Stereoscopic Augmented Reality System for Supervised Training on Minimal Invasive Surgery Robots

Florin Octavian Matu
Aalborg University
Fredrik Bajers Vej 7B, 9220
Aalborg, Denmark
fmatu13@student.aau.dk

Mikkel Thøgersen
Aalborg University
Fredrik Bajers Vej 7B, 9220
Aalborg, Denmark
mthoge10@student.aau.dk

Bo Galsgaard
Aalborg University
Fredrik Bajers Vej 7B, 9220
Aalborg, Denmark
bgalsg13@student.aau.dk

Martin Møller Jensen
Aalborg University
Fredrik Bajers Vej 7B, 9220
Aalborg, Denmark
mmje10@student.aau.dk

Martin Kraus
Aalborg University
Sofiendalsvej 11, 9200,
Aalborg, Denmark
martin@create.aau.dk

ABSTRACT

Training in the use of robot-assisted surgery systems is necessary before a surgeon is able to perform procedures using these systems because the setup is very different from manual procedures. In addition, surgery robots are highly expensive to both acquire and maintain — thereby entailing the need for efficient training. When training with the robot, the communication between the trainer and the trainee is limited, since the trainee often cannot see the trainer.

To overcome this issue, this paper proposes an Augmented Reality (AR) system where the trainer is controlling two virtual robotic arms. These arms are virtually superimposed on the video feed to the trainee, and can therefore be used to demonstrate and perform various tasks for the trainee. Furthermore, the trainer is presented with a 3D image through a stereoscopic display. Because of the added depth perception, this enables the trainer to better guide and help the trainee.

A prototype has been developed using low-cost materials and the system has been evaluated by surgeons at Aalborg University Hospital. User feedback indicated that a 3D display for the trainer is very useful as it enables the trainer to better monitor the procedure, and thereby enhances the training experience. The virtual overlay was also found to work as a good and illustrative approach for enhanced communication. However, the delay of the prototype made it difficult to use for actual training.

Categories and Subject Descriptors

J.3 [**Life and Medical Sciences**]: Health; H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems — *Artificial, augmented, and virtual realities*; I.2.9 [**Artificial Intelligence**]: Robotics — *Operator Interfaces*

Figure 1: The da Vinci Surgical System. From left to right: the surgeon console, the patient/surgical cart with a monitor used for drawing and finally the video cart. ©[2014] Intuitive Surgical, Inc.

General Terms

Experimentation, Human Factors, Performance, Verification

Keywords

Augmented Reality, Medical Training, Virtual Reality, Stereoscopic Displays, Minimal Invasive Surgery

1. INTRODUCTION

Robotic surgical systems are a fairly new technology that enables complex endoscopic surgery in a minimally invasive way [6]. These systems bring higher accuracy and performance compared to that of traditional endoscopic surgeries. In contrast to traditional/manual endoscopic surgery, the surgeon is positioned at a surgeon console. In the console, the surgeon is presented with a 3D-image from a stereoscopic camera on the patient cart and is able to control two robotic arms situated on the patient cart. Using this 3D video feed and the robotic arms, the surgeon can perform surgeries with high precision, Fig. 1.

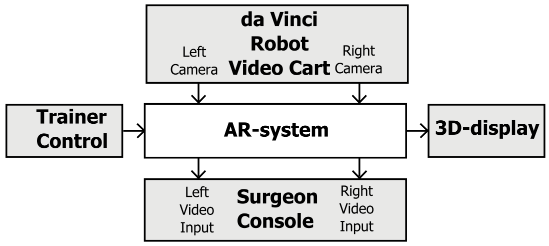

Figure 2: Block diagram of the proposed system

Prior training is necessary for a surgeon to be able to perform robotic surgery as the setup is essentially different from that of manual surgery.

Training is often carried out in multiple steps, where many of these steps include supervision from a skilled surgeon, who therefore acts as a trainer. At Aalborg University Hospital training is split into three stages, that the trainee will go through: 1) training using virtual reality, i.e. the trainee will be moving around and controlling objects in a virtual setup, 2) performing surgery on pigs using the da Vinci Surgical System, often monitored and guided by an experienced robotic surgeon, 3) finally, the trainee will perform surgery on humans while monitored by an experienced robotic surgeon. In step 2) and 3) the trainer is able to give visual input to the trainee by drawing on a monitor, see Fig. 1, which creates a 2D overlay for the trainee. Secondly, an additional surgeon console can be connected during complex surgery training. This makes it possible for the trainer to see the same 3D image as the trainee. Furthermore, the trainer is able to take over control of the surgery robot in cases where the expertise of the trainee is lacking and/or the communication between the two surgeons is lacking. The use of two consoles requires the purchase of an extra console and takes up extra space in the operating room.

This paper presents an enhanced and optimized training solution using AR, and is aimed to function with the da Vinci Surgical System made by Intuitive Surgical Inc. The surgery system is considered the market-leader and is being used widely for different medical procedures all over the world [7]. The da Vinci system is depicted in Fig. 1.

Related work
AR is already being applied in many different training situations [5]; for example, an AR solution performed better than conventional 2D drawings for assembly training [2].

The use of an external pair of controllers to mimic the control of the da Vinci Robot has been researched [4]. This research showed, that a pair of off-the-shelf game controllers can effectively simulate the controls of the da Vinci. The authors of the research paper [4] have kindly supplied their source code. This code includes graphic models of two virtual robotic arms and the interface to the controllers. These parts of their code are used in the implementation of the prototype of this project.

Previous work on superimposing virtual overlay onto live video signals from the da Vinci has been discussed by Asaki Hattori and Naoki Suzuki [1]. They focused on visualizing the inner structure of various organs during surgery. Their solution was used for a phantom operation with varying results and it did not include any immersive training.

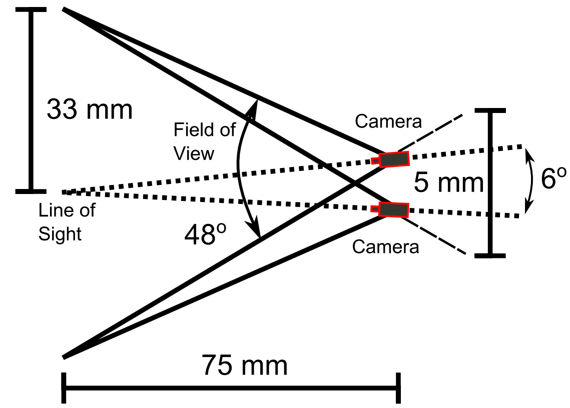

Figure 3: Setup of optical system of the endoscopic camera on the da Vinci patient cart [3].

2. METHODS AND MATERIALS

The concept of the system presented here is to extend on the current da Vinci system by utilizing AR to enhance the communication during training. Communication from the trainer to the trainee is currently done verbally and by visually drawing lines on a monitor which are displayed to the trainees view in the surgeon console. The trainer can only provide instructions based on what he can see on this same monitor. Instead of being limited to simple 2D drawings, the purpose of the proposed system is to enhance the visual communication by giving both the trainer and trainee familiar interfaces and depth perception. This is achieved by two means: 1) by providing a 3D display to the trainer, enabling depth perception and 2) by creating a virtual environment with two virtual robotic arms, similar to those situated on the da Vinci patient cart. These virtual robotic arms are controlled by the trainer and can be seen in both the trainers and the trainees views as an augmentation of reality. By superimposing the virtual environment onto the video footage from the endoscopic camera on the patient cart, an immersive experience is created for both the trainer and trainee, as both can see the combination of the footage from the camera and the virtual robotic arms.

In Fig. 2, a simplified overview of the system is presented, showing the main elements and the arrows indicating the direction of data. The signal being processed in the system is a dual analog video feed from the stereoscopic camera. These video feeds are combined with the virtual overlays created by the computer and are then fed back into the da Vinci console. Furthermore, the video feeds are duplicated to an external stereo display for the trainer. Thus, both the trainee in the console and the trainer are seeing the same stereo image. Hereby the trainer gains depth perception and the trainee gets instructions which are visualized by the trainer using two virtual replicas of the real da Vinci robot arms, see Fig. 5.

2.1 Augmented Stereoscopic Video

AR is obtained by supplementing a video feed with virtual objects. In the case of stereoscopic video, the virtual objects are captured from two different points of views, and then superimposed to their respective video stream. The depth perception obtained by the AR system must match that of the original system, therefore the superimposed layer needs to be consistent with the one introduced by the original stereo system. To obtain this, the cameras of the virtual set-up have to match the exact set-up of the stereoscopic camera. This ensures that the augmentation is consistent with the input video streams.

The system uses a stereoscopic endoscope which can be manipulated by the console user. It consists of two cameras, each of which are rotated inwards by 3 degrees and has a field of view of 48 degrees. The inter-camera separation is 5 mm for a distance to the point of interest of 75 mm. This set-up is presented in a study by Ellis et al. [3] and is illustrated in Fig. 3. A similar set-up is used for computing the augmented overlays.

2.2 Prototype

A prototype has been developed and is the foundation for evaluation and testing the proposed solution. This section describes the prototype hardware used. An overview of the system is presented in Fig. 2.

2.2.1 *Hardware*

The prototype is implemented on a desktop computer with an AMD Radeon HD5770 graphics card, which supports up to three outputs. This is set up using AMD's EyeFinity software, which can turn three screens into one large display in Windows, as displayed in Fig. 4. This is necessary because Unity, which is used for the virtual environment, only supports one output.

The video signal is converted back and forth between digital and analogue through the system to enable image processing on the computer. Standard off-the-shelf components are utilized for this.

A Sony HMZ-T2 Head Mounted Display (HMD) is used as the 3D video display for the trainer. This HMD consists of two small displays — both with a full HD resolution — alike those of the da Vinci. The da Vinci console requires two video inputs while the HMD can receive only one input, via HDMI, which is therefore used for both displays. The displaying devices require three outputs, namely two for the da Vinci console and one for the HMD, which will contain both the left and right output streams.

2.2.2 *User Interaction*

Interaction with the system is enabled with Razer Hydra game controllers. The purpose of these is to mimic the controllers of the da Vinci console, such that interaction with the virtual instruments is similar to that with the real instruments. The system consists of two controllers connected to a base that can compute in real-time the exact location and position of the controllers in the hands by means of magnetic motion sensing, with errors down to one millimeter and one degree. The position, orientation and buttons of the Razer Hydra are mapped to the virtual instruments. This mapping was approved by users, as reported by Grande et al. [4]. The buttons are used for moving the arms and to toggle the virtual objects on and off.

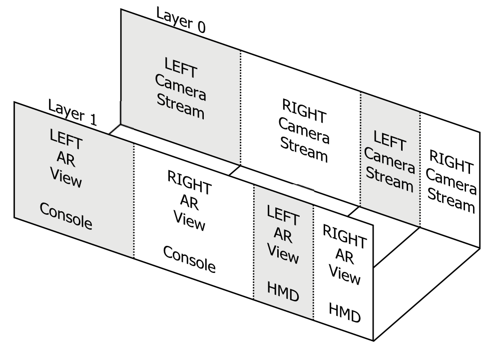

Figure 4: Viewport setup of the system.

2.2.3 *Software Platform*

The virtual overlay is implemented using the Unity game engine which, among other benefits, allows the use of multiple virtual cameras to render the scene. Fig. 4 illustrates eight cameras, configured in two layers: one for the background and one for the overlay. These are output as pairs, consisting of a background camera and an overlay camera. The two layers in each pair are superimposed onto each other and displayed on the four individual viewports. The computer output is split into these viewports, such that the left and right components of the stereoscopic input video are drawn in the corresponding viewports. See Fig. 4.

3. TEST & RESULTS

The presented system is able to acquire two video feeds from the endoscopic cameras, add the virtual overlay and display the output on both the HMD and the surgeon console. Furthermore, the trainer can control the virtual arms using the Razer Hydra controllers, preserving the seven degrees of freedom as in the original da Vinci console. A frame rate of 60 FPS was achieved with the application. It was also found that the proposed system induces a delay of approximately 164 ms.

The system was tested at Aalborg University Hospital by chief surgeon Poulsen, the trainer, during a training session. The test consisted of an informal interview following a tryout of the system. The trainer was introduced to the system and to the controls. By using the virtual arms as visual pointers, the trainer instructed a trainee on how to proceed during an operation. The trainee followed the visual instructions from the trainer while operating on a living pig with the da Vinci robot. The trainer also took over the control of the da Vinci, to try out the system. Both the trainer and the trainee agreed that the delay made the prototype cumbersome to use.

After testing the system as both trainer and trainee, the chief surgeon noted a few drawbacks of the system. The latency of the system made the solution hard to use and there was also a form of salt and pepper noise visible on the left

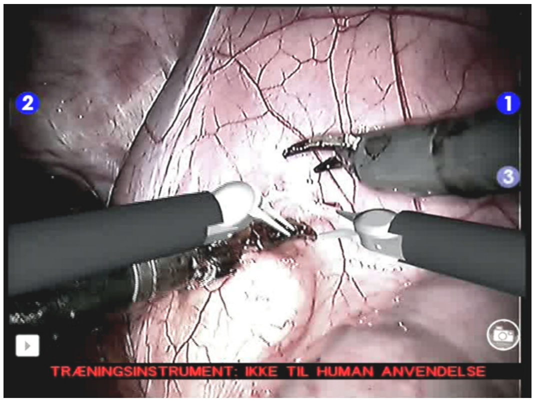

Figure 5: In-system screenshot of proposed system during surgical training. The virtual robotic arms are controlled by the trainer to guide the trainee.

Figure 6: System setup and running during surgical training. The trainer is using the Razer Hydra, and is temporarily using the two external monitors displaying right- and left channel, respectively.

eye on the da Vinci console. The chief surgeon also suggested to make the controls more responsive and the virtual arms more distinguishable from the real arms. The use of a secondary 3D display, such as the Sony HMD, was well received. Furthermore, the concept of instructing using virtual arms were endorsed by the surgeons. A screenshot from the actual training session footage is displayed in Fig. 5.

4. DISCUSSION AND FUTURE WORK

Based on the results of the evaluation, the proposed solution was found to have some shortcomings. The worst of which was the delay, since it causes the system to be cumbersome to use. A solution to this problem could be to implement the system using a compositing device to merge the computer generated overlay and the video streams. Such a solution could be created using a digital video keyer. As mentioned in Section 3, some salt and pepper noise was noted, which has not been seen prior to the user test and might have been caused by the video conversions, unshielded cables, etc.

The test showed, as the assisting nurse pointed out, that it would be beneficial for the assistant to have the depth perception as well, for example by using an HMD. Future work could include a 3D monitor to accommodate the needs of both the trainer and the robot assistant, however such solutions do exists. For improving the system, the chief surgeon brought up several suggestions. In particular, changing the colors of the virtual arms to distinguish them from real ones and maybe adding transparency.

5. CONCLUSION

A system has been presented that uses AR to improve the current training methods with the da Vinci surgical robot. A working prototype has been constructed using low-cost components. The prototype has been tested by skilled surgeons at Aalborg University Hospital and has been found to be a potentially well-functioning solution for training, if the delay would be reduced to an acceptable level.

The AR is done on a desktop computer and the system includes both an analog-to-digital and a digital-to-analog conversion, which introduces delay. The delay limits the effectiveness and usability, but the basic principle of the system has been demonstrated and the purpose has been maintained. The two virtual robotic arms have been found to be a useful tool, as they enable intuitive guidance. The extra 3D display was used by the trainer and this worked well as it provided the surgeon with depth perception. Both the surgeon and the assistant nurse indicated that an extra 3D display was a useful extension for the medical staff, for instance when nurses perform manual suction.

It can be concluded that the use of AR is a useful approach that enables intuitive training and enhances the communication between the trainer and the trainee in this application.

Acknowledgments

The Authors would like to thank chief surgeon Johan Poulsen and robotic surgery assistant Jane Petersson for their help in testing and for their valuable input to the project.

6. REFERENCES

[1] N. S. Asaki Hattori. A robotic surgery system (da vinci) with a data fusion system for navigation surgery in the abdomonal region. *Jikeikai Med*, 50:19–27, 2003.
[2] A. B. D. H. C. Baber and S. Steiner. Virtual reality and augmented reality as a training tool for assembly tasks. *IEEE International Conference on Information Visualization*, 1:32 – 36, 1999.
[3] S. R. Ellis, J. M. Fishman, C. J. Hasser, and J. D. Stern. Effect of reduced stereoscopic camera separation on ring placement with a surgical telerobot. *Surgical Endoscopy*, 22, Issue 11:2396–2400, 2008.
[4] K. Grande, R. Jensen, M. Kibsgaard, and M. Kraus. *Low-Cost Simulation of Robotic Surgery*, pages 6:1–6:4. Association for Computing Machinery, 2013.
[5] K. Lee. Augmented reality in education and training. *InSight: A Journal of Scholarly Teaching*, 7:31–42, 2012.
[6] V. R. Patel. *Robotic Urologic Surgery*. Springer, 2007.
[7] Z. Zabinski. Product quality choice under weak price competition: Hospital adoption of robotic surgery. *http://gradstudents.wcas.northwestern.edu/ zwz708/zabinski-jmp.pdf (last accessed: 13/12/2013)*, 2013.