Studying the Effects of Stereo, Head Tracking, and Field of Regard on a Small-Scale Spatial Judgment Task

Eric D. Ragan, Regis Kopper, Philip Schuchardt, and Doug A. Bowman

Abstract—Spatial judgments are important for many real-world tasks in engineering and scientific visualization. While existing research provides evidence that higher levels of display and interaction fidelity in virtual reality systems offer advantages for spatial understanding, few investigations have focused on small-scale spatial judgments or employed experimental tasks similar to those used in real-world applications. After an earlier study that considered a broad analysis of various spatial understanding tasks, we present the results of a follow-up study focusing on small-scale spatial judgments. In this research, we independently controlled field of regard, stereoscopy, and head-tracked rendering to study their effects on the performance of a task involving precise spatial inspections of complex 3D structures. Measuring time and errors, we asked participants to distinguish between structural gaps and intersections between components of 3D models designed to be similar to real underground cave systems. The overall results suggest that the addition of the higher-fidelity system features support performance improvements in making small-scale spatial judgments. Through analyses of the effects of individual system components, the experiment shows that participants made significantly fewer errors with either an increased field of regard or with the addition of head-tracked rendering. The results also indicate that participants performed significantly faster when the system provided the combination of stereo and head-tracked rendering.

Index Terms— Artificial, augmented, and virtual realities; Graphical user interfaces.

————————————— ◆ —————————————

1 INTRODUCTION

Immersive virtual reality (VR) systems commonly provide advanced features such as stereoscopy, wide field of view, and head-tracked view rendering. Compared to standard desktop displays, immersive VR systems produce visual stimuli with a higher level of similarity to real-world stimuli (we refer to this as the system's level of *fidelity*). Since immersive features support enhanced spatial cues, researchers often point to improved perception and understanding of 3D spatial information as an example of the benefits of VR [e.g., 1, 2, 3]. For instance, Chance, Gaunet, Beall, and Loomis [4] found that rotating using physical head or body rotations (as enabled by head tracking), as opposed to joystick-controlled rotation, can improve the ability to maintain orientation and understanding spatial layout in a virtual environment (VE). Other studies [1, 5] found that the addition of stereoscopy and head tracking improved participant comprehension of 3D graph structures.

Despite the evidence for the advantages of additional spatial cues, relatively few applications take advantage of immersive VR displays to support real-world tasks. One obvious reason for the low number of real-world VR applications is the high cost associated with immersive displays. But a VR system does not have to be viewed as either immersive or non-immersive; that is, individual immersive features can be added to increase the overall level of VR fidelity. In this sense, rather than categorizing immersive and non-immersive systems, the level of immersive fidelity can be viewed along a multidimensional continuum, with different combinations of individual immersive features contributing to the overall level of fidelity [6]. These features account for both the realism of the sensory stimuli output by the display (i.e., *display fidelity*) and the realism of the interaction techniques that provide input to the virtual simulation (i.e., *interaction fidelity*). Unlike subjective outcomes of immersive VR systems, such as engagement or *presence* (i.e., the feeling of being in the simulated environment, rather than merely working with a computer system [7]), the levels of display and interaction fidelity objectively depend on the display's hardware and the supported methods of interaction [6, 8]. As such, studying how different immersive features affect performance both individually and in combination with other features can increase knowledge of how to design VR systems to maximize performance while minimizing costs.

One challenge when applying the results of controlled studies to real-world scenarios is that the experimental tasks may not be similar to the types of real-world tasks that could potentially benefit from improved spatial perception. While numerous previous experiments have considered spatial tasks involving navigation [e.g., 3, 4] or the general understanding of 3D structures [e.g., 9, 10], many real-world tasks require high-precision, relative spatial judgments of specific

————————————————————

- *E. D. Ragan is with Virginia Tech, Center for Human-Computer Interaction, Blacksburg, VA 24060. E-mail: eragan@vt.edu.*
- *R. Kopper is with University of Florida, Gainesville, FL 32611. E-mail: kopper@cise.ufl.edu.*
- *P.Schuchardt is with Cavewhere (www.cavewhere.com).*
- *D. A. Bowman is with Virginia Tech, Center for Human-Computer Interaction, Blacksburg, VA 24060. E-mail: bowman@vt.edu.*

structural sub-components. Rather than focusing solely on spatial perception, our research investigates how a display's spatial cues affect the ability to judge positions and relationships among sub-components. *Small-scale spatial judgments* require careful visual inspections of components that are small relative to the scale of the environment, including tasks such as precise size comparisons, the identification of object intersections, or spatial projections. Such judgments are important for many real-world tasks in engineering and scientific visualization when it is necessary to determine whether different objects are touching each other, whether two paths will cross each other, or whether open spaces exist between objects. Correctly making such judgments is important for collision analysis in architecture and construction, well path planning for oil and gas pipelines [11], design drafting in engineering [12, 13], as well as certain types of scientific visualization [e.g., 10, 14].

In this research, we studied the effects of several components of visual display fidelity and viewing interaction fidelity on small-scale spatial judgments with an experimental task involving the identification of collisions and gaps in complex underground cave systems. In an earlier experiment, we found that participants exhibited significantly better spatial understanding of underground cave systems when the VR system provided more immersive viewing (head-tracked view rendering, stereoscopy, and the additional surrounding screens) [15]. However, because this previous study only compared two conditions (low fidelity vs. high fidelity), it was unable to determine how the individual features affected spatial understanding. Extending this prior research, we present a follow-up study that addresses this limitation by independently controlling each immersive component. This new investigation provides a deeper analysis of the effects of the system components than was possible with the earlier approach, making it possible to generalize the effects to multiple systems. Though our previous study considered a broad analysis of various spatial understanding tasks, our new study focuses on small-scale spatial judgments requiring close inspection of structural components. The results suggest that higher levels of components of display and interaction fidelity—even individually—can increase both the speed and accuracy of collision and gap identifications.

2 RELATED WORK

Our work builds upon the results of many previous studies of how various system characteristics affect spatial understanding in VEs. Following a previous study that evaluated the combined effects of stereoscopy, head tracking, and field of regard on spatial understanding tasks in underground cave systems, the study presented in this paper evaluated the effects of these components (both independently and combined) on tasks requiring high-precision spatial judgments.

2.1 Fidelity in Immersive VR and Spatial Cues

We define *display fidelity* as the objective degree to which the sensory stimuli produced by a system correspond to real-world sensory stimuli [8]. Display fidelity is thus dependent on the display's physical output, rather than the realism of the virtual content. For example, in the real world, we observe the world in stereo with a field of view of approximately 180°. The closer a display is to matching such levels from the real world, the higher the display fidelity will be for these corresponding components. Different systems can have different levels of display fidelity for different components. For example, a computer monitor might have a lower field of view than a large projected display but the monitor could have higher spatial resolution.

We and others have previously used the term *immersion* to refer to display fidelity [7, 16], but we have found that *immersion* can be ambiguous, since it is sometimes used to describe engagement or the sense of presence in another place. Thus, we opt for the use of *display fidelity* to avoid such confusion.

Just as *display fidelity* describes the realism of a display's sensory output, *interaction fidelity* describes the realism of the interaction methods used in a VR system as compared to the actions used in an equivalent real-world scenario [8]. For example, in the real world, we can physically turn in any direction to view more of our surroundings. If we can also physically turn to view more of a VE, then the level of interaction fidelity for view control would be higher than if we could only use mouse and keyboard input to virtually turn. Note that interaction fidelity is specific to the type of action. For the topic of spatial perception, we are most concerned with viewing interactions.

Immersive VR systems often support high visual display fidelity (e.g., stereo, high FOV) in conjunction with high-fidelity viewing interactions (i.e., physical rotation, head tracking). Both types of fidelity affect the realism of the viewing experience and the perception of 3D space. Head tracking, for example, allows users to use natural, physical head and body movements to control motion parallax—a change in the visual location as a result of a change in the viewer's location [17]. Because objects or surfaces that are further away move more slowly across the visual field than those that are closer to the viewer, such movements can help the user to distinguish among objects at different distances [13-14]. Motion parallax can also help viewers judge 3D depths [18] and object orientations [17].

VR systems also often support stereoscopy, which presents slightly different imagery to each eye based on the distance between the eyes. Binocular disparity allows viewers to merge the two images and use the difference to gauge depth information [19]. Just as this aids spatial perception of physical objects in the real world, stereoscopy can help with the spatial processing of virtual objects [e.g., 1, 5]. However, stereoscopy is not perfect, and can also introduce new forms of eye strain [20]. For example, while stereoscopy allows eye convergence on an object, it does not change the distance of the physical display, so it does not support normal eye accommodation for the virtual imagery [21]. Thus, while some display features have the potential to improve spatial perception, empirical studies are needed to assess their value in VEs.

The perception of 3D depth and shape are improved when multiple depth cues are present concurrently, as in immersive VR. Well-documented cue integration effects include improved 3D shape perception from the combination of stereo (binocular disparity) and motion cues [22, 23], visual texture motion cues [24], and for stereo, shape outline and texture cues [25]. The effects of depth cue integration extend beyond 3D shape perception to include more basic phenomena such as the perception of surface slant [26], with some researchers going so far as to posit that stereo and motion cues are intrinsically interdependent in the visual system [27]. The behavioral evidence for depth cue integration has been supported by single-unit neurophysiology studies in macaques. Parietal neurons selective for 3D surface orientation [28-30] and for 3D shape features [31] have been found to respond similarly to multiple kinds of depth cues, including stereo [29-32], texture [30, 31] and monocular perspective cues [28, 31]. Notably, individual neurons were also found that responded selectively only when multiple different depth cues occurred simultaneously [29]. Since VR displays often afford multiple depth cues and make cue integration possible, these findings support the idea that stereo and head tracking will influence performance on spatial tasks in VR. However, controlled studies are needed to show which cues and combinations of cues affect performance on particular tasks, and how large these effects are.

2.2 Spatial Understanding in VR

Many studies have partially addressed this need by investigating how the components of immersive systems affect different types of spatial understanding in VR. Spatial understanding is important for various tasks that require the interpretation and understanding of spatial information. For example, navigating a VE requires knowledge of an environmental layout and maintenance of self-orientation. Pausch, Proffitt, and Williams [33] studied the effects of head tracking on a search task within the space surrounding the user. Comparing conditions using either a head-tracked HMD (head-mounted display) or an HMD with a hand-held input device, the researchers found that head tracking helped participants to more quickly determine when a target item was not in the environment. These results suggest that participants were able to develop better mental models of the 3D environment, allowing more efficient search strategies. Studying navigation and wayfinding using HMDs, previous studies also contributed evidence that the addition of head tracking can help users maintain orientation and better understand the overall spatial layout of a VE [4, 34].

As an example of another type of a task involving spatial understanding, the exploration and interpretation of abstract, 3D information requires understanding positions and recognizing patterns in a data set. In a controlled experiment, Arns, Cook, and Cruz-Neira [35] compared performance differences between an CAVE-type system (having relatively high display and interaction fidelity) and a desktop display (providing lower fidelity) for a statistical visualization task involving the identification of structures in data. Their study found faster performance with the high-fidelity setup, which provided stereoscopic imagery on four large projection screens (three walls and a floor). It can be interpreted that the additional spatial cues provided advantages in perceiving and navigating the 3D visualizations. In similar work, Raja, Bowman, Lucas, and North [36] studied the effects of individual display components on abstract information visualization, controlling the number of display walls and the use of head-tracked rendering. In this experiment, participants tried to determine minimum data values, identify possible outliers, and recognize patterns in a data set. The resulting trends found in the study suggested that the additional higher-fidelity display components were helping participants maintain their orientations within the VE and complete their tasks more quickly.

Other research has also shown that higher levels of visual fidelity can help viewers gain an overall understanding of the shape of 3D structures. For example, Barfield, Hendrix, and Bystrom [9] provided results showing that the addition of either stereo or head tracking (or both) to a desktop display helped participants to better understand the shape of a bent wire structure, though the results were not statistically significant.

Studying the effects of display components through the same approach as we used, Laha et al. [37] independently controlled head tracking, stereo, and FOR (field of regard—the range of the VE that can be viewed with physical rotation) in a CAVE-type display. Studying performance differences on feature searches and general structural understanding for volume data visualization, the researchers found benefits for the enhanced versions of each of the display components. While these results do provide some backing for the hypotheses of our study, the emphasis on volume data visualization and visual search make it difficult to apply the results to tasks involving precise, small-scale spatial judgments.

2.3 Precise Spatial Judgments in VR

While a large number of studies have investigated spatial understanding in VR, fewer studies have focused on spatial judgment tasks that require precise visual inspections and comparisons. Prabhat et al. [10] presented a study using an information visualization task involving not only comparing structures and identifying key features, but also identifying object intersections. Based on a task involving volumetric biological data, this study compared a standard desktop system with no tracking or stereoscopy, a "fish tank VR" system using a desktop monitor with stereoscopy and head tracking, and a CAVE-type system (with three walls and a floor) with stereoscopy and head tracking. Again, the results suggested improvements in task performance due to the addition of higher-fidelity system components. The comparison of the fish tank setup to the CAVE-type setup allowed the researchers to evaluate performance differences due to an increased FOV (field of view) and FOR provided by multiple large screens. In the tasks of this study, the enhanced spatial cues helped participants to more easily explore the visualizations and gain a better understanding of individual features and overall composition. However, this work did not focus on analyzing the results specifically for small-scale spatial judgment tasks. Additionally, because the experi-

ment compared separate display systems and input devices, many factors prevented clearly determining which display components caused performance differences.

Comparing task completion times between participants using an immersive CAVE-like system and those using a desktop computer system, Gruchalla [11] found significant speed improvements for the task of planning paths for oil wells. In this study, the more immersive condition provided stereo and head-tracked viewing, and used a tracked wand for navigation and direct pointing. The desktop version of the application used a stereoscopic computer monitor with a mouse, keyboard, and virtual widgets to support interaction. While this task required judgments of small-scale spatial features, it was not possible to specifically determine what differences in display or interaction techniques caused the performance differences between the two conditions.

Several studies have employed spatial understanding tasks that require participants to visually trace paths within 3D graph structures [1, 5]. While this type of task may require small-scale spatial judgments to distinguish between paths in some places, understanding the general 3D structure of the graph would be expected to be the primary indicator allowing viewers to trace a path through the graph. Several studies employing such a path-tracking path investigated the effects of adding stereo and head tracking to a desktop display [1, 5]. Overall, the results suggest performance improvements due to the addition of either stereo or head tracking, with the best performance achieved with the combination of both stereo and head tracking. Additionally, finding no significant differences between head-tracked motion cues and hand-controlled motion, Ware and Franck [1] concluded that simply having any type of motion cues—not necessarily just through head-tracked viewing—is enough to improve performance in this type of path-tracing task.

Combined with our previous study [15], these past experiments helped us to focus our new study. While previous research has shown advantages for spatial understanding due to increased levels of display and interaction fidelity, few studies have focused on small-scale spatial judgments or employed experimental tasks similar to those used in real-world applications. Additionally, few previous studies have been able to independently control different components of display and interaction fidelity in order to determine both their individual and combined effects.

2.4 Prior Experiment on Spatial Understanding

The study presented in this paper is an extensive follow-up study to an earlier experiment, in which we investigated spatial understanding of underground cave systems [15]. The previous study compared performance on several spatial understanding tasks between two conditions with varying levels of fidelity. The high-fidelity system provided head-tracked viewing and stereoscopy within a CAVE with three walls and a floor, while the low-fidelity condition only used a single wall of the CAVE without stereo or head tracking. The display condition was varied between subjects, so each participant either used the high-fidelity or the low-fidelity display setup. Both conditions also allowed participants to use a wand joystick to translate or rotate the virtual world.

Participants performed a variety of spatial understanding tasks involving the inspection of virtual models of complex, underground cave systems. The 3D models were created based on cave survey data from a real cave. Fig. 1 shows an example of a cave model used for the experimental task. After a training session with a practice model, participants completed a set of tasks, in which they were asked to answer questions about the cave model while they navigated the VE. The spatial understanding tasks included searches for key spatial features, comparisons of relative feature measurements, and absolute measurements of spatial features of the cave model. Task performance was measured based on both time and accuracy.

The analysis of the results showed that the high-fidelity condition supported significantly better performance for both time and accuracy. Additionally, analysis of specific tasks revealed an interaction between the task and the level of fidelity, showing that the effects of the level of display fidelity varied based on the specifics of the spatial understanding task. For example, higher fidelity improved both time and accuracy for certain tasks involving searches for small spatial features (such as identifying connections between portions of the cave or locating pits on the cave floor), but had no significant effect on either metric for certain questions involving measurements of distances or angles. Thus, for the study presented here, we decided on a spatial judgment task involving small-scale features because of the interesting interaction results of the previous study, as well as its relevance to real-world engineering and scientific visualization tasks.
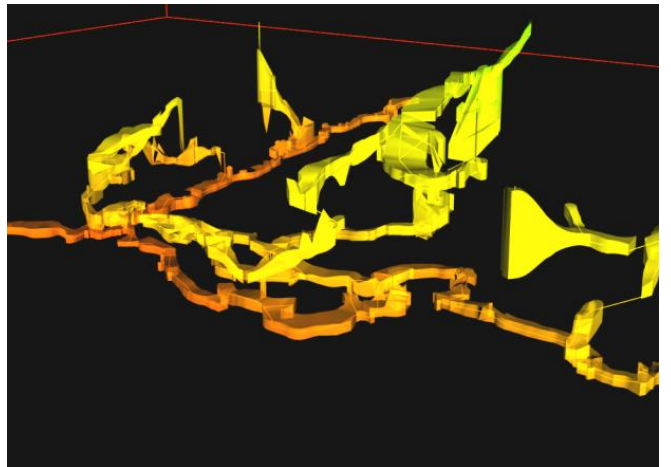

Fig. 1. An example of a 3D cave model used in the previous experiment [15].

While the experimental display conditions of the previous study varied in terms of stereoscopy, head tracking, and FOR, it was not possible to accurately determine which display components (or combinations of components) affected performance. Understanding the effects of individual display features is essential for optimizing systems while observing cost and space constraints. This is also important for generalizing the effects for other applications, and makes it possible to organize the results within the scope of other work that studies the effects of individual components of

display and interaction fidelity.

In the research presented in this paper, we control these components independently to study their effects on the performance of a task requiring spatial judgments of small-scale regions of a 3D structure. Though previous projects have studied the effects of these features in spatial understanding tasks [e.g., 1, 10, 15], our work is novel in that we independently controlled all three components and focus on small-scale spatial judgments. This led to a much larger and more in-depth study than the previous experiment.

3 EXPERIMENT

We conducted a controlled experiment to study the effects of stereoscopy, head tracking, and FOR on performance of a spatial understanding task requiring spatial judgments of small-scale structural features. The results show that either increased FOR or the addition of head-tracked rendering was enough to significantly improve judgment accuracy, and the combination of both stereo and head tracking significantly increased completion speed.

3.1 Hypotheses

We hypothesized that stereoscopy, head tracking, and increased FOR would improve performance on small-scale spatial judgment tasks. Stereoscopy provides the additional spatial cue of binocular disparity, making it easier to perceive spatial depths at close distances. Head tracking enables the user to change the view point using familiar physical movements (e.g., walking, leaning, crouching, or turning) to use motion parallax cues to understand 3D structures. Previous studies have found evidence that stereoscopy and head-tracked viewing can improve spatial understanding [e.g., 1, 11, 34]. Our study looked at whether similar effects are observed for small-scale spatial judgments. Also, compared to either head tracking or stereoscopy alone, we hypothesized better performance with both, as observed by Ware [5].

Increased FOR increases the amount of the VE that can be viewed with physical, bodily rotations. We expected this would make it easier to maintain understanding of position and orientation within the VE, helping users to improve performance.

The combination of stereoscopy, head tracking, and increased FOR afford users the opportunity to bring the virtual 3D model (or at least its smaller structural subcomponents) into the physical workspace of the CAVE, and walk around it to examine it from different sides (corresponding to the four screens of the CAVE). Thus, much less virtual navigation is needed with this higher level of viewing interaction fidelity, and users can take advantage of strong visual and proprioceptive cues to increase spatial understanding. None of the conditions with lower overall fidelity (which lack stereoscopy, head tracking, and/or increased FOR) afford this same level of "natural" viewing of the 3D model. Therefore, we hypothesized that we would find a three-way interaction among these variables, with the best performance at the highest overall level of fidelity.

3.2 Task

For the experimental task, participants inspected virtual 3D models of complex, underground cave structures. Eleven similar models were created based on a real cave-maze layout (a structural layout of multiple intersecting pathways). For each model, the structure was designed with four horizontal layers of interconnecting cave tubes (see Fig. 2). These horizontal layers of networking tubes were connected by several vertical tube paths. For each experimental trial, participants inspected the cave structures and counted the number of vertical tubes that connected the horizontal layers. The models also included vertical tubes that did not connect between levels. The presence of these tubes complicated the task; careful inspection was required in order to determine whether or not a vertical tube actually connected horizontal levels. Fig 3 shows an example of a gap between the end of vertical tube and the horizontal level.
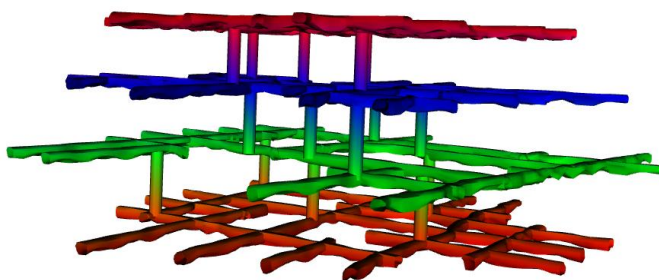

Fig. 2. An example of an entire cave-system structure used in the experiment, as viewed from the side. This side view shows four horizontal levels, connected by vertical tubes
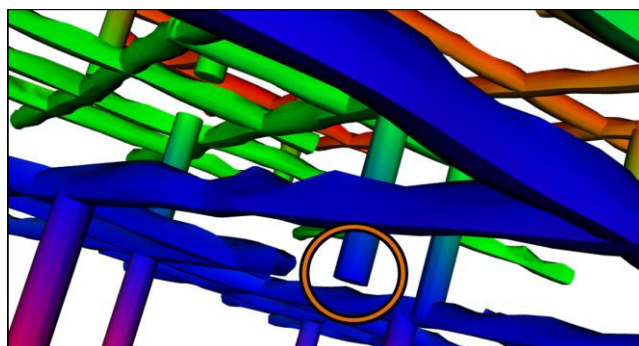

Fig. 3. A view of the cave-system structure as viewed from inside the model. An example of small gap between a vertical tube and a horizontal level of tubes is circled

Because participants were able to control their navigation (see the *Experimental Design* section for details), they were able to view the structures and intersections from varying distances and viewpoints. Consequently, objects' visual sizes varied based on virtual movement. Navigation made it possible for participants to move closer or farther away to change the visual size of the tubes and gaps. Participants were allowed to navigate freely to achieve what they felt were the most advantageous viewing locations.

The structure models were smooth-shaded and colored according to elevation, making each horizontal tube layer a

consistent color. The vertical tubes between horizontal layers were colored with a gradient between the colors of its two enclosing horizontal layers. Participants viewed the structure against a white background. Model dimensions were designed with approximately a 3:1:3 ratio for x:y:z (with y being the vertical dimension). The models were scaled to roughly fit entirely inside the CAVE's 10' by 10' horizontal display space.

The tube-structure models were designed to exhibit approximately equal levels of structural complexity. All structures included 15 vertical tubes. In each model, some of these 15 tubes (between three and 12) made connections between the horizontal levels, while others did not (see Fig. 3). One of the models was used for the training session and the other ten were used for the experiment trials. All participants viewed the models in the same order.

Participants reported the number of connections verbally, with time and the number of errors recorded as performance metrics. This task involved three stages: visual search for a potential gap location, navigation to view the potential gap location from an advantageous viewpoint, and judgment of whether a gap was present or not. We refer to the overall task as a small-scale spatial judgment, however, because success or failure of the task depended on the correctness of the judgment. We do not imply that stereoscopy, head tracking, and FOR affect only the final judgment stage of the task. These components could affect navigation and, to some extent, visual search as well (e.g., head tracking could improve the effectiveness of short-range navigation). We were therefore looking for the effects of these components on the overall task of making small-scale spatial judgments.

We expected, however, that most of the effects would be due to the users' ability to perceive the gaps or lack thereof from close range. The intersections or gaps between structural components were relatively small in comparison to the rest of the structure, making judgments difficult. Though participants had to navigate through the VE to view each small-scale potential collision point, this navigation was relatively easy and always required virtual translation (using a joystick, as explained in section 3.4). On the other hand, judging potential intersections was difficult and required up-close spatial inspections.

3.3 Participants

Fifty-two volunteers (39 male, 13 female) participated in the experiment. Participant ages ranged from 18 to 68, with 68% of participants younger than 30 years of age. The mean age was 27.5 and the median was 21.5 years.

3.4 Apparatus

Participants viewed the 3D structures within a four-screen CAVE projection display using 1280x1024 Electrohome CRT projectors. The CAVE display consisted of three rear-projected walls measuring 10' wide and 9' high and a front-projected floor measuring 10' by 10'. Stereoscopic viewing was possible through active shutter glasses.

Participants used an InterSense tracked six degree-of-freedom wand with a joystick for navigation. Participants could point the wand in the direction they wished to travel and push the joystick forward or backward in order to move in the desired direction. Additionally, participants could rotate the virtual world along the vertical axis by moving the joystick to the left or the right. Navigation was not restricted by collisions with the tube structures; participants were free to navigate through the 3D model. Regardless of condition or positioning in the virtual space, the system's frame rate remained at approximately 60 fps.

3.5 Experimental Design

Stereoscopy, head tracking, and FOR were each varied by two levels, each in a between subjects design. This provided a 2x2x2 design with a total of eight experimental groups.

FOR was varied by two levels: high and low. In the low FOR conditions, the test application only used the CAVE's front wall, providing a 90 degree horizontal FOR and a 90 degree vertical FOR. The high FOR conditions used all four of the CAVE's screens (three walls and the floor), providing a horizontal FOR of 270 degrees and a vertical FOR of 180 degrees. We note that these measurements of FOR are approximate, as the exact FOR can vary with physical translation within the VE and the walls of the CAVE were not perfect squares.

Stereoscopy was varied by two levels: stereoscopic and monoscopic rendering. Active shutter glasses enabled stereoscopic viewing. In order to maintain consistent brightness and field of view among all conditions, all participants (regardless of condition) wore the shutter glasses. The shutter glasses limited FOV to approximately 100 degrees.

Two levels of head tracking were also controlled: head-tracked or not head-tracked. The shutter glasses were tracked to allow head-tracked rendering of the 3D structures, allowing participants to use physical bodily movements (as is possible in the real world) to adjust the view of the model. Note that controlling head tracking as an independent variable did not preclude the use of motion cues for the spatial judgment task because participants in all conditions were able to use the wand's joystick for navigation. Thus, this experiment tested for performance differences based on the supported method of view control, rather than the presence or absence of motion cues.

3.6 Procedure

Participants were introduced to the CAVE environment, taught how to use the wand to navigate through the VE, and briefed on the immersive features provided in the condition. For participants in the head-tracked conditions, this introductory period included explicit instruction of how physical movements (i.e., walking, leaning, crouching, turning) could be used to change the point of view. After the familiarity session, the experimenter then explained the task with the aid of an example cave-structure. The experimenter pointed out examples of vertical tubes that connected horizontal levels and of those that did not, and explained the goal of determining the number of connection tubes as quickly as possible without sacrificing accuracy.

Following this explanation, all participants completed a

practice task using the example structure model. After providing their responses, the experimenter informed the participants of the correct number of connections and showed participants their locations in the structure. After the practice task, participants completed ten trials (all in the same experimental condition, per the between-subjects design). For each trial, participants verbally reported the number of tube connections. The experimenter instructed participants to complete each trial as quickly as possible without sacrificing accuracy. For these ten trials, the experimenter provided no additional feedback regarding performance.

3.6 Results

The collected trial data contained several outlier trials with exceptionally high error levels. Before analysis, we removed trials with error values beyond three standard deviations from the mean, removing 1.5% of all trials. Outliers were distributed among conditions. After outlier removal, the remaining errors and times were averaged to generate the performance metrics for each participant. Because each participant performed ten trials and had at most two outlier trials removed, outlier handling did not remove any complete conditions or eliminate any participants.

Average error values for all conditions are summarized in Table 1. The condition with high FOR, head tracking, and stereo had the overall lowest average error. Table 2 presents average task times for all conditions. Note that the conditions with head tracking and stereo had by far the best average times.

The data met the assumptions of ANOVA (analysis of variance) testing for statistical analysis. Shapiro-Wilk tests of normality suggested that both time and error data were normally distributed, and the results of Levene's tests showed homogeneity of variance across conditions for both metrics. Participants' metrics were independent, as study sessions were conducted individually with a between-subjects design.

We analyzed the error data with an independent factorial ANOVA to test for differences due to FOR, stereo, and head tracking, as we for interactions. There was a significant main effect of head tracking on task errors, with $F(1, 44) = 4.54$ and $p < 0.05$. The number of errors with head tracking (M = 0.66, SD = 0.38) was significantly less than the number of errors without head tracking (M = 0.87, SD = 0.39). The test also indicated a significant main effect of FOR on errors, with $F(1, 44) = 8.95$ and $p < 0.01$. Significantly fewer errors were made in the high FOR conditions (M = 0.61, SD = 0.37) than in the low FOR conditions (M = 0.92, SD = 0.37). These results support our hypotheses of the effects of FOR and head tracking on performance. Either increasing FOR or adding head-tracked viewing was enough to improve performance, significantly reducing errors in the spatial judgment task.

No significant effects on task errors was found for stereo, with $F(1, 44) = 0.44$ and $p = 0.51$. No significant interactions were found for errors, with $F(1, 44) = 1.43$ and $p = 0.24$ for the interaction between head tracking and stereo, $F(1, 44) = 1.89$ and $p = 0.18$ for the interaction between FOR and head tracking, $F(1, 44) = 0.07$ and $p = 0.80$ for the

interaction between stereo and FOR, and $F(1, 44) = 0.93$ and $p = 0.34$ for the interaction among all three components.

We also tested for effects on task time with an independent factorial ANOVA. There was a significant main effect of head tracking on time, with $F(1, 44) = 9.15$ and $p < 0.01$. We also found a significant main effect of stereo on time, with $F(1, 44) = 7.73$ and $p < 0.01$. The analysis also revealed a significant interaction between head tracking and stereo, with $F(1, 44) = 5.43$ and $p < 0.05$, which explained the significant effects of these components indi-

TABLE 1
Mean Errors

	Stereo		Mono (no stereo)	
	Head tracking	No head tracking	Head tracking	No head tracking
High FOR	0.36	0.92	0.53	0.68
Low FOR	0.92	1.02	0.85	0.90

Lower numbers indicate better performance than higher numbers. Both head tracking and higher FOR caused significantly fewer errors. Also note that the condition with high FOR, head tracking, and stereo had the overall lowest average error (though interactions were not statistically significant).

TABLE 2
Mean Task Times

	Stereo		Mono (no stereo)	
	Head tracking	No head tracking	Head tracking	No head tracking
High FOR	47.98	79.32	75.07	79.31
Low FOR	41.73	76.43	77.68	81.99

Lower numbers indicate better performance than higher numbers. Note that the conditions with head tracking and stereo had significantly faster average times than the other six conditions
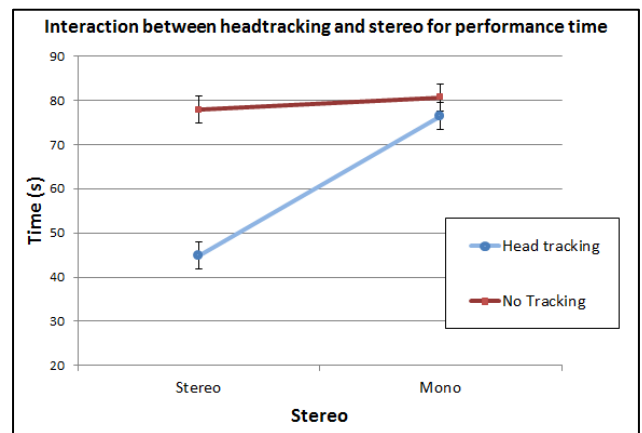

Fig. 4. Significant interaction between head tracking and stereo for task completion times. Stereoscopy and head tracking support significant speed increases when used together, but provide little benefit individually.

vidually. Fig. 4 shows this interaction. A post-hoc Tukey HSD analysis showed that the combination of stereo with head tracking allowed significantly faster performance

than other conditions. This interaction suggests that the addition of either stereo or head tracking individually provided little benefit for performance speed, but a much greater improvement was achieved when both were used together. These results support our hypothesis that the combination of both stereo and head tracking support better performance than either individually, but we reject the hypothesis that stereo alone is enough to improve performance when making small-scale spatial judgments.

We found no significant effect of FOR on task time, with $F(1, 44) = 0.24$ and $p = 0.88$. Interaction effects were not significant between FOR and head tracking for time, with $F(1, 44) = 0.02$ and $p = 0.89$, nor between FOR and stereo, with $F(1, 44) = 0.34$ and $p = 0.56$. There was also no significant interaction among all three components for time, with $F(1, 44) = 0.02$ and $p = 0.89$, so we reject the hypothesis of a three-way interaction.

4 DISCUSSION

Overall, the results do suggest that the addition of the higher-fidelity display features support performance improvements in distinguishing between structural gaps and intersections. The condition with the highest level of overall fidelity, having high FOR, head tracking, and stereo, had the fewest errors among all conditions. Further, by individually controlling each of these features, we are able to further dissect their effects.

4.1 Interpreting the Effects of Display Components

From a practical standpoint, the results of this experiment suggest that either increasing the FOR of a display or enabling both head tracking and stereo can improve a user's ability to identify gaps and collisions. By considering how different combinations of display components correspond to actual real-world types of displays, it is possible to generalize the results of this experiment to other systems. While high FOR, head tracking, and stereo were supported in a CAVE for this study, similar features could also be supported in a stereo-enabled HMD with head-tracked rendering. Of course, this is not a perfect comparison, as HMDs are different from CAVE-type displays in other ways. For example, the distance from the display screen to the eye, latency in updating the display with physical head rotation, or the weight on the head all vary between HMDs and CAVE-type displays. Still, the experimental results demonstrate that increasing the levels of fidelity for these components can positively affect performance, and these results should be taken into consideration when selecting a display for a particular purpose.

As both HMDs and large-screen displays can allow a higher level of fidelity than is supported by a standard computer monitor, the results of our experiment provide evidence that either of these more immersive displays could provide benefits for small-scale spatial judgment tasks. Through a greater understanding of how different system features affect task performance both individually and in combination with other components, it becomes possible to design VR systems that provide the best ratio of benefit to cost. For example, high FOR allows users to use physical head or body rotations to view virtual content. In VR systems, the FOR is affected by the number of display screens surrounding a user or orientation tracking on an HMD. The experiment showed that increasing the FOR significantly reduced errors in a surround-screen display. This result suggests that the ability to physically rotate to control the view of the VE is important to achieve accurate high-precision spatial judgments. This result is consistent with previous studies that have found evidence that physical rotation makes it easier for participants to maintain orientation in the environment [4, 34]. We believe that improved orientation and easier view control allowed for a more thorough inspection of the structure, decreasing task errors. We suspect that the additional display screens made it easier to judge structures quickly and easily from different angles or from different viewpoints. Having a larger display surface reduced the amount of precision needed with virtual movement from the joystick, since larger virtual movements could be made while still keeping the structure visible within the display area.

The addition of head-tracked view control allowed users to adjust the rendering based on the perspective of their actual head locations. We believe that it was the positional head tracking (rather than the rotational tracking) that provided the greatest benefits for the small-scale spatial judgment task. Positional head-tracked rendering allowed users to physically walk, lean, or crouch to adjust the view of the structures, significantly reducing errors. This effect cannot be attributed to the addition of motion cues alone because participants in all conditions were able to move with joystick navigation. Rather, this result was due to the additional method of view control provided. We believe that head tracking improved performance by allowing participants to use the same types of physical movements as used in everyday life to control viewing and achieve motion parallax, allowing an easier and more intuitive spatial investigation. This familiar, physical method of view control may have reduced mental workload by relieving the attention to wand and joystick operations necessary to control viewing.

The time results also showed that head tracking was able to help reduce task completion times, but only when coupled with stereo. The interaction effect between stereo and head tracking also shows that stereo did not provide significant speed benefits without the use of head-tracked viewing. This suggests that depth cues from both stereo and head-tracked viewing were needed to quickly make the difficult depth judgments required in the experimental task. This result demonstrates the importance of having multiple depth cues—in this case, the binocular disparity provided by stereo and the motion parallax from head tracking—for efficient spatial processing. Considering the significant effect of head tracking on task errors, the results indicate that participants were able to make small-scale spatial judgments accurately using head tracking without stereo. However, participants could not make these accurate judgments more quickly, as they needed to spend time moving around the potential collision area in order to find an effective point of view (such

as that shown in Fig. 3). We suspect that the addition of binocular disparity provided by stereoscopy helped to reduce the amount of movement necessary to make accurate judgments, improving performance times.

4.2 Comparisons to Previous Work

While this study reveals benefits of increased display and interaction fidelity for one type of spatial understanding task (spatial judgments of small-scale spatial features), it is important to note that it is not guaranteed that any type of spatial task will be affected by the system's FOR, stereo, and head-tracked rendering. For instance, consider the observed effect of head tracking in our study, in which the addition of head tracking to standard wand-based flying significantly decreased task errors. It is interesting to note how this result differs from that of a previous study by Ware and Franck [1], in which it was observed that the method of controlling motion cues (head-tracked or hand-controlled) made no significant difference in tracing paths in 3D node graphs. The task of that study, however, probably had less emphasis on high-precision spatial judgments. Performance on the path-following task was probably more dependent on the ability to correctly perceive graph shapes and distinguish between separate components of the 3D structures. In contrast, in the task of our study, the components of the structure (that is, the vertical and horizontal tubes) could be easily perceived. The challenge in our task was identifying collisions or gaps between the components. Thus, in our task, motion parallax may have been more important for distinguishing between structural components. The results suggest that head tracking helped participants to more easily control these motion cues and interpret their meaning for spatial objects.

As another example, Barfield, Hendrix, and Bystrom [9] studied the effects of head tracking and stereo on the ability to understand the overall shape of 3D bent-wire structures, but found no significant performance differences. This study involved a 3D-to-2D projection task, requiring participants to select the correct 2D sketch of the wire that corresponded to the 3D structure. The researchers did observe the worst overall performance in conditions with neither head tracking nor stereo, and their results did show performance gains from the addition of head tracking or stereo, but these findings were not statistically significant. They hypothesized that structures may have been too simple for viewers to benefit from the additional spatial cues. They also posited that the projection task may have been too difficult to observe significant performance differences. This example demonstrates that increasing display or viewing interaction fidelity may not necessarily improve performance just because the task is spatial in nature. Specifics of the type of spatial task, the level of complexity of the spatial structures, or the degree of task difficulty could certainly affect the results. We saw this in our earlier study [15], in which the higher levels of display and viewing interaction fidelity significantly improved performance on some spatial tasks, but not others.

Specifically, our previous study [15] did not detect any performance benefits of the high-fidelity condition (stereo, 270° horizontal FOR, and head tracking) over the low-fidelity condition (no stereo, 90° horizontal FOR, and no head tracking) for a few spatial understanding tasks, including 3D projection and simple feature search. In contrast to the difficult projection task in the study by Barfield, Hendrix, and Bystrom [9], this projection task may have been too simple to be affected by the display components [15]. The feature search tasks, which involved judgments about relatively large portions of the 3D structure, were also relatively simple. Consequently, it was easy to investigate the large or overall structural features even without the additional spatial cues provided by the high-fidelity condition.

The high-fidelity condition did evoke significantly better performance for three spatial understanding tasks: collision search and identification, small feature searching, and relative size comparisons [15]. The collision identification task was somewhat similar to the task of the study presented in this paper, involving the identification of colliding structures; however, the task in the previous study was much more dependent on searching for potential collisions before making the spatial judgments. In the new study, the more organized tube structure made the searching trivial, thus shifting the focus to making spatial judgments. At any rate, the results of the previous study for this task do agree with those of our newer study.

Higher fidelity also improved performance searching for small structural details (pits in the cave environment) in the prior study. Since the task was primarily concerned with spatial search, we hypothesize that the increased FOR helped scanning efficiency by allowing physical rotation, and that both stereo and head tracking made it easier to identify spatial features of interest. Significant performance benefits were also found for relative size and distance comparisons, in which participants were required to compare the sizes of components of the 3D model or to determine the shortest path through the cave from one point to another. We suspect that these tasks benefited from increased FOR and head tracking for quick and efficient view control, which would make it easier to compare structures and inspect the spatial layout. Stereo could also be especially useful in conjunction with head tracking to improve spatial perception during these comparisons. Though combined benefits were observed, further research with independent control of the system components would be needed to test these hypotheses—just as we did with the focused analysis of the effects of FOR, stereo, and head tracking on collision identification tasks in our new study.

4.3 Perceptual Cues and Spatial Tasks

With the results that we currently have from our work (this study and Schuchardt and Bowman [15]) and that of others [e.g., 1, 9, 10], we can say that the effects of immersive VR technology appear to be more noticeable when the spatial judgment tasks involve precise and careful inspections, rather than overall or larger-scale shape analyses, which can many times be performed without the need for enhanced spatial cues.

This conclusion also agrees with what we know of the benefits of individual system features. For instance, because a high FOR supports natural view rotation through physical turning, it seems obvious that spatial tasks will benefit more when the viewpoint is within objects or structures. In these cases, in which the spatial structure takes on a larger scale relative to the user, the advantages of easy and efficient rotation is certainly greater than in situations where the majority of the spatial content can be viewed within a limited FOV.

We can also consider what we know about stereoscopy and motion parallax in our interpretation of the types of tasks and spatial inspections that might greater benefit from more immersive displays. While research has provided evidence that binocular disparity can enable improved distance perception for distances as far as 40 meters away [38], the usefulness of stereo for practical spatial tasks is generally limited to 10 meters, and is best within one meter [39]. It follows, then, that stereo would be most helpful for close-range spatial inspection tasks (which our task was, as participants moved up close to potential collision areas in order to make judgments).

Of course, since head tracking complements stereo for the combined benefits of both motion parallax and more realistic retinal disparity from the stereo [1, 5], the addition of motion parallax would also be useful for close range tasks. Additionally, because motion parallax alone is useful for judging object positions in larger areas [39], head tracking could still provide benefits for spatial understanding tasks in larger environments or with larger structures. However, since other types of motion besides head tracking can also enable motion parallax, it is not certain that head tracking would definitely provide benefits in any situation. Thus, system designers should consider the specifics of the task (in addition to the scale of the spatial structures and the complexity of the structure layout) when selecting a display.

4.4 Task Specificity and Generalization

Though task dependence is always an issue when interpreting experimental results, since the inspection task in our experiment is common in many domains, our results can be generalized to a variety of applications. As we previously described, our spatial judgment task focused on high-precision inspections of specific spatial features. The task required participants to first identify potential collision areas for static tube structures and then to decide whether or not the structures were touching or intersecting each other. This type of spatial inspection task is important for real-world applications. For example, in architecture and construction, visualizations are used to help plan plumbing, electrical wiring, and structural components while avoiding collisions or intersections [13]. Judging gaps and intersections would also be important when using virtual environments for structural design [e.g., 12, 40, 41]. In visualizations for oil well-path planning, similar spatial judgments are used to prevent collisions [11]. As another example, scientists identify model intersections and tight spaces when working to better understand complex protein structures through immersive scientific visualizations [14].

While the type of difficult, small-scale spatial judgments that we focused on in our study help make our experimental results somewhat generalizable, application domains would certainly be interested in other types of spatial understanding as well. We note that results could still vary for other types of spatial judgments (e.g., size comparisons or spatial projections), and additional research on other spatial tasks is needed to know for sure.

Further research is also needed into how characteristics of the structures themselves influence the effects of the display components. For instance, perhaps the benefits of increased levels of fidelity could be more pronounced with complicated structures with many intricate spatial cavities and protrusions, but more simplistic structures could be perceived and understood well enough with lower levels of fidelity. Similarly, it could be that the effects for regions that are densely packed with spatial structures would be different from shapes with more open spaces. The types of curvature or sharpness of angles within the structure could also be characteristics of interest. Are smoother, more organic structures affected by varying fidelity differently than objects with a higher frequency of right angles and straight edges? While the structures of the presented study were based on underground cave systems, the models were much more regular than those of the previous study [15]. The structures in this study were clearly divided into separate horizontal levels and intersections between pathways were always close to 90 degrees (see Fig. 2). In contrast, the models of the previous study [15] had greater variety in angles and shapes of their structural components (see Fig. 1).

5 CONCLUSIONS AND FUTURE WORK

While many controlled studies have investigated different types of spatial understanding tasks, it is still important to consider the effects of system characteristics on specific tasks relevant to real-world applications. In this research, we independently controlled FOR, stereoscopy, and head-tracked rendering to study their effects on the performance of a task involving small-scale spatial inspections of 3D structures. Measuring time and errors, we asked participants to distinguish between structural gaps and intersections between components of a model that was designed based on underground cave systems. The analysis of the results shows that participants made significantly fewer errors with either an increased FOR or with the addition of head-tracked rendering. The results also indicate that participants performed significantly faster when the display provided both stereo and head-tracked rendering. Additionally, the results show that the condition with high FOR, head tracking, and stereo had fewer errors than all other conditions, while the condition with low FOR, no head tracking, and no stereo had the overall worst average time (though this last difference was not significant).

Achieving a greater understanding of how different system features affect performance both individually and in combination with other features will contribute to the

design knowledge of how to select an appropriate VR system for a given purpose. Controlling individual components of display or interaction fidelity on a single system provides a means of simulating the fidelity of other display systems. The results of our study show that increasing the levels of fidelity for these components can positively affect performance. By considering what high-fidelity features are provided by different displays, we interpret our results as evidence that both HMDs and surrounding large-screen displays could improve performance over standard computer monitors. However, because other factors vary among different displays (e.g., resolution, brightness, form factor), additional work is needed to validate how well a given display setup can be used to simulate other displays.

ACKNOWLEDGMENT

REFERENCES

[1] C. Ware, and G. Franck, "Evaluating stereo and motion cues for visualizing information nets in three dimensions," *ACM Trans. Graph.,* vol. 15, no. 2, pp. 121-140, 1996.

[2] A. E. Richardson, D. R. Montello, and M. Hegarty, "Spatial knowledge acquisition from maps and from navigation in real and virtual environments," *Memory & Cognition,* vol. 27, no. 4, pp. 741-750, 1999.

[3] D. Waller, E. Hunt, and D. Knapp, "The Transfer of Spatial Knowledge in Virtual Environment Training," *Presence: Teleoperators and Virtual Environments,* vol. 7, no. 2, pp. 129-143, 1998.

[4] S. S. Chance, F. Gaunet, A. C. Beall, and J. M. Loomis, "Locomotion Mode Affects the Updating of Objects Encountered During Travel: The Contribution of Vestibular and Proprioceptive Inputs to Path Integration," *Presence: Teleoperators and Virtual Environments,* vol. 7, no. 2, pp. 168-178, 1998.

[5] C. Ware, K. Arthur, and K. S. Booth, "Fish tank virtual reality," in Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems, Amsterdam, The Netherlands, 1993, pp. 37-42.

[6] D. A. Bowman, and R. P. McMahan, "Virtual Reality: How Much Immersion Is Enough?," *Computer,* vol. 40, no. 7, pp. 36-43, 2007.

[7] M. Slater, "A note on presence terminology," *Presence connect,* vol. 3, no. 3, 2003.

[8] R. P. McMahan, D. A. Bowman, D. J. Zielinski, and R. B. Brady, "Evaluating Display Fidelity and Interaction Fidelity in a Virtual Reality Game," *IEEE Transactions on Visualization and Computer Graphics,* vol. 18, no. 4, pp. 626-633, 2012.

[9] W. Barfield, C. Hendrix, and K. Bystrom, "Visualizing the structure of virtual objects using head tracked stereoscopic displays," in Proceedings of the 1997 Virtual Reality Annual International Symposium (VRAIS '97), 1997, pp. 114.

[10] Prabhat, A. Forsberg, M. Katzourin, K. Wharton, and M. Slater, "A comparative study of desktop, fishtank, and cave systems for the exploration of volume rendered confocal data sets," *Visualization and Computer Graphics, IEEE Transactions on,* vol. 14, no. 3, pp. 551–563, 2008.

[11] K. Gruchalla, "Immersive Well-Path Editing: Investigating the Added Value of Immersion," in Proceedings of the IEEE Virtual Reality, 2004, pp. 157.

[12] A. S. Watson, and C. J. Anumba, "The need for an integrated 2D/3D cad system in structural engineering," *Computers & Structures,* vol. 41, no. 6, pp. 1175-1182, 1991.

[13] A. Khanzode, M. Fisher, and D. Reed, "Challenges and benefits of implementing virtual design and construction technologies for coordination of mechanical, electrical, and plumbing systems on large healthcare project," *Proceedings of CIB 24th W78 Conference,* 2007, pp. 205–212.

[14] N. Akkiraju, H. Edelsbrunner, P. Fu, and J. Qian, "Viewing Geometric Protein Structures From Inside a CAVE," *IEEE Comput. Graph. Appl.,* vol. 16, no. 4, pp. 58-61, 1996.

[15] P. Schuchardt, and D. A. Bowman, "The benefits of immersion for spatial understanding of complex underground cave systems," *Proceedings of the 2007 ACM symposium on Virtual reality software and technology,* ACM, 2007, pp. 121–124.

[16] M. Slater, "Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments," *Philosophical Transactions of the Royal Society B: Biological Sciences,* vol. 364, no. 1535, pp. 3549-3557, 2009.

[17] E. J. Gibson, J. J. Gibson, O. W. Smith, and H. Flock, "Motion parallax as a determinant of perceived depth," *Journal of Experimental Psychology,* vol. 58, no. 1, pp. 40-51, 1959.

[18] B. Rogers, "Motion parallax as an independent cue for depth perception," *Perception (London),* vol. 8, no. 2, pp. 125-134, 1979.

[19] B. Rogers, and M. Graham, "Similarities between motion parallax and stereopsis in human depth perception," *Vision Research,* vol. 22, no. 2, pp. 261-270, 1982.

[20] M. Mon-Williams, J. P. Wann, and S. Rushton, "Binocular vision in a virtual world: Visual deficits following the wearing of a head-mounted display," *Ophthalmic and Physiological Optics,* vol. 13, no. 4, pp. 387-391, 1993.

[21] J. P. Wann, S. Rushton, and M. Mon-Williams, "Natural problems for stereoscopic depth perception in virtual environments," *Vision Research,* vol. 35, no. 19, pp. 2731-2736, 1995.

[22] E. B. Johnston, B. G. Cumming, and M. S. Landy, "Integration of stereopsis and motion shape cues," *Vision Research,* vol. 34, no. 17, pp. 2259-2275, 1994.

[23] P. B. Hibbard, and M. F. Bradshaw, "Isotropic integration of binocular disparity and relative motion in the perception of three-dimensional shape," *Spatial Vision,* vol. 15, no. 2, pp. 205-217, 2002.

[24] R. A. Jacobs, "Optimal integration of texture and motion cues to depth," *Vision Research,* vol. 39, no. 21, pp. 3621-3629, 1999.

[25] D. Buckley, and J. P. Frisby, "Interaction of stereo, texture and outline cues in the shape perception of three-dimensional ridges," *Vision Research,* vol. 33, no. 7, pp. 919-933, 1993.

[26] J. M. Hillis, S. J. Watt, M. S. Landy, and M. S. Banks, "Slant from texture and disparity cues: Optimal cue combination," *Journal of Vision,* vol. 4, no. 12, December 1, 2004.

[27] F. Domini, C. Caudek, and H. Tassinari, "Stereo and motion information are not independently processed by the visual system," *Vision Research,* vol. 46, no. 11, pp. 1707-1723, 2006.

[28] K.-I. Tsutsui, M. Jiang, H. Sakata, and M. Taira, "Short-Term Memory and Perceptual Decision for Three-Dimensional Visual Features in the Caudal Intraparietal Sulcus (Area CIP)," *The Journal of Neuroscience,* vol. 23, no. 13, pp. 5486-5495, July 2, 2003.

[29] K.-I. Tsutsui, M. Jiang, K. Yara, H. Sakata, and M. Taira, "Integration of Perspective and Disparity Cues in Surface-Orientation–Selective Neurons of Area CIP," *Journal of Neurophysiology,* vol. 86, no. 6, pp. 2856-2867, December 1, 2001.

[30] K.-I. Tsutsui, M. Taira, and H. Sakata, "Neural mechanisms of three-dimensional vision," *Neuroscience Research,* vol. 51, no. 3, pp. 221-229, 2005.

[31] H. Sakata, K.-I. Tsutsui, and M. Taira, "Toward an understanding of the neural processing for 3D shape perception," *Neuropsychologia,* vol. 43, no. 2, pp. 151-161, 2005.

[32] M. Taira, K.-I. Tsutsui, M. Jiang, K. Yara, and H. Sakata, "Parietal Neurons Represent Surface Orientation From the Gradient of Binocular Disparity," *Journal of Neurophysiology,* vol. 83, no. 5, pp. 3140-3146, May 1, 2000.

[33] R. Pausch, D. Proffitt, and G. Williams, "Quantifying immersion in virtual reality," in Proceedings of the 24th annual conference on Computer graphics and interactive techniques, 1997, pp. 13-18.

[34] R. A. Ruddle, S. J. Payne, and D. M. Jones, "Navigating Large-Scale Virtual Environments: What Differences Occur Between Helmet-Mounted and Desk-Top Displays?," *Presence: Teleoper. Virtual Environ.,* vol. 8, no. 2, pp. 157-168, 1999.

[35] L. Arns, D. Cook, and C. Cruz-Neira, "The benefits of statistical visualization in an immersive environment," in IEEE Virtual Reality, Houston, TX, 1999, pp. 88-95.

[36] D. Raja, D. Bowman, J. Lucas, and C. North, "Exploring the benefits of immersion in abstract information visualization," *Proc. Of IPT (Immersive Projection Technology),* 2004.

[37] B. Laha, K. Sensharma, J. D. Schiffbauer, and D. A. Bowman, "Effects of Immersion on Visual Analysis of Volume Data," *IEEE Transactions on Visualization and Computer Graphics,* vol. 18, no. 4, pp. 597-606, 2012.

[38] S. Palmisano, B. Gillam, D. G. Govan, R. S. Allison, and J. M. Harris, "Stereoscopic perception of real depths at large distances," *Journal of Vision,* vol. 10, no. 6, 2010.

[39] C. Ware, *Information visualization: Perception for design*, San Francisco: Morgan Kaufman, 1999.

[40] D. A. Bowman, M. Setareh, M. S. Pinho, N. Ali, A. Kalita, Y. Lee, J. Lucas, M. Gracey, M. Kothapalli, Q. Zhu, A. Datey, and P. Tumati, "Virtual-SAP: An Immersive Tool for Visualizing the Response of Building Structures to Environmental Conditions," in Proceedings of the IEEE Virtual Reality 2003, pp. 243.

[41] F. Bacim, N. Polys, J. Chen, M. Setareh, J. Li, and L. Ma, "Cognitive scaffolding in Web3D learning systems: a case study for form and structure," in Proceedings of the 15th International Conference on Web 3D Technology, Los Angeles, California, 2010, pp. 93-100.

Eric D. Ragan is a PhD candidate in computer science at Virginia Tech. His research interests include 3D virtual environments, spatial information visualization, navigation techniques, and educational software. Ragan received an MS in computer science from Virginia Tech. Contact him at eragan@vt.edu.

Regis Kopper is a Post-Doctoral Associate with the Virtual Experiences Research Group at the University of Florida. His research interests include 3D user interfaces, virtual human interaction, novel interaction techniques, and large high-resolution displays. Kopper received his PhD from Virginia Tech. Contact him at kopper@cise.ufl.edu.

Philip Schuchardt is the founder and main developer of Cavewhere (www.cavewhere.com), an underground cave mapping software. Schuchardt received a BS in computer science from Virginia Tech.

Doug A. Bowman is an associate professor of computer science and director of the Center for Human-Computer Interaction at Virginia Tech. His research interests include 3D user interfaces and the effects of immersion in virtual reality. Bowman received a PhD in computer science from the Georgia Institute of Technology. He is a member of the IEEE Computer Society and the ACM. Contact him at bowman@vt.edu.