ITcon
www.itcon.org - Journal of Information Technology in Construction - ISSN 1874-4753

EFFECT OF VARYING TASK ATTRIBUTES ON AUGMENTED REALITY AIDED POINT LAYOUT

SUBMITTED: August 2018
REVISED: February 2019
PUBLISHED: March 2019 at https://www.itcon.org/2019/6
EDITOR: Amor R.

Jad Chalhoub, MS,
Graduate Research Associate, Arizona State University;
jadmchalhoub@outlook.com

Steven K. Ayer, Ph.D.,
Assistant Professor, Arizona State University;
sayer@asu.edu

SUMMARY: *Augmented Reality (AR) has been shown to enable several construction tasks and supplement the use of Building Information Modeling (BIM) on construction sites. However, little effort has been made to evaluate the effect of specific task attributes and model-related factors on the accuracy and performance of current generation AR devices. To address this knowledge gap, the authors identified a commonly used AR device and conducted an experiment related to electrical construction point layout tasks. Furthermore, the effects of several task attributes and content variations were explored. The results suggest that the AR device can display content with a locational accuracy of five centimeters from intended design, equally distributed along the X-axis and Y-axis on the design plane. The location of the content is more accurate the closer it is to the paper marker used to locate content in space. Additionally, increased amounts of content shown and variations in types of content displayed significantly affect the accuracy of the points along the Y-axis (elevation). This paper provides an empirical understanding of certain capabilities and limitations of current AR devices using industry practitioners. The research enables practitioners to better plan for the use of AR in different construction and engineering tasks, and guide future research to develop use cases around the strengths of the technology.*

KEYWORDS: *HoloLens, Augmented Reality, Point Layout, Task Variables.*

REFERENCE: *Jad Chalhoub, Steven K. Ayer (2019). Effect of varying task attributes on augmented reality aided point layout. Journal of Information Technology in Construction (ITcon), Vol. 24, pg. 95-111, http://www.itcon.org/2019/6*

1. INTRODUCTION

According to estimates, adoption of Building Information Modeling (BIM) has reached 82% among contractors (Jonathan Woetzel, Srodhar and Mischke, 2017). Augmented Reality (AR) is one emerging technology that is increasingly researched for its ability to leverage the 3D models generated using BIM, supplementing its use both in design offices and on construction sites (Park *et al.*, 2013). For example, AR has been used to enable the assembly of prefabricated electrical conduit (Chalhoub and Ayer, 2018), enhance urban planning (Cirulis and Brigmanis, 2013), and enable better indoor navigation using natural markers for maintenance purposes (Koch *et al.*, 2014).

While previous research highlights the opportunity to use AR in industry, most current AR research is still in the proof of concept stage. Use cases for the technology are being explored by researchers, where most hardware and software has long been in the prototype stages (Feiner *et al.*, 1997; Wang *et al.*, 2014). Subsequently, the effects of variations in the target tasks, such as increased task complexity, on the performance of AR have not yet been empirically identified.

This research studies the use of AR to enable point layout tasks for electrical construction tasks. Previous research demonstrates that AR can be used to communicate design information that had traditionally been illustrated through paper plans for electrical layout tasks (J. Chalhoub, SK. Ayer, "Augmented Reality for Construction Layout Tasks", submitted, Arizona State University, Tempe, Arizona). While this paper does not present new software or hardware related to AR in construction, it investigates how AR performance is affected by changes in design concept factors related to the construction layout task itself. This research leverages existing AR hardware and software to highlight the strengths and weaknesses of current generation AR devices, enabling researchers to investigate more suitable use cases for the technology that meet the needs of current practitioners. Furthermore, developers may use the findings to address some of the current shortcomings of AR, and engineers would be better equipped when planning whether to use AR for a given task, depending on its specific requirements. This research answers the following research question: How do task variables affect the performance of practitioners using AR from accuracy, time, and mental workload perspectives?

2. BACKGROUND

2.1 Augmented Reality

Augmented Reality (AR) is a visualization technology that integrates 3D virtual content and real environment in the same field of view in real time (Azuma, 1997). Milgram and Kishino proposed a "reality spectrum", ranging from a fully real environment to a fully virtual environment (Milgram and Kishino, 1994). Mixed Reality (MR) is any merging of the real and virtual worlds in a single view, and AR is a subset of MR where the environment is predominantly real with some virtual content (Milgram and Kishino, 1994).

In recent years, due to technological advancements, AR research in the civil engineering and construction industry grew significantly. During design and planning stages, AR was used to facilitate discussion and enhance communication concerning BIM content (Lin *et al.*, 2015), and to provide contextually aware information on sites (Bae, Golparvar-Fard and White, 2013). In construction, AR has been used to enable pipe and conduit assembly (Hou, Wang and Truijens, 2015; Chalhoub and Ayer, 2018) and to provide chronological instructions from automatically generated assembly sequences (Makris *et al.*, 2013). AR was also used to enable non-skilled labor to build complex free-form surfaces (Fazel and Izadi, 2018) and to deliver personalized safety information to workers on site (Kim, Kim and Kim, 2017). Post-construction, AR was used for displacement inspection in tunneling systems (Zhou, Luo and Yang, 2017). In education, AR was shown to contribute to student learning for structural analysis purposes by better visualizing content from different angles (Turkan *et al.*, 2017). Generally, AR research and implementation is gaining traction throughout the different industry sectors.

However, current research efforts are still mainly focused on finding potential use cases of the technology and have not thoroughly studied the effects of variations within the task on the performance of the proposed AR solutions. This research contributes to the body of knowledge by exploring this research gap using a construction layout task in electrical subcontracting.

2.2 Cognitive Workload and NASA-TLX

High cognitive workload has long been associated with lower productivity, increased error rate, and slower task completion (Swain and Guttmann, 1983). The NASA Task Load Index (NASA-TLX) is a survey that quantifies the perceived cognitive workload required from a user (Hart and Staveland, 1988). Although the survey is subjective in nature, NASA-TLX has been used more than a thousand times, and is widely accepted as a measurement of the cognitive workload in users (Hart, 2006). In civil engineering research, the NASA-TLX survey has been used to measure the cognitive workload required for masonry construction and to evaluate different design communication methods (Mitropoulos and Memarian, 2013) and quantify the differences in cognitive workload when using different information delivery methods (G. B Dadi *et al.*, 2014). The survey has also been used to study cognitive workload of AR solutions in the AEC industries (Dunston, 2009; Wang and Dunston, 2011; G. B. Dadi *et al.*, 2014). Table 1 summarizes the questions asked in the NASA-TLX survey.

Table 1. NASA-TLX subcategories and descriptions

Subcategory	Description
Mental Demand	How mentally demanding was the task?
Physical Demand	How physically demanding was the task?
Temporal Demand	How hurried or rushed was the pace of the task?
Performance	How successful were you in accomplishing what you were asked to do?
Effort	How hard did you have to work to accomplish your level of performance?
Frustration	How insecure, discouraged, irritated, stressed, and annoyed were you?

2.3 Human-Computer Interaction

The Human-Computer Interaction (HCI) community has long studied user behavior when interacting with virtual content in both AR and Virtual Reality (VR). Research suggests that, when given different types of data visualized in 3D using AR, participants tended to use the technology to create a suitable bird's eye view to primarily understand the content, and then prioritized spatial interactions compared to using computer input to manipulate the content (Büschel *et al.*, 2017). In fact, research suggests that there is tangible performance benefit of using static interaction with content, compared to active interactions (Herman and Stachon, 2016). Specifically, participants with low spatial abilities benefit best form natural interaction with content, such as walking, since walking is a natural human action (Simpson, Zhao and Klippel, 2017; Lages and Bowman, 2018). From a presentation point of view, recent research suggests that it is important to show participants only task related information and limit interaction with content to increase performance (Gardony *et al.*, 2018). These findings were taken into consideration when designing the AR environment for this experiment, where only task-related content was shown and no computer-based interactions were possible.

2.4 Point Layout and Current Practices

Point layout is a construction activity where an individual locates a point on the construction site that is relevant to a given task. For example, in electrical construction, point layout may refer to the task of identifying where certain electrical devices will be installed in a room. A mark is typically left where the electrical device should be installed, and an installation crew would later follow to build the targeted element at the location of the mark. The same process is used for mechanical installations, plumbing and other construction activities. This process was chosen for this research because it is a commonly used, adaptive task that has multiple distinct levels of complexity stemming from various task specific attributes, as explained thereafter in section 3.1.

Currently, point layout is solely dependent on the spatial capabilities of site workers and managers to map 2D plans onto their 3D surroundings (Kwon, Park and Lim, 2014). The practitioners typically receive sets of plans, where the points are identified through a set of distance measurements to other known points in the space. On BIM projects, the plans are produced by generating 2D projections from the 3D model. Figure 1 shows a typical shop drawing for electrical devices layout.

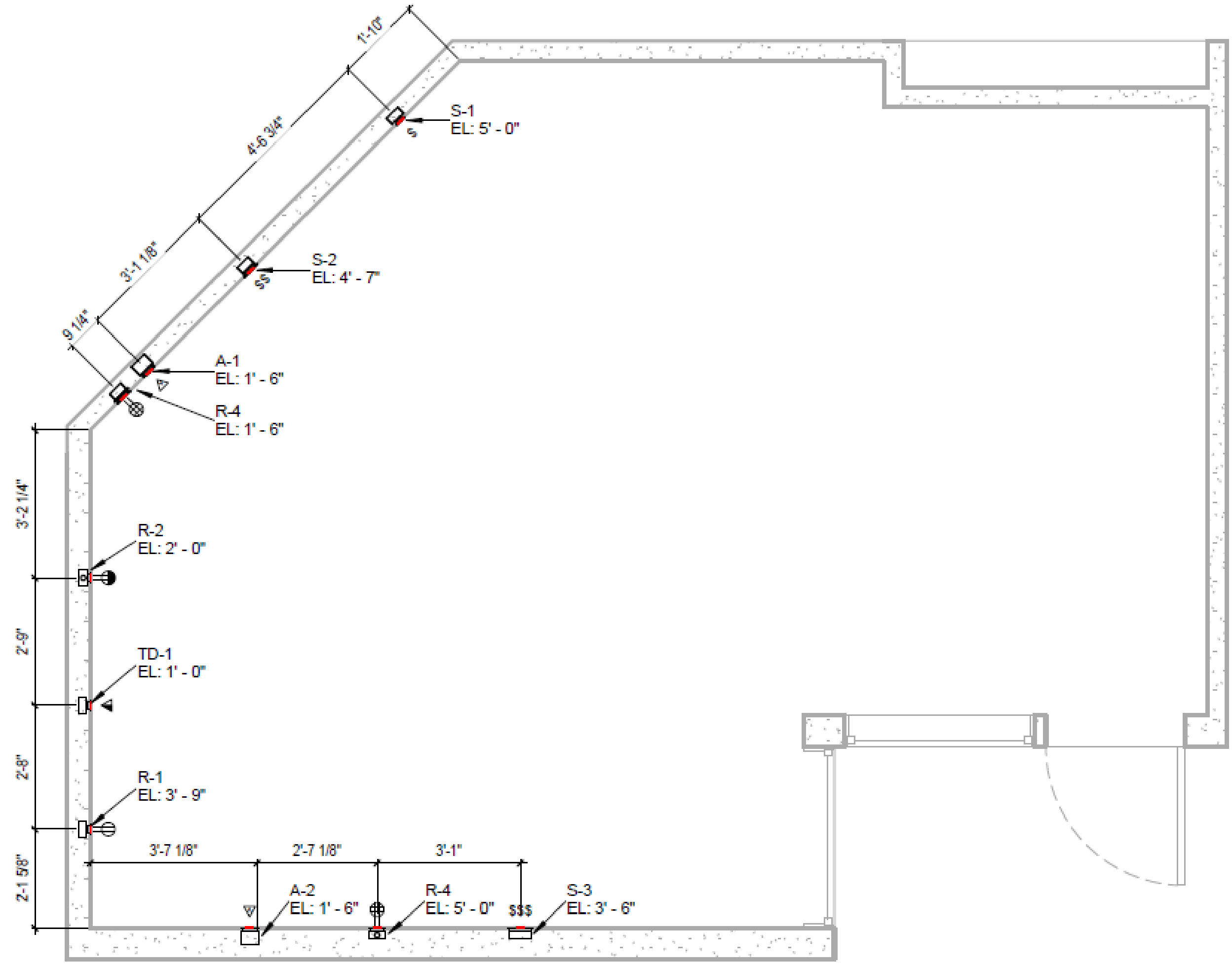

FIG. 1: Standard electrical conduit layout plan.

2.5 Task classification

For most of the twentieth century, research focusing on construction task classification studied the potential for automating those tasks. Porter divided a task into a physical component and an information component (Porter, 1980). Proctor further divides a task into the chronological succession of a perception task, cognitive task and motor task (Van Zandt and Proctor, 2008). Everett theorized that machines are better at physically intensive tasks that require little information exchange and understanding (Everett and Slocum, 1994). Researchers also categorized tasks based on automation potential: Warszawski identified ten "basic activities" that can be performed by robots (Warszawski, 1990); Tucker identified 17 distinct automatable areas (Tucker, 1988); and Kangari created a "robotics feasibility" score by assessing 33 processes in a task (Kangari and Halpin, 1989). Everett proposed a nine-level hierarchical system for classifying all tasks (Everett, 1990). Specifically, construction field operations follow a seven-level hierarchical system, where "project" is the highest level, and "cell", referring to the fiber muscle and nerve stimulated to complete a given action, is the lowest (Everett, 1991).

Recently, some classification efforts have shifted towards the potential of using AR for construction tasks. Unlike robotics and automation, AR was found to be a better fit for information intensive tasks (Wang and Dunston, 2006; Dunston, 2008). Dunston and Wang adapted Everett's hierarchical classification into a five level system, and concluded that the lowest two levels, "composite" and "primitive" tasks are the most appropriate for AR implementation (Dunston and Wang, 2011). Shin and Dunston studied a comprehensive list of construction tasks and theoretically assigned potential AR use cases, including the use of AR for layout tasks (Shin and Dunston, 2008). Because of recent advancements in simulation technologies, more robust, data driven classification systems have arisen. Some research has used smartphone sensors to identify and recognize construction tasks that often produce distinct data signatures (Akhavian and Behzadan, 2016) and utilized machine learning algorithms to better recognize and classify tasks through the collected data (Akhavian and Behzadan, 2018). Different software and coding solutions, such as Dynamic Time Warping techniques, are used to increase the accuracy of the recognition and classification processes (Kim *et al.*, 2018).

Although some research suggests that complexity does not hinder performance when using AR for assembly tasks (Radkowski, Herrema and Oliver, 2015), "mental workload" was mentioned as a limitation for the potential of using AR for a given task (Dunston and Wang, 2011). Furthermore, although prior research theoretically proposes that layout tasks could benefit from AR implementation, little work has been done to assess how task attributes variables may affect the AR implementation. This paper fills this knowledge gap.

3. USE-CASE

The researchers collaborated with a large electrical subcontractor in the Southwest region of the United States. All models were created by the partner company's design team and all the participants were then current practitioners in different roles within the company. The experiment took place in an emptied conference room at the company's regional headquarters, representing a safe environment where participants can work and be effectively monitored.

3.1 Model Variations and Preparations

To test electrical construction layout tasks with AR, several electrical device layout designs were created based on the selected conference room location. The conference room had non-orthogonal walls, making it especially challenging for electrical device layout processes. Figure 1 shows a plan view of the room. Three walls were used for layout in this case, with the devices shown in the figure, and the other portion of the room was used by the researchers to monitor participants and run the experiment.

Although many factors may technically affect the performance of the AR device, the researchers were interested in testing the same variations that currently affect point layout task performance when using paper plans. Several project managers and BIM modelers from the partner company were interviewed, and based on the input provided, three possible variations became apparent: (1) variation in elevation of the devices compared to all devices at the same elevation, (2) low device density compared to high device density in a room and (3) laying out different types of devices (i.e. switches and receptacles) compared to laying out only one type of device.

Four different designs were generated, and the different variables were strategically introduced to allow pairwise comparisons to isolate their effects. Table 2 summarizes the four designs and their various characteristics. All designs were originally created by the partner company using Revit, but the researchers received the models in a 3D AutoCad format.

Table 2. Summary of room designs and factors in each design

Design	Elevation of Devices	Number of Devices	Variety of Devices
1	Same elevation	5	Different Devices
2	Different Elevations	5	Different Devices
3	Different Elevations	10	Different Devices
4	Different Elevations	5	Same Device

The models received included all 3D geometric content, but did not include any embedded information from the original BIM, such as the cost of each element. The room walls, flooring, ceiling, ceiling light fixtures, doors and windows, in addition to the electrical devices to be laid out, were all in the model. Figure 2 shows an isometric view of the received model. The model size varied between 252 Kb and 556 kb, depending on the number of electrical devices in each model.

For the point layout task, only the electrical devices were required to be viewed by the participants through AR, since all other elements physically exist in the space. For example, showing the walls would simply overlay the virtual walls directly on top of the existing walls, which may be disorienting and would further load the AR device. Therefore, all unnecessary elements were removed. Furthermore, the shapes that represent the electrical devices are complex on the "back side", made up of 182 vertices, but are invisible by the user. The shape was simplified to only show the front plate with a cross sign on its center. The cross sign and the name of the device, which is located above the face plate, were both colored in red to create a contrast to the green front plate, enhancing visibility through the AR headset. Other than these minor changes, the original model content was unmodified from what was created by the partner company. Specifically, no content was added, and the points were not moved by the researchers. Figure 3 shows the remaining portions of the model received.

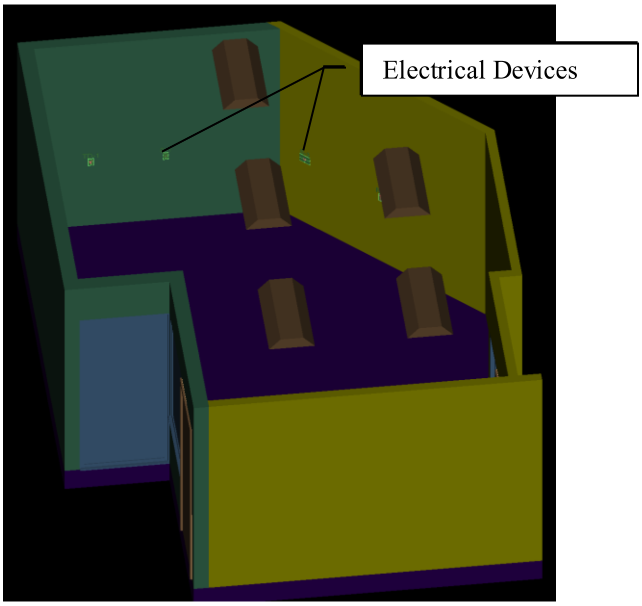

FIG. 2: Design in AutoCad as received from the partner company.

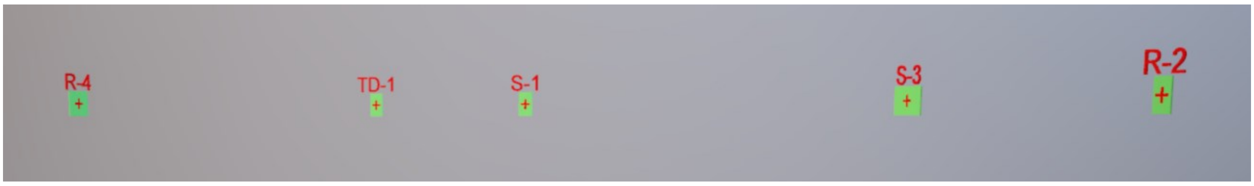

FIG. 3: Design after removing unnecessary elements.

To be viewed through the AR device, the models must be exported from the CAD format to a universal 3D format. FBX format was used in this research because of its broad compatibility, specifically with the game engine used for deployment on the AR device. The exporting method ensured that all shape, texture and color information was retained.

3.2 AR Preparation

The AR device chosen by the researchers was the Microsoft HoloLens, a self-contained computing unit. The unit included 12 total sensors, allowing it to scan and interpret spaces around it. It also has “2 HD 16:9 light engines, with 2.3 M total light points and more than 2,500 light points per radian” to display virtual content, positioned relevant to the scanned space (“HoloLens hardware details”).

In order to correctly display the models on the AR device, three commercial software suites were used: (1) Unity Game Engine, (2) Vuforia SDK and (3) Microsoft Visual Studio.

The Unity game engine enables game development on a variety of software and hardware, including the Microsoft HoloLens (“Unity - Products”). For development, Unity relies heavily on imported content using FBX and provides an Application Programming Interface (API) accessible through JavaScript and C#. Previous visualization efforts in civil engineering research have relied on Unity (Keough, 2009; Pauwels, De Meyer and Campenhout, 2011; Ayer, Messner and Anumba, 2013), proving its suitability for construction focused applications.

The Vuforia Software Development Kit (SDK) is a package that can be installed inside Unity. Vuforia enables advanced computer vision, which allows a broad range of target devices to recognize everyday images and objects using an ordinary built-in camera. A website interface manages a “targets” database, the given set of markers required to be recognized. Once a marker is recognized, the device would display the correct model relevant to the location of the marker in space. Finally, Microsoft Visual Studio compiles and debugs the application created, and then deploys it to the HoloLens. Once deployed, the application is fully contained inside the HoloLens, and does not require external computing power or connection to function.

3.3 The Experiment

The experiment took place over the span of six business days, spread evenly over two weeks. Four to six participants completed the experiment each day. Before starting, the participants were told they would be participating in an electric device room layout exercise using AR technology, but were not given any further information.

Prior to starting the experiment, each participant received two copies of a consent form and a pre-session questionnaire. One signed copy of the consent form was collected, and the other was left with the participant. The pre-session questionnaire asked general questions about each participant, including age, years of experience, average time spent doing point layout, highest education level, prior experience using AR and VR technologies and the participant's perception towards AR use on a construction site. Definitions of point layout and AR were presented at the beginning of the questionnaire for the participants' reference.

In practice, device locations are often indicated with the use of a marker pen or spray paint. Since the experiment was completed in a finished conference room, sticky notes were used as a non-permanent mark of the location of a given point. Figure 4 shows a sample sticky note. To correctly lay out a point, the participant would have to line up the cross on the sticky note to the cross on shown on the device in the model. This allowed the researchers to quickly reset the room to an empty condition between the different exercises and participants.

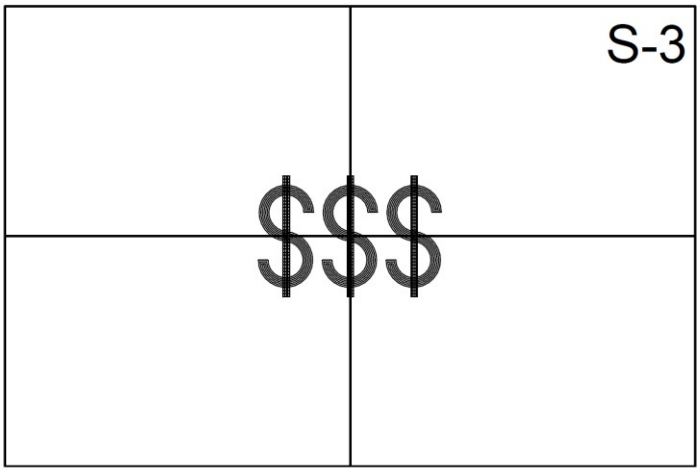

FIG. 4: Sample sticky note (device S3).

Each participant laid out the room using all four designs, but the order of the designs was randomized to mitigate the learning effect. For each run, the content was loaded onto the AR device by the researcher, and the participant was assisted in wearing the device. Once the participant acknowledged that they were able to see the content, they were handed a set of sticky notes corresponding to the devices in the model that they are laying out. The entire session was video recorded from multiple angles to study the behaviors demonstrated during the activity.

Once the layout task was complete, the participant was assisted in removing the headset, and they were handed a NASA-TLX questionnaire to fill. Meanwhile, the researchers measured distances from the center of sticky notes to the walls and floor using a laser measuring tape, quoted by the manufacturer to be accurate to the nearest millimeter. The measurements create a coordinate system for each laid out point, comparable with the coordinate system of the points in the model, enabling a one-to-one accuracy comparison. When the measurements were taken and the NASA-TLX was completed, all sticky notes were removed from the walls, the next design model was loaded, and the process was repeated until all designs were laid out. When the last design was laid out, in addition to the NASA-TLX, the participant received a post-session questionnaire including questions about their comfort level and thoughts for other high-potential applications for the technology in electrical construction based on their experience.

3.4 Analysis Approach

The researchers considered three metrics to assess the performance of the AR solution proposed: accuracy, time, and mental workload.

3.4.1 Accuracy

The main purpose of the layout task is to layout the points accurately where they were designed. Specifically, in electrical construction, depending on the type of the project and contract, accuracy tolerances can be as low as $1/8^{th}$ of an inch (0.003 meter) from intended placement. Each designed and laid out point were assigned a set of coordinates, that represent the distance from a wall on the X-axis and the distance from the floor on the Y-axis. Separate differences between the designed and actual point placements along each axis were calculated. The overall distance (hypotenuse) from the targeted point can be computed using the X and Y values.

3.4.2 Time

The researchers used the videos recorded of the activity to accurately determine the start and end time of each task. The start time was determined as the moment the participant declared he or she can see the content through the AR device, and the end time was determined when he or she declared they were done with the layout task. All times presented in this paper are in seconds.

During some tasks, the participants had technical difficulties viewing the content. Specifically, the content would either shake significantly because of poor spatial tracking, or the application would close, and the content would no longer be viewable. In these instances, the participant had to take off the headset, and the researcher had to reset it. The task times presented in this paper include both times with and without technical difficulties. It is reasonable to expect those times to be reduced as practitioners become more accustomed to using and fixing the device when needed and as the technology matures, but both datasets are included to increase the granularity of reported findings.

3.4.3 NASA-TLX

The collected NASA-TLX questionnaires were digitized and stored in spreadsheet files. Each entry had the responses of the user, the model design it corresponds to, and the order in which that design was laid out for each user. The responses were analyzed using paired statistical analysis to adjust for personal bias from the responders. Additionally, the responses were also analyzed linearly to investigate whether using the AR tool would change the perceived cognitive workload.

4. RESULTS & DISCUSSION

This paper aims to quantify the effect of the varying task attributes on the performance of the participants when using AR for electrical device layout tasks. In the experiment, each participant laid out four different layouts with different factors included in each design. The experiment allows the pairwise comparison of designs to isolate the effect of each task attribute. Table 3 below summarizes the factors included in each design, isolating the effect of each task attribute: comparing designs 1 and 2 isolates the effect of device elevation, comparing designs 2 and 3 isolates the effect of number of devices in a room, and comparing designs 3 and 4 isolates the effect of device diversity in a single space.

Table 3. Summary of effect studied and relevant designs

Effect Isolated	Design 1	Design 2	Design 3	Design 4
Elevation Difference		X	X	X
Number of devices			X	
Diversity of Devices	X	X	X	

4.1 Accuracy

The accuracy was studied along the X-axis and Y-axis separately. Table 4 summarizes the overall accuracy along the X-axis and Y-axis in both data sets. All measurements shown are in meters.

Table 4. Overall accuracy in each design for the X-axis and Y-axis

	Design 1	Design 2	Design 3	Design 4
X-Axis	0.0302	0.0369	0.0357	0.0311
Y-Axis	0.0253	0.0268	0.0344	0.0271

In order to utilize suitable comparative statistical tests, the Shapiro-Wilk test of normality test was used on all datasets tested. The Shapiro-Wilk test of normality is one statistical test that determines whether the population of a dataset follows a normal distribution: the null hypothesis assumes the population is normal, and if the returned *p-value* is less than 0.05, the null hypothesis is rejected, and the population is considered not normally distributed. Table 5 below summarizes the *p-value* for the Shapiro-Wilk test of normality run on each of the cases above. Most of the data was not normally distributed, except for the Y-axis accuracy for designs 2 and 3.

Table 5. Summary of the Shapiro-Wilk test on the datasets

	Design 1	*Design 2*	*Design 3*	*Design 4*
X-Axis	*4.744e-6*	*2.105e-9*	*8.937e-5*	*6.736e-5*
Y-Axis	*2.948e-5*	*0.1404**	*0.8986**	*1.152e-8*

** indicates non-significant values; data is normally distributed*

4.1.1 Task Variations effects

Along the X-axis, none of the task variations had any effects on accuracy. Along the Y-axis, the increased number of devices affected the accuracy. As discussed above, designs 2 and 3 are compared to isolate the effect of increased number of devices and their accuracies along the Y-axis are normally distributed (Shapiro-Wilk test *p-value* = 0.1404 and 0.8986, respectively). A paired t-test can be used, and Table 6 presents the results of the paired t-test. The paired t-test compares the performance of the same set of users under two different circumstances, and if the returned *p-value* is less than 0.05, the performances are considered statistically different. When there are only 5 devices in a room, device placement is 0.00762 meter (22%) more accurate along the Y-axis compared to when a room has 10 devices, and the difference is significant at the 95% confidence level (p-value = 0.01121). In practice, laying out smaller batches of point at a time might maximize the accuracy of the laid-out point along the Y-axis.

Table 6. Summary of the paired t-test on Y-axis accuracy

Testing	Y-axis accuracy (Meter)		Difference (Meter)	t-value	p-value
	Design 2	Design 3			
Number of Devices	0.0268	0.0344	0.00762	2.7225	0.01121

4.1.2 Distance from paper marker

The application developed for this experiment utilized a marker-based approach to accurately place the digital content on site, using the process described in detail in (Chalhoub, Alsafouri and Ayer, 2018). When using marker-based AR, the device stabilizes the content based on the location of the marker. However, as the user gets farther from the marker, the fidelity of the placement of the digital content may also change. The relation between the distance of the point from the marker and the overall point accuracy is studied.

A linear regression approach was used to explain the relation between the distance from the marker and the accuracy of the point placed. First, the distance to the marker was used to explain the variation in accuracy; however, when the model was further analyzed, a power transformation was deemed to be required on the regressor. The model presented in this paper uses the distance to the marker squared as the predictor to explain variation in accuracy. Figure 5 shows a graph of the scatter plot of each point placed, where the Y-axis represents the overall accuracy of the point placed and the X-axis represents the distance from the marker squared, and the regression line passing through them. All distances are in meters.

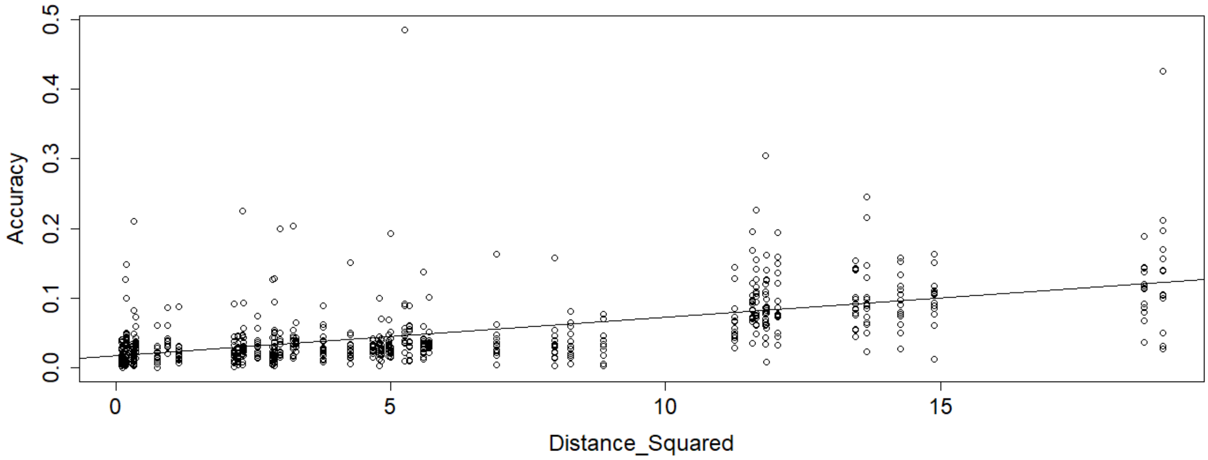

FIG. 5: Plot of accuracy of points vs the squared distance from the marker to the device.

During the experiments, some participants mentioned that the model had significantly shifted from its original location, and he or she either used the new points' locations or tried to place the points by memory in relation to other point. These cases have created several outliers that are clear in Figure 5. However, due to the high number of observations, the data was not adjusted in any way and the outliers did not affect the accuracy findings significantly. Table 7 summarizes the regression and Table 8 presents the corresponding ANOVA table.

Table 7. Summary of the linear regression

	Coefficient	Standard Error	t-value	p-value
Intercept	0. 018162	0.0021517	8.441	<2.2e-16
Distance to Marker ^2	0.005508	0.0002823	19.508	<2.2e-16

A positive coefficient of the square of the distance to the marker indicates that the distance between the placed point and its intended location increases as the distance from the marker increases, and the relation is significant (*p-value* $<$2.2e-16). The Pearson correlation factor between the predictor and variable is 0.5849, and R-square is 0.3421. The regression is significant: The *F-value* is 380.56 with a corresponding *p-value* < 0.05.

Table 8. ANOVA associated with the linear regression

	Degrees of Freedom	Sum of Squares	Mean Sum of Squares	F-value	p-value
Distance to Marker ^2	1	6.2085	6.2085	380.56	<2.2e-16
Residuals	732	11.9419	0.0163		

While the regression would not be necessarily appropriate to predict the exact placement errors of points in future layout jobs when using AR, given the high sample size (734 points), decreased accuracy levels at distant locations from the marker should be expected to follow a parabolic curve in future implementations of this type and generation of technology.

4.1.3 Effects of repetition

The accuracy of point placement on either axis did not change as the participant went through the four exercises. Table 9 shows the mean accuracy along each axis for the different runs (in meters), and the significance of the paired Mann-Whitney comparison of each run and the one that precedes it. On average, accuracy ranged between 0.024 and 0.0358 meter, and all *p-values* are higher than 0.05, indicating no significance at the 95% confidence level. This indicates that the maximum possible accuracy using current technology is effectively achievable from the first use of AR by a participant, reducing required training time.

Table 9. Cumulative paired Mann-Whitney test on the consecutive layout runs concerning accuracy on X-axis and Y-axis

Run	X-Axis			Y-Axis		
	Accuracy	Cumulative V-value	Cumulative significance	Accuracy	Cumulative V-value	Cumulative significance
1	0.0344	N/A	N/A	0.0268	N/A	N/A
2	0.0304	318	0.1757	0.0310	190	0.2639
3	0.0334	281	0.3285	0.0311	172	0.2206
4	0.0358	161	0.9789	0.0244	144	0.6338

4.2 Time:

The effect of varying task attributes on time to complete the layout of the devices was computed. Because some designs have different numbers of devices, the overall time was divided by the number of devices in each run, and the times presented thereafter are times per device in seconds. Table 10 summarizes the Shapiro-Wilk test of normality findings. Since the data is not normally distributed, the paired Mann-Whitney test was used. The paired Mann-Whitney test is similar to the paired t-test: it compares the performance of the same group under two different circumstances, and if the returned *p-value* is less than 0.05, there is a statistically significant difference. However, unlike the paired t-test, the Mann-Whitney does not require normality of the datasets, and so it was used when the samples where not normally distributed.

Table 10. Summary of Shapiro-Wilk test on time datasets

Case	Design	W-value	P-value
With Technical Difficulties	Design 1	0.66201	2.463e-7
	Design 2	0.70821	2.765e-6
	Design 3	0.55525	2.165e-8
	Design 4	0.59086	7.981e-8
Without Technical Difficulties	Design 1	0.65501	1.971e-7
	Design 2	0.68395	1.248e-6
	Design 3	0.54612	1.706e-8
	Design 4	0.54228	2.217e-8

When the devices were designed at different elevations and when the devices designed were themselves different, there was a significant difference in the time required to layout each time. The findings are described below. Notably, the layout time per device did not significantly vary when more devices were in the room (*p-value* = 0.1414).

4.2.1 Effect of Elevation Difference

Time to complete designs '1' and '2' were compared to quantify the effect of difference in devices' elevation on the layout times using AR. Table 11 summarizes the findings of the test for both times with and without technical difficulties.

In both cases, the participants were on average 8 seconds faster per device laid out when all devices were at the same elevation, compared to when they were at different elevations, and the difference is significant at the 95% confidence level (*p-values < 0.05)*. In effect, splitting a design into separate layouts where all devices are at the same height may reduce the time to finish the overall task faster.

Table 11. Summary of Mann-Whitney paired test on effect of elevation difference

Cases	Mean of Design 1 (seconds)	Mean of Design 2 (seconds)	Difference	V-value	P-value
With Technical Difficulties	23.54	32.17	8.63	52	0.0003598
Without Technical Difficulties	23.37	31.49	8.12	53	0.0003907

4.2.2 Effect in variability of devices

Time to complete designs '2' and '4' were compared to quantify the effect of variability of types of devices used on the layout times using AR. Table 12 summarizes the findings.

Table 12. Summary of Mann-Whitney paired test on effect of device diversity

Database	Mean of Design 4 (seconds)	Mean of Design 2 (seconds)	Difference	V-value	p-value
With technical difficulty	24.88	32.17	7.29	103	0.02346
Without technical difficulty	23.61	31.49	7.88	96	0.0153

In both cases, the participants were around 7 seconds faster per device when all the devices in the layout are the same, compared to when different devices are in each room. The difference is significant at the 95% confidence level (*p-values* < 0.05). Similar to the case of elevation difference, splitting a design into separate layouts where all devices are the same type may enable faster productivity in laying out the points.

4.2.3 Effect of Repetition

As previously mentioned, each participant laid out four separate room designs. It is possible that the participants got more comfortable with the AR device and layout task after the first use and may perform better in the second or third runs. Table 13 summarizes the performances of the participants and the comparisons between the first and second, second and third, and third and fourth runs using the paired Mann-Whitney test for the datasets with and without technical difficulties.

Table 13. Cumulative paired Mann-Whitney test on the consecutive layout runs concerning time per device

Cases	Run	Mean Layout Time per Device (seconds)	Cumulative V-value	Cumulative Comparison significance
Case 1: With Technical Difficulties	1	33.57	NA	NA
	2	26.48	415	0.000644
	3	25.12	309	0.1191
	4	24.33	146	0.6668
Case 2: Without Technical Difficulties	1	32.55	NA	NA
	2	25.96	418	0.0004954
	3	23.95	331	0.04265
	4	24.29	121	0.2699

Table 13 summarizes the findings of the cumulative Mann-Whitney test on both datasets. Generally, the participants tend to perform better in each subsequent layout task compared to the one that proceeds it. When considering the dataset with technical difficulties, the performance gains are significant at the 95% confidence level only between the first and second runs (*p-value* = 0.000644). When considering the dataset without technical difficulties, the performance gains are significant in both the second (*p-value* = 0.0004954) and third (*p-value* = 0.04265) runs. Generally, the results indicate that the performances of the participants tend to be enhanced as the participants get more familiar with using the technology.

4.3 Cognitive workload

When considering cognitive workload, each of the six NASA-TLX questions were compared separately. The only difference was between design '2' and '3'. Specifically, participants required an average of 5.43 extra "effort" points to layout 10 devices compared to when laying out 5 devices, and the difference is significant (*p-value* = 0.02663). Table 14 summarizes the findings of the paired Mann-Whitney test. This finding is largely intuitive, as more effort would likely be required to layout more devices.

Table 14. Summary of Mann-Whitney paired test on effort factor in the NASA-TLX questionnaire

Mean of Design 3	Mean of Design 2	Difference	V-value	p-value
23.52	18.09	5.43	34.5	0.02663

Interestingly, none of the cognitive workload factors changed significantly as the participants repeated the tasks. Overall, perceived cognitive workload is independent from repetition and varying task attributes presented in this experiment.

4.4 Limitations

This research explores the effects of varying task attributes on performance when using AR. The limitations of this work are related to the technology, the task attributes studied, and the environment where the work took place.

First, this experiment is based on commercially available hardware and software solutions. The aim of the researchers was not to create a new AR device or a new software suite to display virtual content, but rather to measure the capabilities and limitations of what current technology can afford to any interested party. It is expected that new generations of hardware and software will be developed, and the accuracy may be enhanced. However, the human behaviors involved, especially relating to how participants dealt with more complex situations, is less likely to change.

Second, not all perceivable task variations were studied. The researchers based the designs on discussions with stakeholders from the partner company, in order to quantify the effects of relevant factors. The factors represent the opinions and experience of individuals from a single company in one engineering discipline, and other individuals may consider other task variations, and may require separate studies to understand their effects. Furthermore, when AR becomes more commonly used in the industry, task variables uniquely related to AR may emerge and require separate exploration.

Finally, a conference room was used for the experiment. While the researchers aimed to mimic as closely as possible the layout tasks required on a typical construction site, they did not want to conduct the experiment on an active site because of potential safety concerns. Active construction site conditions, such as varied lighting, noise, congestion, heat or cold, and other conditions may not only affect the AR device, but also the associated human behavior as well. Many of these factors already present challenges to professionals when using traditional paper plans, but their effect on AR remains unknown.

5. CONCLUSION

The work presented in this paper explores the effects of changing various task attributes on the performance of current generation AR hardware and software. The researchers chose an electrical device layout task to complete using AR, and strategically introduced three task attributes variations in four designs: (1) number of devices laid out, (2) difference in elevations of laid out devices, and (3) diversity of the type of devices laid out. Practitioners from the partner company participated in this experiment and completed all four designs in randomized orders. The practitioners also completed NASA-TLX after completing each design to measure their perceived cognitive workload.

First, the accuracy of placement of the points was measured. There is a mild positive correlation between the accuracy of placement of the points and the distance from the paper marker, placed at the center of the marker

(r=0.5849). Points were also laid out more accurately when there were fewer devices in a room compared to when there are many devices. Rooms with more devices also required a significantly higher effort as reported by the NASA-TLX.

The layout completion time per device was computed for each case. In general, the layout process was faster when designs were less complex. Participants required nine seconds less per device when all devices were at the same elevation, and 8 seconds less when devices were all similar and not of different types. Moreover, participants performed significantly faster in the second run compared to the first and faster in the third compared to the second.

The contribution of this paper is in identifying and validating the attributes of a construction layout task that make it advantageous or disadvantageous for using current AR devices with industry practitioners. These findings will allow researchers and practitioners to strategically leverage AR (or avoid its use) to support the needs of a given layout task, and to plan the deployment and implementation of AR in order to maximize its benefits. The findings also highlight the current limitation of AR technology to enable researchers to focus development to help mitigate the current shortcomings. As new AR technologies become more prevalent and powerful, the findings from this work may guide the industry in planning for how to use the new technologies to support the needs of the people who are tasked with using them.

6. ACKNOWLEDGMENTS

The researchers would like to thank Rosendin Electric for participating in this research. The researchers would also like to thank Mr. Fred Meeske, Mr. Andrew Tippit and Mr. Jose Samaniego for their input and support throughout the preparation and testing processes.

This material is based upon work supported by the National Science Foundation under Grant No. IIS-1566274 and ELECTRI International.

7. REFERENCES

Akhavian, R. and Behzadan, A. H. (2016) 'Smartphone-based construction workers' activity recognition and classification', Automation in Construction, 71(Part 2), pp. 198–209. doi: 10.1016/j.autcon.2016.08.015.

Akhavian, R. and Behzadan, A. H. (2018) 'Coupling human activity recognition and wearable sensors for data-driven construction simulation', Journal of Information Technology in Construction, 23, pp. 1–15.

Ayer, S. K., Messner, J. I. and Anumba, C. J. (2013) 'ecoCampus: A new approach to sustainable design education', in itc.scix.net. Available at: http://itc.scix.net/data/works/att/convr-2013-34.pdf (Accessed: 12 March 2018).

Azuma, R. T. A. (1997) 'Survey of Augmented Reality', Presence: Teleoperators and Virtual Environments, pp. 355–385. doi: 10.1162/pres.1997.6.4.355.

Bae, H., Golparvar-Fard, M. and White, J. (2013) 'High-precision vision-based mobile augmented reality system for context-aware architectural, engineering, construction and facility management (AEC/FM) applications', Visualization in Engineering, 1(1), p. 3. doi: 10.1186/2213-7459-1-3.

Büschel, W. et al. (2017) 'Investigating the Use of Spatial Interaction for 3D Data Visualization on Mobile Devices', in Proceedings of the Interactive Surfaces and Spaces on ZZZ - ISS '17. doi: 10.1145/3132272.3134125.

Chalhoub, J., Alsafouri, S. and Ayer, S. K. (2018) 'Leveraging Site Survey Points for Mixed Reality BIM Visualization', in Construction Research Congress. Available at: https://ascelibrary.org/doi/abs/10.1061/9780784481264.032 (Accessed: 5 April 2018).

Chalhoub, J. and Ayer, S. K. (2018) 'Using Mixed Reality for electrical construction design communication', Automation in Construction, 86, pp. 1–10. doi: 10.1016/j.autcon.2017.10.028.

Cirulis, A. and Brigmanis, K. B. (2013) '3D outdoor augmented reality for architecture and urban planning', in Procedia Computer Science, pp. 71–79. doi: 10.1016/j.procs.2013.11.009.

Dadi, G. B. et al. (2014) 'Cognitive Workload Demands Using 2D and 3D Spatial Engineering Information Formats', Journal of Construction Engineering and Management, 140(5), p. 04014001. doi:

10.1061/(ASCE)CO.1943-7862.0000827.

Dadi, G. B. et al. (2014) 'Effectiveness of communication of spatial engineering information through 3D CAD and 3D printed models', Visualization in Engineering, 2(1), p. 9. doi: 10.1186/s40327-014-0009-8.

Dunston, P. (2008) 'Identification of application areas for Augmented Reality in industrial construction based on technology suitability', Automation in Construction. Available at: http://www.sciencedirect.com/science/article/pii/S0926580508000289 (Accessed: 12 February 2017).

Dunston, P. (2009) 'Evaluation of augmented reality in steel column inspection', Automation in Construction. Available at: http://www.sciencedirect.com/science/article/pii/S092658050800085X (Accessed: 4 August 2017).

Dunston, P. S. and Wang, X. (2011) 'A hierarchical taxonomy of aec operations for mixed reality applications', Electronic Journal of Information Technology in Construction, 16, pp. 433–444. Available at: http://itcon.org/paper/2011/25 (Accessed: 8 August 2017).

Everett, J. (1991) 'Construction automation--basic task selection and development of the CRANIUM'. Available at: https://dspace.mit.edu/bitstream/handle/1721.1/13900/24832599-MIT.pdf?sequence=2 (Accessed: 15 March 2018).

Everett, J. G. (1990) 'Back to Basics in Construction Automation', in. doi: 10.22260/ISARC1990/0073.

Everett, J. G. and Slocum, A. H. (1994) 'Automation and Robotics Opportunities: Construction versus Manufacturing', Journal of Construction Engineering and Management, 120(2), pp. 443–452. doi: 10.1061/(ASCE)0733-9364(1994)120:2(443).

Fazel, A. and Izadi, A. (2018) 'An interactive augmented reality tool for constructing free-form modular surfaces', Automation in Construction, 85, pp. 135–145. doi: 10.1016/j.autcon.2017.10.015.

Feiner, S. et al. (1997) 'A Touring Machine : Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment Columbia University Graduate School of Architecture , Planning and Preservation Columbia University', Scenario, 97(4), pp. 74–81. Available at: http://link.springer.com/article/10.1007/BF01682023 (Accessed: 13 February 2017).

Gardony, A. L. et al. (2018) 'Interaction Strategies for Effective Augmented Reality Geo-Visualization: Insights from Spatial Cognition', Human–Computer Interaction, pp. 1–43. doi: 10.1080/07370024.2018.1531001.

Hart, S. G. (2006) 'Nasa-Task Load Index (NASA-TLX); 20 Years Later', Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50(9), pp. 904–908. doi: 10.1177/154193120605000909.

Hart, S. G. and Staveland, L. E. (1988) 'Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research', Advances in Psychology, 52(C), pp. 139–183. doi: 10.1016/S0166-4115(08)62386-9.

Herman, L. and Stachon, Z. (2016) 'Comparison of user performance with interactive and static 3d visualization - Pilot study', in International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives. doi: 10.5194/isprsarchives-XLI-B2-655-2016.

Hou, L., Wang, X. and Truijens, M. (2015) 'Using Augmented Reality to Facilitate Piping Assembly: An Experiment-Based Evaluation', Journal of Computing in Civil Engineering, 29(1), p. 05014007. doi: 10.1061/(ASCE)CP.1943-5487.0000344.

Jonathan Woetzel, Srodhar, M. and Mischke, J. (2017) The construction industry has a productivity problem — and here's how to solve it - MarketWatch. Available at: https://www.marketwatch.com/story/the-construction-industry-has-a-productivity-problem-and-heres-how-to-solve-it-2017-03-04 (Accessed: 1 February 2019).

Kangari, R. and Halpin, D. (1989) 'Potential robotics utilization in construction', Journal of Construction Engineering and Management, 115(23286), pp. 126–143. doi: http://dx.doi.org/10.1061/(ASCE)0733-9364(1989)115:1(126).

Keough, I. (2009) 'goBIM: BIM Review for the iPhone'. Available at: http://papers.cumincad.org/cgi-bin/works/Show?acadia09_273 (Accessed: 12 March 2018).

Kim, H. et al. (2018) 'Application of dynamic time warping to the recognition of mixed equipment activities in cycle time measurement', Automation in Construction, 87, pp. 225–234. doi: 10.1016/j.autcon.2017.12.014.

Kim, K., Kim, H. and Kim, H. (2017) 'Image-based construction hazard avoidance system using augmented reality in wearable device', Automation in Construction, 83, pp. 390–403. doi: 10.1016/j.autcon.2017.06.014.

Koch, C. et al. (2014) 'Natural markers for augmented reality-based indoor navigation and facility maintenance', Automation in Construction, 48, pp. 18–30. doi: 10.1016/j.autcon.2014.08.009.

Kwon, O., Park, C. and Lim, C. (2014) 'A defect management system for reinforced concrete work utilizing BIM, image-matching and augmented reality', Automation in construction. Available at: http://www.sciencedirect.com/science/article/pii/S0926580514001162 (Accessed: 8 August 2017).

Lages, W. S. and Bowman, D. A. (2018) 'Move the Object or Move Myself? Walking vs. Manipulation for the Examination of 3D Scientific Data', Frontiers in ICT. doi: 10.3389/fict.2018.00015.

Lin, T.-H. et al. (2015) 'Using Augmented Reality in a Multiscreen Environment for Construction Discussion', Journal of Computing in Civil Engineering, 29(6), p. 04014088. doi: 10.1061/(ASCE)CP.1943-5487.0000420.

Makris, S. et al. (2013) 'Assembly support using AR technology based on automatic sequence generation', CIRP Annals - Manufacturing Technology, 62(1), pp. 9–12. doi: 10.1016/j.cirp.2013.03.095.

Microsoft (2018) HoloLens hardware details. Available at: https://developer.microsoft.com/en-us/windows/mixed-reality/hololens_hardware_details (Accessed: 18 March 2018).

Milgram, P. and Kishino, F. (1994) 'Taxonomy of mixed reality visual displays', IEICE Transactions on Information and Systems, E77–D(12), pp. 1321–1329. doi: 10.1.1.102.4646.

Mitropoulos, P. and Memarian, B. (2013) 'Task Demands in Masonry Work: Sources, Performance Implications, and Management Strategies', Journal of Construction Engineering and Management, 139(5), pp. 581–590. doi: 10.1061/(ASCE)CO.

Park, C. S. C. et al. (2013) 'A framework for proactive construction defect management using BIM, augmented reality and ontology-based data collection template', Automation in Construction, 33, pp. 61–71. doi: 10.1016/j.autcon.2012.09.010.

Pauwels, P., De Meyer, R. and Campenhout, J. Van (2011) 'Linking a Game Engine Environment to Architectural Information on the Semantic Web', Journal of Civil Engineering and Architecture, 5(46), pp. 787–798. Available at: http://www.davidpublishing.com/davidpublishing/Upfile/4/4/2012/2012040472815913.pdf.

Porter, M. E. (1980) 'Competitive strategy: Techniques for analyzing industries and competition', New York, p. 300. doi: 10.1002/smj.4250020110.

Radkowski, R., Herrema, J. and Oliver, J. (2015) 'Augmented Reality-Based Manual Assembly Support With Visual Features for Different Degrees of Difficulty', International Journal of Human-Computer Interaction. Taylor & Francis, 31(5), pp. 337–349. doi: 10.1080/10447318.2014.994194.

Shin, D. and Dunston, P. (2008) 'Identification of application areas for Augmented Reality in industrial construction based on technology suitability', Automation in Construction, 17(7), pp. 882–894. doi: 10.1016/j.autcon.2008.02.012.

Simpson, M., Zhao, J. and Klippel, A. (2017) 'Take a Walk: Evaluating Movement Types for Data Visualization in Immersive Virtual Reality', in Workshop on Immersive Analytics, IEEE Vis.

Swain, A. D. and Guttmann, H. E. (1983) 'Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications (NUREG/ CR-1278)', U.S. Nuclear Regulatory Commission, (August), p. 728. doi: 10.2172/5752058.

Tucker, R. (1988) 'High payoff areas for automation applications', Proceedings of the 6th International Symposium on. Available at: https://scholar.google.com/scholar?q=―High+payoff+areas+for+automation+applications&btnG=&hl=en&as_sdt=0%2C3 (Accessed: 8 August 2017).

Turkan, Y. et al. (2017) 'Mobile augmented reality for teaching structural analysis', Advanced Engineering

Informatics, 34, pp. 90–100. doi: 10.1016/j.aei.2017.09.005.

Unity (2018) Unity - Products. Available at: https://unity3d.com/unity (Accessed: 18 March 2018).

Wang, X. et al. (2014) 'Integrating Augmented Reality with Building Information Modeling: Onsite construction process controlling for liquefied natural gas industry', Automation in Construction, 40, pp. 96–105. doi: 10.1016/j.autcon.2013.12.003.

Wang, X. and Dunston, P. S. (2006) 'Compatibility issues in Augmented Reality systems for AEC: An experimental prototype study', Automation in Construction, 15(3), pp. 314–326. doi: 10.1016/j.autcon.2005.06.002.

Wang, X. and Dunston, P. S. (2011) 'Comparative effectiveness of mixed reality-based virtual environments in collaborative design', IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews, 41(3), pp. 284–296. doi: 10.1109/TSMCC.2010.2093573.

Warszawski, A. (1990) Industrialization and Robotics in Building: A Managerial Approach. doi: 0060469447, 9780060469443.

Van Zandt, T. and Proctor, R. W. (2008) Human Factors in Simple and Complex Systems. Available at: https://books.google.com/books?hl=en&lr=&id=LfqDZ1VEmyoC&oi=fnd&pg=PP1&dq=Human+Factors+in+Simple+and+Complex+Systems&ots=h-Gv496z_a&sig=1Pz8g9VGwq8bN3AnF3zs6CcaXbo (Accessed: 8 August 2017).

Zhou, Y., Luo, H. and Yang, Y. (2017) 'Implementation of augmented reality for segment displacement inspection during tunneling construction', Automation in Construction, 82, pp. 112–121. doi: 10.1016/j.autcon.2017.02.007.