Accurate Vision-based Flight with Fixed-Wing Drones

Valentin Wüest[1], Enrico Ajanic[1], Matthias Müller[2], and Dario Floreano[1]

***Abstract*— Fixed-wing drones must navigate to the desired location accurately for maneuvers such as picking up objects and perching. However, current GNSS receivers limit their navigation accuracy to several meters in outdoor environments, making such maneuvers impossible. RTK GNSS can improve flight accuracy, but it requires ground stations at the target location and additional communication modules on the drone. Here, we describe a fixed-wing platform with onboard computation that uses positional information from a GNSS receiver and vision from an onboard camera. The drone relies on a GNSS signal for flying towards a point of interest and switches to vision-based information to accurately reach the target. We conducted outdoor experiments to compare the flight accuracy of three navigation methods: GNSS, RTK GNSS, and the proposed GNSS-vision method. We also systematically assessed the robustness of vision-based control to compensate for GNSS errors and quantify the accuracy of the proposed method. Our results show that the accuracy of the proposed GNSS-vision system is on par with RTK GNSS. GNSS-vision reduces the average error of GNSS by over an order of magnitude, from 3.033 m to 0.283 m, and reduces the variance across repeated flights from 2.095 m to 0.309 m. We open-source the software-hardware architecture used in this paper to enable the research community to build on these results and expand the capabilities of fixed-wing drones.**

I. INTRODUCTION

Fixed-wing drones are gaining increasing attention for their larger flight endurance compared with rotary-wing drones of similar mass. However, perching on a precise spot [1], [2], [3], [4] or picking up an object in flight [5] requires accurate localization with respect to the target. In laboratory conditions, highly accurate localization is often accomplished with an optical motion capture system [6]. However, when flying in outdoor environments such systems are no longer available and accurate localization becomes challenging.

GNSS receivers are routinely used for long-range flight, but they result in a position error of approximately 2 m during flight [7]. Therefore, GNSS is well suited to approach a target region from far away, but is not sufficiently precise for the accurate flight leading to perching on a spot, picking up an object, or passing through a narrow gap. Real-time Kinematic Global Navigation Satellite Systems (RTK GNSS) or the beacon-based ultra-wideband positioning systems can support precise flight outdoors, but require the positioning of additional communication hardware on the ground. Furthermore, all control methods based on absolute localization systems assume that the target area is static.

This work was supported by Intel Network on Intelligent Systems, by the Swiss National Science Foundation through the National Centre of Competence in Research (NCCR), and by the European Union's Horizon 2020 research and innovation programme under Grant 871479 AERIAL-CORE.

Valentin Wüest, Enrico Ajanic, and Dario Floreano are with the Laboratory of Intelligent Systems, Ecole Polytechnique Federale de Lausanne (EPFL), CH1015 Lausanne, Switzerland (e-mail: valentin.wueest@epfl.ch; enrico.ajanic@epfl.ch; dario.floreano@epfl.ch)

Matthias Müller is with the Embodied AI Lab, Intel, 85579 Neubiberg, Germany (e-mail: matthias.mueller@intel.com).

Fig. 1. Fixed-wing drone accurately flying towards a yellow balloon using a combination of GNSS and vision.

Here, we describe a fixed-wing platform with embedded sensing and computation that combines GNSS information with visual information from an onboard camera for accurate relative localization with respect to a target. The control method relies on a GNSS signal during long-range flight and uses vision-based navigation for accurate flight towards the target spot.

Accurate vision-based flight has received some attention on rotary-wing drones [8], [9], [10], [11], [12]. Despite the increased range of fixed-wing over rotary-wing vehicles [6], research on accurate vision-based flight for fixed-wing vehicles has been limited [13], [14], [15]. This can be explained by the increased complexity of fixed-wing aerodynamics, the difficulty of creating a controlled environment outdoors, and the comparatively higher speed of fixed-wing drones to prevent stalling. In addition, there is a lack of readily available hardware and software platforms to study vision-based flight of winged drones, resulting in different custom setups. While a few works investigated accurate vision-based flight with fixed-wing drones [13], [14], [15], none of them report statistically representative accuracy values. Furthermore, reproducing those results is difficult because the hardware setup and software is not publicly available.

In this article, we address these gaps by (1) describing a fixed-wing research platform that allows the development and testing of vision-based flight for fixed-wing drones, (2) developing a controller that brings together GNSS and vision to achieve accurate and reliable flight to a target, and (3) characterizing the accuracy of the proposed method against GNSS and RTK GNSS.

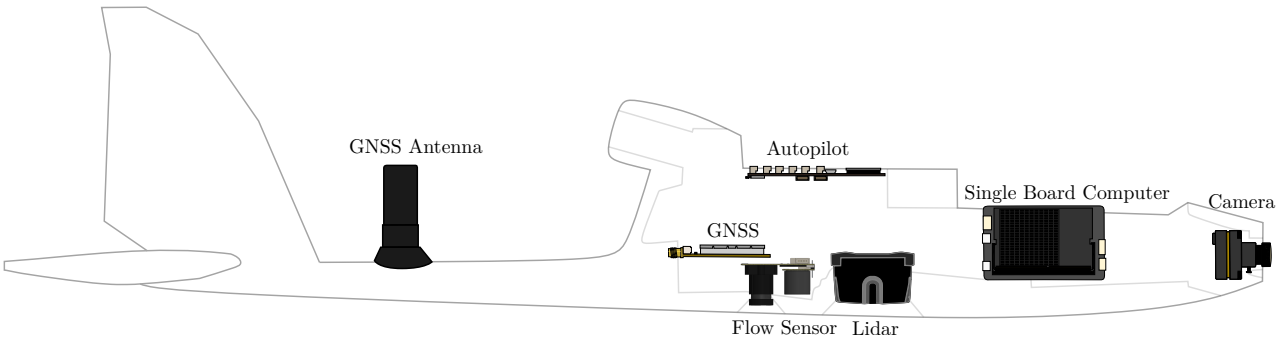


Fig. 2. Figure depicting the key components and their position on the 1.55 m wingspan fixed-wing drone.

A. Related Work

A method for achieving accurate flight towards a target consists of adding sensory equipment at the target location. For example, an optical motion capture system has been used to achieve perching with a flat plate vehicle on a powerline [1], [2]. Similarly, perching on an office chair with a dihedral-wing drone was successfully achieved [3], [4]. Ground cameras have been used outdoors to localize drones and land them on a runway [16], [17], [18], [19].

Onboard cameras have mostly found application on rotary-wing vehicles. Quadrotors relied on a front-facing camera to estimate the position of a window or racing gate to fly through it [9], [10], [20]. In the context of the Mohamed Bin Zayed International Robotics Challenge (MBZIRC), quadrotors achieved successful landing on a slowly moving vehicle by detecting a tag [12], [21], [22], [23]. In all of these applications, the vehicles either fly significantly slower than it is possible with most fixed-wing drones or require temporary hovering.

In fixed-wing vehicles, onboard cameras were mostly used to land on a runway where edge-detection algorithms were used to detect and align the drone with the runway [24], [25], [26] or horizon detection was employed to land in an open field [27]. None of these works quantify how accurately the drone landed on a target line or position. Another approach for accurate localization with onboard sensors on a fixed-wing drone consists of using a 2D LIDAR sensor [28]. However, this method requires prior mapping of the area and was only demonstrated for indoor flight. Scaling such an approach to large outdoor areas requires extensive resources.

Only a few papers report data on vision-based flight with a fixed-wing drone to a target position or object. A large red airbag was used as a landmark and soft landing spot to fly into; color filtering and shape based descriptors called Hu's moments were used to identify the red air bag [29]. Image-based visual servo control then guided the drone to the bag [14], [30]. The control algorithm, however, is platform dependent and is not guaranteed to be stable. Carrying out tests on our platform resulted in unstable behavior, which we describe in section V. Furthermore, the authors only state that the target bag had a radius of 2 m but did not specify an achieved accuracy. Similarly, a red inflatable triangle and moving red truck were landed onto by estimating their position in GNSS coordinates and flying towards it [13], [31]. A success rate of 12 out of 15 was reported for landing on the 2 m^2 truck platform. Finally, an autonomous landing system was demonstrated by detecting and flying into a red net [15]. The approach was reported to achieve a flight accuracy of 2.21 m which was based on GNSS measurements with an accuracy of 2.5 m. Furthermore, in these lastly mentioned works [13], [15], [31], the authors used detections of the net to guide the drone in a global frame, as a result, the method relies on GNSS data even during the final phase.

In summary, the accuracy of vision-based flight for fixed-wing vehicles reported in the literature is in the order of 2 m, although statistically representative accuracy values are missing. The lack of an easily available platform makes the results difficult to reproduce and leverage for additional research.

II. FIXED-WING RESEARCH PLATFORM

We start by describing a simple fixed-wing drone platform to carry out vision-based flight experiments. The platform is composed of a lightweight off-the-shelf drone with fixed wings and is equipped with sensors and processing power for real-time computation during autonomous flight. Here, we employ the open-source hardware and software autopilot PX4 [32], which is supported by an active community that provides regular bug fixes and updates. Along with this article, we make how-to instructions, a list of readily available hardware parts, 3d printed parts, and our software code publicly available at `github.com/lis-epfl/lis-vision-flight`. We hope that this will enable additional research on intelligent, vision-based flight of winged drones. In the following, we briefly describe the hardware and software stack in more detail.

A. Hardware Components

The drone body consists of an H-King Bixler 3 that is widely used in the RC-hobby drone community. With its 1.55 m wingspan and large fuselage cavity, it can carry a large number of components making the setup easily adaptable. The drone contains a 3S, 2400 mAh lithium polymer battery, four servo motors (Corona DS-919MG) to

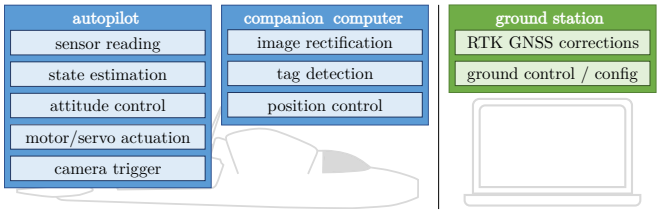


Fig. 3. Scheme of the software stack on the drone and ground station.

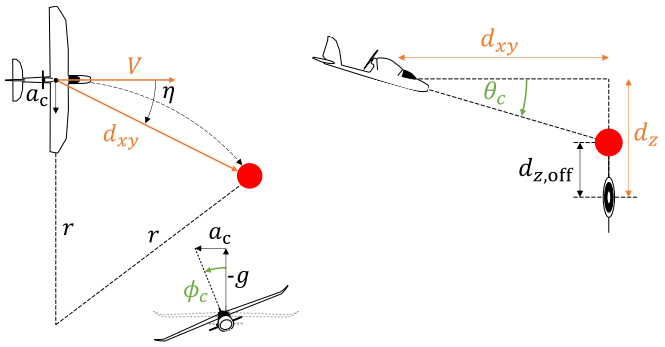
actuate ailerons, elevator, and rudder, and a brushless motor (Scorpion SII-2212-1400KV) to produce thrust.

For the experiments described in this paper, we equipped the drone (see Fig. 2) with a PX4 autopilot (Holybro Pixhawk4) which builds the core of the electronics setup. The drone carries a GNSS receiver with an antenna (Drotek F9P RTK GNSS & DA233) to receive GNSS. A telemetry radio (Holybro SiK Telemetry Radio V3 433 MHz) onboard communicates with the ground station to receive RTK corrections. For vision-based flight, an optic flow sensor (PX4FLOW) and a point lidar (Garmin LIDAR-Lite v3HP) are used to estimate the ground velocity of the drone without the use of GNSS. The drone is capable of measuring airspeed with the use of a pitot tube and a differential pressure sensor (Sensirion SDP3x). To detect the tag, a global shutter camera (Flir Firefly S) with a 56° field of view lens (Arducam 1/2.5" M12 Mount 6mm) captures images at a resolution of 1440x1080 pixels at 20 fps. These images are then processed on a capable single board computer (Nvidia Jetson Nano and a custom carrier board that we open-source) to detect the tag and log data from both the camera and the autopilot.

The ground station is made up of a laptop, the same telemetry module, and the same GNSS receiver with antenna.

B. Software Stack

The autopilot consists of the PX4 software stack on the Pixhawk 4 autopilot. We show an overview of the software in Fig. 3. The PX4 software stack filters and fuses sensor data from the point lidar, optic flow sensor, IMU, GNSS, and differential pressure sensor in an Extended Kalman Filter (EKF). Based on the measurements, the PX4 software runs the attitude control and commands the thrust motor, ailerons, elevator, and rudder servos. Through telemtry link, PX4 communicates with the ground station. The PX4 stack also communicates to the companion computer. The communication is achieved thorugh the communication layer mavros [33], which allows inspection of measurements and setting of commands on the PX4 autopilot. The companion computer is also responsible for image processing. It receives image from the camera through ROS, rectifies it and detects a Whycon tag in it [34]. The guidance control law receives the detection information and state estimation through mavros, calculates a roll and pitch command and sends it back through mavros to PX4. On the ground station, QGroundControl [35] is running to monitor the experiment and set parameters of the PX4 autopilot. Through the telemtry link, QGroundControl sends RTK corrections to the drone for high accuracy positioning.



(a) Lateral control (b) Longitudinal control

Fig. 4. Graphical representation of the control inputs in orange and control outputs in green. Variables are explained in the main text. This graphical representation is adapted from [36].

III. CONTROL AND ESTIMATION

Traditional GNSS-based fixed-wing position control operates in an inertial, world frame. However, onboard cameras measure relative position in a body-fixed frame. Therefore, using vision as a primary sensor requires a change in the design of the control scheme. Concurrently, we estimate the attitude of the body in the world using an Extended Kalman Filter (EKF) based on measurements from the Inertial Measurement Unit (IMU) [32]. Naively rotating camera detections into the world frame, however, does not lead to the desired accuracy. This can be accredited to unreliable yaw estimates. While roll and pitch are observable through accelerometer measurements, the yaw of a drone is commonly observed through the magnetometer, which often suffers from large inaccuracies. This in turn leads to large position estimation errors of the target location.

Within this work, we thus operate in a gravity-oriented frame, *i.e.* a frame whose z direction is aligned with the gravity vector and has the same yaw angle as the plane. This choice allows us to easily integrate camera detections into the control pipeline without the issues of yaw inaccuracies. Additionally, it also allows us to split the dynamics into lateral and longitudinal components.

1) Lateral control: The lateral control is based on the L_1 guidance [37]. This control law relies on the fact that given a constant roll angle ϕ, the fixed-wing will ideally follow a circular trajectory with a fixed radius r. In order to calculate the radius of the trajectory required to reach a target location, we need to know the flight velocity V, the relative horizontal position of the target d_{xy}, and the angle η between the two, as shown in Fig. 4(a). The radius can then be related to the centripetal acceleration with $a_c = V^2/r$ from which follows that the centripetal acceleration a_c and the desired roll angle ϕ_c are:

$$a_c = 2\frac{V^2}{L_1}\sin(\eta)$$

$$\phi_c = \tan^{-1}\left(\frac{a_c}{g}\right)$$

where g is the gravitational constant. In the original paper [37] L_1 is $L_1 = d_{xy}$. Note that both terms a_c and ϕ are independent of the mass of the drone, which makes this control law applicable to a wide range of fixed-wing vehicles.

To keep the target in the camera frame, it is desirable to align the line-of-sight (LOS) of the drone with the target. To achieve such a behavior, we define L_1 different from the original paper. We choose L_1 to be smaller by defining $L_1 = \min(L_{1,\max}, d_{xy})$, where $L_{1,\max}$ is a design parameter. The larger $L_{1,\max}$ is, the slower the lateral behavior gets when the horizontal distance to the target is larger than $L_{1,\max}$. As demonstrated in [38], both choosing a constant $L_1 = L_{1,\max}$ or time dependent $L_1 = d_{xy}$ have been shown to lead to asymptotically stable behavior.

2) Longitudinal control: The longitudinal control aims to fly accurately to a target while keeping it in the field of view. We detect a Whycon tag [34] placed at a known height difference $d_{z,\text{off}}$ below our target. Whenever the tag is detected in the image, we transform the detection from the body frame to the gravity-oriented frame. This allows us to obtain the horizontal distance d_{xy} and the vertical distance d_z to the tag. We obtain the commanded pitch angle θ_c through:

$$\theta_c = \tan^{-1}\left(\frac{d_z - d_{z,\text{off}}}{d_{xy}}\right) - \theta_{\text{off}},$$

where θ_{off} is the angle of attack at trim flight condition.

3) Estimation: In the experiments, we investigate the performance of *GNSS-based control* and *vision-based control*. When we refer to *GNSS-based control*, the drone is using GNSS data for guidance control. However, when we refer to *vision-based control*, the drone is not using any GNSS information for guidance control and thus operates in the gravity-oriented frame. The control system uses the tag detection to obtain the relative position and uses the output of the EKF, which is fusing measurements from the optic flow, the lidar, and the IMU, to obtain an estimate of the ground velocity.

IV. RESULTS

We first define the terms of accuracy and precision which will be used throughout the evaluation and then describe the experimental setup. We proceed by evaluating the vision-based controller for the case with very low initial offsets due to the GNSS. By systematically introducing expected GNSS offsets, we then analyze the effect of lateral and longitudinal offsets on the accuracy and study whether the proposed system is able to compensate for them. Finally, we analyze the accuracy in a scenario that is close to various applications and report on the accuracy and precision that can be expected.

A. Metrics: Accuracy and Precision

We define two metrics to report the findings of our experiments in a consistent and quantitative manner. In the context of target-directed flight, it is both important to have a measure of closeness to the target value and a measure of statistical variability. We thus define the metrics of *accuracy* and *precision*. A simplified representation of the two metrics is shown in Fig. 5.

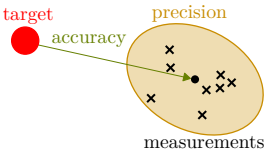


Fig. 5. Schematic graphical representation of accuracy and precision.

We first introduce the *mean error* as $\boldsymbol{\mu} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{r}_i$, where $\mathbf{r}_i \in \mathbb{R}^2$ is the vector from the target to the measurement. We then define the *accuracy* as its 2-norm $\|\boldsymbol{\mu}\|_2$ [39]. Similar to the standard deviation in one dimension, we define the variability measure *precision* as $\sqrt{\frac{1}{n}\sum_{i=1}^{n} \|\mathbf{r}_i - \boldsymbol{\mu}\|_2^2}$ [39].

B. Experimental Setup

The experiments consist of autonomously flying the drone towards a target position indicated by a balloon. The balloon is positioned on top of a 4.8 m long vertical pole to increase the distance to the ground and avoid any crashes during initial testing. Below the balloon, at a distance of 0.75 m, we attach a Whycon [34] tag, which will allow us to estimate the drone's relative position to the tag and subsequently to the target. The tag has a diameter of 0.24 m, enabling the first detection on average 35 m away from to target, which corresponds to a time to impact of around 3 s for our flights at 12 m/s.

Before the experiment begins, we manually fly the drone to line it up with the target in a distance of 80 m. The location of the target is given to the control system in GNSS coordinates and control is handed over to the onboard companion computer. The companion computer steers the drone towards the target by using GNSS-based control. When the drone reaches a distance of 60 m from the target, the experiment flight begins. Each experiment flight is split into three phases: GNSS phase, vision phase, and blind phase as shown in Fig. 6. In the GNSS phase, the drone continues flying towards the coordinate of the target with the GNSS-based control. During this phase, the target cannot yet be detected by the onboard camera. The vision phase is initiated as soon as the target is detected for the first time. In this phase, control is switched to the vision-based control and the drone does not rely on GNSS anymore. When the plane gets close to the target, the tag disappears from the field of view (blind phase) and the drone keeps flying using the last roll and pitch command received from the vision-based controller. In the following, we will refer to the strategy of relying on GNSS for the long range and on vision after the first detection of the tag as *vision-GNSS*.

The experiments were conducted on an open area with trees being 65 m away and low-rise buildings 95 m away while receiving data from at least 20 satellites. In all experiments, we measure the flight accuracy with the onboard RTK GNSS, which reports an accuracy of 1.0-5.0 cm both vertically and horizontally.

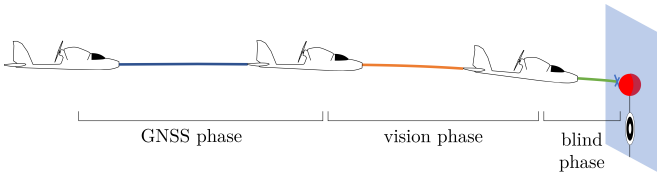
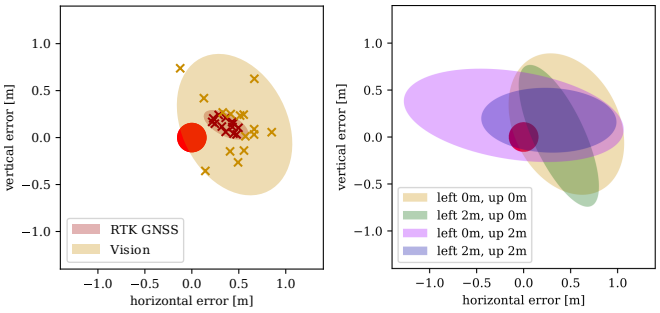


Fig. 6. Experiment phases, where the blue line indicates the GNSS-based approach phase, the orange line the vision range phase, and the green line the blind phase. The drone aims to hit the red sphere, which represents the yellow balloon of Fig. 1. The blue cross shows where the drone passes by the cross plane shown by the shaded blue area.



(a) Comparison of vision-based flight to RTK GNSS flight without offsets

(b) Influence of initial offsets

Fig. 7. Representation of the cross plane for the experiment of low GNSS inaccuracies and the influence of GNSS inaccuracies on the performance of the vision-based control. The crosses indicate the points where the drone passes by the target, shown as a red sphere. The shaded areas shows the 95% confidence ellipses.

C. Comparison of Vision-GNSS to RTK GNSS

Flight accuracy can be affected by multiple factors, such as control and vision estimation. To disentangle the causes of accuracy errors, we perform an experiment with two sets of flights. In the first set, we determine the upper bound for accuracy and precision with the proposed controller by using the RTK GNSS for all three phases of the flight. In the second set of flights, we only use GNSS for the first phase and switch to the vision-based controller as soon as the tag is detected for the first time (see Fig. 6).

The first set of flights consists of 13 approaches to the target. We visualize the positions of where the drone passes through the *cross plane* (blue shaded plane in Fig. 6) in Fig. 7(a). The target is located at (0,0) and is visualized by a red sphere with a radius of 0.15 m – the diameter of a standard sized latex balloon. The red crosses indicate the positions where the drone passes the cross plane. We refer to these positions as *cross points*. Note that these cross points are defined by only a horizontal and vertical distance to the target. We therefore report the error in these two directions. We visually represent the variance of the cross points by the 95% confidence ellipse, where the center of the ellipse corresponds to the mean cross point. During these 13 flights, we achieved an accuracy of 0.361 m, with a horizontal offset of 0.314 m during a side wind of 1.6 m/s to the right of the cross plane (as estimated by the onboard autopilot). The precision is as small as 0.070 m, indicating that the proposed control system can achieve highly accurate flights.

TABLE I

Summary of flight experiments with systematic initial offsets corresponding to expected GNSS errors during the approach phase. We report the mean cross point, side wind, achieved accuracy, and precision.

sensor	# of flights	initial offset (right, above) [m]	mean cross point (right, above) [m]	side wind [m/s]	accuracy [m]	precision [m]
RTK	13	(0, 0)	(0.314, 0.179)	1.6	0.361	0.070
vision	15	(0, 0)	(0.455, 0.138)	1.2	0.475	0.384
vision	5	(-2, 0)	(0.374, 0.013)	1.2	0.374	0.317
vision	5	(0, 2)	(-0.105, 0.231)	1.1	0.254	0.662
vision	5	(-2, 2)	(0.268, 0.179)	1.2	0.322	0.292

The second set of flights consists of 15 approaches to the target. The cross points are shown as yellow crosses in Fig. 7(a). The drone passed by the target with an accuracy of 0.475 m and a precision of 0.384 m with a horizontal offset caused by an estimated lateral wind of 1.2 m/s. In terms of accuracy, vision-GNSS is only about 30% worse than RTK GNSS. The vision-based controller is less precise than the RTK GNSS upper bound but on par in terms of accuracy. The results are summarized in the top part of Table I.

D. Robustness of vision-based control to inaccurate GNSS

In the previous experiment, we used RTK GNSS which is accurate but requires ground stations and is less practical as a result. While raw GNSS only requires a sensor on the plane, it is less accurate and can lead to large offsets between the estimated and actual state of the platform or target. In order to evaluate the robustness of the vision-based controller to this sensor noise, we emulate GNSS errors by introducing both horizontal and vertical *initial offsets* at the beginning of the vision phase.

GNSS errors reported in the literature vary largely. In a study that is closest to the flight conditions of these experiments, the authors report a median error of 2.0 m for a GNSS receiver on a bicycle [7]. In accordance with these data, the u-blox NEO-M8N shipped with the Pixhawk autopilot, reports an average accuracy of 2.2 m. We thus choose a horizontal and vertical offset of 2 m each. The horizontal offset was set to the left of the target only because the behavior for offset to the left and right should be identical. The vertical offset was 2 m above the target for two reasons: firstly, approaching a target from above applies to a large range of situations, such as pick-up, landing, and drop-off; secondly, since the Whycon tag is below the target, setting an offset below the target would artificially facilitate the task. Finally, we apply both the horizontal and vertical offsets at the same time. We approach the target five times for each offset (in meters): (-2, 0), (0, 2), and (-2, 2), where the first number depicts an offset to the right and the second number an offset above. The results are summarized at the bottom of Table I.

TABLE II

Comparison of vision-based flights to flights using GNSS only.

sensor	# of flights	mean cross point (right, top) [m]	side wind [m/s]	accuracy [m]	precision [m]
GNSS	17	(1.181, 2.793)	-1.0	3.033	2.095
vision	15	(-0.175, 0.222)	0.3	0.283	0.309

For all offsets, the vision system was able to reliably detect the target and the vision-based controller was able to correct for the offsets.

We show a visual representation of the flight distributions as confidence ellipses in Fig. 7(b). To increase legibility, we refrain from showing the cross points, but show the confidence ellipse from the vision-based flights in section IV-C for comparison. We report the accuracy, precision, the mean cross point, and the estimated side wind during each of the set of flight approaches in Table I. Taken together, these results indicate that the vision-based control was able to compensate for the 2 m offsets that a GNSS receiver could introduce, and report an accuracy below 0.4 m, which is comparable to the absence of any initial offset. The mean cross points in all sets of flights are horizontally less than 0.5 m and vertically less than 0.25 m from the target. Furthermore, the precision was not greatly affected by the offsets and was only slightly worse for the (0,2) offset. These findings suggest that expected GNSS inaccuracies can in fact be compensated for with the vision-based control without adversely affecting flight accuracy.

E. Comparison of GNSS and vision-GNSS

Finally, we compare flights using exclusively GNSS-based control with flights where vision-based control replaces GNSS-based control as soon as the target is detected. We summarize the comparison between them in Table II.

The GNSS-based control uses the u-blox NEO-M8N GNSS module without differential RTK corrections. In the first set of experiments, we use GNSS in all three phases of the flight experiment. We evaluate the accuracy over a set of 17 target flights and show the cross points in dark red in Fig. 8.

In these experiments, the drone crossed the target plane with an accuracy of 3.033 m and a precision of 2.095 m. As shown in Fig. 8, we observed large errors that confirm previous results. The mean cross point was shifted horizontally by 1.181 m (with an estimated side wind of -1.0 m/s) and vertically by 2.793 m. Such large offsets are prohibitive for tasks such as pick-up and perching.

In the second set of 15 flights, the drone flies with GNSS-based control and switches to vision-based control after the first visual detection; during the final blind phase, the latest control command is applied.

In these experiments, accuracy was improved by more than an order of magnitude from 3.033 m to 0.283 m and the precision was greatly improved as well from 2.095 m to 0.309 m. A similar trend is observable in the mean cross point, which was reduced by an order of magnitude to below 0.25 m horizontally and vertically. These experiments led to significantly more accurate vision-based flight with fixed-wing drones than any previous work in the literature (which reported errors in the order of 2 m). More importantly, the accuracy measured in the vision-GNSS experiment showed no significant loss of accuracy over RTK GNSS. The drastic improvement of vision-based control over GNSS and the comparable performance to RTK GNSS highlights the potential of using fixed-wing vehicles for long-range missions that require accurate flight without ground stations.

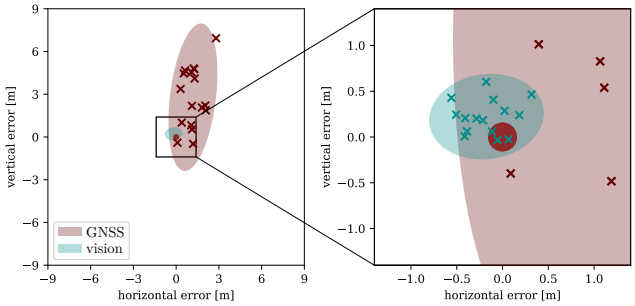


Fig. 8. Accuracy comparison of GNSS-based flight and vision-based flight.

V. CONCLUSION

This paper presented a method for accurate vision-based flight with fixed-wing drones. We also introduced a research platform based on widely available hardware components and an open-source code for further research and development by the robotics and machine learning community. The bill of material, an assembly guide, and the code are available at `github.com/lis-epfl/lis-vision-flight`. The research platform we present here allowed systematic and precise comparisons of navigation methods described in the literature for the first time. We showed that vision-based control could compensate for inaccuracies of commonly used GNSS and improve the achieved accuracy by one order of magnitude. It is thus comparable with RTK GNSS-based flights (albeit with lower precision), which however requires dedicated ground installations. We hope that the fixed-wing platform described here and the open-access code developed for these experiments will pave the way for new research in agile, multi-functional, and vision-based flight with fixed-wing drones.

ACKNOWLEDGMENT

We thank Olexandr Gudozhnik for his valuable contribution to the drone hardware, Davide Zambrano for the helpful discussions and support, and Fabian Schilling for the feedback and help with the experiments.

REFERENCES

[1] J. Moore and R. Tedrake, "Powerline perching with a fixed-wing UAV," in *AIAA Infotech at Aerospace Conference and Exhibit and AIAA Unmanned...Unlimited Conference*, 2009.

[2] J. Moore, R. Cory, and R. Tedrake, "Robust post-stall perching with a simple fixed-wing glider using LQR-Trees," *Bioinspiration and Biomimetics*, vol. 9, no. 2, p. 025013, May 2014.

[3] A. Paranjape, J. Kim, N. Gandhi, and S.-J. Chung, "Experimental Demonstration of Perching by an Articulated Wing MAV," in *AIAA Guidance, Navigation, and Control Conference*. Portland, Oregon: American Institute of Aeronautics and Astronautics, Aug. 2011.

[4] A. A. Paranjape, S.-J. Chung, and J. Kim, "Novel Dihedral-Based Control of Flapping-Wing Aircraft With Application to Perching," *IEEE Transactions on Robotics*, vol. 29, no. 5, pp. 1071–1084, Oct. 2013.

[5] W. Stewart, E. Ajanic, M. Müller, and D. Floreano, "How to Swoop and Grasp Like a Bird With a Passive Claw for a High-Speed Grasping," *IEEE/ASME Transactions on Mechatronics*, pp. 1–9, 2022.

[6] D. Floreano and R. J. Wood, "Science, technology and the future of small autonomous drones," *Nature*, vol. 521, no. 7553, pp. 460–466, May 2015.

[7] J. Schipperijn, J. Kerr, S. Duncan, T. Madsen, C. Klinker, and J. Troelsen, "Dynamic Accuracy of GPS Receivers for Use in Health Research: A Novel Method to Assess GPS Accuracy in Real-World Settings," *Frontiers in Public Health*, vol. 2, 2014.

[8] J. Thomas, G. Loianno, K. Sreenath, and V. Kumar, "Toward image based visual servoing for aerial grasping and perching," in *Proceedings - IEEE International Conference on Robotics and Automation*, 2014, pp. 2113–2118, iSSN: 10504729.

[9] D. Falanga, E. Mueggler, M. Faessler, and D. Scaramuzza, "Aggressive quadrotor flight through narrow gaps with onboard sensing and computing using active vision," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, May 2017, pp. 5774–5781.

[10] T. Bera, A. Sinha, A. K. Sadhu, and R. Dasgupta, "Vision based Autonomous QuadCopter Navigation through Narrow Gaps using Visual Servoing and Monocular SLAM," in *2019 Sixth Indian Control Conference (ICC)*, Dec. 2019, pp. 25–30.

[11] J. Thomas, G. Loianno, K. Daniilidis, and V. Kumar, "Visual Servoing of Quadrotors for Perching by Hanging From Cylindrical Objects," *IEEE Robotics and Automation Letters*, vol. 1, no. 1, pp. 57–64, Jan. 2016.

[12] D. Falanga, A. Zanchettin, A. Simovic, J. Delmerico, and D. Scaramuzza, "Vision-based autonomous quadrotor landing on a moving platform," in *2017 IEEE International Symposium on Safety, Security and Rescue Robotics (SSRR)*, Oct. 2017, pp. 200–207, iSSN: 2475-8426.

[13] B. Barber, T. McLain, and B. Edwards, "Vision-Based Landing of Fixed-Wing Miniature Air Vehicles," *Journal of Aerospace Computing, Information, and Communication*, vol. 6, no. 3, pp. 207–226, Mar. 2009.

[14] S. Huh and D. H. Shim, "A vision-based automatic landing method for fixed-wing UAVs," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 57, no. 1-4, pp. 217–231, 2010.

[15] H. J. Kim, M. Kim, H. Lim, C. Park, S. Yoon, D. Lee, H. Choi, G. Oh, J. Park, and Y. Kim, "Fully Autonomous Vision-Based Net-Recovery Landing System for a Fixed-Wing UAV," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 4, pp. 1320–1333, Aug. 2013.

[16] W. Kong, D. Zhou, Y. Zhang, D. Zhang, X. Wang, B. Zhao, C. Yan, L. Shen, and J. Zhang, "A ground-based optical system for autonomous landing of a fixed wing UAV," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sep. 2014, pp. 4797–4804, iSSN: 2153-0866.

[17] W. Kong, D. Zhang, X. Wang, Z. Xian, and J. Zhang, "Autonomous landing of an UAV with a ground-based actuated infrared stereo vision system," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Nov. 2013, pp. 2963–2970, iSSN: 2153-0866.

[18] D. Tang, T. Hu, L. Shen, D. Zhang, W. Kong, and K. H. Low, "Ground Stereo Vision-Based Navigation for Autonomous Take-off and Landing of UAVs: A Chan-Vese Model Approach," *International Journal of Advanced Robotic Systems*, vol. 13, no. 2, p. 67, Mar. 2016.

[19] Z. Ma, T. Hu, and L. Shen, "Stereo Vision Guiding for the Autonomous Landing of Fixed-Wing UAVs: A Saliency-Inspired Approach," *International Journal of Advanced Robotic Systems*, vol. 13, no. 2, p. 43, Mar. 2016.

[20] E. Kaufmann, M. Gehrig, P. Foehn, R. Ranftl, A. Dosovitskiy, V. Koltun, and D. Scaramuzza, "Beauty and the Beast: Optimal Methods Meet Learning for Drone Racing," in *2019 International Conference on Robotics and Automation (ICRA)*, May 2019, pp. 690–696, iSSN: 2577-087X.

[21] O. Araar, N. Aouf, and I. Vitanov, "Vision Based Autonomous Landing of Multirotor UAV on Moving Platform," *Journal of Intelligent & Robotic Systems*, vol. 85, no. 2, pp. 369–384, Feb. 2017.

[22] M. Beul, S. Houben, M. Nieuwenhuisen, and S. Behnke, "Fast autonomous landing on a moving target at MBZIRC," in *2017 European Conference on Mobile Robots (ECMR)*, Sep. 2017, pp. 1–6.

[23] M. Beul, M. Nieuwenhuisen, J. Quenzel, R. A. Rosu, J. Horn, D. Pavlichenko, S. Houben, and S. Behnke, "Team NimbRo at MBZIRC 2017: Fast Landing on a Moving Target and Treasure Hunting with a Team of MAVs," *Journal of Field Robotics*, vol. 36, no. 1, pp. 204–229, Jan. 2019, arXiv: 1811.05471.

[24] M. Laiacker, K. Kondak, M. Schwarzbach, and T. Muskardin, "Vision aided automatic landing system for fixed wing UAV," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Nov. 2013, pp. 2971–2976, iSSN: 2153-0866.

[25] C. Li, W. Jin, D. Li, and Y. Xi, "Vision-aided Automatic Landing Design for Small Twin-engine Fixed Wing UAV," in *2019 IEEE 15th International Conference on Control and Automation (ICCA)*, Jul. 2019, pp. 435–440, iSSN: 1948-3457.

[26] L. Burlion and H. de Plinval, "Toward vision based landing of a fixed-wing UAV on an unknown runway under some fov constraints," in *2017 International Conference on Unmanned Aircraft Systems (ICUAS)*, Jun. 2017, pp. 1824–1832.

[27] S. Thurrowgood, R. J. D. Moore, D. Soccol, M. Knight, and M. V. Srinivasan, "A Biologically Inspired, Vision-based Guidance System for Automatic Landing of a Fixed-wing Aircraft," *Journal of Field Robotics*, vol. 31, no. 4, pp. 699–727, 2014.

[28] A. Bry, C. Richter, A. Bachrach, and N. Roy, "Aggressive flight of fixed-wing and quadrotor aircraft in dense indoor environments," *The International Journal of Robotics Research*, vol. 34, no. 7, pp. 969–1002, Jun. 2015.

[29] D. H. Shim, S. Huh, and B.-m. Min, "A vision-based automatic landing system for fixed-wing UAVs using an inflated airbag," in *AIAA Guidance, Navigation and Control Conference and Exhibit*, 2008, p. 7450.

[30] K. R. Chandra and S. Ghosh, "Hu-Moment-Based Autonomous Landing of a UAV on a Hemispherical Dome," in *2019 International Conference on Unmanned Aircraft Systems (ICUAS)*, Jun. 2019, pp. 19–25, iSSN: 2575-7296.

[31] D. B. Barber, S. R. Griffiths, T. W. McLain, and R. W. Beard, "Autonomous Landing of Miniature Aerial Vehicles," *Journal of Aerospace Computing, Information, and Communication*, vol. 4, no. 5, pp. 770–784, May 2007, publisher: American Institute of Aeronautics and Astronautics.

[32] "PX4: An open source flight control software for drones and other unmanned vehicles," library Catalog: px4.io. [Online]. Available: https://px4.io/

[33] "mavros." [Online]. Available: http://wiki.ros.org/mavros

[34] T. Krajník, M. Nitsche, J. Faigl, P. Vaněk, M. Saska, L. Přeučil, T. Duckett, and M. Mejail, "A Practical Multirobot Localization System," *Journal of Intelligent & Robotic Systems*, vol. 76, no. 3-4, pp. 539–562, Dec. 2014.

[35] "QGC." [Online]. Available: http://qgroundcontrol.com/

[36] P. Oettershagen, A. Melzer, S. Leutenegger, K. Alexis, and R. Siegwart, "Explicit model predictive control and L1-navigation strategies for fixed-wing UAV path tracking," in *22nd Mediterranean Conference on Control and Automation*. Palermo, Italy: IEEE, Jun. 2014, pp. 1159–1165.

[37] S. Park, J. Deyst, and J. How, "A New Nonlinear Guidance Logic for Trajectory Tracking," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*. Providence, Rhode Island: American Institute of Aeronautics and Astronautics, Aug. 2004.

[38] S. Park, J. Deyst, and J. P. How, "Performance and Lyapunov Stability of a Nonlinear Path Following Guidance Method," *Journal of Guidance, Control, and Dynamics*, vol. 30, no. 6, pp. 1718–1728, 2007.

[39] b. P. D. Bièvre, "The 2012 International Vocabulary of Metrology: "VIM"," *Chemistry International – Newsmagazine for IUPAC*, vol. 34, no. 3, pp. 26–27, May 2012, publisher: De Gruyter Section: Chemistry International.

APPENDIX

We additionally demonstrate the results of our outdoor experiments implementing the visual-servoing approach presented in [29] on our proposed fixed-wing drone. In our tests, we simulated visual detections by projecting the relative position vector from the drone to the target into the body frame. This allowed us to calculate the heading and pitch angle deviations ($\Delta\psi$, $\Delta\theta$) that were used in [29] to servo the drone to the target.

In our first experiments, we applied the control as it was described in the paper. While the control applies corrective actions, it introduces oscillations that lead to an unstable behavior on our drone. We show a demonstrative example of the resulting flight behavior in Fig. 9 as a red dashed line ([29], P gain 1.0). The figure shows the top view of the trajectory flying towards the target at (0,0), indicated by the red circle with a radius of 1.5 m. We rotated the trajectory such that the the starting point is aligned with the x-axis. The trajectory ends, when our safety pilot took over the control because the oscillations were leading to dangerous flight behavior.

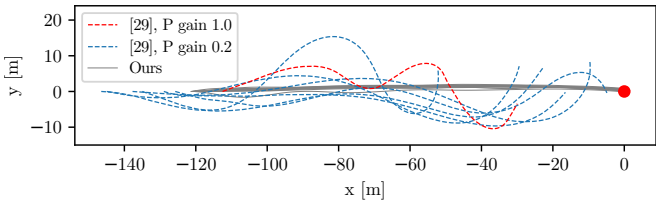


Fig. 9. Top down view of flight experiments with the vision-based control approach presented in [29] of flying towards a target at (0,0) depicted by the red circle. The red dashed trajectory depicts a flight paths following the control as described in [29], while the dashed blue trajectories represent flights where an additional proportional gain (P gain) of 0.2 was introduced. For comparison, we add 15 flights of our proposed flight approach in grey.

In an attempt to reduce the oscillations, we introduced a proportional gain (P gain) term and set it to 0.2 which should reduce the direct effect an error has on the guidance control. We recorded six flights but still observed unstable behavior with increasing oscillation over the duration of the trajectory for all flights.

In our experiments with a P gain of 0.2, we observed oscillations on the length scale of about 70 m. Decreasing the P gain further may potentially lead to stable behavior, however, in turn it would lead to a less reactive control. Considering that the first target detections on our platform are on average at a distance of 35 m, further decreasing the P gain would likely lead to a control system that is not reactive enough.

For comparison, we additionally visualize the 15 flights from the experiment described in section IV-E using our proposed vision-GNSS approach.

Since the introduced P gain was not part of the original work [29] and all shown flights using their approach resulted in unstable flights, we conclude that the control proposed in [29] is platform dependent and may generally lead to unstable behavior.