Augmented Reality Interfaces for Semi-Autonomous Drones

Courtney Hutton*

Illusioneering Lab
University of Minnesota

ABSTRACT

With recent advances in hardware technology, drones are increasingly present in research activities outside of robotics, performing a multitude of tasks such as image capture and sample collection. However, user interfaces for task-oriented drones have not kept pace with hardware breakthroughs. Current planning and control interfaces for drones are not intuitive and often place a large cognitive burden on those who are not highly trained in their use. This paper proposes a course of research that seeks to leverage natural user interfaces in augmented reality (AR) for controlling drones completing task-oriented objectives.

Index Terms: Human-centered computing—Mixed/augmented reality—Visualization techniques—Treemaps; Human-centered computing—Visualization—Visualization design and evaluation methods

1 INTRODUCTION

As drones proliferate through the research sector and public sphere, the challenge of designing interfaces for human-drone interaction arises.[1] While position-based drone control is often solved through the use of a joystick or touchscreen interface, task-based control remains a more complex problem. Natural user interfaces, which seek to exploit innate skills such as gesture and voice, offer an intuitive method for configuring paths during task-based manipulations. They also have been demonstrated to reduce the cognitive load required when performing complex tasks [1]. This makes these interfaces an ideal tool to simplify the control of task-oriented drones.

Task-based control of drones can be viewed as an extension of path planning capabilities, where a drone is performing an action, such as image capture, in addition to traveling between points [1]. Path planning is a thoroughly studied task in robotics, and is known as the multi-agent pathfinding (MAPF) problem. Solutions generally incorporate some level of AI planning [2, 3, 6, 8], and have also been extended for multiple robots in obstacle-rich environments [10]. These approaches typically generate road maps to identify potential collisions, establish execution schedules, and refine these plans into smooth trajectories. Simulation is commonly used to validate this research for both robots and drones [3].

Augmented reality has not been formally researched as a medium for controlling drones completing task-based objectives, although the idea has been proposed [4]. However, mixed reality has been deployed as an interaction test bed [13] and has been proposed as a medium to test interactions between drones and virtual humans [7]. The open-source software for controlling a quadrotor drone swarm, Crazyswarm, [14] has been successfully implemented in a mixed-reality environment [13], although users did not have control over the paths of the drones.

*e-mail: hutto070@umn.edu

[1]For the purpose of this paper, *drone* refers to an unmanned aerial vehicle, while *robot* refers to an unmanned ground vehicle. Military drones are not referenced in this paper.

There have been investigations into novel user interfaces for controlling drones. The Microsoft Kinect has previously been implemented as one method for supporting gesture control of drones [11, 16], but this required users to memorize a set of full body poses to interact with the drone. Fernandez et al. [4] developed a graphical user interface (GUI) for drone control and tested a limited number of natural user interfaces. These included having the drone follow the user, or fly along a path defined by physical markers. Unfortunately, limitations exist with this approach, including the requirement that the user and the drone be co-located, and that the user can physically navigate the environment the drone occupies. Voice commands have also been proposed as a user interface [12], but as with previous approaches, this requires the user to memorize a predefined set of non-intuitive commands before operating the system. One research group proposed a using swarm components as the user interface: a new class of human-computer interfaces comprised of many autonomous robots that handle both display and interaction [9]. While certainly novel, this approach is unlikely to scale well.

While the interfaces described above could be described as more natural than traditional computer interfaces, none of them are truly natural interfaces. They also fail to address the complexities that accompany designing an interface for task-oriented drone control. Truly intuitive natural user interfaces leverage innate human skills, rather than requiring users to memorize a set of new interactions. Augmented reality offers a suite of tools that could be leveraged towards drone control, including intuitive, natural gestures and non-obstructive visualization of potential paths. This type of system requires minimal training for users and requires a lower cognitive load compared to traditional, controller- or graph-based planning systems. The remainder of this work will outline current progress and future steps that are necessary to realize AR as a natural user interface for task-based drone control.

2 OUTLINE OF RESEARCH

To accomplish the goals outlined in this paper, I will be working at the University of Minnesota with my thesis advisor, Dr. Evan Suma Rosenberg, and I will be collaborating with the Applied Motion Lab at the University of Minnesota.

2.1 Initial System Design

To begin developing an AR system for control task-oriented drones, we have started by developing a simple "point-and-go" system for robots. This has the advantage of confining mistakes to two axes for a slow moving vehicle, as opposed to three axes for a free-flying drone. In work accepted as a research demonstration for IEEE VR 2019, we demonstrate the basic capabilities of an augmented reality user interface for drone control using the Magic Leap Creator One. Design and implementation is occurring in Unity. This system is a proof-of-concept that marries path-planning algorithms such as RRT* from the Applied Motion Lab with gesture-based user input received from the Magic Leap. The path planning module generates a viable path offline, which can be visualized and modified within the environment prior to execution by the robot or virtual drone. The cognitive burden is minimized for the user, and the system requires requires minimal training. Subsequently, we intend to add fine-grained velocity and rotation controls for users seeking advanced applications.

2.2 User Studies and Interface Refinement

The next step in this research is to conduct user studies on the usability and effectiveness of this interface compared to traditional systems. The first user study is set to be conducted at the 2019 Minnesota State Fair. Participants will be asked to guide the drone through a simple obstacle course using either an augmented reality interface, or a traditional controller-based interface. Quantitative data regarding prior video game experience, in addition to qualitative feedback, will be collected from participants. We hope to analyze the time it takes to plan and complete the course and to examine the number of errors, such as crashes or misdirection, that occur for each interface.

Following the conclusion of this first experiment, we expect to refine the AR interface based on the results seen in the user study. We will then conduct a second study designed to examine the effectiveness of the interface for task-based objectives. The task performed will be photogrammetric capture of multiple objects, which has previously been conducted and evaluated for drones [5, 15]. The AR interface will be used to plan a photogrammetric capture path for multiple objects, and the resulting completion times and paths will be compared with those created using traditional methods.

The interface was designed in accordance with best practices for 3-D virtual interfaces. We hope that the results and feedback from these two studies will illuminate how these practices translate from virtual reality to augmented reality. After refining the interface, and assuming the studies yield insight into the design of AR interfaces, we expect to publish results detailing best practices for drone and robot control in AR.

2.3 Further Extensions

Following the refinement of the interface, we would like to leverage the real-time spatial mapping capabilities of devices like the Microsoft Hololens and the Magic Leap. We believe that this technology can be used by the general public and academics to easily plan drone flight paths. For instance, a user can plan a flight path for a drone to capture images or video of an event. Alternatively, a researcher could use the interface to construct a path for a drone to capture images of a sculpture to be reconstructed in 3-D. Ideally, we want this interface to lower the barrier to entry for using drones and other robots.

3 Contribution and Conclusion

Research in both augmented reality and drone hardware has been advancing rapidly, but best practices and applications have yet to emerge when designing interfaces for complex tasks. This doctoral research seeks to quantify the advantages offered by an augmented reality natural user interface and to develop a schema for how to implement these techniques in an interface for task-oriented drones. It also seeks to produce an system that can be used by even the most novice pilot, due to the intuitive design. Because this research is still in its infancy, I am seeking feedback on the design of the user studies, and on the metrics that will be used to evaluate this new system.

References

[1] 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 - UbiComp '15*, pp. 361–365. ACM Press, New York, New York, USA, 2015. doi: 10.1145/2750858.2805823

[2] L. Cohen, T. Uras, T. K. S. Kumar, H. Xu, N. Ayanian, and S. Koenig. Improved Solvers for Bounded-Suboptimal Multi-Agent Path Finding. *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 3067–3074, 2016.

[3] M. Debord, W. Hönig, and N. Ayanian. Trajectory Planning for Heterogeneous Robot Teams. *International Conference on Intelligent Robots and Systems*, 2018.

[4] R. A. Fernandez, J. L. Sanchez-Lopez, C. Sampedro, H. Bavle, M. Molina, and P. Campoy. Natural user interfaces for human-drone multi-modal interaction. In *2016 International Conference on Unmanned Aircraft Systems, ICUAS 2016*, pp. 1013–1022. IEEE, jun 2016. doi: 10.1109/ICUAS.2016.7502665

[5] J. Fernández-Hernandez, D. González-Aguilera, P. Rodríguez-Gonzálvez, and J. Mancera-Taboada. Image-Based Modelling from Unmanned Aerial Vehicle (UAV) Photogrammetry: An Effective, Low-Cost Tool for Archaeological Applications. *Archaeometry*, 57(1):128–145, feb 2015. doi: 10.1111/arcm.12078

[6] W. Hönig, T. K. S. Kumar, L. Cohen, H. Ma, H. Xu, N. Ayanian, and S. Koenig. Multi-agent path finding with kinematic constraints. In *Proceedings of the International Conference on Automated Planning Scheduling*, number 26, pp. 477–485, 2016.

[7] W. Hönig, C. Milanes, L. Scaria, T. Phan, M. Bolas, and N. Ayanian. Mixed reality for robotics. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5382–5387. IEEE, sep 2015. doi: 10.1109/IROS.2015.7354138

[8] W. Hönig, J. A. Preiss, T. K. S. Kumar, G. S. Sukhatme, N. Ayanian, W. Hönig, J. A. Preiss, T. K. S. Kumar, G. S. Sukhatme, and N. Ayanian. Trajectory Planning for Quadrotor Swarms. *IEEE Transactions on Robotics*, 34(4):856–869, aug 2018. doi: 10.1109/TRO.2018.2853613

[9] M. Le Goc, L. H. Kim, A. Parsaei, J.-D. Fekete, P. Dragicevic, and S. Follmer. Zooids. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology - UIST '16*, pp. 97–109. ACM Press, New York, New York, USA, 2016. doi: 10.1145/2984511.2984547

[10] H. Ma, J. Li, T. K. S. Kumar, and S. Koenig. Lifelong Multi-Agent Path Finding for Online Pickup and Delivery Tasks. *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 837–845, may 2017.

[11] A. Mashood, H. Noura, I. Jawhar, and N. Mohamed. A gesture based kinect for quadrotor control. In *2015 International Conference on Information and Communication Technology Research, ICTRC 2015*, pp. 298–301. IEEE, may 2015. doi: 10.1109/ICTRC.2015.7156481

[12] E. Peshkova, M. Hitz, and D. Ahlström. Exploring user-defined gestures and voice commands to control an unmanned aerial vehicle. In *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, vol. 178, pp. 47–62. 2017. doi: 10.1007/978-3-319-49616-0_5

[13] T. Phan, W. Hönig, and N. Ayanian. Mixed Reality Collaboration between Human-Agent Teams. In *25th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2018 - Proceedings*, pp. 659–660, 2018. doi: 10.1109/VR.2018.8446542

[14] J. A. Preiss, W. Hönig, G. S. Sukhatme, and N. Ayanian. Crazyswarm: A large nano-quadcopter swarm. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3299–3304. IEEE, may 2017. doi: 10.1109/ICRA.2017.7989376

[15] F. Remondino, L. Barazzetti, F. Nex, M. Scaioni, and D. Sarazzi. UAV photogrammetry for mapping and 3D modeling-Current status and future perspectives. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII-1/C22:25–31, sep 2012. doi: 10.5194/isprsarchives-XXXVIII-1-C22-25-2011

[16] A. Sanna, F. Lamberti, G. Paravati, and F. Manuri. A Kinect-based natural interface for quadrotor control. *Entertainment Computing*, 4(3):179–186, aug 2013. doi: 10.1016/j.entcom.2013.01.001