Effective Remote Drone Control Using Augmented Virtuality

Kamil Sedlmajer[1], Daniel Bambušek[1] and Vítězslav Beran[1]
[1]*Brno University of Technology, Faculty of Information Technology, Centre of Excellence IT4Innovations, Bozetechova 1/2, Brno, 612 66, Czech Republic.*
kamilsedlmajer@gmail.com, {bambusekd, beranv}@fit.vutbr.cz

Keywords: Augmented Virtuality, UAV, Drone Piloting, Virtual Scene, Navigation Elements, First Person View, Third Person View

Abstract: Since the remote drones control is mentally very demanding, supporting the pilot with both, first person view (FPV) and third person view (TPV) of the drone may help the pilot with orientation capability during the mission. Therefore, we present a system that is based on augmented virtuality technology, where real data from the drone are integrated into the virtual 3D environment model (video-stream, 3D structures, location information). In our system, the pilot is mostly piloting the drone using FPV, but can whenever switch to TPV in order to freely look around the situation of poor orientation. The proposed system also enables efficient mission planning, where the pilot can define 3D areas with different potential security risks or set navigation waypoints, which will be used during the mission to navigate in defined zones and visualize the overall situation in the virtual scene augmented by online real data.

1 INTRODUCTION

Nowadays, the use of unmanned aerial vehicles (UAVs) extends to a wide range of areas, from rescue services and police forces to the commercial sector. Drones are used to monitor the quality of high voltage structures, the development of infrastructure outages, or to support complex interventions by rescue or police units. In all cases, the use of a drone requires high skill and mental demand for the drone operator.

Recent research let arise to various autonomous modes, where the drones are able to perform a precisely predefined mission independently and without the need for operator intervention. Unfortunately, today there is no problem in solving a wide range of tasks with the autonomous capability of the drone, but with the operator's legal constraints. For this reason, it is necessary to look for a different solution. This article deals with solving this situation by linking autonomous drone functions to operator control. The article seeks to reduce the operator's mental load in controlling drone in action using semi-autonomous drone functions.

Legal constraints today do not allow the full use of autonomous drone functions. Similarly, existing drone control solutions are now extremely burdensome for the operator. Orientation in space, keeping in safe zones, tracking key mission points, all of this makes the operation of the drone operator quite challenging.

Based on the study of existing tools and published results in the field, supported by own drone control experience, this work aims to define the key attributes of attention to the drone's operator and proposes a range of visualization and interactive features to reduce the drone operator's mental load, using augmented virtuality.

The key elements, that this study builds on, are the use of available topography maps, elevation maps, 3D data and virtual objects to enhance mission navigation clarity. The proposed solution combines existing data sources into a 3D virtual scene, augmented by online drone camera video-stream and other drone sensor data, as well as user-defined virtual objects such as safe zone boundaries or key points of a planned mission.

2 RELATED WORK

With an increasing popularity of drones, new effective methods for piloting them are emerging. Some are concentrating solely on autonomous flights with autonomous obstacle avoidance algorithms (Gageik et al., 2015) and (Devos et al., 2018), others are focusing on various user interfaces for manual drone controlling, such as gesture-based, voice-based, or remote controllers along with head-mounted displays (HMD). Recently, there were some experiments for controlling drones using brain-computer interfaces (Nourmohammadi et al., 2018) and (Mamani and Yanyachi, 2017). More common are attempts of direct drone control using gestures, which are detected with use of vision based methods from either drone attached cameras (Fernández et al., 2016), (Natarajan et al., 2018) or from additional hand tracking devices, such as the Leap Motion (Gubcsi and Zsedrovits, 2018). Unfortunately, such approaches are usable only for piloting the drone within a close distance rather than piloting it remotely. When piloting the drone remotely, but still from a visible distance, using some remote controller, it is often difficult to distinguish the drone's front face, which causes high workload on the operator. In order to lower this workload, Cho et al. introduced an egocentric drone control approach, which keeps the drone's back face automatically rotating towards the operator, who is piloting the drone from his/hers perspective (Cho et al., 2017).

To enable the user to control the drone completely remotely, without directly observing it, it is crucial to provide the user some sort of vision from the drone perspective. Currently, there are many commercially available products that are able to transmit the image from the drone attached camera into either the HMD, handheld display or regular screen in order to provide the first person view (FPV) for the operator. However, most of them only sends monocular video feed, which is insufficient in terms of perceiving distances, depths and proportions inside the FPV. The solution to overcome this problem has proved to be the attachment of two cameras to the drone to enable the stereoscopic vision inside the HMD (Smolyanskiy and Gonzalez-Franco, 2017).

Using remote controllers require attention and developed skills during piloting. Replacing them with wearable interfaces, such as exoskeleton suit with smart glove, could enable a more natural and intuitive drone control, where the operator feels more immersed (Rognon et al., 2018). Unfortunately, the operator could still struggle with awareness of drone surroundings that are out of drone camera's field of view (FoV). There comes in place the idea of placing the camera image into a virtual environment model, in order to extend the limited FoV (Calhoun et al., 2005). The image, along with the virtual model, can be augmented with waypoints, danger zones or points of interest. We believe that enabling the operator to whenever switch from the FPV into a third person view (TPV), where he/she can see whole drone situated in a virtual environment, should improve operator's situation awareness and reduce the mental load.

3 PROPOSED INTERFACE

The main goal of this work is to design a visualization system and navigation elements that will reduce the operator's mental strain when piloting the drone over longer distances. We define the following key attributes and how to resolve them:

1. Off-line and on-line spatial and sensory data fusion.
2. Virtual cameras.
3. Navigation or security structures.

Operator orientation might be improved by the TPV. The presented solution is based on virtual 3D scene augmented by real on-line data registered and visualized in the virtual scene (*Augmented Virtuality*). The 3D scene is created from the available free map data sources (topological, elevation). The 3D model of the drone inserted to the scene is controlled (position and orientation) by the data transmitted from the real drone. A video-stream from the camera on the drone is also rendered to the virtual scene. All combined spatial data must be properly registered into a local coordination frame. The pilot is then able to control the drone as usual (FPV), but can also *unlock* the camera and move to the TPV and see the drone and its surroundings. The operator's FoV is significantly expanded. This concept is depicted in the Figure 1.

The **rendering of the video-stream to the virtual scene** can be realized in several different ways. The first is to project it into a huge *virtual screen* (see Figure 1). The screen size depends on the FoV of the physical camera with either a plane or a suitably curved surface. The *screen* is moving at a fixed distance from the drone and follows the drones' position and orientation. The other option is to project the image directly onto objects in the scene. The usability of the video projection onto virtual scene requires very precise registration between the 3D virtual scene and the drone.

Virtual cameras in the 3D scene provides the user with ability to look into the scene from any viewpoint and to manipulate the camera freely. E.g. the operator can stop the drone in a safe position (safe distance from the obstacle) and move around the scene by *flying with the virtual camera*. When the virtual camera is fixed, the operator can switch back to con-

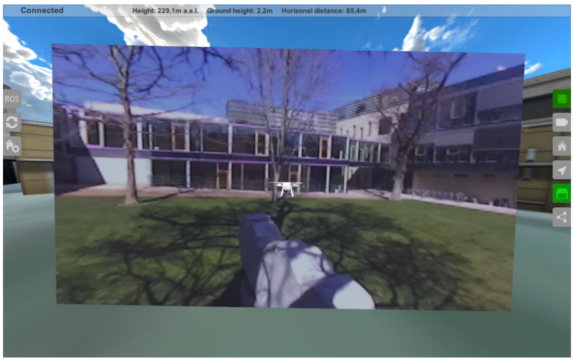

Figure 1: Concept of controlling the drone from the third person view. The virtual screen (with the video feed) moves at a certain distance in front of the drone and is synchronized with the physical gimbal configuration.

trol the drone and observe the scene from a new virtual position.

Another way to increase the safety of remote drone control is visualization of other sensor data, such as depth data, usually presented by 3D point cloud, that samples outer surfaces of objects around. This data significantly refine and complement the world model created from off-line data and allow the pilot to have a much better overview of the static obstacles around the drone, such as trees, buildings, cars, etc.

The use of augmented virtuality further provides the operator with the ability to add **navigation or security structures** to the scene. One such element may be **virtual walls**. These can define a space in the scene, such as unauthorized entry or a safe zone. Since such zones are often up to a certain height, displaying them in 3D scene is far more perspicuous than displaying them on a 2D map.

Another elements, e.g. for a flight in a more complex environment with a number of obstacles, are **navigation arrows** that point to waypoints. As a result, the pilot will always know which direction to fly and the task is just to avoid the obstacles. These waypoins can be placed not only on the ground, but also at a specific height in the air. The application should also be able to navigate back to the starting point and the pilot should be able to display both navigation arrows at once. One will point to the starting point, the other to the current waypoint. In addition to the arrows, the distance to a given point may also be displayed.

4 System Architecture

The testing application was developed using the Unity game engine. For creating the 3D virtual environment, we used the plugin MapBox[1]. The plugin automatically loads the environment maps and heightmaps in order to create the terrain in Unity. MapBox also includes several map layers with 3D building models, which are not perfect (rough building shapes, approximate heights, unrealistic textures or fake rooftop shapes), but they are still sufficient for our proof of concept. For a better quality virtual environment, virtual objects and textures from Google Maps, which are currently the best on the market, could be used. However, Google does not provide them to third parties yet.

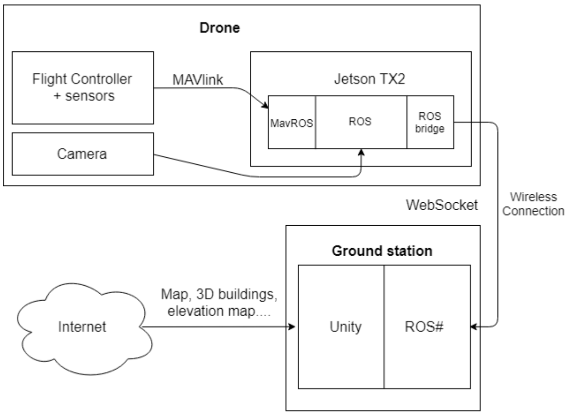

Figure 2: Communication between the drone and the ground station. On the application side, the ROS# library communicates with the drone via Rosbridge, where the data are transmitted via the WebSocket protocol. The application also automatically downloads maps, 3D buildings and heightmaps from the Internet.

For a successful combination of the virtual scene with the real drone, it is necessary to transmit following data from the drone into the ground station:

- Drone's position (pair of coordinates in WSG84 format).
- Drone's rotation (yaw – angle obtained from an electronic compass on the drone; pitch and roll).
- Compressed drone-attached camera stream.
- Drone-attached camera rotation (in respect to the drone).
- Other available sensor data (e.g. point cloud, battery status, flight speed, flight mode).

The method is based on 3D virtual scene model. The 3D virtual data are integrated from existing data sources and the coordination system is registered to actual real drone position. The GPS drone coordinates and orientation data is used for registration with virtual 3D frame. The augmentation of the virtual 3D scene is achieved by rendering the live video-stream

---

[1] https://docs.mapbox.com/unity/maps/overview/

from the drone front camera to projection plane in front of the virtual drone. The video latency is an important issue and highly influence the quality of the interaction with the proposed system. The video latency hardly depends on the quality of the WiFi signal, i.e. distance between the drone and the station. The additional navigation UI elements, like mission points, direction to next point or virtual walls, are rendered into virtual scene and presented to the user. The registration method might be improved by computer vision techniques, but this step will be considered later according to user tests, when the effect of rough GPS and compass based registration and caused video latency will be considered by professional pilots as an important issue.

The communication protocol is dependent on the drone manufacturer, drone's control unit and used software. In order to have customization possibilities, we built our custom experimental drone (see Figure 3), which comprises of the Pixhawk control unit with PX4 Autopilot software, Nvidia Jetson TX2 with the Ubuntu OS, stereoscopic camera, GPS and a compass. We chose the communication over WiFi between the drone and the base station. WiFi communication is limited in transmit distance, but has a high data throughput. The PX4 Autopilot uses the MAVlink communication protocol, which is transferred into the MavRos protocol of the ROS operating system (running on the Nvidia Jetson) over the Rosbridge tool. The communication scheme is depicted in the Figure 2. The system architecture is described in more detail in (Sedlmajer, 2019).

5 RESULTS

A test application was created with the implementation of several basic elements described in the previous chapter; and several real drone tests were performed. Since it was only an experimental platform with a gimbal-free camera and connection to a ground station via WiFi, which was limited in scope, it was only possible to perform the testing with limited restrictions. In spite of this, it was possible to try out a few basic use cases where the tests showed that the concept works and it is possible to control the drone with it. The application was built and tested on a laptop.

5.1 Test 1: Monitoring the area where the drone must not fly

When using a drone, for example, in the service of the police, it is often necessary to monitor an area (road, demonstration area, etc.). However, the police must also comply with the legal constraints, therefore their drones must not be flown, for example, near a highway or over a square full of people. If a pilot wants to use the drone to track the area, but to keep it out of its protective zone, he must constantly check where the drone is.

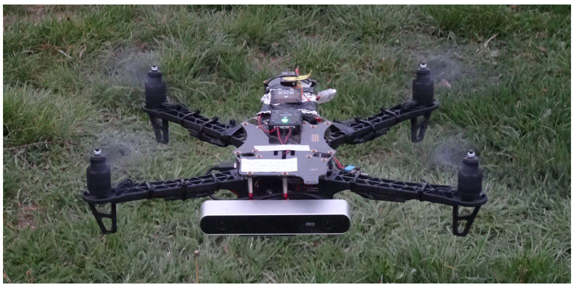

Figure 3: Test drone with Pixhawk control unit, ZED stereoscopic camera and Nvidia Jetson supercomputer, shot just above ground.

The first designed test flight was supposed to check whether adding virtual walls, representing borders of such areas, would help the pilot stay on their edge while observing what is happening around.

Virtual transparent walls were very useful in this test when flying inside the area, because the pilot sees clearly that he is approaching the border, or that he has just flown through it (see Figure 4). But the flight at the walls' border proved surprisingly difficult. When flying near these walls (about 5-8 m) it is really hard to estimate their distance. Therefore, it would be useful to hide them and display only the nearest part of the border when the drone really approaches it. The second option could be to gradually make the walls more transparent if the drone moves away from them. Most clearly, the part of the border that is closest to the drone would shine, the distant parts would be completely hidden. So if the pilot saw a glowing grille in front of him, he would know with absolute certainty that the border is very close. The distance, at which the boundary would appear, would be good to adjust to the speed at which the drone is approaching it, so that the pilot always has enough time to react, but at the same time the boundary does not unnecessarily interfere with the pilot's vision.

Another problem was the blending of the virtual wall with a virtual screen that was distracting and unpleasant. But even the proposed solution could at least partially solve this problem, because only the part of the boundary that is really needed at that moment would always be displayed.

5.2 Test 2: Exploration of a distant object and flight between obstacles

This test mimicked another fairly common task – exploring a more distant object (such as a house or a parked car) that is too far away from the drone, causing the pilot to fly closer to it. Here, the aim was to test whether the application really could help to improve the pilot's spatial orientation during the flight to this location, the object being explored, and the return to the starting point. In this test, it was assumed that the pilot knows the position of the object in advance and can create a waypoint on the object's position and navigate to it. The second part of the test was a low altitude flight between obstacles, when it is not possible to use an autonomous flight, but the pilot must manually get through a lot of obstacles (e.g. trees) and not to lose the spatial orientation and direction of the flight even with no landmarks around. It is not possible to use map data in this mode, because it is necessary to fly using a video only in order to watch obstacles carefully. Here, however, navigation arrows could help, because they keep the pilot informed about the direction to the waypoint and back to the starting point. This allows the pilot to accomplish the task faster.

When testing a remote object survey, an object was placed on the ground at a distance of approximately $60m$ from the starting point, and a waypoint was created at that point. Between the starting point and this point was an asphalt cycle path and several rows of freshly planted young trees that created natural obstacles. Since there were only meadows and low trees nearby, there were hardly any natural landmarks. Orientation only by the poor quality video stream was very demanding. However, during this test, the navigation arrows, which performed perfectly in their role, had greatly facilitated spatial orientation and significantly helped to reduce cognitive stress during driving. It was not necessary to search for natural landmarks, it was enough to observe nearby obstacles, a navigation arrow and a gradually decreasing distance to the destination. After a while, a target point indicator appeared on the ground.

But at this stage, there was a problem, because the target waypoint was constantly traveling the ground a few meters in all directions from a relatively small target. This was probably due to the inaccuracy of GPS navigation. In an effort to circumvent the examined object and explore it from all sides, the constant movement of this point and the confused rotation of the arrow was very annoying and confusing. Surprisingly, this subject was complicated.

It was much easier to do the fly around and survey with the navigation arrow turned off, by video only. At this point, it would probably be better if the navigation elements were automatically hidden after arriving close to the target and appeared again only if the drone moved away from the target again.

Figure 4: Screen of the implemented application based on augmented virtuality. The virtual environment model is augmented with the data from the real drone – position and orientation, which are used to render the virtual drone in the scene; camera image that is aligned with the virtual environment model (marked with the red circles); and other sensor data. This environment model can be enriched by virtual walls (the green grille) that can mark restricted areas or with a waypoints and direction arrows.

On the other hand, the second arrow pointing to the starting point could be used for orientation, as a some sort of compass that makes clear how the drone is being turned. In addition, for a pilot, a pointer to the starting point is somewhat more natural than an ordinary compass pointing to the north.

5.3 Discussion

Overall, the implemented application was relatively pleasant to use and the fact that it was able to partially compensate the absence of gimbal was positively appreciated. Indeed, by moving the video according to the current tilt of the drone, the objects on the video were still displayed at approximately the same location in the scene. Of course, the convenience of using the application was reduced by controlling the camera's rotation with the arrow keys of the laptop keyboard. This convenience could be increased by connection of VR glasses with head-tracker, which would allow the pilot to naturally look around the scene.

On the other hand, the application did not work very well at very low altitudes (up to $3m$), where the flight altitude and the distance from smaller objects were very poorly estimated. However, this problem also occurs in the first person view (FPV). Surprisingly, adding a drone model to the scene did not reduce the problem, but slightly increased it.

Quite surprisingly, the virtual screen with the camera image was relatively well-connected to the virtual scene most of the time, which even slightly exceeded expectations, especially given that the test drone certainly did not have the most accurate sensors available. Professional drone data is likely to be even more accurate.

The further development will be primarily focused on solving problems that were discovered during testing and on other designed ideas that were not implemented (e.g. the visualization of other sensor data, the point cloud, and the completion of area boundary visualization). Then, VR glasses with a head-tracker, which allows natural looking around the scene, will be connected to the application. Another such thing is to implement a free camera and test its capabilities.

6 CONCLUSIONS

The aim of this work was to improve pilot's orientation and to reduce his mental load during the drone remote control. Based on research and experience, a system has been designed that is based on augmented virtuality, where on-line data from drone sensors (video-stream, flight data, etc.) are integrated into the virtual environment model. The 3D virtual model consists of the data from external data sources like topography maps, elevation maps and 3D building models. The model also includes the user-specified planned mission information like waypoints, safe zone boundaries or flight directions.

The system architecture is designed to be scalable to communicate with multiple drones simultaneously. This could be useful in situations where more pilots are simultaneously carrying out a mission and have to work together.

The preliminary user tests proved that the proposed concept and technical implementation of the entire system improves the operator's orientation and navigation skills and so reducing the mental load. More user tests are planned in future work. The professional pilots will test the system to refine the concept, to improve or include more UI elements and for further development based on their needs.

ACKNOWLEDGEMENTS

The work was supported by Czech Ministry of Education, Youth and Sports from the National Programme of Sustainability (NPU II) project "IT4Innovations excellence in science – LQ1602" and by Ministry of the Interior of the Czech Republic project VRASSEO (VI20172020068, Tools and methods for video and image processing to improve effectivity of rescue and security services operations).

REFERENCES

Calhoun, G., H. Draper, M., F Abernathy, M., Delgado, F., and Patzek, M. (2005). Synthetic vision system for improving unmanned aerial vehicle operator situation awareness. *Proceedings of SPIE - The International Society for Optical Engineering*, 5802.

Cho, K., Cho, M., and Jeon, J. (2017). Fly a drone safely: Evaluation of an embodied egocentric drone controller interface. *Interacting with Computers*, 29(3):345–354.

Devos, A., Ebeid, E., and Manoonpong, P. (2018). Development of autonomous drones for adaptive obstacle avoidance in real world environments. In *2018 21st Euromicro Conference on Digital System Design (DSD)*, pages 707–710.

Fernández, R. A. S., Sanchez-Lopez, J. L., Sampedro, C., Bavle, H., Molina, M., and Campoy, P. (2016). Natural user interfaces for human-drone multi-modal interaction. In *2016 International Conference on Unmanned Aircraft Systems (ICUAS)*, pages 1013–1022.

Gageik, N., Benz, P., and Montenegro, S. (2015). Obstacle detection and collision avoidance for a uav with complementary low-cost sensors. *IEEE Access*, 3:599–609.

Gubcsi, G. and Zsedrovits, T. (2018). Ergonomic quadcopter control using the leap motion controller. In *2018 IEEE International Conference on Sensing, Communication and Networking (SECON Workshops)*, pages 1–5.

Mamani, M. A. and Yanyachi, P. R. (2017). Design of computer brain interface for flight control of unmanned air vehicle using cerebral signals through headset electroencephalograph. In *2017 IEEE International Conference on Aerospace and Signals (INCAS)*, pages 1–4.

Natarajan, K., Nguyen, T. D., and Mete, M. (2018). Hand gesture controlled drones: An open source library. In *2018 1st International Conference on Data Intelligence and Security (ICDIS)*, pages 168–175.

Nourmohammadi, A., Jafari, M., and Zander, T. O. (2018). A survey on unmanned aerial vehicle remote control using brain–computer interface. *IEEE Transactions on Human-Machine Systems*, 48(4):337–348.

Rognon, C., Mintchev, S., Dell'Agnola, F., Cherpillod, A., Atienza, D., and Floreano, D. (2018). Flyjacket: An upper body soft exoskeleton for immersive drone control. *IEEE Robotics and Automation Letters*, 3(3):2362–2369.

Sedlmajer, K. (2019). User interface for drone control using augmented virtuality. Master's thesis, Brno University of Technology, Faculty of Information Technology.

Smolyanskiy, N. and Gonzalez-Franco, M. (2017). Stereoscopic first person view system for drone navigation. *Frontiers in Robotics and AI*, 4.