Augmented Reality for Infrastructure Inspection with Semi-autonomous Aerial Systems: An Examination of User Performance, Workload, and System Trust

Jared Van Dam*
Virginia Tech

Alexander Krasner
Virginia Tech

Joseph L. Gabbard
Virginia Tech

ABSTRACT

The use of augmented reality (AR) with drones in infrastructure inspection can increase human capabilities by helping workers access hard-to-reach areas and supplementing their field of view with useful information. Still unknown though is how these aids impact performance when they are imperfect. A total of 28 participants flew as an autonomous drone while completing a target detection task around a simulated bridge. Results indicated significant differences between cued and un-cued trials but not between the four cue types: none, bounding box, corner-bound box, and outline. Differences in trust amongst the four cues indicate that participants may trust some cue styles more than others.

Keywords: Augmented reality, signal detection, infrastructure inspection, workload, unmanned aerial system.

Index Terms: Multimedia information systems—artificial, augmented, and virtual realities, User interfaces—ergonomics, evaluation/ methodology, screen design, style guides

1 Introduction

Current methods of civil infrastructure methods rely on labor intensive methods that require inspectors visually analyze each area of the bridge. Given that these inspections must occur biennially, there arises a need for a timely remote inspection solution that still provides the operator with adequate information but allows for remote inspections [1]. The use of unmanned aerial systems (UASs) during inspections provides easier access to these hard-to-reach areas [2, 3]. In addition, by using stereoscopic depth cameras on the UAS, computer vision algorithms, and augmented reality (AR), we can detect defects on bridge surfaces and relay those areas to operators while augmenting their view with useful information (e.g., world-fixed rulers and annotations) that can aid in task completion . The use of these technologies may help operators beyond what a simple monoscopic camera would provide by displaying relevant information to the operator as well as stereoscopic views needed for depth perception [4]. It is still unknown though what performance decrements may be associated with these AR interface elements [5].

To test these possible decrements and for ease of rapid prototyping and interface development, we created a virtual testbed (Figure 1) for the present study that allows us to simulate video-see-through AR in virtual reality [6]. With this testbed, we can examine how different interfaces impact users during bridge inspection tasks without the associated costs of real-world testing as the testing of these designs in the real-world often incurs a great deal of expense in both time and energy.

* Jaredvd@vt.edu

2 Empirical user study

2.1 Study purpose and hypotheses

The proposed study aims to extend the previous work to the virtual reality space and determine the effect of cue type, target saliency, and automation aid reliability on signal detection rates. Additionally, signal detection rates with different cue types were measured to see if feature bound AR cues afford similar detection rates compared to fully bound and corner bound AR cues while also granting operators a better view and understanding of target features. Thus, we hypothesize that signal detection (percentage of hits, misses, and false alarms) in the low target saliency trials will not be significantly different amongst the three AR cue types but will differ significantly non-cued trials. Also hypothesized is that workload will be highest in the low target saliency trials and lowest in the high saliency trials and will also be change based on cuing type with workload highest in trials with no cues present and lowest in those trials with AR cues. Similarly, trust will be highest in the high saliency trials and lowest in the low saliency trials.

2.2 Experimental design

We assigned participants (n = 28) to experimental conditions via a Latin square where participants experienced all three levels of target saliency and all four AR cue types. Specifically, our study employed a four (AR cue type: none, bounding box, corner-bound box, and outline – see Figure 1) × three (target saliency: low (0.12), medium (0.17), and high (0.22)) within-subjects design. After each trial, we asked participants to remove the head-mounted display (HMD) and complete the survey packet which contained the NASA-TLX and trust measures. On average, 60% of the targets were hits, 20% were false alarms, and 20% were misses. Researchers modeled the virtual bridge in the Unity game engine and presented it to participants via a wireless HTC Vive Pro headset.

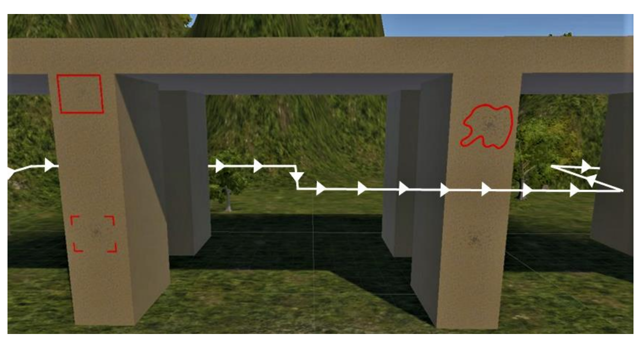

Figure 1: Simulated bridge environment with all saliency levels and cue types represented. The top left target is the low saliency target (0.12) with the bounding box; the bottom left target is the medium saliency target (0.17) with the corner-bound cue; the top right target is the high saliency target (0.22) with the outline type cue; and the bottom right target is the non-cued, high saliency target.

978-1-7281-6532-5/20/$31.00 ©2020 IEEE
DOI 10.1109/VRW50115.2020.00222

3 Results

Significant pairwise differences in hits per trial were present between the no-cue condition and bounding box ($p = 0.009$) and no-cue and corner bound cue ($p = 0.033$) but were not significant for the no-cue and outline cue pair ($p = 0.074$). Significant differences existed between the low and medium ($p < 0.001$) categories, low and high ($p < 0.001$) categories, and medium and high ($p < 0.001$) categories.

While mean misses for the no cue trials were higher relative to the bounding box, corner bounded box, and outline cues, none reached statistical significance ($p > 0.09$). Pairwise comparisons with the three saliency levels though indicated significant differences between low and medium ($p < 0.001$), low and high ($p < 0.001$), and medium and high ($p = 0.002$). As shown in figure 2, results indicate that overall performance did differ significantly based on saliency but pairwise differences between the no-cue condition and the cued conditions did not reach significance.

Pairwise comparisons for false alarms within the cue types led to no significant differences with ($p > 0.7$). Pairwise comparisons between low and medium ($p = 0.083$) saliency as well as medium and high ($p = 1$) levels of saliency did not reach significance but did exist when comparing low to high saliency ($p = 0.03$). Thus, participants had significantly more false alarms in the high saliency (0.22) condition when compared to the low saliency (0.12) condition.

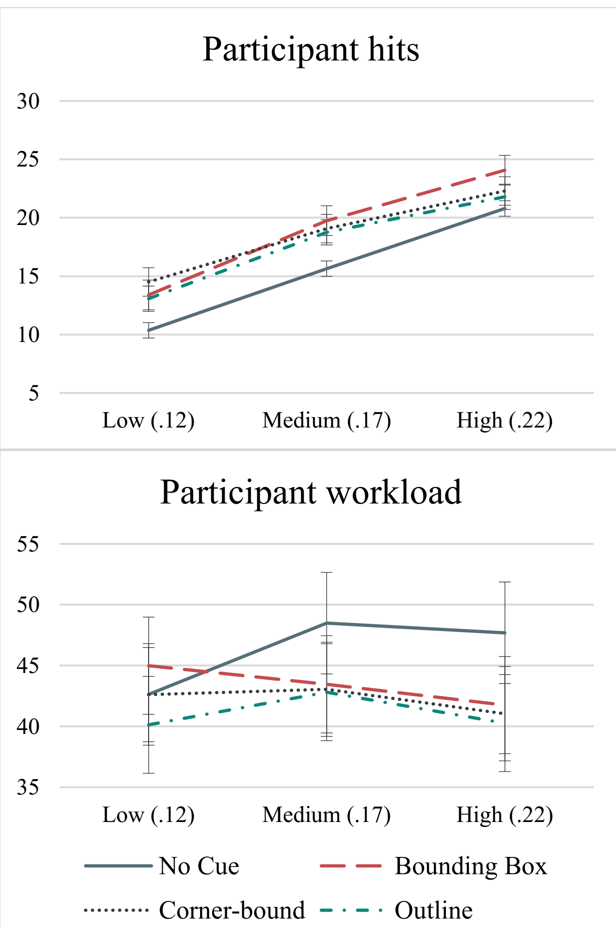

Figure 2: Group means for number of hits (top) and workload (bottom) per trial categorized by cue type and saliency level. Error bars represent standard error of the mean.

Pairwise comparisons of workload across cue types indicate significant differences between no cues and corner-bound cues ($p = 0.03$), and outline cues ($p < 0.001$), but not for the no cue and bounding box ($p = 0.07$) pair. Pairwise comparisons of the low, medium, and high saliency levels were not significant ($p > 0.6$).

Pairwise comparisons of trust showed that while the cues did not differ significantly amongst themselves, the outline cue did differ significantly from no cue ($p = 0.30$) while the bounding box, and corner-bound cue did not differ significantly from having no cues.

4 Discussion

Results showed that performance in the signal detection task was generally better in the cued conditions relative to the non-cued conditions except in the case of false alarms where participants had more false alarms in the highly salient conditions. Given the work of Maltz and D. Shinar [5] who found that cuing was generally helpful during complex tasks, but hurtful during easy tasks, and the finding of the present study where users had more false alarms during the highly salient trials, care should be taken to only use cuing systems during complex tasks or those tasks where user misses carry significant weight. Future research should also examine this area further by studying cuing systems that adjust cue saliency to the target difficulty so that, when target saliency is low and the target is difficult to detect, the augmented cue is very salient but, when target saliency is high and easy to detect, the augmented cue saliency is either low and unobtrusive, or not present.

This study also shows that the corner-bound cue may be useful when the user's attention is needed without the increased clutter that accompanies the typical bounding box. Similarly, the outline cue did not hinder performance beyond the other cue types, so this cue type may be useful when the exact shape of the target area is useful for the operator to know. The outline type cue also elicited greater levels of trust relative to no cues, so the use of this cue type could be useful when designers want to engender more trust in their system. Future work in this research space entails analyzing the effect that imperfect cuing aids have on user signal detection when the cuing aid augments targets that are near un-cued targets.

Acknowledgements

This work is supported by National Science Foundation grant #1840044

References

[1] J. L. Wells, B. Lovelace, and T. Kalar, "Use of Unmanned Aircraft Systems for Bridge Inspections," Transp. Res. Rec. J. Transp. Res. Board, vol. 2612, no. 1, pp. 60–66, 2017.

[2] Y. Ham, K. K. Han, J. J. Lin, and M. Golparvar-Fard, "Visual monitoring of civil infrastructure systems via camera-equipped Unmanned Aerial Vehicles (UAVs): A review of related works," Vis. Eng., vol. 4, no. 1, pp. 1–8, 2011.

[3] A. H. Behzadan, S. Dong, and V. R. Kamat, "Augmented reality visualization: A review of civil infrastructure system applications," Adv. Eng. Informatics, vol. 29, no. 2, pp. 252–267, Apr. 2015.

[4] D. H. Shin and P. S. Dunston, "Evaluation of Augmented Reality in steel column inspection," Autom. Constr., vol. 18, no. 2, pp. 118–129, Mar. 2009.

[5] M. Maltz and D. Shinar, "New Alternative Methods of Analyzing Human Behavior in Cued Target Acquisition," Hum. Factors J. Hum. Factors Ergon. Soc., 2004.

[6] C. Merenda, C. Suga, J. Gabbard, and T. Misu, "Effects of Vehicle Simulation Visual Fidelity on Assessing Driver Performance and Behavior," *2019 IEEE Intell. Veh. Symp.*, no. Iv, pp. 1679–1686, 2019.