This is an Accepted Manuscript of an article published by Taylor & Francis in *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization* 10(4), p. 357-365, available at: https://doi.org/10.1080/21681163.2021.1999332

ARTICLE

The Effect of Interactive Cues on the Perception of Angiographic Volumes in Virtual Reality

Andrey Titov[ab], Marta Kersten-Oertel[b] and Simon Drouin[a]

[a]Multimedia Lab, Software and Information Technology Engineering , École de technologie supérieure, Montreal, Canada;
[b]Applied Perception Lab, Gina Cody School of Computer Science and Engineering, Concordia University, Montreal, Canada

ARTICLE HISTORY
Compiled April 19, 2023

ABSTRACT
In this paper, we evaluate the effect of depth cues on the perception of three-dimensional cerebral angiographic data in a virtual reality (VR) environment. Specifically, a user study was conducted to evaluate the effectiveness of shading, pseudo-chromadepth and aerial perspective, both with and without a dynamic component, where the volume rendering parameters are modified based on the motion of the VR controllers. The results of the study showed that the type of cue that is used has little impact on decision time or relative depth perception, contrary to what has been previously observed in related works using 2D displays. However, shading resulted in better spatial understanding of local vascular structures. In terms of the effect of the dynamic cues, although they resulted in worse depth perception, they also resulted in less head movement which may provide a more ergonomic and intuitive solution for data exploration. This work is a first step towards gaining a better understanding of the interplay between interactive perceptual rendering and the impact it has on spatial understanding of volume rendered medical data within VR contexts.

KEYWORDS
Virtual reality; depth cues; angiography; volume rendering; interaction techniques

1. Introduction

Interaction is an important aspect in medical data exploration and one that has not been sufficiently studied despite the complex nature of medical scans and the limitations of 2D displays (McGuffin et al. 2003; Drouin et al. 2018). It has also been shown that recognizing objects from different angles is easier if the angle change results from self-movement (Simons et al. 2002) or active user input (Harman et al. 1999) rather than from passive object movement (e.g. a rotating scan). Thus virtual reality (VR) environments where the user can explore complex data by actively interacting with it offer a promising solution for medical data exploration.

Three-dimensional cerebral angiographic scans, which are used for diagnosis, neurosurgical planning and intraoperative guidance (Abhari et al. 2012), are particularly complex due to the intricate branching and frequent overlapping between vessels. This

CONTACT Andrey Titov. Email: andrey.titov.1@ens.etsmtl.ca

makes these types of data scans particularly difficult to visualize and understand spatially (Kersten-Oertel et al. 2014). Researchers have studied the use of various perceptual cues, illustrative methods and interaction techniques (Ropinski et al. 2006; Preim et al. 2016) to improve depth perception and spatial understanding of 3D vascular data. However, most of this work has targeted classic 2D monitors for displays coupled with keyboard and mouse for interaction. This setup suffers from a limited capacity to represent visual information and a lack of degrees of freedom for interaction and thus may not be optimal for exploring angiographic datasets (Heinrich et al. 2020).

With the recent advancements in VR hardware and controllers, new human-computer interaction paradigms have become possible. Modern VR headsets allow for *binocular disparity* (slight differences between the right and left eye images that allows an observer to see 3D) and *motion parallax* (a change in position of an object caused by the movement of an observer), which provide strong cues for depth perception (Vienne et al. 2020). A recent study by Heinrich et al. showed that VR significantly improves perception of depth order with angiographic data (2021). Additionally, efforts are being made to develop varifocal VR displays to improve depth perception even further (Toulouse et al. 2019). Moreover, the input devices that can be used in VR such as handheld controllers or hand tracking devices allow more natural gestures and significantly more degrees of freedom (Laha and Bowman 2013; Laha et al. 2016) enabling more intuitive and natural exploration of 3D datasets.

In this paper, we revisit several of the cues previously studied for the visualization of angiographic scans using VR in the context of pre-operative planning, simulation and in general for anatomical exploration. It is important in VR that the surgeon uses the best performing depth visualization techniques to improve their perception (Heinrich et al. 2021). In particular, we aim at determining whether the effectiveness of the selected depth cues is affected by the increased expressivity provided by VR devices. Building on the work of Drouin et al. (2020) we aimed to determine the impact of dynamic and interactive depth cues within a VR context. Thus, this work addresses the need for a better understanding of the interplay between interactive perceptual rendering and the impact it has on spatial understanding of volume rendered medical data within VR contexts.

2. Related Work

Various perceptual visualization methods have been developed with the goal of improving the perception of angiographic data. The effectiveness of such techniques was extensively studied previously by performing user studies. In task of these user studies is often to determine the depth relationship of specific vessels, and performance is typically measured using correctness and decision time. These studies have been typically done using a 2D monitor, and to the best of our knowledge, only one such study was done so far in a VR environment with a head mounted display (HMD).

Ropinski et al. (2006) performed a user study where they compared the effectiveness in depth perception of previously studied techniques such as edge ehnancement, Phong shading, depth of field (DoF), stereoscopy, chromadepth, and pseudo-chromadepth. It was determined that pseudo-chromadepth performs better than full chroma, and is the best shading technique in terms of decision time. However, in terms of correctness, the overlaid edges technique was found to be better than pseudo-chromadepth. In a similar study, Kersten-Oertel et al. (2014) compared the effectiveness of kinetic depth, stereoscopy, edge, pseudo-chromadepth and fog for novices and experts. It was determined that pseudo-chromadpeth generally resulted in the best decision time and correctness, similar to the results obtained by Ropinski et al. (2006). It was also found that the edge cue was more helpful to experts than novices. Kreiser et al. (2018) introduced the concept of Void Space Surfaces (VSS) that encode the depth of vessels in the scene using the surrounding backround. They then compared the chromadepth and pseudo-chroma versions of VSS to directly applied versions of these cues (to the surface of the vessels) using standard Phong shading as the base case. Although directly applied cues and VSS resulted in similar correctness, they both performed better than Phong. In terms of decision time, direct cues performed better than Phong, while VSS performed worse. The results of this study regarding the effectiveness of pseudo-chromadepth slightly differ from those obtained by Ropinski et al. (2006) and Kersten-Oertel et al. (2014) in that pseudo-chromadepth was not found to be more effective than chromadepth.

Dynamic versions of these cues were studied by Drouin et al. (2020). They introduced a dynamic aspect to Phong shading, pseudo-chromadepth and fog where the user could use a surgical pointer to control certain rendering parameters. The new dynamic versions of cues resulted in a better understanding of the local structures but not global ones, when compared to their static counterparts. Additionally, it was found that dynamic pseudo-chromadepth and dynamic fog had higher decision times compared to their static versions. In a post-study questionnaire, users said that they preferred pseudo-chroma and that dynamic cues allowed them to perform better in terms of time, although this was not confirmed by the time measures.

We are aware of only one study that has been done using a VR environment for angiographic data exploration. Heinrich et al. (2021) conducted a user study where the effectiveness of Phong, pseudo-chromadepth and fog was compared in a monoscopic desktop environment versus a steroscopic VR enironment. They found that the environment did not have a significant impact on decision time and that fog resulted in the quickest trials, while pseudo-chromadepth resulted in the longest ones. In terms of correctness however, VR resulted in the least amount of sorting errors, and participants were also more certain about their decisions compared to the 2D desktop environment. Another interesting finding was that the type of cue that was used had a lesser impact in VR than the monoscopic desktop environment. This suggests that VR offers powerful perception cues that allow for correct depth judgements regardless of the shading technique.

3. Materials and Methods

The effectiveness of three depth cues in VR was studied: shading, pseudo-chromadepth and aerial perspective, as well as the dynamic version of each of these techniques as described by Drouin et al. (2020). In the dynamic version, the user modifies the rendering parameters of the cue by moving a virtual 3D surgical pointer using the VR controller. Our goal was not only to identify which visualization methods provide the best spatial understanding in a VR environment but also to see if generally, adding an interactive aspect to the depth cues significantly improves depth and spatial perception of cerebral vascular volumes in VR. As in the study by Drouin et al. (2020), all cues were implemented using standard ray casting and shaded with the Blinn-Phong reflection model (Blinn 1977). A video demonstrating the implemented visualizations is provided in the supplemental material.

3.1. Depth Cues

In this work, we studied shading, pseudo-chromadepth and aerial perspective cues and their dynamic counterparts. The reason behind this choice is that these cues have been extensively studied in 2D environments. Shading has often been used as the base case (eg. Abhari et al. (2012); Lawonn et al. (2017)), while pseudo-chromadepth and aerial perspective generally perform well in terms of correctness and decision time (Kersten-Oertel et al. 2014). Furthermore, the dynamic component for these cues was previously proposed by Drouin et al. (2020) and studied in a 2D environment. We thus chose these cues to better understand the impact of dynamic perceptual cues in a VR environment.

Each cue was tested twice, in two different variants: static and dynamic. Static cues require no user input and are calculated automatically using the 3D position of the VR headset. In the case of dynamic cues, the position of the tip of the handheld pointer controlled by the user modifies some parameters in the rendering.

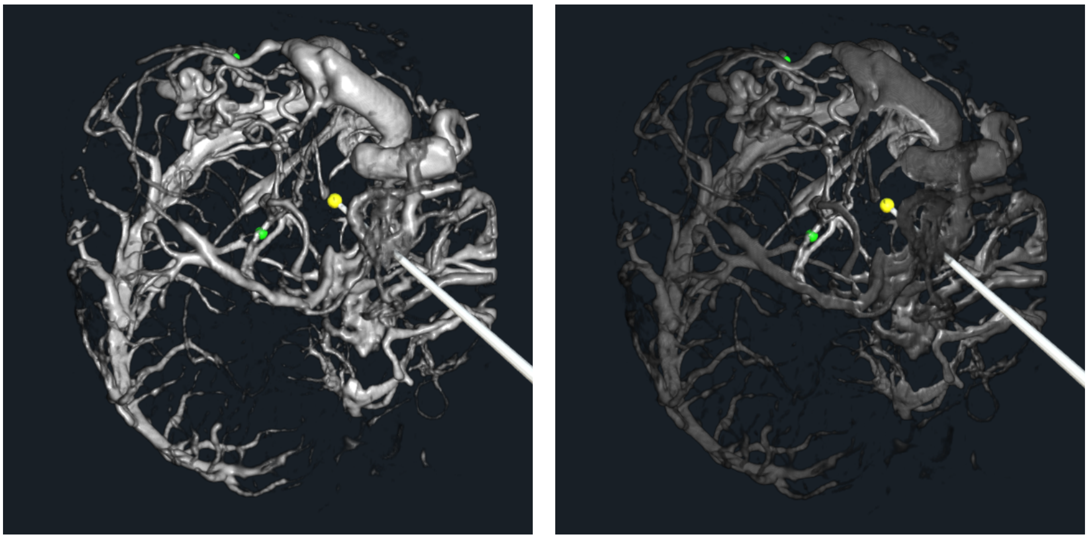
Shading

Shading is a photorealistic cue that simulates how an object reflects light when illuminated by a light source. One of the main advantages of this cue is that it is intuitive since it lets the viewer use knowledge about how objects are illuminated in real life. In the static version of shading, the light source is located at the midpoint between the two eyes, similarly to the light on a miner's helmet. This way, the volume is always fully illuminated, no matter the position of the head in the virtual environment (see Figure 1 (a)). In the dynamic version of shading, a point light source is attached to the tip of the pointer. Thus, as the user moves the pointer around the volume the anatomy around the pointer tip is illuminated and other parts of the volume fall into shadows. In other words, the light source works similarly to a match or flashlight that illuminates an object in a dark environment (see Figure 1 (b)). The light source has a linear decay in intensity, which provides an additional cue for localization of the pointer inside the volume. Shading was implemented using the Blinn-Phong reflection model (Blinn 1977)).

Chromadepth

Chromadepth is a non-photorealistic depth cue that uses colour to encode depth from the position of the viewer. The most popular chromadepth techniques are full chromadepth presented by Bailey and Clark (1998) and two-color (red to blue) chromadepth presented by Ropinski et al. (2006), also called pseudo-chromadepth. Kersten-Oertel et al. (2014) and Ropinski et al. (2006) determined that pseudo-chromadepth generally leads to better depth perception, thus we used this version in our study. For brevity, we will further refer to pseudo-chromadepth as chromadepth or simply chroma. In the static version, the interpolation is defined by the closest and furthest visible parts of the volume and adjusted to follow the view direction, so that the closest parts to the viewer appear in red and the furthest ones appear in blue. The closest and furthest visible points of the angiographic volume were estimated using a bounding sphere whose diameter was calculated as the average between the physical measures of width, height and depth of the volume.

In the dynamic version, the tip of the handheld pointer defines the near plane which corresponds to the start of the chroma interval and the red color. The far plane is located at 5 *cm* from the near plane and corresponds to the end of the chroma interval



(a) Static Shading (b) Dynamic Shading

Figure 1. Illustration of static and dynamic shading. With static shading (a), the volume is fully illuminated from the point of view of the user, while with dynamic shading (b), the tip of the handheld pointer controls the position of the light source, whose intensity decays with further distances.

and the blue color. Any part of the object closer than the close plane is shaded in red, while parts further than the far plane are shaded in blue. The distance between the near and far planes was defined manually using a trial and error approach, with 5 *cm* resulting in the best distance. Figure 2 illustrates static and dynamic chromadepth.

Aerial Perspective

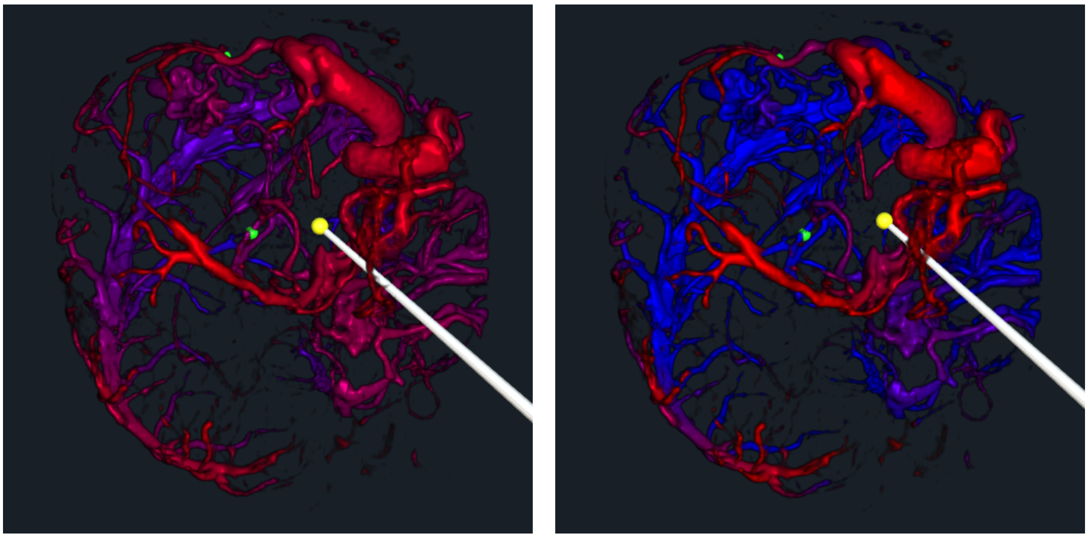
Aerial perspective encodes depth similar to the natural phenomenon of fog in real life, i.e. using a loss of contrast (Preim et al. 2016). In other words, parts of the volume which are close are highly contrasted from the background and are simply shaded with Blinn-Phong (1977), while further points fade into the color of the background. The encoding of this cue is very similar to chromadepth described previously. In the static version of aerial perspective, the interpolation is done across the full visible depth range of the volume. In the dynamic version, the user controls the depth at which the loss of contrast starts with the tip of the pointer. For brevity, we will sometimes refer to aerial perspective as fog. Figure 3 illustrates static and dynamic aerial perspective.

3.2. Experiment

A preliminary user study was conducted with the goal of comparing the effectiveness of visualization of angiographic medical scans using the three previously described cues in their static and dynamic version in VR. The goal of the study was to determine which techniques resulted in the best spatial and depth understanding of the angiographic vasculature, when judged using objective metrics.

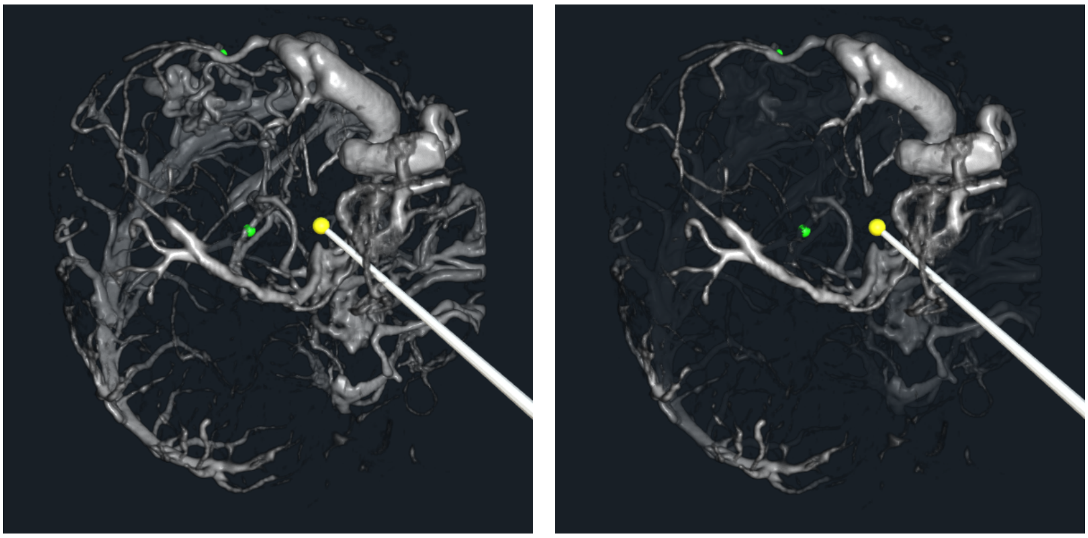
Data

The data consisted of 10 computed tomography angiographies (CTA) with 8 predefined views for each. Thus the study consisted of 80 trials. Half of the views had a distance



(a) Static Chromadepth (b) Dynamic Chromadepth

Figure 2. Illustration of static and dynamic shading. In static chromadepth (a), the color-coded depth interval spans through the entire volume. With dynamic chromadepth (b), the tip of the handheld pointer positions the close plane, which marks the start of the color gradient. This gradient spans for a short predefined distance.



(a) Static Aerial Perspective (b) Dynamic Aerial Perspective

Figure 3. Illustration of static and dynamic aerial perspective. In static aerial perspective (a), the loss of contrast spans through the entire volume. With dynamic aerial perspective (b), the tip of the handheld pointer positions the close plane, which marks the start of the contrast decay. The loss of contrast then spans for a short distance, and further parts become almost invisible.

of 20 *mm* between the targets, while the other half had a distance of 60 *mm*. The positions of the targets were chosen randomly, with the limitation that they had to be attached to some vessel in the volume. To achieve more variety for a single CTA scan, all views presented a different rotation. Additionally, the 80 trials were presented in a random order to each participant with a random cue associated to each to avoid any biases.

The angiographic scans have a resolution of 504×509×207 *voxels* on average and are rendered based on the actual size, which is 218×220×152 *mm* on average for all volumes. To achieve real-time direct volume rendering, the two-pass rendering algorithm described by Kruger and Westermann (2003) was used. Each participant saw each of the 3 cues 13 or 14 times, both in their static and dynamic version.

Setup

The study was designed as a sitting VR experience and was performed with the Valve Index headset (Valve Corporation, United States, Bellevue). The participants interacted with the virtual environment using the Valve Index controllers which were secured to their hands. In the VR environment, the medical volumes are rendered in 3D in a stereoscopic manner, so that binocular disparity may be used by the participant as an additional cue to understand the volume. Interocular distance was adjusted for each subject.

Experimental Task

Prior to performing the experiment the users were given a presentation about the study and how each of the cues works and gave their informed consent. After this, each participant did a tutorial within the VR environment to get used to the system.

The experimental task was designed to replicate as closely as possible the work of Drouin et al. (2020) to enable a direct comparison of the effect of dynamic and static cues in a VR environment versus when using a 2D display. Minor adjustments were made given the context of the VR display.

The task, similar to previous works in this research area (Ropinski et al. 2006; Kersten-Oertel et al. 2014; Kreiser et al. 2018), required the participant to determine which of two target vessels on the CTA volume is closer to the participant. The volume was fixed at the midpoint between the user's hands within the VR environment and the participant could move their head around to better understand the volume or to better localize the targets. The task of the user was to position the tip of a virtual pointer as close as possible to the closest of the two targets, without touching it. Each trial began with the user positioning the pointer on an orange spherical marker situated slightly below the volume (see Figure 4). This ensures that for all trials, the controller is positioned at approximately the same position at the start. Once the orange marker is touched, a new trial begins. Each trial ends in one of two ways: (1) the participant presses a button on the control that is not representing the surgical pointer or (2) if, by accident, one of the two targets is touched with the tip of the pointer. Performing the trials in this manner allowed us to measure spatial and depth understanding at both a global and local level. To determine which of the targets is closer to them, the participant must have a good global understanding of the angiographic volume. Whereas, to position the pointer as close as possible to the target without touching the target, the user must have a good understanding of the local structures of the volumes. As well as requiring precise localization, the limitation of not letting the user

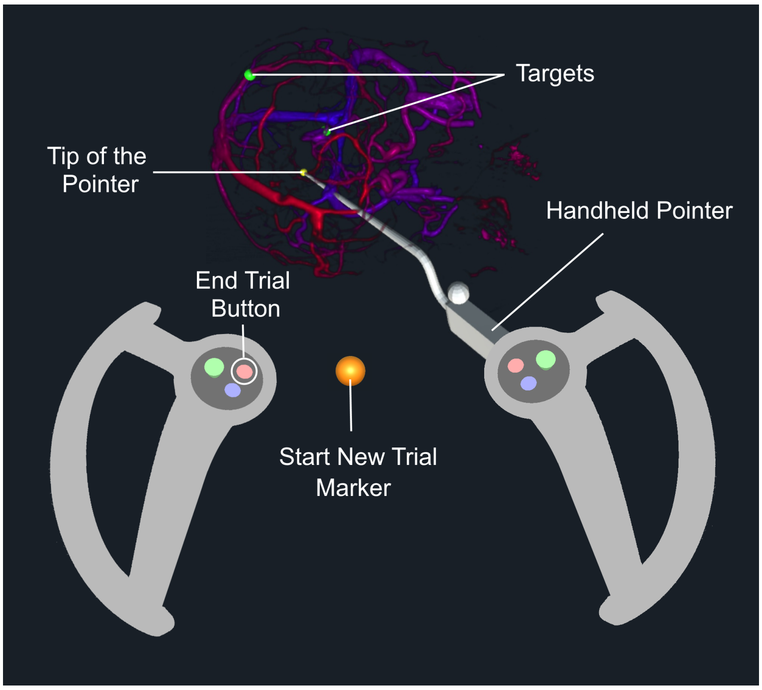


Figure 4. Illustration of the physical setup superposed with the image observed in the VR headset. The right controller modified the position and rotation of the handheld pointer, while the left controller is used uniquely to press a button to end the trial. Targets are indicated in green and are always attached to a vessel of the angiographic scan.

touch the target with the pointer also avoids having an additional occlusion cue which may affect the results.

Implementation

The study was implemented on the Unity Engine (Unity Technologies, Unites States, San Francisco) with a custom GPU ray casting implementation using HLSL shaders. The volume rendering component was implemented based on the direct volume rendering algorithm described by Drouin and Collins (2018). For the VR component, the Unity XR SDK[1] was used to allow hardware-agnostic VR. The implementation is currently closed-source.

The Valve Index VR Kit was used for VR during the user study. The rendering was done on a Windows 10 machine with an AMD Ryzen 7 3700X processor, 32 Gb of RAM and an NVidia RTX 2080 Ti video card with 11 Gb of video memory. The Valve Index headset provides a 1440 × 1600 LCD panel for each eye with 80 Hz, 90 Hz, 120 Hz and 144 Hz as possible refresh rates[2]. In our case, the rendering was done at 90 Hz in the virtual environment to ensure a smooth image.

[1] https://docs.unity3d.com/Manual/XR.html
[2] https://www.valvesoftware.com/en/index/headset

4. Results

A total of 12 subjects (7 male, 5 female, age 19-45y) participated in the study. All subjects had normal or corrected to normal vision and none of them were color-blind. In the case of corrected eyesight, subjects were wearing glasses while using the headset.

Although none of the participants were clinicians, they were doing research related to medical image visualization or computer image perception enhancement. It was determined with a post-test questionnaire that 92% of the participants had at least some previous experience with either virtual or augmented reality, with 42% rating their level of experience with either as "Very High". Furthermore, 83% of the participants had at least some experience with volume rendering, 92% had at least some experience with medical images and 67% had at least some experience with neurovascular anatomy. However, only 25% of the participants said that they at least somewhat used their knowledge about neurosurgical anatomy to execute the tasks of the study.

4.1. Objective Measures

The following metrics were recorded for each trial of the experiment:

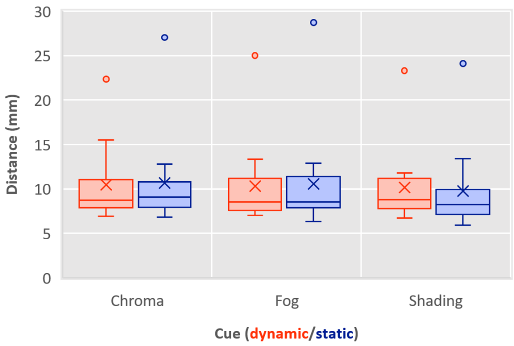
(1) **Correctness**: whether the participant properly determined which target was closer.
(2) **Accuracy**: the Euclidean distance between the tip of the handheld pointer and the chosen target.
(3) **Trial time**: the time that the participant took to complete the trial.
(4) **Head path**: the total distance that the head has traveled during the trial calculated as the sum of Euclidean distances between the positions of the head at each frame. The position of the head is defined as the 3D position of the headset in the real environment, which is measured using the Base Stations[3] in the Valve Index VR Kit.

We performed a 2-way repeated measures ANOVA on each of the four dependent variables, with the two discrete independent variables being the cue that was used (shading, chroma, fog) and whether the cue was static or dynamic. We will refer to these discrete independent variables as "cue" and "dynamics" correspondingly. The results are plotted in Figure 5.

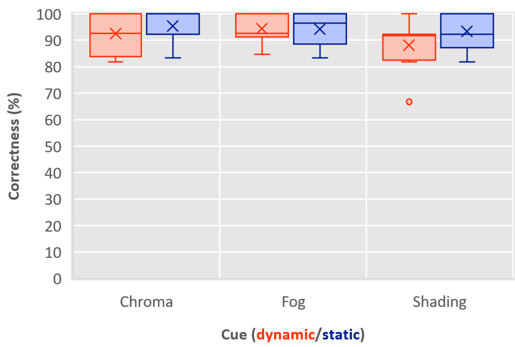
Effect on Accuracy

An ANOVA indicated that cue has a significant main effect on accuracy ($F(2, 22) = 3.484, p = 0.048$). Mauchly's test showed that the sphericity assumption was not violated ($\chi^2(2) = 1.48, p = 0.477$), so Greenhouse–Geisser correction was not applied. The Post-hoc Bonferroni-corrected pairwise t-tests have revealed that there exists a significant difference between the shading and chroma cues, and that shading results in 0.60 mm closer distance to the target (95% CI $(0.06, 1.15), p = 0.030$). No statistically significant difference was found between fog and other cues. Dynamics ($F(1, 11) = 0.006, p = 0.94$) and the interaction between cue and dynamics ($F(2, 22) = 0.867, p = 0.43$) did not have a significant impact on accuracy.

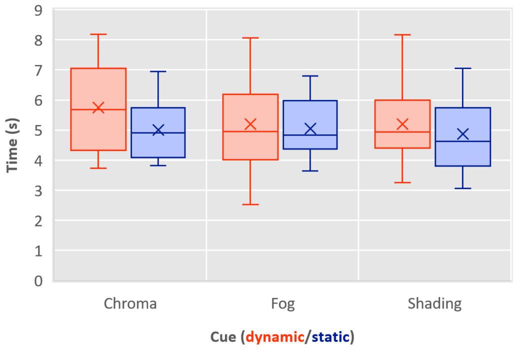
[3]https://www.valvesoftware.com/en/index/base-stations



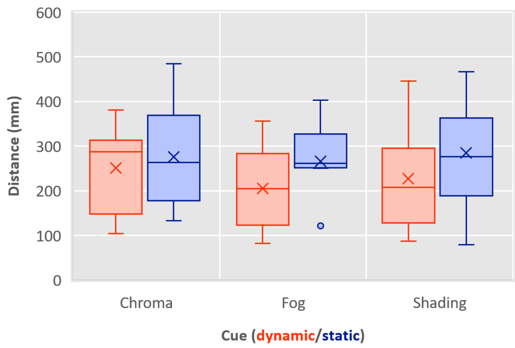
(a) Distance Pointer-Target (Accuracy)



(b) Correctness



(c) Trial Time



(d) Head Path Length

Figure 5. Collected data for each metric. Box plot represents the distribution of data in 4 quartiles. Dots represent outliers. Crosses represent the means.

Effect on Correctness

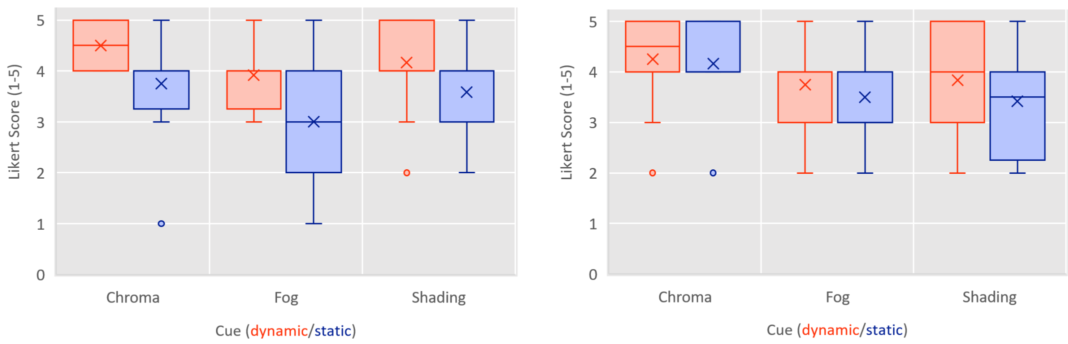
A significant main effect of dynamics was found on correctness ($F(1, 11) = 7.163, p = 0.022$). With static cues, participants properly guessed the closest target 2.6% more often (95% CI (0.5, 4.7)). There was no significant main effect of cue ($F(2, 22) = 1.686, p = 0.208$) or the interaction between cue and dynamics ($F(2, 22) = 0.997, p = 0.385$) on correctness.

Effect on Trial Time

In terms of trial time, no statistically significant main effect was found for dynamics ($F(1, 11) = 3.41, p = 0.092$) or cue ($F(2, 22) = 0.98, p = 0.392$). Additionally, no significant two-way interaction between dynamics and cue was found ($F(2, 22) = 1.563, p = 0.232$).

Effect on Head Path

A statistically significant main effect on head path was found for dynamics ($F(1, 11) = 6.058, p = 0.032$). It was determined that dynamic cues resulted in a 47.5 mm shorter head path than static ones (95% CI (5.0, 89.9)). In terms of cue ($F(2, 22) = 0.67, p = 0.52$) or the interaction between dynamics and cue ($F(2, 22) = 0.51, p = 0.61$), no statistically significant effect was found.



(a) Subjective Evaluation of Accuracy (b) Subjective Evaluation of Correctness

Figure 6. Collected subjective data for each metric. For each cue-dynamic combination, the subject was asked to rate how easy it was to reach the target (a) and how easy it was to determine the closest target (b).

4.2. Subjective Measures

Every participant completed a post-test questionnaire where they were asked to rate, using a Likert scale from 1 to 5, how easy it was to (1) determine which target is closer to them and (2) reach the chosen target accurately. The first is a subjective evaluation of **correctness** for each cue-dynamics combination. The second is a subjective evaluation of **accuracy** for each cue-dynamics combination. A two-way repeated measures ANOVA was calculated using the acquired information from the questionnaire where the first discrete variable is the cue and the second one is dynamics. The results are plotted in Figure 6.

Effect on Accuracy

In terms of user perception of accuracy, a statistically significant main effect was found on dynamics ($F(1,11) = 8.13, p = 0.016$). The post-hoc t-test showed that subjects rated dynamic cues 0.75 points higher than static ones (95% $CI(0.17, 1.33), p = 0.016$). Additionally, a statistically significant effect was found on cue ($F(2,22) = 9.14, p = 0.001$). The pairwise t-tests, Bonferroni-corrected, determined that subjects have rated chroma significantly higher than fog by 0.67 points (95% CI $(0.27, 1.07), p = 0.002$).

Effect on Correctness

In terms of user perception of correctness, a statistically significant main effect on cue was found (F(2, 22) = 5.34, p = 0.013). The post-hoc pairwise t-tests (Bonferroni-corrected) indicated that there exists a significant difference between perception of chroma and fog, and that participants rated chroma 0.58 points higher than fog on average (95% CI $(0.07, 1.10), p = 0.026$).

5. Discussion

In our study, we found that cue had no impact on relative depth perception or on decision time, while in many of the previous studies (performed using a 2D display) cues had a significant effect on depth perception (Ropinski et al. 2006; Kersten-Oertel et al. 2014; Kreiser et al. 2018). In these studies, chromadepth or its two-color variation were generally found to be more efficient than shading, both in terms of correctness and trial time. However, it seems that in VR, cue has little effect on those two variables. This result is very similar to what was observed by Heinrich et al. (2021), where cue had a significant impact on depth perception in the desktop 2D environment, while having little impact in the VR environment. Similarly to what was suggested in that study, we posit that this discrepancy may be due to the fact that in VR offers additional perception cues such as stereoscopy, which provide a good spatial and depth perception even with simple rendering. Thus, the specific rendering parameters in VR have less impact than on 2D displays.

The results of our study suggest that in terms of appreciation of local structures of the CT angiographic data (i.e. accuracy), shading performs better than chromadepth. Participants were able to get significantly closer to the target using shading than when using chromadepth. This differs from the results obtained by Drouin et al. (2020), where chroma and fog resulted in better appreciation of local structures than shading. This difference may suggest that in VR, shading provides a stronger depth understanding particularly if the light source is coupled with head movement. Even in static shading, the light is attached to the position of the viewer, so it is possible that by moving their head, viewers reveal some additional information about the geometry. However, more research is needed to confirm this hypothesis.

In terms of the effects of dynamics, we found that dynamic cues did not have a significant impact on understanding of local structures. However, in terms of understanding of global structures by determining the closer target (measured by the "correctness" variable), we determined that static cues performed better than their dynamic counterparts.

When compared to the study by Drouin et al. (2020), we found that for local and global understanding of the structures, dynamic cues performed worse in VR than when using a 2D display. On a 2D display, dynamic cues resulted in better understanding of local structures, while not having an effect on understanding of global structures (Drouin et al. 2020). In our study, dynamic cues performed equally or worse than static ones. This may suggest that in VR, binocular disparity and motion parallax play an important role in depth perception, which could sometimes negate the advantages of the dynamic cues.

Another difference that we observed between our and Drouin et al.'s (2020) study is that we didn't have a two-way interaction for any of the variables. In our results, there was no cue-dynamics combination which performs better than others in any metric.

When it comes to head movement, participants moved their head significantly less with dynamic cues than with static cues. This is an interesting result which could potentially be a strong argument in favor of dynamic cues. The more head movement in static cues suggests that the reason why participants were able to achieve better global appreciation of the vessel structure is because they relied on the motion parallax cue more. Dynamic cues could be particularly interesting if they are used in a context where the point of view is fixed, for example with a surgical microscope in the operating room (Drouin et al. 2015). We posit that the dynamic aspect could compensate for the absence of motion parallax which could have a significant impact on bringing these technologies in an intuitive and ergonomic way into an operating room context. We plan to explore this in future work.

When it comes to the comparison of effectiveness between static and dynamic cues, a discrepancy between objective and subjective results was found. This discrepancy is very similar to the one noted by Drouin et al. (2020), in that participants were overly confident about the effectiveness of dynamic cues. Even though subjects got similar accuracy with both static and dynamic cues, they thought that they were more effective with dynamic cues. A similar pattern could be observed for correctness; although subjects were able to correctly determine the closest target better with static cues, they thought that they performed equally well with both static and dynamic cues.

Overall, 75% of subjects said that they prefer dynamic cues over static ones. Even though this did not translate into better appreciation of local or global structures with static cues. We believe that this is due to the fact that with static cues participants had to move their head significantly more (perhaps to use the motion parallax cue), which may be more energy consuming for the user.

One of the limitations of the study was a relatively small number of participants (12) which was caused by the lockdown restrictions of the COVID-19 pandemic. However, this study size is similar to previous vascular volume visualization studies performed in a research lab environment where the number of participants varies from 10 (Abhari et al. 2012) to 30 (Heinrich et al. 2021). To ensure a high quality of data with subjects that managed to participate, the users were asked to remove their mask while using the VR headset to ensure their comfort. The headset and controllers were cleaned between each participant.

6. Conclusion

The goal of this study was to determine how medical volumes rendered with specific depth cues, initially studied on 2D displays rendering, could translate into a virtual environment. To do so, we performed a user study where we implemented three widely studied perceptual cues (shading, chromadepth and aerial perspective) and their dynamic counterparts proposed by Drouin et al. (2020).

Unlike with 2D displays, chromadepth and aerial perspective had little effect on previously widely studied metrics such as trial time and correctness. It seems to us that other cues in VR such as stereoscopy and motion parallax overshadow the transfer function-based ones if judged by these two metrics. However, in terms of local appreciation of the volume, we determined that shading results in better accuracy than chromadepth. This is a surprising result which might suggest that there exists an interaction between shading and stereoscopy or motion parallax. We plan to explore this in future work.

In terms of dynamics, dynamic cues perform equally or worse than static ones if we judge them using the same metrics that exist in 2D. However, if we look at observer movement during the study, the total path that the head travelled for each participant during the study was significantly less for dynamic cues. This may be considered as an advantage of dynamic cues. It is possible that for static cues, participants have to rely significantly more on the motion parallax cue, which explains the additional movement. Another advantage that dynamic cues have is that participants generally like them more, and they also think that they perform better with them than what can be observed using objective measures. Determining if there exists a link between higher preference for dynamic cues and smaller head movement could be an interesting research topic in the future.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Abhari K, Baxter JSH, Eagleson R, Peters T, Ribaupierre Sd. 2012. Perceptual enhancement of arteriovenous malformation in MRI angiography displays. In: Medical Imaging 2012: Image Perception, Observer Performance, and Technology Assessment; vol. 8318; Feb. International Society for Optics and Photonics. p. 831809. [accessed 2021-04-26]. Available from: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/8318/831809/Perceptual-enhancement-of-arteriovenous-malformation-in-MRI-angiography-displays/10.1117/12.911687.short.

Bailey M, Clark D. 1998. Using ChromaDepth to Obtain Inexpensive Single-image Stereovision for Scientific Visualization. Journal of Graphics Tools. 3(3):1–9. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10867651.1998.10487491; [accessed 2021-04-27]. Available from: https://doi.org/10.1080/10867651.1998.10487491.

Blinn JF. 1977. Models of light reflection for computer synthesized pictures. In: Proceedings of the 4th annual conference on Computer graphics and interactive techniques; Jul; New York, NY, USA. Association for Computing Machinery. p. 192–198. SIGGRAPH '77; [accessed 2021-04-25]. Available from: https://doi.org/10.1145/563858.563893.

Drouin S, Collins DL. 2018. PRISM: An open source framework for the interactive design of GPU volume rendering shaders. PLOS ONE. 13(3):e0193636. Publisher: Public Library of Science; [accessed 2021-04-26]. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0193636.

Drouin S, Collins L, Kersten-Oertel M. 2018. Interaction in Augmented Reality Image-Guided Surgery. In: Mixed and Augmented Reality in Medicine. 1st ed. CRC Press; p. 99–114.

Drouin S, Giovanni DAD, Kersten-Oertel M, Collins DL. 2020. Interaction Driven Enhancement of Depth Perception in Angiographic Volumes. IEEE Transactions on Visualization and Computer Graphics. 26(6):2247–2257. Conference Name: IEEE Transactions on Visualization and Computer Graphics.

Drouin S, Kersten-Oertel M, Louis Collins D. 2015. Interaction-Based Registration Correction for Improved Augmented Reality Overlay in Neurosurgery. In: Linte CA, Yaniv Z, Fallavollita P, editors. Augmented Environments for Computer-Assisted Interventions; Cham. Springer International Publishing. p. 21–29. Lecture Notes in Computer Science.

Harman KL, Humphrey GK, Goodale MA. 1999. Active manual control of object views facilitates visual recognition. Current Biology. 9(22):1315–1318. [accessed 2021-07-15]. Available from: https://www.sciencedirect.com/science/article/pii/S0960982200800536.

Heinrich F, Apilla V, Lawonn K, Hansen C, Preim B, Meuschke M. 2021. Estimating depth information of vascular models: A comparative user study between a virtual reality and a desktop application. Computers & Graphics. 98:210–217. [accessed 2021-07-14]. Available from: https://www.sciencedirect.com/science/article/pii/S0097849321001138.

Heinrich F, Bornemann K, Lawonn K, Hansen C. 2020. Interacting with Medical Volume Data in Projective Augmented Reality. In: Martel AL, Abolmaesumi P, Stoyanov D, Mateus D, Zuluaga MA, Zhou SK, Racoceanu D, Joskowicz L, editors. Medical Image Computing and Computer Assisted Intervention – MICCAI 2020; Cham. Springer International Publishing. p. 429–439. Lecture Notes in Computer Science.

Kersten-Oertel M, Chen SJS, Collins DL. 2014. An Evaluation of Depth Enhancing Perceptual Cues for Vascular Volume Visualization in Neurosurgery. IEEE Transactions on Visualization and Computer Graphics. 20(3):391–403. Conference Name: IEEE Transactions on Visualization and Computer Graphics.

Kreiser J, Hermosilla P, Ropinski T. 2018. Void Space Surfaces to Convey Depth in Vessel Visualizations. arXiv:180607729 [cs]. ArXiv: 1806.07729; [accessed 2021-07-14]. Available from: http://arxiv.org/abs/1806.07729.

Kruger J, Westermann R. 2003. Acceleration techniques for GPU-based volume rendering. In: IEEE Visualization, 2003. VIS 2003.; Oct. p. 287–292.

Laha B, Bowman DA. 2013. Volume cracker: a bimanual 3D interaction technique for analysis of raw volumetric data. In: Proceedings of the 1st symposium on Spatial user interaction;

Jul; New York, NY, USA. Association for Computing Machinery. p. 61–68. SUI '13; [accessed 2021-04-23]. Available from: https://doi.org/10.1145/2491367.2491368.
Laha B, Bowman DA, Socha JJ. 2016. Bare-Hand Volume Cracker for Raw Volume Data Analysis. Frontiers in Robotics and AI. 3. Publisher: Frontiers; [accessed 2021-04-23]. Available from: https://www.frontiersin.org/articles/10.3389/frobt.2016.00056/full.
Lawonn K, Luz M, Hansen C. 2017. Improving spatial perception of vascular models using supporting anchors and illustrative visualization. Computers & Graphics. 63:37–49. [accessed 2021-08-25]. Available from: https://www.sciencedirect.com/science/article/pii/S0097849317300171.
McGuffin M, Tancau L, Balakrishnan R. 2003. Using deformations for browsing volumetric data. In: IEEE Visualization, 2003. VIS 2003.; Oct. p. 401–408.
Preim B, Baer A, Cunningham D, Isenberg T, Ropinski T. 2016. A Survey of Perceptually Motivated 3D Visualization of Medical Image Data. Computer Graphics Forum. 35(3):501–525. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.12927; [accessed 2021-04-26]. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.12927.
Ropinski T, Steinicke F, Hinrichs K. 2006. Visually Supporting Depth Perception in Angiography Imaging. In: Butz A, Fisher B, Krüger A, Olivier P, editors. Smart Graphics; Berlin, Heidelberg. Springer. p. 93–104. Lecture Notes in Computer Science.
Simons DJ, Wang RF, Roddenberry D. 2002. Object recognition is mediated by extraretinal information. Perception & Psychophysics. 64(4):521–530. [accessed 2021-07-15]. Available from: https://doi.org/10.3758/BF03194723.
Toulouse A, Thiele S, Herkommer A. 2019. Virtual reality headset using a gaze-synchronized display system. In: Optical Design Challenge 2019; vol. 11040; Feb. International Society for Optics and Photonics. p. 1104009. [accessed 2021-04-26]. Available from: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11040/1104009/Virtual-reality-headset-using-a-gaze-synchronized-display-system/10.1117/12.2523920.short.
Vienne C, Masfrand S, Bourdin C, Vercher JL. 2020. Depth Perception in Virtual Reality Systems: Effect of Screen Distance, Environment Richness and Display Factors. IEEE Access. 8:29099–29110. Conference Name: IEEE Access.