11,12} however, it has not yet been applied to endoscopic surgery, which inherently suffers from additional problems, such as strong camera distortion and a moving light source. Clinical applications of panoramic mosaics have been shown by posing an additional assumption that images can be projected on a flat 2-D model of the placenta¹³ or a spherical model of the bladder.^{14,15} Of course, such approaches are limited to applications with known and simple geometry and cannot be used in general for more complex shapes in larger environments. A recent review mentions that optical intraoperative 3-D reconstruction in laparoscopic surgery should be feasible.¹⁶ This review, however, does not show examples and results of incrementally generated wide panoramic overviews based on laparoscopic video images.

During this study, we investigated whether the complex mental spatial-orientation task to reconstruct a 3-D panorama from video images can be achieved with computer technology for different generations of the Da Vinci robot and a variety of different laparoscopic procedures.

The outline of this paper is as follows: Sec. 2 describes our methods and system architecture for display, recording, and 3-D reconstruction. Section 3 presents the experiment and results of the auxiliary 3-D monitor for depth perception and Sec. 4 presents the results of the 3-D panorama for overview and spatial orientation. Finally, Sec. 5 summarizes the overall conclusions.

2 Method

This section describes the methods that were used to obtain our results. Section 2.1 describes the hardware of the system and Sec. 2.2 focuses on the software for 3-D reconstruction.

2.1 Hardware of the System

We realized coupling of live 3-D monitors for all three generations of Da Vinci robots that are presently on the market. An overview of the generalized system architecture is shown in Fig. 1. The existing Da Vinci system has separate outputs for left and right (L&R) camera signals that enter the camera control

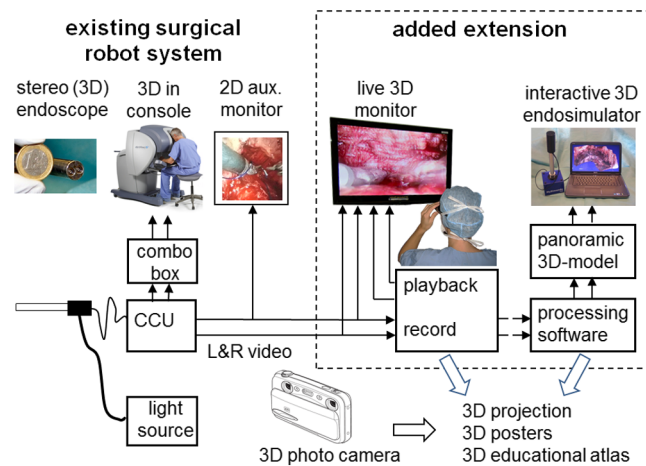


Fig. 1 Overview of the system architecture.

unit (CCU) and pass via the combo box to the surgical console for display in 3-D. The extension that was added by us routes these L&R auxiliary video output signals to a live 3-D monitor and a 3-D record/playback system. 3-D videos are recorded while moving the endoscope around in selected (still) scenes of interest. These recordings are processed offline into a panoramic 3-D model and coupled to an interactive laparoscopic camera simulator (Simendo) and so could be interactively explored for educational purposes as if freely manipulating the camera during a real laparoscopy (of course without surgical intervention).

The Da Vinci system that generates analog video streams (first generation) is connected to another hardware extension than the ones that generate digital video streams (second and third generations).

2.1.1 Hardware for the analog video streams (first generation Da Vinci)

For the analog video streams, a 3-D video converter (Image Tek Corp., Laguna Niguel, California) was used to connect a 3-D monitor (Panasonic TX-P50VT20E, utilizing active LCD-shutter glasses, Osaka, Japan) to the left and right S-video outputs of a first generation Da Vinci Robot. For routine 3-D recording by users, an high-definition digital versatile disc (HD-DVD) recorder was connected (Panasonic DMR-8H63). The resulting recordings are directly playable on a Blu-ray DVD system, but maintain only about half of the resolution due to interlacing the L&R video signal. Three-dimensional recordings for research on 3-D panoramic reconstructions were made using two synchronized Pleora iPORT digitizing engines (PT1000-ANL-1-6-V2-E) connected to a DELL XPS laptop with RAID hard disk system and 2 × GB-LAN (1 connector and 1 express card). The latter method preserves maximum resolution without video compression but is less user friendly and not Blu-ray compatible.

2.1.2 Hardware for the digital video streams (second and third generations Da Vinci)

For digital video streams, a medical grade 3-D monitor (Sony LMD-2451MT, utilizing passive rotational polarization glasses, Tokyo, Japan) was directly connected to the L&R spare high-definition serial digital interface (HD-SDI) outputs of the L&R

CCUs from a second generation Da Vinci S robot. The same type 3-D monitor was also connected to a third generation Da Vinci Si robot. Since the CCUs of the Da Vinci Si did not have a spare HD-SDI video output but only a spare YPbPr video output instead, we applied two converters for the L&R signal (AJA video, HD10A-R0). Three-dimensional video recordings were performed using two synchronized Panasonic P2 broadcast studio recorders (AG-HPD24E) which store L&R in full HD without any video compression.

Note: Achieving a live 3-D view on the auxiliary monitor does not require 3-D recorder hardware nor the software for 3-D panoramic reconstruction.

2.2 Software for 3-D Reconstruction

The applied 3-D reconstruction method is an add-on to existing endoscopy systems and basically compatible with any video endoscope. By distilling spatial information from the video stream while the endoscope moves around, it builds a much wider cumulative overview than the normal momentary endoscopic FOV (just as humans do in their mind). Thus by scanning a bodily cavity, a much wider 3-D panorama can be reconstructed (even if the input data comes from a conventional 2-D endoscope). There is no need for position tracking, additional markers, or prescans from magnetic resonance (MR) or computed tomography (CT), because reconstruction is purely based on information from the endoscopic video-stream. Calibration is simply achieved from filming a black and white checkerboard pattern at various angles.

The (offline) 3-D reconstruction processing consists of four main steps, which are shown in Fig. 2. Recorded streaming video (monoscopic or stereoscopic) from the endoscope system

forms the system input. The first processing step is motion estimation, which computes the endoscope's changes in position and orientation. Based on this information, the second processing step can reconstruct dense sets of visible surface points in 3-D. In the next step, these 3-D points are converted into 3-D surfaces (mesh pattern). Finally, texture mapping is applied with a photorealistic detailed 3-D panorama as a result. An extensive further technical description of the applied image processing method is described in "Streaming video-based 3-D reconstruction method compatible with existing monoscopic and stereoscopic endoscopy systems."¹⁷

3 Results of the Auxiliary 3-D Monitor for Depth Perception

This section describes the experiments and results related to the auxiliary 3-D monitor for depth perception, which includes the experimental setup for the monitor for surgical procedures (Sec. 3.1), results of the questionnaire (Sec. 3.2), console times for partial tumor nephrectomy (Sec. 3.3), the *in vitro* experimental setup for time measurement for a delicate cooperation task (Sec. 3.4), and finally, the results of this time measurement (Sec. 3.5).

3.1 Experimental Setup for the Auxiliary 3-D Monitor for Depth Perception

We connected a live 3-D monitor plus 3-D video record/playback system to all three Da Vinci robot generations presently available. This setup was used for two years, following which we probed user image perception by questionnaire. In total, 925 consecutive RAL procedures in three hospitals were

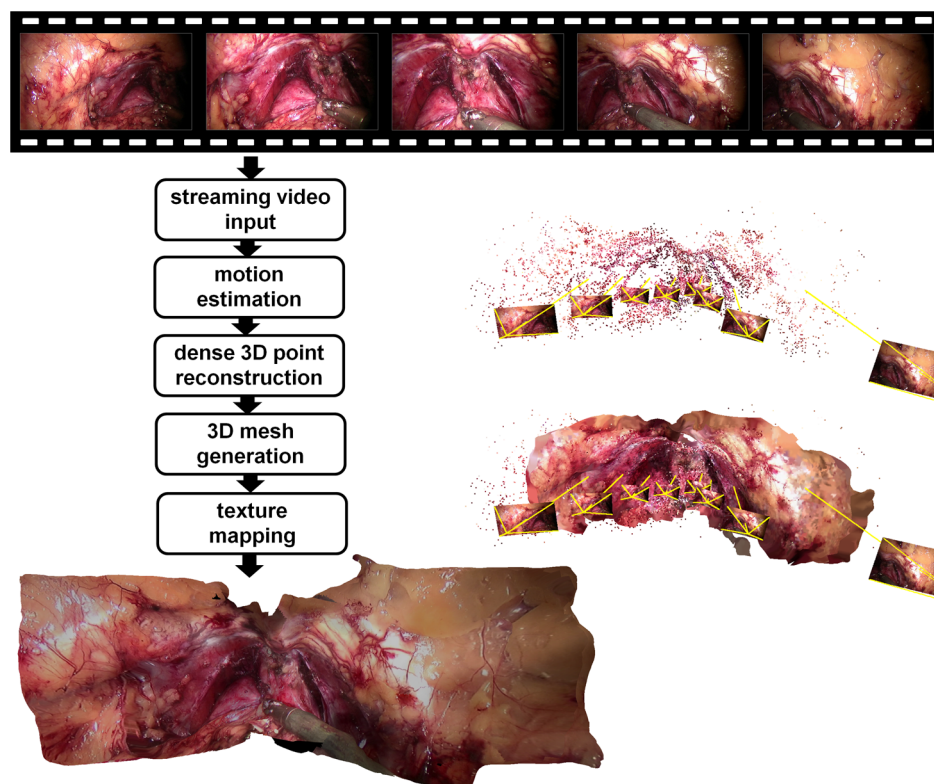


Fig. 2 Block diagram of software processing steps to construct a three-dimensional (3-D) panoramic model from video streams of camera moving around in a scene of interest.

included (41% with an analog Da Vinci and 59% digital). Surgical procedures involved RAL-prostatectomy (RALP) ($n = 748$), RAL-cystectomy (radical $n = 66$ radical, partial $n = 3$), RAL-partial nephrectomy ($n = 58$), RAL-pyeloplasty ($n = 23$), RAL-nephro-ureterectomy ($n = 13$), RAL-implantation of brachytherapy¹⁸ cannulas ($n = 7$), and others ($n = 7$).

Users were free to use or not use 3-D monitors. Conventional video monitors were kept operational alongside 3-D monitors (3-D glasses did not impair viewing normal monitors). After two years of experience, a questionnaire was distributed among 31 users. Standard surgical procedures were unchanged.

3.2 Results of the Questionnaire about the Auxiliary 3-D Monitor

All participating urologists within the three participating hospitals indicated quicker performance of tasks requiring delicate spatial co-operation, like clip placement and handing over of suture needles. In order to obtain more quantitative information, a survey on perceived 3-D image quality was held among all users. The results are listed in Table 1. This table contains an almost literal translation into English from the original questions in Dutch.

Thirty-one users participated in our questionnaire. The results show that in the role of scrub nurse, 16% preferred using 2-D monitors and the other 84% preferred 3-D monitors.

Fischer's exact test showed that the chance to find 26 responses in favor of the 3-D monitor out of 31 by mere coincidence is $<8 \times 10^{-5}$, which is why the null hypothesis that there is no difference can be rejected. Furthermore, all users (100%) reported qualitative spatial-perception improvement compared to normal 2-D monitors, which is even more significant ($p = 5 \times 10^{-10}$).

All 31 respondents of the questionnaire indicated that the live 3-D monitor offered a clearly enhanced visual depth perception in comparison to a normal 2-D monitor, thus indicating that all respondents had functional stereopsis (although we did not measure this). Yet five of these 31 respondents (16%) indicated they preferred using the 2-D monitor when assisting surgery. Based on the free text from the questionnaire, their underlying reasons could be divided into two subgroups.

The first group mentions practical issues related to the 3-D glasses. This subgroup consisted of two respondents (6%). Both used active glasses (LCD-shutter type, first generation Da Vinci). One respondent indicated that the active glasses were quite heavy and she could not reposition them when working sterile at the operating table. The other respondent indicated that when looking down on the instruments and then back up to the 3-D monitor, the active glasses had to resynchronize with the monitor, which was distracting. Hence, both did not use the 3-D monitor when working as a scrub nurse. They preferred, however, using the 3-D monitor when assisting as

Table 1 User experience with auxiliary 3-D monitor over 2 years (925 RAL procedures). Note that between procedures, people can switch between multiple roles.

Question	Response options	Number
In which role do you work with the auxiliary 3-D monitor? (in case of multiple roles check all relevant categories)	Assistant (circulating)	16
	Assistant (scrub)	24
	Viewer (for education)	2
	Console surgeon	10
Via which auxiliary monitor do you prefer to work?	Normal monitor	5 (16%)
	3-D monitor	26 (84%)
	No preference	0
Does the 3-D monitor offer you improved spatial perception compared with the normal monitor?	Yes, very clearly	31 (100%)
	To some extent	0
	I do not see a difference	0
	The 3-D monitor offers worse spatial perception	0
Please indicate any issues with regard to the 3-D monitor that you find relevant. You may fill out your ID, but this is NOT mandatory. Please do not return more than one form per person	Received free text is summarized in Sec. 3.2	
Respondents per category	Assistants	20
	Educational viewers	1
	Urologists	9
	Oncologists	1
	Grand total respondents	31

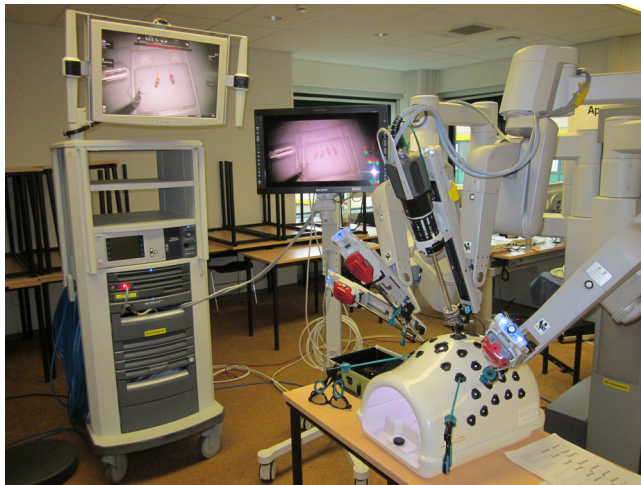


Fig. 3 Overview of the experimental setup that was used to measure the time needed to perform a task with two-dimensional (2-D) and 3-D monitors.

a circulating nurse. In particular, the passive glasses (used with the digital monitor) allowed multiple assistants to work with the 3-D monitor in a wide viewing range. For active glasses, the viewing range was less wide, but perceived as acceptable.

The second group mentions physiologic discomfort issues. This subgroup contained three respondents (10%) that reported uncomfortable feelings like headache, dizziness, and forthcoming slight nausea. Although this is a well-known phenomenon, scientific reports on incidences are scarce and, for complex matters, such effects vary between different (continuously evolving) 3-D display technologies. Estimations indicate that up to 30% of 3-D cinema attendants experience some discomfort due to accommodation-vergence conflicts and 1/500 even has to vomit due to visual-vestibular conflicts, which compares to 1/5000 of regular 2-D cinema attendants.¹⁹ Moreover,

approximately up to 30% of the population has impaired stereopsis and 3% has no stereopsis at all, often without knowing this or being troubled by it in daily life.²⁰

Further, a few normal recommendations about monitor height and tilt adjustments were received.

3.3 Console Times for Partial Tumor Nephrectomy with the Auxiliary 3-D Monitor

We inspected the console times of partial tumor nephrectomies closely before and after switching from 2-D to 3-D, because this is a very time-critical procedure with intense spatial 3-D teamwork. Here, we assume the learning curve to be negligible, because all procedures were performed by an experienced urologist and an experienced team. All kidney-preserving operations were performed by the same urologist in the same hospital (Rijnstate), which avoids additional factors that could influence the experimental outcome. We inspected the times of the last 13 patients that were treated with a 2-D monitor and the first 13 patients that were treated with the 3-D monitor. The console time in 2-D appears to be (on average \pm std) 138 ± 25 min and in 3-D 108 ± 26 min, which is an improvement of 22%. A student *T*-Test was performed using the one-sided distribution and unequal variances. The test shows that the console times in this *in vivo* experiment are significantly lower with the 3-D monitor than with the 2-D monitor ($p = 0.003$).

3.4 Experimental Setup to Measure Time to Perform an *In Vitro* Task with the Auxiliary Monitor

We also performed an *in vitro* experiment to compare the time that was needed to perform a defined task viewing a 2-D auxiliary monitor with the time needed to perform the same task viewing a 3-D monitor.

We connected a live 3-D monitor to a Da Vinci Si robot that generated HD-SDI video. All participating assistants performed the same task twice; once in 2-D (using 2-D mode of the 3-D

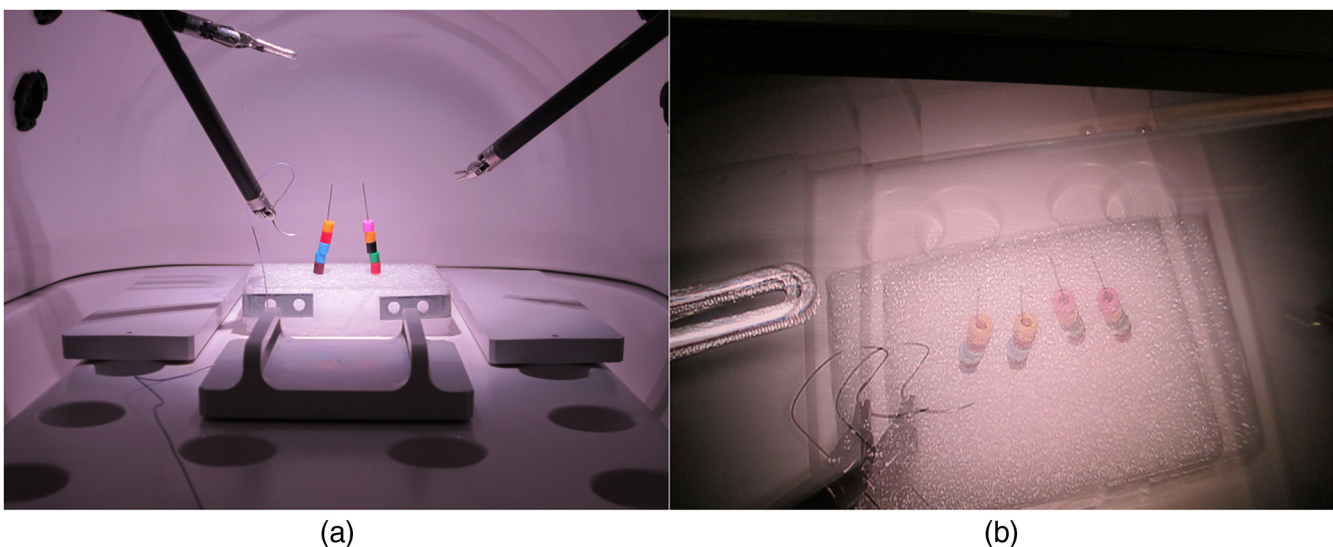


Fig. 4 (a) Details of the experimental setup that was used to measure the time. The beads were available on long straight needles inside a surgical test box and must be picked up and moved toward the curved needle. (b) 2-D photo of the 3-D monitor in 3-D mode (hence the “double” picture). This monitor was switched between 2-D and 3-D mode during the experiment.

monitor) and once in 3-D (using 3-D mode on the same monitor). The same monitor was used in all experiments to avoid the influence of changing monitor quality. To compensate for the effects of a learning curve, assistants started alternating with either 2-D view or 3-D view to obtain balanced results. An overview of the setup is shown in Fig. 3 and the details are shown in Fig. 4.

Laparoscopic or robot assisted transfer of beads or pegs is a well-determined validated test to assess bimanual dexterity, hand-eye coordination, and spatial awareness, as has been described in a validation study of Tjiam et al.²¹ The procedure that we used was performed as follows. Before the experiment started, the assistant was allowed to practice grasping one bead and putting it on the needle. In the starting position of the experiment, the assistant had the curved needle in the grasper. When the time measurement started, the following subsequent steps were repeated: First, the assistant presents the curved needle to the grasper of the Da Vinci (operated by the console surgeon), then the assistant picks up a bead from the stock-pile on long straight needles, the assistant puts the bead on the curved needle (the console surgeon was not allowed to help), the assistant takes the curved needle at the tip and the console surgeon detaches the needle (so that the bead can move further on the wire), and finally, the assistant returns the curved needle to the grasper of the Da Vinci console. Then the cycle is repeated with the next bead. The stop criterion is reached when three min passed, or when 10 beads were successfully put on the curved needle, or when no more beads were available on the long straight needles. When the three min were used, the assistant was allowed to finish this last bead. The number of beads on the curved needle and thread were counted and the time it took to place those beads was reported to avoid discretization effects due to “half” beads. In this way, the average time per bead was computed for each user.

3.5 Results of In Vitro Timing Experiment with the Auxiliary Monitor

In total, 18 volunteers participated as assistant, all performed the experiment twice ($N = 18$ in 2-D mode and $N = 18$ in 3-D

mode), and in total, 280 beads were successfully placed on the curved needle (121 beads in 2-D mode and 159 beads in 3-D mode). The results are summarized in Tables 2 and 3.

Before the experiment, we had two expectations. The first and main expectation was that people are faster with a 3-D monitor than with a 2-D monitor. Second, there may be a learning effect which must be taken into account for a fair comparison. We used analysis of variance to analyze these factors in this balanced dataset and the assumptions were tested at the threshold of $p = 0.01$. The difference between the 2-D mode and the 3-D mode appears to be highly significant [$F(1, 32) = 13.6$, $p = 0.0008$]. On average, people are 40% faster when they use the 3-D monitor for a complex interaction task. Neither the difference between first and last appears to be significant [$F(1, 32) = 0.78$, $p = 0.38$] nor the interaction between these factors [$F(1, 32) = 0.05$, $p = 0.82$]. Because no order effect was found (no main effect nor interaction), it can be concluded that the main effect of 3-D versus 2-D was independent of the order in which the tests were applied. Of course, a longer task may lead to different learning effects and we do not expect that it will be impossible to improve skills.

Table 3 shows that for a low number of participants (11%) there is hardly a difference between using 2-D and 3-D. However, a large majority (89%) performs >10% faster and some (28%) even become twice as fast by using the 3-D monitor.

Table 2 already showed that (on average) people are 40% faster when they use the 3-D monitor for a complex interaction task. It is important to note that this does not mean that a 3-D monitor will necessarily reduce the whole surgical procedure (from skin to skin) to 60% of the total time, since only a fraction of the surgical procedures will require these complex tasks of the assistant. The *in-vitro* results match well with the impressions of the console surgeons that teamwork was quicker with the use of the 3-D facility (Sec. 3.2) and with the observed improvement *in-vivo* console time for partial tumor nephrectomies (Sec. 3.3).

During the *in vitro* experiment with beads, we noticed a similar behavior as during *in vivo* clip placement. When using a 2-D monitor, assistants placed the bead and clip by “trial-and-error.”

Table 2 Average time needed per bead (sec/bead) and the standard deviation in 2-D or 3-D mode and as first or last task. In the 2-D column, “First” indicates that the volunteer started with 2-D mode. The difference between 2-D and 3-D mode appears to be significant.

	2-D mode	3-D mode	2-D + 3-D mode
First	30.5 ± 15.5 ($N = 9$)	16.1 ± 3.9 ($N = 9$)	23.3 ± 13.2 ($N = 18$)
Last	27.0 ± 7.4 ($N = 9$)	18.2 ± 6.7 ($N = 9$)	22.6 ± 8.2 ($N = 18$)
First + last	28.8 ± 11.9 ($N = 18$)	17.1 ± 5.4 ($N = 18$)	23.0 ± 10.9 ($N = 36$)

Table 3 Histogram of people that obtain a relative time gain (gain = t_{3-D}/t_{2-D} , where time is defined as the average time per bead (s/bead) for 2-D (t_{2-D}) or 3-D (t_{3-D}). This table shows that 28% is at least twice as fast with 3-D, 89% has a benefit of >10% in using 3-D, and nobody has a benefit of >10% in using 2-D.

Time gain	$t_{3-D}/t_{2-D} < 0.5$ (%)	$0.5 < t_{3-D}/t_{2-D} < 0.7$ (%)	$0.7 < t_{3-D}/t_{2-D} < 0.9$ (%)	$0.9 < t_{3-D}/t_{2-D} < 1.1$ (%)	$1.1 < t_{3-D}/t_{2-D}$ (%)
People	28	39	22	11	0
Cumulative	28	67	89	100	100

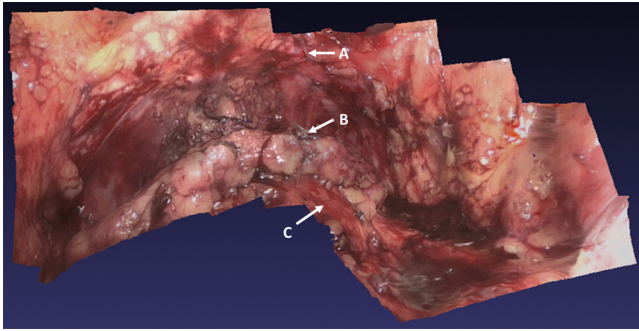


Fig. 5 Three-dimensional panorama of a prostatectomy based on an analogue video stream showing an intraoperative view of the (a) symphysis and the (b) subsymphyseal area postprostatectomy. The (c) bladder is visualized at the bottom.

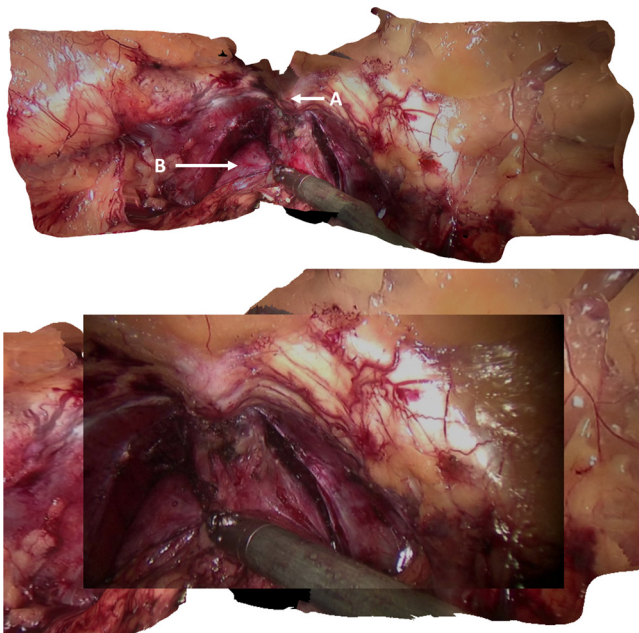


Fig. 6 (Top) 3-D panorama of a prostatectomy based on a digital video stream showing a view on the (a) small pelvis just after opening the (b) endopelvic fascia. (Bottom) Magnified portion of the same panorama with one original video frame overlaid upon the model. This clearly illustrates the large field-of-view and fine details of the panoramic reconstruction compared to the normal image.



Fig. 7 (Left) 3-D panorama showing a (A) tumor in the upper pole of the right kidney, prior to resection. A vessel loop pulls the (B) renal veins forward, disclosing the renal artery branch. (Right) The same 3-D panoramic model, but now rotated to a different viewing angle. This clearly illustrates the 3-D shape information in the model. When loaded into the surgical simulator, all models could be freely explored with the camera joystick, with a display setting that could be set to 3-D or 2-D.

By this, we mean that the assistant often aims to put the bead over the needle, but fails a few times due to incorrect depth estimation, and thus moves in empty space. After observing this, the assistant then pulls back and tries again at another distance to the camera. This behavior was seen during the *in vitro* experiment and also during *in vivo* clip placement using a 2-D monitor. We did not see this “trial-and-error” approach with the 3-D monitor, but observed a first-time-right placement of the bead over the needle. Also, here, there was a parallel with typical *in vivo* behavior during clip placements: Using a 3-D monitor, first-time-right spatial positioning of clips was observed.

4 Results of the 3-D Panoramas for Overview and Spatial Orientation

This section describes the experimental setup (Sec. 4.1), qualitative results (Sec. 4.2), and discussion (Sec. 4.3) of the 3-D panoramas for overview and spatial orientation.

4.1 Experimental Setup of the 3-D Panoramas for Overview and Spatial Orientation

In order to create the panoramas, synchronized L&R video was recorded during surgery and reconstructed offline after surgery. In this paper, we focus on reconstructions from three different hospitals. We describe a prostatectomy with an analogue (first generation) Da Vinci, a prostatectomy with a digital (second generation) Da Vinci S, and the removal of a kidney tumor with a digital (third generation) Da Vinci Si.

4.2 Results of the 3-D Panoramas for Spatial Orientation and Overview

As for panoramic 3-D reconstruction, Figs. 5–7 provide examples of the wide overviews that can be obtained with the present embodiment. Figure 5 is based on 15 s of video and shows an intraoperative view of the symphysis (top) and the subsymphyseal area postprostatectomy. Figure 6 is based on 20 s of video and shows a view of the small pelvis during RALP, just after opening the endo-pelvic fascia. Within Fig. 6 (bottom) one original full-size video frame is overlaid on its original location within the zoomed-in reconstructed panorama. This clearly illustrates that the complete panorama of Fig. 6 (top) contains a much wider FOV than the individual original live video frames. In Figs. 5 and 6 the FOV of the reconstructed panoramas are approximately 4 respectively 3 times larger than the individual live video frames. Note that the live video image loses brightness toward the edges, whereas the panoramic 3-D reconstruction offers overall well-lighted contrasts. The reconstruction software can achieve this by combining the image regions

with optimum contrasts as the scope tip is moved around to scan the entire area. Figure 7 is based on 18 s of video and shows a kidney tumor in the upper pole prior to resection. A vessel loop pulls the renal veins forward, disclosing the renal artery branch. The rotation in Fig. 7 (left and right) clearly shows the 3-D information in the panoramic model.

4.3 Discussion of the 3-D Panoramas for Overview and Spatial Orientation

Interactive 3-D panoramas may lead to new possibilities for mentoring and teaching. During practice with an endoscopic simulator, the method can, for example, visualize position and orientation of the simulated endoscope relative to the environment as viewed from a virtual mentor viewpoint outside the simulated endoscope (but inside the patient), which may help the viewers to better see the anatomical structures and improve situational awareness among the surgical team.

A limitation of the present embodiment for 3-D reconstructions is that these presently are generated offline from recorded data. This, however, is not a principal limitation but merely an engineering issue of implementing the embedded processing. Movement artifacts can also arise from objects (e.g., instruments) that move within the FOV while recording. For some inspection procedures, the latter can be avoided relatively easily. The dynamic update of the existing model after motion in the environment is not part of the current embodiment. If extended to real-time processing, the method may well be applied to, for example, 3-D mapping of bladder wall inspections [e.g., in repeated follow-ups after transurethral resection of tumor (TUR)].

The 3-D panoramic model is not yet currently used during the forms of laparoscopic surgery mentioned in this paper. In other clinical procedures, such as twin to twin transfusion syndrome (TTTS) or bladder operations, there is a clear benefit in first creating an overview before subsequently starting the actual treatment. We learned that the technique can be relevant in laparoscopy for educational purposes and reporting, and it is expected that the described technique will become even more useful for laparoscopic applications when it is combined with other 3-D modalities (e.g., US-echoprobe or CT/MR-prescan). For 3-D reconstructions, broad educational applications (also outside the scope of urology) are envisaged. Simulators are rapidly evolving and we are entering a new era of proficiency-based endosurgical training.^{22,23} 3-D environments from real patients can be imported into interactive endoscopic simulators and used for e-learning, showing 3-D views that respond to manipulation of an endoscopic camera simulator joystick, which can enhance surgical training.²⁴

5 Conclusions

Provision of a 3-D auxiliary monitor improves the motor co-ordination between console surgeon and assisting staff: 84% of team members preferred 3-D monitors and perceived them as enhancing teamwork, all respondents (100%) indicate that the live 3-D monitor offered a clearly enhanced visual depth perception in comparison to a normal 2-D monitor, and all participating urologists indicated quicker performance of tasks requiring delicate spatial co-operation (e.g., clip placement). This is supported by console times of the partial tumor nephrectomy (*in vivo*) and the timing experiment that was performed *in vitro*. The results of this experiment showed that assistants were,

on average, 40% faster using a 3-D monitor, which is a highly significant improvement.

Furthermore, we reconstructed photorealistic interactive 3-D environments generated from laparoscopic video data of different hospitals, different generations of the Da Vinci robot, and different surgical procedures. These 3-D reconstructions offered very wide interactive panoramas, but the present embodiment is vulnerable to movement artifacts.

Future work may include the measurement of the difference between analog and digital systems, or the measurement of latency issues with the different systems.

Acknowledgments

The authors acknowledge the support of OR-staff and Medical Technology Department from Canisius Wilhelmina Hospital (in particular, John Jansen), Jeroen Bosch Hospital (in particular, Bas Kraneveld), and Rijnstate Hospital (in particular, Cor Evers and Carl Wijburg). The authors also thank the Dutch foundation for endourology (SWEN) as well as the Dutch urological association (NVU). Furthermore, the authors thank the companies Mpluz, Sony Medical Europe, and Intuitive Surgical for providing us with the necessary information about hardware aspects and video connectivity issues. The authors also thank the firm Simendo for assistance in connecting their virtual laparoscopic camera manipulator joystick and the firm Bourgonje for facilitating 3-D congress projection.

References

1. R. Berguer, "Surgery and ergonomics," *Arch. Surg.* **134**(9), 1011–1016 (1999).
2. J. E. Capó-Aponte et al., "Visual perception and cognitive performance," in *Helmet-Mounted Displays: Sensation, Perception and Cognition Issues*, C. E. Rash et al., Eds., pp. 335–390, U.S. Army Aeromedical Research Laboratory, Fort Rucker, Alabama (2009).
3. M. Beurden, W. IJsselstein, and J. Juola, "Effectiveness of stereoscopic displays in medicine: a review," *3D Res.* **3**(3), 1–13 (2012).
4. M. Beurden et al., "Stereoscopic displays in medical domains: a review of perception and performance effects," *Proc. SPIE* **7240**, 72400A (2009).
5. A. Blavier et al., "Impact of 2D and 3D vision on performance of novice subjects using Da Vinci robotic system," *Acta Chir. Belg.* **106**(6), 662–664 (2006).
6. J. Byrn et al., "Three-dimensional imaging improves surgical performance for both novice and experienced operators using the Da Vinci Robot System," *Am. J. Surg.* **193**(4), 519–522 (2007).
7. V. Falk et al., "Influence of three-dimensional vision on surgical telemanipulator performance," *Surg. Endosc.* **15**(11), 1282–1288 (2001).
8. S. Kong et al., "Comparison of two- and three-dimensional camera systems in laparoscopic performance: a novel 3D system with one camera," *Surg. Endosc.* **24**(5), 1132–1143 (2010).
9. Y. Munz et al., "The benefits of stereoscopic vision in robotic-assisted performance on bench models," *Surg. Endosc.* **18**(4), 611–616 (2004).
10. O. J. Wagner et al., "Three-dimensional vision enhances task performance independently of the surgical method," *Surg. Endosc.* **26**(10), 2961–2968 (2012).
11. Y. Chen et al., "Live accurate and dense reconstruction from a handheld camera," *Comp. Anim. Virtual Worlds* **24**(3–4), 387–397 (2013).
12. R. Newcombe, S. Lovegrove, and J. Davison, "DTAM: dense tracking and mapping in real-time," in *Proc. IEEE ICIP*, pp. 2320–2327, IEEE, Brussels, Belgium (2011).
13. M. Reef et al., "Mosaicing of endoscopic placenta images," *Informatik P-93*, 467–474 (2006).
14. A. Behrens et al., "2-D and 3-D visualization methods of endoscopic panoramic bladder images," *Proc. SPIE* **7964**, 796408 (2011).
15. T. D. Soper, M. P. Porter, and E. J. Seibel, "Surface mosaics of the bladder reconstructed from endoscopic video for automated surveillance," *IEEE Trans. Biomed. Eng.* **59**(6), 1670–1680 (2012).

16. L. Maier-Hein et al., "Optical techniques for 3D surface reconstruction in computer-assisted laparoscopic surgery," *Med. Image Analysis* **17**(8), 974–996 (2013).
17. H. Bouma et al., "Streaming video-based 3D reconstruction method compatible with existing monoscopic and stereoscopic endoscopy systems," *Proc. SPIE* **8371**, 837112 (2012).
18. K. Hendricksen et al., "Technique of robot-assisted brachy therapy catheter implantation in the bladder," *BJU Int.* **110**(Suppl. 3), 148–149 (2012).
19. J. E. Bos et al., "Cinera sickness and postural instability," *Ergonomics* **56**(9), 1430–1436 (2013).
20. R. van Ee and W. Richards, "A planar and a volumetric test for stereo-anomaly," *Perception* **31**(1), 51–64 (2002).
21. I. M. Tjiam et al., "Program for laparoscopic urological skills: a newly developed and validated educational program," *Urology* **79**(4), 815–820 (2012).
22. R. M. Satava, "Surgical education, simulation, and the revolution in training: personal observations and predictions," *Laparosc. Today* **3**(2), 7–8 (2005).
23. F. P. Wieringa et al., "Visualization in the operation room: from looking good to looking better," in *Congress of the Dutch Association for Endoscopic Surgery NVEC*, NVEC, Arnhem, The Netherlands (2012).
24. J. M. Luursema et al., "The role of stereopsis in virtual anatomical learning," *Interact. Comput.* **20**(4–5), 455–460 (2008).

Fokko P. Wieringa received his PhD degree in biomedical engineering at Erasmus University in The Netherlands. He is a senior scientist at TNO. He is a serial innovator, with various patents on medical devices and semiconductor manufacturing tools. Presently, he works in the Van 't Hoff Program, performing shared research in medical photonics. He is also involved in writing standards on medical devices for IEC and CEN and heavily interested in human factor engineering.

Henri Bouma received his PhD degree in biomedical image analysis at the TU Eindhoven in The Netherlands. He is a research scientist at

TNO in the field of computer vision and pattern recognition for military and medical applications. He is experienced with accurate blood vessel quantification, computer-aided diagnosis and medical 3-D reconstructions. He led and participated in projects about object recognition, behavior analysis, and person reidentification for the Dutch Police and Ministry of Defense.

Pieter T. Eendebak received his PhD degree in mathematics at Utrecht University in The Netherlands. He is a research scientist at TNO in the field of computer vision and 3-D data processing.

Jean-Paul A. van Basten is a urologist at CWZ hospital in Nijmegen, The Netherlands. He is a specialist in robotic and laparoscopic urology in the division of minimal invasive surgery. He is a console surgeon and proctor of laparoscopic urological surgery and robotic surgery.

Harrie P. Beerlage received his PhD on the thesis titled "Diagnosis and minimal invasive treatment of prostata carcinoma" in 1999. He is a urologist at Jeroen Bosch Hospital in 's-Hertogenbosch, The Netherlands. After his training, he worked in the Urology Department at Radboud University in Nijmegen, where he further specialized in oncology and laparoscopy. Since 2001, he has worked in 's-Hertogenbosch and is focusing on robotic surgery in oncological urology.

Geert A. H. J. Smits, PhD, FEBU, is a urologist at Rijnstate hospital in Arnhem, The Netherlands. He is specialized in robot-assisted laparoscopy and uro-oncology. He is a chairman of the Dutch Robotic Urology Section and Education Section within SWEN.

Jelte E. Bos is a researcher at TNO and professor at the Faculty of Human Movement Sciences of VU University in Amsterdam. His expertise concerns applied and scientific knowledge regarding effects of motions on humans in general, and spatial orientation and motion sickness in particular. This concerns physical and visual motion, and the effects concern health, safety, performance, and comfort.