

# Contextual Anatomic Mimesis

## Hybrid In-Situ Visualization Method for Improving Multi-Sensory Depth Perception in Medical Augmented Reality

Christoph Bichlmeier<sub>1</sub>\*

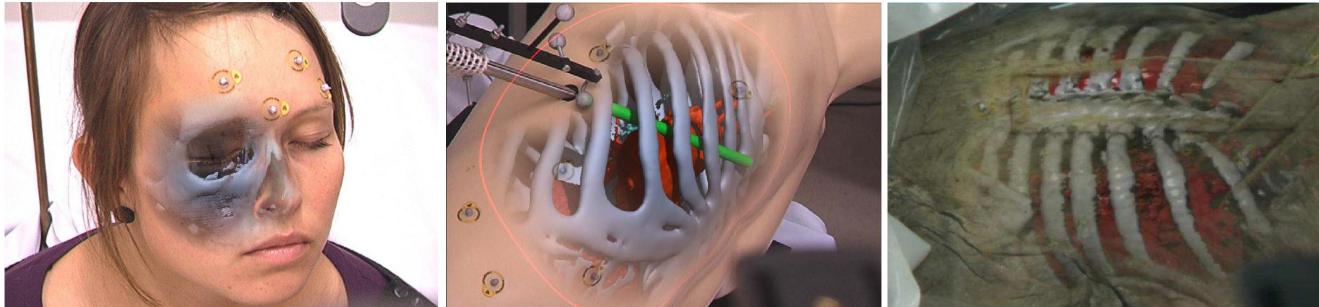
Felix Wimmer<sub>1</sub>†

Sandro Michael Heinig<sub>2</sub>‡

Nassir Navab<sub>1</sub>§

<sub>1</sub>Computer Aided Medical Procedures & Augmented Reality (CAMP), TUM, Munich, Germany

<sub>2</sub>Trauma Surgery Department, Klinikum Innenstadt, LMU, Munich, Germany



### ABSTRACT

The need to improve medical diagnosis and reduce invasive surgery is dependent upon seeing into a living human system. The use of diverse types of medical imaging and endoscopic instruments has provided significant breakthroughs, but not without limiting the surgeon's natural, intuitive and direct 3D perception into the human body. This paper presents a method for the use of Augmented Reality (AR) for the convergence of improved perception of 3D medical imaging data (mimesis) in context to the patient's own anatomy (in-situ) incorporating the physician's intuitive multi-sensory interaction and integrating direct manipulation with endoscopic instruments. Transparency of the video images recorded by the color cameras of a video see-through, stereoscopic Head-Mounted-Display (HMD) is adjusted according to the position and line of sight of the observer, the shape of the patient's skin and the location of the instrument. The modified video image of the real scene is then blended with the previously rendered virtual anatomy. The effectiveness has been demonstrated in a series of experiments at the Chirurgische Klinik in Munich, Germany with cadaver and in-vivo studies. The results can be applied for designing medical AR training and educational applications.

**Index Terms:** H.5.1 [Information Interfaces and Presentation]; Multimedia Information Systems—Artificial, augmented, and virtual realities; H.5.2 [Information Interfaces and Presentation]; User Interfaces—Interaction styles; I.3.6 [Computer Graphics]; Methodology and Techniques—Interaction techniques; J.3 [Life and Medical Sciences]

### 1 INTRODUCTION

Since the days of Leonardo da Vinci (see figure 1(a)), the ability to see into a living human system has been tremendously valuable in providing decisive insight for the medical community for diagnosis, surgical operation and post-operative examination. However, there is equal concern for the application of new methods that need to consider the patient's comfort as well as the risk of life and death consequences. The growing application of Augmented Reality (AR) provides the transfer of volumetric data in context to a live patient in real-time. The contextual accuracy of the augmentation in real-time and the intuitive representation of layered, complex interoperable biological systems can provide critical insight. However, there is equal concern of misdiagnosis, due to ineffective visualization methods when mixing realities with such layered asymmetrical complexity. There is a need to develop, evaluate and refine new methods for in-situ visualization to improve depth perception with AR and utilize the natural and intuitive multi-sensory manipulation in context with the actual living patient.

AR for medical in-situ visualization has been the subject of intensive research in the last decade. The objective of this research is the use of AR technology for preoperative diagnoses and surgical planning as well as intraoperative navigation. A medical AR scene can be presented to the observer using various display devices such as the "AR window" [24], endoscopic cameras [9] and different types of head mounted displays [14, 4]. Within the scope of this project, we employ the stereoscopic video see-through head mounted display presented by Sauer et al. [22]. Medical in-situ visualization of anatomy has to fulfill several conditions:

1. The visualization method has to provide a non-restricted view on the anatomical region of interest, which includes concealing obstructing anatomy and emphasizing the operation site.
2. Visualized imaging data has to be integrated in a way so that relative and absolute distances of objects within the AR scene can be perceived intuitively.
3. Using the AR technology intraoperatively, surgical instruments have to be integrated to enable the observer to perceive relative distances between anatomy and instrument and to get visual feedback due to this interaction.

\*e-mail: bichlmei@cs.tum.edu

†e-mail: wimmerf@in.tum.de

‡e-mail: Sandro-Michael.Heining@med.uni-muenchen.de

§e-mail:navab@cs.tum.edu

## 2 RELATED WORK

Various approaches for visualization of volumetric, medical imaging data such as *Computed Tomography (CT)*, *Magnetic Resonance Imaging (MRI)* and *Ultrasound* have been investigated in the field of computer graphics [16, 8] and evaluated for purpose in the medical workflow [21, 19]. The need for techniques to explore medical volume data maintaining the ability to understand spatial relations of visualized objects within the volume data is lively discussed. Burns et al.[6] presented an importance-driven rendering method to highlight focus regions in visualized medical imaging data. Bruckner et al.[5] presented an impressive "context preserving volume rendering model" [5] working with transparency effects. Their method uses "a function of shading intensity, gradient magnitude, distance to the eye point" [5] to define transparency of occluding layers. In particular using the distance to eye point is interesting to be adapted by in-situ visualization for medical augmented reality. Distance could be intuitively defined by the user wearing an HMD by moving slightly towards the augmented operation site. Krueger et al.[15] introduced *ClearView* an "interactive framework for texture-based volume ray-casting" [15]. Their method uses transparency combined with different shading effects to emphasize focus and context information, which supports the perception of spatial information within a region of interest inside the visualized volume data. Their visualization method uses a highlighted ring painted on the skin surface around the region of interest, which was adopted by our visualization method for augmented reality. Diepstraten et al. [7] propose a small set of guidelines for using transparency for the automatic generation of illustrative industrial visualizations to communicate interior structure of objects. According to the rules "faces of transparent context objects never shine through other transparent faces". Opaque objects which are occluded by two transparent objects do not shine through. Transparency falls off close to the edges of transparent context objects and increases with the distance to edges" [7]. Despite of providing a clear view onto the focus, contour rendering has the drawback that details of the transparent objects are lost and that it ignores material and surface information of the context objects [7]. Furthermore it only provides "two transparency states: fully opaque or fully non-opaque; semi-transparency cannot be visualized" [7]. Interrante et al. [12] determined "ridge and valley lines" providing visual cues to "communicate the shape and position of the transparent skin" and occluded anatomical objects. Another interesting approach presented by Levoy et al. [17]. They map a solid, simple shaped texture on an outer semi-transparent surface model occluding an inner opaque surface model to "improve the ability to distinguish the shape of the surfaces and their relationship to each other" [17].

However, the mentioned methods deal within virtual environments and do not consider the requirements for a medical AR scene. Improving data presentation for correct perception of depth, relative position and layout of "hidden and occluded objects", for instance the position and shape of anatomy inside the human body from the surgeons point of view, is a major issue in augmented reality [10, 1, 25]. Bajura et al., in the first publication [2] about medical AR, identified the problem of misleading depth perception when virtual anatomy occludes the patient. To handle the problem, they render a "synthetic hole [...] around ultrasound images in an attempt to avoid conflicting visual cues". An interesting approach to get interactively visual access to occluded objects is based on "magic lenses" [20]. The lens is used to discover hidden structures, which are presented distinctively compared to the objects laying outside the field of the lens. However, the lens requires in most cases an additional tracked device the guide the virtual tool interactively above the region of interest. Integration of such additional devices is not the objective of this work as hands of a surgeon wearing the HMD are occupied with the navigation of surgical instruments.

Sielhorst et al. [25] described an evaluation comparing the perceptual efficiency of seven different render modes visualizing the spinal column in situ. Their results show that a virtual window overlaid on the surface of the patient's skin is capable of overcoming the lack of misleading depth perception. The virtual window can be positioned interactively by head motion of the user wearing an HMD, which is described in detail in [3]. Both methods described in [3, 2] provide visual cues to perceive anatomical objects inside the patient, however the visualizations utilize highlighted, abstract objects, which are not smoothly integrated into the scene to create an intuitive view on the operation site. In contrast to our previous work [25], [3] we focus in the present paper on a more intuitive and elegant method of data presentation. We try to avoid additional abstract objects and rather use visual cues composed of natural properties of the individual anatomy and the observer's point of view. Several artists dealt with the problem of presenting anatomical structures and publish their images. Vesalius published the anatomy book *De Humanis Corporis Fabrica* in 1543 (figure 1(b)). In 1510 Leonardo da Vinci prepared his famous anatomical sketches presenting inner organs and muscles within the context of the human body (figure 1(a)). Development of the present visualization pipeline was highly inspired by the work of Alexander Tsiaras [26] illustrating the anatomy in an impressive way also from a point of view located outside the human body.

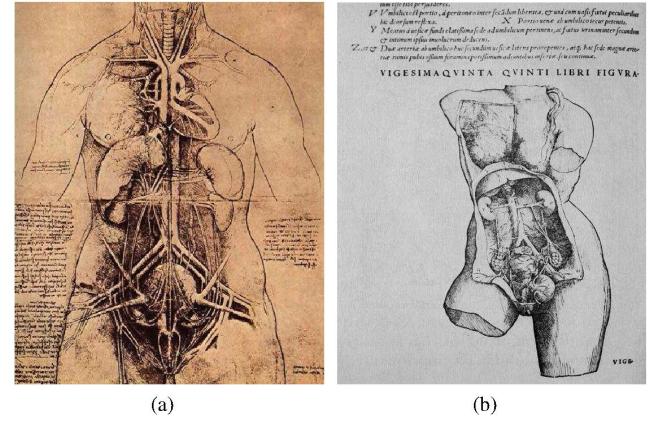


Figure 1: Anatomical drawing by Leonardo da Vinci(a) and Vesalius(b) 16th century .

This paper focuses on handling the problem of misleading perception of depth and spatial layout in medical AR (see figure 2) by adjusting the transparency of parts of the video image recorded by the color cameras of the HMD according to the shape of the patient's skin. The modified video image of the real scene is then blended with the previously rendered virtual anatomy. The technique is described in detail in section 3.2. In section 3.3 we present



Figure 2: Virtual anatomy, for instance the brain or the spinal column, is presented superimposed on the patient's head and seems to be located in front of the patient.

a method for the smooth integration of surgical instruments into the medical AR scene. Regarding keyhole surgery, the method highlights the port to the inside of the patient. Inside the body, the instrument is virtually extended to distinguish the position of the penetration through the skin surface. In addition the method employs the virtual shadow to provide visual feedback for user interaction and enable additional depth cues. For this reason, surgical instruments are virtually extended as soon as they penetrate into the patient's body.

Shadow is a natural illumination effect and provides information about the order of objects in an illuminated field of view when some of these objects, the *occluders*, cast shadows on others, the *receivers*. Kersten et al. [13] arranged a psychological experiment presenting the so-called ball-in-box scene. The authors claim: "The results support the hypothesis that the human visual system incorporates a stationary light source constraint in the perceptual processing of spatial layout of scenes" and "the information provided by the motion of an object's shadow overrides other strong sources of information and perceptual biases, such as the assumption of constant object size and a general viewpoint". The authors examine the perception of the 3D trajectory due to the shadow caused by the moving objects. The same group shows that an observer is able to gain information about "the shape of the object, the shape of the background surface and the spatial arrangement of the object relative to the background" [18]. They found out that "shadows were perceptually most relevant for the recovery of spatial arrangement, especially when the shadow is in motion". When the surgeon guides a surgical instrument close to the visualized anatomy, the shadow can be recognized on the virtual bones and tissue. According to Kersten et al., the human visual system is able to estimate the position of the light source and the relative distances and in this case, between the instrument and the spinal column.

### 3 METHOD

This section describes the AR system setup and the technique enabling an intuitive vision channel through the "real" skin onto virtual anatomy and an appropriate visualization of virtually extended surgical instruments.

#### 3.1 Tracking System

For superior registration quality, the system uses two synchronized tracking systems.

The single camera inside-out tracking RAMP system [22] allows for a high rotational precision [11] necessary for tracking the stereoscopic video see-through head mounted display (HMD). Two color cameras rigidly attached to the HMD simulate the eye's view. An additional infrared camera, mounted on the HMD, tracks a reference frame, a static set of retro reflective markers, for head pose estimation. 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" [23]. Second the system allows for more options 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 external, optical outside-in tracking system from A.R.T GmbH (Weilheim, Germany) with four infrared cameras fixed to the ceiling covers a large working area (around  $2.5m^3$ ). Once this system is calibrated, it provides stable tracking data for a long term unless cameras are moved. The system is capable of tracking the targets in our setup with an accuracy of  $0.35[mm]$ .

Both systems use the same kind of retro reflective fiducial markers offering a registration free transformation from one tracking system to the other. 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 every target that should be tracked. This could be for example the body of a patient (respectively a phantom) or surgical instruments. The reference 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 for the reference frame. All augmentations on targets, which are tracked by the optical outside-in tracking system, have to be positioned with respect to the reference frame of the inside-out tracking system. The following equation calculates the transformation  ${}^{target}T_{ref}$  from the reference frame to an exemplary target ( ${}^{target}T_{from}$ ).

$${}^{target}T_{ref} = {}^{target}T_{ext} * ({}^{ref}T_{ext})^{-1} \quad (1)$$

Transformations  ${}^{target}T_{ext}$  and  ${}^{ref}T_{ext}$  are provided by the external, optical outside-in tracking system. The former describes the transformation with respect to the origin of the tracking system to a target, the latter one is the transformation from the origin of the tracking system to the marker frame for inside-out tracking. Figure 3 provides a schematic overview of the AR system and figure 4 provides a view on the observer wearing the HMD, the reference frame, a thorax phantom and one of four infrared cameras of the outside-in tracking system. A PC based computer is used to render 3D

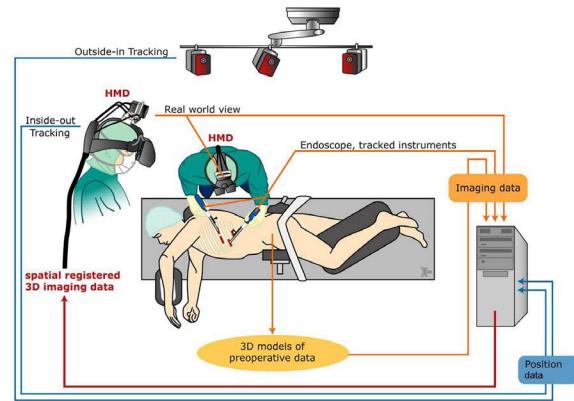


Figure 3: Schematic overview of the AR system including an external, optical tracking system and a stereoscopic, video see-through head mounted display with a single camera inside-out tracking system.

graphics, to compute and include tracking data, 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 HMD has a resolution of  $1024x768$ . Visualizations of virtual entities are implemented in C++ using OpenGL<sup>1</sup>.

The transformation for in-situ visualization of volumetric, medical imaging data can be described by  ${}^{CT}T_{ref}$ .

$${}^{CT}T_{ref} = {}^{CT}T * {}^{patient}T_{ext} * ({}^{ref}T_{ext})^{-1} \quad (2)$$

The optical outside-in tracking system provides the transformations  ${}^{patient}T_{ext}$  and  ${}^{ref}T_{ext}$ . Markers are segmented automatically from the volumetric imaging data. Correspondence of segmented positions and tracking data results in a registration matrix  ${}^{CT}T_{patient}$  that aligns the imaging data volume with the tracked object.

Our AR system provides capability for different visualization techniques such as 3D Texture & GPU-based volume rendering or polygonal surface models. The present approach works with surface models to reach real-time rendering of the anatomy.

<sup>1</sup>Open Graphics Library - www.OpenGL.org



Figure 4: The reference frame (see manually inserted red arrow), the thorax phantom and the instrument are tracked by the infrared camera attached to a frame work above the scene (see manually inserted red circle).

### 3.2 Fusion of Real and Virtual Anatomy

In this section we describe a new technique to manipulate the transparency of the camera image is manipulated as a function of the geometry of the skin and the line of sight of the observer wearing the HMD. By using this approach, the anatomy of interest becomes totally visible, except at positions where partial occlusion by the skin results in better perception of the focussed anatomy.

The visualization method divides the skin surface into two domains,

- transparent and semi-transparent skin within the vision channel - *TransDom*
- opaque skin outside the vision channel - *OpaDom*

The vision channel to the inside of the patient can be interactively defined by the line of sight of the observer wearing the HMD. Within *TransDom* only parts of the skin providing important contextual information are shown. Visualized anatomy within *OpaDom* are completely occluded by the skin. Selection of the remaining fragments on the skin is performed by three parameters (described in sections 3.2.1, 3.2.2, 3.2.3) calculated from geometrical properties of the skin surface and the line of sight of the observer to provide visual cues for correct perception of depth and spatial layout.

The transparency value of each single pixel of the video image is determined by off-screen-rendering of a skin model of the human body. This transparency calculation is done by a fragment shader

program exploring the geometrical properties of the skin. The computed transparency values are stored to a buffer, a designated texture called transparency map (render-to-texture approach)). This map is used to blend the video image with the previously rendered virtual entities of the scene, which is performed by alpha blending. The alpha values of the video image are determined by lookups into the generated transparency map using a second shader program. Three different parameters are taken into account to calculate the transparency (respectively opacity) of a pixel in the video image (see figure 5):

1. The curvature within the region around a fragment on the skin surface,
2. the angle between the fragments normal and the vector from the position of the fragment to the cameras of the HMD and
3. the distance of a fragment on the skin surface to the intersection point of the line of sight with the skin surface.

Finally these three parameters are combined and stored to the transparency map, which is used to determine the transparency of pixels of the camera image. The following sections describe the method in detail.

#### 3.2.1 Curvature

The first parameter *curvature* (see figure 5(a)) refers to the curvature around a particular 3D position of a fragment on the surface model of the skin [15]. A curvature value is determined for every vertex of the surface model of the skin and interpolated for every fragment during the rasterization process in the rendering pipeline. The curvature value for a vertex is calculated by comparing the normal vector  $\vec{n}$  of the vertex with the normal vectors  $\vec{n}_i$  of all neighboring vertices.  $N$  represents the set of normals of the neighboring vertices.

$$curv = 1 - \left( 1 - \frac{\sum_{n_i \in N} ||\vec{n} - \vec{n}_i||}{2 \cdot |N|} \right)^\alpha \quad (3)$$

The sum of length values resulting from  $||\vec{n} - \vec{n}_i||$  are calculated of every neighboring normal  $\vec{n}_i$ , whereas one length value corresponds to the angle between two normals. High length values implicate a high curvature value within the region of the vertex position. The division of the sum by two-times  $|N|$  clamps the curvature value to the range of  $[0, 1]$ .  $\alpha \in \mathbb{R}_+$  can be modified interactively by the user to adjust the parameter *curvature*. Increasing  $\alpha$  approximates the parameter *curvature* to 1, decreasing approximates it to 0. While increasing  $\alpha$ , regions with lower curvature contribute more to the opacity of the pixel. Thanks to the parameter *curvature* regions of the skin providing high curvature become more opaque than regions with low curvature. For this reason wrinkled, bumpy and sinuous regions on the skin can still be seen, while flat parts become transparent. A high weighting of the parameter can be suitable for the visualization within the face regions as anatomy provides high curvature due to nose, eyes and mouth (see figures 5). The calculation of the parameter has to be done only once, since the direction of normals on the surface model of the skin does not change.

#### 3.2.2 Angle of Incidence Factor

The second parameter *angle of incidence factor* (see figure 5(b)) is defined as a function of the vector pointing from the position of the fragment to the eye  $\vec{v}$  (view vector) and the normal  $\vec{n}$  on the skin surface. It provides contour information, since it calculates the dot-product of the normal and the view vector.

$$angOfInci = 1 - (\vec{n} \cdot \vec{v})^\beta \quad (4)$$

For approximately parallel vectors  $\vec{n}$  and  $\vec{v}$ , this parameter is close to 0 and results in low opacity. For almost perpendicular vectors  $\vec{n}$

and  $\vec{v}$  the parameter converges to 1 resulting in high opacity. Similar to the curvature, this parameter can be adjusted interactively by  $\beta \in \mathbb{R}_+$ . The parameter *angle of incidence factor* is modified by the observer wearing the HMD and defining the camera positions and consequently the view vector for every fragment on the skin surface. For this reason patches of the skin facing directly to the observer appear more transparent than others. In contrast to the curvature the parameter is view-dependent.

### 3.2.3 Distance Falloff

The third parameter *distance falloff* (see figure 5(c)) is a function of the distance between each surface fragment and the intersection point of the line of sight and the skin surface. The parameter can be calculated using any monotonically increasing falloff function. The transparency of skin patches decreases with respect to the distance to the intersection point. Skin patches having a greater distance to the intersection point than a chosen maximum distance are rendered completely opaque. Thus this parameter divides the skin in the *TransDom* and *OpaDom*, whereas the transparent region can be considered as the vision channel to the region of interest inside the body. The falloff function leads to a smooth border between the domains *TransDom* and *OpaDom*.

$$\text{distFalloff} = \left( \text{saturate}\left(\frac{\text{distToViewIntersecPoint}}{\text{radiusOfTransparentRegion}}\right) \right)^\gamma \quad (5)$$

*Saturate* is a function that clamps its input parameter to  $[0, 1]$ . The falloff is calculated by a division of the distance to the intersection point of the line of sight with the skin surface by the defined radius of the *TransDom*. The parameter *distance falloff* implements a restriction of the observer's vision channel to the important region of the virtual entities. The adjustable exponent  $\gamma \in \mathbb{R}_+$  defines the smoothness of the transition between the domains *TransDom* and *OpaDom*. A smaller value for  $\gamma$  results in a smoother transition and a higher degree of opaqueness close to the intersection point of the line of sight with the skin surface. Vice versa, a high value for  $\gamma$  results in a sharp transition of the domains and a low degree of opaqueness close to the intersection point. The position of the vision channel is interactively defined by the head motion of the observer.

### 3.2.4 Final Transparency/Opacity Determination

In order to calculate the opacity value of a pixel on the video image, the maximum of the three parameters is determined. In addition to the respective exponents  $\alpha$ ,  $\beta$  and  $\gamma$ , the parameters can be interactively weighted by  $w_1, w_2, w_3 \in \mathbb{R}^+$ .

$$\text{opacity} = \text{saturate}(\max(w_1 \cdot \text{curv}, w_2 \cdot \text{angOfInci}, w_3 \cdot \text{distFalloff})) \quad (6)$$

The resulting opacity value is then stored to the mentioned buffer called transparency map. A second fragment shader uses this map to define the alpha values of the pixels in the video image. Finally the colors of the previously rendered virtual objects are blended with the manipulated video image.

### 3.2.5 Highlighting the Border of the Semi-Transparent Domain

In order to attract the user's attention to the region of interest, the border between *TransDom* and *OpaDom* can be enhanced with a colored line on the skin surface having user-defined thickness. For this purpose the distance of a fragment to the intersection point of the line of sight and the skin surface, which was already determined for calculating the *distance falloff* parameter, is used. The border is rendered by setting the red fraction of colors of those fragments to 1 having a distance to the intersection point within the range  $[(1 - \text{thickness}) \cdot \text{maxdist}, \text{maxdist}]$ . *thickness* denotes a user-defined thickness value for the border and *maxdist* denotes the user-defined radius of *TransDom*.

### 3.2.6 Virtual Window to the Inside of the Patient

Initially the center of *TransDom* on the skin surface is the intersection point of the line of sight with the skin surface model. The observer wearing the HMD is able to navigate this domain to a region of interest by moving his/her head to get the desired vision channel into the patient. The vision channel can be fixed at a particular region. Enhancing the transition between *OpaDom* and *TransDom* by a colored borderline results in a round window overlaid on the skin surface (see figure 6) providing useful depth cues. While moving relative to the window, the observer gets an intuitive view on the anatomy of the patient and depth cues occlusion and motion parallax correct depth perception of objects within the field of view.

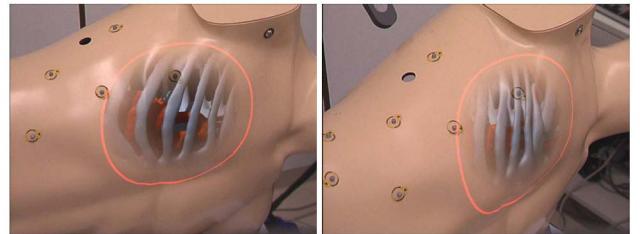


Figure 6: Moving around the fixed virtual window enables the depth cue motion parallax and improves depth perception of objects within the medical AR scene.

### 3.3 Integration of Surgical Instrument

In keyhole surgery, appropriate surgical instruments are inserted into the patient's body through a trocar<sup>2</sup> serving as a port to the operation site. In state-of-the-art systems, video images recorded by an endoscopic camera device are presented on a monitor viewing the operation site inside the patient and involved instruments. However, surgeons have to deal with visual cues from an additional environment presented on the monitor and integrate this spatial information into their mental model of the operation site.

Thanks to an HMD for in-situ visualization in minimally invasive surgery the fields of view can be reduced to only workspace namely the patient lying on the operation table. Furthermore such a system does not require the mental mapping of medical imaging data presented on a 2D display or attached to the wall with the patient. It integrates such data three dimensionally into the mentioned field of view. However, this requires new techniques for a smooth integration of such surgical tools into the AR scene to enable the surgeon to perceive relative distances between instruments and the operation site. For this reason, the augmentation of the instrument is presented only inside the patient's body. This is performed by rendering the instrument in two rendering passes after rendering the skin model. In the first pass the stencil<sup>3</sup> bit is set for those fragments of the instrument failing the depth test, that means fragments of the instruments laying under the skin surface inside the body. During the first rendering pass, drawing to the color buffer is disabled. In the second pass, fragments are stored to the color buffer if the stencil bit is set on their respective screen position.

The transparency settings of the video image around the penetrating instrument can be adjusted using the previously calculated intersection point of the instrument and the skin surface. Within the video images, the region around the port remains opaque, so

<sup>2</sup>(a) A surgical plastic device used to protect the port through the skin to the inside of the body for insertion of endoscopic instruments.

<sup>3</sup>(b) A sharp-pointed metal surgical instrument to puncture a body cavity for insertion of further surgical devices or substances.

<sup>3</sup>The stencil buffer is an additional buffer beside the color buffer and depth buffer on modern computer graphics hardware.

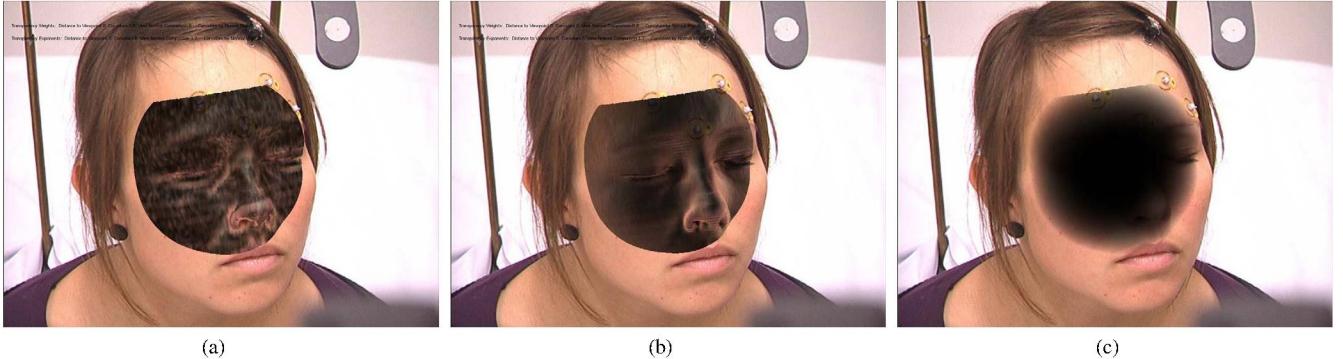


Figure 5: Figures show the three parameters, (a) *curvature*, (b) *angle of incidence factor*, (c) *distance falloff*, without any virtual objects in the black background. According to the particular parameter, different regions of the skin on the video images remain.

that the real instrument entering the patient is not occluded by the visualized anatomy. The transparency of pixels within this region depends on the distance to the estimated intersection point. In order to do this, a fourth parameter is used for transparency calculation of the skin around the region of the trocar, which is decreasing linearly with respect to the distance to the intersection point:

$$instrumentIntersectionFactor = 1 - \frac{distanceToInstrIntersecPoint}{maxdist} \quad (7)$$

In addition to phong shading for the illumination of the virtual anatomy, a shadow mapping algorithm was integrated to provide improved spatial information of virtual objects [27]. Shadow cast by the surgical instrument onto the anatomy provides helpful visual feedback due to user interaction, which can be used to navigate the tip of the instrument to the operation site.

### 3.4 Visualization Pipeline

This section summarizes all necessary rendering steps for the present visualization technique:

1. Generation of curvature values for each vertex of the skin model - this has to be done only once.
2. Calculation of the intersection point between skin and instrument - if an instrument is used.
3. Determination of the intersection point of the line of sight with the skin - if the vision channel is not fixed.
4. Calculation of the transparency map by rendering the skin model - parameters: center point of transparent region, maximum distance to the center point of the transparent region, "falloff distance" weight, "curvature" weight, weight of "angle of incidence factor", "falloff distance" exponent, "curvature" exponent, "angle of incidence factor" exponent, intersection point of the skin and the instrument, thickness of focus border.
5. Rendering the instrument and increasing the stencil bit when z-test fails.
6. Calculation of the shadow map - rendering instrument and datasets from the view geometry of the light source.
7. Rendering the virtual datasets - phong illumination, projection of fragment positions into the shadow map for shadow calculation.
8. Drawing the instrument where stencil bit is set.
9. Rendering the camera image with alpha blending manipulated with the transparency map.

## 4 RESULTS

The effectiveness of the presented visualization technique is demonstrated with the resulting images of a cadaver study, a thorax phantom and an in-vivo study.

### 4.1 Cadaver Study

A cadaver was scanned with a CT scanner. The resulting volumetric data set includes imaging data starting from the pelvis and ending at cervical vertebrae. Before data acquisition, CT markers were attached to the skin around the region of the spinal column to automatically register the visualization to the body. Figures 7 and 9 show the results of the cadaver study. We did not insert a surgical instrument (see figure 8) into the body, however figure 7(b) shows the shadow cast onto visualized ribs caused by a user guided surgical instrument and a virtual light source positioned above the scene. Figure 7(a) shows an expedient combination of the three transparency parameters described in section 3.2 including the highlighted borderline to create a virtual window (section 3.2.6). In



Figure 8: Tracked surgical instrument.

addition to the bones we rendered further tissue, basically the lung, shown in figures 9.

### 4.2 Phantom

Images of an augmented thorax phantom showcase the integration of the surgical instrument for keyhole surgery. Originally the phantom consisted only of a spinal column installed inside the phantom. However, we extended the phantom's virtual anatomy by surface models segmented from the Visible Korean Human (VKH)<sup>4</sup> data set. Virtual models are registered manually to the thorax phantom. This is sufficient for the evaluation of our visualization method. The region of the skin around the port for inserting the surgical instrument into the phantom gets opaque when the instrument is penetrating the body. Thanks to this effect, the view on the port is not

<sup>4</sup>The Visible Korean Human project provides different full body data sets: CT, MRI and anatomical images <http://vkh3.kisti.re.kr/new>



Figure 7: Spinal column and ribs can be visualized with a suitable transparency map for the video images (a). Shadow is cast by a surgical instrument onto the virtual ribs (b). The instrument was not inserted into the body during this study.

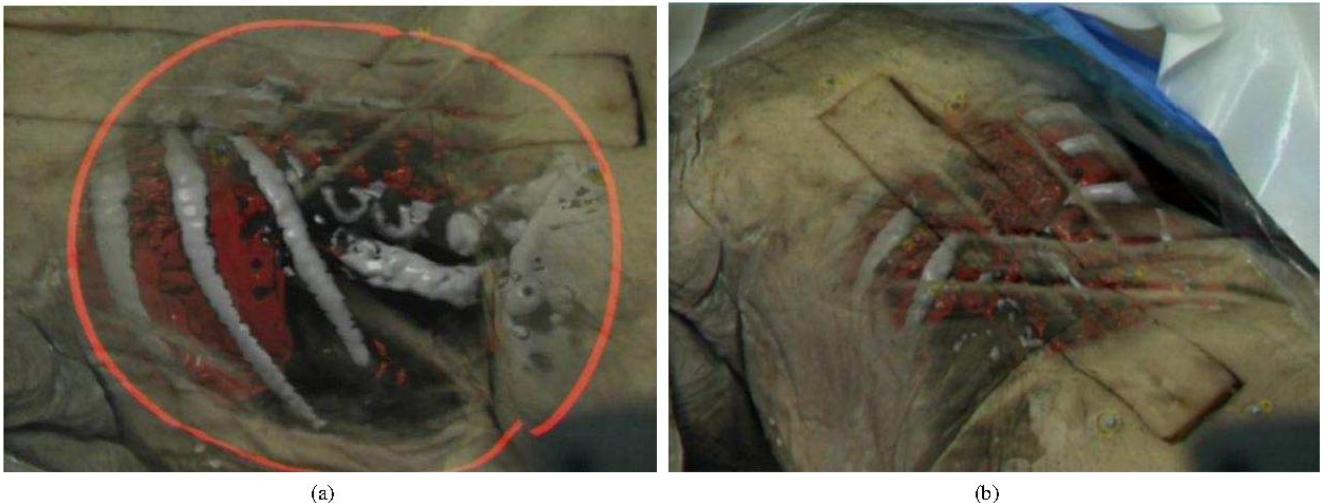


Figure 9: In addition to the visualization of the bones further segmented organs - basically the lung - are visualized. Lumbar vertebrae (a) and partially removed ribs using clipping volumes(b) to get better view on the lung.

restricted since it is not occluded by the virtual anatomy (see figures 10). The instrument and the anatomical surface models cast shadow on each other and thus provide helpful visual feedback while navigating the instrument to the operation site. Furthermore, the surgical instrument is virtually extended only inside the phantom.

#### 4.3 In-Vivo Study

We also had an opportunity to conduct a an in-vivo study. Susanne<sup>5</sup>, had a ski accident and got a CT scan of parts of her head. She is fine again and agreed to participate in the study.

The face structure provides high curvature due to nose, mouth and eyes and therefore is perfect to showcase the effectiveness of the present visualization method. At the day of the accident, no CT markers were attached before the CT scan. For this reason, visualized anatomy was registered manually with her head using natural landmarks. Figure 11 shows different combinations of the three parameters to calculate the transparency map. The vision channel

through the face surface follows the line of sight of the observer. Users can intuitively navigate the extended field of view to observe the patient and determine a reasonable visual port to the operation site (figures 11). When the region of interest is defined, the vision channel can be locked. It then turns to a virtual window overlaid on the skin surface. The window borders can be highlighted and broadened to enhance the visual cues occlusion and motion parallax for improved depth perception while the observer is moving around the virtual window. Also the patient can be moved while the virtual window remains at its previously defined position. We integrated a tracked spoon into the scene to interact with the in-situ visualization. In order to do this, we created an exact virtual counterpart of the spoon to visualize it also beneath the skin surface and to cast a shadow on visualized anatomy. We do not insert the spoon into the mouth because the data set did not include this region. The spoon was positioned under the skin surface by pressing it against the cheek, which makes the greenish virtual spoon visible (see figure 12(a)).

Shadow is also cast onto the virtual skull while moving the spoon

<sup>5</sup>name changed

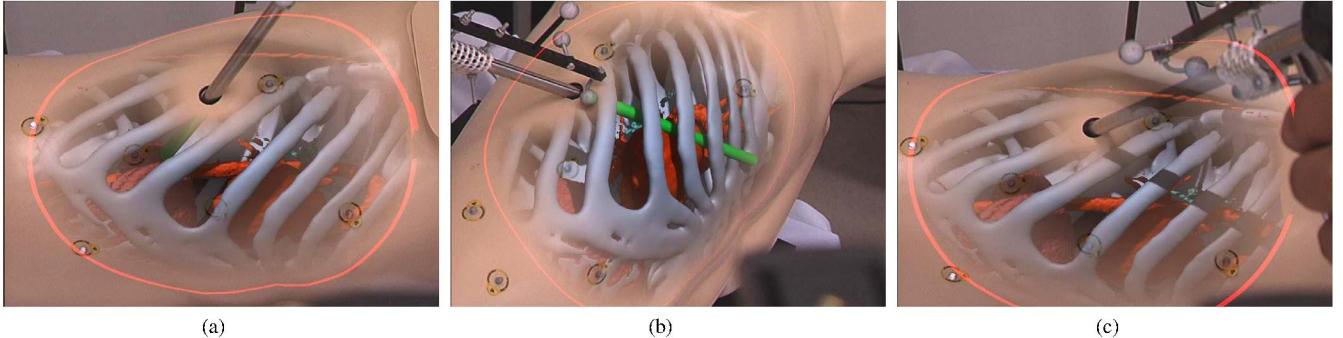


Figure 10: When the instrument is inserted through the port the skin remains opaque within the region of penetration (a)(c). Furthermore objects cast shadow on each other and therefore facilitate the navigation of the instrument (b)(c). The instrument is virtually extended inside the thorax phantom (b).

around the head space (see figure 12(b)(c)).

## 5 DISCUSSION

Our current approach uses polygonal surface models to allow for real time visualization. However, surface models have to be segmented, triangulated and smoothed before the visualization can be performed. The quality of the transparency effect, in particular created by the parameters *curvature* and *angle of incidence factor*, strongly depends on the accuracy level of the surface model of the skin. Rendering speed is regulated by the number of processed triangles of the anatomical surface models. For this reason high quality of data visualization suffers from low performance speed. Both factors are essential regarding the acceptance and ease of use for the AR system.

Future work will include the integration of the method into a ray cast based volume renderer to be able to render volumetric CT or MRI data directly without time wasting, preparative steps. Using such a rendering technique requires powerful hardware to obtain real-time rendering. Next step will be the exploitation of new graphics hardware and the customization of the present methods according to the hardware specification and specific requirements for medical AR.

In addition, we plan on integrating importance driven rendering [6] to be able to move the focus region to more deep-seated anatomical regions hidden by bones and further tissue. For instance blood vessel supplying a tumor are often embedded inside organs. Such blood vessel structure, for instance rips and occluding tissue, have to be presented in a way that information about relative position of anatomical entities as well as information about objects of interest at the operation site, such as the tumor, is optimally provided. Those blood vessels can be highlighted within CT data due to a previously injected contrast agent. Hence, unimportant anatomy has to be suppressed, however, their contextual function for perception of depth and spatial layout has to be remained.

Our method did not include effects like the ridge and valley lines [12] or textures [17] to communicate the position and shape of transparent or semi-transparent occluding layers. We plan to integrate such effects as their perceptive advantage for the medical AR scenario seems quite promising. Anyhow, it has to be evaluated, whether such additional abstract fragments can improve perception or rather hinder the view on the region of interest.

Within this scope, we plan to rearrange an earlier experiment [25], in order to evaluate the efficiency of the present visualization method regarding the correct perception of relative and absolute distances of objects within the AR scene. We recently started a project aiming at the integration of the present AR system into the operating room (OR). We determined the clinical intervention

*vertebroplasty* as a promising application in spine surgery to bring the system into the OR. Objective of the intervention is to stabilize porous vertebra. Cement is injected through a trocar, which is inserted into the vertebrae before. The intervention is performed in the trauma room right under a CT scanner to be able to control permanently position of the trocar and amount of cement injected into the vertebra. We plan to track the patient and the trocar and observe the whole medical procedure with an HMD. Although, in most cases the CT scan volume is limited to only a few affected vertebrae we plan to present the CT data with an extended version of the present method. The actual approach still need to much time for data preparation to present the system in the OR during an intervention. Once the method is integrated into the mentioned ray cast based volume renderer, volume data can be rendered directly and presented as shown in figure 7.

Using the present visualization method for a real patient in particular organs within the thorax and abdominal region, breathing affects the registration quality. This topic was not addressed by the present work. However, we assume that for spine surgery, for instance *vertebroplasty*, registration quality of vertebrae is sufficient, if the tracking target localizing the patient and registering the operation site is attached close to the operation site on the skin. Even if the patient is in prone position and the spine is moved due to breathing, relative distance of a suitable arranged CT marker set to the vertebrae remains adequately.

## 6 CONCLUSION

This paper presents a new method for medical in-situ visualization that allows for improved perception of 3D medical imaging data and navigated surgical instruments relative to the patient's anatomy. We describe a technique to modify the transparency of video images recorded by the color cameras of a stereoscopic, video see-through HMD. Transparency of the video images depends on the topology of the skin surface of the patient and the viewing geometry of the observer wearing the HMD. The presented method allows for an intuitive view on the deep-seated anatomy of the patient providing visual cues to correctly perceive absolute and relative distances of objects within an AR scene. In addition, we describe a method for integrating surgical tools into the medical AR scene resulting in improved navigation. The effectiveness has been demonstrated with a cadaver study and a thorax phantom, both visualizing the anatomical region around the spinal column, and an in-vivo study visualizing the head.

## ACKNOWLEDGEMENTS

We would like to thank Frank Sauer, Ali Khamene, and Sebastian Vogt from Siemens Corporate Research (SCR) for the design,

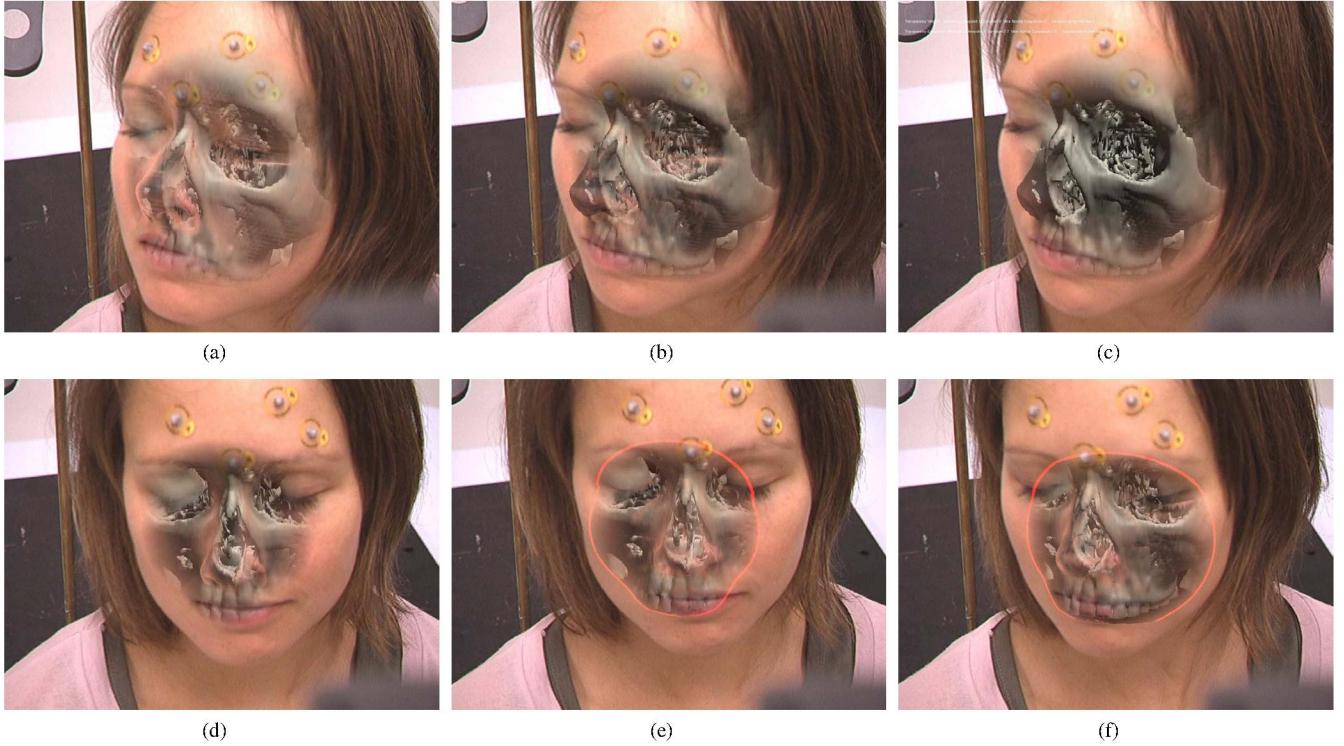
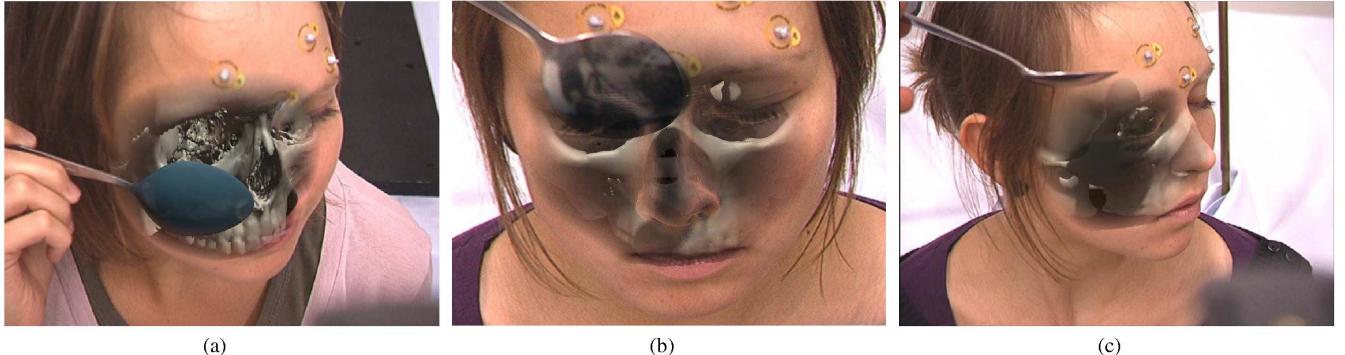


Figure 11: High "angle of incidence factor" (a). Combination of "angle of incidence factor" with the "curvature" parameter (b). High weighting of the "distance falloff" parameter (c). The observer navigates his or her extended field of view over patient (d) and defines a reasonable vision channel to the operation site. The vision channel can be locked and turns to a virtual window fixed on the skin surface (e). Also the patient can move relatively to the observer and the vision channel slides over the anatomy (f).

setup, and implementation of the RAMP system they provided us. Thanks to A.R.T. GmbH for providing cameras and software for the outside-in tracking system. We also want to express our gratitude to the radiologists and surgeons of *Klinikum Innenstadt Muenchen* for their precious contribution in obtaining medical data and evaluating our systems. We would like to thank Susanne Stettner for taking her time for the in-vivo study and Dr. J. Ehmann of the *Klinik für Chirurgie, Spital Saanen AG* to provide the CT data of Susannes head. Thanks also to Joerg Traub, Marco Feuerstein, Tobias Sielhorst, Mohammad Rustaee of the NARVIS group and Christopher Stapleton for their inspiration and support. This work was granted by the BFS within the NARVIS project ([www.narvis.org](http://www.narvis.org)).

## REFERENCES

- [1] R. T. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre. Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, 21(6):34–47, 2001.
- [2] M. Bajura, H. Fuchs, and R. Ohbuchi. Merging virtual objects with the real world: seeing ultrasound imagery within the patient. In *SIGGRAPH '92: Proceedings of the 19th annual conference on Computer graphics and interactive techniques*, pages 203–210, New York, NY, USA, 1992. ACM Press.
- [3] C. Bichlmeier, T. Sielhorst, S. M. Heining, and N. Navab. Improving depth perception in medical ar: A virtual vision panel to the inside of the patient. In *Proceedings of BVM 2007*. Springer, Mar. 2007.
- [4] W. Birkfellner, M. Figl, K. Huber, F. Watzinger, F. Wanschitz, J. Hummel, R. Hanel, W. Greimel, P. Homolka, R. Ewers, and H. 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.
- [5] S. Bruckner, S. Grimm, and A. Kanitsar. Illustrative context-preserving exploration of volume data. *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1559–1569, 2006. Member-M. Eduard Gröller.
- [6] M. Burns, M. Haidacher, W. Wein, I. Viola, and E. Groeller. Feature emphasis and contextual cutaways for multimodal medical visualization. In *EuroVis 2007 Proceedings*, May 2007.
- [7] J. Diepstraten, D. Weiskopf, and T. Ertl. Transparency in interactive technical illustrations, 2002.
- [8] K. Engel, M. Hadwiger, J. M. Kniss, A. E. Lefohn, C. R. Salama, and D. Weiskopf. Real-time volume graphics. In *SIGGRAPH '04: ACM SIGGRAPH 2004 Course Notes*, page 29, New York, NY, USA, 2004. ACM Press.
- [9] M. Feuerstein, T. Mussack, S. M. Heining, and N. Navab. Registration-free laparoscope augmentation for intra-operative liver resection planning. In *SPIE Medical Imaging*, San Diego, California, USA, February 2007.
- [10] C. Furmanski, R. Azuma, and M. Daily. Augmented-reality visualizations guided by cognition: Perceptual heuristics for combining visible and obscured information. In *Proc. IEEE and ACM Int'l Symp. on Mixed and Augmented Reality (ISMAR)*, page 215, Washington, DC, USA, 2002. IEEE Computer Society.
- [11] W. Hoff and T. Vincent. Analysis of head pose accuracy in augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 6(4):319–334, 2000.
- [12] V. Interrante, H. Fuchs, and S. Pizer. Enhancing transparent skin surfaces with ridge and valley lines. In *VIS '95: Proceedings of the 6th conference on Visualization '95*, page 52, Washington, DC, USA, 1995. IEEE Computer Society.
- [13] D. Kersten, P. Mamassian, and D. Knill. Moving cast shadows and the perception of relative depth, 1994.
- [14] 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–



(a)

(b)

(c)

Figure 12: A tracked spoon that is virtually extended inside the body. To position the spoon under the skin surface we press the spoon against the cheek (a). Shadow is cast by a tracked spoon onto the virtual skull (b)(c).

1093, 2000.

- [15] J. Krueger, J. Schneider, and R. Westermann. Clearview: An interactive context preserving hotspot visualization technique. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):941–948, 2006.
- [16] J. Krueger and R. Westermann. Acceleration techniques for gpu-based volume rendering. In *VIS '03: Proceedings of the 14th IEEE Visualization 2003 (VIS'03)*, page 38, Washington, DC, USA, 2003. IEEE Computer Society.
- [17] M. Levoy, H. Fuchs, S. Pizer, J. Rosenman, E. Chaney, G. Sherouse, V. Interrante, and J. Kiel. Volume rendering in radiation treatment planning. In *Proceedings of the First Conference on Visualization in Biomedical Computing*, pages 4–10, Los Alamitos, CA, USA, 1990. IEEE Computer Society Press.
- [18] P. Mamassian, D. Knill, and D. Kersten. The perception of cast shadows. *Trends in Cognitive Sciences*, 2(8):288–295, 1998.
- [19] M. Meissner, J. Huang, D. Bartz, K. Mueller, and R. Crawfis. A practical evaluation of popular volume rendering algorithms. In *VVS '00: Proceedings of the 2000 IEEE symposium on Volume visualization*, pages 81–90, New York, NY, USA, 2000. ACM Press.
- [20] E. Méndez, D. Kalkofen, and D. Schmalstieg. Interactive context-driven visualization tools for augmented reality. In *Proc. IEEE and ACM Int'l Symp. on Mixed and Augmented Reality (ISMAR)*, pages 209–218, 2006.
- [21] A. Pommert and K. H. Hoehne. Validation of medical volume visualization: a literature review. In *Computer Assisted Radiology and Surgery (CARS 2003)*, volume 1256, pages 571–576, 2003.
- [22] F. Sauer, A. Khamene, B. Bascle, S. Vogt, and G. 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.
- [23] F. Sauer, U. J. Schoepf, A. Khamene, S. Vogt, M. Das, and S. G. Silberman. Augmented reality system for ct-guided interventions: System description and initial phantom trials. In *Medical Imaging: Visualization, Image-Guided Procedures, and Display*, 2003.
- [24] B. Schwald, H. Seibert, and T. Weller. A flexible tracking concept applied to medical scenarios using an ar window. In *Proc. IEEE and ACM Int'l Symp. on Mixed and Augmented Reality (ISMAR)*, page 261, Washington, DC, USA, 2002. IEEE Computer Society.
- [25] T. Sielhorst, C. Bichlmeier, S. Heining, and N. Navab. Depth perception a major issue in medical ar: Evaluation study by twenty surgeons. In *Proceedings of MICCAI 2006*, LNCS, pages 364–372, Copenhagen, Denmark, Oct. 2006. MICCAI Society, Springer.
- [26] A. Tsiaras and B. Werth. *The Architecture and Design of Man and Woman: The Marvel of the Human Body, Revealed*. Doubleday, firsts edition, 2004.
- [27] Y. Uralsky. Efficient soft-edged shadows using pixel shader branching. In *GPU Gems 2*. Ed. NVIDIA Corporation, 2005.