Efficient Stereoscopic Rendering in Virtual Endoscopy Applications

A. del Río, D. Bartz, R. Jäger,
WSI/GRIS
University of Tübingen
Sand 14,
D-72076 Tübingen, Germany
{anxo, bartz}@gris.uni-tuebingen.de

Ö. Gürvit, D. Freudenstein,
Dept. of Neuroradiology, Dept. of Neurosurgery
University Hospital Tübingen
Hoppe-Seyler-Str. 3,
D-72076 Tübingen, Germany
ozlem.gurvit@med.uni-marburg.de
dkfreude@med.uni-tuebingen.de

ABSTRACT

Optical endoscopy suffers from several problems which increase the difficulties of a successful intervention. Among these problems are the limited spatial or depth perception, and the fish-eye effect which virtually flattens the geometry of the anatomical structures. Standard virtual endoscopy is inflicted by similar problems. In this paper, we present stereoscopic VIVENDI, an approach for the endoscopic visualization of complex anatomical data which increases the depth and spatial perception significantly. The stereoscopic examinations can be viewed with standard shutter-glasses, or with simple red/blue anaglyph glasses which additionally enable stereoscopic viewing with low-end PC hardware. Efficient rendering has been achieved by modifying an existing culling approach to accommodate stereoscopic imaging.

Keywords
Stereoscopic rendering, virutual endoscopy, visibility driven rendering.

1. INTRODUCTION

The relevance of computer-based planning systems for technical challenging medical procedures, has significantly increased in recent years. Virtual endoscopy systems like VIVENDI [BS99] demonstrated their usability for a wide range of different minimally-invasive interventions, which are among the most challenging medical procedures. These interventions include colonoscopy [VSG+94], bronchoscopy [VSH+94], angioscopy [BSSW99], and neuroendoscopy [BS99; BSG+01].

Journal of WSCG, Vol.11, No.1., ISSN 1213-6972
WSCG 2003, February 3-7, 2003, Plzen, Czech Republic.

Uncommon viewing angles and large magnifications increase the difficulties of the spatial depth perception in optical endoscopic procedures. This problem is aggravated by the fish-eye effect, which describes the flattening of displayed structures with the large viewing angle of common optical endoscopes. Combining these viewing limitations with the limited flexibility of the movement of the endoscope results in very little spatial depth perception.

Virtual endoscopy mimics an optical endoscope, while providing significantly more movement flexibility. This flexibility and the resulting motion parallax are used as an additional cue for a better sense of the orientation and location of anatomical features. However, complex and narrow features are difficult to understand when only the motion parallax cue is available. Stereoscopic rendering techniques exploit the ability of the human visual system to integrate two slightly translated (and rotated) perspective images of a scene – representing the left and the right eye – into a three-dimensional representation. With this mental three-dimensional representation, we can increase the spatial depth perception in virtual endoscopy applications, which in turn enhances the orientation in an organ system and distance estimation between the various anatomical features. Therefore, we can improve

the quality of the planning of a minimally-invasive intervention.

Unfortunately, only few virtual endoscopy systems provide stereoscopic visualization, which requires a two pass perspective rendering from the slightly transformed view-points. Basically, only the SGI OpenInventor-based FreeFlight system potentially enables stereoscopic rendering [VSA+97] at reasonable speed on high-end graphics systems (at that time SGI InfiniteReality), while no direct volume rendering approach currently provides sufficient performance. VIVENDI in contrast uses a visibility driven rendering which removes the geometry of anatomical features that are occluded by other anatomical features from the current view-point. With a small relaxation of the viewing parameters (enlarging the field of view), we can use the generated visibility information for the rendering of the left and right image, thus significantly reducing the visibility overhead and hence increase the framerate.

While we focused on the medical usefulness in [BJG+02], we concentrate here on efficient rendering issues. In the remaining paper, we will briefly review the related work in the field of stereoscopic rendering (Section 1.1). In Section 2, we describe the rendering techniques we used to generate the stereoscopic images (Section 2.1) and discuss the visibility driven rendering technique (Section 2.2). In Section 3, we summarize the results of stereoscopic rendering for virtual endoscopy and draw a conclusion in Section 4.

1.1. Related Work

There has been significant previous work in the field of stereoscopic rendering, and in virtual endoscopy. Here, we focus on a brief overview on stereoscopic rendering and stereoscopic virtual endoscopy, while we refer to [BSSW99; BSG+01] for a recent overview on virtual endoscopy. A good overview on stereoscopic basics is provided by Lipton [Lip97] and Hodges [Hod92].

Among many different approaches, two main avenues for stereoscopic rendering are well established; head-mounted-displays (HMDs) where two (LCD) screens are mounted right in front of the eyes, and head-tracked-displays (HTD) [WHR99]. Here, we focus on HTDs. Early stereoscopic rendering assumed a fixed eye position which leads to distortions if the actual eye position of an observer is too far away from the assumed position [Lip91]. Later, tracking of the observer position was introduced [Dee92] which was extended to the virtual table paradigm [KF94], and to the CAVE [CNSD93]. In particular the original GMD implementation of the virtual table, the *responsive workbench*, was later extended for two users [ABF+97]. Most of this work uses StereoGraphics CrystalEyes shutter glasses [Lip91], where a transmitter synchronizes the shutting of the left and right glasses with the interleaved rendering of the right and left images. Many researchers also explored the parameter space for good stereoscopic perception [YS90; Ros93; WRMW95; WHR99], where in particular parameters like eye separation, disparity, and zero-parallax distance (distance to the focus plane) were examined.

The only virtual endoscopy system known which potentially includes stereoscopic rendering capabilities is the OpenInventor-based FreeFlight system [VSA+97]. Like all OpenInventor-based applications, two images are generated from the geometry, which are synchronized with StereoGraphics CrystalEyes shutter glasses [Lip91]. Unfortunately, these techniques neither enable offline viewing of stereoscopic animations, as discussed later, nor do they reduce the amount of geometry to be rendered to a feasible quantity (for interactive rendering).

2. METHODS

2.1. Stereoscopic Rendering

Stereoscopic rendering is an important cue for depth perception among others like motion parallax, perspective projection, depth cuing, shading, or shadows. During motion, the motion parallax provides the strongest depth perception. However, if motion is limited or even not possible at all (ie., in virtual endoscopy), the perception of depth is limited too. In contrast, stereoscopic rendering provides an effective depth perception without additional movement. Therefore, it also can be used to generate still images with good depth perception.

For stereoscopic rendering, two images for the left and the right eye of a viewer are generated. Both images are rendered from view-points which are slightly translated (and rotated to keep the focus point) from the original monocular view-point according to the interocular distance, or eye separation of the viewer (e in Fig. 1). While measured eye separation of the users provides an accurate parameterization, often an average eye separation is used [Ros93] which usually leads to an under- or overestimation of the eye separation. In some applications, this under- or overestimation can be very helpful to enhance the depth perception; overestimation increases the depth perception, while it also increases the difficulties of the human visual system to fuse the left and right image to a stereoscopic image. In contrast, underestimation of the eye separation reduces the difficulties of image fusion, but it also reduces the depth perception. However, the appropriate (but not necessarily accurate) parameterization depends on the specific application, and on the individ-

ual abilities of the users [YS90][1]. Note that inaccurate eye separation also leads to distortion of the projected geometry [WHR99].

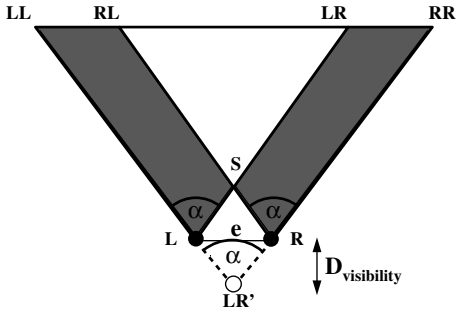


Figure 1: Modified field of view: The left field of view (L,LR,LL) and the right field of view (R,RR,RL) are combined to the extended field of view (LR',RR,LL). The white area of (S,LR,RL) is the field of stereoscopic view; e is the interocular distance (eye separation).

The view-points of the users of virtual endoscopy applications are usually located within the scene. Consequently, the geometry should be perceived as being behind the screen – the *stereo window* –, or in other words, in *CRT space*. If parts of the scene extend into the *viewer space* (and hence are cut off), between the screen and the viewer, stereoscopic artifacts due to the conflicting depth cues with the two-dimensional surrounding will be perceived as blurred or out-of-focus [Lip97]. Therefore, we reduce these artifacts by choosing a positive parallax where the focus plane is located in the foreground.

For our standard stereoscopic rendering, we support two different variants. On high-end graphics systems with quad-buffer support (front/back, left/right), we use synchronized shutter glasses [Lip91; Lip97] to provide the left and right images to the eyes. This standard approach in stereoscopic rendering has the advantage that colors can still be used to highlight multi-modal features. One of the drawbacks, however, is that it is difficult to provide the synchronization information for the shutter glasses in video animations. Therefore, we also apply the simple red/blue stereo technique, where the left image is rendered using only the red channel of the RGBA color buffer, and the right image is rendered using only the blue channel of the RGBA color buffer. Finally, these two images are combined to the final image. Red/blue stereoscopic images can be viewed with very limited technical equipment; basically, only red/blue cardboard glasses (approximately 0.50 US$ per piece) are required as additional hardware. Red/blue stereo videos can also be broadcasted/provided via the Internet, totally detached from the potentially expensive graphics system which rendered the images.

[1] Yeh and Silverstein describe "large individual differences in stereoposis in the population".

2.2. Visibility Driven Rendering

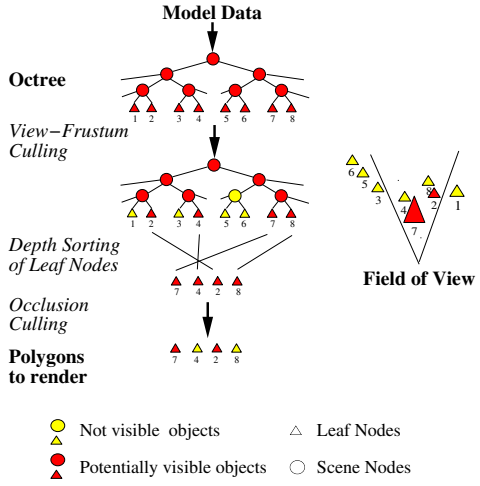


Figure 2: Hierarchical occlusion culling

The data of virtual endoscopy applications is usually generated by three-dimensional tomographic scanners like CT, MRI, or rotational angiography. We decompose the volume datasets into smaller entities using an octree approach [BS99]. Based on this decomposition, we extract the isosurfaces that are associated with the organs of interest using the Marching Cubes algorithm [LC87], resulting in octree leaf blocks which contain a similar number of triangles. This surface reconstruction leads to a large number of small triangles, which can range from 500K to several million triangles. This large polygonal complexity prevents the interactive visualization (at least 10 fps) of the extracted polygonal model. Therefore, we remove the geometry which is associated with the octree blocks that are definitely not visible from the current view-point using a hierarchical occlusion culling approach (see Fig. 2).

As already described in previous publications [BMH99; BS99], we apply an OpenGL-based view-frustum culling test hierarchically to the octree which decomposes the volume dataset and the extracted polygonal model [Cla76; BMH99]. This operation results in a depth-sorted list of octree leaf blocks (according to the near depth value of the respective bounding box) which are at least partially located within the view-frustum. Thereafter, we render the geometry of the front most 10% octree leaf blocks without any occlusion test, since they are virtually al-

ways visible[2]. All the remaining geometry is rendered using the HP occlusion culling flag [SOG98]. An axis-aligned bounding box which contains the geometry of an octree leaf block is rendered in a hardware-supported occlusion mode which does not contribute to the framebuffer. If the bounding box would modify the depth buffer, the HP occlusion culling flag returns TRUE and FALSE otherwise. Depending on this result, the geometry associated with the not occluded bounding box is rendered [BS99].

Naïve stereoscopic visibility driven rendering runs two individual rendering and culling passes for the left and the right image which doubles the rendering costs. Here, we modify the visibility driven rendering to exploit the occlusion coherence between the left and right image with a simple method. Specifically, we increase the field of view by moving the view-point backwards such that the new field of view contains the left and right field of views ((LR',RR,LL) in Fig. 1), while keeping the near plane of the original view-frusta. The amount $D_{visibility}$ of this backward movement of the view-point is computed by evaluating Equation 1, according to Figure 1. Therefore, the visibility

$$D_{visibility} = \frac{e}{2}\sqrt{\frac{1}{\sin^2\frac{\alpha}{2}} - 1} \qquad (1)$$

information is required only once which saves a significant amount of time consumed by the otherwise necessary second pass of visibility tests. However, the larger field of view also increases the number of visible octree leaf blocks, and hence it also increases the number of triangles to be rendered. Fortunately, measurements show that this increase is less than 0.1% of the total geometry (with the associated costs of approximately 0.01 fps), while the additional costs for a second visibility test pass with the original field of view would account for approximately two fps. The total framerate is on average 10 fps on an HP X class PC with a 750MHz PIII CPU and a VISUALIZE fx6 graphics card running LINUX. This compares to 16 fps of the non-stereoscopic standard rendering.

Red/blue stereoscopic rendering (see Section 2.1) requires the compositing of the red left and the blue right rendered images of the scene. For the multi-modal stereoscopic rendering of endoscopic images [BSG+01], red/blue stereoscopic rendering cannot be used, since we require (at least) two easy distinguishable colors for the rendering of the various anatomical structures (ie., blood vessels in red, ventricular system in white), which is not really possible with red/blue (or red/green) rendering. Therefore, we apply the standard shutter-glass interleaved rendering using a quad-buffer and StereoGraphics CrystalEyes shutter glasses [Lip91]. This requires a sufficiently correct transparent rendering of the geometry. Unfortunately, all the geometry located within the view-frustum is potentially visible with transparent rendering. Hence, there is virtually no occluded geometry. Due to the blending attenuation however, the visual impact of the geometry is smaller the farther away the geometry is. Our experiments exposed some popping artifacts during the walk-throughs of the geometry because of changes of the visibility status of the respective octree leaf block[3]. Furthermore, the blending operation is slightly more expensive than non-blended rendering which accounts for approximately 0.5 fps.

Note that the correct blending of the geometry requires a primitive depth sorting on triangle level. Although the number of triangles has been reduced tremendously by occlusion culling, it still represents significant costs which prevent interactive performance (sorting costs and display lists cannot be used). For this reason, we only sort the geometry on octree leaf level, resulting from the view-frustum culling step. In our walk-throughs, we noticed only minor visual artifacts in comparison to the full depth sorting of the triangles. Similar to the popping artifacts, the block sorting artifacts are attenuated by the blending of the left and right image.

3. STEOROSCOPIC VIRTUAL ENDOSCOPY

We applied stereoscopic VIVENDI to several virtual endoscopy applications. Specifically, we examined the complex anatomical topology of aneurysms of cerebral blood vessels, where the connections between the neck of the aneurysms and the supplying blood vessels need to be examined (see Fig. 3a). Other applications included the technical challenging ventriculoscopy, which was planned with multi-modal VIVENDI [BSG+01] (see Fig. 3c). Our medical collaborators assessed the additional depth perception with stereoscopic rendering in comparison to the usual rendering with monocular cues only. Although we

[2]We performed several experiments that measure the distribution of visible octree leaf blocks over a set of typical walk-throughs through the organ of interest. In approximately 98% of the test, the 10% front most octree leaf blocks were always visible.

[3]With transparent rendering, visibility is only reduced by the attenuated display (due to the blending) of the geometry, but not by occlusion of the geometry. In contrast, occlusion computations are usually based on opaque surfaces. This results in minor popping artifacts when the geometry of an octree leaf block (or its bounding box) becomes not occluded (as opaque object), although it could have already been (attenuated) visible before (as semi-transparent object). Also note that correct transparent rendering actually requires back-to-front rendering, while we are using front-to-back rendering to achieve reasonable occlusion. However, the visual effects due to this combination of culling techniques and transparent blending are negligible.

did not conduct experiments which compared the performance of depth perception like in [Ros93; YS90; WRMW95], our medical users noticed a significant improvement of the understanding of the complex anatomical structures and for the navigation through these structures. The lack of a multi-color rendering using red/blue techniques did not pose additional problems for the virtual explorations of datasets from a single modality.

Popping and sorting artifacts of the transparent multi-modal representation did not have a significant visual impact, specially considering the typical size of the considered pathologies. Whereas the popping artifacts were noticeable while moving through the virtual organ, they were not perceived as annoying or disturbing, and in particular, they did not lead to overseeing important features. Sorting artifacts were even less apparent, mostly due to the small section of the object surface a single leaf block is occupying, which in turn resulted in very little self-overlapping – the potential problem of leaf-block based sorting – of the projected surface. Furthermore, the effects of the both types of artifacts were reduced by the blending attenuation of the transparent rendering.

The stereoscopic parameters (mostly eye separation) were predefined for each specific applications. Users who were not familiarized with stereoscopic images experienced fusing the left and right images as quite strenuous, if the overestimation of the eye separation was too large. However, we could reduce this effect with the user-specific reduction of the eye separation. Another problem arose with the increasing polygonal complexity of the scene of the online synchronized shutter glasses version. The reduction of the framerate led to a flickering display which resulted in an increased fatigue.

4. CONCLUSIONS AND FUTURE WORK

In this paper, we presented stereoscopic VIVENDI, a virtual endoscopy system which integrated stereoscopic rendering to increase the depth perception for virtual endoscopy applications. Two different stereoscopic rendering techniques were applied; synchronized shutter glasses-based rendering on graphics systems with quad-buffer support, which enables a multi-color representation of the scene, and simple red/blue monochrome stereoscopic rendering which is also viewable through video animations via the Internet. Besides the very limited color representation, red/blue stereoscopic rendering also provided only a limited (perceived) contrast in comparison to the possible black/white representation of the synchronized approach.

Rendering at sufficient framerates was achieved with a visibility driven rendering approach which reduces the polygonal complexity of the scene to be rendered significantly. Furthermore, we exploited occlusion coherence by relaxing the view-frustum parameters to significantly reduce the visibility computation overhead for stereoscopic rendering .

With stereoscopic rendering, our users experienced a significantly increased depth perception which helped with the understanding of the complex anatomy and the navigation through this anatomy, while the associated artifacts did not introduce any significant obstructions.

Future work will also focus on other stereoscopic rendering technique that do not require complex additional hardware and synchronization information, but enable a higher quality multi-color depth perception.

5. ACKNOWLEDGEMENTS

This work has been supported by the Hewlett-Packard Workstations Systems Lab, Ft. Collins, USA, by European Project DynCT, and by DFG Project CatTrain. Datasets were provided by the Department of Neuroradiology of the University Hospital in Tübingen. Especially, we would like to thank Martin Skalej of the Department of Neuroradiology for his help and support in this project, and Stan Stoev for valuable discussions.

6. REFERENCES

[ABF+97] M. Agrawala, A. Beers, B. Fröhlich, P. Hanrahan, I. McDowall, and M. Bolas. The Two-User Responsive Worksbench: Support for Collaboration Through Independent Views of a Shared Space. In *Proc. of ACM SIGGRAPH*, pages 327–332, 1997.

[BJG+02] D. Bartz, R. Jäger, Ö. Gürvit, D. Freudenstein, and W. Straßer. Räumliche Tiefenwahrnehmung in virtuell-endoskopischen Anwendungen durch stereoskopisches Rendern. In *Proc. of Workshop Bildverarbeitung in der Medizin*, Informatik Aktuell, pages 393–396, 2002.

[BMH99] D. Bartz, M. Meißner, and T. Hüttner. OpenGL-assisted Occlusion Culling of Large Polygonal Models. *Computers & Graphics*, 23(5):667–679, 1999.

[BS99] D. Bartz and M. Skalej. VIVENDI - A Virtual Ventricle Endoscopy System for Virtual Medicine. In *Proc. of Symposium on Visualization*, pages 155–166,324, 1999.

[BSG+01] D. Bartz, W. Straßer, Ö. Gürvit, D. Freudenstein, and M. Skalej. Interactive

and Multi-modal Visualization for Neuroendoscopic Interventions. In *Proc. of Symposium on Visualization*, pages 157–164, 2001.

[BSSW99] D. Bartz, W. Straßer, M. Skalej, and D. Welte. Interactive Exploration of Extra- and Intracranial Blood Vessels. In *Proc. of IEEE Visualization*, pages 389–392,547, 1999.

[Cla76] J. Clark. Hierarchical Geometric Models for Visible Surface Algorithms. *Communications of the ACM*, 19(10):547–554, 1976.

[CNSD93] C. Cruz-Neira, D. Sandin, and T. DeFanti. Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE. In *Proc. of ACM SIGGRAPH*, pages 135–142, 1993.

[Dee92] M. Deering. High Resolution Virtual Reality. In *Proc. of ACM SIGGRAPH*, pages 195–201, 1992.

[Hod92] L. Hodges. Tutorial: Time-Multiplexed Stereoscopic Computer Graphics. *IEEE Computer Graphics & Applications*, 12(2):20–30, 1992.

[KF94] W. Krüger and B. Fröhlich. The Responsive Workbench. *IEEE Computer Graphics & Applications*, 4:12–15, 1994.

[LC87] W. Lorensen and H. Cline. Marching Cubes: A High Resolution 3D Surface Construction Algorithm. In *Proc. of ACM SIGGRAPH*, pages 163–169, 1987.

[Lip91] L. Lipton. *The CrystalEyes Handbook*. StereoGraphics Corp., San Rafael, 1991.

[Lip97] L. Lipton. *Stereo3D Handbook*. StereoGraphics Corp., San Rafael, available from http://www.stereographics.com/html/whtpaprs .html, 1997.

[Ros93] L. Rosenberg. The Effect of Interocular Distance upon Operator Performance Using Stereoscopic Displays to Perform Visual Depth Tasks. In *Proc. of IEEE Virtual Reality Annual International Symposium*, pages 27–32, 1993.

[SOG98] N. Scott, D. Olsen, and E. Gannett. An Overview of the VISUALIZE fx Graphics Accelerator Hardware. *The Hewlett-Packard Journal*, (May):28–34, 1998.

[VSA+97] D. Vining, D. Stelts, D. Ahn, P. Hemler, Y. Ge, G. Hunt, C. Siege, D. McCorquodale, M. Sarojak, and G. Ferretti. FreeFlight: A Virtual Endoscopy System. In *CVRMed-MRCAS*, volume LNCS 1205, pages 413–416, 1997.

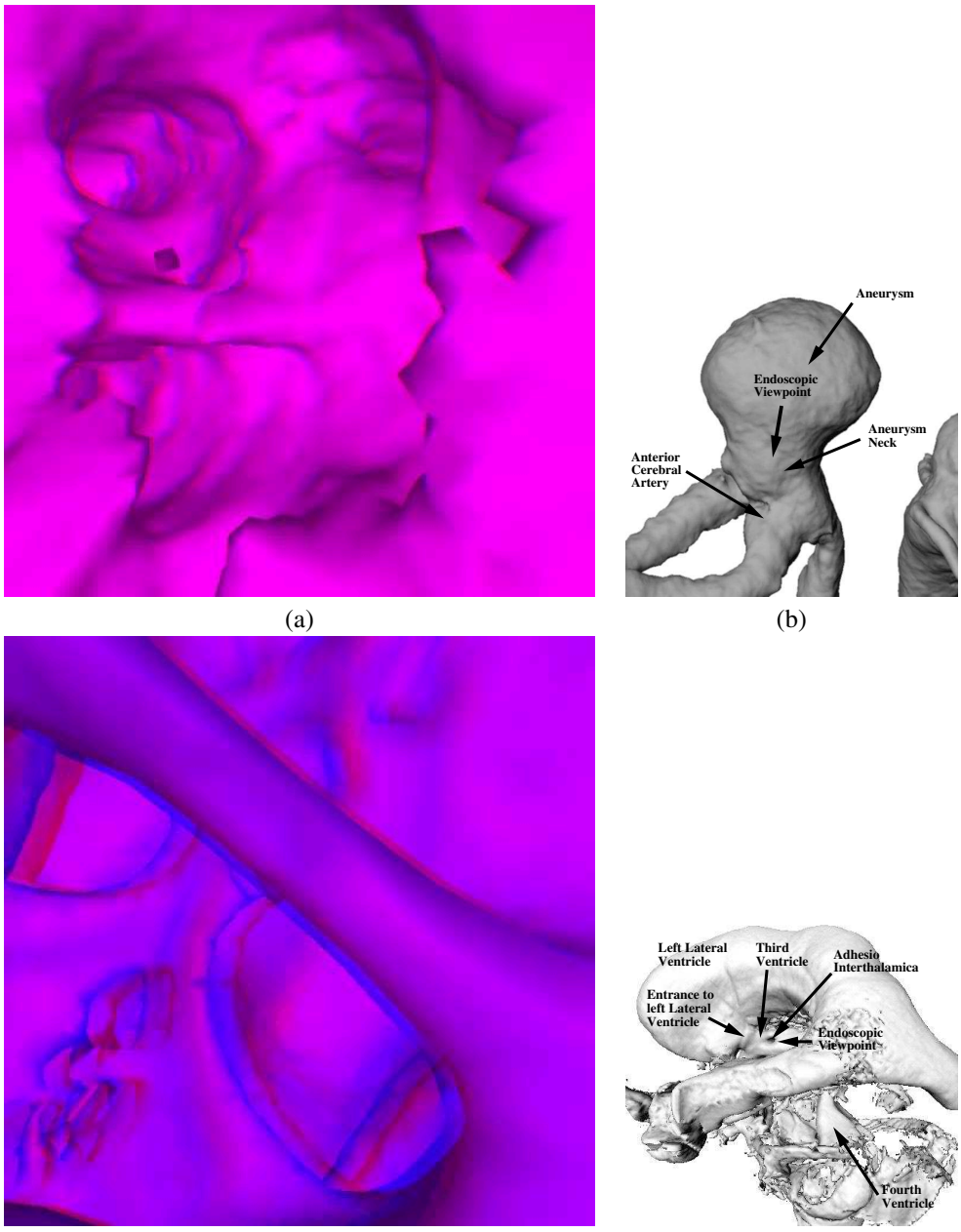
[VSG+94] D. Vining, R. Shifrin, E. Grishaw, K. Liu, and R. Choplin. Virtual Colonoscopy (abstract). In *Radiology*, volume 193(P), page 446, 1994.

[VSH+94] D. Vining, R. Shifrin, E. Haponik, K. Liu, and R. Choplin. Virtual Bronchoscopy (abstract). In *Radiology*, volume 193(P), page 261, 1994.

[WHR99] Z. Wartell, L. Hodges, and W. Ribarsky. Balancing Fusion, Image Depth and Distorsion in Stereoscopic Head-Tracked Displays. In *Proc. of ACM SIGGRAPH*, pages 351–358, 1999.

[WRMW95] J. Wann, S. Rushton, and M. Mon-Williams. Natural Problems for Stereoscopic Depth Perception in Virtual Environments. *Vision Research*, 35(19):2731–2736, 1995.

[YS90] Y. Yeh and L. Silverstein. Limits of Fusion and Depth Judgements in Stereoscopic Color Displays. *Human Factors*, 32(1):45–60, 1990.



(a) (b)

(c) (d)

Figure 3: Stereoscopic snapshots from virtual endoscopy (use red/blue anaglyph glasses; red = left eye, blue = right eye, see also website at: *http://www.gris.uni-tuebingen.de/people/staff/bartz/proj/endos/stereo*): (a) view through the neck of an cerebral aneurysm to three supplying arteries; (b) overview over cerebral aneurysm; (c) view through third cerebral ventricle; the entrances to the lateral ventricles can be seen in the back, while the pipe-like connection of the thalamus through the third ventricle, the adhesio interthalamica, can be seen in the foreground; (d) overview over cerebral ventricle system.