# A Case for Enhancing UAV Ground Control Stations with Cross Reality

BRYSON LAWTON, University of Calgary, Canada
FRANK MAURER, University of Calgary, Canada


Fig. 1. An example of what a cross-reality GCS experience could look like. Here a conventional GCS controlling UAVs from a remote location is run on a laptop (bottom-center) and is paired with a mixed reality headset to expand its capabilities to display extra virtual screens (upper-middle & bottom-right), UAV information panels (upper-right & upper-left), and a 3D map of the operational environment (bottom-left). Images on the laptop and virtual screens are of the VCSi GCS by Lockheed Martin [8].

Ground control stations are widely used to enable humans to monitor and control one or many unmanned aerial vehicles within operational environments. Standard approaches to ground control station software utilize screen-based user interfaces, which due to the intrinsic limitations of physical screens results in constraints on workspace scalability, information visualization, system mobility, and operator collaboration. This paper proposes that a cross-reality user experience utilizing mixed reality head-mounted displays alongside conventional ground control stations is a promising way to move past prior restrictions imposed by screen-based approaches while maintaining access to the strengths of preexisting hardware and software.

## 1 INTRODUCTION

When unmanned aerial vehicles (UAVs) are deployed to accomplish tasks, ground control stations (GCSs) are often used to facilitate the control and monitoring of such vehicles via one or multiple users. This is especially the case when a large number of UAVs, colloquially referred to as drones, need to be controlled simultaneously by relatively few users in an effective and efficient manner. GCS software traditionally employs a 2D screen-based user interface (UI), targeting user interaction for just one user through a computer monitor or screen of custom portable hardware [4, 6, 8, 11, 12, 17]. State-of-the-art GCSs, such as UAV Navigation's Visionair [11] and Lockheed Martin's VCSi [8], even include support for collaborative control via several users and the utilization of multiple displays for increased screen real-estate.

### 1.1 Limitations of Conventional GCSs

Regardless of how advanced a conventional GCS is, their reliance on physical screens puts restrictions on the capabilities of such systems due to the inherent limitations of screen-based displays. For instance, the amount a GCS UI can scale to the needs of operators is directly linked to the physical display hardware available and how feasible this hardware is to set up around a user. Due to this dependability, adding more display real-estate forces increased costs, power consumption, weight, and physical space required which in turn reduces the mobility of such systems. As a result, GCSs for larger and more complex operations are often stuck to one location [4, 6]. With such workstations, it is also usually not easy to add or move around display hardware to customize a workstation to how an operator prefers. Although these stationary workspaces may be fine in some applications, GCS workspaces incorporating many screens are often not realistic, especially when required to be deployed in the drones' operating environment where a user may need to move around the real world, or they are operating with limited power or space, such as from within a vehicle.

Another issue arises from how screens present 3D information to users. Since screens present flat, 2D images to a viewer, all information that is inherently 3D needs to be flattened to a pseudo-3D view to be displayed. When it comes to GCSs, the terrain information and 3D geospatial data of UAVs within their operating environment is often visualized using 2D or pseudo-3D maps, examples of which can be seen on the laptop and virtual screens in Figure 1 and Figure 2. This flattened perspective, however, results in a loss of depth cues important for quickly and intuitively understanding natively 3D spatial information [14] such as a drone's position above ground level or 3D terrain features, potentially leading to misinterpretations of the visualized data and errors of judgment.

Multi-user collaboration is also restricted by the screen-based approach of traditional GCSs. In order for many operators to collaborate and effectively distribute their attention over a large number of UAVs, typically they work with their own separate workspaces while communicating verbally to coordinate their efforts and best handle the ongoing operation together. These operators could be located locally within the same room or outdoor environment, or situated remotely across vast distances. Verbally describing what one is seeing, however, can sometimes be hard for others to understand without visual context [2]. There are also important non-verbal aspects to communication, such as gestures or gaze, which are not supported if one can only interact verbally. Even when working in close physical proximity, non-verbal methods of communicating could involve removing focus from one's own workstation or abandoning it to temporarily shift attention to another operator or their workspace. When located remotely, this kind of in-person, non-verbal communication is of course not possible at all with screen-based user experiences. Thus, collaboration across both local and remote users may at times be challenging with screen-based UIs, as natural non-verbal communication can be partially or entirely restricted, potentially causing teamwork inefficiencies, frustration, and delayed responses.

### 1.2 Strengths of Conventional GCSs

Despite these issues, there are still several advantages to using traditional GCS approaches, with Figure 2 showing one such example. Conventional screen-based interfaces have a familiar design that experienced GCS operators may prefer over new interfaces or technologies which they would have to learn. Physical methods of direct interaction with UIs, such as touchscreens, mice, and keyboards, also provide haptic feedback to users which can be better for confirming that their actions have been successfully detected than auditory or visual feedback [13]. Also, physical keyboards are a well-known method of quick and accurate text entry which is sometimes quite cumbersome with virtual keyboards in mixed reality. Lastly, custom-built, screen-based GCS hardware like that in Figure 2 is already widespread for a variety of mobile deployment applications and includes the onboard hardware necessary not only for running complex GCS software, but for communicating with UAVs and satellites [1, 12, 15, 16]. Therefore, since the main problem with current GCSs is the screen-based user experience and not the rest of the hardware or backend software, an ideal solution for enhancing conventional GCSs would be one which doesn't unnecessarily replace all preexisting hardware and software but rather works together with them, utilizing their strengths while improving upon their current weaknesses.

Fig. 2. Worthington Sharpe's Wing GCS [12], an example of a mobile GCS used for deployment in outdoor operational environments.

## 2 ENHANCING GROUND CONTROL STATIONS WITH CROSS-REALITY

In order to move past the current limitations of screen-based GCSs, a cross-reality approach may be the best solution. Specifically, expanding the capabilities of screen-based GCSs with stereoscopic mixed reality (MR) head-mounted displays (HMDs), such as the Microsoft Hololens 2 [9], would allow one to maintain access to the strengths and familiarity of conventional GCSs while enhancing many of their weaknesses with the inherent benefits immersive MR technologies have over screen-based displays.

### 2.1 Single-User Experience Improvements

Concerning the impact a cross-reality experience could have on individual GCS operators, first consider that the working space available in MR is far beyond what is capable with screens [7]. In immersive MR workspaces, not only can one place virtual 2D panels and 3D objects around themselves in any size and direction, but since MR HMDs display virtual surroundings in what one perceives to be true 3D, one can also take advantage of depth to place these objects closer and farther away. As such, MR displays remove the physical constraints imposed by screen-based hardware on how much can be displayed at once to users, instead offering them much more usable working space to utilize. Therefore if used in tandem with a screen-based interface as a cross-reality experience, one no longer needs more than one physical screen and instead can expand the display real-estate of preexisting GCSs using as many holographic screens as are needed. Demonstrated in Figure 1, these could be used exactly the same as physical screens, but in addition be dynamically created and destroyed, scaled to any desired size, and positioned around the operator to create a customized workspace that scales with their needs and provides much more display real-estate than was previously viable.

Information and user interface elements could also be passed seamlessly between these real and virtual screens and interacted with either via input device hardware (touchscreen, mouse, keyboard, etc.) or MR interaction methods like hand-tracking, gestures, and eye-tracking. This would allow one to use whatever interaction method is most intuitive and comfortable for them. For instance, cross-reality enables one to use both hands to interact with UI elements instead of a single cursor, which could improve the speed, efficiency, and natural feel of navigating the GCS and issuing commands. On the other hand, input with a mouse and keyboard could grant others a familiar experience with more rapid text entry and better fine interaction than with one's hands in MR. The UI design of these virtual screens could even employ a combination of traditional and MR UI elements to both be familiar to new users with prior GCS experience and ease users into what new ways of interacting with GCSs are possible with MR technologies.

In addition, one also can utilize MR to visualize the drones' geospatial information within their operating environment using virtual holographic 3D maps like the one in Figure 1. Able to be created, scaled, and placed in one's cross-reality workspace however they wish, these maps have the potential to more intuitively convey 3D spatial and terrain information than their pseudo-3D counterparts as it is now presented in its naturally perceived dimensions [5]. Removing the reliance on physical screens also allows for any number of 3D maps one wishes to be created alongside the real and virtual screens, each customizable for varying zoom levels to place focus on important drones or locations.

These improvements overall increase the scalability of such systems while maintaining their mobility. No longer would increased screen real-estate make a portable GCS less mobile when one can instead place as many weightless holograms as one wants around it and have them track with it, such as in Figure 1. Furthermore, if the capabilities of the remaining physical hardware permits, operators could now even have the option to easily transition between remote and local operation of a GCS, as whether working remotely from a room, moving vehicle, or locally in the operating environment, the scalability and mobility of the GCS remains the same.

### 2.2 Multi-User Experience Improvements

In situations where multiple GCS operators are working together, there is also the potential for further advancement. To offer improved methods of team communication and collaboration, MR HMDs can be utilized to grant co-located operators large, shared virtual workspaces in which they can work together alongside remote collaborators being represented virtually by humanoid avatars. Such a workspace could even support cooperation from operators across Milgram's reality-virtuality continuum [10], whether they are using MR, virtual reality (VR), or traditional screen-based technologies. This shared workspace could either be laid out to resemble conventional in-person shared GCS workspaces to keep operator interaction familiar, or be entirely new and customizable to their shared needs and preferences. Individual workspace elements, like maps or panels, could either be shared and interacted with by all or be in a private task space [3] and seen only by one operator to reduce unnecessary workspace clutter for others. In addition, MR and VR users could utilize hand tracking or hand-held controllers to map hand and arm gestures onto their avatar's hands to provide a richer reference space [3] for more naturally and clearly expressing themselves to others through a combination of physical and verbal communication methods. For instance, operators could virtually gather in the middle of a shared semi-spherical layout of many virtual screens, information panels, and 3D maps and gesture towards various details while talking to effectively and efficiently communicate how best to distribute tasks amongst themselves and what they should next prioritize together.

## 3 CONCLUSION

By incorporating a MR HMD to communicate and operate together with preexisting GCS systems as a cross-reality experience, their capabilities have the potential to be enhanced with mixed reality's ability to support larger, more customizable workspaces, hands-on interaction, and true 3D visualization. Multi-user collaboration could also be positively impacted, as local and remote operators could more naturally communicate and interact in shared virtual workspaces. Such improvements would make scaling up GCSs easier, while maintaining their mobility and utilizing the advantages of both technologies for an enriched experience overall.

## REFERENCES


[1] Edge Autonomy. 2022. Portable Ground Control Station. Retrieved April 3, 2022 from https://uavfactory.com/en/products/ground-control-station/portable-ground-control-station

[2] Eliza Bobek and Barbara Tversky. 2016. Creating visual explanations improves learning. *Cognitive Research: Principles and Implications* 1, 27 (12 2016). https://doi.org/10.1186/s41235-016-0031-6

[3] Bill Buxton. 2009. Mediaspace – Meaningspace – Meetingspace. In *Media Space 20 + Years of Mediated Life*, Steve Harrison (Ed.). Springer, London, 217–231. https://doi.org/10.1007/978-1-84882-483-6_13

[4] Jinyeong Heo, Suhwan Kim, and Yongjin James Kwon. 2016. Design of Ground Control Station for Operation of Multiple Combat Entities. *Journal of Computer and Communications* 4, 5 (May 2016). https://doi.org/10.4236/jcc.2016.45010

[5] Mark St. John, Michael B. Cowen, Harvey S. Smallman, and Heather M. Oonk. 2001. The Use of 2D and 3D Displays for Shape-Understanding versus Relative-Position Tasks. *Human Factors* 43, 1 (2001), 79–98. https://doi.org/10.1518/001872001775992534 arXiv:https://doi.org/10.1518/001872001775992534 PMID: 11482314.

[6] Yongjin (James) Kwon, Jinyeong Heo, Sekyung Jeong, Seunghee Yu, and Suhwan Kim. 2016. Analysis of Design Directions for Ground Control Station (GCS). *Journal of Computer and Communications* 4, 15 (2016), 1–7. https://doi.org/10.4236/jcc.2016.415001

[7] Lee Lisle, Xiaoyu Chen, J.K. Edward Gitre, Chris North, and Doug A. Bowman. 2020. Evaluating the Benefits of the Immersive Space to Think. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. Institute of Electrical and Electronics Engineers, New York, NY, USA, 331–337. https://doi.org/10.1109/VRW50115.2020.00073

[8] Lockheed Martin. 2022. VCSi. Retrieved March 19, 2022 from https://www.lockheedmartin.com/en-us/products/cdl-systems/vcsi.html

[9] Microsoft. 2022. Microsoft HoloLens | Mixed Reality Technology for Business. Retrieved April 2, 2022 from https://www.microsoft.com/en-us/hololens

[10] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. 1994. Augmented reality: A class of displays on the reality-virtuality continuum. *Telemanipulator and Telepresence Technologies* 2351 (01 1994). https://doi.org/10.1117/12.197321

[11] UAV Navigation. 2022. Visionair - GCS Software: UAV Navigation. Retrieved March 19, 2022 from https://www.uavnavigation.com/products/ground-control-station/visionair-gcs-software

[12] Worthington Sharpe. 2022. Wing GCS - ERL Robotics League. Retrieved April 1, 2022 from https://www.worthingtonsharpe.com/?wix-vod-video-id=fa711618c69b4c3fba670e9e3e5ad93e&wix-vod-comp-id=comp-kngggd67

[13] Roland Sigrist, Georg Rauter, Robert Riener, and Peter Wolf. 2013. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review. *Psychonomic Bulletin and Review* 20 (02 2013), 21–53. https://doi.org/10.3758/s13423-012-0333-8

[14] Harvey Smallman, Mark St John, Heather Oonk, and M. Cowen. 2001. Information Availability in 2D and 3D Displays. *IEEE Computer Graphics and Applications* 21 (09 2001), 51–57. https://doi.org/10.1109/38.946631

[15] Alpha Unmanned Systems. 2022. Ground Control Station. Retrieved April 3, 2022 from https://alphaunmannedsystems.com/ground-control-station-2/

[16] UAVOS. 2019. UAVOS Expands Its Family of Ground Control Stations. Retrieved April 3, 2022 from https://www.uavos.com/uavos-expands-its-family-of-ground-control-stations/

[17] Zhang Wen and Qin Guo-Shun. 2012. Design of UAV Ground Control Station. In *Software Engineering and Knowledge Engineering: Theory and Practice (Advances in Intelligent and Soft Computing)*, Wei Zhang (Ed.). Springer, Berlin, Heidelberg, 249–256. https://doi.org/10.1007/978-3-642-29455-6_36