

Embodied Flight with a Drone

Alexandre Cherpillod
Laboratory of Intelligent Systems (LIS)
Ecole Polytechnique Fédérale de
Lausanne (EPFL)
 Lausanne, Switzerland
 alexandre.cherpillod@epfl.ch

Stefano Mintchev
Laboratory of Intelligent Systems (LIS)
Ecole Polytechnique Fédérale de
Lausanne (EPFL)
 Lausanne, Switzerland
 stefano.mintchev@epfl.ch

Dario Floreano
Laboratory of Intelligent Systems (LIS)
Ecole Polytechnique Fédérale de
Lausanne (EPFL)
 Lausanne, Switzerland
 dario.floreano@epfl.ch

Abstract—Most human-robot interfaces, such as joysticks and keyboards, require training and constant cognitive effort and provide a limited degree of awareness of the robots' state and its environment. Embodied interactions, that is the bidirectional link between the physical bodies and control systems of the robot and of the human, could not only enable a more intuitive control of robots, even for novices, but also provide users with more immersive sensations. But providing an embodied interaction by mapping human movements into a non-anthropomorphic robot is particularly challenging. In this paper, we describe a natural and immersive embodied interaction that allows users to control and experience drone flight with their own bodies. The setup uses a commercial flight simulator that tracks hand movements and provides haptic and visual feedback. The paper discusses how to map body movement with drone motion, and how the resulting embodied interaction provides a more natural and immersive flight experience to unskilled users with respect to a conventional RC remote controller.

Keywords—Aerial Robotics, Telerobotics and Teleoperation, Virtual Reality and Interfaces

I. INTRODUCTION

There are several situations where humans are required to control distal robots, for example in exploration of remote areas, inspection, or monitoring of disaster areas. Human-robot interfaces (HRI) can significantly improve the symbiosis between the human and machine [1]. However, most current HRIs—such as joysticks, keyboards, and touch screens—require user training and concentration during operation; therefore, they are limited by a physical and cognitive effort that is often a barrier between the user and the robot. Force feedback through a joystick or through an exoskeleton [2] can improve human robot interaction, but is often limited to direct mappings between human body parts and corresponding body parts of an anthropomorphic robot (a fingered gripper, a robotic arm, a humanoid, e.g.).

A different approach relies on embodied interactions rather than interfaces, which is the bidirectional link between the physical bodies and control systems of the robot and of the human. The ultimate goal is not only a more natural and effective control of distal robots, but also the physical transformation of a human into a sensory-motor system (robot) with a different morphology and behavior. With embodied interactions based on a more natural form of control for humans, for example, through gestures, and an immersive and rich multi-modal feedback, the user can ultimately embody a robot with non-anthropomorphic morphology; hence, the user's body and perception can be seamlessly blended with a



Fig. 1. The main components of the setup that allows users to “fly like a drone”. The Birdly® platform, has visual, haptic and vestibular (pitch and roll, yellow and orange arrow respectively) feedback. The pronation and supination movements of user's hands (red arrows) are transformed into drone motions and the movement of her head is reproduced by the camera gimbal of the drone. The pitch and roll angles of the drone are transformed into pitch and roll angles of the Birdly® platform, its speed is transformed into speed of the air fan, and the image from the drone camera is projected in the visual goggles.

distal machine. Embodiment, also known as “experiencing presence” in the field of virtual reality [3], results from a combination of self-localization—when the user projects himself onto the distal robot—and self-ownership—when the user feels the artificial body as his/her own [4], [5]. Embodiment is usually triggered by tailored sensory-motor stimulations that enhance the feelings of self-localization and self-ownership [4], [5]. In proper embodied interactions, the distal robot becomes transparent to the user, whose actions will be defined by only the environment [6]. For example, a user embodying a drone should experience a sensation of flight.

Among the various mobile robots with non-anthropomorphic morphology, drones are interesting candidates for embodied interactions because of their remarkable capability to extend human perception and range of action.

This paper first reviews the state-of-the-art research in human-drone interaction. Then it demonstrates through different mapping strategies the advantage of replacing a standard control interface, an RC remote controller with a natural and immersive setup, Birdly® (Somniacs^{SA}, Zurich, Switzerland) (Fig. 1). Birdly® is a flight simulator that allows the user to fly in a virtual environment with hand gestures, reproducing the flight of a bird.

II. RELATED WORK

In human-drone interaction, we discuss two types of control methods: non-gestural and gestural control.

A. Non-gestural Control

Non-gestural controllers are widely used in drones. RC remote controller, joysticks, and touch screens are widely used by all major drone manufactures (DJI, Parrot, 3D Robotics). Despite their popularity, these interfaces are not immersive and sometimes difficult to use. For example, it was found that people could not complete their flying tasks using joysticks and touch interfaces (iPhone), whereas they could successfully complete these tasks using upper body gestural control [7]. Brain-machine interaction (BMI) has also been used to control drones. LaFleur *et al.* [8] developed a controller based on electroencephalographic (EEG) signals. The user had to think left, right, up or down to steer the drone. However, the use of EEG signals requires the users to maintain a high degree of concentration during the flight to reduce the noise in the EEGs signals. Moreover, the use of EEG signals requires time-consuming user calibration. Gaze gesture control [9] has also been used in drones. However, this technique requires the user to focus his/her gaze for the control of the drone and prevents the user from exploring the environment.

B. Gestural Control

Studies on the gestural control of drones can be divided into two main categories: third-person view and first-person view (FPV). In the third-person view, the user acts on the drone as an external viewer, whereas in the first-person view, the user acts from the perspective of the drone.

In the third-person view, the user usually exploits body gestures to interact with single [10] or multiple drones that are within the line of sight [11]. The inherent issue with third-person view is the lack of immersion as the user acts on the drone from an external perspective.

In the first-person view, head motion is commonly used to directly control the drone [12] (Fig. 2(a)). The user often wears goggles to receive feedback from the video captured by the drone. Nevertheless, head control systems have limitations because they do not allow the user to explore by looking around; the user could lose track of his/her desired trajectory or even crash by just looking at the environment.

Another approach is to fly a drone using upper body gestures. Pfeil *et al.* [13] developed a controller based on torso and arm gestures, but they defined a set of five different gestural controls (three of them are shown on Fig. 2(b)). Sakamoto *et al.* [14] tested an adapted version of the *first person* method developed by Pfeil *et al.* [13] but in the laying position. These gestural controls have two main drawbacks. First, they all use discrete mapping (i.e., every motion of the drone is defined based on the threshold of the gestures). This leads to discontinuous movements of the drone, which significantly diminishes the experience of embodiment because the *immediacy of control* is limited [3], that is, subspaces of the control input have no appropriate consequences.



(a)



(b)

Fig. 2. Different types of gestural controls found in the literature. (a) The user controls the position and orientation of the drone with the movement of his head [12]. (b) Different flight styles to control the drone [13].

Some platforms allow users to simulate flight using gesture commands while receiving both visual and haptic feedbacks. The *Humphrey* developed by Formquadrat [15] and the aforementioned Birdly[®] are two examples.

III. IMPLEMENTATION

We aim to compare the naturalness and immersion of flying between the use of Birdly[®] platform and a standard RC remote controller. Birdly[®] allows users to fly in simulation using hands gestures as shown in Fig. 1. Tilting both hands in the same direction allows to control pitch, while tilting them oppositely allows the user to roll. In this experiment, Birdly[®] was interfaced with a fixed wing drone simulator as this type of drone is more compatible than a hovering drone with the gestures and the vestibular feedback provided by Birdly[®].

The comparison is done in simulation. The user wears an Oculus Rift DK2 for the visual feedback. The flight simulator is based on the dynamic of a fixed wing drone. This type of drone has been selected as it is more compatible with the gestures and the vestibular feedback provided by Birdly[®] with respect to a hovering drone.

For the mapping strategies, the hand rotations on Birdly[®] wing paddles must be mapped into fixed-wing controls. One possibility is to consider the hands as the flaps of the fixed-wing drone. In this case, a given angular deflection of the paddles, will translate into an angular velocity for the fixed wing. Thus, we could associate this mapping with an *angular velocity controller (rate controller)*. This controller is usually called *Acro mode*, e.g., in 3DR products and PX4 documentation. It is known to be used by experienced pilots, for example, using remote controllers in drone racing. Another strategy is to map the angle of the paddles into corresponding

angles, or *attitude*, of the fixed-wing drone. Thus, we could associate this mapping with an *attitude controller*. This control mode is called the *Stabilize* mode in 3DR products and in PX4 documentation. These two mapping strategies will be quantified and compared with respect to an RC remote controller.

IV. FROM RC REMOTE TO GESTURAL CONTROL

We compare four different flight conditions: RC remote with attitude or rate controller, and Birdly® with attitude or rate controller. To precisely compare the four combinations, we used a custom-designed virtual simulator of a fixed-wing drone. Fig. 3 shows the simulated environment developed in Unity3D [16]. This comparison aims to determine the most natural and immersive controller by evaluating users' capability to fly through a series of waypoints, which are visualized as small clouds scattered in the sky at different altitudes and directions.

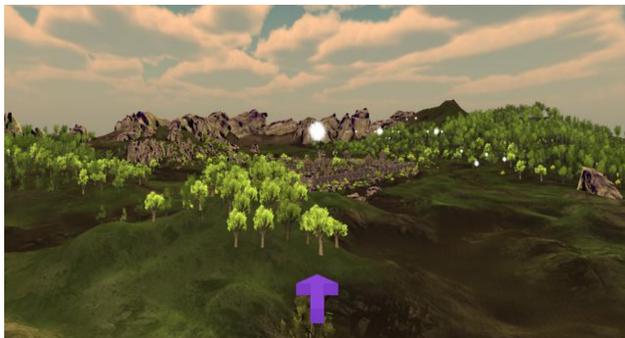


Fig. 3. Flight Environment. The user is required to fly through a series of white points visualized as small clouds in the sky. The arrow on the bottom of the screen always shows the direction of the next waypoint, which is helpful if the user turned away from the waypoint and the cloud is not visible.

Human subjects are first explained how the simulated drone is controlled through Birdly® or the RC remote controller and are then exposed to three successive flight phases: passive flight, training and evaluation. The first phase is a 1-min *passive flight* where the subject sees the environment and gets familiar with it. The subject cannot control the drone but receives feedback (visual, haptic, vestibular, and auditory for the Birdly®, visual feedback only for the remote controller). During this first phase, the simulated drone flies autonomously through a sequence of clouds and the subject is explained the task he/she will have to do in the next phases which is to “follow the clouds and to be as close as possible to the center”. The subject then starts a 9-min *training phase*, which involves flying through the clouds one after the other. The subject can ask for breaks, during which the simulation is paused. The training phase is followed by an *evaluation phase*, which involves flying through 84 waypoints. The size and flight dynamics of the simulated drone reproduce the fixed-wing drone eBee (senseFly SA). Since the drone speed is maintained constant at 12 m/s which is the cruising speed of the eBee, the average duration of the evaluation phase is approximately 5 min. The spacing between the waypoints is approximately 40 m. Whenever the subject crashes the drone, he/she is repositioned in front of the next waypoint.

The experiments were evaluated with quantitative measures and with a questionnaire. The quantitative measure was based on the distance between the drone and the center of each waypoint, as presented in [17]. 100% performance was obtained when the drone was crossing the center of the waypoint; this value decreases with distance from the center according to a Gaussian function and reached a value of 1% at 38.4 m. This distance has been computed with the data of the evaluation phase and corresponds to the average distance to the waypoints center plus 2.5 times the standard deviation. This is assumed to be an outlier distance [18]. For the qualitative measure, after the experiments subjects were asked to rate several statements shown in Table II to indicate their degree of agreement (from 1 to 7). The statements were intended to assess the level of immersion and the naturalness of the control strategy.

A. Results

42 subjects (37 male, 5 female, age range 19-51 years, average age 28 years) participated in the experiment. 2 subjects felt nauseous (1 male, 1 female) and discontinued the experiment. Among the 40 valid subjects, 10 reported to already have directly or remotely piloted an aircraft for more than one hour. Out of the 40 subjects, 20 tested the RC remote controller and 20 the Birdly®. For both interface 10 were presented with the angular velocity controller and the other 10 were presented with the attitude controller, leading to four different combinations of interface and mapping strategy. Fig. 4 illustrates the performance of the subjects with the two interfaces and the two mapping strategies during the training phase. For ease of visualization, the curves represent an average window over 20 waypoints.

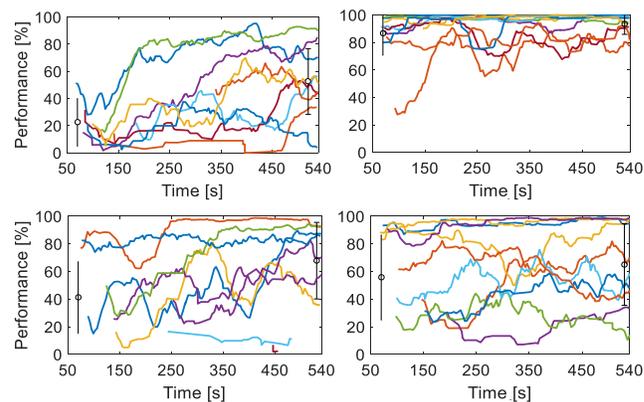


Fig. 4. Training performance. Curves represent an average window over 20 waypoints. The mean and standard deviation of the starting and ending performance are shown with error bars. Birdly® with angular velocity controller (a) and with attitude controller (b). RC remote controller with angular velocity controller (c) and with attitude controller (d).

From Fig. 4, we notice that globally subjects improved their skills all along the training. For each combination, the Mann-Whitney U-test confirmed that there is a significant difference between the starting and ending performance. According to Wigdor [19], a natural controller must make the user an expert in a short training period. From the training, we can see that the attitude controller using Birdly® has an average starting and

ending performance higher than the other combinations. Indeed, its starting and ending performance are significantly higher than the attitude and angular velocity controller using the RC remote controller and the angular velocity controllers using Birdly® ($p < .001$).

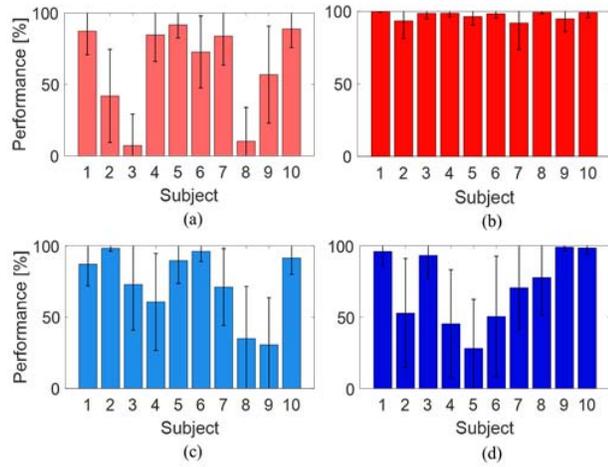


Fig. 5. Average performances of each subject during the evaluation phase with their standard deviations. (a) Birdly® with angular velocity controller. (b) Birdly® with attitude controller (c) RC remote controller with angular velocity controller. (d) RC remote controller with attitude controller.

Fig. 5 shows the average performance levels for each subject during the evaluation phase. Table I shows the average performances of the subjects after the evaluation for each type of interface and control strategy. The results show that except the attitude controller using Birdly® the other combination have a higher variability in term of performance and contain subjects with a performance below 50%. The attitude controller using Birdly® presents an average performance which is significantly higher compared to other combination ($p < .01$). These data from the evaluation confirm that using Birdly® with an attitude controller is more effective and more natural, according to Wigdor [19] as the level of expertise is higher. Interestingly, the performances of both mapping strategies using the RC remote controller are not significantly different ($p > 0.1$) in the case of using Birdly® they are significantly different ($p < .01$). This demonstrates the difficulty encountered by the users using an RC remote controller independently of the mapping strategy.

Table I. Average performance during the evaluation phase

Control	Mapping strategy	Average performance [mean ± SD]
Birdly	Angular velocity controller	63.2% ± 38.1
	Attitude controller	97.3% ± 8.1
RC remote control	Angular velocity controller	74.7% ± 33.5
	Attitude controller	71.1% ± 37.0

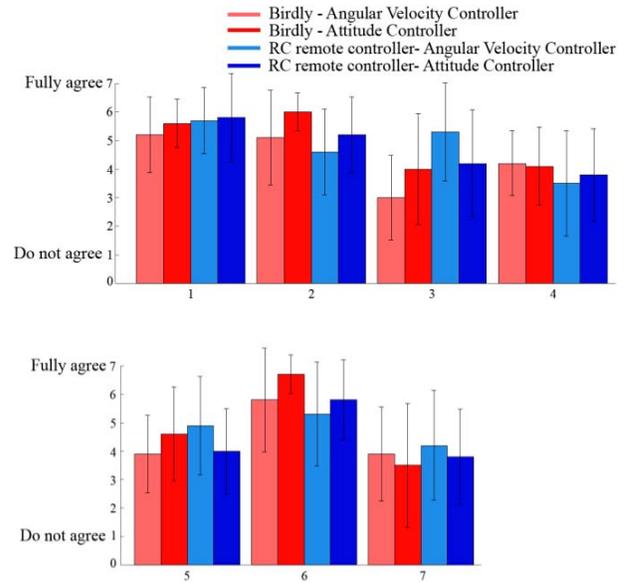


Fig. 6. Questionnaire results. Mean and standard deviation. The questions are presented in Table II. Higher agreement is better, except for question 7.

The results of the questionnaire are presented in Fig. 6. The respective questions are shown in Table II. From the training and the evaluation, we saw that the angular velocity controller using the RC remote controller has in both cases the second best performance, see Table I. But regarding the questionnaire, people found that using Birdly® with the attitude controller was giving a better sensation of controlling the flight trajectory (question index #2) ($p < .5$) and they enjoyed it more (question index #6) ($p < .05$). This confirmed that the attitude controller using Birdly® is more immersive and natural. Indeed according to Wigdor [19], the enjoyability of the interface is another factor of its naturalness.

Table II. Questionnaire filled out by the subjects at the end of the experiment. Subjects provide ratings between 1 (Disagree) and 7 (Strongly agree).

#	Question
1	I felt as if I was flying myself.
2	I had the sensation of controlling the flight trajectory.
3	I felt some physical discomfort.
4	If I could fly in the air, I would use the same gesture.
5	The proposed gesture control was natural for me.
6	I enjoyed the experiment.
7	I found the experiment tiring.

V. CONCLUSION

In this paper, we used of Birdly® to fly like a drone exploiting embodied interactions. We found that using a fixed-wing drone and a mapping between user gestures and drone's commands based on an attitude controller enable an immersive and natural flight experience. We also found that using a gestural control

based on Birdly[®] is more natural and immersive than using standard RC remote controller.

Preliminary experiments with a real drone are available in an extended version of the paper available at arXiv¹. In this extended version we also addressed the technical challenges of interfacing Birdly[®] with a real drone.

One future challenge is to develop a system more portable than the Birdly[®] platform but having similar gestural controls and feedback. A promising avenue is to use a lightweight exosuit for gesture recognition and haptic feedback [20], [21]. Another challenge is to add a shared control that adaptively shift the control between the user and the autopilot of the drone based on the skills or cognitive fatigue of the user.

ACKNOWLEDGMENTS

The authors thank the Swiss NCCR Robotics Symbiotic Drone Project team for their advice on the design of the experiments protocol. We also thank the MAVRIC team for the software development of the drone and Grégoire Heitz for his useful advice and devotion as a safety pilot during the outdoor experiments. This work was supported by the Swiss National Science Foundation through the National Centre of Competence in Research Robotics.

REFERENCES

- [1] T. C. Spotlight, "Benefits of Human – Robot Interaction," *IEEE Robotics and Automation Magazine*, pp. 98–99, 2011.
- [2] A. Wang, J. Ramos, J. Mayo, W. Ubellacker, C. Justin, and S. Kim, "The HERMES Humanoid System : A Platform for Full-body Teleoperation with Balance Feedback," *Humanoid Robot. (Humanoids), 2015 IEEE-RAS Int. Conf.*, pp. 730–737, 2015.
- [3] B. G. Witmer and M. J. Singer, "Measuring Presence in Virtual Environments: A Presence Questionnaire," *Presence: Teleoper. Virtual Environ.*, vol. 7, no. 3, pp. 225–240, 1998.
- [4] B. Lenggenhager, T. Tadi, T. Metzinger, and O. Blanke, "Video ergo sum: manipulating bodily self-consciousness," *Science*, vol. 317, no. 5841, pp. 1096–1099, 2007.
- [5] C. Lopez, P. Halje, and O. Blanke, "Body ownership and embodiment: Vestibular and multisensory mechanisms," *J. Neurophysiol. Clin.*, vol. 38, no. 3, pp. 149–161, 2008.
- [6] P. Dourish, *Where the Action Is*, MIT Press. 2004.
- [7] A. Sanna, F. Lamberti, G. Paravati, and F. Manuri, "A Kinect-based natural interface for quadrotor control," *J. Entertain. Comput.*, vol. 4, no. 3, pp. 179–186, 2013.
- [8] K. LaFleur, K. Cassady, A. Doud, K. Shades, E. Rogin, and B. He, "Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain-computer interface," *J. Neural Eng.*, vol. 10, no. 4, p. 046003, 2013.
- [9] M. Yu, Y. Lin, X. Wang, D. Schmidt, and Y. Wang, "Human-Robot Interaction Based on Gaze Gestures for the Drone Teleoperation," *J. Eye Mov. Res.*, vol. 7, no. 4, pp. 1–14, 2014.
- [10] V. M. Monajjemi, J. Wawerla, R. Vaughan, and G. Mori, "HRI in the sky: Creating and commanding teams of UAVs with a vision-mediated gestural interface," *IEEE Int. Conf. Intell. Robot. Syst.*, pp. 617–623, 2013.
- [11] A. Stoica, F. Salvioli, C. Flowers, and L. Lane, "Remote Control of Quadrotor Teams , Using Hand Gestures," in *International Conference on Human-Robot Interaction (HRI)*, 2014, pp. 296–297.
- [12] H. Hayakawa, C. Lasantha Fernando, M. Yamen Sarajji, M. Kouta, and S. Tachi, "Telexistence Drone : Design of a Flight Telexistence System for Immersive Aerial Sports Experience," in *Augmented Human International Conference*, 2015, pp. 171–172.
- [13] K. Pfeil, S. L. Koh, and J. LaViola, "Exploring 3d gesture metaphors for interaction with unmanned aerial vehicles," in *2013 international conference on Intelligent User Interfaces (IUI 13)*, 2013, pp. 257–266.
- [14] M. Sakamoto *et al.*, "Human Interaction Issues in a Digital-Physical Hybrid World," *2014 IEEE Int. Conf. Cyber-Physical Syst. Networks, Appl.*, pp. 49–54, 2014.
- [15] "Immersive Flight - formquadrat." [Online]. Available: <http://www.aec.at/futurelab/referenzen/alle-jahre/2003/humphrey-ii/>. [Accessed: 11-Oct-2018].
- [16] "Unity3D." [Online]. Available: <https://unity3d.com/>. [Accessed: 11-Oct-2018].
- [17] C. Rognon, A. R. Wu, S. Mintchev, A. Ijspeert, and D. Floreano, "Haptic Guidance with a Soft Exoskeleton Reduces Error in Drone Teleoperation," *Haptics Sci. Technol. Appl.*, vol. 10894 LNCS, pp. 404–415, 2018.
- [18] J. Miller, "Reaction time analysis with outlier exclusion: bias varies with sample size," *Q. J. Exp. Psychol. Sect. A Hum. Exp. Psychol.*, vol. 43, no. February 2015, pp. 907–912, 1991.
- [19] D. Wigdor, Daniel; Wixon, *BRAVE NUI WORLD designing natural user interfaces for touch and gesture*. Morgan Kaufman Publishers, 2010.
- [20] C. Rognon, S. Mintchev, F. Dell'Agnola, A. Cherpillod, D. Atienza, and D. Floreano, "FlyJacket: an upper-body soft exoskeleton for immersive drone control," *IEEE Robot. Autom. Lett.*, vol. 3, no. 3, pp. 1–1, 2018.
- [21] J. Miehlebradt *et al.*, "Data-driven body-machine interface for the accurate control of drones," *Proc. Natl. Acad. Sci.*, vol. 115, no. 31, pp. 7913–7918, 2018.

¹ <https://arxiv.org/abs/1707.01788>