

Evaluation of X-ray visualization techniques for vertical depth judgments in underground exploration

Mustafa Tolga Eren¹  · Selim Balcisoy¹

Published online: 12 January 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract This paper investigates depth judgment-related performances of X-ray visualization techniques for rendering fully occluded geometries in augmented reality. The techniques we selected for this evaluation are careless overlay (*CO*), edge overlay (*EO*), excavation box (*EB*) and a cross-sectional visualization technique (*CS*). We have designed and conducted a comprehensive user study with 16 participants to examine and analyze the effects related to visualization techniques, having additional virtual objects and the scale of the vertical depths. To the best of our knowledge, this is the first user study on judged vertical depth distances that these techniques were compared against each other. We report our findings using four dependent variables: accuracy, signed error, absolute error and response time to shed some light into real-world performances and also to reveal estimation tendencies of each technique. Our findings suggest similar and better performance for *EB*, *CS* compared to *CO* and *EO*. We also observed significantly better results for *EB* and *CS* techniques when judging *Top* and *Bottom* distances compared to *Middle* distances. Derived from our findings, we proposed a new visualization technique for underground investigation with multiple views. The multi-view technique is our own implementation inspired by magic lens and cross-sectional visualizations with correlating displays.

Keywords Augmented reality · Depth perception · User study · X-ray visualization

✉ Mustafa Tolga Eren
tolgaeren@gmail.com

Selim Balcisoy
balcisoy@sabanciuniv.edu

¹ Sabanci University, Istanbul, Turkey

1 Introduction

Outdoor augmented reality is a wide research field with a large set of application areas spanning from defense to entertainment. A topic of interest for many researchers is the visualization of occluded objects and the preservation of spatial relationships between physical and rendered objects. Many of these studies are focused on displaying information hidden behind other surfaces such as walls, buildings and natural formations using X-ray visualization techniques [1, 24]. We are interested in evaluating the existing state-of-the-art visualization techniques while providing information on what is hidden beneath objects such as floors, streets and terrain.

In this paper, we aim to evaluate several techniques that can be used to visualize subterranean objects. There are several techniques on exploring existing urban infrastructure and archaeological artifacts such as ground-penetrating radar [32], radio frequency [25, 26] and electrical resistance tomography [5]. New and existing pipe networks and other geo-referenced subterranean data are documented using geographical information systems. We believe there is also a need for in situ visualization of documented data on a mobile device such as a smartphone or tablet in AR fashion.

The goal of our investigations was to compare the perceptual performance of four X-ray visualization techniques of different complexities: careless overlay (*CO*), edge overlay (*EO*) [1], excavation box (*EB*) [31] and a cross-sectional (*CS*) visualization techniques. All of four techniques have already been applied to underground exploration to a certain extent. All techniques except *CO* provide a sense of occlusion for subterranean objects. The latter two of these methods are task-specialized visualizations and provide graphical depth cues that can help to measure distances (*EB* and *CS*).

Although horizontal depth perception is a very active research area, vertical depth distance judgments is an under-investigated topic, especially for fully occluded geometries. We have designed and conducted a user study to examine and analyze these X-ray visualization techniques' perceptual performance through three hypotheses for identifying and comparing vertical depth distances at close range (0–1m). These hypotheses are related to:

1. The effect of X-ray visualization techniques.
2. The effect of having additional virtual objects.
3. The effect of the scale of vertical depth.

The evaluation of these four X-ray visualization techniques by a comprehensive user study on perceived vertical depths is our main contribution. Furthermore, we propose a new multi-view technique that is based on the results derived from the experiment. The multi-view technique is our own implementation inspired by magic lens [4] and cross-sectional visualizations [21].

2 Previous work

Depth perception is the recognition and interpretation of visual sensory stimuli to understand depth [12]. The human visual system utilizes multiple depth cues to derive a vivid three-dimensional perceptual world from two-dimensional retinal images of a scene [33]. Landy et al. [19] describe this procedure as cue theory and explain how depth cues interact and combine with each other. Lappin et al. [20] explain the influence of context to perceived distances by experimenting in different indoor and outdoor settings. Two comprehensive surveys, one on X-ray vision techniques [9] and another on evaluation of these techniques [23], are provided by Dey et al. and Livingston et al., respectively.

The notion of depth perception is studied extensively in AR and VR [7,18]. Jones et al. [14] provide a comparative analysis of egocentric depth perception between real world, VR and AR. They report conventional underestimation problem is considerably low in AR. Livingston et al. [22] compare AR depth perception in outdoor and indoor settings and analyze the effects of supplying user with linear virtual depth cues. They report that although they found evidence for underestimation in indoor, subjects overestimate depth values at outdoors.

X-ray visualization techniques are used for viewing occluded objects while preserving important features in an AR scene [28]. Exploding diagrams [15], ghosting [35] and cutaways [6] are examples of such techniques. Moreover, Diepstraten et al. [10] investigate cutaway visualization by utilizing artistic illustration techniques to enhance perception. Bane et al. [2] propose several tools of X-ray vision to

be used in AR context. Avery et al. [1] discuss how overlaying edge features of the occluding structure would give better depth cues to the viewer and describe three tools for further improving spatial perception. In the multi-view technique, we used a similar approach for promoting sense of occlusion for subterranean structures.

Zollmann et al. [35] employ ghosting techniques for solving single-layer occlusion problems between the surface and the infrastructure system. They use panoramic images from the viewed site for calculating a ghosting map and then use the features on this map to preserve the above ground context. Although they demonstrate occlusion clearly for a single layer of subsurface system, in the real-world subsurface systems may consist of multiple layers that are occluding each other.

In addition, Livingston et al. [24] propose an algorithm that solves multilayer occlusion problem on Z axis by changing the opacity values of virtual objects. The attacked problem is similar to ours but on a different domain of distinguishing occluded buildings. In our multi-view approach, we employed a second view to explicitly indicate separate layers instead of modulating opacity values of virtual objects. Elmqvist et al. [11] provide an in-depth taxonomy for occlusion management.

There are a number of X-ray visualization techniques addressing subsurface occlusion problems. Schall et al. [4,27,34] introduce an excavation tool inspired from magic lens techniques that virtually digs the ground letting viewer to see underground pipes [30]. This technique requires viewer to be close to the location to effectively perceive the hidden structure (see Fig. 1). In certain situations where there is a large distance between the observer and the excavation site, it may suffer from the long-flat view problem described in [16].

3 X-ray visualization overview

Occluded geometry visualization is studied extensively in the AR domain. We have selected to include four X-ray visualization techniques in our evaluation. A careless overlay (*CO*) of occluded geometry as shown in Fig. 1a is not sufficient for visualizing these objects [35]. Although simple, the inclusion of *CO* in the evaluation study is important to provide a comprehensive study for X-ray visualization. We believe we cover a large partition of X-ray visualization techniques and hope to both show and evaluate the evolutionary progress of the field.

Previous studies enhance the scene by employing X-ray visualization via using ghosting map or simple edge overlay techniques to give the sense of occlusion while visualizing hidden geometry [1]. Figure 1b presents an edge overlay technique (*EO*). Detected edges of the background image are

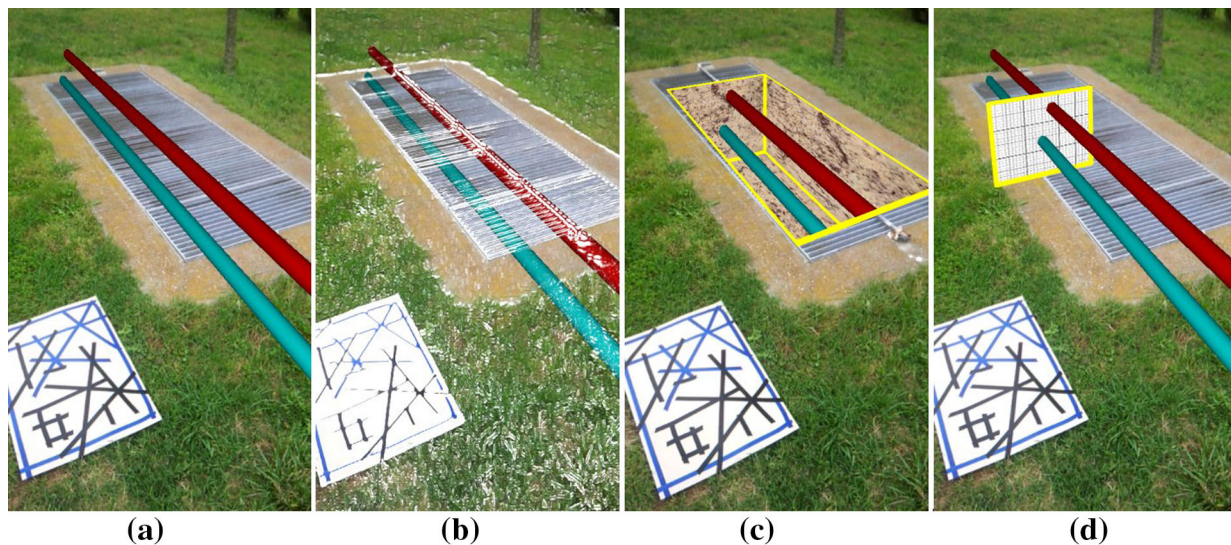


Fig. 1 Visualization of underground pipe networks using different techniques. **a** Careless overlay, **b** edge overlay, **c** excavation box and **d** cross-sectional techniques

overlaid on Top of infrastructure pipes. This technique provides user a depth cue for occluded objects. However, the user is not supplied with any focus cues. Without additional cues, observers are just presented with the information that the shown objects are behind (or in this case below) the object with edges overlaid. However, how much further or deeper the virtual object is located can only be guessed by the projected size of the virtual object.

In the Smart Vidente project, researchers utilized an excavation box (*EB*), to present focused visualization for an excavation site. We implemented a similar visualization that is shown in Fig. 1c. The rendering of underground objects is restricted to the volume covered by this rectangular excavation box. In this work flow, the excavation box is created and fixed to user-defined geo-location. This technique is tailored for examining a specific excavation location [30]. The excavation box technique provides a context for the virtual objects. Figure 1d demonstrates a fourth and the final visualization technique where only the back plane of an excavation box is drawn on the scene (*CS*). In this visualization, the plane is oriented to demonstrate the pipes cutting through itself to emphasize the pipe's spatial relation with respect to each other as well as the ground.

3.1 Vertical depth judgments

Distance judgments have been investigated heavily both in the AR and in the VR domain [22,33]. There are also studies that report X-ray visualization's effect on depth perception [1,7]. A common focus of these studies is the experiments were carried out to investigate horizontal distance judgments; "how far away is an object from the observer?" As an object

moves along the *Z* axis, two main difficulties may arise, namely perspective distortion and long-flat view problems. The perspective distortion problem can be defined as the optical illusion of projecting a distant large object and nearby smaller object to similar screen areas. Long-flat view problem is caused when viewing flat virtual objects at a distance [16,33].

On the other hand, vertical depth judgment is an underinvestigated topic, especially for fully occluded geometries. We believe there are two main issues for vertical distances: orientation of the problem and the occlusion of objects. Although perspective is a function of distance, the orientation of the objects' layout has been a point of interest. One of our motivations is to investigate the vertical displacement's effect on depth perception. Moreover, the occlusion of objects makes things more interesting and has not been investigated for vertical orientation. We hope to fill this gap in field. We believe the main issue worthy of further investigation in underground X-ray visualization domain is vertical depth judgments:

- Are state-of-the-art methods perceptually and numerically accurate for estimating vertical depths? Moreover, do they have estimation tendencies, such as over or underestimation? Furthermore, are these tendencies in line with horizontal distance estimations reported in recent studies?

In order to answer these questions, we have conducted a vertical depth judgment experiment with four visualization techniques, two conditions related to having additional objects and three depth intervals.

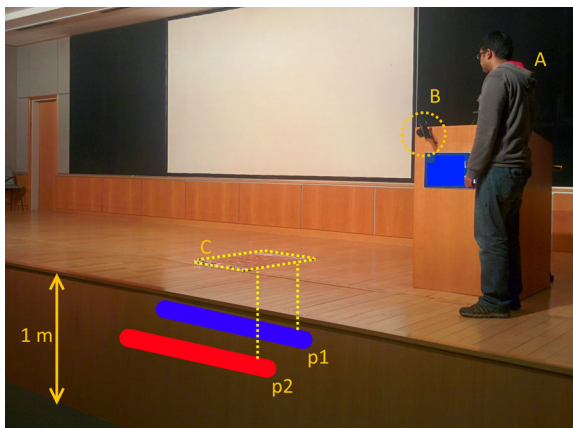


Fig. 2 Experiment scene and setup are demonstrated. *A* The participant, *B* fixed position and orientation of the mobile device, *C* the marker used for tracking, *p1*, *p2* two underground pipes

4 Experiment

4.1 Setup and task

We have performed our experiment in a controlled indoor setting, specifically a theater stage. Although our main concern is exploratory analysis in outdoor settings, the advantages of using a controlled environment (light, temperature, etc.) convinced us to perform the experiment in an indoor setting.

A mobile phone mounted on a stationary object was used as the primary interaction object. The height was fixed at 150 cm. The participants were asked to judge the vertical distance between an underground pipe and the ground as shown in Fig. 2. The scenes contained either one or two pipes, colored red or blue. There was a five-minute break in the Middle of the experiment (after 48 judgments).

A LG Nexus 5 smartphone running Android version 5.0 was used in all experiments. Average frame rate was around 30.

Underground pipes were placed about horizontally 2.5 m away from the participant. Diameter of the pipes was 10 cm. The localization was carried out using a marker that was placed on the ground. One-on-one training sessions were conducted for each participant. No participant reported having difficulty in understanding the task.

4.2 Participants

A total of 16 undergraduate- or graduate-level university students, of which 7 were female and 9 were male, participated in the study (age $\mu = 21.38$, $s = 2.50$). The height difference between participants was an issue of concern since if we allowed participants to control the position of the mobile device, they might be able to observe the scene from different heights and orientations. To overcome the height variance and

Table 1 Independent variables for the vertical depth experiment

Name	No. of levels	Description
Visualization	4	Careless overlay (<i>CO</i>), edge overlay (<i>EO</i>), excavation box (<i>EB</i>), cross-sectional visualization (<i>CS</i>)
Pipe count	2	1, 2
Vertical depth	3	Top, Middle, Bottom
Repetition	4	1, 2, 3, 4

disparities in viewing angles between observers, we attached the mobile device to a stationary object using a dock, fixing the viewing position and angle. Upon completing the tasks, the participants were offered prepaid gift cards to be used at campus coffee shop, approximately valued at four cups of coffee.

4.3 Independent variables

- Visualization Technique (within subjects):

Participants were presented with all of the four visualization techniques: careless overlay, edge overlay, excavation box and cross-sectional visualization (Table 1).

We have omitted additional depth cues for *EB* and *CS* such as layered textures and numerical depth scales in order to analyze effects of the inherently available depth cues in these methods. In *EB* technique, participants can judge distances utilizing the box's own dimensions as a depth cue. For the cross-sectional technique, the dimensions of rectangular plane can act as a depth cue for judging distances more accurately.

For each of the techniques, we supplied some verbal explanations: For *CO* and *EO*, we have acknowledged that the pipes are underground objects. For *EB* technique, participants were supplied with the real-world dimensions of the virtual dig box ($w \times h \times d = 1.0, 1.25, 0.85$ m). For *CS* group, we informed the participants with the functionality of the rectangle, as well as its corresponding real-world dimensions ($w \times h = 1.0, 0.85$ m).

- Pipe Count (within subjects):

For each task, the participant was shown either one or two pipes.

- Vertical Depth (within subjects):

The participants were asked to provide a judgment for the vertical distance between the ground and the pipe. This kind of vertical distance can also be referred as a type of egocentric distance, since the participants judged the vertical distances between their feet and shown objects. The depths were grouped into three sections: *Top*, *Middle*, *Bottom*. Each depth interval had four distinct values which were multiples of five.

Depth interval	Values
Top depths:	15, 20, 25, 30 cm
Middle depths	35, 40, 45, 50 cm
Bottom depths	55, 60, 65, 70 cm

The reasoning behind grouping the vertical depths is not trivial. For *EB* and *CS*, we expected the participants to judge distances by observing the relation between the sizes of the objects in the scene. Furthermore, one of our aims is to analyze this interaction with respect to placement of the pipe being closer to the Top or Bottom edge of the object. Grouping the vertical depth distances into these three categories enables us the investigate this interaction further.

- Repetition (within subjects):
Each condition was repeated two times.

4.4 Dependent variables

We report four quantitative variables as follows:

- Accuracy ACC: A normalized perceived signed distance metric is calculated as:

$$ACC = \left(1 - \left| \frac{PD - AD}{AD} \right| \right) \times 100 \tag{1}$$

where ACC is the normalized perceived distance accuracy, PD is the perceived distance and AD is the actual distance.

- Signed error: *SE* is the difference between *PD* and *AD*. We utilize this metric to analyze whether the conditions show under or overestimation tendencies.
- Absolute error is calculated as $|PD - AD|$. This metric is utilized to analyze sum of errors for each condition.
- Response time is the elapsed time between the start and end times of each case in millisecond. This metric also includes the time where the participants report their judgments verbally.

4.5 Controlled variables

- Variance in participants heights and movement of display: We have utilized a stationary and fixed height experiment setup as shown in Fig. 2. The height of the mobile device was fixed at 150 cm. The fixed viewport also allowed us to investigate differences in visualization rather than user experience.
- Environmental conditions: Approximately, half of the participants were invited in the mornings and other half in the afternoons. All experiments were conducted at the same place. The lighting was set to the same level for

each participant, and the brightness of the mobile device was adjusted appropriately.

- Orientation of pipes: In order to focus on effects of visualization techniques, we opted to display simple scenes. There was not any occlusion or collision between the pipes. The pipes were parallel to each other and perpendicular to the back plane of *EB* (similarly, perpendicular to the *CS* plane). We discuss the pipe placement procedure in the experiment design subsection in detail.
- Viewing angle: We chose to visualize the underground pipes parallel to each other as well as parallel to the viewing direction of the observer. In the *EB* technique, we have placed the box so that the underground pipes went through it (from the front plane to the back plane) perpendicularly. Similarly, in the *CS* technique the pipes went through the visualization plane perpendicularly.
- Experiment site: The physical features of the terrain play important role especially in one of the four visualization techniques (*EO*). We chose an experiment site that has adequate features for edge detection and visualization. In this particular theater stage, hardwood floors were used, and we adjusted the lighting to avoid any reflection from the floor. Experiment site is shown in Fig. 2.
- Discrete distances: The participants were informed that the possible depths for pipes are multiples of five and were asked to provide one of the following values:

<input type="checkbox"/> 0	<input type="checkbox"/> 5	<input type="checkbox"/> 10	<input type="checkbox"/> 15	<input type="checkbox"/> 20	<input type="checkbox"/> 25	<input type="checkbox"/> 30
<input type="checkbox"/> 35	<input type="checkbox"/> 40	<input type="checkbox"/> 45	<input type="checkbox"/> 50	<input type="checkbox"/> 55	<input type="checkbox"/> 60	<input type="checkbox"/> 65
<input type="checkbox"/> 70	<input type="checkbox"/> 75	<input type="checkbox"/> 80	<input type="checkbox"/> 85	<input type="checkbox"/> NA		

Verbally reported values are recorded by the experiment conductor. Although the actual range of the pipes was [15–70] cm, additional values are appended in the answer sheet.

4.6 Experiment design

The order of presentation was a concern both for visualization and for pipe count factors. In order to minimize these effects, we chose a latin square-based experiment design. We crossed these factors to create the following conditions:

Label	Visualization technique	Pipe count
A	Careless overlay	1-Pipe
B	Careless overlay	2-Pipes
C	Edge overlay	1-Pipe
D	Edge overlay	2-Pipes
E	Excavation box	1-Pipe
F	Excavation box	2-Pipes
G	Cross section	1-Pipe
H	Cross section	2-Pipes

The final latin square ordering can be seen here:

→	A	B	H	C	G	D	F	E
→	B	C	A	D	H	E	G	F
→	C	D	B	E	A	F	H	G
→	D	E	C	F	B	G	A	H
→	E	F	D	G	C	H	B	A
→	F	G	E	H	D	A	C	B
→	G	H	F	A	E	B	D	C
→	H	A	G	B	F	C	E	D

Each participant is assigned to a row and shown the conditions in order from left to right. There is also another issue that needs to be considered about data points related to pipe count factor. Conditions with 2-pipes provide two data points (one judgment for each pipe) and conditions with 1-pipe provide one data point. In order to create a balanced data set, we have opted to show 1-pipe conditions two consecutive times. For example, a participant assigned to the first row would be asked to provide judgments in the following order:

A, A, B, H, C, C, G, G, D, F, E, E

This ordering results in 16 depth judgments, 8 from conditions with 1-pipe ($[A, C, G, E] \times 2$) and another 8 from conditions with 2-pipes ($[B, H, D, F]$).

Moreover, we want to analyze these conditions over three depth intervals (Top, Middle and Bottom). To achieve this, we repeat the ordering three times. The participant observes her respective row from the latin square three times while the ordering of depth intervals as well as the possible values in each interval is randomized. We have enforced two constraints for this randomization:

- (i) All conditions must be observed for all depth intervals exactly one time, and
- (ii) The participant cannot be shown the same vertical depth value at the same time or consecutively.

This allows for a participant to observe each condition for all three depth intervals. To include our final factor, repetition, we simply repeat the whole process two times. In the second pass, we keep the orderings (including the depth interval orderings); however, we randomize the values within each depth interval.

The randomization approach and the pipe placement strategy are closely related. Two of the visualization techniques (*EB* and *CS*) have constraints on where the pipes can be placed for a meaningful measurement. For *EB*, the pipe or pipes should be located inside the box and similarly for *CS* the pipes should go through the measurement plane without touching the edges. Figure 3 demonstrates pipe

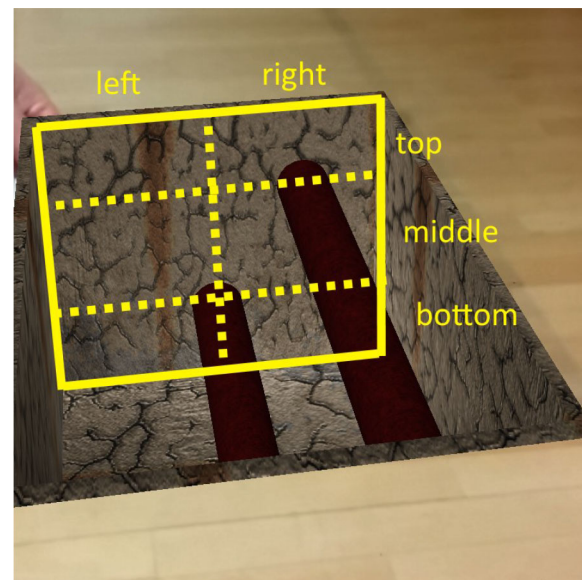


Fig. 3 Pipe placement options are shown on the excavation *box* visualization

placement options on a scene visualized with *EB*. The parameters for placing a pipe into the scene can be defined as: $\{x_i, y_i, c_i\}$ where x is the horizontal displacement, y is the vertical displacement and c is the assigned color for pipe i .

Our placement approach can be summarized as:

1. For each pipe, a side is randomly selected (either left or right). If the condition has 2-pipes, then both sides are used.
2. c is decided in this step. Red or blue color is randomly assigned to the pipe. If the condition has 2-pipes, the other color is assigned to the second pipe:

$$c_1 \neq c_2 \text{ and } c_1, c_2 \in \{\text{blue, red}\}$$

3. One of the *Top*, *Middle* or *Bottom* placement options is used with respect to depth interval variable of the specific condition. If the condition has 2-pipes, then the second pipe which is located on the other side of the plane is also placed on the same depth interval.
4. As shown in Fig. 4, x_1, y_1 and x_2, y_2 variables are the actual parameters that define the placement of the pipes. x_1, x_2 variables are randomly sampled, but should fulfill the following conditions:

$$r < x_1 < \frac{w_{\text{box}}}{2} - r$$

$$\frac{w_{\text{box}}}{2} + r < x_2 < w_{\text{box}} - r$$

where r is the radius of the pipe and w_{box} is the width of the excavation box (which is the same width as the plane in the *CS* visualization).

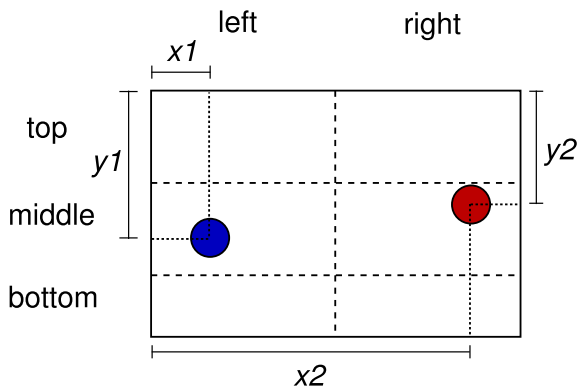


Fig. 4 Pipe placement variables are visualized

5. Similarly y_1, y_2 should comply to:

$$y_1 \neq y_2 \text{ and } y_1, y_2 \in d_i$$

where d_i is the set of values in the desired depth interval, $i \in \{Top, Middle, Bottom\}$. With these constraints, we guarantee that even the same interval is used for p_1 and p_2 , different y_1 and y_2 vertical depth values will be used.

6. Finally, if the condition has 2-pipes, then both pipes are shown with x_1, y_1, c_1 and x_2, y_2, c_2 parameters. If the condition has 1-pipe, firstly a pipe is shown with x_1, y_1 parameters, and after the participant has made a judgment, in a separate scene, the second pipe is visualized with x_2, y_2 parameters.

The number of data points can be computed as (with 16 participants):

$$\text{participant} \times \text{visualization} \times \text{pipe count} \times \text{depth interval} \times \text{repetition} = 16 \times 4 \times 2 \times 3 \times 4 = 1536$$

In summary, we have designed a three-factor (visualization, pipe count and vertical depth) repeated measures, within subjects experiment with latin square balancing.

4.7 Hypothesis

- [H1] Task-specialized techniques that have graphical depth cues (excavation box and cross section) will outperform other techniques in both accuracy and absolute error metrics.
- [H2] There will be more underestimated results compared to overestimated ones. Moreover, as the distance becomes larger the underestimation will increase. We expect our results to be in line with Dey et al.'s [8] observations for horizontal depth judgments.

[H3] Middle depth interval will produce larger absolute errors in EB and CS technique compared to Top and Bottom intervals.

5 Results

We analyzed our results with respect to our four defined dependent variables. Descriptive statistics for each measure is shown in Table 2.

Some of the participants had difficulties when judging distances using CO and EO techniques. In our questionnaire, the participants had the option to report "I cannot decide" (NA) for each judgment. Specifically one user reported NA option for 80 % of CO measurements. We chose to leave this user out of the experiment since it would create a bias toward this technique. For other participants, a total of 10.48 % of all judgments were reported as NA, 6.25 % for CO and 4.23 % for EO. For EB and CS, no participants reported NA option for any of the measurements.

5.1 Accuracy

We have run a three-way repeated measures ANOVA test and found visualization technique to be a main effect; $F(3, 45) = 944.618, p < 0.01, \eta_p^2 = .98$. The NA responses in CO and EO techniques were assumed to be the worst answer which was given in the available depth range. Post hoc Tukey HSD analysis for visualization techniques revealed that EB ($\mu = 88.94, s = 10.45$) and CS ($\mu = 83.13, s = 15.36$) techniques had significantly higher means; $CO < EO < CS < EB (p < .05)$ supporting [H1]. This was expected and was mainly due to graphical depth cues that are inherently available in those techniques but lacking in CO ($\mu = 16.33, s = 43.52$) and EO ($\mu = 36.27, s = 38.48$).

Pipe count was not observed to be a main effect with $F(1, 15) = 666.315, p > .05, \eta_p^2 = .05$.

On the other hand, depth interval was observed as a main effect; $F(2, 30) = 365.831, p < 0.01, \eta_p^2 = .96$. Post hoc tests revealed an ordering: $Middle < Top = Bottom (p < 0.05)$. We attribute this result to the participants' tendency to utilize visualization technique-related features in the scene. For example, for EB technique the back plane (where the pipes cut through) can provide a better understanding of the vertical depth distances. When the pipes are closer to the Top or the Bottom of this plane, the participants performed better as shown in Fig. 5.

We also observed an interaction between visualization and depth interval factors, $F(6, 90) = 45.08, p < 0.01, \eta_p^2 = .75$. Pairwise comparison revealed significant differences in all pairs except (EB, CS) at all depth levels ($p < 0.03$). We can interpret this result as EB and CS performing very similarly with respect to accuracy metric.

Table 2 Mean and standard deviation values of each dependent variable and corresponding visualization, pipe count and depth interval are shown

Visualization technique	Pipe count	Depth Interval	Dependent variable							
			Accuracy		Signed error		Absolute error		Response time	
			μ	std	μ	std	μ	std	μ	std
Careless overlay	1-Pipe	Top	-25.10	55.19	-26.09	8.57	26.09	8.57	10282.34	3177.85
		Middle	28.18	25.00	-30.00	9.47	30.00	9.47	9811.23	3130.78
		Bottom	44.23	10.66	-34.77	5.38	34.77	5.38	9367.52	3098.04
	2-Pipes	Top	-17.21	50.51	-24.30	7.28	24.30	7.28	9860.50	3308.49
		Middle	26.22	22.65	-30.70	8.16	30.70	8.16	10327.64	3649.15
		Bottom	41.64	9.65	-36.09	5.08	36.09	5.08	10493.45	3415.05
Edge overlay	1-Pipe	Top	0.26	42.95	-21.33	7.78	21.33	7.78	8780.95	2762.63
		Middle	48.99	21.96	-21.17	8.25	21.17	8.25	8718.36	3262.71
		Bottom	57.89	14.59	-26.09	7.94	26.09	7.94	8593.84	2700.60
	2-Pipes	Top	3.36	47.96	-20.23	7.94	20.23	7.94	9433.23	3051.36
		Middle	47.55	16.20	-21.95	6.34	21.95	6.34	9361.48	3630.20
		Bottom	59.57	12.73	-25.00	7.61	25.00	7.61	9178.52	3586.17
Excavation box	1-Pipe	Top	86.22	12.65	-2.89	2.49	2.89	2.49	8080.91	3044.68
		Middle	85.35	8.66	-6.02	3.47	6.02	3.47	9479.47	3088.79
		Bottom	96.19	4.19	-1.48	3.29	2.42	2.67	8082.28	3158.49
	2-Pipes	Top	84.79	12.42	-3.28	2.55	3.28	2.55	7997.89	2933.12
		Middle	85.92	9.46	-5.39	4.39	5.86	3.73	9971.30	3251.00
		Bottom	95.17	5.20	-2.19	3.77	2.97	3.18	9063.42	3513.14
Cross-sectional visualization	1-Pipe	Top	70.91	19.07	-5.78	3.36	5.94	3.07	7686.45	2580.94
		Middle	82.62	10.52	-7.34	4.79	7.50	4.54	8871.05	2532.15
		Bottom	95.17	4.07	-2.97	2.48	2.97	2.48	7672.28	3041.05
	2-Pipes	Top	75.47	17.56	-5.23	3.50	5.23	3.50	8695.08	3227.13
		Middle	79.79	10.48	-8.44	4.26	8.44	4.26	9977.42	2958.47
		Bottom	94.84	3.94	-3.20	2.42	3.20	2.42	8877.88	3109.36

5.2 Signed and absolute errors

An analysis on signed errors is used to evaluate under- or overestimation tendencies related to our factors. *NA* answers in *CO* and *EO* technique are omitted for this analysis. Overall, vertical depth distances were consistently underestimated with 1033 underestimations, 116 overestimations and 226 exact estimations. Observed estimation tendencies are in line with [H2]. Exact estimations were expected as the participants were asked to report in multiples of five and there were 18 distinct values for answers. We have run a separate three-way repeated measures ANOVA including *only EB* and *CS* techniques and found depth interval to be a main effect, $F(1, 15) = 25.08$, $p < 0.05$, $\eta_p^2 = .65$. A post hoc test revealed $Middle > Top = Bottom$, $p < 0.05$ for signed errors. For *EB* and *CS* techniques, participants performed significantly better for *Top* ($\mu = -4.30$, $s = 3.24$) and *Bottom* ($\mu = -2.46$, $s = 3.10$) intervals compared to *Middle* ($\mu = -6.80$, $s = 4.39$) depth interval as shown in Fig. 6.

For *CO* and *EO* techniques, we observe increasing underestimation for larger distances. This is a similar effect observed by Dey et al. [8] in horizontal orientation and greater distances.

As for absolute errors, boxplots for visualization technique and depth interval show lower errors for *EB* and *CS* techniques especially in *Top* and *Bottom* distances in Fig. 7. For *CO* and *EO*, we observe increasing errors with larger distances. Related to [H3], as expected we observed larger errors for *Middle* ($\mu = 6.95$, $s = 4.14$), compared to *Top* ($\mu = 4.34$, $s = 3.19$) and *Bottom* ($\mu = 2.89$, $s = 2.70$) in *EB* and *CS* techniques. For *EB* and *CS* techniques, the sum of errors for *Middle* was 1780 cms where the sum of absolute errors of *Top* and *Bottom* was 1110 and 740, respectively.

5.3 Response time

In design stage, we expected to observe a learning effect with respect to response count. Figure 8 shows a scatter plot for

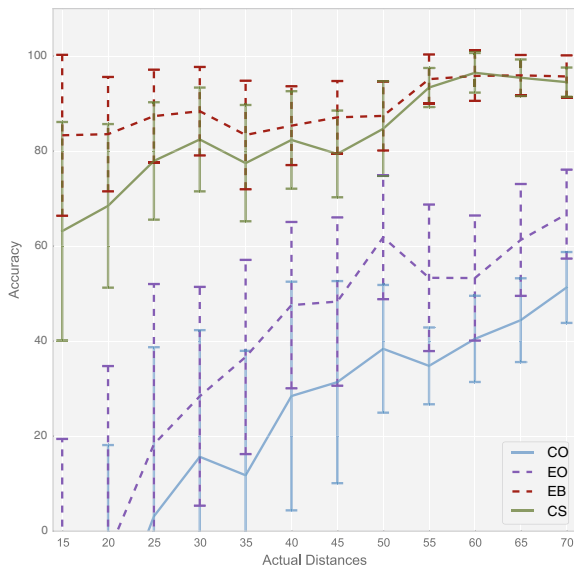


Fig. 5 Accuracy metric results are plotted through mean values for measurement distances on distances for vertical depth judgment experiment. In this metric, we observe EB and CS techniques distinguish themselves from the rest as expected, supporting [H1]. We contribute this difference to the depth cues that are inherently available in EB and CS. Error bars denote one standard deviation

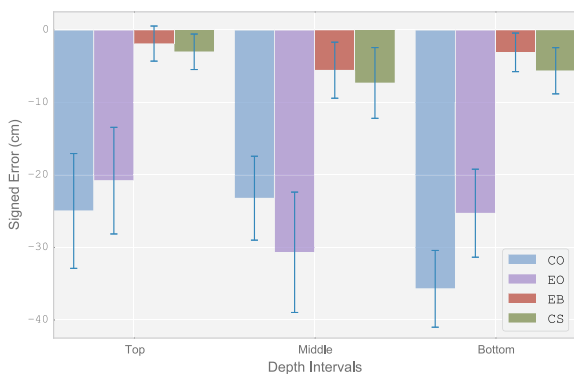


Fig. 6 Means of signed error values for depth intervals and visualization techniques are shown. Error bars denote one standard deviation. We observed generally increasing errors for CO and EO. On the other hand for EB and CS, the Middle distances resulted the highest errors

all trials, and a trend analysis indicates that the participants may responded faster as the experiment progressed. However, we did not observe a statistically significant learning effect. We attribute this lack of evidence to the task being simple enough where average response time was around 10 seconds.

The pipe count metric was observed as a main effect for response time, $F(1, 15) = 29.171, p < 0.01, \eta_p^2 = .66$. Participants responded faster for 1-pipe condition ($\mu = 8785.56, s = 3058.64$) compared 2-pipe ($\mu = 9436.48, s = 3361.01$) condition. It is possible to attribute this observation to (i) requiring to report one more measurement verbally or

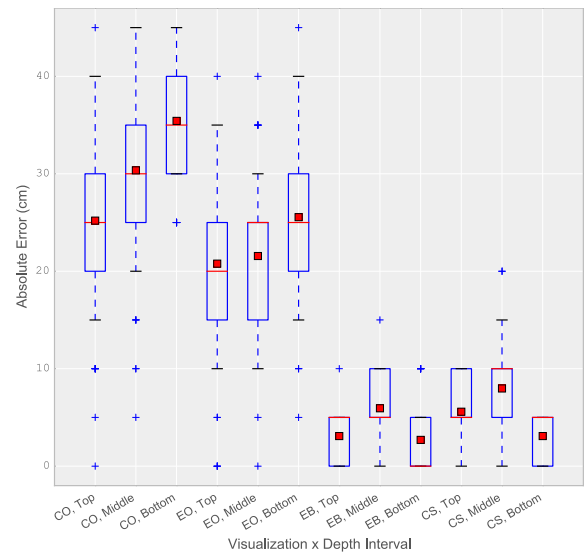


Fig. 7 Absolute errors observed from each technique is plotted for the experiment. Red line denotes median values, error bars represent one standard deviation, plus signs are outliers, and red squares are mean values

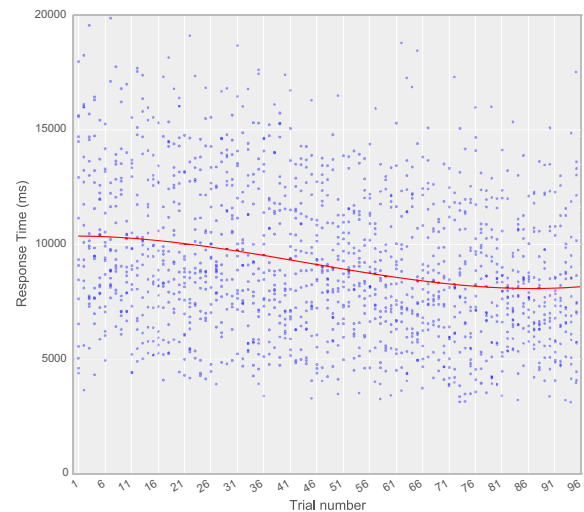


Fig. 8 Response times for each data point are visualized in a scatter plot. A trend line is shown with red

(ii) requiring a higher level of mental load for judging multiple objects.

We have also observed an interaction between visualization techniques and depth intervals, $F(3, 45) = 6.26, p < 0.05, \eta_p^2 = .3$. For EB and CS techniques, users responded to Top ($\mu = 8115.08, s = 2961.20$) and Bottom ($\mu = 8423.96, s = 3242.34$) depths faster compared to Middle ($\mu = 9574.81, s = 2986.88$) depths. We did not observe similarly significant differences for other techniques.

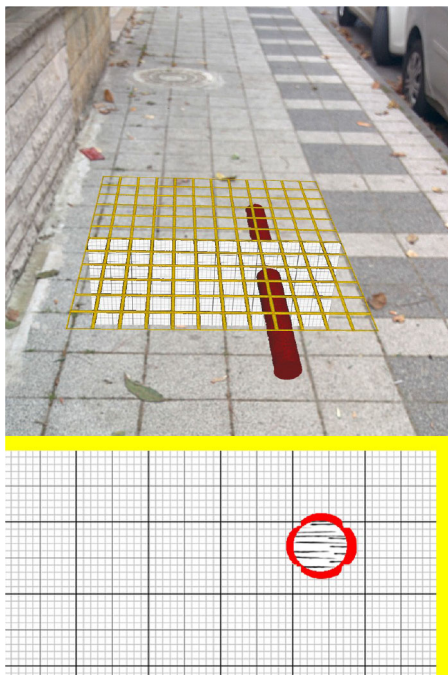


Fig. 9 Proposed multi-view technique is shown. The two correlating displays, one perspective (*Top*) and one orthographic (*Bottom*), visualize a pipe that is 12.5 cm below the ground. A ground grid is also added in order to create a spatial relation between the cutting plane and the ground

6 Discussion

Our investigation into X-ray visualization techniques having additional objects and several depth intervals led us to several conclusions.

We expected a statistically significant improvement when *EO* is utilized instead of *CO*. But our findings did not support this claim. Although the edges allowed participants to comprehend the spatial relations in the scene in a better way, there was no statistically significant difference between the performances of *CO* and *EO* for *ACC* metric. Participants chose the *NA* option for *EO* less frequently, leading us to believe *EO* can only be used to convey the information about the virtual objects being occluded. As no participant reported an *NA* measurement for *EB* and *CS* techniques and they provided better performance in nearly every condition, we believe they should be focused on for visualizing fully occluded objects, if the users are needed to measure vertical distances.

Having an additional pipe in the scene did not show a statistical difference for *ACC* and *SE* metrics. We have reported significance in response times; however, this can simply be attributed to having more than one objects to report. A more focused study on this phenomena can provide useful facts. Recently Jansen et al. [13] showed that multiple objects and variety of shapes can have profound effect on distance and scale judgments.

One of the more interesting results we have obtained is the fact that *EB* and *CS* perform very similarly in almost every condition. We have attributed this to several facts: (i) *CS* can be seen as a subset of *EB* where the planes of the box are removed except the back plane and (ii) when the participants judge distances, they make use of the planes that the pipes cut through (a feature exists in both *EB* and *CS*).

We derive three main conclusions from this finding. Firstly, given that the performances are similar, they may have different uses for appropriate situations. While *EB* provides a context for the AR scene and makes it easier to comprehend the spatial relations, it also occupies a large portion of the screen. It is possible that the box itself may occlude other important features or due to viewing angle and distance may hide the actual objects being investigated. On the other hand, *CS* visualization technique provides little context for the scene. It provides spatial relation between the virtual objects, but there is little information about the spatial relation between the virtual and real-world objects. *CS* does not occlude the scene or the objects. Depending on the needs of specific applications, more appropriate one can be selected. Dey et al. [9] also argue that a suite of complementary visualization should be provided to be used in different conditions. Consequently, we believe it is also possible to utilize both visualizations and perform a switch between them, depending on the task or viewing orientation.

Secondly, related to the usage of back plane for *EB* as a measurement tool, our investigation validates the usage of layered textures in later publications related to Smart Vidente project [30]. We have analyzed three depth intervals and found out that when there is a smaller distance between the underground objects and the Top or Bottom of the excavation box, the participants performed better. Using a texture with proportional layers may improve the distance judgments for Middle distances as well.

Thirdly, we believe for vertical depth judgments the most important feature is the intersection point between underground objects and the visualization objects. For our task, these features were the back plane of *EB* and rectangular visualization plane of *CS*. To emphasize this feature to a full extent, we propose to include an additional view that shows the cross section of said plane. We believe a multi-view approach has advantages for several reasons. During the exploration process, single-view approaches often lead to misinterpretation of the underlying data [29]. By displaying the same data using different techniques and from different angles, viewers are encouraged to match correlated elements. This approach can be used to overcome the depth perception shortcomings of three-point perspective visualizations. In our implementation, we have utilized two views: the first one is mainly for navigation where the second one is tailored to help users to judge distances more accurately.

The first view is a calibrated 3D virtual scene overlaid on Top of camera imagery. This view can be seen on the upper side of Fig. 9 and provides navigational support through the scene. Similar visualizations can be found in several studies in AR domain [3, 30]. Secondly, we generate a cross-sectional view which visualizes the subterranean layers as shown in the lower part of Fig. 9.

The orthographic view is generated with an orthographic camera that is positioned facing the rectangle located underground. This projection has a very narrow near to far-field range. We employ this narrow range in order to mimic volumetric cross-sectional visualization. The cross-sectional slice of the infrastructure that sits immediately below the anchor (above and under rectangles) is displayed on the orthographic view.

When the anchor position changes, both views change in a consistent manner. Orthographic visualization is a supplementary view to perspective view for providing a user with additional depth and shape cues which are consistent to the perspective view. The ideal usage of this visualization can be stated as follows: the user explores and navigates through the scene using the anchor in the upper view and then uses the second view to analyze the underground structures that are currently being cut by the rectangular plane.

These two views consistently visualize the same data with shape and color cues correlating with each other. Preserving spatial and *Focus and Context* relationships are crucial for improving the user's perception of the scene. In technical illustrations, 2D visualization techniques are favored [17]. Cross-sectional visualization of complex 3D objects is also investigated in AR context [21].

Although we have built the multi-view visualization technique on Top of CS, it is possible to use the secondary view as a complement to EB as well.

We have conducted our experiments through a marker-based visual tracking setup. Marker-based techniques are robust to environmental factors and much more suitable to create similar conditions between trials compared to markerless techniques such as SLAM-based approaches. The effect of tracking technique on vertical depth perception is an open question to be investigated further.

7 Conclusion and future work

We have designed and conducted a user study to evaluate four X-ray visualization techniques with different complexities, when judging vertical distances. Our observation had led us to come up with a set of design guidelines when working with underground objects. We observed and reported estimation tendencies which were previously investigated for horizontal orientation.

Derived from our findings we proposed a new visualization technique for underground investigation with multiple views. We plan to explore the performance of the multi-view technique in a future study.

References

1. Avery, B., Sandor, C., Thomas, B.H.: Improving spatial perception for augmented reality X-ray vision. In: Virtual Reality Conference, 2009. VR 2009. IEEE, pp. 79–82 (2009)
2. Bane, R., Höllerer, T.: Interactive tools for virtual X-ray vision in mobile augmented reality. In: Mixed and Augmented Reality, 2004. ISMAR 2004. Third IEEE and ACM International Symposium on IEEE, pp. 231–239 (2004)
3. Behzadan, A.H., Kamat, V.R.: Visualization of construction graphics in outdoor augmented reality. In: Proceedings of the 37th Conference on Winter Simulation, pp. 1914–1920. Winter Simulation Conference (2005)
4. Bier, E.A., Stone, M.C., Pier, K., Buxton, W., DeRose, T.D.: Toolglass and magic lenses: the see-through interface. In: Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques, pp. 73–80. ACM (1993)
5. Boubaki, N., Saintenoy, A., Tucholka, P.: Gpr profiling and electrical resistivity tomography for buried cavity detection: a test site at the abbaye de l'ouye (france). In: Advanced Ground Penetrating Radar (IWAGPR), 2011 6th International Workshop on IEEE, pp. 1–5 (2011)
6. Coffin, C., Hollerer, T.: Interactive perspective cut-away views for general 3d scenes. In: 3D User Interfaces, 2006. 3DUI 2006. IEEE Symposium on IEEE, pp. 25–28 (2006)
7. Dey, A., Cunningham, A., Sandor, C.: Evaluating depth perception of photorealistic mixed reality visualizations for occluded objects in outdoor environments. In: Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology, pp. 211–218. ACM (2010)
8. Dey, A., Jarvis, G., Sandor, C., Reitmayr, G.: Tablet versus phone: depth perception in handheld augmented reality. In: Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on IEEE, pp. 187–196 (2012)
9. Dey, A., Sandor, C.: Lessons learned: evaluating visualizations for occluded objects in handheld augmented reality. *Int. J. Hum. Comput. Stud.* **72**(10), 704–716 (2014)
10. Diepstraten, J., Weiskopf, D., Ertl, T.: Interactive cutaway illustrations. *Comput. Graph. Forum.* **22**, 523–532 (2013). doi:[10.1111/1467-8659.t01-3-00700](https://doi.org/10.1111/1467-8659.t01-3-00700)
11. Elmqvist, N., Tsigas, P.: A taxonomy of 3d occlusion management for visualization. *Vis. Comput. Graph. IEEE Trans.* **14**(5), 1095–1109 (2008)
12. Goldstein, E.B.: *Sensation and Perception*. Cengage Learning, Delhi (2013)
13. Jansen, Y., Hornbæk, K.: A Psychophysical Investigation of Size as a Physical Variable. In: IEEE Transactions on Visualization and Computer Graphics, vol. 22, pp. 479–488 (2016). doi:[10.1109/TVCG.2015.2467951](https://doi.org/10.1109/TVCG.2015.2467951)
14. Jones, J.A., Swan II, J.E., Singh, G., Kolstad, E., Ellis, S.R.: The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. In: Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization, pp. 9–14. ACM (2008)
15. Kalkofen, D., Tatzgern, M., Schmalstieg, D.: Explosion diagrams in augmented reality. In: Virtual Reality Conference, 2009. VR 2009, pp. 71–78. IEEE (2009)

16. King, G.R., Piekarski, W., Thomas, B.H.: Arvino-outdoor augmented reality visualisation of viticulture gis data. In: *Mixed and Augmented Reality, 2005. Proceedings. Fourth IEEE and ACM International Symposium on IEEE*, pp. 52–55 (2005)
17. Kruger, J., Schneider, J., Westermann, R.: Clearview: an interactive context preserving hotspot visualization technique. *Vis. Comput. Graph. IEEE Trans.* **12**(5), 941–948 (2006)
18. Kruijff, E., Swan, J., Feiner, S.: Perceptual issues in augmented reality revisited. In: *Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on IEEE*, pp. 3–12 (2010)
19. Landy, M.S., Maloney, L.T., Johnston, E.B., Young, M.: Measurement and modeling of depth cue combination: in defense of weak fusion. *Vis. Res.* **35**(3), 389–412 (1995)
20. Lappin, J.S., Shelton, A.L., Rieser, J.J.: Environmental context influences visually perceived distance. *Percept. Psychophys.* **68**(4), 571–581 (2006)
21. Li, W., Ritter, L., Agrawala, M., Curless, B., Salesin, D.: Interactive cutaway illustrations of complex 3d models. *ACM Trans. Graph.* **26**(3), 31 (2007)
22. Livingston, M.A., Ai, Z., Swan, J.E., Smallman, H.S.: Indoor versus outdoor depth perception for mobile augmented reality. In: *Virtual Reality Conference, 2009. VR 2009. IEEE*, pp. 55–62. IEEE (2009)
23. Livingston, M.A., Dey, A., Sandor, C., Thomas, B.H.: *Pursuit of “X-Ray Vision” for Augmented Reality*. Springer, Berlin (2013)
24. Livingston, M.A., Swan II, J.E., Gabbard, J.L., Höllerer, T.H., Hix, D., Julier, S.J., Baillet, Y., Brown, D.: Resolving multiple occluded layers in augmented reality. In: *Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 56. IEEE Computer Society (2003)
25. Lo Monte, L., Erricolo, D., Picco, V., Soldovieri, F., Wicks, M.C.: Distributed rf tomography for tunnel detection: suitable inversion schemes. In: *Aerospace and Electronics Conference (NAECON), Proceedings of the IEEE 2009 National*, pp. 182–189. IEEE (2009)
26. Lo Monte, L., Erricolo, D., Soldovieri, F., Wicks, M.C.: Radio frequency tomography for tunnel detection. *Geosci. Remote Sens. IEEE Trans.* **48**(3), 1128–1137 (2010)
27. Looser, J., Billingham, M., Cockburn, A.: Through the looking glass: the use of lenses as an interface tool for augmented reality interfaces. In: *Proceedings of the 2nd International Conference on Computer*, pp. 204–211. ACM (2004)
28. Mendez, E., Schmalstieg, D.: Importance masks for revealing occluded objects in augmented reality. In: *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*, pp. 247–248. ACM (2009)
29. Roberts, J.C.: On encouraging multiple views for visualization. In: *Information Visualization, 1998. Proceedings. 1998 IEEE Conference on IEEE*, pp. 8–14. IEEE (1998)
30. Schall, G., Mendez, E., Kruijff, E., Veas, E., Junghanns, S., Reiting, B., Schmalstieg, D.: Handheld augmented reality for underground infrastructure visualization. *Personal Ubiquitous Comput.* **13**(4), 281–291 (2009)
31. Schall, G., Schmalstieg, D., Junghanns, S.: *Vidente-3d visualization of underground infrastructure using handheld augmented reality*. In: *Geohydroinformatics-Integrating GIS and Water Engineering*, vol. 1, pp. 1–17. CRC Press, Boca Raton (2010)
32. Simi, A., Bracciali, S., Manacorda, G.: Hough transform based automatic pipe detection for array gpr: algorithm development and on-site tests. In: *Radar Conference, 2008. RADAR’08. IEEE*, pp. 1–6. IEEE (2008)
33. Swan, J., Livingston, M.A., Smallman, H.S., Brown, D., Baillet, Y., Gabbard, J.L., Hix, D.: A perceptual matching technique for depth judgments in optical, see-through augmented reality. In: *Virtual Reality Conference, 2006*, pp. 19–26. IEEE (2006)
34. Viega, J., Conway, M.J., Williams, G., Pausch, R.: 3d magic lenses. In: *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology*, pp. 51–58. ACM (1996)
35. Zollmann, S., Kalkofen, D., Mendez, E., Reitmayr, G.: Image-based ghostings for single layer occlusions in augmented reality. In: *Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on IEEE*, pp. 19–26. IEEE (2010)



Mustafa Tolga Eren He was born in Ankara, Turkey, in 1984. He received his B.Sc. and Ph.D. degrees from Sabanci University on Computer Science and Engineering in 2006 and 2013, respectively. His interests include real-time visualization, big data analytics, data visualization, interactive projection mapping and augmented reality. He is currently employed at Huawei R&D, Turkey, as a Senior Research Engineer.



Selim Balcisoy Dr. Selim Balcisoy received his Ph.D. on Computer Science in 2001 from Swiss Federal Institute of Technology, Lausanne (EPFL). Between 2001 and 2004, he was Senior Research Engineer at Nokia Research Center Dallas. Dr. Balcisoy co-authored over 50 publications and has been granted with one US patent. His research interests are augmented reality, data visualization and cultural heritage. He is an Associated Professor at Sabanci University, Istanbul, since 2004 and founded VisioThink, a spatial business intelligence company in 2006.