Using Pose Estimation for Cultivating Subtleties in Human-Drone Interaction

JOSEPH LA DELFA, RMIT University
MEHMET AYDIN BAYTAŞ, Chalmers University of Technology
NICK HUPPERT, RMIT University
IAN PEAKE, RMIT University
LEAH HEISS, RMIT University

Human-drone interaction based on direct manipulation is a burgeoning concept in HCI, offering a tightly coupled movement-based experience with a drone. The technologies behind these interactions can also generate data on the movements of various parts of the body, which accumulate into a rich dataset over time. How can we leverage the richness of these datasets to tailor how the robotic interaction responds to an individual's movements? Towards this end, we are experimenting with PoseNet, a camera-based pose recognition algorithm, which can track various features of the body. By moving with the system, we are looking for ways to measure how our movements change over time. We hope to design a system that encourages reflection on past movements and exploration of new movements.

CCS Concepts: • **Computer systems organization** → Robotics.

Additional Key Words and Phrases: deep learning, drones, human-drone interaction, motion capture, neural networks, pose estimation

ACM Reference Format:
Joseph La Delfa, Mehmet Aydın Baytaş, Nick Huppert, Ian Peake, and Leah Heiss. 2020. Using Pose Estimation for Cultivating Subtleties in Human-Drone Interaction. In *Programming for Moving Bodies, October 25–26, 2018, Talinn, Estonia.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Human drone interaction (HDI) that is closely coupled to the human body is a growing topic at the intersections of HCI, design, and research [1, 2, 4, 5]. Through these designs, researchers have found that tightly coupling bodily movement to a drone enables new possibilities for interactive digital experiences, whilst simultaneously getting in touch with and reflecting on how our own bodies move. One example is the 2019 research project "Dancing with Drones" by Eriksson et al., which documents the experience and morphology of human-drone co-movement through the design and choreography of stage performance [7]. Over months of exchanges between design, engineering, and performance specialists, the project demonstrates how it is not only the drone technology that changes throughout the process, but the participants' own bodies as well.[1] Another example is La Delfa et al.'s Drone Chi, a Tai Chi-inspired HDI design which encourages people to reflect on how they are moving in relation to a drone [10–13]. During the process of

[1]In later analysis, Eriksson et al. (2020) also document "how ethicality is shaped in interaction" between the participants and the robotic agents, and "how movements enabled by the human-drone assemblage may limit or liberate artistic expressions" [6].

creating Drone Chi, the design team also had experiences of closely relating to how the movements of the drones evolved daily, and their own ever-changing bodies contributed and responded to this process.

Even after the aforementioned designs are 'finished,' the ever-changing human body still registers a slightly different experience with every encounter. Each performance with the choir of drones is slightly different from the last, while each meditation with Drone Chi reveals something new and different about the body. Yet, the robotic material and the interactions remain the same by design. We could argue that there is value in an engagement with such systems while exerting reflective effort; that there is value, in being willing to look for and experience changes in their body and by extension the drone's. The design of these systems can already provide a space for the person to exert those efforts without distraction, but we wonder if we can design to offer more. HDIs, for example, can be tailored to evoke a sense of curiosity in an evolving robotic interaction, and how the body relates to it – a curiosity to influence the development of the interaction itself through the body, to tinker with the relationship between the two, just as we did when we were designing these systems. How might we capture the experiences we had while designing and iterating on these systems, and imbue the experiences of the end-users or participants with these qualities?

In this position paper we introduce the early-stage concept How to Train Your Drone (HTTYD), which responds to the motivations above. We discuss the design process, outline the current technical experiments, and lay out relevant challenges. Thus we aim to contribute to the discussions on the potentials of new tools and methods for programming with and for movement, mapping bodily movements to computational composite materials, working with bodily data as a design material, and catering for changing bodies.

2 CONCEPT: HOW TO TRAIN YOUR DRONE

How to Train your Drone (HTTYD) builds on our experiences in designing and studying Drone Chi. HTTYD combines a focus on reflective movement (which defined the essence of Drone Chi) with support for temporality and dynamism: for changing and shaping the behaviour of the drone and the body over time. While designing Drone Chi, we observed that the evolving relationship between our bodies and the technological materials influenced the design, as well as our selves. The aim of HTTYD, is to capture this kind of long-term, subtle, evolutionary relationship experienced by the designers, and to make it available to the end-users or participants.

The context where we envision people will encounter HTTYD can be similar to a gymnasium: an open space, where a person can move freely and experiment. As a design research project, we envision conducting an experience study where we recruit a small number (<10) of participants to regularly visit this gymnasium over a multi-month period and conduct prescribed exercises, as well as free explorations. At the end of this period, we would ask the participants to try each other's drones. Through data sources that can include interviews, videos, and body sheets, we aim to cultivate a picture of how these drones reflect and affect an individual's movement. As much as we are interested in questions around how to design such systems, we are interested in what a person learns about themselves through training their own drone, and what a person learns about another person through trying their drone.

While designing Drone Chi, we developed a certain "offset-midpoint" mapping that determines how a drone can responds to the movements of the hands [12]:

> Based on motion capture data, we first compute two offset points that reside on imaginary rays shooting perpendicularly out of the palms, at 20 cm away from them. We then compute the midpoint of the line segment straddling the two offset points – the offset-midpoint. The vector Δ between the offset-midpoint and the drone drives the interaction.

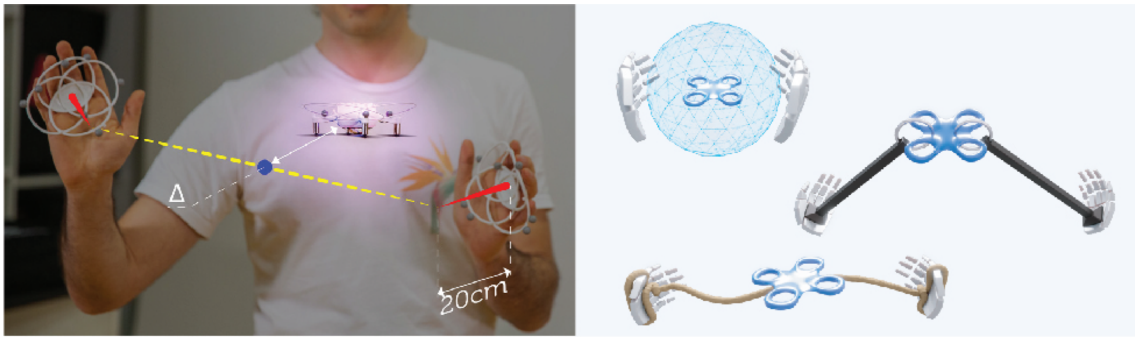

Fig. 1. The offset midpoint mapping of Drone Chi (left), and alternative mappings explored in the design process (right).

The Drone Chi drone thus follows this offset-midpoint, with a limited top speed. During the creative process, by varying aspects of the mapping like the lengths of the 'imaginary rays' or the speed of the drone, we could change the feel of the interaction. Similarly, the different types of mappings explored during the early design stages each had a different feel as well (Figure 1).

During the design process, in order to change the mapping in any way, the designer first imagined and acted them out (informed by "embodied sketching" [8, 9, 15]), which was followed by sitting down and changing lines of code. With HTTYD, we aim to alter this particular workflow in altering the design, enabling alterations on the mapping in a more fluid, embodied workflow.

One direction to explore is to utilize metrics or 'computable descriptors' for how the whole body is moving [14], and utilize these quantities to modulate variables in the mapping scheme, thus modulating the "feel" of the interaction. As an example: the average speed of each hand over the course of a time period can be taken into account to move the offset-midpoint (blue dot in Figure 1) – and therefore the drone – towards the slower hand. Or: we could take into account that a person's center of mass (or geometric centroid of tracked limb nodes) whilst they are moving with the drone will be different from when they are standing still. This difference can then be mapped to the "top speed" of the drone, encouraging a person to perform more dynamic and elaborate movements over time. Figure 2 depicts the latter concept, and illustrates how this information can be sensed and computed via pose estimation using a webcam. At the workshop on Programming for Moving Bodies, we are interested in discussing possibilities such as these, and establishing collaborations for future experiments.

3 IMPLEMENTATION

HTTYD builds on our experiences with technological materials which we utilized to build Drone Chi – drones and motion tracking [3, 10, 12]. One of the main goals in HTTYD is to reduce the cost and engineering complexity of these materials by employing commodity components instead of high-precision systems. For example, tracking and control of the drone in Drone Chi relied on motion capture cameras, while for HTTYD we are currently experimenting with commodity tracking beacons which were originally developed for virtual applications (e.g. HTC Vive).

For human pose estimation, we are experimenting with the PoseNet model[2] which is able to perform real-time pose estimation from monocular RGB video in a desktop application or browser environment. PoseNet is an open-source model that is distributed with the TensorFlow machine learning platform, which is well-documented and enables rapid

[2]https://github.com/tensorflow/tfjs-models/tree/master/posenet

Fig. 2. (Top) Celebrated Artist Bill T Jones using Google's PoseNet during a collaboration. © Google.com (Bottom) We visualise the centroid of the data points created by Bill's body with a large red dot.

iteration on design parameters relevant to our purpose. It should be made clear that the pose estimation technology is not being considered as the primary mode of controlling the drone's flight path. This task is still left to a more precise and accurate system such as motion capture cameras or virtual reality tracking beacons. At the workshop, we are aiming to demonstrate a running application which interfaces PoseNet with a virtual drone.

4 CHALLENGES

PoseNet delivers a set of points in a 2D plane, while it is possible to coarsely calculate 3D data from this, refraining from doing so may lead to interesting results from the metrics mentioned above. We see this as a restriction of the technology that could yield a novel interactive outcome.

The key usability challenge is to balance various considerations around intuitiveness and ease of use with capacities for supporting unique relationships between people and their drones. We noted during experiments with Drone Chi that it was important to design for "first contact," making basic control mappings and system behavior explicit through onboarding [12]. Ideally, we would like the system to possess a level of intuitiveness that allows the majority of people to begin using the drones with little to no instruction. As well as a level of complexity allows an individual to design a mapping scheme unique to them. We look forward to discussing how to address these challenges and others at the workshop.

REFERENCES

[1] Mehmet Aydın Baytaş, Damla Çay, Yuchong Zhang, Mohammad Obaid, Asım Evren Yantaç, and Morten Fjeld. 2019. The Design of Social Drones: A Review of Studies on Autonomous Flyers in Inhabited Environments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. ACM, New York, NY, USA, Article 250, 13 pages. https://doi.org/10.1145/3290605.3300480

[2] Mehmet Aydin Baytas, Markus Funk, Sara Ljungblad, Jérémie Garcia, Joseph La Delfa, and Florian 'Floyd' Mueller. 2020. IHDI 2020: Interdisciplinary Workshop on Human-Drone Interaction. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3375149

[3] Mehmet Aydın Baytaş, Mohammad Obaid, Joseph La Delfa, Asım Evren Yantaç, and Morten Fjeld. 2019. Integrated Apparatus for Empirical Studies with Embodied Autonomous Social Drones. In *1st International Workshop on Human-Drone Interaction*. Ecole Nationale de l'Aviation Civile [ENAC], Glasgow, United Kingdom. https://hal.archives-ouvertes.fr/hal-02128387

[4] Mehmet Aydın Baytaş, Markus Funk, Sara Ljungblad, Jérémie Garcia, Joseph La Delfa, and Florian 'Floyd' Mueller (Eds.). 2020. *Proceedings of the Interdisciplinary Workshop on Human-Drone Interaction (iHDI 2020)*.

[5] Anke M. Brock, Jessica Cauchard, Markus Funk, Jérémie Garcia, Mohamed Khamis, and Matjaž Kljun. 2019. IHDI: International Workshop on Human-Drone Interaction. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–7. https://doi.org/10.1145/3290607.3299001

[6] Sara Eriksson, Kristina Höök, Richard Shusterman, Dag Svanes, Carl Unander-Scharin, and Åsa Unander-Scharin. 2020. Ethics in Movement: Shaping and Being Shaped in Human-Drone Interaction. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376678

[7] Sara Eriksson, Åsa Unander-Scharin, Vincent Trichon, Carl Unander-Scharin, Hedvig Kjellström, and Kristina Höök. 2019. Dancing With Drones: Crafting Novel Artistic Expressions Through Intercorporeality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. ACM, New York, NY, USA, Article 617, 12 pages. https://doi.org/10.1145/3290605.3300847

[8] Kristina Höök. 2018. *Designing with the Body: Somaesthetic Interaction Design*. MIT Press.

[9] Caroline Hummels, Kees C. Overbeeke, and Sietske Klooster. 2007. Move to Get Moved: A Search for Methods, Tools and Knowledge to Design for Expressive and Rich Movement-based Interaction. *Personal Ubiquitous Comput.* 11, 8 (Dec. 2007), 677–690. https://doi.org/10.1007/s00779-006-0135-y

[10] Joseph La Delfa, Mehmet Aydin Baytaş, Emma Luke, Ben Koder, and Florian 'Floyd' Mueller. 2020. Designing Drone Chi: Unpacking the Thinking and Making of Somaesthetic Human-Drone Interaction. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS '20)*. Association for Computing Machinery, New York, NY, USA, 575–586. https://doi.org/10.1145/3357236.3395589

[11] Joseph La Delfa, Mehmet Aydın Baytaş, Olivia Wichtowski, Rohit Ashok Khot, and Florian "Floyd" Mueller. 2019. Are Drones Meditative?. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. ACM, New York, NY, USA, Article INT046, 4 pages. https://doi.org/10.1145/3290607.3313274

[12] Joseph La Delfa, Mehmet Aydin Baytas, Rakesh Patibanda, Hazel Ngari, Rohit Ashok Khot, and Florian 'Floyd' Mueller. 2020. Drone Chi: Somaesthetic Human-Drone Interaction. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376786

[13] Joseph La Delfa, Robert Jarvis, Rohit Ashok Khot, and Florian 'Floyd' Mueller. 2018. Tai Chi In The Clouds: Using Micro UAV's To Support Tai Chi Practice. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts* (Melbourne, VIC, Australia) *(CHI PLAY '18 Extended Abstracts)*. ACM, New York, NY, USA, 513–519. https://doi.org/10.1145/3270316.3271511

[14] Caroline Larboulette and Sylvie Gibet. 2015. A Review of Computable Expressive Descriptors of Human Motion. In *Proceedings of the 2nd International Workshop on Movement and Computing* (Vancouver, British Columbia, Canada) *(MOCO '15)*. Association for Computing Machinery, New York, NY, USA, 21–28. https://doi.org/10.1145/2790994.2790998

[15] Elena Márquez Segura, Laia Turmo Vidal, Asreen Rostami, and Annika Waern. 2016. Embodied Sketching. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. ACM, New York, NY, USA, 6014–6027. https://doi.org/10.1145/2858036.2858486