

Virtual Window for Improved Depth Perception in Medical AR

Christoph Bichlmeier*

Nassir Navab†

Chair for Computer Aided Medical Procedures (CAMP), TU Munich, Germany



ABSTRACT

Real-time in-situ visualization has been a subject of intensive research and development during the last decade [2], [6], [8]. Besides accuracy and speed of the systems one of the challenges to improve acceptance of medical AR is to overcome the misleading depth perception caused by superimposed virtual entities of the AR scene onto real imagery, e.g. virtual tissue and bones occlude real skin. Occlusion is the most effective depth cue [3] and let e.g. the visualized spinal column appear in front of the real skin. We present a technique to tackle this problem. A virtual window overlaid onto the real skin of the patient creates the feeling of getting a view on the inside of the patient. This view is restricted by the frame of the window, however, due to motion of the observer the frame covers and uncovers fragments of the visualized bones and tissue and enables the depth cues motion parallax and occlusion, which rectify the perceptive misinformation. An earlier experiment has shown the perceptive advantage of the window. Therefore seven different visualization modes for the spinal column were evaluated regarding depth perception. This paper describes the technical realization of the window.

Keywords: augmented reality, depth perception, proprioception, HMD

1 INTRODUCTION

Real-time in-situ visualization of medical data is getting increasing attention. Watching a stack of radiography is time and space consuming within the firm work flow in an OR. Physicians have to associate the imagery of anatomical regions with their proper position on the patient. Medical augmented reality allows for the examination of medical imagery like radiography right on the patient. Three dimensional visualizations can be observed by moving with a head mounted display HMD around the AR scene. Several systems [9, 2, 5] that are custom made for medical procedures tend to meet the requirements for accuracy and to integrate their display devices seamlessly into the operational work flow. Also depth perception has become a major issue of current research in medical AR. Virtual data is superimposed on real imagery and visual depth perception is disturbed as shown in figure 1. This problem

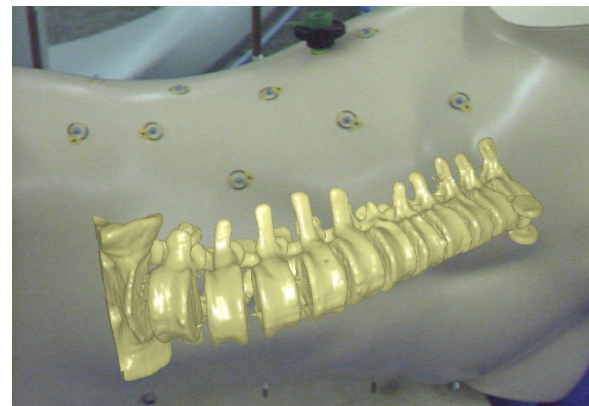


Figure 1: Opaque surface model occludes real thorax. Therefore it is perceived in front of the body although the vertebrae is positioned correctly.

has been identified as early as 14 years ago in the first publication about medical augmented reality [1]. This group tasked the problem by rendering a "synthetic hole" ... "around ultrasound images in an attempt to avoid conflicting visual cues." In an earlier paper our group described an experiment that evaluated seven different visualization modes for the spinal column regarding depth perception.

This paper describes the technical realization of one of the winners of the evaluation. This is a virtual window that can be overlaid onto the skin and provides a bordered view onto the spinal column inside the patient.

Section 2 presents our AR system. Section 3 describes the setup and design of the virtual window. In section 4, we analyze the perceptive advantage of the window.

2 HARDWARE SETUP

This section describes our AR system that consists of an optical outside-in tracking system for target tracking and an inside-out tracking system for head pose estimation.

2.1 AR System

First of all, we like to introduce our AR system that allows for in-situ visualization. Figure 3 gives a complete overview about an AR system in surgical use. For superior registration quality the system

*e-mail: bichlmei@in.tum.de

†e-mail: navab@cs.tum.edu

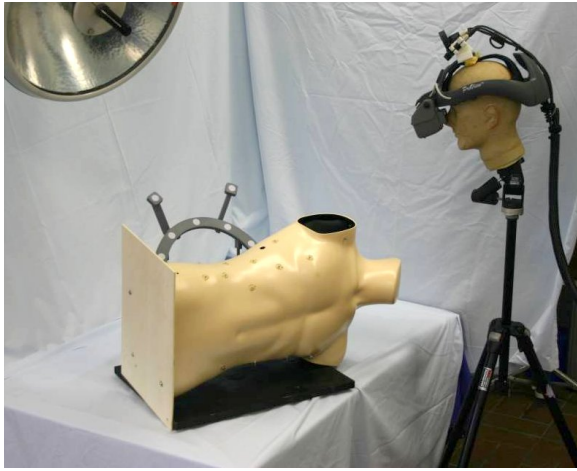


Figure 2: Plastic thorax model, HMD and a marker frame for inside-out tracking are some of the components of our experimental AR application.

uses two synchronized tracking systems.

The single camera inside-out tracking system allows for a high rotational precision [4], which is necessary for tracking the stereoscopic video-see-through head mounted display (HMD). The hardware setup is similar to the one proposed by Sauer et al. [8] for medical augmented reality. Two color cameras rigidly attached to the HMD simulate the eye's view. An additional infrared camera mounted on the HMD tracks a marker frame (figure 2) for head pose estimation [11]. There are two major reasons why to choose a video see-through system. Real and virtual imagery can be optimally synchronized to avoid time lags between the images of the camera, which would lead to undesirable and for the user fatiguing effects like "perceivable jitter or swimming" [10]. Second the system allows for more options how to combine real and virtual imagery like occluding real objects since we have full control over the real images while optical systems offer only a brightening augmentation. The optical outside-in tracking system from A.R.T GmbH (Weilheim, Germany) with four cameras fixed to the ceiling covers a large working area, i.e. 3x3x2 m. The system is capable of tracking the targets in our setup with an accuracy of $< 0.35[mm]$ RMS. Both of the systems use the same kind of retroreflective fiducial markers offering a registration free transformation from one tracking system to the other. In order to recover the six degrees of freedom of a rigid body, the external optical tracking system requires at least four rigidly attached markers. Fiducial markers are attached to the patient lying on the operating table (see figure 2) and further surgical instruments. The marker frame target has an exceptional function as it enables the transition between the inside-out and the outside-in tracking systems. Both tracking systems calculate the same coordinate system respective the reference target. All augmentations of targets, which are tracked by the optical outside-in tracking system, have to be positioned respectively the marker frame of the inside-out tracking system. The following equation calculates the transformation $^{anytarget}H_{frame}$ from the marker frame to an exemplary target (${}^{to}H_{from}$).

$$^{anytarget}H_{frame} = ^{anytarget}H_{ext} * (^{frame}H_{ext})^{-1} \quad (1)$$

$^{anytarget}H_{ext}$ and $^{frame}H_{ext}$ are the transformations provided by the optical outside-in tracking system. The former describes the transformation respective the origin of the tracking system to a target, the latter is the transformation from the origin of the tracking system to the marker frame for inside-out tracking.

A PC based computer is used to render 3D graphics, to compute and include tracking data and to synchronize and combine imagery of virtual and real entities. The specification is Intel Xeon(TM), CPU 3,20 GHz, 1,80 GB RAM, NVIDIA Quadro FX 3400/4400. The window is implemented in C++ using the OpenGL.

2.2 In-Situ Visualization

Our system allows for different kinds of visualization techniques such as direct and indirect (isosurfacing) volume rendering (see figures 5). In-situ visualization requires the following preparations.

1. At least four fiducial markers have to be attached to the object of interest, e.g. thorax. These markers have to be visible for the tracking cameras in the OR.
2. The object of interest, e.g. part of the thorax, has to be scanned by CT or MRI to get a three dimensional data volume.
3. Registration: The fiducial markers are segmented automatically from the data volume to be able to align the virtual data with the real tracked object.
4. Choice of visualization technique.

The technique of direct **volume rendering** is able to display every part of the data volume with a certain value for color and transparency. Therefore a predefined number of planes parallel to the image plane are clipped against the volume boundaries. All planes are rendered by interpolating within the volume and blending appropriately. Intensity values in the volume domain are mapped to the three dimensional color space using transfer functions. This enables accentuation of interesting structures.

Indirect volume rendering concerns the extraction of **surface models** from the data volume. Areas of interest, e.g. bones or blood vessels can be determined due to their intensity values in the volume domain. The marching cube algorithm is parameterized with a certain threshold to segment a homogeneous area within the data volume and generates a surface model. Surface models can be designed with color, transparency and textures.

The presentation of volume rendered objects is more computationally expensive than display of surface models. Our system renders the volume rendered spinal column with 5-6 fps and its surface model with 30 fps.

Positioning the visualization of the spinal column inside the thorax within our AR scenario can be described by the transformation $^{visual}H_{frame}$.

$$^{visual}H_{frame} = ^{visual}H_{thorax} * ^{thorax}H_{ext} * (^{frame}H_{ext})^{-1} \quad (2)$$

$^{thorax}H_{ext}$ and $^{frame}H_{ext}$ are the transformations provided by the optical outside-in tracking system. $^{visual}H_{thorax}$ represents the registration matrix to align virtual data with the real tracked object.

3 VIRTUAL WINDOW

The following section introduces a virtual window, which can be overlaid onto the skin of the patient. The user of the HMD observes the inside of the patient through this vision panel.

3.1 Position the Window

Placing the window to get the desired view into the patient can be performed without touching or moving the patient. While positioning the window, the observer wearing the HMD views a frame as shown in figures 4. The observer guides the frame to the area of interest by moving his or her head as shown in figures 4. If the frame

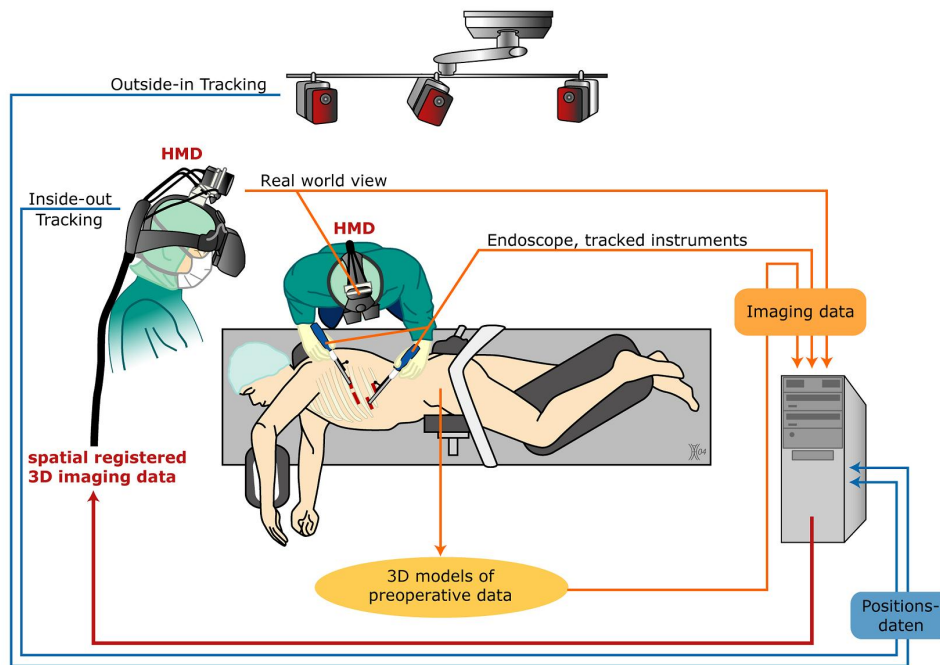


Figure 3: Our augmented reality tracking system consists of an outside-in tracking system, and a video see-through inside-out tracking system.

is red, the window cannot be set because some of the grid points are not located on the skin. When the frame is green the window can be set by key press. The size is adjustable by mouse interaction, which

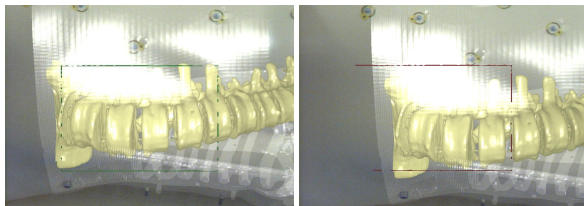


Figure 4: Window can be set when the frame is green, i.e. when all grid points lie on the surface of the virtual skin.

can be performed by an assistant on an external monitor that shows a copy of the imagery presented by the displays of the HMD. The window adopts the shape of the skin. Therefore a marching cube algorithm [7] segments the medical data volume and triangulates a surface model of the skin. The frame of the window defines the borders of a structured 2D grid consisting of a certain number of grid points. For every grid point a so-called *picking algorithm* examines the depth buffer at its corresponding pixel and recalculates three dimensional information of the nearest virtual object. If some of the grid points lie on the far plane of the frustum and not on the surface of the skin the color of the frame turns to red and the window can not be set. Discrimination of grid points on the far plane and triangulation of the remaining grid points on the skin would be another approach. However, we decided to allow for setting up the window completely or not at all because a rectangular frame helps to perceive the position, shape and orientation of the window. Picking in OpenGL can be realized with the following two functions.

```
void glReadPixels(GLint x,
                 GLint y,
                 GLsizei width,
                 GLsizei height,
                 GLenum format,
                 GLenum type,
```

```
GLvoid *pixels)
```

```
int gluUnProject(GLdouble winx,
                GLdouble winy,
                GLdouble winz,
                const GLdouble modelMatrix[16],
                const GLdouble projMatrix[16],
                const GLint viewport[4],
                GLdouble *objx,
                GLdouble *objy,
                GLdouble *objz)
```

After determination of their position in 3D space, the grid points are connected to compose a transparent surface.

When the position of the window is defined, it is used to mask the part of the scene, which is inside the thorax. Therefore we employ the so-called stencil buffer. The stencil buffer is an additional buffer besides the color buffer and depth buffer found on modern computer graphics hardware. One of the major application of the stencil buffer is to limit the area of rendering. In our application the area is limited to the window when the visualized tissue or bones are drawn. After that the window surface and all other objects, which are partially or completely outside the body, e.g. augmented surgical instruments, are rendered.

3.2 Window Design

The following design features help to intensify the depth cues provided by the window.

- The window is overlaid onto the skin. Its shape adapts the shape of the skin. Accuracy of overlay is defined by the level of detail of the surface model of the skin and the number of grid points of the window.
- Certain material parameters let the window appear like glass. Highlight effects due to the virtual light conditions support depth perception. Highlights on the window change the color

of objects behind the window or even partially occlude these objects.

- The window plane is mapped with a simply structured texture, which enhances the depth cue motion parallax. Due to motion of the observer the texture on the window seems to move relatively faster than objects behind the window. Furthermore, the texture helps to perceive the shape of the window [13].
- The frame of the window is colored. A second frame parallel to the outer one simulates the thickness of the window. Borders of the inner frame are visible depended to the current position of the observer.
- The background of the virtual objects seen through the window can be set to transparent or opaque.

Volume rendered objects as well as surface models can be positioned behind the window. Figure 5 and figure 7 illustrate these rendering modes regarding the exemplary application dorsal surgery. The following section describes the perceptive advantages of the virtual window.

4 PERCEPTIVE GAIN

Depth perception is still a major problem in many AR systems when virtual entities can only be displayed superimposed on real imagery. Cutting et al. summarized the most important binocular and monocular depth cues [3]. Our AR scene is perceived binocularly with the two color cameras mounted on the HMD. Stereopsis is realized by the slightly different perspectives of the two cameras. Convergence is predefined by the orientation of the cameras. Regarding pictorial and monocular motion induced depth cues, the most effective cue, occlusion, is responsible for the misleading depth perception when, in our case, the virtual spinal column occludes the real thorax. Adding the augmentation of the skin enables the observer to visually reorder the group of involved objects. The skin could be drawn transparent and positioned around the vertebrae. However, transparency provides only few information about relative position of the objects if one object is placed completely behind or in front of the other object as shown in figure 5. The window enhances perceptive information about depth because it partially occludes the vertebrae and the frame of the window covers and uncovers parts of the spinal column while the observer is moving. The latter depth cue motion parallax is after occlusion and stereopsis the third most effective source of information about depth [3].

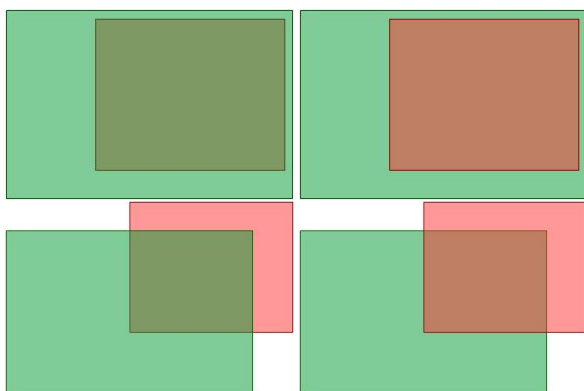


Figure 5: If the observer does not know the color of the planes and one of the objects completely occludes the other one, he or she cannot be sure about the relative position of the planes.

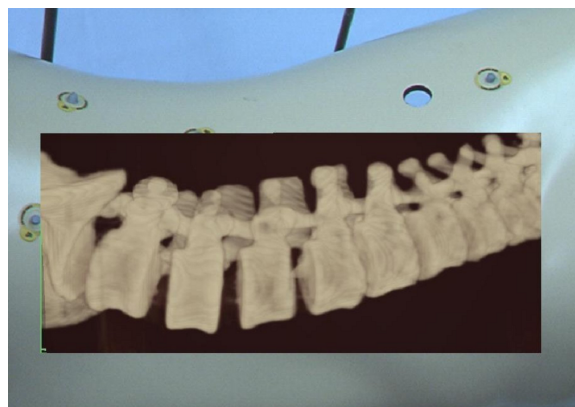


Figure 7: Volume rendered spinal column behind the window.

5 RESULTS AND CONCLUSION

We present a virtual window within a medical AR scenario that helps to overcome the misleading depth perception caused by the superimposed virtual spinal column onto the real thorax. An earlier experiment [12] compared seven different visualization modes of the spinal column regarding depth perception. We evaluated the visualization of a surface model of the spinal column behind a virtual window that is overlaid onto the skin as one of the two best methods. The method of posing the window interactively into the scene has the advantage that the surgeon or personnel of the OR do not have to use a further instrument that has to be kept sterile and wastes space. The observer wearing the HMD can easily position and reposition the window by moving his or her head. Figures 5 show a sequence while the observer is moving the HMD respective the thorax with the attached window. Beside surface models (figure 8) also volume rendered objects can be presented behind the window as shown in figure 7. The window provides the effective depth

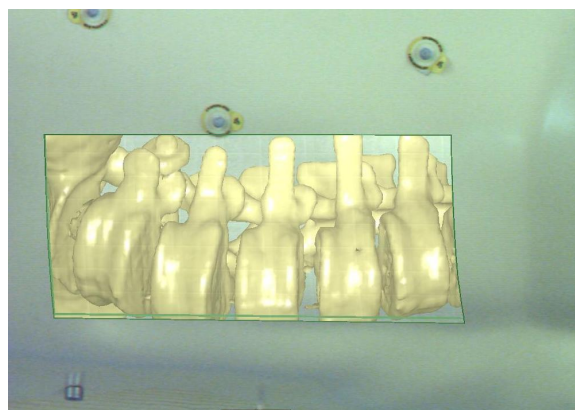


Figure 8: Surface model of the spinal column behind the window.

cues occlusion and motion parallax. However, if the surgeon moves his or her hand in front of his field of vision through the area of the window, this perceptive advantage is lost as shown in figure 9.

6 ACKNOWLEDGMENT

Special thanks to Tobias Sielhorst, Joerg Traub, Marco Feuerstein, Stefan Wiesner and Philipp Stefan for their support. This work was granted by the BFS within the NARVIS project (www.narvis.org).



Figure 6: Sequence shows the window from different perspectives. Each perspective provides another view on the inside of the throat.

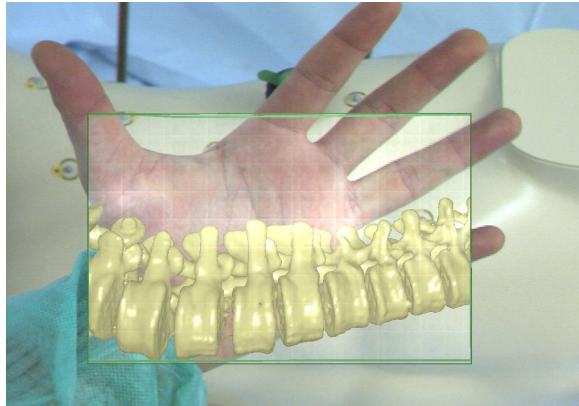


Figure 9: When the observer moves real objects within his or her field of vision like his hand or surgical instruments, the perceptive advantages due to the virtual window is lost again.

REFERENCES

- [1] Michael Bajura, Henry Fuchs, and Ryutarou Ohbuchi. Merging virtual objects with the real world: seeing ultrasound imagery within the patient. In *Proceedings of the 19th annual conference on Computer graphics and interactive techniques*, pages 203–210. ACM Press, 1992.
- [2] Wolfgang Birkfellner, Michael Figl, Klaus Huber, Franz Watzinger, Felix Wanschitz, Johann Hummel, Rudolf Hanel, Wolfgang Greimel, Peter Homolka, Rolf Ewers, and Helmar Bergmann. A head-mounted operating binocular for augmented reality visualization in medicine—design and initial evaluation. *IEEE Trans. Med. Imag.*, 21(8):991–997, 2002.
- [3] James E. Cutting and Peter M. Vishton. Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In *W. Epstein & S. Rogers (Eds.), Perception of Space and Motion*, pages 69–117, 1995.
- [4] W. A. Hoff and T. L. Vincent. Analysis of head pose accuracy in augmented reality. *IEEE Trans. Visualization and Computer Graphics*, 6, 2000.
- [5] A. P. King, P. J. Edwards, C. R. Maurer, Jr., D. A. de Cunha, D. J. Hawkes, D. L. G. Hill, R. P. Gaston, M. R. Fenlon, A. J. Strong, C. L. Chandler, A. Richards, and M. J. Gleeson. A system for microscope-assisted guided interventions. *IEEE Trans. Med. Imag.*, 19(11):1082–1093, 2000.
- [6] A.P. King, P.J. Edwards, C.R. Maurer, Jr., D.A. de Cunha, R.P. Gaston, M. Clarkson, D.L.G. Hill, D.J. Hawkes, M.R. Fenlon, A.J. Strong, T.C.S. Cox, and M.J. Gleeson. Stereo augmented reality in the surgical microscope. *Presence: Teleoperators and Virtual Env.*, 9(4):360–368, 2000.
- [7] William E. Lorensen and Harvey E. Cline. Marching cubes: A high resolution 3d surface construction algorithm. In *SIGGRAPH '87: Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, pages 163–169, New York, NY, USA, 1987. ACM Press.
- [8] Frank Sauer, Ali Khamene, Benedicte Bascle, Sebastian Vogt, and Gregory J. Rubinob. Augmented reality visualization in imri operating room: System description and pre-clinical testing. In *Proceedings of SPIE, Medical Imaging*, volume 4681, pages 446–454, 2002.
- [9] Frank Sauer, Ali Khamene, and Sebastian Vogt. An augmented reality navigation system with a single-camera tracker: System design and needle biopsy phantom trial. In *Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 2002.
- [10] Frank Sauer, Uwe J. Schoepf, Ali Khamene, Sebastian Vogt, Marco Das, and Stuart G. Silverman. Augmented reality system for ct-guided interventions: System description and initial phantom trials. In *Medical Imaging: Visualization, Image-Guided Procedures, and Display*, 2003.
- [11] Frank Sauer, Fabian Wenzel, Sebastian Vogt, Yiyang Tao, Yakup Genc, and Ali Bani-Hashemi. Augmented workspace: designing an ar testbed. In *Proc. IEEE and ACM International Symposium on Augmented Reality*, pages 47–53, 2000.
- [12] Tobias Sielhorst, Christoph Bichlmeier, Sandro Heining, and Nassir Navab. Depth perception a major issue in medical ar: Evaluation study by twenty surgeons. In *MICCAI*, 2006.
- [13] James T. Todd, Lore Thaler, and Tjeerd M.H. Dijkstr. The effects of field of view on the perception of 3d slant from texture. *Vision Research*, 45(12):1501–1517, 2005.