PinpointFly: An Egocentric Position-pointing Drone Interface using Mobile AR

Linfeng Chen
clinfeng@riec.tohoku.ac.jp
Tohoku University, Japan

Akiyuki Ebi
ebi-mcs@riec.tohoku.ac.jp
Tohoku University, Japan

Kazuki Takashima
takashima@riec.tohoku.ac.jp
Tohoku University, Japan

Kazuyuki Fujita
k-fujita@riec.tohoku.ac.jp
Tohoku University, Japan

Yoshifumi Kitamura
kitamura@riec.tohoku.ac.jp
Tohoku University, Japan

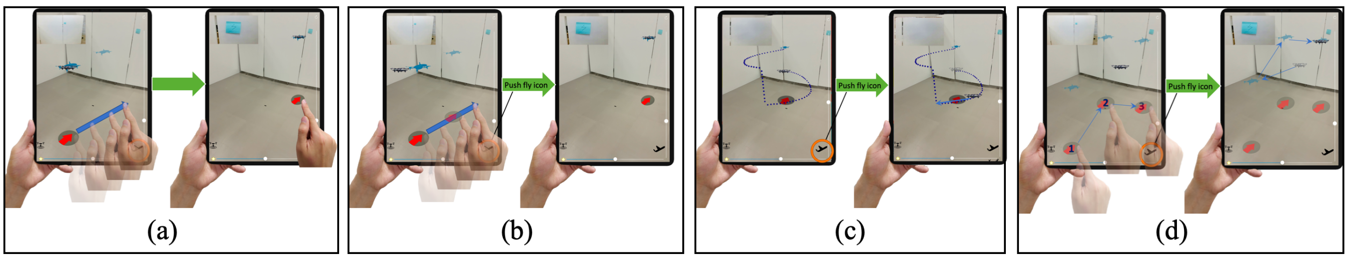

Figure 1: Four interaction techniques of PinpointFly: (a) DragFly, (b) PointFly, (c) TrajectoryFly, (d) WaypointFly

ABSTRACT

We propose PinpointFly, an egocentric drone interface that allows users to arbitrarily position and rotate a flying drone using position control interactions on a see-through mobile AR where the position and direction of the drone are visually enhanced with a virtual cast shadow. Unlike traditional speed control methods, users hold a smartphone and precisely edit the drone's motions and directions by dragging the cast shadow or a slider bar on the touchscreen.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**.

KEYWORDS

Visualization, Depth Perception, Targeting, Videography

ACM Reference Format:
Linfeng Chen, Akiyuki Ebi, Kazuki Takashima, Kazuyuki Fujita, and Yoshifumi Kitamura. 2019. PinpointFly: An Egocentric Position-pointing Drone Interface using Mobile AR. In *SA '19 Emerging Technologies, November 17-20, 2019, Brisbane, QLD, Australia.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3355049.3360534

SA '19 Emerging Technologies, November 17-20, 2019, Brisbane, QLD, Australia

ACM ISBN 978-1-4503-6942-8/19/11.
https://doi.org/10.1145/3355049.3360534

1 INTRODUCTION

Quadcopter drones with high-quality video cameras have been increasingly applied to many domains. Examples include troubleshooting, architecture inspection, and nature observation. Although automated drone operations are becoming more intelligent and prevalent, drone operations that rely on visual line of sight (VLOS) remain the leading operation methods. With such methods, users simultaneously look at the flying drone and the first-person view (FPV) of the drone camera from a screen. For successful VLOS operations, precise control of drone position and direction is critical but it remains challenging due to two issues: (1) Users have insufficient spatial perception of the position and the direction of the flying drone from their own perspective. (2) The speed control interface with a joystick is not intuitive for beginners and ill-suited for precise positioning tasks [Zhai 2002]. These two issues often appear together and complicate the mapping from the two different stick inputs to the drone movements. For example, if a user fails to clearly estimate the flying drone's depth distance, it is hard to precisely position it to target locations. These issues and difficulties considerably decrease piloting performance and cause potential dangers and risks. Recently, to improve the controllability of drones, previous work employed AR technologies [Chen et al. 2018; Yonezawa and Ogawa 2015] to make the interaction more natural and direct. However, there is still no attempt improving user's spatial perception and control intuitiveness during drone piloting. exTouch [Kasahara et al. 2013] uses a spatially aware egocentric approach to manipulate moving physical objects (e.g., drones) through an AR-mediated mobile interface. The system enables spatial projection of the user's input into the actuated target's motions with an egocentric interaction through a see-through mobile AR. However, it fails to achieve precise drone positioning and did not completely discuss direction control for flying drones.

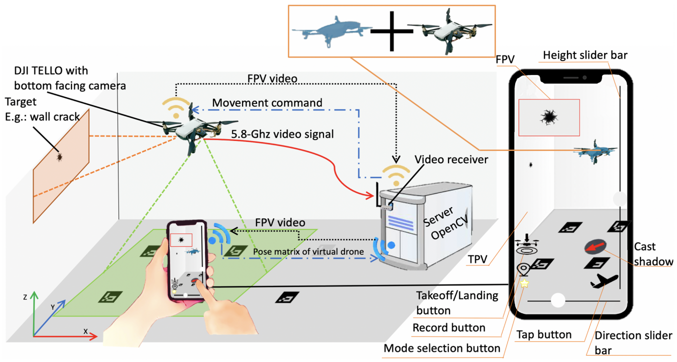

Figure 2: System workflow and user interface

In this demo, we propose PinpointFly, an egocentric drone interface which allows users to arbitrarily edit the flying drone's next position and direction using touch-based interactions over a see-through mobile AR where the drone position and direction are visually enhanced and indicated. PinpointFly employs an off-the-shelf motion-tracking system and see-through mobile AR technologies to achieve highly accurate drone control by position control, intuitive egocentric interaction using a spatially consistent coordinate system between the drone and user, and her enhanced spatial perception of the drone by showing its cast shadow. Fig. 1 shows four representative interaction techniques based on our proposal. DragFly shown in Fig. 1(a) offers a most basic real-time drone control by manipulating the virtual drone on the mobile device. PointFly , TrajectoryFly, and Waypoints, shown in Fig. 1 (b), (c), and (d) respectively, provide different predefined drone manipulations by allowing users to edit the movements of the flying drone on the screen.

2 SYSTEM DESIGN

2.1 Workflow and user interface

Fig. 2 shows the workflow of our prototype. The setup was designed for an indoor drone operation (e.g., wall inspection). This figure shows an example where a small drone is manipulated to check where the wall is damaged. To track the drone, this example uses an ad-hoc tracking system with QR code-based visual markers on the floor and a drone-mounted bottom-facing camera (whose FOV is shown in green). The camera streams the captured video to a server using 5.8-Ghz wireless transmission (red line). The server then calculates the drone's position and direction based on the predefined marker arrangement using the OpenCV Acuco algorithm [Garrido-Jurado et al. 2015; Romero-Ramirez et al. 2018]. Additionally, a WiFi connection (yellow WiFi icon) is established between the server and the drone to exchange motion commands (blue dotted line) and the FPV video stream (black dotted line). Another WiFi connection is created (blue WiFi icon) to send the FPV video stream from the server to the mobile device in real time. A user holds a mobile device running our see-through mobile AR software that employs augmented reality recognition technology to obtain the virtual drone's 3D pose matrix (homogeneous transformation matrix) using its camera and IMU sensor. The virtual drone's pose matrix is sent to the server by WiFi.

2.2 Coordinate system

As shown in Fig. 2, we used AR markers to define the drone's coordinate system. To make the coordinate systems of the drone and mobile device identical, we did a one-time calibration in which the user must capture these markers on the floor using the camera of the mobile device and define the origin with the same algorithm [Garrido-Jurado et al. 2015; Romero-Ramirez et al. 2018]. This process defines the shared coordinate system and origin. After this, instead of using the markers, the virtual drone pose matrix is updated using the IMU information provided by ARkit. Then we use the feedback control algorithm [Kiam Heong Ang et al. 2005] to make the flying drone approach the virtual drone on the mobile screen. Also, we can use GPS as the tracking system to meet outdoor operation requirements.

3 OUR INSIGHT AND USER EXPERIENCE

In our informal testing involving targeting and videography tasks, we found that Pinpoint Fly significantly improves indoor drone operation performance in terms of accuracy, efficiency and subjective satisfaction. Our system outperforms the traditional joystick controller especially when the drone operation path involves depth movement or multiple rotations.

At SIGGRAPH 2019 Emerging Technology exhibition we will demonstrate the four type of interaction techniques in Fig. 1. Visitors can use the interaction techniques to control the drone to complete a simple shooting game and take a pleasant video clip. We will also show an item delivery scenario using the drone. We will use a small toy drone with an AR-enabled mobile device. All equipment will operate within a safety net, ensuring that there will be no collision or other safety risk.

4 CONCLUSIONS

We proposed PinpointFly, a novel AR-based egocentric drone manipulation interface that increases spatial perception and manipulation accuracy by overlaying a cast shadow on the ground. We designed and implemented a proof-of-concept prototype using a motion tracking system, see-through AR techniques and a programmable drone.

REFERENCES

Yu-An Chen, Mike Y. Chen, Te-Yen Wu, Tim Chang, Jun You Liu, Yuan-Chang Hsieh, Leon Yulun Hsu, Ming-Wei Hsu, Paul Taele, and Neng-Hao Yu. 2018. ARPilot: Designing and Investigating AR Shooting Interfaces on Mobile Devices for Drone Videography. In *Proceedings of MobileHCI '18*. 42:1–42:8.

Sergio Garrido-Jurado, Rafael Muñoz-Salinas, Francisco Madrid-Cuevas, and Rafael Medina-Carnicer. 2015. Generation of fiducial marker dictionaries using Mixed Integer Linear Programming. *Pattern Recognition* 51 (10 2015).

Shunichi Kasahara, Ryuma Niiyama, Valentin Heun, and Hiroshi Ishii. 2013. exTouch: Spatially-aware Embodied Manipulation of Actuated Objects Mediated by Augmented Reality. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction (TEI '13)*. ACM, 223–228. https://doi.org/10.1145/2460625.2460661

Kiam Heong Ang, G. Chong, and Yun Li. 2005. PID control system analysis, design, and technology. *IEEE Transactions on Control Systems Technology* 13, 4 (July 2005), 559–576.

Francisco Romero-Ramirez, Rafael Muñoz-Salinas, and Rafael Medina-Carnicer. 2018. Speeded Up Detection of Squared Fiducial Markers. *Image and Vision Computing* 76 (06 2018). https://doi.org/10.1016/j.imavis.2018.05.004

K. Yonezawa and T. Ogawa. 2015. Flying robot manipulation system using a virtual plane. In *2015 IEEE Virtual Reality (VR)*. 313–314. https://doi.org/10.1109/VR.2015.7223421

Shumin Zhai. 2002. Human Performance in Six Degree of Freedom Input Control.