Augmented Reality-Mediated Human-Drone Interaction

Balasaravanan Thoravi Kumaravel
UC Berkeley
Berkeley, CA 94704, USA
tkbala@berkeley.edu

Bjoern Hartmann
UC Berkeley
Berkeley, CA 94704, USA
bjoern@eecs.berkeley.edu

Abstract
With an increase in use of drones for different purposes, scenarios arise in which humans will need to navigate shared environments with drones. In the worst case, this can lead to dangerous collisions. To avoid this, people have to anticipate the drone's behavior, which can be cognitively demanding if not impossible. One approach to solve this problem is to enable humans to visualize drone trajectories using a 3D augmented reality (AR) system, helping them to apprehend the drone's plan and act accordingly. We study the differences in human behavior in presence and absence of these trajectory visualizations. Our informal user study indicate that these visualizations ease the task load, reduce task time and help humans to envisage the drone's plan thereby enabling humans to adapt to it. Furthermore, we enable the user to create and modify the drone's trajectory and evaluate the interface. The work opens up research questions of using AR as a means to study Human-Robot Interactions as well as to facilitate robot control.

Author Keywords
Mixed Reality; Augmented Reality; Human Robot Interaction; Drones; Holographic User Interface;

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g.,HCI)]: Miscellaneous

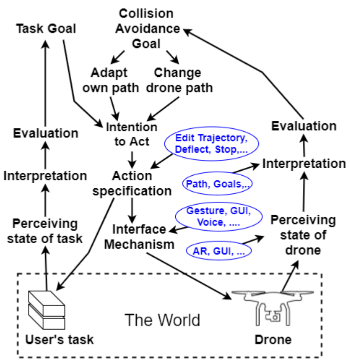

Figure 2: Frame work illustrating the gulfs of execution and evaluation present in the Human-Drone Interaction scenario carried out in our informal user study. The categories in Blue represent different possible techniques in which Human-Drone Interaction can be achieved.

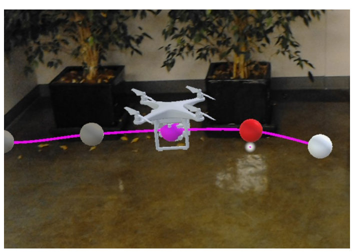

Figure 3: Visualization of the trajectory of the virtual drone. Visited waypoints are greyed, the immediately following waypoint is colored pink, the one following that is red, and rest are white.

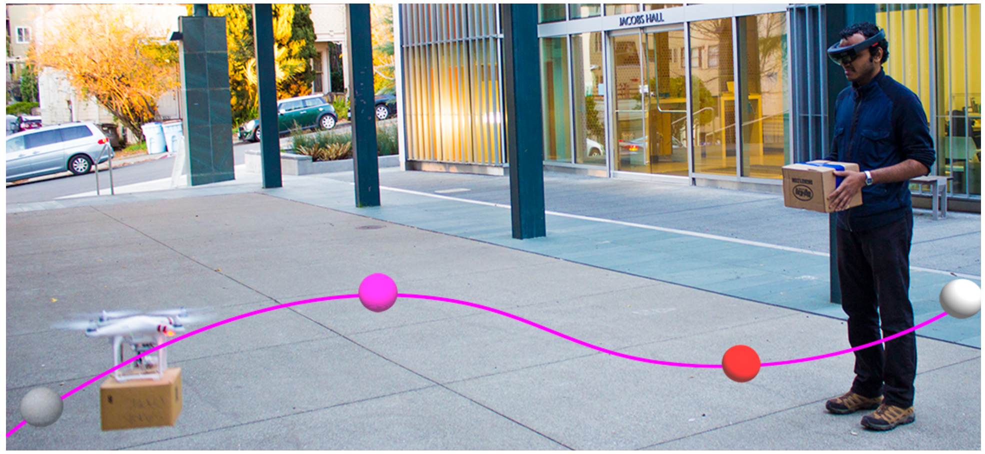

Figure 1: Augmented reality based trajectory visualization enabling navigation of humans in a shared space with drones

Introduction

Drones are flying robots that take different forms and sizes. Recently, they are being used for a variety of tasks like disaster management, geo-spatial mapping, navigation, product delivery, photography etc. It is possible that in near future we interact with them on a daily basis, which provides a strong case for Human-Robot Interaction(HRI) research specific to drones, commonly called Human-Drone Interaction (HDI). A common HRI problem is the expression of robot intent and a fair amount of prior works have dealt with this issue in non-aerial robots. In case of drones, this problem takes a new dimension, since drones are free-flyers [10] having six unconstrained degrees of freedom. This makes it challenging to navigate alongside drones in a shared environment. Some prior researches have contributed to a similar space for drones [6, 3, 4], but these rely only on the drone's responsibility to autonomously navigate, and for the most part treat humans as just another obstacle. In our work we take the approach where in a shared environment, people may want to avoid a drone by changing their own behavior, or they may want to actively influence what the drone is doing. Both require an understanding of the drone's flight plan as depicted by the gulfs of execution and evaluation [7] framework in Fig. 2. One possible way of communicating this is through visualization of its trajectories as pictured in Fig. 1. This work analyzes the difference such a visualization could have in Human behavior while interacting with drones.

Related Work

This work takes inspiration and builds on prior attempts to get drones to express some degree of intent to humans so as to aid in a better collaboration. Szafir et. al [10] modified drone trajectories to include elements from animation principles to convey information regarding drone's intended path and subsequently [11] embedded peripheral LED strips to convey drone's subsequent action. Cauchard et. al [2] tried to express drone's 'emotion' using particularly

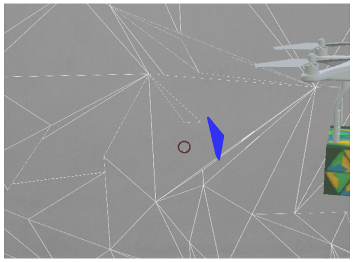

Figure 4: The user visual consists of the mesh of the space mapped by the Hololens which is represented by the white triangles. The blue drone direction pointer points to the direction of the drone and appears whenever the virtual drone is outside a certain margin of FOV of the Hololens.

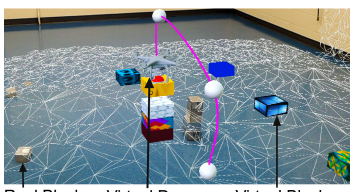

Figure 5: The figure shows the First person view of the user's Hololens feed during the user study. Colored blocks are virtual blocks and silver blocks are the real blocks. The feed shows the virtual drone picking and placing virtual blocks. Simultaneously the human is in the process of doing something similar with the real blocks.

characterized motion paths. Megha et. al [9] used a Laban effort based system for designing affective trajectories for quadrotors. These works convey drone's intent through indirect and subtle visual cues about the trajectory. We take a different stance to this by enabling the user a direct visual access to Drone's trajectory using Augmented Reality.

A closely related work by Rosen et. al [8] showed that using a head mounted mixed-reality system, one can infer quickly and accurately if a particular robot arm motion plan is collision-free. Though the work provides evidence for effectiveness of a mixed-reality system visualization, the task in the study did not require humans to be performing an active task which is usually the case in a real-life setting. This is a fundamental difference from our work, where we evaluate the effect of visualizations on Human-Drone Interaction, when each of them have their own independent task to be carried out in a shared physical space. This implies that human has to ensure completion of their own task while staying collision-free. The associated gulfs of execution and evaluation of the interaction [7] is illustrated by the block diagram in Fig. 2. Our work is targeted for drones which are mobile and have different interaction mechanism than that of a robot arm.

In our work, we also introduce an Augmented Reality based multi-modal control mechanism where human can tangibly manipulate the drone's path in an exocentric [10] passive manner. Though this is not the main focus of the work, it is a novel form of HDI that has not been explored previously.

ARMeHDI system design

A Head-mounted AR platform is used to enable the trajectory visualization. Microsoft HoloLens was chosen for the same due to its robust spatial mapping, multi-modal input and its ability to generate an immersive AR experience. Other methods such as using flat screen GUI, AR tablets, displays are not hands-free, require higher cognitive load from the user and have reduced usability [8]. In our system, the visualization consists of colored spheres and a 3D line passing through the centers of these spheres. The spheres represent the waypoints and the line represents the trajectory of the center of the drone as shown in Fig. 3. The visuals of the trajectory and waypoints are cleared once the drone reaches a desired goal, and a next set of visuals is shown for the subsequent goal. This reduces waypoint clutter and cognitive load of the user as compared to a scenario where the entire path of the drone for all the goals are shown at all times. The visualization includes a drone direction pointer which appears and points to the direction of the virtual drone when it is not well within the narrow 35° Field of View (FOV) of the Hololens, as seen in Fig. 4.

In order to collect quantitative information for analyzing differences in the human-drone interaction, a wireless channel was used to acquire important data. The data included the time series of human head's position and pose as measured by the IMU in the Hololens and the drone position. This data allows partial recreation of the interaction and enables us to calculate parameters that we use to evaluate the hypotheses. A typical recreation of the interaction using the acquired data is visualized in Fig. 7. The data can also be used to extract other behavioral traits of the interaction, though this has not been explored in this work.

User Study

The goal of the user study is to analyze the effect of trajectory visualization on human behavior while they perform a task which requires them to move in a shared space with the drone. A virtual DJI phantom drone projected into the AR headset is used in the study and spatial audio generated using real propeller audio sequences is used as an

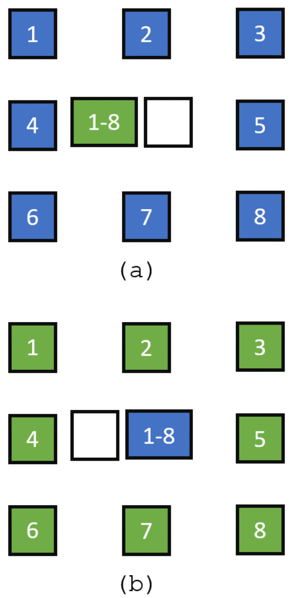

Figure 6: Initially in (a), there are eight virtual blocks (in Blue) placed in the grid and eight real blocks (in green) stacked over each other, placed close to the center. The drone picks up virtual blocks from the edge and stacks it in the center. Simultaneously the human is required to pick up the real blocks, place and arrange them in edge positions of the grid without colliding with the drone while trying to finish the task (b) in minimum time.

additional cue. This is to eliminate any potential safety hazards in the user studies. Though it may not entirely substitute a real drone like that envisioned in Fig. 3, efforts were taken to offer a close-to-real user experience. A feed of the first person view while performing the task is seen in Fig. 5.

The user study carried out was informal and had ten participants. Out of the ten participants, four had prior experience using AR/VR headsets. The task involved users to move in a space in the presence of the virtual drone and carry out some task while the drone itself performs some motion/task. The participants need to avoid collision with the virtual drone while simultaneously achieving their task in a minimum time. Our study involved a task where the users have to pick, place and arrange rectangular blocks in a the edges of a 3x3 grid as shown in Fig. 6

The conditions of the experiment are (1) *Vis* - Full trajectory visualization present and (2) *NoVis* - No visualization. *Vis* had the entire trajectory visualized in form of colored 3D lines and waypoints. The experiment was designed in a within-subjects manner for *Vis* and *NoVis*.

For analyzing effect of the two conditions, four parameters were used - (1) Time taken to complete the entire task, (2) Distance (time-averaged) maintained from the drone, (3) User's rating on adapting to drone trajectories and (4) Work load as measured by NASA Task Load Index [5]. Data on participant's rating on perceived 'realness' of the drone, rating on usability, qualitative feedback on control of trajectories and the overall experience was also collected.

Hypotheses

Visualizations enable more clarity of the drone's path. Thus we can expect users to identify the drone's goals quickly and accurately, thereby planning and executing their own actions easily and quickly. This leads us to our hypotheses H1: $T_{Vis} < T_{NoVis}$; where T_{Vis} denotes time taken to complete the task under *Vis*.

Additionally, we can also expect this to reduce task load on the user. In this study, we measured this using the NASA-Task Load Index. Emphasis was on mental demand, effort and frustration with these parameters weighted two times as compared to physical, temporal demand and performance. Hence our second hypothesis takes the form H2: $L_{Vis} < L_{NoVis}$; where L_{Vis} denotes task load in *Vis*.

Since visualization informs users about the drone's intended trajectory, we can expect users to keep their distance from it resulting in increased distance from the drone. Further, this could increase user tendency to adapt their actions in response to drone's path visualization. These form the basis of our hypotheses H3: $D_{Vis} > D_{NoVis}$ and H4: $A_{Vis} > A_{NoVis}$; Where D_{Vis} denotes average drone-user distance in *Vis*, A_{Vis} the user's rating (on a 5-point Likert scale) to adapt to the drone under *Vis*.

Findings

In the informal user study carried out, the following was found in the context of the formulated hypotheses. It is to be noted that since the study was carried out with just ten participants, the findings are to be treated as preliminary results only. The time taken as well as the task load index to complete the tasks was found to be consistently less in *Vis* as compared to *NoVis* for all participants. The time in *Vis* was reduced by around 28% on average when compared to *NoVis*. Measured task load index for reduced from 66.33 in *NoVis* to 49.05 in *Vis*. These results provide strong evidence in favor of hypotheses H1 and H2. User interviews revealed that the biggest advantages of visualizations were that it removed uncertainties in drone's actions thereby letting go of actively observing the drone, enabling them to focus on their own task.

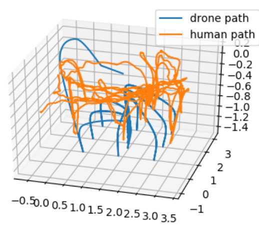

Figure 7: The plot shows the recreation in 3D space of a Human-Drone Interaction from an informal user study, using the time series data that was collected. The orange lines and the blue lines represent the positions of the human head and the drone respectively.

Nine of ten participants rated *Vis* to be better than *NoVis* in terms of their ability to adapt to drone's trajectory with an average rating of 4.2 for *Vis* as compared to 3.11 for *NoVis*, which is in favor of H4. They felt that visualizations helped them figure out the drone's goals ahead of time which encouraged them to be adaptive to drone's actions. This was also strongly correlated to participant's rating of their comfort level, usefulness of visualizations as well as overall experience of the interaction. Two participants felt that the limited Field of View (FOV) led to occlusion and distraction thereby countering the benefit of intent communication to some degree.

While testing for the third hypothesis, observations were were not consistent with all users. Visualizations encouraged some users to keep an additional distance from the drone's path in case of *Vis*. While at same time, for other users the visualizations increased their confidence on drone's path enabling them to get much closer in *Vis* as opposed to *NoVis*. Hence, H3 could not be validated.

Discussion, Limitations and Future work

Results on hypothesis 3 indicate that some participants tend to get closer to the drone in *Vis*. This raises an important question of whether the participants behave in a similar manner if the drone was real? In case of real drones, there is some uncertainty associated with its position control due to various real-world factors. Hence to be safe, the user may not be willing to get as close with real drone. This necessitates the requirement of a study to compare human behavioral difference as compared to AR versus real robots. Additionally, this also brings up the need to include visualizations of the trajectory uncertainties coupled with that of the path, which needs to be done in tandem with deploying better trajectory generation algorithms based on the drone's control algorithm.

Despite the presence of a drone-direction pointer shown in Fig. 4, the limited FOV of the Hololens proved to be a challenging hurdle for the interaction and was explicitly mentioned by all the participants. Some participants felt that the line could be more information than required and could possibly distract users from their tasks. These factors need to be addressed and subsequently the system is to be studied further with more number of participants.

The system also allows users to add, remove waypoints using speech and moving waypoints using hand gestures. The gestures are designed to control and place waypoints beyond user's arm space as well. This modification of trajectory happens in real-time which simulates an interaction experience of being able to modify the path of the drone while it moves in the space. The task in user study did not involve any manipulations of waypoints, though after the completion of the study participants were allowed to manipulate the waypoints and change the trajectory of the drone. The participants rated the developed system's usage based on the SUS questionnaire [1] with an average score of 81.25. In the future, we intend to explore how this AR interaction mechanism coupled with visualization can affect the HDI. Such an interface after addressing the problems mentioned in this section could serve as a tool for studying Human-Robot Interaction problems using simulated virtual robots. As depicted by the framework in Fig. 2, it is to be noted that many techniques exist for perceiving drone states and executing user actions. Though in this work we use AR based path visualization for the former and propose AR for the latter, there is scope for exploring other combinations for the different tasks that fit into this framework. These may lead to varying levels and types of human behavior adaptations and the robot control actions by the human interacting with the robot.

Conclusion

In this work we have designed and deployed an Augmented Reality (AR) based Human-Drone Interaction system that enables users to visualize and manipulate drone trajectories. Informal user studies gave some insightful information on how such a visualization affects human interaction with drones. It was seen that the visualization enabled humans in presence of drone to deduce the drone's goal and intent, which helped reduce the time taken by humans to complete a specific task. This work also developed novel holographic interactions which enables users to act by directly manipulating the drone trajectories in real-time.

REFERENCES

1. John Brooke and others. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.
2. Jessica Rebecca Cauchard, Kevin Y Zhai, Marco Spadafora, and James A Landay. 2016. Emotion encoding in human-drone interaction. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*. IEEE Press, 263–270.
3. Francisco de Babo Martins, Luis F Teixeira, and Rui Nóbrega. 2016. Visual-Inertial Based Autonomous Navigation. In *Robot 2015: Second Iberian Robotics Conference*. Springer, 561–572.
4. Allen Ferrick, Jesse Fish, Edward Venator, and Gregory S Lee. 2012. UAV obstacle avoidance using image processing techniques. In *Technologies for Practical Robot Applications (TePRA), 2012 IEEE International Conference on*. IEEE, 73–78.
5. Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology* 52 (1988), 139–183.
6. Alexandros Lioulemes, Georgios Galatas, Vangelis Metsis, Gian Luca Mariottini, and Fillia Makedon. 2014. Safety challenges in using AR. Drone to collaborate with humans in indoor environments. In *Proceedings of the 7th International Conference on PErvasive Technologies Related to Assistive Environments*. ACM, 33.
7. Donald A. Norman. 2002. *The Design of Everyday Things*. Basic Books, Inc., New York, NY, USA.
8. Eric Rosen, David Whitney, Elizabeth Phillips, Gary Chien, James Tompkin, George Konidaris, and Stefanie Tellex. 2017. Communicating robot arm motion intent through mixed reality head-mounted displays. *arXiv preprint arXiv:1708.03655* (2017).
9. Megha Sharma, Dale Hildebrandt, Gem Newman, James E Young, and Rasit Eskicioglu. 2013. Communicating affect via flight path Exploring use of the Laban Effort System for designing affective locomotion paths. In *Human-Robot Interaction (HRI), 2013 8th ACM/IEEE International Conference on*. IEEE, 293–300.
10. Daniel Szafir, Bilge Mutlu, and Terrence Fong. 2014. Communication of intent in assistive free flyers. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. ACM, 358–365.
11. Daniel Szafir, Bilge Mutlu, and Terry Fong. 2015. Communicating directionality in flying robots. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 19–26.