This paper has been presented at the Robotics: Science and Systems (RSS) Workshop in Close Proximity Human-Robot Collaboration, New York, 2022.

# Aerial Manipulation System for Safe Human-Robot Handover in Power Line Maintenance

F.J. Gañán, A. Suarez, R. Tapia, J.R. Martínez-de Dios and A. Ollero
GRVC Robotics Lab., Escuela Técnica Superior de Ingeniería, Universidad de Sevilla, Seville, Spain.
Email: fjganan14@gmail.com, {raultapia, asuarezfm, jdedios, aollero}@us.es

***Abstract*—Human workers conducting inspection and maintenance (I&M) operations on high altitude infrastructures like power lines or industrial facilities face significant difficulties to get tools or devices once they are deployed on this kind of workspaces. In this sense, aerial manipulation robots can be employed to deliver quickly objects to the operator, considering long reach configurations to improve safety and the feeling of comfort for the operator during the handover. This paper presents a dual arm aerial manipulation robot in cable suspended configuration intended to conduct fast and safe aerial delivery, considering a human-centered approach relying on an on-board perception system in which the aerial robot accommodates its pose to the worker. Preliminary experimental results in an indoor testbed validate the proposed system design.**

## I. Introduction

Many industries and civil infrastructures require a wide variety of complex inspection and maintenance (I&M) tasks that are performed periodically by workers operating at height. Many of them require physical contact such as the installation of devices like bird flight diverters, insulators, or spacers on power lines [1]. These activities require extensive man power in extremely dangerous conditions. The need to automatize these tasks has motivated intense R&D efforts in aerial robotic manipulators such as the EU-funded H2020 AERIAL-CORE project (https://aerial-core.eu/), which aims to develop aerial cognitive robotic systems with unprecedented operational ranges and safety for I&M of large linear infrastructures.

Aerial manipulation robots [2, 3] have demonstrated their ability to conduct diverse operations involving physical interactions on flight with objects or the environment, proposing a wide variety of robotic manipulators integrated in multirotor platforms developed according to different design criteria, like the payload or force capacity, degree of dexterity, size, or weight. The use of aerial manipulators is well motivated in applications involving contact-based inspection typically conducted in bridges, viaducts, oil and gas refineries, or chemical plants, as well as in other extensive infrastructures like the power grid. However, the close interaction of aerial manipulators with humans has been, until recently, considered hazardous and avoided due to the possible injuries that the propellers of these platforms could cause on people [4]. To improve safety and reduce the risk of collision and the downwash effect due to the proximity of the multirotor propellers, work [5] compares the standard and long reach configurations, considering an anthropomorphic dual arm system suspended from a long reach link attached at the multirotor base.

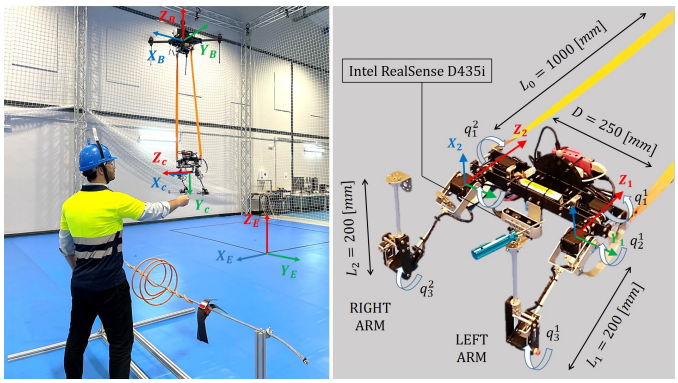


Fig. 1: Dual arm aerial manipulation robot in cable suspended configuration delivering a tool to a human worker operating on a power line testbed (left). Detailed view of the manipulator: links length, joint angles, and reference frames (right).

This paper presents a system for autonomous human-robot handover in which aerial manipulation robots deliver tools to workers at height, see Figure 1. The aerial robot is equipped with an anthropomorphic dual-arm manipulator in cable suspended configuration to enlarge workers separation with the propellers. The handover system includes onboard perception to ensure safe human-robot collaboration including techniques for the detection of workers' face, hands, and hand gestures. The system is validated in a representative indoor scenario.

The rest of the paper is organized as follows. Section II presents the anthropomorphic dual-arm manipulator. The system operation and main onboard perception methods are described in Section III. Preliminary experimental results are presented in Section IV, and the conclusions, in Section V.

## II. Anthropomorphic Dual-Arm Aerial Manipulator in Cable Suspended Configuration

The handover will be conducted by a dual arm aerial manipulation robot in cable-suspended configuration –a.k.a. long-reach configuration [5, 6]– consisting of a lightweight and compliant, anthropomorphic dual arm manipulator [7] (1.5 kg) attached to a multirotor robot through a pair of $L_0 = 1$ m length harness. This cable-suspended configuration [8, 9, 10] presents three main benefits w.r.t. the usual aerial manipulation configuration in which the arms are integrated at the base of the multirotor. First, it enhances safety by increasing the separation between the propellers and the worker, reducing collision risks [5] and downwash effects.

This paper has been presented at the Robotics: Science and Systems (RSS) Workshop in Close Proximity Human-Robot Collaboration, New York, 2022.

It allows the electrical insulation of the multirotor from the manipulator when the arms touch the operator on live power lines, avoiding electrostatic discharges [6]. Additionally, it facilitates the passive accommodation of the manipulator [7] to the pulling/pushing forces exerted by the worker during the handover, preventing instability.

The anthropomorphic dual arm kinematics [7] facilitates the simultaneous manipulation of two objects, or long objects that cannot be handled with a single arm. The human-like design ($L_1 = L_2 = 0.2$ m forearm-upper arm links length, $D = 0.25$ m arms separation), along with the low stiffness (5 Nm/rad) of the compliant joints [5], contributes to achieve a comfortable and intuitive interaction for the human worker. A pair of Pololu aluminium mounting hubs are used as flange at the tip of the forearm links, employing permanent magnets as end effectors for holding and retrieving tools from the worker like screwdrivers or pliers.

The arms integrate an Intel Realsense D435i RGB-D camera at the shoulder structure pointing forward such that the aerial robot can detect and localize the human worker during the approach stages, and maintain his/her hands within the field of view during the handover. The control programme of the arms, as well as the perception methods described in next section, are implemented on a Raspberry Pi 4 B computer board, feeding the system with a 1500 mAh 2S LiPo battery. The software architecture of the arms [5] is developed in C/C++, using UDP sockets for communication between on-board software modules and with the ground control station.

The kinematic model of the aerial robot is represented in Figure 1. Four reference frames are considered in the definition of the task: the Earth fixed frame $\{\mathbf{E}\}$ (inertial), the multirotor base frame $\{\mathbf{B}\}$, the camera frame $\{\mathbf{C}\}$, and the arms frames $\{\mathbf{i}\}$, with $i = \{1,\ 2\}$ for the left/right arms. The rotation angle of the $j$-th joint of the $i$-th arm is denoted as $q^i_j$, $j = \{1, 2, 3\}$.

Two assumptions are considered: 1) the multirotor velocity during approach is relatively low, so the dynamics of the suspended manipulator [10] does not introduce a significant disturbance on the attitude/position controller; 2) the arms are controlled in position/trajectory, relying on the mechanical joint compliance and passivity of the suspended configuration to support the physical interaction with the operator.

## III. System operation and Methods

The objective of the proposed system is to support the realization of I&M tasks by providing tools to workers operating in high altitude areas, reducing risks and increasing well-being. The handover system identifies the worker, moves around him/her to achieve a safety distance and suitable orientation, and performs the handover safely, using gesture detection to enable suitable collaboration and command the robot. Its operation is illustrated in Figure 2 and Figure 3.

In state *Init*, the robot is ready to take off. When it receives the start command with the worker's approximate location, the mission starts. The system operation is summarized as follows.

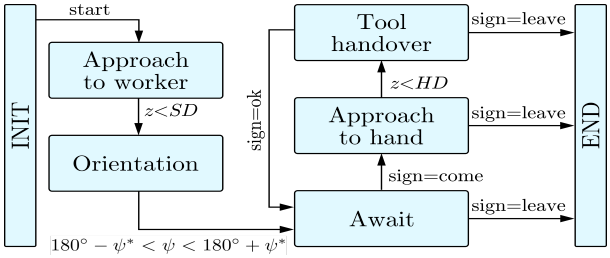


Fig. 2: General scheme of the autonomous handover system.

### A. Approach to worker

After take-off, the robot follows an obstacle-free trajectory that uses the worker's location as goal. The robot navigation system is not the objective of the paper, and for brevity is omitted. The robot navigates until its distance to the worker is lower than $SD$. To accomplish this task, YOLOv5n is used to detect the worker, and his/her 3D position is estimated by combining the RGB image and depth field from the camera. YOLOv5n was proven to be highly accurate to detect the worker and also computationally light to be executed in the Raspberry Pi at approximately 0.6 fps. This rate is considered suitable for the application where the person is assumed static.

### B. Orientation

To increase worker's safety and comfort, the robot changes the orientation w.r.t. the worker ensuring that he/she sees the robot approaching and receives the tool comfortably. The robot defines and performs a circular horizontal trajectory of radius $SD$ that keeps the robot oriented towards the worker. The circular motion is kept until the relative face-camera yaw angle is suitable for handover, $180° - \psi^* < \psi < 180° + \psi^*$. We choose $\psi^* = 30°$.

YOLO5Face [11] is used to detect the worker's face and estimate the relative face-camera yaw $\psi$ using the position of the worker nose and the face bounding box, both provided by YOLO5Face. First, we define the deviation $\delta$ of the nose w.r.t. the center of the bounding box as $\delta = 2x/L$, where $x$ is the horizontal coordinate of the nose on the image plane w.r.t. the center of the worker face bounding box, and $L$ is the width of the bounding box. The relative face-camera yaw $\psi$ can be estimated from $\delta$ as $\psi = -\frac{1}{b}\log\left(\frac{a}{\delta} - 1\right) + \psi_0$, where $a$, $b$ and $\psi_0$ can be obtained by fitting, as described in Section IV. The method is shown to be robust and has very high face detection rates over a wide range of relative face-camera pitch and yaw angles. Additionally, it is robust to partial occlusions e.g., due to the worker's helmet.

### C. Handover

This stage includes modules *Await*, *Approach to hand*, and *Tool handover*, in which the worker communicates with the robot using the 5 signs shown in Figure 4. In *Await*, the robot hovers at distance $SD$ facing the worker while waiting for commands. The detection of sign *come* triggers module *Approach to hand*, in which the robot moves towards the worker until the distance is lower than $HD$. During this phase, the hand is detected and its 3D position is estimated using

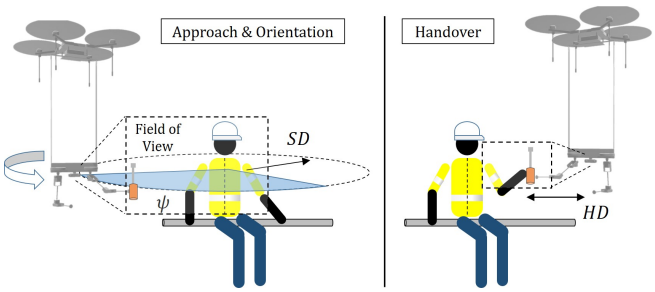


Fig. 3: Approaching, orientation and handover phases involved in the aerial delivery operation to a worker on a power line.

the RGB-D camera to conduct the handover. During *Tool handover*, the robot keeps hovering and if the worker's hand is recognized as sign *open*, the manipulator end effector is commanded towards the worker's hand solving the inverse kinematics. If more than one *open* hands are detected, the closest one to the camera is chosen to guide the robot. If no *open* hand is detected, the handover is paused. It will be resumed when at least one *open* hand is detected. When the worker confirms correct handover with sign *ok*, the robot returns to the waiting pose. If sign *leave* is detected, the robot cancels the mission. The on-board image processing is executed at 6 fps on a Rapsberry Pi 4, which were proven to be sufficient for this application.

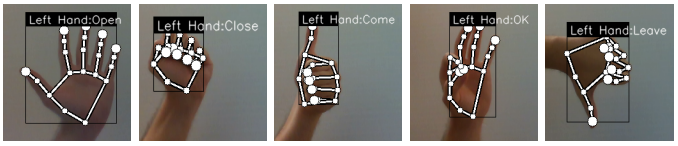


Fig. 4: Signs considered during *Handover*. From left to right: *open*, *close*, *come*, *ok*, and *leave*.

The onboard perception system uses MediaPipe Hands Python Framework [12] to detect the worker's hand and the hand skeleton. It adopts a single-shot detector (SSD) to recognize the hand and a Convolutional Neural Network (CNN) to extract the 21 landmarks of the hand skeleton. Using the depth field of the RGB-D camera, the landmarks' 3D positions are estimated. Next, gesture recognition is performed using the landmark positions w.r.t. the wrist normalized considering the greatest landmark-wrist distance. These are the input of a Feed-forward Neural Network, with 2 hidden layers and ReLU activation functions, trained to classify into the signs in Figure 4. Both hand detection & localization and gesture recognition have accurate and robust performance as shown in Section IV.

## IV. Experiments

The proposed system was experimentally validated at the GRVC Robotics Lab flight arena, equipped with a motion capture system with 24 *OptiTrack Prime$^x$13* cameras for millimeter-accuracy pose estimation. The multirotor platform was a 3 kg payload Proskytec LM (Light Multirotor) quadrotor equipped with a CUAV v5 autopilot running Arducopter 3.6. The quadrotor has attached the lightweight long-reach dual-arm manipulator described in Section II (Figure 1). Perception

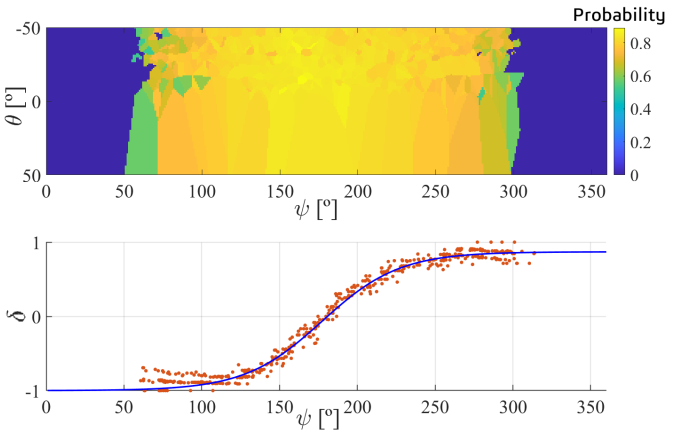


Fig. 5: Face recognition validation: Top) output score from the face detection network for different pitch and yaw face-camera relative angles. Bottom) $\delta$ versus the face-camera relative yaw angle $\psi$ (red) and sigmoid fitting (blue) to formally relate them and determining $\psi^*$.

algorithms were implemented in Python3 on a Rapsberry Pi 4 using RGB-D images from an Intel Realsense D435i .

First, preliminary experiments are conducted to validate the performance of the face detector. The camera was placed at a distance of $SD$ m from the worker, chosen to be 2.5 m due to the available space and safety reasons, with different relative orientations w.r.t. the worker's face while the camera described a circular trajectory. The worker's head was kept in the center of the image along the experiment. Figure 5-top shows the output score of the face detection neural network for a wide range of different face-camera relative pitch $\theta$ and yaw $\psi$ angles. The face is detected if the output score of the face detection network is higher than a threshold percentage, taken as 40%. The results show that the face is recognised with an output score $> 0.75$ for $150° < \psi < 210°$ and that $\theta$ has a negligible effect on the detection for $|\theta| < 50°$. This evidences the robustness of face detection against the workers moving its head during handover. The relationship between the relative face-camera yaw angle $\psi$ (measured with the *Optitrack*) and $\delta$ is presented in Fig. 5-bottom, showing that the relative face-camera yaw angle $\psi$ can be estimated from $\delta$.

The hand detector is also evaluated. First, for proper handover, the worker's hand should be detected in the distance range [0.15, $SD$] m. In the experiments, the worker's hand has a detection probability higher than 99.5% within this range of distances. Also, the workers hand's should be accurately localized in 3D for stage *Tool handover* within range [0, 1.5] m. This distance range is defined based on the possible hand-arms distances due to the worker movements, being $HD$ equal to 0.3 m since the maximum arm reach is 0.4 m. The 3D position estimation results (using the *Optitrack* as ground truth) in one experiment are presented in Figure 6. The mean and the standard deviation of the errors at different camera-hand distances are shown in Table I. Localization errors reduced to lower than 1.6 cm for distances within [0.2, 0.5] m, enabling accurate handover.

| | [0.2, 0.5] m | | | [0.5, 1.5] m | | | [1.5, 3] m | | |
|---|---|---|---|---|---|---|---|---|---|
| | $X$ | $Y$ | $Z$ | $X$ | $Y$ | $Z$ | $X$ | $Y$ | $Z$ |
| $\mu$ (cm) | 1.57 | 1.39 | 0.34 | 3.57 | 2.99 | 0.79 | 10.82 | 11.25 | 12.94 |
| $\sigma$ (cm) | 1.09 | 0.91 | 0.32 | 2.83 | 2.04 | 0.67 | 6.44 | 6.07 | 9.98 |

TABLE I: Mean error $\mu$ and standard deviation $\sigma$ of $X$, $Y$ and $Z$ coordinates of hand detection for different distances.

Finally, the complete handover system was evaluated during autonomous flight. The aerial robot with the dual-arm manipulator took off at a distance of 4 m from the worker, who was at the centre of the testbed. The manipulator held a screwdriver at its end effector. The handover was performed as expected (described in Section III), see video[1]. A set of 20 flights was performed to test repeatability. In stages *Approach to worker* and *Orientation*, the worker and his face were detected in $>$99% of the frames. In stages *Await* and *Tool handover*, the signs were successfully recognised in $>$95% of the images. The hand 3D position localization during the complete flight mission was proven to be very accurate (mean of $<$ 2 cm for distances $<$ 1.5 m) and enable successful execution of the handover mission. Although in some few images the person, hand, or gesture recognition fails, it successfully triggers the transitions between the stages, and the full handover operated as expected in all the conducted experiments.

## V. Conclusions

This paper presented a system for safe autonomous human-robot handover in applications where human workers are deployed in high altitude areas. The aerial robot used an anthropomorphic dual-arm manipulator in cable suspended configuration to enhance safety by increasing the separation distance between the propellers and the worker. The handover system ensures safe and efficient human-robot collaboration using onboard perception techniques to detect workers, ensure safety distances, ensure correct human-robot relative orientation, and enable human-robot communication. The system was implemented and executed onboard and validated in indoor experiments.

The extension of the system to a full aerial robot co-worker including perception methods to recognize tools and a wider range of worker's signs as well as the validation in outdoor experiments are object of current development.

## Acknowledgments

This work was supported by the EU project H2020 AERIAL-CORE (871479). Partial funding was obtained from the Spanish Project ROBMIND (PDC2021-121524-I00) and the Plan Estatal de Investigación Científica y Técnica y de Innovación of the Ministerio de Universidades del Gobierno de España (FPU19/04692).

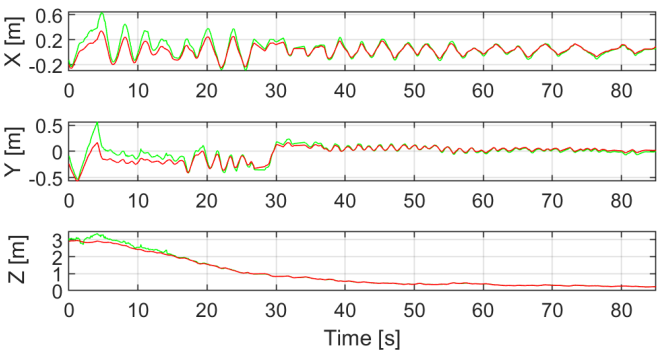


Fig. 6: Estimation of hand position (green) versus ground truth values using *OptiTrack* motion capture (red). Notice that Z-distance is decreasing over time since the camera is approaching to the worker.

## References

[1] M. A. Trujillo, J. R. Martínez-de Dios, C. Martín, A. Viguria, and A. Ollero. Novel aerial manipulator for accurate and robust industrial ndt contact inspection: A new tool for the oil and gas inspection industry. *Sensors*, 2019.

[2] A. Ollero, M. Tognon, A. Suarez, D. Lee, and A. Franchi. Past, present, and future of aerial robotic manipulators. *IEEE Transactions on Robotics*, 2022.

[3] X. Meng, Y. He, and J. Han. Survey on aerial manipulator: System, modeling, and control. *Robotica*, 38, 2020.

[4] P. E. I. Pounds and W. Deer. The safety rotor—an electromechanical rotor safety system for drones. *IEEE Robotics and Automation Letters*, 2018.

[5] A. Suarez, F. Real, V. M. Vega, G. Heredia, A. Rodriguez-Castaño, and A. Ollero. Compliant bimanual aerial manipulation: Standard and long reach configurations. *IEEE Access*, 2020.

[6] A. Suarez, R. Salmoral, P. J. Zarco-Periñan, and A. Ollero. Experimental evaluation of aerial manipulation robot in contact with 15 kv power line: Shielded and long reach configurations. *IEEE Access*, 2021.

[7] A. Suarez, P. J. Sanchez-Cuevas, G. Heredia, and A. Ollero. Aerial physical interaction in grabbing conditions with lightweight and compliant dual arms. *Applied Sciences*, 2020.

[8] Y. S. Sarkisov, M. J. Kim, D. Bicego, D. Tsetserukou, C. Ott, A. Franchi, and K. Kondak. Development of sam: cable-suspended aerial manipulator. In *IEEE Int. Conf. on Robotics and Automation*, 2019.

[9] R. Miyazaki, H. Paul, T. Kominami, and K. Shimonomura. Wire-suspended device control based on wireless communication with multirotor for long reach-aerial manipulation. *IEEE Access*, 2020.

[10] G. Li, A. Tunchez, and G. Loianno. Pcmpc: Perception-constrained model predictive control for quadrotors with suspended loads using a single camera and imu. In *IEEE Int. Conf. on Robotics and Automation*, 2021.

[11] D. Qi, W. Tan, Q. Yao, and J. Liu. Yolo5face: why reinventing a face detector. *arXiv:2105.12931*, 2021.

[12] F. Zhang, V. Bazarevsky, A. Vakunov, A. Tkachenka, G. Sung, C.L. Chang, and M. Grundmann. Mediapipe hands: On-device real-time hand tracking. *arXiv:2006.10214*, 2020.

[1] A video showing one of the flights to validate the complete experiment can be found at https://youtu.be/KZ8iNqub9QE.