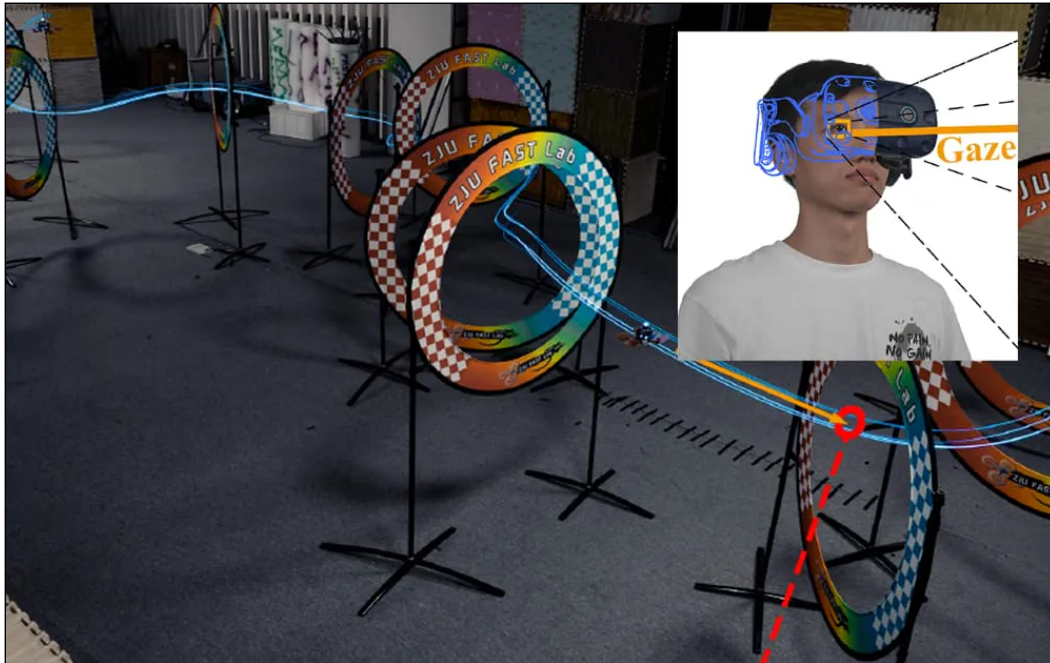


Controlling Drones Through Direct Vision

Originally published at <https://www.unite.ai/controlling-drones-through-direct-vision/>



Researchers from China have developed a new algorithm that can enable the control of drone flight by directly interpreting user vision. Effectively, the human operator 'becomes' the drone, and guides its trajectory based on the user's gaze direction.



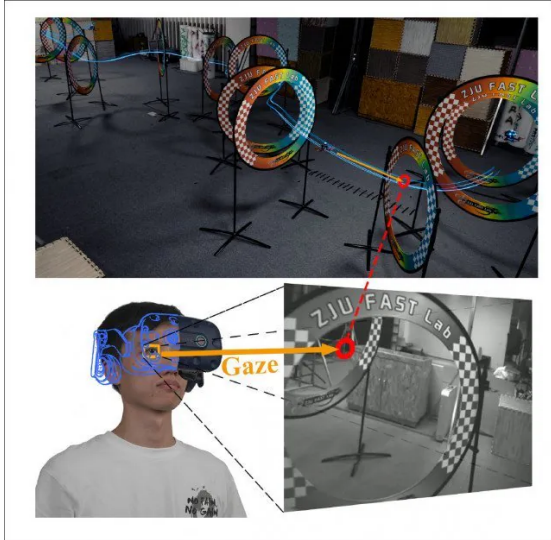
The user's POV is seen at bottom left, with the drone's flight path captured externally by a shadow device. See video at end of article for extended full-motion footage. Source: <https://www.youtube.com/watch?v=WYujLePQwB8>

The [paper](#) is called *GPA-Teleoperation: Gaze Enhanced Perception-aware Safe Assistive Aerial Teleoperation*, and comes from researchers at the Institute of Cyber-Systems and Control at Zhejiang University, and the School of Automation at the Nanjing Institute of Technology. The researchers have also released a video today demonstrating the capabilities of the system (see end of article).

Beyond Abstract Control

The researchers are seeking to remove the layer of abstraction for drone control, arguing that secondary control units require training and are only a rough abstraction of the user's intent, leading to unpredictable maneuvering and misinterpretation of guidance movements.

A paper earlier in the year, from the same researchers, emphasized the importance of line-of-sight visibility in drone navigation, and the current work is a development of the findings from [that research](#).



Above, a composite of the trajectory of the drone testing lab 'assault course' (see end video for exterior tests in a natural open air environment). Bottom, the operator wears an eye tracker which feeds through the direct view of the quadrotor drone's forward camera (bottom right). Source: <https://arxiv.org/pdf/2109.04907.pdf>

Algorithm

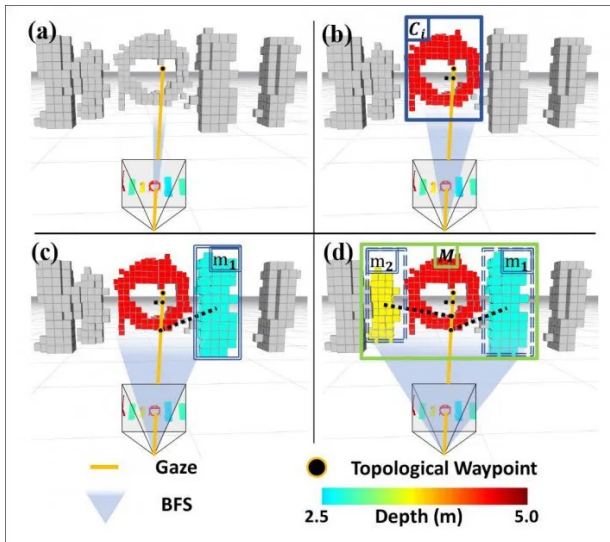
GPA uses a back-end optimizer which refines the user's gaze into the safest optimal path, arguably equivalent to 'auto aim' in video games, at practically zero latency (for obvious reasons).

The UAV subsystem modules are installed directly into the drone, including facilities for state estimation, planning, mapping and control modules. The local system receives eye-gaze data from a single-eye unit mounted in a head-worn harness by the end user, which supplies an initial topological path, which the system must sanitize on-the-fly.

To create a coherent experience for the remote controller, the monochrome view that the user receives is auto-centered by the on-board drone system, not least because without this it would be hard to interpret new intended deviations of route (as indicated by a change of gaze-direction).

The system first parses estimated vector coordinates from the image stream. Since the video input of the researchers' system is currently monocular, the camera's depth-of-perception boundary is used to obtain a second (depth) vector that's imposed onto the 2-D vector derived from the image. In theory, later iterations could use stereo cameras to improve this pipeline, though it remains to be seen whether the additional processing overhead would leave the benefit of hardware-based 3D perception intact.

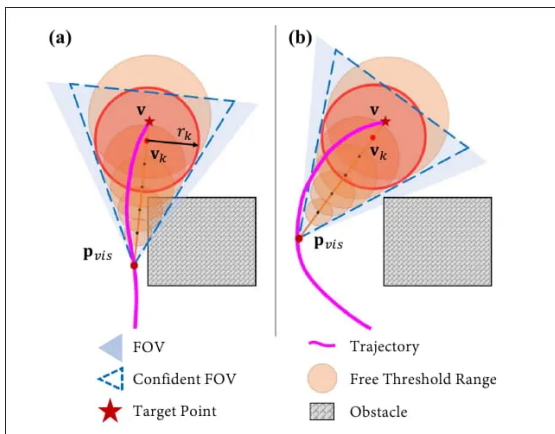
In any case, with the 3D values obtained, the calculation is used as an origin for a [Breadth First Search](#) (BFS). Pixels which would otherwise be obviated by BFS (i.e. pixels identified as already within bounds) are used as the anchor point for [DBSCAN clustering](#) (if not already clustered), and the routine reverts to BFS evaluation from the last break-point.



The architecture of GPA.

The process iterates until an object is identified and labeled within margin parameters correspondent to a field of view (FOV – which, in this case, must be absolutely clear in order to avoid collisions).

Finally, the vector calculations are used to generate clear paths, or else to validate that the user's gaze direction is already a safe path through or past an obstacle.

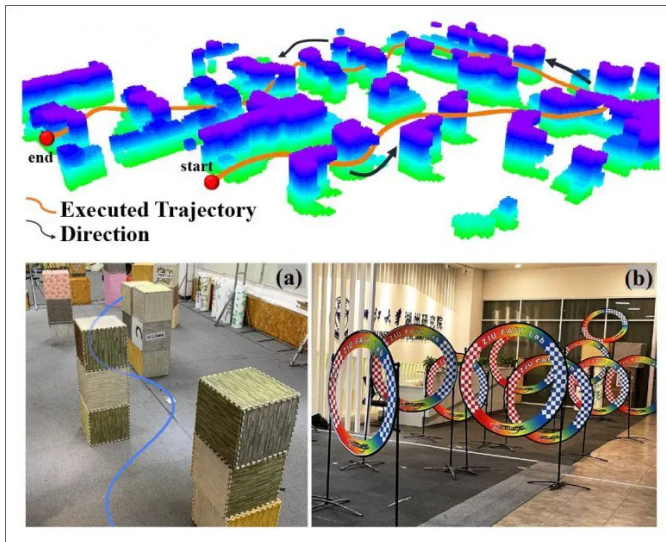


Tests where the system ignores visibility (left) and where the trajectory is recalculated to consider visibility as critical to a flight path (right).

Testing

To test the gaze-based drone control system, the Chinese researchers used a series of volunteers with zero knowledge of the system, and with no experience in controlling drone flight. The subjects were required to navigate obstacle courses in closed and exterior environments with only three brief orientation routines with which to familiarize themselves with the basic operation of the system.

Additionally, having briefed the volunteers as to the basic topology of the obstacles, the researchers added 'surprise obstacles' not included in the briefing.



Above, the trajectories of the online quadrotor drone, colored by height. Below, the navigable obstacles, beginning with boxes and moving on to rings.

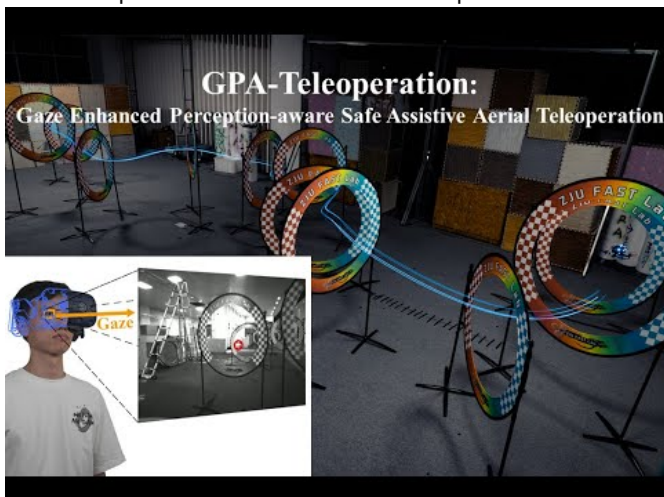
In practice, the system was able to effectively correct the gaze data so that the space-critical drones used were able to pass by (or through) the ring and box-shaped obstacles without collision, and the researchers have concluded that their system is both intuitive and safe, with a high margin of safety in operation.

The researchers also compared the performance of their approach to the FocusTrack architecture in the [Mavic Air 2](#) system, concluding that it outperforms the latter by being able to gauge and act on precise user intention.

Eye tracking technology has been extensively researched in fields such as autonomous vehicle data gathering for machine learning-based SDV systems, and in research into the attention patterns of pilots, among other sectors. In July of this year, a research team from Bulgaria [published](#) findings from observations of Unmanned Aerial Vehicle (UAV) pilots that established the landing stage of a flight as the most challenging for novice learners.

Check out the researchers' official video for GPA, below.

GPA-Teleoperation: Gaze Enhanced Perception-aware Safe Assistive Aerial Teleoperation



Writer on machine learning, domain specialist in human image synthesis. Former head of research content at Metaphysic.ai, until its dissolution into DNEG's Brahma.ai.

Portfolio site: martinanderson.ai

Contact: martin@martinanderson.ai

