Telescopied aerial manipulator and its avatar. Part 1: Communication, system’s interconnection, control, and virtual world

Rodolfo Verdín, Germán Ramírez, Carlos Rivera, and Gerardo Flores

Abstract—The tasks that an aerial manipulator can perform are incredibly diverse. However, nowadays the technology is not completely developed to achieve complex tasks autonomously. That’s why we propose a human-in-the-loop system that can control a semi-autonomous aerial manipulator to accomplish these kinds of tasks. Furthermore, motivated by the growing trend of virtual reality systems, together with teleoperation, we develop a system composed of: an aerial manipulator model programmed in PX4 and modeled in Gazebo, a virtual reality immersion with an interactive controller, and the interconnection between the aforementioned systems via the Internet. This research is the first part of a broader project. In this part, we present experiments in the software in the loop simulation. The code of this work is liberated on our GitHub page. Also, a video shows the conducted experiments.

Supplementary Material

The implementation of our system is released on GitHub and is available under the following link: https://github.com/Rodolfo9706/VR-teleoperated-aerial-manipulator.git

In addition, this letter has a supplementary video material available at https://youtu.be/Ur4sNFR9U-Y, provided by the authors.

I. INTRODUCTION

Some of the most important problems for aerial manipulators are the control, and the complexity in achieving a variety of grasping and manipulation tasks [1]–[3]. For that, several efforts have been conducted for robot teleoperation , [4] [5]. One of the major problems of teleoperated systems is the difficulties encountered by the human pilot while he/she is performing a task due to its visualization restriction [6]. This is especially true for moving robots [7]. In this work, we aim to solve the problem of interaction between the human operator and the scene seeing by an aerial manipulator robot by using a robot’s avatar and a virtual reality world which is a copy of the real one.

A. Problem

Performing a given task with a teleoperated aerial manipulator seeing the scene through mounted-cameras can be a complicated task for the human operator due to the limitation of vision on the scene. In this approach, the operator can see only what the mounted-camera can see. Using virtual reality is an alternative since you can have a virtual model in a ground control that replicates the movements and orientations of the robot [8]. Thus, the operator is able to visualize and control the robot by means of a reconstructed virtual world, knowing at any moment all the scenes of the environment.

B. Contribution

This paper presents the development and construction of the following system: an aerial manipulator and its avatar which is semi-autonomously commanded via the Internet. The avatar means a system that transfers the robot’s states, actions, presence, and environment to a virtual location in real-time. In this part of the research the contribution is as follows: a) the control algorithm for the robot; b) the software development for the interconnection via the Internet of all the parts of the complete system; c) test in the SITL simulation environment using the PX4 autopilot firmware; d) a demonstration of the experiments in a virtual world created in Unity; e) teleoperated control with a virtual reality headset. The aforementioned establish the first part of a bigger project. The part II will include the experiments with the real aerial manipulator depicted in Fig. 1 and 2) the SLAM algorithm for reconstructing the environment via SLAM, that we have developed previously. This scene reconstruction will be performed in real-time building the
virtual world, which is a copy of the real world seeing by the aerial manipulator.

C. Content

The remainder of this paper is as follows. Section II presents the related work of our contribution. In Section III the problem statement is given with some remarks and stating the scope of the article. In addition, we describe the main problems tackled by our approach. Then, in Section IV we develop the modeling and control approach we used for the aerial robot. Section V describes the complete system architecture of our approach including the communication, interconnection, and virtual world. In Section VI the experiments that corroborate our approach are presented and explained. Finally, in Section VII the future work and the conclusions are discussed.

II. Related work

With regard of control, there are several recent efforts in controlling aerial manipulators [9], [10], [11], [12]. Some of them considering complex coupled dynamics. We propose a geometric control programmed in the PX4 autopilot and tested in software in the loop simulations (SITL). Such control has demonstrated to be robust enough under the forces and torques exerted by the arm.

With the aim to create a dynamic and intuitive interface between the human and the robot for object manipulation, several applications have been developed in virtual environments. Certain applications include flying UAV systems [13] and simulations where the 3D world is reconstructed and gives dimensional feedback while the user is teleoperating [14]. In reference [15] a monitoring system for old buildings is created through virtual reality through a 3D reconstruction to detect possible structural damage using a UAV. On the other hand in [16] vision-based navigation algorithms for UAVs are developed to monitor dependent people through a virtual environment.

Tasks involving manipulation are presented in [17] for maintenance or repairing industrial robots in a VR environment, using an HTC vive device, and [18] for complex manipulating tasks using Virtual Reality (VR) sets. For high-risk tasks in [19] intuitive and effective control methods based on virtual reality implemented in ROS packages are proposed to teleoperate an underwater robotic arm to manipulate potentially dangerous objects. Regarding aerial manipulation systems win with virtual reality in reference [20] a 3D virtual simulator is implemented for collaborative tasks of autonomous and teleoperated navigation in a virtual room. The simulator allows the user to manipulate an object using two robotic arms attached to an aerial vehicle. Haptic and Virtual Reality-Based Shared Control for MAV is presented in [21] including an interface that allows safe operation in addition to provide a stable interaction and control between the aerial manipulator and the remote environment. A combined feedback system for an aerial manipulator is presented in [8] using VR trackers set in the user’s arm and tracking gloves. This framework provides vibrating feedback to control the robotic arm and a Head-Mounted Display to visualize the object. In [6] the authors propose a new interaction paradigm that provides adaptive views for improving drone teleoperation. This adaptive view provides a user with environment-adaptive viewpoints that are automatically configured to improve safety and provide smooth operation. However, in [6] the authors only focus their efforts on 3D reconstruction and virtual navigation with the human, not considering an aerial manipulator neither an avatar of it. Unlike these works, our contribution offers a solution for teleoperating systems in which an avatar recreates the movement of the real robot while a human operator controls the real robot. The interconnection of the robot, the human interface, and the avatar is via the Internet. In addition, this work was built using open source code free for the community.

III. Problem setting

The problem statement is as follows:

Problem statement 1: The problem can be devised in two layers. The first layer is that of proposing a way to perform complex tasks in a human-assisted aerial manipulator. The second layer is finding a solution for an enhanced visualization of the environment seeing by the robot and transmitted to the human-operator. In addition, the aerial manipulator must be stable during all the tasks performed by the human operator.

For solving this, we propose developing an avatar of a real aerial manipulator endowed with sensors and cameras. Such an avatar receives all the aerial robot’s states, together with the information captured by the embedded cameras. With this, a virtual environment is constructed in real-time identical to that seeing by the aerial manipulator.

The goal of the avatar and constructing a virtual world being a copy of the real world is twofold: 1) for repetitive tasks, one obtains a copy of the real scenario in which several human pilots can be trained to perform desired tasks, and at the same time, recovering all the data [robot states and environment] for eventually training an artificial-intelligence-based control; and 2) with the reconstructed virtual world, the robot can easily navigate and perform tasks that in the case of not having this virtual world could be impossible. If one thinks of drones with mounted cameras commanded by First Person View with a human pilot, one notices that these systems could be an option for our approach. However, those systems cannot reconstruct a copy of the navigated environment and as a consequence they lack of a virtual world. Using FPV systems can cause loose of orientation in the pilot, or even damage in the robot. We are trying to avoid such possible catastrophes with our proposal.

Motivated by the growing trend of virtual reality systems together with teleoperation, we develop a system composed of:

1. The aerial manipulator. In this part of the research we construct a Gazebo model connected and controlled to PX4 firmware.
2. Control algorithm to stabilize the UAV during all the scenarios. This is a geometric control programmed in the PX4 firmware.

3. A virtual world constructed in Unity that simulates the copy of the real world represented in Gazebo. The virtual world includes a virtual copy of the aerial manipulator.

4. Communication system. This is responsible to send all the control commands from the HTC vive to the aerial manipulator via Internet. Then, this system also sends such signals to the avatar, which copies the behavior of the aerial manipulator presented in Gazebo.

IV. MODELING AND CONTROL

A. Modeling

We model the aerial manipulator considering the following assumptions:

Assumption 1: The movement of the manipulator produces unknown moments in the quadrotor frame.

Assumption 2: The quadrotor serves as a moving platform for the position of the arm. Thus, the arm position is partially controlled by the quadrotor’s pose, in this way the forces exerted by the quadrotor frame to the arm are always controlled.

Assumption 3: The arm is totally controlled by the operator, only a simple PD control is implemented in each of the arm joints.

Thus, the aerial manipulator is modeled as follows

\[
\begin{align*}
\dot{x} &= v \\
\dot{v} &= ge_3 - \frac{f}{m}Re_3 + F_a(t) \\
\dot{R} &= R\hat{\Omega} \\
\dot{\Omega} &= -J^{-1}\Omega \times J\Omega + J^{-1}\tau + T(t)
\end{align*}
\]  

(1)

(2)

(3)

(4)

where \(\hat{\cdot} : \mathbb{R}^3 \rightarrow \mathfrak{so}(3)\) is \(\hat{x} = \begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix}\) with \(x = [x_1, x_2, x_3]^T\) in which \(\mathfrak{so}(3)\) is a Lie algebra in \(SO(3)\). Such operation has an inverse given by \(\vee : \mathfrak{so}(3) \rightarrow \mathbb{R}^3\) which basically transforms a skew-symmetric matrix into a vector in \(\mathbb{R}^3\). \((F_a, T_a)\) are the unknown forces and moment vectors exerted by the arm to the quadrotor frame. The rotation matrix \(R\) is given by [22]

\[
R = \begin{pmatrix}
c_\phi c_\psi & s_\phi s_\psi - c_\phi s_\psi & c_\phi s_\psi + s_\phi c_\psi \\
c_\phi s_\psi & s_\phi s_\psi + c_\phi c_\psi & -c_\phi s_\psi + s_\phi c_\psi \\
-s_\phi & c_\phi s_\psi & c_\phi c_\psi
\end{pmatrix}
\]  

(5)

B. Control

Let define

\[
e_p = x - x_d, \ e_v = v - \dot{v}_d
\]  

(6)

then we implement the control

\[
f = m \|\dot{e}_3 + K_v e_v + K_p e_p - \ddot{x}_d\|
\]  

(7)

The attitude control is given by

\[
\tau = -k_R e_R - K_\Omega e_\Omega
\]  

(8)

with

\[
e_R = \frac{1}{2} (R^\top_i R - R^\top_i R_d) \in \mathbb{R}^3, \ e_\Omega = \Omega - R^\top R_d \Omega_d \in \mathbb{R}^3.
\]  

V. SYSTEM ARCHITECTURE

The first element that needs to be added in both virtual environments (Gazebo, Unity3D) is the aerial manipulator 3D CAD model. Gazebo works on Ubuntu 18.04 and Unity works on Windows 10 OS and both of the environments communicate through common MAVROS messages which are included inside their respective topics. The topics to be employed are LocalPosition and MountControl, inside the LocalPosition topic the element “pose” collects the local quadrotor positions through GPS and this information is sent through a string type message. The MountControl topic is used to publish and subscribe to the robotic arm’s orientation or any other actuator that can be included in the vehicle model. Attitude information is published from Gazebo and the PX4 works as a subscriber to display that information in the Gazebo simulation. The visual system structure is described in Fig. 3 The system architecture is divided in three subsections: A) Unity; B) Gazebo & PX4; and C) communication (mavros).

A. Unity

The purpose of designing this environment in Unity is to create a special remote teleoperation ground station. It allows the use of virtual reality devices to visualize the environment
and get a robot’s position feedback in the real world. The aerial manipulator model in Unity contains several apps. These apps work individually and interact with each other to create and send the robot’s states depending on the input and the force exerted over the aerial manipulator model. The objects inside Unity are divided by hierarchy keeping the reference over the same objects of the model to place them or create the necessary force to move them in a local or global form as it is required. Therefore, the main reference of the model is placed on the UAV body followed by each link of the robotic arm and each revolute joint representing every degree of freedom that moves individually. The main reference of the model contains one app to get the position of the vehicle and the other app is to communicate Unity and Gazebo. Depending on the position input the first code generates the necessary force to be applied on each rotor to get the desired attitude \((\phi, \theta, \psi)\) leading to the desired position obtained from the position data in the Gazebo simulation. This virtual environment in Unity3D is a recreation of the real world where the aerial manipulator is moving involving the information concerning building dimensions and obstacles that might need to be avoided in the real world. In Fig. 4 it is shown the virtual reality world and its two-screen views. These views aim to facilitate the tasks performed by the robot through the operator and represent the avatar’s view.

B. The Gazebo model

Gazebo is an open-source 3D robotics simulator that provides elements such as sensors, actuators control, cameras, simulation tools, and realistic dynamics of each model. The purpose to make this virtual model is to work on a SITL simulation (software in the loop) that allows testing and simulating the aerial manipulator before being tested in the real world (which is the subject of the part 2 of this project). The Gazebo model of the aerial manipulator is programmed by us with the PX4 firmware, which is one of the most used autopilots by the robotics community. PX4 firmware contains packages that integrate gazebo to perform SITL that facilitates the design of the vehicle and the control implementation. To communicate with some external software/hardware, a communication protocol capable of sending and receiving messages between different topics is needed. Therefore, in this case MAVROS (mavlink / ROS) was the best choice to create Python or C++ scripts to extract or send information through the PX4 - Gazebo network we have created. In parallel, the ground station Qgroundcontrol is employed to work as a command center, where technical information about the aerial vehicle is displayed.

We construct the aerial manipulator in Gazebo based on an existing model called Typhoon \((\text{typhoon}_480\text{h})\) in the PX4 firmware. The Typhoon model is a Hexa-rotor with an embedded gimbal. The model \(\text{typhoon}_480\text{h}\) is in the PX4 firmware directory \(\text{src/tools/sitl_gazebo/models/typhoon}_480\) where a document meets SDF format. This contains all the aerial vehicle’s features such as color, collisions, visual, sensors, actuators, etc. In the same Firmware, a folder called Meshes contains the frame, the body link, and impellers in SITL format. To transform into a different vehicle, it is necessary to add the Collada files (*.dae) to the meshes folder. For the dimensions in the propellers and the manipulator, the collisions were modified. The created model is depicted in Fig. 5.

All the necessary commands and prompts to run and install the environment can be found in our GitHub page referenced at the beginning of the document.

C. MAVROS & communication

MAVROS is a ROS node that allows communication through mavlink protocols containing several topics, each of them contains specific information about sensors and actuators from the aerial robot. As was previously mentioned, the required topics for the aerial manipulator are the \(\text{LocalPosition}\) and \(\text{MountControl}\) for the position output (subscriber), and the robotic arm position inputs (publisher), respectively.

To get data from \(\text{LocalPosition}\) a python script was created using rospy. It allows linking python with ROS. Then, the geometry message was defined to establish communication with the aerial robot and be published in the \(\text{servo}\) message. To allow Unity to get the message, it is used a ROSBridge protocol, which is a WebSocket. Then, it subscribes to \(\text{servo}\) message to get the positions messages via the Internet. To get and read the message from Unity, a script is created to communicate with Ubuntu using an IP address, and then receiving the data subscribing to the \(\text{servo}\) message. To
do this process, JSON-formatted codes were used. Once Unity gets the positions, it moves the vehicle in the virtual environment to get to the same position from the message.

To get the positions from the robotic arm, a similar process is made. The topic used in this case is MountControl. Another python script is created to publish 3 variables for each joint of the robotic arm. Gazebo reads those 3 variables and moves the robotic arm to the indicated position given by the human operator. Then, Unity subscribes to a message called data, created inside the same script, to get the robotic arm information to move the manipulator in the virtual environment. A complete diagram of the communication structure is shown in Fig. 6.

VI. EXPERIMENTAL RESULTS

This section presents experimental results of the approach described in the last sections. The experiments consist in teleoperating an aerial manipulator trough the HTC vive VR headset via the Internet connection. Furthermore, we show the avatar of the aerial robot mimics the dynamic behavior of the aerial manipulator. Also, we test the control law given by (7) and (8).

The computer used for Gazebo simulation is an Intel Core i7-7820HK laptop with 32 Gb RAM and a GPU GeForceGTX 1070. For the virtual reality environment, we used an AMD A12-97209 laptop with 12 Gb RAM at 2.70 GHz, with no GPU.

A. Control algorithm

The objective of the first experiment is to measure the error between the setpoint position and the actual position of the vehicle, under the action of controls (7) and (8). We test the control response under the following teleoperated movements: left, right, up, down and several rotations. While we move the manipulator in several directions. The results are as depicted in Fig. 7, where robot’s pose and velocities are given. Angular positions and angular velocities are shown in Fig. 8. Notice how the attitude is
Fig. 8: Aerial manipulator’s attitude and angular velocities during software in the loop simulation conducted in Gazebo under control (8). The control law is programmed in PX4 firmware.

C. Pick and place experiment

In this experiment, the aerial robot is teleoperated through the HTC vive. The task to perform is that of taking an object and transporting it to a given point chosen by the operator. This is performed in the Gazebo environment while running the control algorithm. The avatar copies the task and the environment. The results demonstrate the control performance and robustness to a mass variation and under forces and moments generated by the arm and exerted to the drone. The mass of the object is 160 grams.

A video of the experiments is available in the following link [https://youtu.be/Ur4sNFR90-Y](https://youtu.be/Ur4sNFR90-Y)

Fig. 10: Pick and place experiment

VII. CONCLUSIONS

A teleoperated aerial manipulator was simulated in a virtual environment in Unity 3D and Gazebo. The simulation contains the vehicle dynamics and the kinematics of the manipulator, as well as the control programmed in the PX4 firmware in Gazebo. Also, it is teleoperated remotely by commands transmitted via ROSBridge protocol (WebSockets). This allowed the VR application to visualize in real-time the states of the aerial manipulator. The time response can be improved using high-performance computer equipment.

A. Future work

For the second part of this research, the experiments will be conducted in the real aerial manipulator developed at the lab and depicted in Fig. 11 In that work, a SLAM system will be implemented to reconstruct the virtual environment with real dimensional and imaging data. Also, a more intuitive sensorial virtual reality system to control the aerial manipulator will be included.

ACKNOWLEDGMENTS

The authors greatly appreciate the comments and time taken by the editor and the anonymous reviewers in evaluating this paper.
Fig. 11: Aerial manipulator prototype constructed in the LAB. This robot will be used in the second part of this research.

REFERENCES

[1] Watkins S, Burry J, Mohamed A, Marino M, Prudden S, Fisher A, Kloet N, Jakobi T and Clothier R, “Ten questions concerning the use of drones in urban environments,” Journal Pre-proof, 2020.

[2] S. Suzuki, “Recent researches on innovative drone technologies in robotics field,” Advanced Robotics, 2018.

[3] Hossein Bonyan Khamseh, Farrokh Janabi-Sharifi, Abdelkader Abbassameud, “Aerial manipulation - a literature survey,” EL SEVIER, 2018.

[4] D. Zhang, G. Yang, and R. P. Khurshid, “Haptic teleoperation of uavs through control barrier functions,” IEEE Transactions on Haptics, vol. 13, no. 1, pp. 109–115, 2020.

[5] M. Wang and H. Voos, “An integrated teleoperation assistance system for collision avoidance of high-speed uavs in complex environments,” in 2020 17th International Conference on Ubiquitous Robots (UR), pp. 290–296, 2020.

[6] J. Thomason, P. Ratsamee, J. Orlosky, K. Kiyokawa, T. Mashita, Y. Uranishi, and H. Takeamura, “A comparison of adaptive view techniques for exploratory 3d drone teleoperation,” ACM Trans. Interact. Intell. Syst., vol. 9, Mar. 2019.

[7] Ladislav Jurišica, František Duchoň, Martin Dekan, Andrej Babinec, Peter Pásztó, “General concepts of teleoperated systems,” Institute of Robotics and Cybernetics Slovak University of Technology, 2018.

[8] G. A. Yashin, D. Trinitatova, R. T. Agishev, R. Ibrahimov, and D. Tsetserukou, “Aerovr: Virtual reality-based teleoperation with tactile feedback for aerial manipulation,” in 2019 19th International Conference on Advanced Robotics (ICAR), pp. 767–772, 2019.

[9] Ran Jiao, Wusheng Chou, Rui Ding and Mingjie Dong, “Adaptive robust control of quadrotor with a 2-degree-of-freedom robotic arm,” Advances in Mechanical Engineering, 2018.

[10] J. Alvarez-Munoz, Nicolas Marchand, J.F. Guerrero-Castellanos, J. Tellez-Guzman, J. Escareño, M. Rakotondraha, “Rotorcraft with a 3dof rigid manipulator: Quaternion-based modeling and real-time control tolerant to multi-body couplings,” International Journal of Automation and Computing, 2018.

[11] R. Naldi, A. Macchelli, N. Mimmo, L. Marconi, “Robust control of an aerial manipulator interacting with the environment,” International Federation of Automatic Control, 2018.

[12] J. Lee, R. Balachandran, Y. S. Sarkisov, M. De Stefano, A. Coelho, K. Shinde, M. J. Kim, R. Triebel, and K. Kondak, “Visual-inertial telepresence for aerial manipulation,” in 2020 IEEE International Conference on Robotics and Automation (ICRA), pp. 1222–1229, 2020.

[13] Yuchen Wang and Wenhui Zhang, “Four-rotor uav virtual simulation system for agricultural sowing,” Information Technology and Mechatronics Engineering Conference, 2018.