Eurographics / IEEE Symposium on Visualization 2011 (EuroVis 2011)
H. Hauser, H. Pfister, and J. J. van Wijk
(Guest Editors)

Volume 30 (2011), Number 3

Perceptual Evaluation of Ghosted View Techniques for the Exploration of Vascular Structures and Embedded Flow

Alexandra Baer[1], Rocco Gasteiger[1], Douglas Cunningham[2], Bernhard Preim[1]

[1]Department of Simulation and Graphics, University of Magdeburg, Germany
[2]Institute of Graphical Systems, University of Cottbus, Germany

Abstract

This paper presents three controlled perceptual studies investigating the visualization of the cerebral aneurysm anatomy with embedded flow visualization. We evaluate and compare the common semitransparent visualization technique with a ghosted view and a ghosted view with depth enhancement technique. We analyze the techniques' ability to facilitate and support the shape and spatial representation of the aneurysm models as well as evaluating the smart visibility characteristics. The techniques are evaluated with respect to the participants accuracy, response time and their personal preferences. We used as stimuli 3D aneurysm models of five clinical datasets. There was overwhelming preference for the two ghosted view techniques over the semitransparent technique. Since smart visibility techniques are rarely evaluated, this paper may serve as orientation for further studies.

Categories and Subject Descriptors (according to ACM CCS): I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction Techniques; J.3 [Life and Medical Sciences]: General—

1. Introduction

In medical research and treatment planning it is often necessary to visualize multiple superimposed layers of spatial information. Multiple layers may represent different anatomic structures, e.g. an organ as outer layer and its vascular supply as inner layer. In other situations, the anatomy should be displayed at the same time as derived information such as biomechanical simulation [DGB*09] or blood flow simulation [GNK*10]. Since the internal structures are spatially embedded in the surrounding structure, occlusion occurs. Thereby, a trade-off between the visibility of internal information and the simultaneous depiction of the 3D shape of the enclosing surface has to be found. Moreover, to avoid the misleading interpretation of spatial relationships, enhancement of depth is important.

Smart visibility techniques, like cutaway, ghosted, and section views may be used as focus-context illustrations. When exploring blood flow data, streamlines, representing the flow, are focus objects and the vascular surface is the context object. Thus, we employ ghosted views to reveal flow while showing important characteristics of the vascular anatomy, e.g. pathologic variations. There are, however, no clear guidelines to decide which technique should be used in different situations. Smart visibility techniques have rarely been systematically evaluated. The few existing user studies are primarily informal or questionnaire-based, and thus reflect more the personal beliefs and preferences of the participants than actual objective measurement of task performance. This paper presents three controlled perceptual experiments quantitatively evaluating the ghosted view techniques developed by Gasteiger et al. [GNK*10]. The techniques are evaluated in the context of the visualization of cerebral aneurysms, where the vascular anatomy should be displayed along with the internal flow to support the exploration of flow-vessel correlations.

2. Previous and Related Work

2.1. Smart Visibility

The basic strategy of smart visibility techniques is to emphasize the most relevant visual information of an object by means of local modifications of visual attributes or changes in spatial arrangement according to an importance value of the object. Viola and Gröller [VFS*06] did pioneering work on smart visibility techniques, e.g. developing the importance-driven visualization technique. Moreover, they

discussed these techniques and their application for effective anatomical visualizations [VG05]. Further, medical applications can be found in [DWE03].

2.2. Perceptual Experiments

In computer graphics, findings from psychophysical studies are increasingly used to improve virtual and augmented environments and to enhance the effectiveness of 3D and 2D visualizations [BCF*08]. Perceptual experiments have examined the efficiency of depth cues (such as occlusion, silhouette, and transparency in volume rendered images [BBD*07, CWM*09]) that contribute to the mental reconstruction of 3D objects from 2D images. Additional cues, such as textures for nested surfaces [KHSI04, BH07], shape from shading [CF07] or specular reflections [FTA04] were analyzed by means of experimental studies.

Smart Visibility Techniques. So far, no comprehensive and rigorous evaluation of smart visibility techniques has been performed. Some informal studies, however, do exist.Li et al. [LRA*07] demonstrated their system for authoring and viewing interactive cutaway illustrations to medical educators, illustrators, and architects to gather informal feedback. Similarly, Tietjen et al. [TIP05] presented anatomical visualizations to medical experts and analyzed the subjective preference of the participants in a questionnaire-based form. Gasteiger et al. [GNK*10] presented ghosted view visualizations pairwise and asked the participants to decide which they preferred regarding a specific criterion, e.g. depth perception. One of the most rigorous studies to date is from Chan et al. [CWM*09], who asked participants to rank the degree of perceived transparency and quality of the generated images.

These findings are, however, strictly bound to particular applications and it is difficult to generalize from them. There is considerable evidence that performance on such preference or meta-tasks – where one's beliefs and opinions are surveyed – does not always correlate well with actual task performance [WBC*08]. Controlled perceptual experiments can be performed such that the results are more general and enable a quantitative analysis of specific visualization techniques. Baer et al. [BAL*09] presented experimental studies, where participants were asked to examine the focus parts of focus-context illustrations. Accuracy and response time were measured to evaluate different visualization techniques.

Spatial Perception. Studies investigating spatial perception by means of depth ordering or relationships for anatomic surface visualizations were presented by Ritter et al. [RHD*06], Kuß et al. [KGM*10] and Weigle et al. [WB08]. Ritter et al. [RHD*06] analyzed visualization of vascular branches in blood vessel tree illustrations. Participants had to either specify the correct depth order of marks on vascular structures, or to determine the depth distance order between vascular segments. Similarly, Weigle et al. [WB08] presented filamentous structures using various shadowing methods and asked participants to identify which of several overlapping structures was in front. Kuß et al. [KGM*10] presented a complex scene consisting of three volumetric objects and one transparent filamentous structure for a few seconds and then asked participants whether a filament runs through a transparent structure. The different visualization techniques were evaluated with respect to both accuracy and response time.

Shape Perception. Several researchers have examined the 3D shape perception of structures by asking the participants to orient prepositioned gauges to coincide with the perceived surface normal at several positions on the objects' surface [KHSI04, BH07, CSD*09, CRD10]. The normal estimates of the participants were compared with ground truth data provided by a registered 3D surface model to analyze the participants' accuracy. The visualization techniques were evaluated according to the recorded average angle errors in normal estimation. This gauge technique was pioneered by Koendrink et al. [KvDK92]. Even though they tested 3D visualization techniques, in the majority of experiments static images of 3D models are presented. Gauge studies can document not only shape interpretation but also the priors, bias, and information used by the human visual system. Caniard et al. [CF07] studied the direction of illumination and Fleming et al. [FTA04] the specular reflections using the gauge-figure technique.

3. Visualization Techniques

Our perceptual studies focus on three visualization techniques to depict the surface of vascular structures (see Fig. 1). In the exploration of blood flow data, the internal flow is the focus object, since this is the most important visual information and the vascular structure serves as context. Gasteiger et al. [GNK*10] introduced two ghosted view techniques applied to the enclosing aneurysm surface. Both ghosting techniques were developed to depict shape and spatial perception, whilst simultaneously gaining maximum visibility of embedded flow visualization and thus, to improve the previously applied semitransparent visualization technique. Cerebral aneurysms represent a local widening of cerebral vessels, which is a serious disease (high rupture risk). The medical background for cerebral aneurysm diagnosis were discussed by Augsburger et al. [ARF*09].

The data flow pipeline consists of three processes. First, the acquisition step, where clinical image data (MRA, CTA, or 3DRA) of the brain anatomy with the aneurysm is obtained. The overall segmentation takes about 10 minutes. Based on the segmented mask, the surface morphology of the aneurysm is then reconstructed and the mesh is used for constructing a computational grid, on which a CFD (Computational Fluid Dynamic) simulation is performed. The surface mesh and the flow information form the input for the final visualization.

Semitransparency (S). Semitransparent surface rendering for whole enclosing surfaces is a common visualization method. The transparency is often set to 0.5 and is implemented with depth peeling to get a correct blending (see Fig. 1 (a)). As surface color, we use a bright brown which is distinguishable from the color scale of the streamlines.

(a) Semitransparency (b) Ghosting (c) Ghosting with Depth

Figure 1: *Three visualization techniques for displaying blood flow along with the vascular anatomy.*

Ghosting (G). We apply G by means of a view-dependent transparency rendering according to [GNK*10]. We employ an approximation of the Fresnel-reflection model [Sch93] on the front faces of the aneurysm surface and replace "reflection" with "opacity". The front face color is the same as in the S technique and the back face color is a cool color according to Gooch et al. [GGS*98]. A better shape enhancement (due to more opacity) at surface regions facing away from the viewer as well as a maximal visibility (due to less opacity) of the embedded streamlines facing to the viewer (see Fig. 1 (b)) is achieved. Additionally, specular reflections on the front faces are integrated.

Ghosting with Depth Enhancement (GD). The third visualization extends G by means of shadow and atmospheric attenuation added to the surface and the streamlines (see Fig. 1 (c)). By applying the method of Luft et al. [LCD06], we approximate shadow casting in a non-physically correct way. Atmospheric attenuation is introduced by applying fog which makes the objects fade with increasing distance.

All three techniques depict the flow information with color-coded streamlines where the color represents the local velocity. An optimized color scale is used to enhance the quantitative character of the velocity data.

4. Controlled Perceptual Study

We designed and conducted three perceptual experiments to evaluate the three visualization techniques S, G, and GD introduced in Section 3. The central aim is to measure whether and how the techniques G and GD facilitate the assessment process of cerebral aneurysms and the internal blood flow compared to S. Thus, we evaluate the techniques with respect to their ability for:

1. **Perceptually effective shape representation**: Accurate perception of the shape and curvature of the various surfaces is essential for estimating the aneurysm's risk of rupture, and options for treatment planning. Together with the internal flow characteristics, the surface morphology is essential for risk assessment.
2. **Showing embedded flow:** This defines the smart visibility characteristics of the technique and refers to the visual perception of the flow visualization, since current medical research focuses on the simulations of intravascular blood flow [ARF*09]. The additional information supports the decision-making process and disease understanding.
3. **Perceptually effective spatial representation of the aneurysm's parent vessels:** A special subset of scene perception is the understanding of relative location in depth. This is particularly critical in smart visualization techniques. Moreover, understanding the spatial arrangement (depth ordering) of the aneurysm's blood vessel structure helps to improve understanding of the overall flow characteristics (inflow and outflow regions) and this supports aneurysm rupture risk assessment.

4.1. Hypotheses

Since these three criteria are effectively independent of each other, we separated the evaluation and design into three individual experiments. Thus, we defined specific hypotheses (H) to define the parameters that have to be measured.

Shape-H1: G and GD facilitate the vascular structure's shape perception – as measured by **accuracy** – better than S.

Shape-H2: G and GD facilitate the vascular structure's shape perception – as measured by **response time** – better than S.

Smart-H1: G and GD facilitate the assessment of embedded flow – as measured by **accuracy** – better than S.

Smart-H2: G and GD facilitate the assessment of embedded flow – as measured by **response time** – better than S.

Spatial-H1: G and GD facilitate the vascular structure's spatial perception (depth ordering) – as measured by **accuracy** – better than S.

Spatial-H2: G and GD facilitate the vascular structure's spatial perception (depth ordering) – as measured by **response time** – better than S.

Moreover, we postulate **H-GD**: GD will be more effective in each category than G and S. To test these hypotheses, statistical significance tests were run on the results of the experiments (see Section 8).

4.2. Experimental Setup

In order to investigate accuracy and response time, the participants were asked to fulfill different tasks referring to the

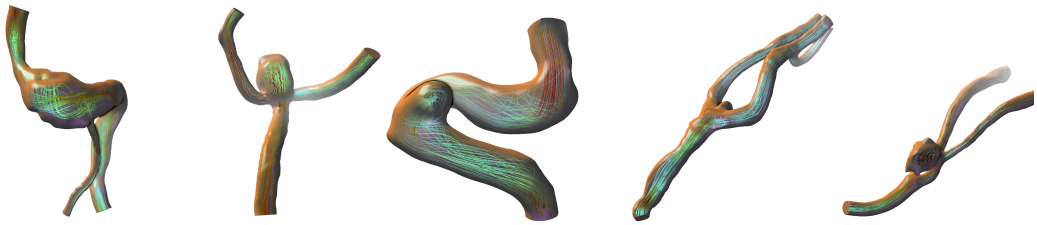


Figure 2: *Our visualizations techniques are evaluated by means of five datasets. These five 3D aneurysm models are visualized with the ghosted view with depth enhancement.*

appropriate experiment type. The stimuli were carefully chosen to be representative renderings of patient-individual 3D aneurysm models. They were visualized with each of the three techniques (see Fig. 2). Thus, the first factor in all three experiments is a *within-participant factor* (visualization technique) which has three *levels* (S, G, and GD). Furthermore, the first experiment (shape perception) had a second factor, which was a *between-participant factor* (interaction): One group saw all stimuli as static models, while the other group could rotate the models within limits.

All participants were tested under the same conditions. They performed the experiment alone by daylight on a $26''$ monitor at a resolution of 1920×1200 pixels. The stimuli were viewed from a distance of approximately 0.7 m (each stimulus image subtended 17.8° of a visual angle on average). Before starting each experimental session, all observers were instructed in written form. One practice trial followed the instruction, to familiarize each participant with the specific task. The practice trials stimuli were randomly chosen (from the aneurysm models and the three visualization techniques) and were not used during the experiments. As soon as they understood the task, approximately after 2-4 minutes, the experiment started. Each experiment consisted of stimuli images being shown one at a time on a white background and the participants were asked to perform a task according to the experiment type (see Sections 5, 6, and 7). Each stimulus was presented until the participants pressed a "Ready" button to indicate that they were satisfied with their answer and ready to move on to the next stimulus. The "Ready" button was used to determine the individual response time for each stimulus. All techniques as well as the aneurysm models were presented in random order to avoid expectation effects. We developed a stand-alone application based on VTK and Qt for the stimuli presentation, handling of the user input, and storage of the user response. At the end, the participants were asked to evaluate the techniques using a bipolar 5 point Likert scale, with each pole representing the performance of one technique compared to the other. This allows us to subjectively compare specific technique properties. Afterwards, a short questionnaire, asking for some personal details, had to be filled out.

4.3. Participants

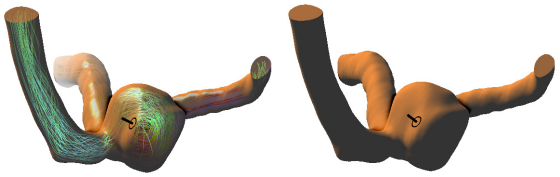
We recruited participants from various parts of the university, each participating in only one experiment. The exact number of the participants for each experiment can be found in the corresponding experiment sections. Although it is recommended to recruit prospective users, in our case medical doctors, participants from the general population can also provide useful insights. Moreover, it is likely that the measured results can be applied to prospective users concerning the perceptual effectiveness, even though medical doctors may achieve better accuracy, because of their clinical experience. Regardless of the difference in clinical expertise, medical doctors and the current participants would both have to become familiar with any new visualization techniques.

5. Shape Perception Experiment

In choosing a task, it was crucial that it allows us to measure both reaction time as well as accuracy. We employed the gauge figure technique from visual psychophysics to obtain local estimates of surface orientation and thus, analyze the perceived surface shape [KvDK92].

Gauge Methodology. Participants were shown 3D aneurysm models with a gauge figure placed on the front side of the model's surface. Each gauge was drawn as a small ellipse representing a disc and a single line indicating the normal of the disc. Participants were asked to orient the gauges to coincide with the apparent surface normal at that specific surface point. They had no control over gauge positions and each gauge initially pointed to the viewer. The orientation of the gauge was controlled by the mouse. To avoid cueing the participants to shape, the gauges did not penetrate or interact with the 3D model.

Stimuli. For the stimuli images, we choose viewpoints corresponding to preferred views of neuroradiologists. Each stimulus consists of a single gauge placed on a single surface model, visualized with either G, GD or S (see Fig. 3). Moreover, to effectively increase the number of our models, we are changing viewpoints, which can help to reduce the ability to recognize an object as being identical to one already seen. These changed viewpoints are still similar to preferred views of neuroradiologists. Therefore, we had nine instead



(a) (b)

Figure 3: *(a) The ghosting with depth enhancement (GD) and (b) the full-color shaded opaque surface. The gauge is a small ellipse representing a disc and a single line.*

of the original five models each presented in the three styles with a single gauge for 27 stimuli. Since we are aiming for perceptual validity, not physical validity, we had to define a perceptual "gold standard" for the gauge figures. To determine this "gold standard", we used for each presented model a full-color rendered stimulus shown at the end of the experiment (see Fig. 3 (b)). Thus, we had nine more stimuli and a total of 36 stimuli. Additionally, we carefully chose the gauge figure positions on the surface. Ghosting techniques are characterized by three different surface areas:

- opaque surface areas (see Fig. 3 (a)),
- semitransparent surface areas that occur between transparency and opaque regions, and
- areas, primarily in the focal view of the user, where the surface is completely transparent.

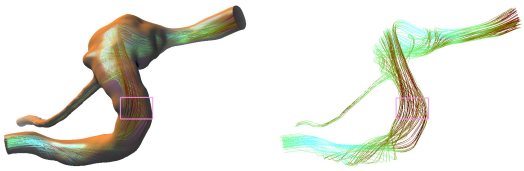
To enable an analysis of each technique and each characteristic region, we presented three of these nine models with gauge figures placed in opaque regions, three in semitransparent, and three models with a gauge placed in fully transparent regions.

Between-Participant Design. For the *between-participant* factor of interaction, here were two equal-sized groups of 17 participants each. Six women and eleven men participated in the first and ten women and seven men in the second experiment, all aged between 20 and 35. While the gauge interaction was the same for both groups, one group received static 3D renderings of the aneurysm models (group -RO) and one had the opportunity to rotate the scene (group +RO). A pilot study with scene rotation showed that participants used the rotation to orientate the model so that the gauge figure was on the model's silhouette (thus at a 90° angle to the viewer and vastly simplifying the task). Since we want to evaluate the perception for the given view (which is the preferred view for specialists) and not the perfect surface normal vector, we decided to restrict the range of possible rotations to 15°.

6. Smart Visibility Experiment

This experiment is conducted to assesses the techniques' ability and effectiveness to show embedded structures, in our case the blood flow. We refer to this as the techniques' smart visibility characteristics.

Visibility Methodology. To validate the techniques' smart visibility characteristics, we use color-coded streamlines that represent the blood flow and quantitatively evaluate the embedded flow perception. We do not evaluate the flow visualization. Participants were asked to define the average flow color at different regions (see Fig. 4) using the CColorDialog of the Microsoft Foundation Class. These individual color results are used to determine the average color error compared to a perceptual color "gold standard", which is represented by the flow at that specific region (see Fig. 4 (b)). This experiment follows a *within-participants* design, where each of the participants is given the same kind of stimuli.



(a) (b)

Figure 4: *A stimulus with (a) the ghosted view with depth enhancement (GD) and (b) just the flow visualization. The pink rectangle defines the region for which the participant had to estimate the average flow color.*

Stimuli. As illustrated in Figure 4, the stimuli are static renderings of our 3D aneurysm models. Each stimulus is overlaid with a pink rectangle selecting a certain surface and the corresponding flow region. We chose the semitransparent and fully transparent surface regions according to the defined regions in Section 5. Since, by definition, an opaque surface will fully occlude the embedded structure, it will not be possible to determine the flow color in opaque regions. Obviously, then, the S technique is superior to the G and GD technique in this regard: The S technique allows you to see the flow everywhere. Since, however, the visualizations were designed with a specific region of interest in mind (the region that is transparent in the ghosting techniques), it is also logical to focus the experiments on the perception of flow color in those areas. Thus, only the semitransparent and fully transparent regions are to be investigated.

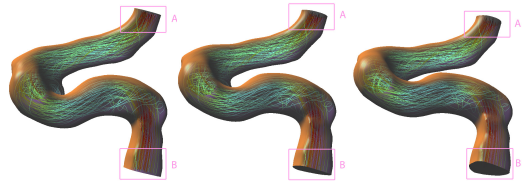
We used eight models, whereas four models were overlaid with a rectangle positioned in transparent and four models with a rectangle positioned in semitransparent regions. Each model was visualized with either the S, G or GD technique, which accounts for 24 stimuli. Since we are aiming for perceptual validity, we had to determine a perceptual color "gold standard" for each region, too. Therefore, we used stimuli images illustrating only the flow visualization on a white background (see Figure 4 (b)). Those eight stimuli were presented at the end of the experiment, so that each

participant saw a total of 32 stimuli. We recruited 27 participants for this experiment. Twelve women and fifteen men aged between 22 and 36 participated in the study. The pilot study results showed that the participants chose the color – especially for the S technique – referring to their expected color and not to the seen color. Thus, the participants were asked to talk while adjusting and to verbally explain their resulting color during the experiment, too.

7. Spatial Relation Experiment

A common strategy to analyze spatial perception is to ask the participants to determine the perceived depth ordering (see, e.g. Weigle et al. [WB08] and Ritter et al. [RHD*06]). We employ this depth-judgment strategy and use the vessel branches of the aneurysm models to investigate the techniques' accuracy according to the hypothesis Spatial-H1.

Depth Methodology. Similar to the previous experiment, this study follows a *within-participant* design with 25 participants aged between 18 and 30 (fifteen women and ten men). For each stimulus, the participants had to determine which branch is closer to the viewer. Possible answers were branch *A*, branch *B* or *None*, if both branches have the same distance to the viewer.



(a) (b) (c)

Figure 5: *Spatial stimuli for one model and one technique. (a) Both vessel branches have the same distance to the viewer. (b) The model rotated by 10 degrees and (c) 20 degrees, whereas the vessel branch B is closer to the viewer.*

Stimuli. We generated stimuli images illustrating the aneurysm models with two branches vertically aligned. Furthermore, the models were rotated around the aneurysms' x-axis clock- or counterclockwise by 0, 10 or 20 degrees (see Fig. 5) to provide depth ordering of the branches. The specific vessel branches were marked with pink rectangles labeled with *A* and *B* to focus the participants on these vessel regions. We left out stimuli images of the fourth model illustrated in Figure 2, since the two branches were too close to separate them. Each participant saw four models, each rotated with three different angles and visualized with three different rendering styles, a total of 36 stimuli.

8. Analysis and Results

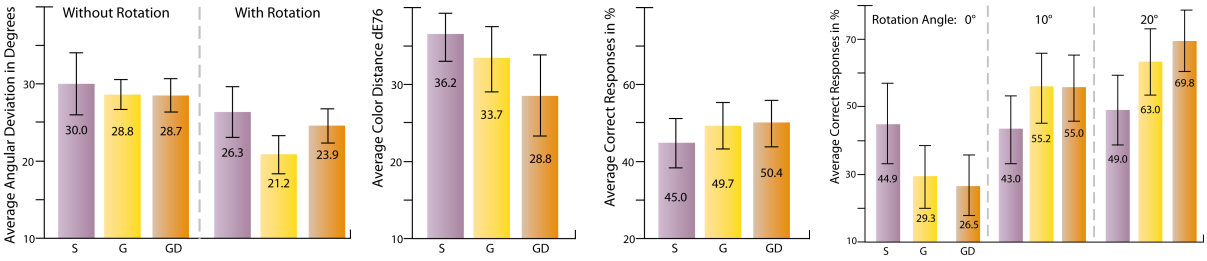
We analyzed the accuracy, the response times, and the subjective technique evaluation. Accuracy for shape perception was determined by analyzing the perceived surface normal vectors, for smart visibility by analyzing the perceived blood flow color, and for the spatial experiment by determining the correct and false answers of the depth-judgment task.

Initially, we tested all data for a normal distribution with the Shapiro Wilk test, since this is a major requirement for choosing an appropriate significance test and consequently achieving valid results. According to these results, we applied standard statistical methods to analyze significances for each factor level. We applied the non-parametric Friedman combined with the Wilcoxon signed rank test for the pairwise comparison, since our results are not normally distributed. Moreover for the shape experiment, we used the Wilcoxon-Mann-Whitney test to compare group -RO and group +RO, since this non-parametric test assesses whether there is a significant difference between two independent samples.

8.1. Shape Analysis

We stored the xyz-components of the apparent surface normal vector and the response time for each participant and each gauge figure. Accuracy was analyzed by comparing the surface normal estimate with the perceptual "gold standard" for the corresponding gauge. We determined these "gold standard" vectors (gs-vectors) using the full-color shaded stimuli results, since we expect the participants perform most accurately on shaded models. All parameters and the knowledge of the shown stimuli enable a quantitative analysis for each group.

Accuracy. The average angle deviation of the normal estimates compared to the gs-vectors, for group -RO was between 28.7° and 30° and for group +RO between 21.2° and 26.3° (see Fig. 6 (a)). Participants of group -RO were more precise with GD and of group +RO with G. The Wilcoxon-Mann-Whitney test confirmed significantly ($p \leq 0.05$) more precise results for group +RO. In detail, no significant difference ($p > 0.05$, $\chi^2(2) = 0.118$) was found between any two visualization techniques for group -RO. In contrast, participants of group +RO were more precise with G ($p \leq 0.05$, $z = 1.73$) and GD ($p \leq 0.05$, $z = 1.82$) than with S. Thus, Shape-H1 can be confirmed. Moreover, G ($p \leq 0.05$, $z = 2.07$) enables significantly more accurate shape perception than GD. Additionally, we analyzed the participants' accuracy referring to the three different surface regions (opaque, semitransparent, transparent) for group +RO (see Table 1, right hand side). Compared to S, G enables a significantly more accurate shape perception within semitransparent regions ($p \leq 0.05$, $z = 1.71$) and GD within opaque ($p \leq 0.05$, $z = 2.02$) regions. Furthermore, a significant difference exists between G & GD for semitransparent regions ($p \leq 0.05$), whereas the results for G are more accurate. The results show that as long as the participants could rotate the models, Shape-H1 is confirmed with the exception of transparent regions. In contrast,



(a) Shape (b) Smart Visibility (c) Spatial (d) Specific Angles of Spatial Study

Figure 6: *The average results and error bars (95% confidence intervals) for (a) the shape experiment (group -RO and group +RO), (b) the color distance of the smart visibility experiment, (c) all correct responses for the spatial experiment, and (d) separated for each rotation angle.*

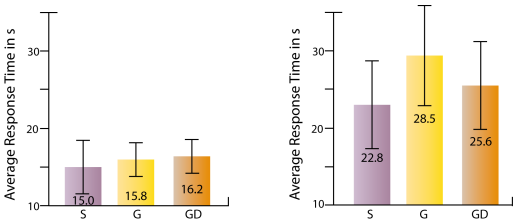
when the participants could not rotate the model, no significant difference was found and Shape-H1 has to be rejected.

Pair	group +RO	Opaque	Semi	Transp
S&**G**	$p \leq 0.05$ $z = 1.73$	$p > 0.05$ $z = 1.09$	$p \leq 0.05$ $z = 1.71$	$p > 0.05$ $z = 0.99$
S&**GD**	$p \leq 0.05$ $z = 1.82$	$p \leq 0.05$ $z = 2.02$	$p > 0.05$ $z = 0.40$	$p > 0.05$ $z = 0.60$
G&*GD*	$p \leq 0.05$ $z = 2.07$	$p > 0.05$ $z = 0.52$	$p \leq 0.05$ $z = 1.81$	$p > 0.05$ $z = 0.21$

Table 1: *Accuracy results for the pairwise technique comparisons and the specific region results for group +RO. Any difference between two techniques is considered significant if $p \leq 0.05$ (probability of error) and $z > 1.64$ (z-score of the Wilcoxon signed-rank test). The significantly more precise technique is respectively emphasized.*

Response Time. The average response times for group -RO are 15*s* for S, 15.8*s* for G, and 16.2*s* for GD. Group +RO required 22.8*s* for S, 28.5*s* for G and 25.6*s* for GD (see Fig. 7). Participants of group +RO required more time orienting the gauges but had smaller angular errors. Both groups performed significantly faster with S than with G (group -RO with $p \leq 0.05$, $z = 1.66$ and group +RO with $p \leq 0.05$, $z = 2.82$), even though participants performed more accurate with the ghosting techniques. Furthermore, group -RO achieved significantly shorter response times for S compared to GD ($p \leq 0.05$, $z = 1.68$) but no significant difference between G and GD. In contrast, for group +RO exists no statistical difference between S and GD ($p = 0.054$, $z = 1.63$) but GD is significantly faster than G ($p \leq 0.05$, $z = 2,87$). Nevertheless, we had to reject Shape-H2, since the participants were faster with S, even though GD achieved better results than G.

The Wilcoxon-Mann-Whitney test compared the techniques of both groups pairwise. The results of each technique of group -RO compared with the results of the same technique in group +RO are significantly different ($p \leq 0.05$). In summary, group +RO were significantly more accurate but slower than group -RO.



(a) Group -RO (b) Group +RO

Figure 7: *Average response times and error bars (95% confidence intervals) of the shape experiment.*

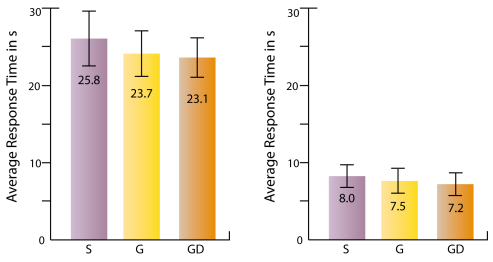
8.2. Smart Visibility Analysis

The accuracy of flow perception indicates the smart visibility characteristics of the illustration techniques. All three techniques were evaluated by the perceptual difference between the estimated color and the perceptual "gold standard" color (gs-color). The gs-color is determined of the stimuli images illustrating only the blood flow (see Fig. 4 (b)), similar to the gs-vector for the gauges. Additionally, the gs-color describes the average perceived color error compared to the histogram results for each region. Representing the participants' color choices in L*a*b*-space allows to compute color differences as Euclidean distances (ΔE), since this color space is perceptually linearized. A difference of 1.0 ΔE means that the color difference between two colors is perceptual distinguishable.

Accuracy. Participants' estimated colors more precise, in terms of smaller distances ΔE to the gs-colors, for G and GD (see Fig. 6 (b)). In contrast to the average results of S (36.2 ΔE), especially, GD enabled smaller distances (28 ΔE). The participants performed significantly better with

GD ($p \leq 0.001$, $z = 3.06$) than with S and with G ($p \leq 0.01$, $z = 2.78$). No significant difference ($p = 0.08$, $z = 1.37$) was confirmed for S & G. One reason might be that the participants defined for S the flow color, which they expected to be the right one and not the color they really saw. Since the participants were asked to explain their adjusted colors during the experiment, they confirmed that assumption. This happened, especially after they saw the first G stimulus. This might be eliminated by designing a *between-subject* design for this experiment, too. Therefore, we have to reject Smart-H1 but we are able to confirm that GD enables a more accurate assessment of embedded flow than S. Nevertheless, the majority of the participants preferred both ghosting techniques over the S technique (see Fig. 9 (a) and (b)).

Response Time. The average response times are very close to each other (see Fig. 8 (a)). A just significant difference ($p \leq 0.05$, $z = 1.65$) exists, however, between the S & GD, whereas GD achieved faster results. No significant difference exists between G and GD ($p > 0.05$, $z = 0.44$) and between S and G ($p > 0.05$, $z = 1.37$). Even though, G is better than S and a significant difference may occur with further results, Smart-H2 has to be rejected, too. We are only able to confirm that GD accelerates the assessment of embedded flow compared to S.



(a) (b)

Figure 8: *Average (a) smart visibility and (b) spatial response times and error bars (95% confidence intervals).*

8.3. Spatial Analysis

The knowledge of each stimulus and the recorded answers enable the analysis of correct and false responses with respect to the depth order. Additionally, we evaluate S, G, and GD referring to the minimal perceived rotation angle.

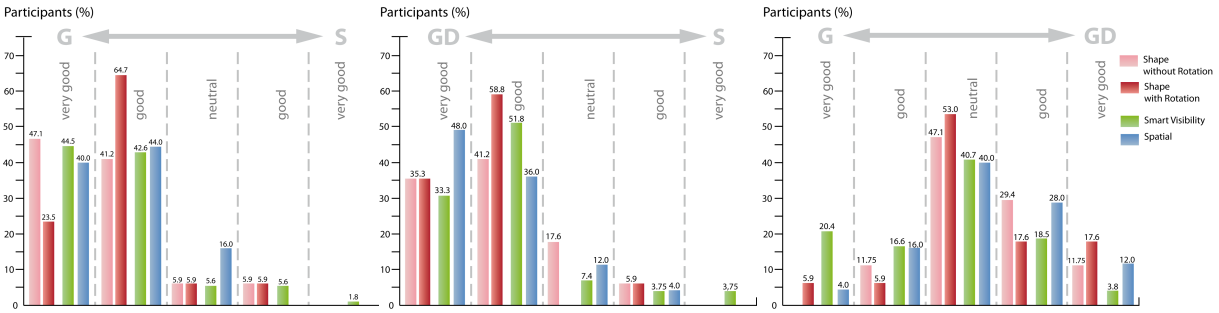
Accuracy. A statistically significant correlation exists between the techniques and the number of correct responses. 45% correct responses were achieved for S, 49.7% for G, and 50.4% for GD (see Fig. 6 (c)). The paired test confirmed significantly more correct answers for G ($p \leq 0.05$, $z = 1.86$) and for GD ($p \leq 0.05$, $z = 2.76$) compared to S. Both ghosting techniques exhibited no significant difference. We can, however, confirm Spatial-H1.

Furthermore, we analyzed the correctness of the depth order individually for 0°, 10°, and 20° (recall Fig. 5). Figure 6 (d) depicts the average correct responses for each technique and each angle. Stimuli with both vessel branches at the same distance (0°) to the participant were correctly perceived more often with S (44.9%) than with G (29.3%) or with GD (26.5%). Thus, a significant difference exists for S & G ($p \leq 0.05$, $z = 1.86$) and for S & GD ($p \leq 0.01$, $z = 2.76$). The ghosting techniques showed no significant difference ($p > 0.05$, $z = 0.67$). This difference may, however, reflect a bias towards a response of "no separation in depth" for S. In such a case, lower performance for the two larger rotations is to be expected, which is precisely what we find. In contrast, stimuli showing a rotation angle of 10° and 20° were correctly perceived more often with G and GD than with S (Fig. 6 (d)). Significant differences exist between S & G (for 10°: $p \leq 0.05$, $z = 1.65$ and for 20°: $p \leq 0.01$, $z = 2.59$) and between S & GD (for 10°: $p \leq 0.01$, $z = 2.55$ and for 20°: $p \leq 0.001$, $z = 3.35$). The larger the rotation angle, the higher the significant difference and the larger the difference of correct responses between G (63%) and GD (69.8%). Even though no significant difference was confirmed between G and GD, Spatial-H2 can be confirmed for 10° and 20°.

Response Time. The average response times are between 7.2 *s* for GD and 8 *s* for S (see Fig. 8 (b)). In contrast to the other experiments, the participants were faster selecting the closest vessel branch than orienting gauges or defining colors. Overall, response times in this experiment showed no significant effects and we have to reject Spatial-H2. G and GD do not accelerate the spatial perception.

8.4. Questionnaire Results

Our questionnaire analysis measures participants' attitude towards a technique. The results of each pairwise technique comparison for each experiment (colored bars) are illustrated in Figure 9. We had a total of 86 participants. As illustrated in Figure 9 (a), 34 participants (39.5%) rated G compared with S with *very good* and 41 (47.7%) with *good*. Thus, 75 participants (87.2%) preferred G over S. The comparison of GD and S (see Fig. 9 (b)) showed that 73 participants (84.9%) rated GD as *very good* or *good*, and thus, preferred this technique over S, too. Finally, the participants should compare G and GD. Figure 9 (c) depicts a small preference for GD. 19 participants (22.1%) rated G with either *very good* or *good*. In contrast, 29 participants (33.7%) preferred GD and chose *very good* or *good*, and 38 participants (44.2%) liked both kinds of visualization techniques and chose *neutral*. Thus, the majority of all participants preferred G and GD over S. Additionally, the attitude towards GD was revealed.



(a) G vs. S (b) GD vs. S (c) G vs. GD

Figure 9: *Participants were asked to rate and compare the techniques pairwise.*

9. Discussion

The shape experiment revealed, that G and GD enable a more accurate shape perception as long as the participants had the opportunity to rotate the models, group -RO showed no significant difference. This result shows the importance of an appropriate experimental design. The achieved results for group -RO are not expressive enough. Since G and GD are developed for interactive 3D visualizations, they have to be evaluated in 3D, too. The estimated surface normals of group +RO had smaller angular errors than of group -RO, even though they required more time. As a standard speed-accuracy trade-off the participants required more time for both ghosting visualizations, since they perceived the surface shape better, and thus, aimed at orienting the gauge as close as possible to the gs-vector. Even the specific surface regions revealed more accurate perception using G and GD for opaque and semitransparent regions. In general, we can confirm Shape-H1 for group +RO and reject it for group -RO as well as Shape-H2 for both groups. Ghosting techniques do not accelerate but enable more precise shape perception. Moreover, we have to reject Smart-H1 but we are able to confirm that GD enables a more accurate assessment of embedded flow than S and G. So far, Smart-H2 has to be rejected but G achieved faster responses than S and this difference is close to a significant difference. Further results may confirm this assumption. Participants are, however, faster with G and GD than with S. The spatial experiment revealed that G and GD facilitate depth-judgement compared to S and thus, Spatial-H1 is true. The analysis showed that depth-judgement especially benefits from G and GD for 10° and 20°, even though both techniques do not accelerate this process and Spatial-H2 is rejected. The hypothesis H has to be rejected, due to the individual results.

Regardless of significant differences and confirmed or rejected hypotheses, the results showed that G and GD enable more accurate results than S. Furthermore, G and GD accelerated the individual assessment task, except for the shape experiment. In this experiment more accurate shape perception is maintained at the expenses of response time. This general result is supported by the qualitative analysis (see Fig. 9 (a) and (b)). We found overwhelming preference for the two ghosting techniques over S. There was also a small trend towards a preference of GD over the simple G.

10. Conclusion

This paper presented three controlled, task-based experiments investigating the visualization of vascular anatomy with embedded flow. Two smart visibility techniques and a common semitransparent visualization technique were compared with respect to shape and spatial representation of the aneurysm models as well as the smart visibility characteristics of the ghosting techniques. Five 3D aneurysm models generated from clinical datasets were used. We presented the individual technique analysis and confirmed significant differences for various tasks based on the participants' accuracy and response time. The quantitative analysis reveal the advantage of both ghosting techniques and clearly show that both techniques support more accurate analysis of aneurysms than the traditional S technique. Additionally, we qualitatively analyzed the participants' preferences for the three techniques and found overwhelming preference for the two ghosting techniques. For further information, we refer to our project page www.vismd.de/Ghosting-Study2011. Although, we selected specific examples from cerebral vasculature, the results are likely to be generalizable to other kinds of vasculature, e.g. coronary arteries, where blood flow simulation is also essential. The results of these experiments, however, should not be overgeneralized. The fact that these results may be specific to the stimuli used here is highlighted by the differences found between the different models in the third experiment. It is unclear, how the results generalize to other, complex anatomic or botanic shapes.

Acknowledgments: We thank Dr. G. Janiga and S. Seshadhri (ISUT Magdeburg) for providing the simulation data and M. Scherbinsky. This work has been funded by the Land Saxony-Anhalt Project MOBESTAN (No: 5161AD/0308M) in Germany.

References

[ARF*09] AUGSBURGER L., REYMOND P., FONCK E., KULCSAR Z., FARHAT M., OHTA M., STERGIOPULOS N., RÜFENACHT D.: Methodologies to assess blood flow in cerebral aneurysms: Current state of research and perspectives. *Journal of Neuroradiology 36*, 5 (2009), 270. 2, 3

[BAL*09] BAER A., ADLER F., LENZ D., ET AL.: Perception-based evaluation of emphasis techniques used in 3d medical visualization. In *VMV* (2009), pp. 295–304. 2

[BBD*07] BOUCHENY C., BONNEAU G.-P., DROULEZ J., ET AL.: A perceptive evaluation of volume rendering techniques. *ACM Trans. on Appl. Percept. 5*, 4 (2007), 1–24. 2

[BCF*08] BARTZ D., CUNNINGHAM D., FISCHER J., ET AL.: The role of perception for computer graphics. In *Eurographics State-of-the-Art Report 4* (2008). 2

[BH07] BAIR A., HOUSE D.: Grid with a view: Optimal texturing for perception of layered surface shape. *IEEE TVCG 13*, 6 (2007), 1656–1663. 2

[CF07] CANIARD F., FLEMING R. W.: Distortion in 3d shape estimation with changes in illumination. In *APGV '07* (2007), ACM, pp. 99–105. 2

[CRD10] CORCORAN A., REDMOND N., DINGLIANA J.: Perceptual enhancement of two-level volume rendering. *Computer&Graphics 34*, 4 (2010), 432 – 438. 2

[CSD*09] COLE F., SANIK K., DECARLO D., ET AL.: How well do line drawings depict shape? *ACM SIGGRAPH 28*, 3 (2009), 1–9. 2

[CWM*09] CHAN M.-Y., WU Y., MAK W.-H., ET AL.: Perception-based transparency optimization for direct volume rendering. *TVCG 15* (2009), 1283–1290. 2

[DGB*09] DICK C., GEORGII J., BURGKART R., ET AL.: Stress tensor field visualization for implant planning in orthopedics. *IEEE TVCG 15*, 6 (2009), 1399–1406. 1

[DWE03] DIEPSTRATEN J., WEISKOPF D., ERTL T.: Interactive cutaway illustrations. *Comp. Graph. Forum 22*, 3 (2003), 523–532. 2

[FTA04] FLEMING R. W., TORRALBA A., ADELSON E. H.: Specular reflections and the perception of shape. *Journal of Vision 4*, 9 (2004), 798–820. 2

[GGS*98] GOOCH A., GOOCH B., SHIRLEY P., ET AL.: A non-photorealistic lighting model for automatic technical illustration. In *ACM SIGGRAPH* (1998), pp. 447–452. 3

[GNK*10] GASTEIGER R., NEUGEBAUER M., KUBISCH C., ET AL.: Adapted surface visualization of cerebral aneurysms with embedded blood flow information. In *EG Workshop on VCBM* (2010), pp. 25–32. 1, 2, 3

[KGM*10] KUSS A., GENSEL M., MEYER B., ET AL.: Effective techniques to visualize filament-surface relationships. *Comp. Graph. Forum 29* (2010), 1003–1012. 2

[KHSI04] KIM S., HAGH-SHENAS H., INTERRANTE V.: Conveying shape with texture: Experimental investigations of texture's effects on shape categorization judgments. *IEEE TVCG 10*, 4 (2004), 471–483. 2

[KvDK92] KOENDRINK J. J., VAN DOORN A. J., KAPPERS A. M. L.: Surface perception in pictures. *Perception & Psychophysics 52*, 5 (1992), 487–496. 2, 4

[LCD06] LUFT T., COLDITZ C., DEUSSEN O.: Image enhancement by unsharp masking the depth buffer. In *ACM Trans. on Graphics* (2006), vol. 25, pp. 1206–1213. 3

[LRA*07] LI W., RITTER L., AGRAWALA M., ET AL.: Interactive cutaway illustrations of complex 3d models. In *ACM SIGGRAPH* (2007), pp. 31–1–31–10. 2

[RHD*06] RITTER F., HANSEN C., DICKEN V., ET AL.: Real-time illustration of vascular structures. *IEEE TVCG 12*, 5 (2006), 877–884. 2, 6

[Sch93] SCHLICK C.: A customizable reflectance model for everyday rendering. In *Fourth Eurographics Workshop on Rendering* (1993), pp. 73–83. 3

[TIP05] TIETJEN C., ISENBERG T., PREIM B.: Combining silhouettes, surface, and volume rendering for surgery education and planning. In *EUROVIS 2005* (2005), pp. 303–310. 2

[VFS*06] VIOLA I., FEIXAS M., SBERT M., ET AL.: Importance-driven focus of attention. In *IEEE TVCG* (2006), pp. 933–940. 1

[VG05] VIOLA I., GRÖLLER E.: Smart visibility in visualization. In *Proc. of EG Workshop on Computational Aesthetics in Graphics, Visualization and Imaging* (2005), pp. 87–94. 2

[WB08] WEIGLE C., BANKS D. C.: A comparison of the perceptual benefits of linear perspective and physically-based illumination for display of dense 3d streamtubes. *IEEE TVCG 14*, 6 (2008), 1723–1730. 2, 6

[WBC*08] WALLRAVEN C., BREIDT M., CUNNINGHAM D. W., ET AL.: Evaluating the perceptual realism of animated facial expressions. *ACM Trans. Appl. Percept. 4*, 4 (2008), 1–20. 2