An Approach for Projector-based Surgeon-Computer Interaction using Tracked Instruments

Bojan Kocev, Darko Ojdanić, Heinz-Otto Peitgen
Fraunhofer MEVIS
Universitätsallee 29, 28359 Bremen, Germany
{bojan.kocev@mevis.fraunhofer.de}

Abstract: Providing an intuitive and easy to operate interaction mechanism for use in medical augmented reality applications is one of the very crucial aspects for accepting these applications for use in the operating room. Currently, soft tissue (e.g. liver) navigation information is displayed on the monitor, requiring the operating surgeon to change focus from the monitor to the surgical site and vice versa during navigation. Projector-based augmented reality has the potential to solve this problem. The aim of this work was to use projection not only for the visualization, but also for interaction. As a development platform, we used a soft-tissue navigation system which contains an optical (infrared) tracking system. A consumer-grade projector was added to visualize preoperatively defined surgical planning data. Moreover, the projection was used as a virtual touch screen with which the surgeon could not only see information, but also interact with it. We performed the registration of the projection in the camera coordinate system by using either three or four points. The former case assumed a rectangular screen projection, while the latter allowed projecting a random quadrilateral shape. In this fashion, we were able to decide if the surgeon tried to interact with the projected virtual information while tracking the surgical tool. To interact with the virtual information, the surgeon could use a predefined set of surgical tool gestures that emulated a computer mouse ($left\_click\_up$, $left\_click\_down$, $mouse\_drag$). These were successfully recognized by the system and mapped to application-specific commands to modify the displayed virtual data.

Keywords: navigation, tracking surgical tools, projector-based augmented reality, intraoperative visualization, intraoperative interaction, gesture recognition, liver surgery

1 Introduction

Nowadays, computer technology and controlling software have reached high-quality values which never fail to fascinate in many different ways. One of the most prominent fields for which the incorporation of computer support is very beneficial is modern medical technology together with its different applications for medical diagnosis and therapy. There are many different applications that enable surgeons to preoperatively diagnose their subjects, while analysing recorded CT or MRI data, for example. This way, the surgeons, having diagnosed the patient, are able in the next step to plan the complex surgical interventions prior to the actual surgery. The next step is to make the different virtual planning data available to the surgeon during the surgery in an easy and intuitive way with the aim of maximizing surgeon efficiency and success in the surgical intervention.

erschienen im Tagungsband der INFORMATIK 2011
Lecture Notes in Informatics, Band P192
ISBN 978-3-88579-286-4

Last but not least, we need to provide an easy, robust, and fast enough mechanism for real-time interaction with the virtual planning data during the surgery. At the moment, the most widespread realization of providing the planning data during the surgery, together with providing a way for interaction with the same data, is to use a station that includes a sterilized monitor with a touch-screen display. The incorporation of a monitor with a touch-screen display provides the surgeon with all the planning data during the surgery, but with a great disadvantage of having to turn the view from the surgical site and great difficulty in the interaction process. The change of view causes decreased concentration for the surgeon, and the interaction is almost impossible due to having the system usually away from the surgeon and difficulty in ensuring a 100 % sterilized touch-screen display. One common solution for the interaction problem in the previous setting is to have surgical assistant that will follow the surgeon instructions, although this is time consuming and usually ambiguous and difficult for both the surgeon and the assistant. C. Graetzel et al. [GFGB04] observed the following very interesting phenomena of how the surgeon and the assistant usually communicate during the interaction process:

> surgeon: Move the mouse to the third button down
> assistant: This one?
> surgeon: No, the next one down.
> assistant: This one?
> surgeon: No, the other one... Yes, that's it.

An alternative way for providing the surgeon with the virtual planning data during surgery is through augmented reality realized with immersive display devices like see-through head-mounted displays (HMDs), as demonstrated by Bornik et al. [BBR$^+$03]. However, this way the field of view of the surgeon is limited and he or she might experience the surrounding real information differently in a way that decreases the quality of the surgeon performance. One of the most recent ideas, as demonstrated by Hansen et al. [HWR$^+$10] and their corporate partners, is to incorporate a projector for the visualization of all the planning data, for example, at a chosen nearby location on the patient or directly on the patient organ that is being examined or an intervention performed on the organ during the surgery. This way, the disadvantage of having the surgeon change his view from the patient to the monitor is eliminated, and in addition, all team members see the same virtual information, and the field of view is not affected as when using HMDs. However, currently to our knowledge, there is no current projector-based medical system that is capable of providing the same touch-screen functionality. In other words, what would be helpful for the surgeon is to be able to interact with the projected visual information in an intuitive and easy way that will not affect his or her performance in a bad way, but in contrast possibly enhance it significantly. We developed a system that adds exactly this functionality to the projector-based visualization system. Our system provides the surgeon with the option of choosing any sterilized flat region where he or she wants to project the virtual information, and then we register this chosen projection area in the camera coordinate system. This way, we enable the surgeon to interact with the projected information using one of his tracked tools. We use an optical measurement system to track the surgical tools that are used for interaction with the projected virtual information. The system is capable of recognizing a set of different surgical-tool-based gestures that are mapped to different mouse events, thus enabling the surgeon to interact with the virtual information. In addition to this, it is important to mention that the functionality of our virtual touch-screen system is equivalent to any standard touch-screen display, making it very easy and intuitive to use.

2 Related Work

Hoppe et al. [HBD+01] developed a system using projector-based augmented reality for the interoperative visualization of preoperatively defined surgical planning data as well as additional information (numerics, distances, etc.) in arbitrary colors. Their system contains a surface scanner which is used to generate a 3D point cloud of the patient's skin surface by projecting a sequence of stripe patterns (coded light) on top of the region of interest immediately before the intervention starts. These images are acquired by the cameras, analyzed in consideration of shifting gray values, and yield a 3D point cloud of the scanned area. The generated point cloud is matched, using the well known iterative closest point (ICP) algorithm [AHB87, FWMJ98], to the preopertively segmented surface of the diagnostic image data (CT, MRI) on which the surgical plan was defined. Their system is particularly attractive due to the fact that both the video projector and the cameras are used for different purposes (registration/visualization and registration/tracking respectively), enabling the surgeon to visualize planning data on top of any preoperatively segmented and triangulated surface (skin, bone, etc.). However, this system lacks the possibility for the surgeon to easily interact with the projected planning data. We believe that introducing an easy and intuitive interaction mechanism will further enhance the power of their projector-based augmented reality system.

Graetzel et al. [GFGB04] developed a system that uses computer vision to replace standard computer mouse functions with hand gestures, so that surgeons will be able to make more effective use of computers during surgery. They allow the surgeon to choose a 3D interaction zone (workspace), usually situated just above the surgical zone, which typically measures $50\times50\times50$ cm, where he or she afterwards can use bare hands or colored surgical gloves to control the computer. Their set of available hand gestures includes the *("pick-up") the non-contact mouse* hand gesture which is realized by placing the surgeon's hand in the workspace and keeping it stationary for a moment. After the tracking is activated, they map hand motion to pointer movement using non-linear gains. They generate mouse clicks by either moving the cursor to a certain screen location and holding the hand stationary for a short time *("wait to click")*, or by detecting a hand motion (around 20 cm) in the direction of the camera *("push to click")*. However, even if we accept that the system is without false positives, there is still another thing that in our opinion is causing the surgeon to lose concentration while using the system: the surgeon needs to change his view from the operating zone to the monitor (around 90 degrees change in the viewing direction).

Fischer et al. [FNFB04] developed a video see-through system for medical augmented reality, which is based on a VectorVision image-guided surgery (IGS) device including an optical tracking system. They use a standard camera with attached infrared marker instrument clamp and calibrated in a one-time calibration step, so that the generation of augmented video stream is possible. The live video stream from the camera, containing the patient organ, is augmented with some planning data or other medical information in the form of graphical scene elements and supports the surgeon during the surgical intervention. In our opinion, this system lacks the possibility for the surgeon to easily interact with the augmenting graphical elements, in the sense to be able to change the displayed medical data information easily or give any other relevant command input to the system.

In their next paper, Fischer et al. [FBS05] presented an extension of their previous work [FNFB04] which revealed how the capabilities of the so-called image-guided surgery system for tracking surgical instruments can be utilized for user interaction. They process the consecutive pose data of the tracked surgical tools and recognize different surgical tool gestures, which are used as an input to their configurable menu system to trigger different menu actions in immediate proximity to the patient. Their system recognizes a so-called *"still click"* gesture which is realized by keeping the movement of the tool below a given threshold. Their second gesture *"angle click"* requires in addition to keeping the surgical tool still, performing a certain rotating motion. In addition, they use the surgical tool for the definition of points or more complex shapes in 3D. As already mentioned in

their paper, the *"still click"* is easily prone to false positives. The second gesture could be difficult to perform, especially during a stressful surgical intervention. Overall, their medical AR application [FNFB04] supported by the configurable menu system together with the tool gestures input, represents an important step in bringing the medical data interoperatively and letting the surgeon interact with it without greatly affecting his concentration, but rather improving his performance.
Hartung et al. [HGS$^+$09] presented a region-of-interest (ROI) two-step calibration method, enabling a projector-based visualization during computer-assisted coronary artery bypass grafting (CABG) surgery at the open heart. In the first step, they determine the paths of the light rays emitted by the projector, while in the second step, they determine the mapping between the camera and projector for correct projection onto the surface.
What we found particularly interesting and motivating for our work was the Wearable Gestural Interface designed by Pranav at al. [MM09]. They use a tiny projector and a camera coupled in a pendant-like mobile wearable device that sees what the user sees and visually augments surfaces, walls or physical objects the user is interacting with using hand gestures.

3 The Navigation Platform

3.1 System Overview

- Stereo camera (NDI Polaris Vicra)
- Touch screen interface
- Navigated ultrasound (Terason T3000)
- Navigated CUSA
- Interface to MeVis planning data
- Desktop computer
- Integrated and transportable system for flexible use in the OR

Figure 1: The soft-tissue navigation system CASOne from CASCination GmbH, Switzerland. (Image taken from Matthias Peterhans et al. [PDN$^+$]. For more information about the system please refer to their web page [CAS].)

For demonstration of our interaction and visualization system, we use our own GUI, not the standard CASOne GUI. In addition to these components, we have included a consumer-grade projector

placed next to the camera. This way, we are able to project the different medical data at a chosen location.
We have developed a MeVisLab C++ module called MLNDIPolarisServer (based on the Virtual Reality Peripheral Network (VRPN) public-domain software, released by the Department of Computer Science at the University of North Carolina at Chapel Hill) that is responsible for the communication with the NDI Polaris Vicra Camera. This module loads all tracked tools configuration files that define the tracked tools uniquely. Later, we are able to retrieve the pose of any of the tracked tools seen by the camera, and the user can choose the number of updates per second received from the camera.

3.2 One-time CUSA calibration

The NDI Polaris Vicra infrared camera, which is part of the navigation system shown in Figure 1, tracks tools with attached infra-red marker clamps as shown in Figure 5. The position sensor reports the tool origin transformations in the camera coordinate system. The tool origin is part of the tool's local coordinate system. A problem arises in the sense that we are not interested in tracking the CUSA tool origin, but the tip of the CUSA tool. For that reason, we use the developed calibration procedure by CASCination GmbH, Switzerland, that determines the CUSA tip offset, which is the vector between the CUSA tip and origin. They provide us with a calibration tool, and we are able to compute the calibration matrix using the CASCination software and store it in an XML file that is loaded upon starting our presented independent MeVisLab application.

4 Projected Screen Registration

In order to be able to interact with the projected screen information, we need to register the projected screen in the camera coordinate system. For that reason, we first need to record either three or four corner points (depending on whether we assume rectangular projection or a random quadrilateral) of the projected screen quadrilateral, as shown in Figure 2:

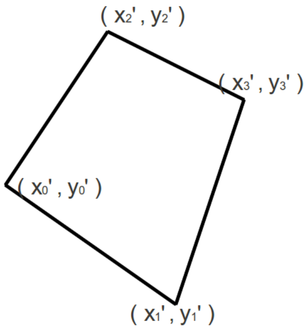

Figure 2: A quadrilateral screen projection.

The point-recording procedure uses a user-defined distance and a variance threshold, so that we are able to check if the distance between the recorded points, and the variances among the points, is

within the defined thresholds and record the points if everything is fulfilled.
Instead of tracking the CUSA tip in the camera coordinate system, we decided to perform the tracking of the CUSA tip positions in our own defined coordinate system, and this way make the tracking more intuitive. Even if we use three or four points for the registration, we always use the first three registration points $(x_0^{'}, y_0^{'})$, $(x_1^{'}, y_1^{'})$, and $(x_2^{'}, y_2^{'})$ to define our own coordinate system that has its x and y axes in the plane where the screen is projected. In addition to this, it is assured that the z-axis is always pointing upwards, which is in accordance with our developed gesture recognition algorithm. We compute a 4×4 transformation *T* in a homogeneous matrix form:

$$T = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}$$

that we use to transform all tracked CUSA tip positions from the camera coordinate system to our own defined coordinate system. In addition, we also transform all recorded projected screen registration points from the camera to our own defined coordinate system using the transformation matrix *T*.

5 Mapping CUSA Tip Positions to Screen Coordinates

The main goal of this work is to use the tracked surgical tool (e.g. CUSA) to control the cursor and generate mouse events at the desired location of the projected screen. Having transformed the tool tip position in our own defined coordinate system, it is very important to correctly map this position to its corresponding position coordinates in screen space. In the case that we assume a rectangular screen projection, it is fairly simple to check if we are inside the projection area, scale the transformed CUSA tip positions according to the screen resolution, and least discretize it.
In the more general case when we assume a random convex quadrialteral projection, we first perform an *inside-convex-quad* test to decide if the transformed CUSA tip position is inside the convex quadrilateral projection area. If we are inside the projection area, we map the transformed CUSA tip position to its corresponding screen coordinates using *Quadrilateral-to-Rectangle Mapping* M^{-1} [KJH02, Hec89]. The *Quadrilateral-to-Rectangle Mapping* M^{-1} is computed using the *Rectangle-to-Quadrilateral Mapping M* as in Figure 3,

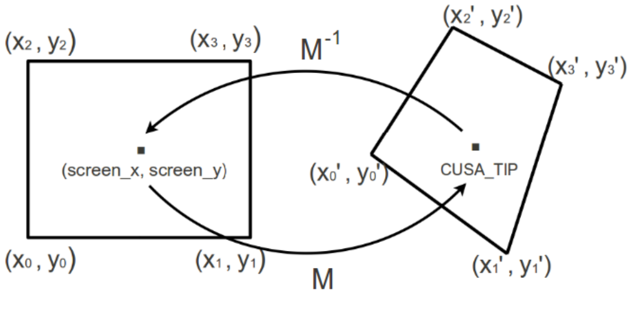

Figure 3: Mapping the transformed $CUSA\_TIP$ to its corresponding cursor $(screen\_x, screen\_y)$ position.

which on the other hand uses the *Unit Square-to-Quadrilateral Mapping A* shown in Figure 4.

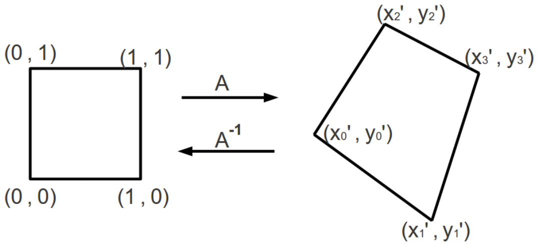

Figure 4: Unit Square-to-Quadrilateral Mapping.

Note that the width and the height of the rectangle to which we map the registered quadrilateral, are defined so that they correspond to the screen resolution with which we are working.

6 Surgical Tool Gestures Recognition and Mapping to Mouse Events

In the previous sections, we described all the necessary tools and methods that are being used to support our tracking and gesture recognition algorithm.

Algorithm 1 Tool Gestures Recognition and Mapping to Mouse Events

```
if $CUSA\_TIP$ is on the projected screen plane then
    if $CUSA\_TIP$ is inside the projected screen quad then
        $(screen\_x, screen\_y) \leftarrow quad\_to\_rect(CUSA\_TIP\_X,\ CUSA\_TIP\_Y)$
        if $not\ LEFT\_MOUSE\_BUTTON\_PRESSED$ then
            $leftClickDown\_at\ (screen\_x,\ screen\_y)$
            $leftClickDown\_soundNotifcation\ ()$
            $LEFT\_MOUSE\_BUTTON\_PRESSED \leftarrow true$
        else
            $setCursorPos\ (screen\_x,\ screen\_y)$
        end if
    end if
else
    if $LEFT\_MOUSE\_BUTTON\_PRESSED$ then
        $leftClickUp\_at\ (screen\_x,\ screen\_y)$
        $leftClickUp\_soundNotifcation\ ()$
        $LEFT\_MOUSE\_BUTTON\_PRESSED \leftarrow false$
    end if
end if
```

The set of recognized surgical tool gestures includes:

- Left Click Down
- Left Click Up
- Mouse Drag (Positioning the Cursor)

We provide in Algorithm 1 the pseudo-code short version of our surgical tool tracking and gesture recognition method. Note that we have incorporated two different sound notifications for which one mimics the same sound as a real mouse button being pressed down, and the other mimics the sound of releasing the mouse button. Black et al. [BL10] in their work, for example, integrated concepts from the field of auditory display to enhance an electronic surgical navigation assistant for image-guided liver surgery, and found that this reduces the overburdened dependence on visual input during surgery and allow the surgeons to keep their eyes in the situs. We use the sound notification mechanism with the hope of making the system more intuitive and more realistic. This is usually not the case with the real touch screens but rather with the standard mouse, and we can think of this as an optional enhancement of the system that can be easily turned off upon user request.

7 Results and Discussion

A glimpse at Figure 5 reveals our new interaction mechanism in action, targeted for any medical application developed for use on a navigation system that includes a tool-tracking system. The same virtual information that one would have to look at on the navigation system standard monitor is now projected at a location chosen by the user who, in our case, would ideally be the surgeon. The surgeon uses his tracked instrument (e.g. the CUSA tool) to interact with the projected medical information. This way, all medical preoperative planning data or any other relevant medical information is provided to the surgeon and his team interoparatively in an easy and intuitive way in the proximity to the patient.

The developed new interaction mechanism has not been tested in the operating room, but the initial tests in our lab at Fraunhofer MEVIS in Bremen, showed very good results. We observed that any user who tried to use the system did not not need much time to get used to the interaction mechanism, but rather found it very familiar to what they were used to in the case of standard touch screens, especially with those that include a pen-like tool for interaction. Due to the fact that our system relies on a very accurate optical tracking system and the well-defined gesture-detection algorithm, we are happy to claim that our system is free from false positives. In addition to this, because we perform optical tracking, and almost all of our computations are matrix-based, we assure that the system runs in real-time without many visual mismatches. Some visual mismatches between the displayed cursor position and the actual tool tip position might occur if the tool is moved very quickly, because the camera will not be able to capture all different positions that quickly or due to the delay caused by the processing of the tracking data until the actual update of the cursor position is performed. Another factor that we need to take into account is the fact that our tracking system is sensitive to occlusion of the tool marker clamps from the camera viewpoint, but our gesture recognition system is capable to easily recover after the occlusion is eliminated without any unusual behaviour, which is very important for fulfilling the safety requirements during any surgical intervention. Another key factor that still requires further investigation, are the different lights in the operating room, which even though are controlled, still might disturb the quality of the visualization.

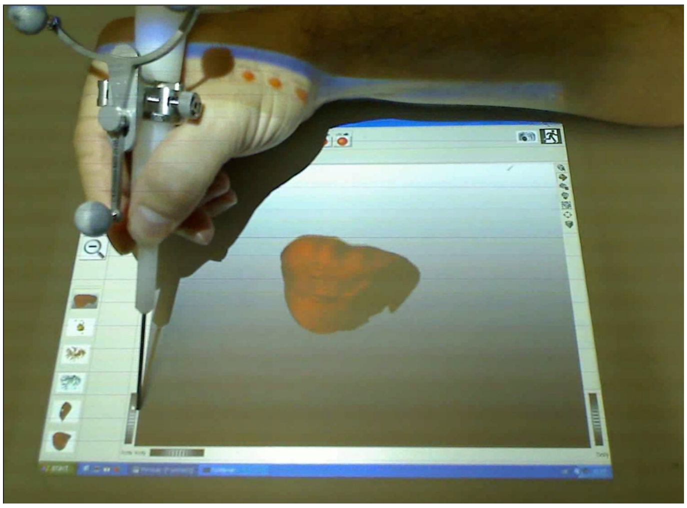

Figure 5: A projected screen, and the surgical tool CUSA interacting with the projected virtual information.

8 Conclusion and Future Work

We have developed a novel wireless interaction mechanism for use in different medical augmented reality applications. The new interaction mechanism eliminates the burden for the surgeon to have to turn his view from the surgical site to the monitor, additionally eliminates unpleasant distractions and enhances the surgical performance. Not only is the surgeon provided with the medical data during the intervention, but he is also given all the power and freedom for interaction with this information with the hope to fulfill all his requirements.

There are still some open questions that we plan to address in the near future. We need to perform more thorough tests, especially in the operating room (OR) to prove the compatibility of our interaction and visualization mechanism with the real OR scenarios and their requirements. We plan to look into the issue of providing even more freedom for the possibility of easily changing the projection area during the surgery. Currently, this is handled by repeating the registration procedure which takes less than a minute. Another solution would be to have an extra web camera that will capture the projection area and will be calibrated with the rest of the system, so that after processing the images and segmenting the projection area, will be possible to recover the new projection area position without needing to repeat the registration procedure. Indeed, this still requires some more research, and it is not clear how fast and robust it would be. Last but not least, we are interested in extending the system in such a way that the projection and the same interaction mechanism is possible on a curved surface, e.g. registered to the target organ, which may offer additional flexibility for the surgeon.

Because our system is based on existing medical equipment, it is assured that it should have an easy transition into the clinical practice once all the necessary testing is completed.

References

[AHB87] K.S. Arun, T.S. Huang, and S.D. Blostein. Least-squares fitting of two 3-D point sets. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (5):698–700, 1987.

[BBR+03] A. Bornik, R. Beichel, B. Reitinger, C. Gotschuli, E. Sorantin, F. Leberl, and M. Sonka. Computer aided liver surgery planning based on augmented reality techniques. In *Bildverarbeitung für die Medizin: Algorithmen, Systeme, Anwendungen. Proceedings des Workshops vom 9.-11. März 2003 in Erlangen*, page 249. Springer, 2003.

[BL10] Hansen C. Black, D. and J. Lovisach. Auditory Support for Image-Guided Liver Surgery. *International journal of computer assisted radiology and surgery*, 5(1):187–188, 2010.

[CAS] Computer Assisted Soft Tissue Surgery. `http://www.cascination.ch`.

[FBS05] J. Fischer, D. Bartz, and W. Straßer. Intuitive and lightweight user interaction for medical augmented reality. *Vision, Modeling, and Visualization. Erlangen*, pages 375–382, 2005.

[FNFB04] J. Fischer, M. Neff, D. Freudenstein, and D. Bartz. Medical augmented reality based on commercial image guided surgery. In *Eurographics Symposium on Virtual Environments (EGVE)*, pages 83–86. Citeseer, 2004.

[FWMJ98] J.M. Fitzpatrick, J.B. West, and C.R. Maurer Jr. Predicting error in rigid-body point-based registration. *Medical Imaging, IEEE Transactions on*, 17(5):694–702, 1998.

[GFGB04] C. Graetzel, T. Fong, S. Grange, and C. Baur. A non-contact mouse for surgeon-computer interaction. *Technology and Health Care*, 12(3):245–257, 2004.

[HBD+01] H. Hoppe, J. Brief, S. Däuber, J. Raczkowsky, S. Haßfeld, and H. Wörn. Projector Based Intraoperative Visualization of Surgical Planning Data. *Proceedings of ISRA-CAS*, 2001.

[Hec89] P. Heckbert. Projective Mappings for Image Warping. In *in Fundamentals of Texture Mapping and Image Warping (Paul Heckbert, Masterâs Thesis), UC Berkeley*. Citeseer, 1989.

[HGS+09] C. Hartung, C. Gnahm, S. Sailer, M. Schenderlein, R. Friedl, M. Hoffmann, and K. Dietmayer. Towards Projector-Based Visualization for Computer-Assisted CABG at the Open Heart. *Bildverarbeitung für die Medizin 2009*, pages 376–380, 2009.

[HWR+10] C. Hansen, J. Wieferich, F. Ritter, C. Rieder, and H.O. Peitgen. Illustrative visualization of 3D planning models for augmented reality in liver surgery. *International journal of computer assisted radiology and surgery*, 5(2):133–141, 2010.

[KJH02] D.K. Kim, B.T. Jang, and C.J. Hwang. A planar perspective image matching using point correspondences and rectangle-to-quadrilateral mapping. In *Proceedings of the Fifth IEEE Southwest Symposium on Image Analysis and Interpretation*, page 87. IEEE Computer Society, 2002.

[MM09] P. Mistry and P. Maes. SixthSense: a wearable gestural interface. In *ACM SIGGRAPH ASIA 2009 Sketches*, pages 1–1. ACM, 2009.

[PDN+] Matthias Peterhans, Benot Dagon, Lutz-Peter Nolte, Charles Baur, Anne Vom Berg, Daniel Interbitzin, and Stefan Weber. Soft tissue navigation: Transferring pre-operative planning data into the operation room. In *European Congress of Radiology*.

erschienen im Tagungsband der INFORMATIK 2011
Lecture Notes in Informatics, Band P192
ISBN 978-3-88579-286-4