Assessing the Impact of VR Interfaces in Human-Drone Interaction

Maitrey Gramopadhye, *Computer Science*, *UNC Chapel Hill*, Chapel Hill, NC USA, maitrey@cs.unc.edu

Arran Zeyu Wang, *Computer Science*, *UNC Chapel Hill*, Chapel Hill, NC USA, zeyuwang@cs.unc.edu

Leonard Shearer, *Computer Science*, *UNC Chapel Hill*, Chapel Hill, NC USA, leojshe@unc.edu

Tony Qin, *Computer Science*, *UNC Chapel Hill*, Chapel Hill, NC USA, tonyqin@cs.unc.edu

Daniel Szafir, *Computer Science*, *UNC Chapel Hill*, Chapel Hill, NC USA, dszafir@cs.unc.edu

***Abstract*—Drones are a uniquely useful type of robot that allows aerial exploration and surveying. However, they require substantial training to operate effectively. Previous research has introduced novel drone interfaces, but have not substantively compared these interfaces. In this paper, we explore the impact and differences of VR and 2D interfaces on layman human-drone interaction. We measured performance on three environmental exploration tasks in an indoor environment. Participants were introduced to the interfaces and then asked to read random sequences of digits from 1 meter away. The simple task had the participant read one sequence, while the complex task had the participant read four sequences in positions around the room. This required participants to demonstrate moderate levels of competence while operating the drone. Our results suggest that our VR interface has a comparable performance to a smartphone interface across all three tasks. Despite this, VR interfaces show potential to reduce barriers to drone operation and should be the focus of future research.**

***Index Terms*—Drones, aerial robots, human-drone interaction, VR interface, drone navigation, robot teleoperation.**

I. Introduction

Drone hardware has greatly improved over the past decade, and drones are more practical and useful than ever before. The emergence of commercially available drones has cemented the ubiquity of these robots. Drones have numerous applications such as search and rescue, photography/videography, traffic monitoring, and surveillance. Although there have been many advances in autonomous drone operation, human controlled teleoperation still remains the default for complex tasks. Teleoperation of drones is necessary for precise or ill-defined tasks better suited to human control. Drone control is a nontrivial challenge to layman operation, as it is difficult to precisely navigate a drone without extensive training. Due to some drones having 6 degrees of freedom (DOF), there is an abundance of information for the user to track. Despite these challenges, straightforward drone interfaces may be able to reduce the barriers to entry.

Previously, interfaces used 2D displays and keyboard inputs [1], but now virtual reality (VR) [2] and joysticks can be used. VR, camera, and computing technologies continue to improve, allowing for higher fidelity displays that can show more information with higher levels of processing. Moreover, the drone has the potential to create more meaningful outputs. For instance, drones have been used to map archaeological sites [3], perform marine surveillance [4], and implement technologies like simultaneous localization and mapping (SLAM) [5] and point clouds [6]. However, there are few analyses that compare these new interfaces to prior versions. It is important to establish that incorporating these new technologies into teleoperation interfaces actually improves navigational and task performance. For an overview on related work, please see Appendix A.

This paper presents two contributions to drone teleoperation interfaces. First, we develop a novel VR interface to control a 6-DOF drone. The interface uses commonly available technology like a mono-camera drone, VR headset, joystick controllers, and a smartphone, that users may already be familiar with. In contrast to highly-sophisticated drones and interfaces used for mission-critical tasks that can afford high levels of operator training, these basic technologies are more suitable for layman operation. Furthermore, they are less expensive and widely available, allowing the propagation of drone usage to fields outside of robotics research. In turn, other fields will be able to access the benefits of drones and introduce new research methods.

Second, we present a comparative analysis of 2D and VR interfaces for drone teleoperation. We compare a stock interface created by Parrot with our novel interface to determine which interface is better suited for layman operation with minimal training. This will shape how future interfaces are designed and refined to further optimize navigational performance. Ideally, a user would be able to use a new interface with no training and successfully operate the drone.

II. System Implementation

For our drone platform, we employ a 6-DOF Parrot ANAFI drone.[1] By using a monocamera, this drone enables users to take pictures and scan the environment. For interfaces, we employ the FreeFlight 6[2] app as our baseline 2D interface, which enables users to control drones on their mobile phones. Using the Meta Quest 2[3] as the VR headset, we implemented a Mixed Reality cyber-physical control room interface to view the real-time video feed from the drone's camera and control

[1]https://www.parrot.com/us/drones/anafi

[2]https://www.parrot.com/us/apps-and-services

[3]https://www.meta.com/quest/products/quest-2

the drone (See Fig. 1). The control room is circular in shape with a railing surrounding it. A video screen mounted on the railing streams the live video from the drone. The screen is designed to follow the headset's orientation by default, and the users have the control to fix the location of the screen. We also use the passthrough feature of the Quest to replace the black background of the interface with the users' surroundings. This allows them to also have a direct line of sight to the drone. The users can interact with the interface and remotely operate the drone using the Quest's joysticks. The location and orientation of the drone is controlled by pose commands sent by the users. We designed the interface using Unity[4], and used ROS nodes [7] running on a laptop along with Olympe (Parrot SDK)[5] to connect the drone to the headset over WiFi. Fig. 2(a) illustrates the employed drone and Fig. 2(b) shows our VR headset.

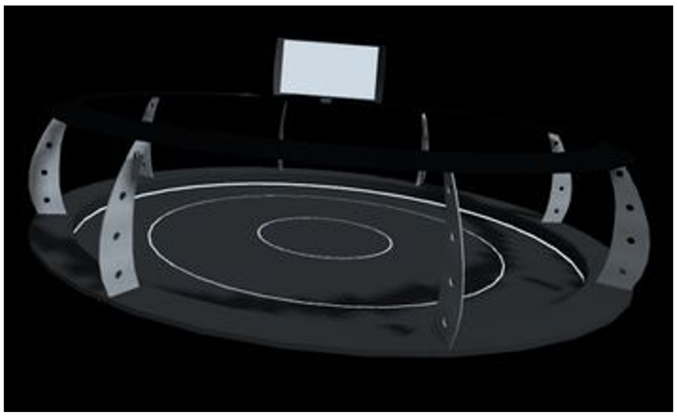


Fig. 1. Our Mixed Reality Cyber-Physical Control Room Interface

(a) Parrot ANAFI drone (b) Meta Quest 2

Fig. 2. The Drone and VR headset used in our study

III. Experiment Design

We performed an in-lab within-participants study measuring how well people were able to finish some basic and complex drone exploration tasks using the baseline 2D interface and our VR interface.

[4]https://unity.com/

[5]https://developer.parrot.com/docs/olympe/overview.html

A. Tasks

To better assess the performance of different interfaces, we employed a three-stage experiment design procedure which includes three tasks with different difficulties. Prior to formal tasks, users were explained the controls for how to fly drones with these interfaces.

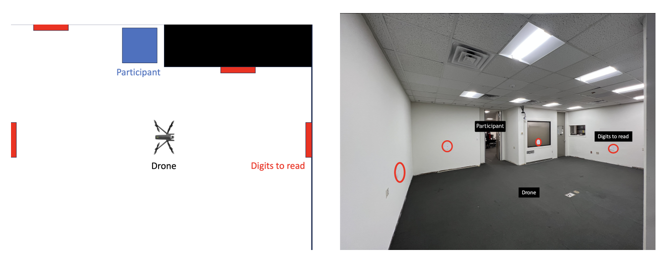
Task 1: Drone Navigation. In the first stage, we asked participants to achieve the most basic task to follow instructions for flying the drone without crashing. This task is simple to achieve and is designed to assess the basic ability of the interfaces.

Task 2: Drone Exploration. In the second stage, we asked participants to achieve a complex exploration task to read a single paper with 5 random digits that are located across the room from the drone. This task further orients participants with the robot platform and clarifies the subsequent task.

Task 3: Complex Drone Exploration. For the third stage, users are asked to finish a more complex task where they are asked to fly the drone and read 4 papers with 5 random digits in various positions around the room. We used this task as it required participants to combine multiple kinds of interactions to navigate complex paths of drones which might comprehensively evaluate the ability of interfaces. It also represents a basic statistical quantity that most lay participants would be able to complete.

B. Experiment Setup

The experiment was conducted in a room in the UNC Chapel Hill computer science department (See Fig. 3). The drone is located in the center of the room at the beginning of each task and the participant is located at the entrance. Each paper of digits has 36 pt Arial font, which forces the participant to move the drone within 1 meter of the paper to read it. We also ensure that the participants can't see the digits directly but have a direct line of sight to the drone at all times.



(a) Top down view (b) Actual room

Fig. 3. Our Experiment setup.

C. Hypotheses

This study allowed us to characterize the effect of types of interfaces in human-drone interaction. We hypothesized that:

H1: The VR-based interface will not reduce user performance when conducting basic tasks.

H2: The VR-based interface will be most effective for complex tasks.

H3: Users will prefer to use the VR-based interface compared to the 2D mobile app.

D. Procedure

Our experiment consisted of six phases: (1) Informed Consent, (2) Demographics Questionnaire, (3) Drone Navigation, (4) Drone Exploration, (5) Complex Drone Exploration, and (6) Exit Questionnaire.

At the beginning of the study, participants were provided and introduced with informed consent in accordance with our IRB protocol. After reading and accepting the informed consent, participants responded to a demographics questionnaire on robotics experience, VR usage, and weekly video game playtime. Then, they were read a description of the controls for each of the interfaces. Each participant was asked to complete all the three tasks twice using both the interfaces. Participants were asked to complete the tasks in order each time. The first interface of a given participant was counterbalanced between VR and smartphone interfaces to counteract transfer effects. Prior to each task, participants were read a script of instructions. We recorded the time elapsed and success rate (percentage of correct digits) of the tasks. Between tasks and interfaces, all sequences of digits were unique. Finally, participants were asked to answer an exit questionnaire about their opinions on the effectiveness and preference of the two interfaces on a 5-point Likert scale.

E. Participants

We recruited 6 participants from UNC-Chapel Hill using convenience sampling. 4 of them reported robotic experience, 4 of them reported VR experience, and 2 of them were hardcore video game players. All of our participants were found in the computer science building and reported normal vision or were wearing corrected glasses. Our experiment took about 20 minutes on average.

F. Analysis

We measured performance as both success rate and completion time spent on both of our tasks. However, since the overall success rate is 100% for our users, we analyzed the resulting completion time using a mixed-factors analysis of variance (ANOVA), with the type of interfaces and users' demographics, i.e., robot experience and VR headset experience.

IV. Results

We discuss significant results and statistical analysis based on the independent factors considered in this paper (see subsection III-F) using both traditional inferential measures and 95% bootstrapped confidence intervals (± 95% CI) for fair statistical communication [8]. Table I summarizes our ANOVA results for completion time as dependent variables respectively, where the *Source* column shows the independent variables, and the rest of columns are the $p-values$ in which T_1, T_2, and T_3 denote the number of tasks in order.

We denote $p-value < 0.01$ to be a significant effect as denoted by bold text in Table I. As we can see, the interface type would significantly impact the performance of Drone Exploration, and both interface type, VR experience, and robot experience would significantly impact the performance of Complex Drone Exploration.

TABLE I
ANOVA results for completion time.

Source	p-value T_1	p-value T_2	p-value T_3
Interface Type	0.15	**< .0001**	0.01
VR Experience	0.09	0.10	**0.006**
Robot Experience	0.18	0.07	**< .0001**

A. Drone Navigation

According to Table I, interface type would not significantly impact users' performance in Drone Navigation. As we can see from Figure 4 (a), the average task completion time of 2D interfaces, which is about 60 seconds, is shorter than those of VR interfaces which is 80 seconds. However, since we counted the time cost to learn to navigate drones, this performance difference might be caused by users' familiarity with mobile phones.

Hence, our results can partially support **H1**: we found that the overall performances of the VR interfaces are worse than 2D interfaces but without significant differences.

B. Drone Exploration

Table I indicates that interface type would significantly impact users' performance of this task with $p-value < 0.0001$. Similarly, Figure 4 (b) also indicates a similar trend where the average completion time of the VR interface users was about 30 seconds, while 2D users took more than 70 seconds for completing the task. VR performs significantly better than those with 2D interfaces.

As a consequence, we conclude these results can support **H2**: we found that for drone exploration, participants using VR interfaces performed significantly better than those using basic 2D interfaces in terms of decreasing processing time.

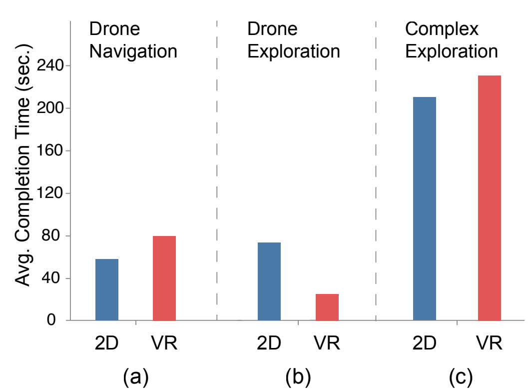


Fig. 4. The overall participants' performance (average completion time) of 2 interfaces in 3 tasks, (a) drone navigation, (b) drone exploration, and (c) complex drone exploration.

C. Complex Drone Exploration

As we can see from Table I, the $p - value$ of interaction type is 0.01, which indicates it might have a marginal effect but not a significant impact on the performance of Complex Drone Exploration. Besides, as it is shown Figure 4 (c), the performance of these two interfaces are similar, where 2D users took about 210 seconds and VR users took about 230 seconds on average.

According to the above discussion, we would say our results of Complex Drone Exploration cannot support **H2**: we found that for complex drone exploration, participants using VR interfaces might not spend less completion time compared to 2D interfaces. However, since this task is quite complex, these results might also be impacted by users' personalities, so we further explore their demographics' impact in subsection IV-E.

D. User Preference Analysis

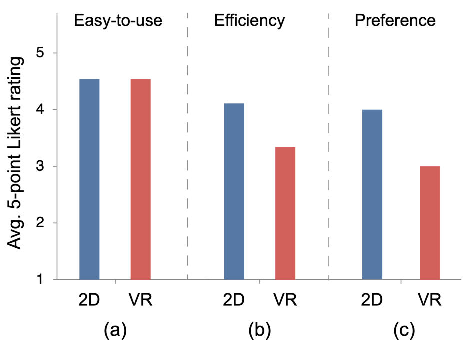


Fig. 5. Participants' average feedback scores of easy-to-use (a), efficiency (b), and preference (c) of interfaces. The three measurements were reported by each user on a 5-point Likert scale.

Figure 5 shows participants' feedback on their point of views of easy-to-use, efficiency, and preference. It's obvious participants think 2D interfaces are more efficient (although it performs a lot worse in Task 2) and they would prefer to use 2D interfaces. This finding indicates that users' preferences might not be faithfully correlated to their performance. But our experiment neither considered nor assumed such effects, so we would have chances to further study this impact with a larger group of users in the future.

Our results can not support **H3**: we found that our participants would agree VR interfaces are the same easy-to-use as 2D interfaces, but they would prefer to use 2D mobile applications compared to VR interfaces.

E. Exploratory Analysis of Users' Demographics

To better analyze other potential impact factors of drone teleoperation performances for Complex Drone Exploration, we further explore how participants' demographics (i.e., their robotic and VR headset experience) impact the performance of interactions.

As shown in Table I, we can find obvious significant impacts of VR and robot experience on participants' performance of Complex Drone Exploration. Figure 6 shows details of the average completion time categorized by users' demographics. This result shows that robot and VR experience significantly impact participants' performance of Complex Drone Exploration. Particularly, we can see participants with VR experience perform the best overall (see Figure 6 (b)), and take only about half time compared to those without VR experience.

Hence we conclude that the performance of Complex Drone Exploration (see Figure 4 (c)) was affected by users' own experience. Under such an assumption, our results still can support **H3**. However, we would not assert it as a finding, and we need future work to confirm this assumption.

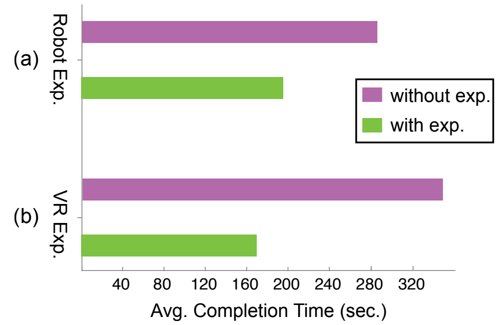


Fig. 6. The participants' performance (average completion time) of the complex exploration task categorized by their background of robot experience (a) and VR experience (b).

V. Conclusion

In this paper, we measured how the differences between 2D and VR interfaces impact people's ability to complete basic and complex teleoperation tasks using a mono-camera drone. Firstly, we implemented a VR interface that enables users to control a 6-DOF drone using a Meta Quest 2 Head Mounted Display. Secondly, we conducted a comparative user study to evaluate the performance of our VR and a 2D flat-screen interface on three drone teleoperation tasks.

Our results suggest that both interfaces perform similarly on the basic Drone Navigation task. But we found VR interfaces significantly outperform 2D flat-screen interfaces in processing time for the Drone Exploration task, and perform slightly lower than 2D interfaces for the Complex Drone Exploration task. However, our user preference analysis indicates that users prefer 2D interfaces more than VR ones. In addition, we reported a potential effect that users' personal experience in using both robots and VR headsets will significantly increase their performance, which might explain its lower performance in complex exploration tasks.

Based on the experimental results, we discussed our consistencies and inconsistencies with past findings on augmented drone interfaces and derived design implications for future VR interfaces. Please see Appendix B for further discussion about design choices and future work. We hope our work will inform future studies to understand VR performance in human-drone interactions deeply and construct more general guidelines for designing human-robot interfaces.

References

[1] D. Kent, C. Saldanha, and S. Chernova, "A comparison of remote robot teleoperation interfaces for general object manipulation," in *Proceed-*

ings of the 2017 ACM/IEEE international conference on human-robot interaction, 2017, pp. 371–379.

[2] F. Covaciu and A.-E. Iordan, "Control of a drone in virtual reality using mems sensor technology and machine learning," *Micromachines*, vol. 13, no. 4, p. 521, 2022.

[3] A. C. Hill, "Economical drone mapping for archaeology: Comparisons of efficiency and accuracy," *Journal of Archaeological Science: Reports*, vol. 24, pp. 80–91, 2019.

[4] I. Jeon, S. Ham, J. Cheon, A. M. Klimkowska, H. Kim, K. Choi, and I. Lee, "A real-time drone mapping platform for marine surveillance," *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, pp. 385–391, 2019.

[5] N. Dijkshoorn, "Simultaneous localization and mapping with the ar. drone," *PhD diss., Masters thesis, Universiteit van Amsterdam*, 2012.

[6] J. Callow, S. May, and M. Leopold, "Drone photogrammetry and kmeans point cloud filtering to create high resolution topographic and inundation models of coastal sediment archives," *Earth Surface Processes and Landforms*, vol. 43, no. 12, pp. 2603–2615, 2018.

[7] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, A. Y. Ng *et al.*, "Ros: an open-source robot operating system," in *ICRA workshop on open source software*, vol. 3, no. 3.2. Kobe, Japan, 2009, p. 5.

[8] P. Dragicevic, "Fair statistical communication in hci," in *Modern statistical methods for HCI*. Springer, 2016, pp. 291–330.

[9] M. Quigley, M. A. Goodrich, and R. W. Beard, "Semi-autonomous human-uav interfaces for fixed-wing mini-uavs," in *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566)*, vol. 3. IEEE, 2004, pp. 2457–2462.

[10] M. Quigley, M. A. Goodrich, S. Griffiths, A. Eldredge, and R. W. Beard, "Target acquisition, localization, and surveillance using a fixed-wing mini-uav and gimbaled camera," in *Proceedings of the 2005 IEEE international conference on robotics and automation*. IEEE, 2005, pp. 2600–2605.

[11] M. A. Goodrich, B. S. Morse, D. Gerhardt, J. L. Cooper, M. Quigley, J. A. Adams, and C. Humphrey, "Supporting wilderness search and rescue using a camera-equipped mini uav," *Journal of Field Robotics*, vol. 25, no. 1-2, pp. 89–110, 2008.

[12] Z. Lan, M. Shridhar, D. Hsu, and S. Zhao, "Xpose: Reinventing user interaction with flying cameras." in *Robotics: Science and Systems*, 2017, pp. 1–9.

[13] H. Kang, H. Li, J. Zhang, X. Lu, and B. Benes, "Flycam: Multitouch gesture controlled drone gimbal photography," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3717–3724, 2018.

[14] W. A. Isop, C. Gebhardt, T. Nägeli, F. Fraundorfer, O. Hilliges, and D. Schmalstieg, "High-level teleoperation system for aerial exploration of indoor environments," *Frontiers in Robotics and AI*, vol. 6, p. 95, 2019.

[15] V. M. Monajjemi, J. Wawerla, R. Vaughan, and G. Mori, "Hri in the sky: Creating and commanding teams of uavs with a vision-mediated gestural interface," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2013, pp. 617–623.

[16] T. Naseer, J. Sturm, and D. Cremers, "Followme: Person following and gesture recognition with a quadrocopter," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2013, pp. 624–630.

[17] J. Nagi, A. Giusti, G. A. Di Caro, and L. M. Gambardella, "Human control of uavs using face pose estimates and hand gestures," in *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2014, pp. 1–2.

[18] J. R. Cauchard, A. Tamkin, C. Y. Wang, L. Vink, M. Park, T. Fang, and J. A. Landay, "Drone. io: A gestural and visual interface for human-drone interaction," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2019, pp. 153–162.

[19] J. R. Cauchard, J. L. E, K. Y. Zhai, and J. A. Landay, "Drone & me: an exploration into natural human-drone interaction," in *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, 2015, pp. 361–365.

[20] L. Jane, L. Ilene, J. A. Landay, and J. R. Cauchard, "Drone & wo: Cultural influences on human-drone interaction techniques." in *CHI*, vol. 17, 2017, pp. 6794–6799.

[21] J. Bruce, J. Perron, and R. Vaughan, "Ready—aim—fly! hands-free face-based hri for 3d trajectory control of uavs," in *2017 14th Conference on Computer and Robot Vision (CRV)*. IEEE, 2017, pp. 307–313.

[22] D. Szafir, B. Mutlu, and T. Fong, "Communication of intent in assistive free flyers," in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, 2014, pp. 358–365.

[23] ——, "Communicating directionality in flying robots," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 2015, pp. 19–26.

[24] J. R. Cauchard, K. Y. Zhai, M. Spadafora, and J. A. Landay, "Emotion encoding in human-drone interaction," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2016, pp. 263–270.

[25] J. Thomason, P. Ratsamee, K. Kiyokawa, P. Kriangkomol, J. Orlosky, T. Mashita, Y. Uranishi, and H. Takemura, "Adaptive view management for drone teleoperation in complex 3d structures," in *Proceedings of the 22nd International Conference on Intelligent User Interfaces*, 2017, pp. 419–426.

[26] J. Thomason, P. Ratsamee, J. Orlosky, K. Kiyokawa, T. Mashita, Y. Uranishi, and H. Takemura, "A comparison of adaptive view techniques for exploratory 3d drone teleoperation," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 9, no. 2-3, pp. 1–19, 2019.

[27] C. Liu and S. Shen, "An augmented reality interaction interface for autonomous drone," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 11 419–11 424.

[28] D. Sainidis, D. Tsiakmakis, K. Konstantoudakis, G. Albanis, A. Dimou, and P. Daras, "Single-handed gesture uav control and video feed ar visualization for first responders," in *Proceedings of the International Conference on Information Systems for Crisis Response and Management (ISCRAM), Blacksburg, VA, USA*, 2021, pp. 23–26.

[29] J.-H. Wei, J.-J. Chen, and Y.-C. Tseng, "Mr-based uav route planning for the coverage task," in *International Summit Smart City 360°*. Springer, 2022, pp. 293–306.

[30] M. Walker, H. Hedayati, J. Lee, and D. Szafir, "Communicating robot motion intent with augmented reality," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018, pp. 316–324.

[31] M. E. Walker, H. Hedayati, and D. Szafir, "Robot teleoperation with augmented reality virtual surrogates," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2019, pp. 202–210.

[32] H. Hedayati, M. Walker, and D. Szafir, "Improving Collocated Robot Teleoperation with Augmented Reality," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '18. New York, NY, USA: Association for Computing Machinery, 2018, pp. 78–86.

[33] D. Whitney, E. Rosen, E. Phillips, G. Konidaris, and S. Tellex, "Comparing robot grasping teleoperation across desktop and virtual reality with ros reality," in *Robotics Research*. Springer, 2020, pp. 335–350.

[34] D. Whitney, E. Rosen, D. Ullman, E. Phillips, and S. Tellex, "Ros reality: A virtual reality framework using consumer-grade hardware for ros-enabled robots," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 1–9.

[35] R. Hetrick, N. Amerson, B. Kim, E. Rosen, E. J. de Visser, and E. Phillips, "Comparing virtual reality interfaces for the teleoperation of robots," in *2020 Systems and Information Engineering Design Symposium (SIEDS)*. IEEE, 2020, pp. 1–7.

[36] K. J. Mimnaugh, M. Suomalainen, I. Becerra, E. Lozano, R. Murrieta-Cid, and S. M. LaValle, "Analysis of user preferences for robot motions in immersive telepresence," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 4252–4259.

[37] L. Tang and G. Shao, "Drone remote sensing for forestry research and practices," *Journal of Forestry Research*, vol. 26, no. 4, pp. 791–797, 2015.

[38] W. R. Sherman and A. B. Craig, *Understanding virtual reality: Interface, application, and design*. Morgan Kaufmann, 2018.

[39] H. Kharoub, M. Lataifeh, and N. Ahmed, "3d user interface design and usability for immersive vr," *Applied sciences*, vol. 9, no. 22, p. 4861, 2019.

APPENDIX

A. Related Work

1) Traditional Human-Drone Interaction: Human-drone interaction has been studied for decades. Most traditional human-drone interaction interfaces employ joystick controllers or 2D flat screens. For example, Quigley et al. [9], [10] created PDA interfaces and TrackPoint controllers[6] to control fixed-wing UAVs and applied them to the wilderness search and rescue task [11]. Xpose [12] and Flycam [13] employed 2D touchable screens to provide complex functionalities such as screen gestures and multiple facet views for drone teleoperation. Isop et al. [14] proposed a 3D visual interface on laptop screens for automatic aerial exploration in indoor environments.

Additionally, gesture-based methods [15]–[18] have also been widely applied in human-drone interaction. Cauchard et al. [19] and Jane at al. [20] conducted user studies for gesture-based human-drone interactions and reported culturally-specific differences that would lead to performance variations in using gesture interactions between US and Chinese participants. Bruce et al. [21] introduced a face pose-based interaction method for drone teleoperation.

Under those interaction mechanisms, drones have been demonstrated to gain more functionalities and increase users' perception of safety through some user-centric studies, such as communicating intent using manipulations [22], communicating direction with LED visual signaling [23], and reflecting emotional states via different flight paths [24].

2) Augmented Human-Drone Interaction: Beyond the above traditional interactions, prior studies have introduced novel augmented interfaces such as VR and AR that incorporate multiple sources of information to ease teleoperation. For example, Thomason et al. [25], [26] introduced an automatic viewpoint selection algorithm from an omnidirectional camera using VR interfaces in 3D drone teleoperation. Liu and Shen [27] implemented an AR-based human-drone interface that allows users to explore space with an autonomous drone. Sainidis et al. [28] proposed an AR-based visualization system enabling single-hand gesture interaction to control UAVs for first responders. Wei et al. [29] developed a spray light-based MR interface that allows drone route planning for achieving the coverage task.

In addition to new interfaces, evaluating the impact of these interactions is also important. Walker et al. [30], [31] and Hedayati et al. [32] investigated the effectiveness of AR interfaces to impact the intent communication in human-drone interactions. They reported three of their designs significantly improve the intent communication performance and found implications that user prefers more explicit information during AR interactions.

However, it's still not clear how those new human-drone interfaces can impact the accuracy and preferences of users' usage scenarios. In this paper, we specifically focus on the impact of VR interfaces on users' performance in navigating drones for achieving tasks.

[6] http://www.handykey.com

3) VR Interface for Robot Teleoperation: Additionally, there are several comparative analyses of 2D vs VR interfaces for non-aerial robot teleoperation. For example, Whitney et al. [33] compared the performance of desktop interfaces and the ROS Reality VR interface [34] in achieving grasping teleoperation tasks using a Baxter robot. Hetrick et al. [35] studied the performance of the ROS Reality interface with different control settings on motor movement tasks with Baxter. Mimnaugh et al. [36] reported the impact of telepresence robots' autonomous movement on users' preferences immersed in the robot through VR headsets.

However, there are still few comparative analyses and limited understandings of the impact of VR interfaces on human-drone interaction and teleoperation. We aim to fill this gap to evaluate the performance difference between VR and flat-screen interfaces.

B. Discussion

We measured the impact of the differences between 2D and VR user interfaces on people's understanding, recognition, and usage of achieving tasks with drones. We find that the performances of 2D and VR interfaces are comparable for the basic Drone Navigation task. However, VR-based interfaces prove more effective for learning the Drone Exploration task. However, we see that this improved learning effect does not transfer to the Complex Drone Exploration task, proving that while VR interfaces might help learn simple exploration tasks faster, there still exists a learning gap for more complicated tasks. Our results provide new perspectives on prior findings and offer both actionable design guidance and opportunities for future research.

1) VR vs AR: Our study indicates that the VR interface does no harm to human-drone interaction efficiency and would lead to increased performance in exploration tasks. Though there lacks an existing study about VR-based drone interfaces, the results are comparable with previous studies about AR-based interfaces.

Our results (especially for the Drone Navigation Task) are consistent with the work of Liu and Shen [27] which found that AR interfaces would reduce command counts, but sacrifice accuracy rate and cost longer task completion time. But for Drone Exploration, our results indicate that VR would be better than 2D interfaces. However, Liu and Shen didn't introduce a learning phase in their user study and reported that their users still need to learn how to use AR during their formal tasks. So our users might benefit from our Drone Navigation task, which could be regarded as a learning phase for the later tasks. Another reason might be they show a 3D depth map in their 2D interface where as we let participants track the drone with their eyes.

Overall, those results might indicate that the performance of VR interfaces is comparable with AR interfaces. However, the prices of VR equipment (e.g., Quest 2 costs $349) are a lot cheaper than AR (e.g., HoloLens costs $3,500). So designing VR interfaces for human-drone interactions would be more achievable commercially. Hence we would recommend considering more VR interface designs for drone interactions.

2) Design Implications: Our results reveal that users' preferences vary from the design of VR interfaces which might influence their abilities to navigate drones. We provide preliminary guidance for VR-based drone interfaces design more broadly:

- **Reducing the drones' speed may improve the interaction performance in the indoor environment.**
  From our formal experiments and feedback, we found that in the indoor study environment, users cannot faithfully catch up with drones' paths when they're at a relatively high speed for both interfaces. And such high speed might lead to malfunctions too. This finding suggests we may need to reduce the maximum speed of drones when designing VR interfaces for indoor tasks.
- **Choose your interaction types that can best fit your own task.**
  Our experiments reveal that the VR interface could be a good choice to teleoperate drones for exploration tasks. While designers can use our results to directly compare with their demanded drone teleoperation tasks, however, as shown in subsubsection B1, our results might be slightly different from previous results focused on other teleoperation tasks. Since we didn't examine too many different tasks, we would recommend considering various interaction types when working on a largely different drone task. Our future work will assess whether these results can be replicated in a wider variety of drone teleoperation tasks.

3) Limitations and Future Work: We studied the impact of the VR interfaces on two drone teleoperation tasks. However, drones offer lots of teleoperation opportunities to achieve a wide variety of different tasks. Different tasks might provide different trade-offs than traditional drone-related tasks, such as exploring 3D spatial structures [25], [26], and domain-specific applications, such as forest resource monitoring [37]. Our future work will explore more details of the performance of different drone teleoperation tasks.

We considered the most basic interface implemented on a VR headset, and didn't explore the details of many design choices. However, VR environments provide us with a large number of design possibilities for applications [38]. We found that the abundance of information when teleoperating the drone still overwhelms novice users. Design choices like gesture control of the drone and using additional input like the orientation of the VR headset can be used to make drone flying more intuitive [39]. In future work, we will explore more design choices on different VR interfaces to provide more insights into design implications.

During our experiments, we also found some leakages and additional details to be expressed. The drone sometimes malfunctioned which could temporarily interrupt the experiment. The VR interface had a 0.5-1s delay in response to users' interaction. This is mostly because we were constrained by the network bandwidth and can be resolved by further improving the interface. However, users mentioned it as one of the reasons they had a hard time using the VR interface to fly the drone. We also set up an emergency control strategy to avoid collisions and hurt, but this seemed to confuse the participants when their controls were interrupted. Finally, propellers got damaged over time due to collisions which might impact their stability. For our future work plan to develop safety measures to prevent crashes or warn participants before time.

We discussed the difference between our work with some previous studies about AR-based interfaces (see subsubsection B1). We summarized the similarity and inconsistencies between their works and indicated that VR would have a similar performance with AR interfaces in these tasks. However, due to the project capacity limit, we didn't actually design an AR-based interface to faithfully explore the actual differences in performance and users' preferences between AR and VR interfaces. And we also didn't recruit a large group of users which can minimize the learning effect and demographics' impact. We plan to fix these gaps in our future work.