Immersive View and Interface Design for Teleoperated Aerial Manipulation

Basaran Bahadir Kocer[1*], Harvey Stedman[2*], Patryk Kulik[1*], Izaak Caves[1*],
Nejra Van Zalk[3],
Vijay M. Pawar[2], and Mirko Kovac[1,4]

Abstract

The recent momentum in aerial manipulation has led to an interest in developing virtual reality interfaces for aerial physical interaction tasks with simple, intuitive, and reliable control and perception. However, this requires the use of expensive subsystems and there is still a research gap between interface design, user evaluations and the effect on aerial manipulation tasks. Here, we present a methodology for low-cost available drone systems with a Unity-based interface for immersive FPV teleoperation. We applied our approach in a flight track where a cluttered environment is used to simulate a demanding aerial manipulation task inspired by forestry drones and canopy sampling. Through objective measures of teleoperation performance and subjective questionnaires, we found that operators performed worse using the FPV interface and had higher perceived levels of cognitive load when compared to traditional interface design. Additional analysis of physiological measures highlighted that objective stress levels and cognitive load were also influenced by task duration and perceived performance, providing an insight into what interfaces could target to support teleoperator requirements during aerial manipulation tasks.

Virtual reality and interfaces, telerobotics and teleoperation, human-centered robotics

I. Introduction

Unmanned aerial vehicles (UAVs) are being deployed in an increasing number of industries as they provide a means to interact and collect data from environments that are hazardous or

[1]Authors are with the Aerial Robotics Laboratory, Imperial College London, London SW7 2AZ, UK. [2]Authors are with the Autonomous Manufacturing Laboratory within the Department of Computer Science, University College London, Gower Street, WC1E 6BT, UK. [3]Nejra Van Zalk is with Dyson School of Design Engineering, Imperial College London, London SW7 1AL, UK. [4]Mirko Kovac is with Materials and Technology Center of Robotics at the Swiss Federal Laboratories for Materials Science and Technology, Switzerland. *Authors contributed equally to this work.

impassable to humans. For example, UAVs have been implemented mostly for observation and monitoring tasks [1], nuclear facilities for detection of radiation sources [2], and for search and rescue tasks. [3]. Generally, UAVs are used as a method of remote sensing; however, recent research has begun exploring the area of aerial manipulation in which UAVs physically interact with their environment [4]. Aerial manipulation can be broadly grouped into two major categories: 1) free-flight operation, where the UAV is able to fly with full degrees of freedom (DOF) and contact forces are small for negligible time, and 2) motion-restricted operation, where the system is required to physically interact with the environment for a sustained period of time with restricted DOF [5]. Examples of free-flight operation include grasping [6], transporting [7] and assembling [8], whilst examples of motion-restricted operation include physical structure inspection [9], canopy sampling [10], and interacting with physical components such as valves and switches [11], [12].

The research for teleoperated aerial manipulators is still in its infancy. Previous work has identified that traditional manual control of aerial manipulators is difficult even for experienced operators [13] and that completely new interface designs are required for optimal human-machine interaction (HMI) [14]. In addition, regardless of the UAVs level of autonomy, there will always be a need for operators to interface with robotic systems in scenarios with significant safety concerns or to solve functional errors [15]. Therefore, understanding optimal interface designs for aerial manipulation tasks is a high priority for their deployment in the wild.

Teleoperation research has identified the importance of understanding the relationship between interface design, operator situational awareness, and cognitive load. Situational awareness relates operators understanding of the interface to different elements in the remote environment and their capability to build an accurate model to make informed predictions and decisions [16]. Low levels of situational awareness and high cognitive load from information sources in the interface have been found to negatively impact operator performance in teleoperation tasks [17]. This can result in confusing the operator, reduce operational efficiency, and potentially result in operation mistakes [18].

In general, visual feedback for teleoperation is generated through streaming real-time video from a mounted camera on the UAV to a monitor to provide an egocentric view of the UAV. However, recent studies have shown that implementing more sophisticated interfaces can improve teleoperation performance, increase situational awareness, and reduce cognitive load by providing better representations of the robot's relative position in the remote environment. Examples of this

include exocentric viewpoints [19], [20], high FOV cameras [17], ecological interface designs that fuse multiple sensors into one viewpoint [21], and dense environment reconstruction [22].

A popular choice among UAV operators is streaming a video feed from a mounted camera to a head-mounted display (HMD) to provide a first-person view (FPV), as this yields a robust method of long-range non-line of site operation with high levels of immersion [23]. It has been previously hypothesised that FPV flight could be adapted for aerial manipulation tasks [24]; however, to the best of the authors' knowledge, an analysis of teleoperation performance and situational awareness of this interface design for aerial manipulation tasks has not been examined considering physiological metrics of operators.

To this end, this study presents a realised FPV teleoperation platform and a user study analysing its applicability in aerial manipulation tasks with self-reported and physiological measures. Specifically, the task is inspired by forestry sampling UAVs [25], [26], which can be considered a motion-restricted operation with the following operational requirements [27]:

- The UAV should include an onboard video feed to allow collecting samples beyond visual line of site.
- Operation of the UAV should be precise to ensure accurate sample collection.
- The operator should be able to safely navigate and sample from all sections of the tree depending on the sample type that is required.
- Once collected, the sample should have minimal exposure to its environment or the ground to avoid contamination.

The contributions of this work are as follows:

- Development of a low cost UAV teleoperation framework.
- Design of an intuitive Unity-based FPV VR teleoperation interface.
- User study to compare the effectiveness of the designed interface to a traditional monitor-based interface for canopy sampling style aerial manipulation tasks adhering to the above operation criteria.

II. System Design

As the majority of development took place in the spring of 2021 during COVID 19 restrictions in the UK, the selected UAV platform had to allow for safe development and testing in private residences. Therefore, the DJI Tello EDU was chosen due to its size, availability, low cost, and open-source API. A custom server was developed to interface with the Tello locally and open

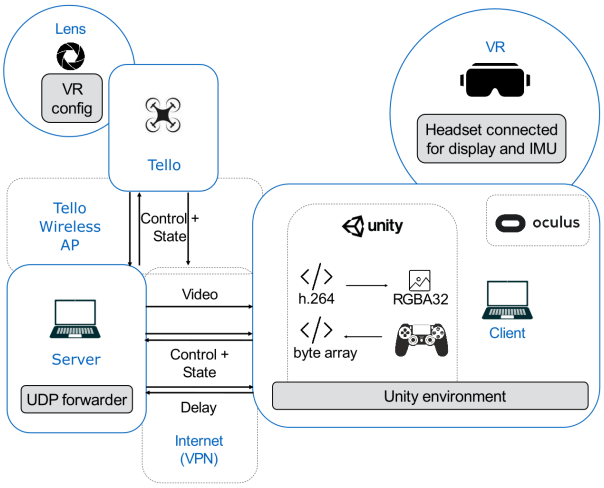


Fig. 1: Overview of system.

commands and its video stream to external operators using sockets, which then connected to a C# client created for Unity. Then, a custom Unity interface was developed to provide both traditional and immersive teleoperation capabilities which operators interacted with through a monitor or Oculus Quest 2, respectively. An overview of the system framework is shown in Fig. 1.

A. *Server-Client Architecture*

As the video feed obtained with Tello SDK was not suitable to conduct teleoperation, the low-level communication protocol used in the Tello's smartphone application was used instead [28]. The main server-client architecture is as follows: A base station runs locally to the Tello with two network access ports, one establishing a wireless connection to the Tello and one for external connections via the internet. The server running on this machine acts as a UDP proxy for all incoming command data packets and outgoing video data packets from the Tello. A custom C# client then connects to the server through external sockets and UDP packets delivered between the two machines provide bidirectional feedback for the Unity teleoperation interface (see Fig. 1). To maximise performance during operation, all processes in the server-client architecture are separated into different threads:

- *Command Thread:* The server listens to commands packets sent from the client and forwards them to the Tello.

Fig. 2: Tello drone with all accessories mounted.

- *State Thread:* The server listens to state packets sent from the Tello and forwards them to the client.
- *Video Thread:* An internal timer is started to track the delay between repeated video frame packets being sent from the Tello to the server. The server listens for video frame packets from the Tello and forwards them to the client.
- *Delay Thread:* When a new video frame is received, the client issues a ping packet to the server. The server responds with the recorded video delay time.

B. System Modifications

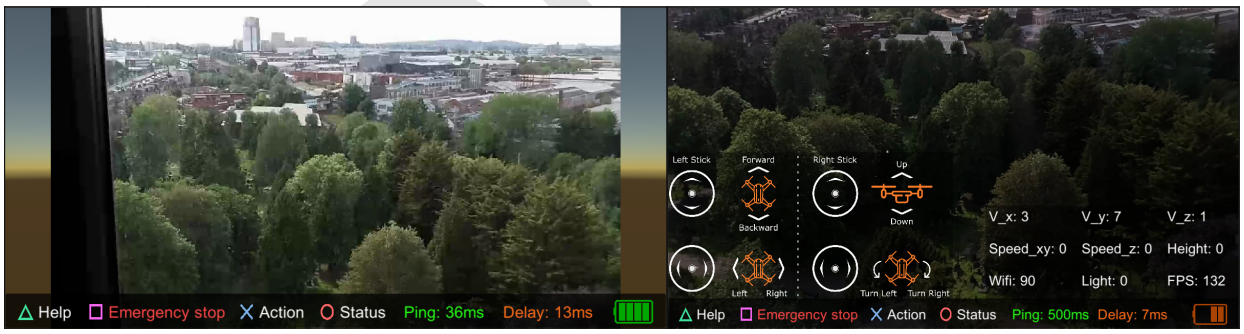


Fig. 3: The interface of the system (left) and the panel view to see the control scheme and drone states (right). Visualisation of the interface's control scheme and the state of the drone are Tello are off by default and can be viewed by pressing the "Help" and "Status" command respectively.

To improve situational awareness and provide a method to simulate forestry tasks, sampling modifications were designed for the Tello drone. A wide angle fish-eye lens was mounted to increase the FOV from a default 82.6°to approximately 120°. Additionally, a low weight sponge was mounted to the top of the Tello to simulate a sampling procedure during operation. More information on this can be found in Section III. The modified Tello drone is presented in Fig. 2.

C. Unity Interface

The Unity interface is built on top of the C# client, providing a means for representing the video and state data from the Tello and using Unity's input system for issuing commands to the Tello.

A Dualshock 4 controller was chosen as the input method for the interface. Using Unity's input system, button and joystick presses were configured using Unity events to connect the controller to the command thread in the underlying C# client. The controller's L1 and R1 buttons were mapped to take off and land respectively, the left joystick for lateral motion and the right joystick for ascending/descending and yaw rotations. Additional buttons on the controller were mapped for emergency landings and displaying the interface's control scheme and information on the Tello's current state. H.264 encoded video frames forwarded to the client are first decoded using FFMpeg to obtain the raw image data. This is then set to a texture and rendered in the Unity environment. Additionally, received state information is visualised by updating position and velocity information in the form of strings and visualising battery percentages as an icon in the interface. Finally, received ping and video delay data from the delay thread are also illustrated to represent connection quality during operation. The full user interface is presented in Fig. 3.

D. Immersive Interface Design

For the VR interface, extra utilities were added to increase the operator's immersion. As the fisheye lens increases the FOV of the mounted camera, it is possible to render the video feed on a curved screen and allow for the operator's HMD gaze to observe different parts of the environment to increase immersion. This was achieved by measuring the FOV properties of the mounted fisheye lens along the horizontal, vertical and diagonal axis, creating a mesh curved screen to match these properties and then using UV mapping to correctly render textures onto the new screen.

III. Experiment Setup

A user study was conducted to compare the presented VR teleoperation interface against a traditional monitor-based interface for aerial manipulation tasks requiring motion restricted operation. Inspired by forestry tasks and drone racing competitions, a course was designed to test the user's control, accuracy and perception of navigating obstacles during a simulated sample collection procedure considering the operational requirements as detailed in Section I.

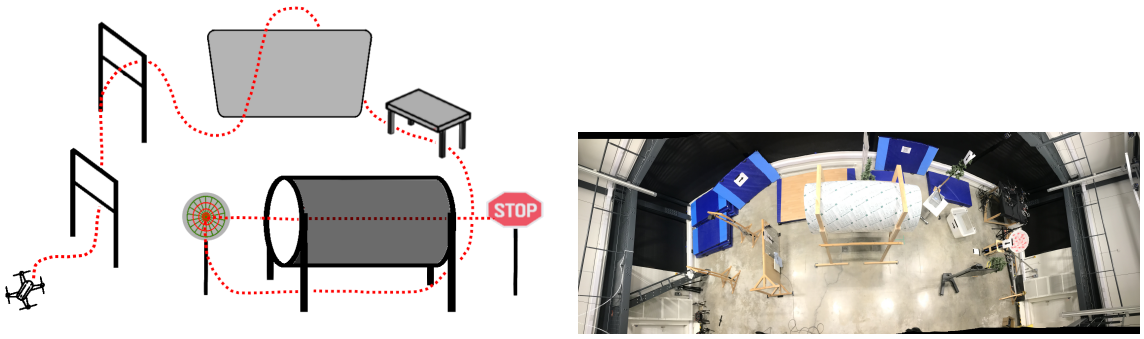
A. User Study Settings

The course had the following distinct steps that were completed in series:

- Obstacle avoidance:
  - Fly through two wooden frames.
  - Fly above a tree.
  - Fly below a table.
  - Fly below a tunnel.
- Physical manipulation:
  - Mark the target with the mounted sponge as accurately as possible (simulating a canopy sampling procedure).
  - Fly through a tunnel (simulating unpredictable flight conditions after taking a physical sample).
  - Land on a target.

For the experimental procedure, users first answered questions based on their demographic and experience with UAVs, VR and video games. Then, using the Polar H10 heart rate sensor, the baseline heart rate for each user was collected. Users then completed the experiment for both VR and non-VR cases in alternating orders to account for training bias. They were first introduced to the interface by completing a brief exploration task prior to the experiment. Each user then completed the course a minimum of three times whilst ensuring that at least two successful attempts of the course were completed, where a successful test is defined as landing at the end of the course without an abrupt end due to a significant crash. The data collected including validated psychometrics are as follows:

- Success rates at completing the experiment.
- Time to complete the course.
- Number and location of crashes and collisions.
- Location of sample collection on target.
- Heart rate and heart rate variability [29].
- NASA TLX Questionnaire [30].
- Situation Awareness Rating Technique (SART) Questionnaire [31].
- Usability Questionnaire with sections to provide qualitative feedback.



(a) (b)

Fig. 4: Course for user study experiment. a) Diagram of the course. b) Picture of the course.

B. *Psychometrics*

Heart rate variability is the measurement of the variance in time between heartbeats. HRV has been shown to be an effective metric for determining mental workload [32]. In short, a decrease in the HRV is a sign of an increase in mental workload. For this study, HRV was only available as an average of measurements taken at a low sampling rate of 1 Hz. As the heart rate monitor used in this study only reports an average bpm at low intervals, we conducted post-processing to extract the variability data:

1) For each measurement, the average R-R interval (RR) from the reported bpm is extracted.
2) Next, the root means square of successive differences (RMSSD) is calculated since it can quantify short term HRV [33]. For this, we find the mean of the squares of RR successive differences. After the mean is calculated, we find its square root to extract the RMSSD [34].

This will result in an approximate value of HRV to estimate trends in the variability throughout the study.

IV. Results and Discussions

A. *System Implementation*

A Dell XPS 15 laptop was used for the Tello server that facilitates the Wifi connection. For the client, a Zephyrus G14 laptop was used equipped with 16GB DDR4 RAM, AMD Ryzen 7 4800HS CPU and NVIDIA GTX 1660 Ti Max-Q GPU with 6GB VRAM. Throughout the

experiment, the DJI Tello was connected to the server through the Tello's Wifi access point, and the server and client were connected through an additional Wifi connection.

1) Participants: A total of 17 participants volunteered to participate in the user study (11 males and 6 females between 21-27 years old, mean age = 21.88). The recruitment process of the users followed after completing the ethical procedure from Imperial College London where the study is identified as low risk. The associated risks are mitigated with in-place precautions and preperations for the study. The call for recruitment is announced in associated email lists in the undergraduate departments and halls. Users were asked about their previous experience with UAVs, VR and video games which are presented in Table I.

B. Evaluation of User Study

1) Overall Results: The general trends in the results suggest that operators had worse performance, lower situational awareness and higher cognitive load when using the immersive FPV interface when compared to the traditional monitor-based interface. The performance metrics in Fig. 5 and statistical analysis presented in Table II show that generally operators took longer to complete the task ($M = 75.58 \pm 29.52$ s, $P < .001$), were less successful at landing on the designated target ($M = -0.35 \pm 0.39$, $P = .002$) and had an increased number of collisions ($M = 0.49 \pm 0.81$, $P = .029$) and crashes ($M = 0.36 \pm 0.28$, $P < .001$). The reduction in performance variables for the VR condition indicates that users had lower levels of situational awareness and higher cognitive load. It can be seen from Fig. 8 and Fig. 9 that in many of the measured parameters the VR condition performed worse. Notably, higher scores on the NASA TLX mental and physical demand, effort and frustration scored were associated with lower perceived performance. Full calculations of the TLX score [30] and situational awareness score [31] in Table II shows that there was a higher perceived cognitive load ($M = 7.47 \pm 4.99$, $P < .001$) and a lower perceived situational awareness ($M = -4.23 \pm 6.07$, $P = .011$) when

TABLE I: Summary of participant experience with drones, VR and gaming. Participants answered on a Likert scale with 1 representing no experience and 7 representing the highest level of experience.

Metric	Mean
Drone	2.29 ± 1.57
VR	2.00 ± 0.93
Gaming	3.88 ± 1.93

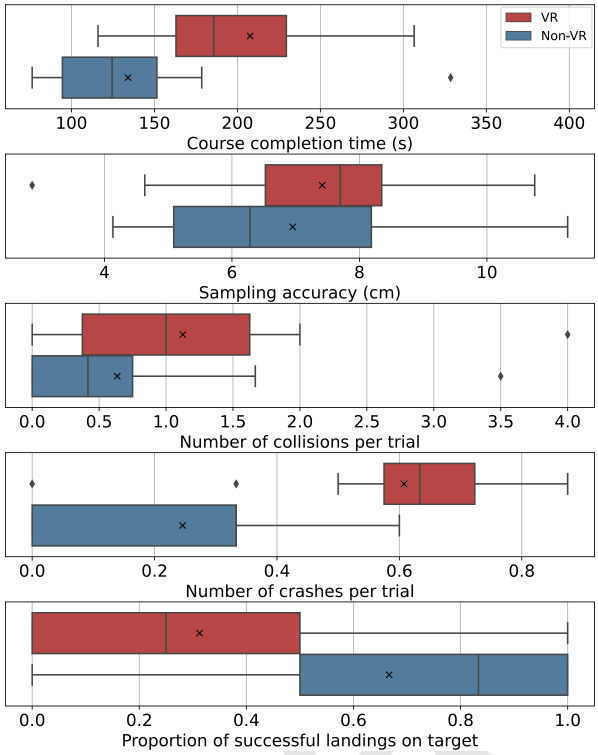


Fig. 5: Box plots of teleoperation performance metrics. Includes mean marked with a $\times$, median, lower and upper quartile, lower and upper whiskers and outliers marked with a $\diamond$.

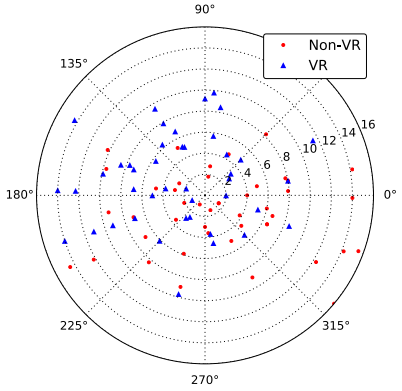


Fig. 6: Visual representation of target marking using the Non-VR and VR platform.

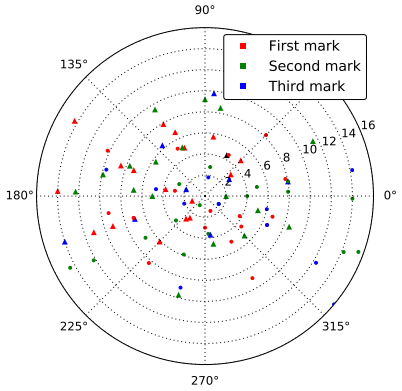


Fig. 7: Visual representation of target accuracy depending on the number of attempt.

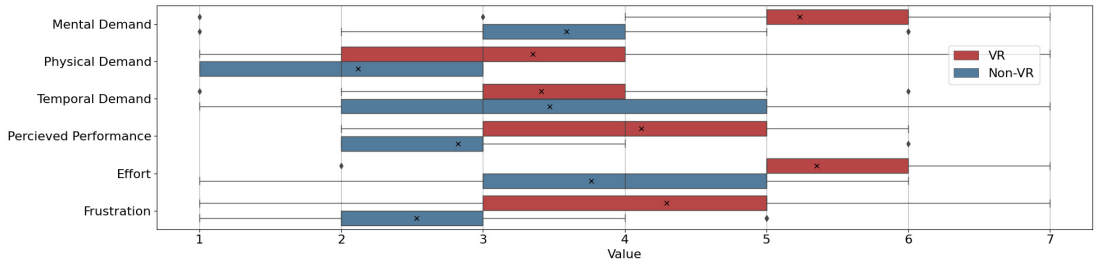


Fig. 8: NASA TLX Results. Includes mean marked with a $\times$, median, lower and upper quartile, lower and upper whiskers and outliers marked with a $\diamond$.

using the VR interface compared to the traditional interface. Objective variables of heart rate and heart rate variability were also collected; however, the analysis showed no significant differences in heart rate ($M = -0.11 \pm 5.73$ bpm, $P = .941$) or heart rate variability ($M = -0.33 \pm 1.61$ ms, $P = .410$) between conditions. This mirrors previous research showing a disassociation

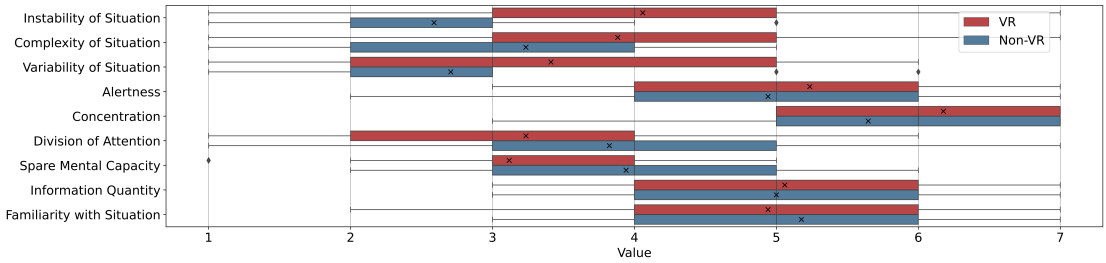


Fig. 9: SART Results. Includes mean marked with a $\times$, median, lower and upper quartile, lower and upper whiskers and outliers marked with a $\diamond$.

TABLE II: Summary of experiment results. All values represent the mean difference between VR and Non-VR performance.

Metric	Mean	p-value
Completion time (s)	75.58 ± 29.52	$< .001$
Sampling Accuracy (cm)	0.46 ± 2.57	.482
Collisions per trial	0.49 ± 0.81	.029
Crashes per trial	0.36 ± 0.28	$< .001$
Proportion successfully landed on target	-0.35 ± 0.39	.002
Heart Rate (bpm)	-0.11 ± 5.73	.941
Heart Rate Variability (ms)	-0.33 ± 1.61	.410
SART score	-4.23 ± 6.07	.011
TLX score	7.47 ± 4.99	$< .001$

between objective and subjective measures of cognitive load [35]. Interestingly, the results of the sampling procedure presented in Fig. 6 and Table II showed no significant difference in sampling accuracy between the conditions ($M = 0.46 \pm 2.57$ cm, $P = .482$). This suggests that the perceived changes in situational awareness and cognitive load between the VR and non-VR conditions did not affect the accuracy of the aerial manipulation task.

Collected feedback can explain the above trends. In the VR condition, many users commented that they found it difficult to orientate themselves in relation to the UAV, which could have negatively influenced cognitive load and situational awareness during the task. A decrease in video quality was also identified due to distortion from the mounted fisheye lens and the curved screen. Previous research has linked resolution of images for visual feedback to cognitive load in teleoperation tasks [17], and therefore it is possible that the increased immersion of the VR interface was overshadowed by the decrease in image quality leading to an increase in perceived cognitive load. Additionally, many users mentioned that the video feed from the drone had increased latency and delay when compared to the non-VR condition. Whilst this wasn't confirmed through system diagnosis, it is possible that operators in the VR interface were influenced by natural connection issues, as this would not only impact teleoperation capabilities but also affect immersion in the interface [36], [37]. Finally, it was mentioned that the control scheme for the VR interface felt more delayed than the traditional interface, which could explain the higher numbers of collisions and crashes. Potential reasons for this include the mounted fisheye lens influencing the dynamics of the Tello controller, and that operators found the control scheme more confusing within the VR interface.

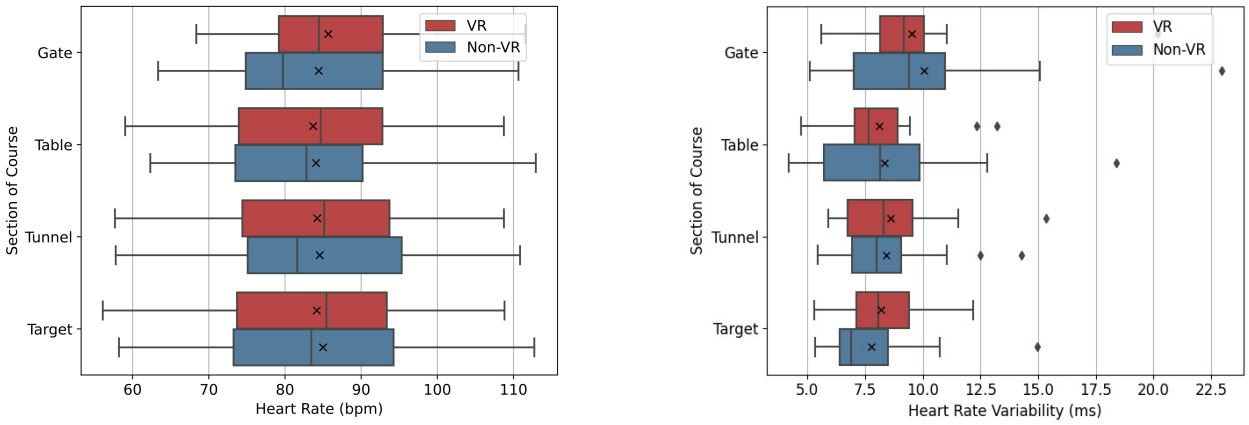
2) Differences in Physiological Metrics Between Experiments: Heart rate and heart rate variability were measured for each operator over the course of the experiment allowing for the metrics to be analysed in distinct sections. The distribution of mean measurements for heart rate and heart rate variability across each section is presented in Fig. 10a and 10b respectively.

From Fig. 10a, it can be seen that no significant difference is found in the heart rate measurements across the experiment for both conditions. However, heart rate variability results presented in Fig. 10b show a general trend, with measured heart rate variability decreasing between the first section "Gate" and final section "Target". For the Non-VR condition, it is noted that the difference between the Gate ($M = 10.07 \pm 4.24$ ms) and Table ($M = 8.35 \pm 3.51$ ms) sections (1.72 ± 3.07 ms difference, $P = .035$), and Gate and Target ($M = 7.78 \pm 2.42$ ms) sections (2.29 ± 2.84 ms difference, $P = .004$) can be considered significant ($P < .05$), although the difference between Gate and Tunnel ($M = 8.43 \pm 2.34$ ms) sections for the Non-VR condition is not considered significant ($M = 1.63 \pm 4.62$, $P = .0164$). Whilst the VR condition also follows this trend, the differences were not significant. This suggests that operators felt increasing levels of cognitive load and stress as the experiment went on, perhaps due to the concentration required to successfully navigate and complete the teleoperation task. The suspected reason that this effect is particularly notable in the Non-VR condition is because of the operator demographics and perceived increase in performance when compared to the VR condition, leading to higher HRV values early in the experiment. As aerial manipulation tasks could require significant operation time for deployed missions, exploring interface designs that reduce cognitive load over prolonged periods of operation is important for future work in the field.

3) Differences in Metrics Over Experiment Trials: Inspecting the net change of performance metrics presented in Table III, operators improved the course completion time in both VR ($M = -24.00$ s change, $p = .031$) and non-VR ($M = -22.62$ s improvement, $p = .003$) conditions.

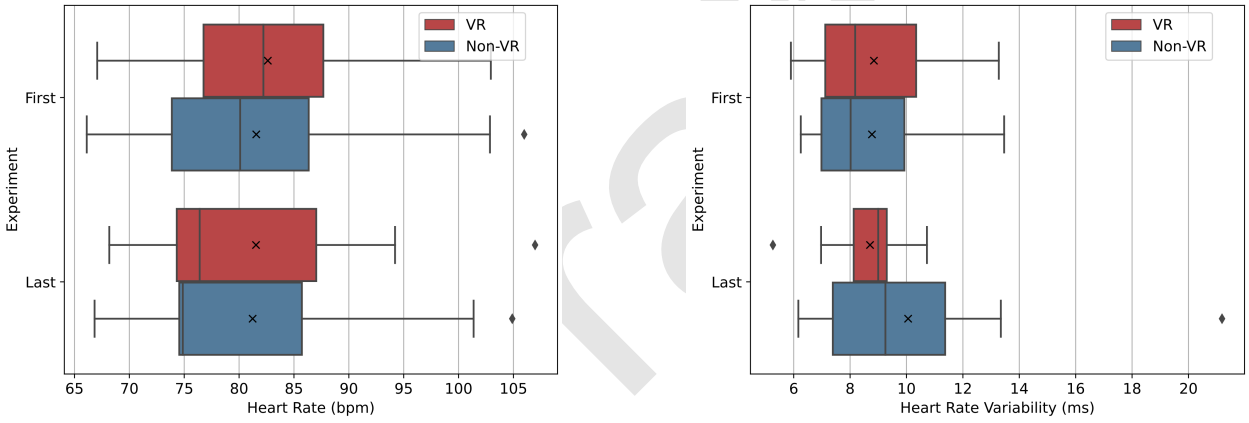
TABLE III: Mean change in metrics between the first and final tests.

Metric	Platform	Change	p-value
Completion Time [s]	VR	-24.00 ± 35.42	.031
	Non-VR	-22.62 ± 22.14	.003
Distance from target [cm]	VR	-0.08 ± 5.41	.954
	Non-VR	0.647 ± 4.55	.566



(a) (b)

Fig. 10: Physiological metrics across the individual sections of the experiment. a) Heart rate b) Heart Rate Variability. Includes mean marked with a $\times$, median, lower and upper quartile, lower and upper whiskers and outliers marked with a $\diamond$.



(a) (b)

Fig. 11: Difference in physiological metrics between the first and final experiment trials. a) Heart rate b) Heart Rate Variability. Includes mean marked with a $\times$, median, lower and upper quartile, lower and upper whiskers and outliers marked with a $\diamond$.

In contrast, despite the improvement in operation over the course, the sampling accuracy was largely the same for both conditions (VR: $M = 0.08$ cm change, $p = .954$. Non-VR: $M = -0.64$ cm, $p = .566$).

Exploring the changes in physiological metrics presented in Fig. 11, no changes in the heart rate are found for either condition (VR: -1.08 ± 3.62bpm change, $P = .305$. Non-VR: -0.33 ± 3.87 bpm change, $P = .763$). Heart rate variability measures presented in Fig. 11b also show no change for the VR condition ($M = -0.14 \pm 1.77$ ms change, $P = .783$) although a small increase is found for the VR condition (1.28 ± 2.61 ms change, $P = .102$). HRV has been

correlated with performance in repeated tasks [38], therefore it is reasoned that the operator's increased performance in the Non-VR condition influenced an increase in HRV, lower stress levels and cognitive load across the experiment trials. On the other hand, the VR interface saw no such change as it was more difficult to lean and use effectively. Perhaps the experiment did not adequately train the operators for complex aerial manipulation missions or FPV interfaces, which are considered to have a high learning curve. Therefore, a future user study might consider using a population of trained FPV UAV operators or operators accustomed to a specific aerial manipulation task to understand the effect of FPV interface design on expert users. Nevertheless, the results presented in this paper are useful in understanding optimal interface designs for aerial manipulation tasks with untrained operators, and also highlight the importance of improving training procedures for such tasks. Further investigation studies with a fully established VR interface and clinically approved sensors for psychometrics (e.g. E4 Empatica) are left for the extension of this study.

V. Conclusion

In this paper, a low-cost VR based FPV teleoperation platform for UAVs was presented. The system, based on the DJI Tello EDU and Unity game engine, was demonstrated to provide adequate teleoperation performance in VR over long distances with real-time video streaming and effective control. Using this platform, a user study was completed to compare FPV and traditional teleoperation interface design for aerial manipulation tasks inspired by forestry drones and canopy sampling. The results showed no notable performance change in the aerial sampling procedure between the two conditions. Despite heightened levels of VR immersion, interface operators had worse teleoperation performance, higher cognitive load and lower situational awareness when compared to the traditional interface. It seems that users' lack of experience of UAV operation and VR systems could have contributed to these results, as FPV interfaces require more training and practice to master. Additional results and heart rate variability demonstrated that subjective cognitive load levels may not match with objective cognitive load metrics and that objective cognitive load was influenced by the task duration and perceived performance. Future research should explore the discrepancy between people's perception of cognitive load and their actual physiological stress response, and how this information could be used to improve teleoperation training. Furthermore, a reactive-based controller will be implemented on the aerial system (e.g.,

optical flow or learning based) considering the data loss between the user and the robot [39], [40].

ACKNOWLEDGMENT

We acknowledge the funding of EPSRC (award no. EP/R009953/1, EP/L016230/1 and EP/R026173/1), NERC (award no. NE/R012229/1) and the EU H2020 AeroTwin project (grant ID 810321). Mirko Kovac is supported by the Royal Society Wolfson fellowship (RSWF/R1/18003).

REFERENCES

[1] B. B. Kocer, L. Orr, *et al.*, "An intelligent aerial manipulator for wind turbine inspection and repair," in *2022 UKACC 13th International Conference on Control (CONTROL)*. IEEE, 2022, pp. 226–227.

[2] J. Aleotti, G. Micconi, *et al.*, "Detection of nuclear sources by UAV teleoperation using a visuo-haptic augmented reality interface," *Sensors (Switzerland)*, vol. 17, no. 10, pp. 1–22, 2017.

[3] M. A. Goodrich, J. L. Cooper, *et al.*, "Using a mini-UAV to support wilderness search and rescue: Practices for human-robot teaming," *SSRR2007 - IEEE International Workshop on Safety, Security and Rescue Robotics Proceedings*, 2007.

[4] B. Stephens, L. Orr, *et al.*, "An aerial parallel manipulator with shared compliance," *IEEE Robotics and Automation Letters*, pp. –, 2022.

[5] X. Meng, X. Meng, *et al.*, "Survey on Aerial Manipulator: System, Modeling, and Control," *Robotica*, vol. 38, no. 7, pp. 1288–1317, 2020.

[6] R. Spica, A. Franchi, *et al.*, "Aerial grasping of a moving target with a quadrotor UAV," *IEEE International Conference on Intelligent Robots and Systems*, pp. 4985–4992, 2012.

[7] C. Wu, J. Qi, *et al.*, "Development of an unmanned helicopter automatic barrels transportation system," *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2015-June, no. June, pp. 4686–4691, 2015.

[8] A. E. Jimenez-Cano, J. Martin, *et al.*, "Control of an aerial robot with multi-link arm for assembly tasks," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 4916–4921, 2013.

[9] A. E. Jimenez-Cano, J. Braga, *et al.*, "Aerial manipulator for structure inspection by contact from the underside," *IEEE International Conference on Intelligent Robots and Systems*, vol. 2015-Decem, pp. 1879–1884, 2015.

[10] J. R. Kutia, W. Xu, *et al.*, "Modeling and characterization of a canopy sampling aerial manipulator," *2016 IEEE International Conference on Robotics and Biomimetics, ROBIO 2016*, pp. 679–684, 2016.

[11] C. Korpela, M. Orsag, *et al.*, "Towards valve turning using a dual-arm aerial manipulator," *IEEE International Conference on Intelligent Robots and Systems*, pp. 3411–3416, 2014.

[12] X. Meng, Y. He, *et al.*, "Contact Force Control of an Aerial Manipulator in Pressing an Emergency Switch Process," *IEEE International Conference on Intelligent Robots and Systems*, pp. 2107–2113, 2018.

[13] M. Orsag, C. Korpela, *et al.*, "Modeling and control of MM-UAV: Mobile manipulating unmanned aerial vehicle," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 69, no. 1-4, pp. 227–240, 2013.

[14] ——, "Valve Turning using a Dual-Arm Aerial Manipulator," in *2014 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, sep 2014, pp. 836–841.

[15] J. Y. Chen, E. C. Haas, *et al.*, "Human performance issues and user interface design for teleoperated robots," *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, vol. 37, no. 6, pp. 1231–1245, 2007.

[16] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors*, vol. 37, no. 1, pp. 32–64, 1995.

[17] M. S. Prewett, R. C. Johnson, *et al.*, "Managing workload in humanârobot interaction: A review of empirical studies," *Computers in Human Behavior*, vol. 26, no. 5, pp. 840–856, 2010.

[18] S. Shao, Q. Zhou, *et al.*, "Study of mental workload imposed by different tasks based on teleoperation," *International Journal of Occupational Safety and Ergonomics*, vol. 0, no. 0, pp. 1–11, 2020. [Online]. Available: https://doi.org/10803548.2019.1675259

[19] D. Saakes, V. Choudhary, *et al.*, "A teleoperating interface for ground vehicles using autonomous flying cameras," *Proceedings of 23rd International Conference on Artificial Reality and Telexistence, ICAT 2013*, pp. 13–19, 2013.

[20] F. Ferland, F. Pomerleau, *et al.*, "Egocentric and exocentric teleoperation interface using real-time, 3D video projection," *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction, HRI'09*, pp. 37–44, 2008.

[21] C. Nielsen, M. Goodrich, *et al.*, "Ecological Interfaces for Improving Mobile Robot Teleoperation," *IEEE Transactions on Robotics*, vol. 23, no. 5, pp. 927–941, oct 2007. [Online]. Available: http://ieeexplore.ieee.org/document/4339541/

[22] J. Du, C. Mouser, *et al.*, "Design and Evaluation of a Teleoperated Robotic 3-D Mapping System using an RGB-D Sensor," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46, no. 5, pp. 718–724, 2016.

[23] M. Jablonowski, "Beyond drone vision: the embodied telepresence of first-person-view drone flight," *Senses and Society*, vol. 15, no. 3, pp. 344–358, 2020. [Online]. Available: https://doi.org/10.1080/17458927.2020.1814571

[24] H. Bonyan Khamseh, F. Janabi-Sharifi, *et al.*, "Aerial manipulationâA literature survey," *Robotics and Autonomous Systems*, vol. 107, pp. 221–235, 2018. [Online]. Available: https://doi.org/10.1016/j.robot.2018.06.012

[25] B. B. Kocer, B. Ho, *et al.*, "Forest Drones for Environmental Sensing and Nature Conservation," *AIRPHARO 2021 - 1st AIRPHARO Workshop on Aerial Robotic Systems Physically Interacting with the Environment*, 2021.

[26] B. Ho, B. B. Kocer, *et al.*, "Vision based crown loss estimation for individual trees with remote aerial robots," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 188, pp. 75–88, 2022.

[27] G. Charron, T. Robichaud-Courteau, *et al.*, "The Deleaves: A UAV device for efficient tree canopy sampling," *Journal of Unmanned Vehicle Systems*, vol. 8, no. 3, pp. 245–264, 2020.

[28] "Low-level protocol," Feb 2021, accessed on Feb 16, 2022. [Online]. Available: https://tellopilots.com/wiki/protocol/

[29] H. M. Melo, L. M. Nascimento, *et al.*, "Mental fatigue and heart rate variability (hrv): The time-on-task effect." *Psychology & Neuroscience*, vol. 10, no. 4, p. 428, 2017.

[30] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in *Advances in psychology*. Elsevier, 1988, vol. 52, pp. 139–183.

[31] R. M. Taylor, "Situational awareness rating technique (sart): The development of a tool for aircrew systems design," in *Situational awareness*. Routledge, 2017, pp. 111–128.

[32] S. Delliaux, A. Delaforge, *et al.*, "Mental workload alters heart rate variability, lowering non-linear dynamics," *Frontiers in physiology*, vol. 10, p. 565, 2019.

[33] F. Shaffer and J. P. Ginsberg, "An overview of heart rate variability metrics and norms," *Frontiers in public health*, p. 258, 2017.

[34] A. B. Ciccone, J. A. Siedlik, *et al.*, "Reminder: Rmssd and sd1 are identical heart rate variability metrics," *Muscle & nerve*, vol. 56, no. 4, pp. 674–678, 2017.

[35] A. Luque-Casado, J. C. Perales, *et al.*, "Heart rate variability and cognitive processing: The autonomic response to task demands," *Biological Psychology*, vol. 113, pp. 83–90, 2016.

[36] B. Blissing, F. Bruzelius, *et al.*, "Effects of visual latency on vehicle driving behavior," *ACM Transactions on Applied Perception*, vol. 14, no. 1, 2016.

[37] M. Meehan, S. Razzaque, *et al.*, "Effect of latency on presence in stressful virtual environments," *Proceedings - IEEE Virtual Reality*, vol. 2003-Janua, pp. 141–148, 2003.

[38] R. Castaldo, L. Montesinos, *et al.*, "Heart rate variability analysis and performance during a repeated mental workload task," in *Embec & nbc 2017*. Springer, 2017, pp. 69–72.

[39] B. B. Kocer, M. A. Hady, *et al.*, "Deep neuromorphic controller with dynamic topology for aerial robots," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 110–116.

[40] F. Xiao, P. Zheng, *et al.*, "Optic flow-based reactive collision prevention for mavs using the fictitious obstacle hypothesis," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3144–3151, 2021.