A Real-Time Vision-Based method to overlay Medical Imaging Data onto a moving patient

Matthew Lanaro, Jonathan Roberts and Peter Corke
Australian Centre for Robotic Vision
School of Electrical Engineering and Computer Science
Queensland University of Technology
2 George Street, Queensland, 4000 Australia
matthew.lanaro@live.com.au

Abstract

This paper will present a novel movement invariant approach for overlaying surgical navigational data on to a live, and moving patient. A pico projector class image projector was used to project an x-ray onto a human forearm, which was sensed using a novel computer vision algorithm. The computer vision algorithm segmented the forearm and found a vector of surrounding points that was then tested for differing forearm thicknesses. Once the forearm was located, an x-ray image was transformed and projected. The projected x-ray was accurate to within 7-10 pixels or 14 to 20mm at a rate of 0.73 seconds per iteration using python and Opencv on an Intel generation 4 i7 laptop. The results could be improved by a projector with a shorter focal length and larger field of view so that the forearm could be much closer to the camera.

1 Introduction

Augmented Reality (AR) has long been regarded as a futuristic technology that could benefit medical and/or surgical applications. The current use of AR for medical or surgical applications involves displaying anatomical models for planning operative procedures. This may also involve tracking and displaying tool orientation in relation to critical structures when conducting procedures to increase awareness of the situation and assist the medical practitioner's decision making. However, there has been little use of AR for surgical navigation applications in clinical settings. The reason for this is related to the high value of space in a surgical environment where many systems have been considered too large, inaccurate or timely to set up [Gavaghan *et al.*, 2012; Sugimoto *et al.*, 2010].

With the miniaturization of technology, many systems that had previously been too bulky could be made smaller and therefore be integrated into a clinical medical setting without compromising or impeding on space. For example, smaller cameras and projectors have been used successfully in a number of procedures. A study by [Gavaghan *et al.*, 2012] presented the use of a small pico projector to overlay 2D data on a patient with high accuracy and required minimal space. Another study by [Fuchs *et al.*, 2009] explored the use of miniaturised cameras to sense small ceramic plates for miniaturised optical tracking and calibration. More recent trials are finding AR systems can be accurate, safe and feasible systems [Onda *et al.*, 2014; Fuchs *et al.*, 2009].

All systems that display surgical navigation data use some way of tracking the movements of surgical tools and/or target areas. Most, if not all surgical/intervention navigation systems use geographic shapes for visual processing. Typically a number of spheres are fixed to a tool so that a vision system can determine the orientation and movements of relevant objects. While these systems are typically accurate, problems often arise from the limited line of sight to the targets and the method of attachment of these visual fiducials can be troublesome.

In this project we investigate the use of small, low cost, readily available pico projectors and develop a computer vision algorithm that does not utilise any markers or additional identifying features. A goal for this project was to develop a system that was invariant to patient movement and maintained an aligned x-ray image projected onto a patient's limb.

The remaining sections are organized as follows. Section 2 describes methods for surgical navigation and highlight recent work using 2D projection techniques as well as methods for tracking. Section 3 outlines how our system and computer vision algorithm were designed. Section 4 presents our experimental setup and Section 5 present results from our preliminary experiments. Finally, Section 6 outlines our conclusions and future work direction.

2 Related Work

The purpose of incorporating AR into surgical navigation is to support and assist in the safe and efficient work of medical professionals. With the introduction of 3D imaging technologies, patient-specific anatomical models can be constructed and displayed next to a surgeon on a 2D screen. However, constantly diverting attention between screen and patient has been noted to consume an unacceptable amount of time [Hansen *et al.*, 2010; Sugimoto *et al.*, 2010]. This excess time reduces patient attention inducing errors that arise due to a difference between the 3D modelling in the displays and the real worksite calling for different AR approaches to be used [Hansen *et al.*, 2010].

In the following Sections (2.1 and 2.2) we briefly review current AR techniques for surgical navigation and part of the projector system design.

2.1 Techniques for surgical Navigation

A number of AR systems have been studied such as semi-transparent displays [Blackwell *et al.*, 1998; Nikou *et al.*, 2000], 2D Projection [Gavaghan *et al.*, 2012; Kobler *et al.*, 2012] and head mounted displays [Fuchs *et al.*, 2009]. [Gavaghan *et al.*, 2012; Sugimoto *et al.*, 2010] conclude that obtrusive equipment reduced surgical vision, long set-up times and difficult calibration have prevented the widespread acceptance of an AR approach as an alternative to monitor displays.

Image overlay is beneficial as it does not require obtrusive head-mounted equipment, therefore taking up minimal space in the immediate area. However, this approach can introduce a significant parallax error.

A study by [Kobler *et al.*, 2012] investigated the feasibility of mounting a laser pico projector onto a surgical drill. The system featured tracking targets for calibration. Evidence revealed that the combination of the surgical instrument and image projector negated the need to consistently reorientate themselves from the patient to a nearby screen. The setup was successful and the researchers concluded that "the technological solution can be easily transferred to other surgical instruments that require intra operative guidance" [Kobler *et al.*, 2012]

2.2 2D Projector Overlay system design

The main area of research in surgical/intervention navigation however is within the field of surgical tool and target area tracking, which has become a state of the art method for a number of interventions [Zakani *et al.*, 2012; Kobler *et al.*, 2012]. Typically an optical system is used to sense the surgical tools and target site such as those used with the pico projector systems described in [Kobler *et al.*, 2012] and [Gavaghan *et al.*, 2012]. However, some researchers observed that these systems have the disadvantage of consuming invaluable space on or around the operating site and evaluated using tiny ceramic plates with tiny cameras [Coigny *et al.*, 2012]. This drive to miniaturize equipment is also used by [Gavaghan *et al.*, 2012] and [Kobler *et al.*, 2012] with a pico projector. Another widely hyped technological advancement has been the use of Electromagnetic (EM) tracking [Lugez *et al.*, 2013; Nijkamp *et al.*, 2014]. The advantage of which is that the targets do not need line of site and they can be much smaller, these developments look promising but suffer from distortions caused by CT machines [Lugez *et al.*, 2013].

Our work will consider a forearm and attempt to display an x-ray onto that forearm. We will evaluate and explore the use of a miniature projector. The pico projector we use is low cost, small and readily available from ebay (Mini LED Projector, Model no: RD-802). We will also design a computer vision algorithm that will not require the use of any visual fiducials.

3 Approach and Methodology

In this section we will describe our approach for developing an AR system for use in medical image projection onto a live moving limb. Firstly we will discuss our general AR system and how it functions. Secondly our computer vision algorithm will be explained.

The AR system used in this project features a camera, projector and a workspace. The system senses the forearm with the camera to find a number of predetermined locations to conduct a homographic transform and project an x-ray of a forearm.

To conduct the homographic transform and project an x-ray, four points will need to be known in the camera coordinate frame and the projector coordinate frame. These can be found in a calibration stage where the projector will project four points in secession and a user will be instructed to click on these points through the camera feed. In Equation 1, the four project points are labelled as xp1, yp1-xp4,yp4 while the camera points are xc1,yc1-xc4,yc4. The values of h00-h21 can then be found.

$$\begin{bmatrix} x_0 & y_0 & 1 & 0 & 0 & 0 & -x_0 \cdot x'_0 & -y_0 \cdot x'_0 \\ 0 & 0 & 0 & x_0 & y_0 & 1 & -x_0 \cdot y'_0 & -y_0 \cdot y'_0 \\ x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1 \cdot x'_1 & -y_1 \cdot x'_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1 \cdot y'_1 & -y_1 \cdot y'_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{N-1} & y_{N-1} & 1 & 0 & 0 & 0 & -x_{N-1} \cdot x'_{N-1} & -y_{N-1} \cdot x'_{N-1} \\ 0 & 0 & 0 & x_{N-1} & y_{N-1} & 1 & -x_{N-1} \cdot y'_{N-1} & -y_{N-1} \cdot y'_{N-1} \end{bmatrix} \begin{bmatrix} h_{00} \\ h_{01} \\ h_{02} \\ h_{10} \\ h_{11} \\ h_{12} \\ h_{20} \\ h_{21} \end{bmatrix} = \begin{bmatrix} x'_0 \\ y'_0 \\ x'_1 \\ y'_1 \\ \vdots \\ x'_{N-1} \\ y'_{N-1} \end{bmatrix} \quad (1)$$

Once the h values have been found we can find the equivalent x,y points in the projectors coordinate frame

with the following Equations 2 and 3.

$$\frac{h_{oo} * x_i + h_{01} * y_i + h_{02}}{h_{20} * x_i + h_{21} * y_i + 1} = x_i' \tag{2}$$

$$\frac{h_{10} * x_i + h_{11} * y_i + h_{12}}{h_{20} * x_i + h_{21} * y_i + 1} = y_i' \tag{3}$$

Figure 1 shows how this process looks when applied to our x-ray.

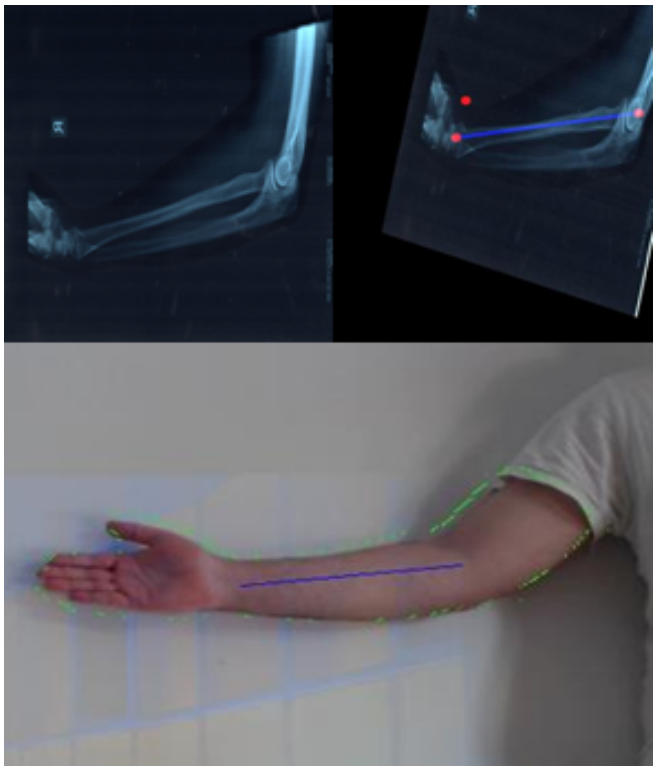

Figure 1: Beginning with the X-ray in the top right the arm segment is found and shown as a blue line in the bottom photo. The X-ray is then homographicly transformed into the photo seen in the top left. The three red dots that are shown in the top right image indicate the location points that were revealed when the arm segment was found, thus allowing for the transform. It is evident that the X-ray in the top right image is aligned to the blue arm segment in the bottom image

3.1 Computer Vision Algorithm

The images for our vision system were collected from a webcam. These were processed to find the arm segment, defined as a line starting in the middle of the wrist to the middle of the elbow. To configure the homographic transform, three points are required and the third point is defined as the top of the wrist.

One option for finding the arm that is widely used with computer vision is feature detection. Depending on the genetic makeup of an arm, it does not contain much texture resulting in a dozen or so features to detect, most of which occur around the hand. If the hand were to be clasped into a fist or the fingers were to be curled slightly many of these features would be altered.

Therefore a simpler algorithm was devised, since the arm will be in front of a plain wall a simple HSV segment will obtain the arm. Of which a Gaussian filter is applied and a canny filter used to detect the edges. It was assumed that the largest edge is the edge that surrounds the entire arm and this edge is saved as a vector of 2D x-y points.

To find these joints the vector was searched for the Euclidean distance of the wrist and the elbow. A number of potential lines were found and have been superimposed in Figures 2 and 3 as green lines. The average was found and is displayed as a red line. The wrist section is a very robust joint to find since it is the thinnest section in the arm, however the elbow is harder to find since there are many possible points that match the equal distance assumption. If we make a new assumption, that the centre of the elbow joint lies across the centre of the arm, then if the length of the arm is known we have then found the arm segment Figures 4 and 5.

Note that this algorithm is valid if the arm is held at a fixed distance perpendicular to the camera, however there is a slight tolerance when finding the Euclidean distance which means that the arm can move freely parallel to the camera.

Figure 2: Green lines show possible distances of wrist measured by Euclidean distance between vector points, red line shows averages of all points from green lines and represents best estimate

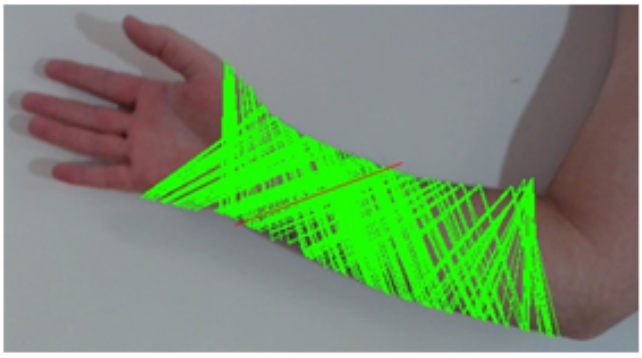

Figure 3: Green lines show possible distances of elbow measured by Euclidean distance between vector points, red line shows averages of all points from green lines and represents best estimate

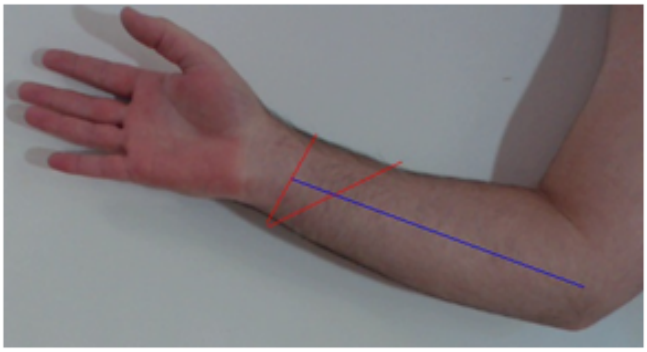

Figure 4: The two red lines are found in fig 2 and fig 3, the blue line can then be found which represents the arm segment. The arm segment starts at the wrist joint and exists through the centre of the found elbow joint

4 Experimental Setup

This section will outline the experimental setup and design. A dataset of 100 images were taken of an arm in front of a plain white wall. In the series the arm appeared in various positions and natural light was used so slightly differing levels of light are present. The series was fed through our algorithm to test for the success rate, iteration time and accuracy.

The algorithm was run with the most up to date version of python and Opencv while the operating system was an Intel generation 4, i7 laptop with 8gb of RAM.

Success rate and iteration time measurements were easily obtainable while running the algorithm Table 1. However, to find the accuracy of the algorithm required some post processing. The series was iterated through and the arm segment was superimposed as a blue line onto the image and saved. The images were then iterated through again with the real and detected positions of the wrist and elbow recorded. Figures 6 and 7 show

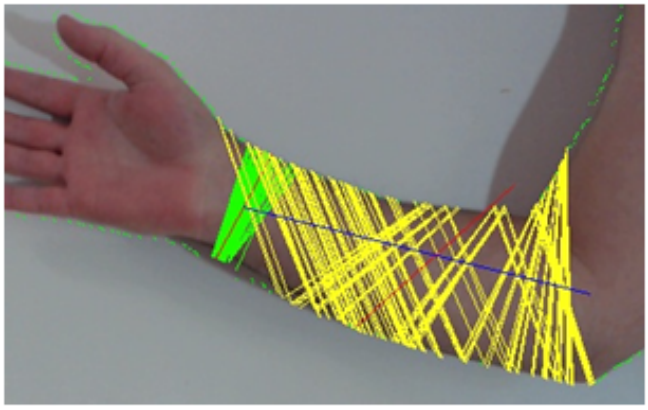

Figure 5: Green dots show outlining vector of points, green line shows possible wist position, yellow lines show possible elbow location, red lines show wrist and elbow average location and the blue line shows the arm segment represented as a line between two points

the error in pixels where 1 pixel equals approximately 2mm. Table 2 also shows some global information from these figures such as mean and variance.

Sucsess Rate	69	percent
Iteration Time	.73	seconds

Table 1: Success Rate and Iteration time

Joint	Mean pixels	Variance
Wrist	10.05	120.84
Elbow	7.5	114.38

Table 2: Mean and Variance of error between actual and detected wrist and elbow positions

5 Results

For the 100 images tested only 69% were successfully processed by our algorithm. However this statistic may be misleading since a failure could occur at any stage of the process such as a bad image segment or a false line being detected. A rate of iteration of 0.728 seconds is acceptable in a real-time system such as this. The accuracy of the algorithm is a little harder to evaluate, both the wrist and elbow joints were accurate to within 7-10 pixels or approximately 14-20mm. As can been seen in Figure 7 however there are a number of large errors in the first 10 images which distorts the mean and variance values for the elbow joint. The mean error for the wrist location was 10 pixels where the elbow location was an average of 7 pixels off however considering the large discrepancy in the first 10 images of the elbow joint it would

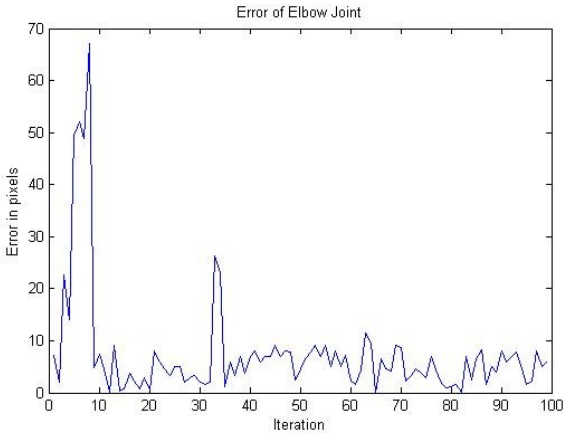

Figure 6: Error between the actually wrist and detected wrist joint.

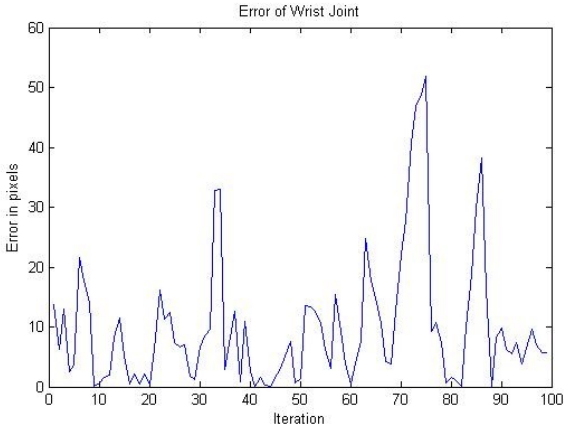

Figure 7: Error between the actually elbow and detected elbow joint.

be expected that the detected position of the elbow is actually far more accurate than the 7 pixels measured.

Since the wrist and elbow positions are mutually dependent it does not make sense that one can be more accurate than the other. The reason that the elbow was more accurate than the wrist was due to the algorithm having failed. If the algorithm failed it returned the previously found locations of the elbow and wrist. This led to the elbow being more accurate since if the arm was to move from the last found position it would do so about the elbow. If we consider this detail and remove the photos for which the algorithm failed then both the wrist and elbow joint would have an extremely similar accuracy with less than 7 pixels error.

In Figure 6 there are a number of spikes which correspond to the locations when the algorithm failed. For example around the 70 and 85 images the error gets relatively high at 53 and 39 pixels respectively, this is due to a number of consecutive algorithm fails.

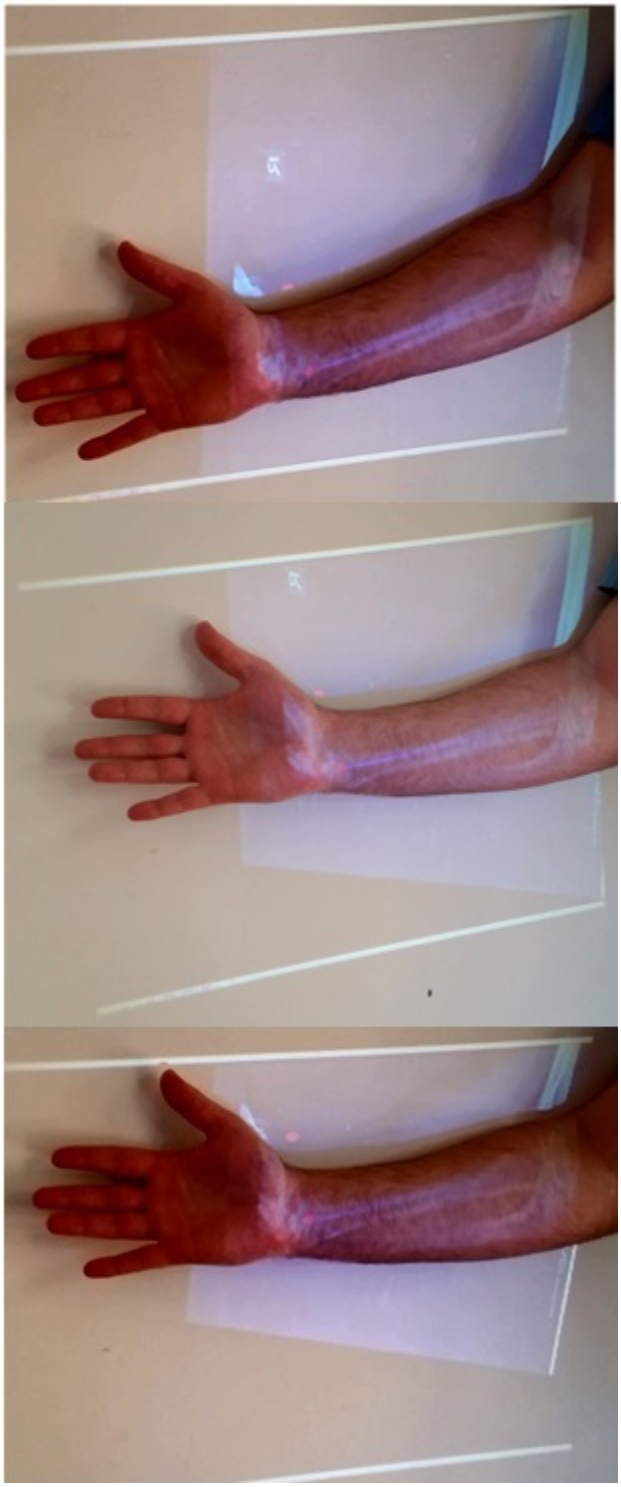

Figure 8: Example images of how the system performed, contrast increased by 10 percent.

The projected x-ray was shown in Figure 8 which displays two images taken from a recording of the system functioning. The x-ray is visible along with the blue line representing the arm segment and 3 red dots to representing the centre wrist, centre elbow and top of wrist. A grey square can also be seen which shows the original x-ray being manipulated to align with the arm. The output lumens of the projector are low and leave the final result less clear than it could be. In the two figures it can be seen that the arm is held in different positions, this displays the real-time capibility of the system which occurs with no extra calibration.

6 Conclusion and Future Work

Throughout this paper we have presented a novel system to display surgical navigation data on to the moving forearm of a live patient. We used a pico projector to create the smallest possible system size in an attempt to gain acceptance from medical professionals with already cramped work spaces. The computer vision algorithm developed was fast, relatively reliable and allowed for the movement of the forearm in the scene.

Future work will explore the use of a projector system with a larger field of projection (a shorter focal length) which would allow the forearm to be closer the the projector and camera. The system as it stands also has the problem that the environment should be dark to allow the projector image to be clear seen, but needs to be bright enough so that the camera can track the forearm. A potential solution to these opposing requirements is to use an near-IR illuminator and near-IR capable camera, that would enable the room to remain relatively dark in the visible spectrum.

References

[Blackwell *et al.*, 1998] M. Blackwell, F. Morgan, A M 3rd DiGioia *Augmented reality and its future in orthopaedics.* Clinical orthopaedics and related research, 1998, 354, 111-122.

[Coigny *et al.*, 2012] F. Coigny, G. Imboden, S. Hemm, J. Beinemann, P. Jurgens, B. Knobel and E. Schkommodau. *Miniaturized surgical navigation.* International Journal of Computer Assisted Radiology and Surgery, 2012, Volume 7, Issue S1, 155-160.

[Fuchs *et al.*, 2009] H. Fuchs, A. State, H. Yang, T. Peck, S. Woo Lee, M. Rosenthal, A. Bulysheva and C. Burke. *Optimizing a Head-Tracked Stereo Display System to Guide Hepatic Tumor Ablation.* Studies in Health Technology and Informatics, Volume 132: Medicine Meets Virtual Reality 16, 126 - 131.

[Gavaghan *et al.*, 2012] K. Gavaghan, T. Oilveira-Santos, M. Peterhands, M. Reyes, H. Kim, S. Anderegg and S. Weber. *Evaluation of a portable image overlay projector for the visualisation of surgical navigation data: Phantom studies.* International Journal of Computer Assisted Radiology and Surgery, 2012, 7(4), 547-556

[Hansen *et al.*, 2010] C. Hansen, J. Wieferich, F. Ritter, C. Rieder, HO. Peitgen *Illustrative visualization of 3D planning models for augmented reality in liver surgery.* International Journal of Computer Assisted Radiology and Surgery, 5(2), 133-41

[Kobler *et al.*, 2012] J.-P. Kobler, M. Kuhlemann and T. Ortmaier. *A laser projector augmented drill for surgical applications: design and feasibility study.* International Journal of Computer Assisted Radiology and Surgery, 2012, 7, S155-S160

[Lugez *et al.*, 2013] E. Lugez, D. Pichora, S. Akl and R. Ellis. *Accuracy of electromagnetic tracking in an operating-room setting.* International Journal of Computer Assisted Radiology and Surgery, 2013, 8, S145-S154

[Nijkamp *et al.*, 2014] J. Nijkamp, S. Schmitz, S. D. Jonge, K. Kuhlmann, J.-J. Sonke and T. Ruers. *Feasibility of surgical navigation for rectal cancer using EM tracking with a tabletop feild generator.* International Journal of Computer Assisted Radiology and Surgery, 2014, 9, S103-S115

[Nikou *et al.*, 2000] C. Nikou, A. M. Digioia, M. Blackwell, B. Jaramaz and T. Kanade. *Augmented reality imaging technology for orthopaedic surgery.* Operative Techniques in Orthopaedics, Volume 10, Issue 1, 82-86

[Onda *et al.*, 2014] S. Onda, T. Okamoto, J. Yasuda l, F. Suzuki, M. Matsumoto, N. Funamizu, R. Ito, Y. Futagawa, S. Fujioka, N. Suzuki, A. Hattori, K. Yanaga. *Augmented-reality based navigation surgery for hepatobiliary diseases.* International Journal of Computer Assisted Radiology and Surgery, 9, S103-S115

[Sugimoto *et al.*, 2010] M. Sugimoto, H. Yasuda, K. Koda, M. Suzuki, M. Yamazaki, T. Tezuka, C. Kosuqi, R. Higuchi, Y. Watayo, Y. Yagawa, S. Uemura, H. Tsuchiya and T. Azuma. *Image overlay navigation by markerless surface registration in gastrointestinal, hepatobiliary and pancreatic surgery.* Journal of HepatoBiliary Pancreatic Sciences, 17(5), 629-36

[Zakani *et al.*, 2012] S. Zakani, G. Venne, E. J. Smith, R. Bicknell and R. E. Ellis. *Analyzing shoulder translation with navigation technology.* International Journal of Computer Assisted Radiology and Surgery, 7, 853-860