User-Defined Gesture for Flying Object

ABSTRACT
Nowadays, ways of interaction between human and computer are more diversified; however, these smart products are mostly immobile. With the development in the field of robotics and automation, intelligent products are no longer stationary. Although the present study of interactions are appropriate for tablet or smartphone, they have not yet been fully discussed in the field of flying objects. We presents a user-defined UAV gesture control study in outside without wearable device to collect user behavior. In all, 176 gestures from 16 participants in user study 1 were logged, analyzed and pair with think-aloud data for 11 commands performed. Our insights indicated BLAHBLAHB LAHBLAHBLA HBLAHBLAHBL AHBLAHBLA HBLAHBLAHBL AHBLAHBLAH BLAHBLAHB LAHBLAHBLA LAHBLAHBLA LAHBLAHBLA LAHBLAHBLA LAHBLAHBLA LAHBLAHBLA LAHBLAHBLA LAHBLAHBLA LAHBLAHBLA LAHBLAHBLA LAHBLAHBLA LAHBLAHBLA LAHBLAHBLA LAHBLAHBLA LAHBLAHBLA LAHBLAHBLA.

Our results will help designers and developers create better gesture sets informed by user behavior.

Author Keywords
AR.Drone; UAV; Gestures; User-defined gesture

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation: :User Interfaces - *Interaction styles, evaluation/methodology, user-centered design.*

INTRODUCTION
Unmanned aerial vehicles (UAVs), also known as drones, are aircrafts whose flight is controlled remotely by a pilot or controlled autonomously by computers. Up to date, most people do not possess drones, aircrafts, flying objects, and so forth. Thus, the manipulation over or the interaction with a flying object remains unfamiliar to most people. The goal of this research is to cull, classify, and analyze intuitive gestures created by participants and to conclude the most favorable gestures for each function set of a UAV. We believe that drone manipulation by user-defined gestures should be fun and greatly reduce the learning curve resulting in improvements of human-drone interaction. Moreover, the results of

Figure 1. A user performing a gesture to move the Parrot AR.Drone forward.

this work provide drone designers, developers, and engineers an insight into such novel human-drone gestural interactions.

RELATED WORK
Wobbrock et al. [5] conducted a study of gesture taxonomy for surface computing. His work provides important insight into the design of gestures for tablets. However, the gestures and its taxonomy concluded in this work were subject to 2D touch-screen devices without any further discussion involving 3D in-air gestures.

Vafaei [4] offered more detailed dimensions of gesture taxonomy. He proposed two main groups of the dimensions: gesture mapping and physical characteristics. Although Wobbrock et al.'s and Vafaei's works put stress on touch screens where no in-air gestures were found, the idea of the classification of gestures was quite inspirational for our study.

Hansen et al. [2] explored four different gaze control modes to control a drone. Gaze interaction provides accessibility for disabled people. Nonetheless, the results of the work indicated that the participants mostly favored the mouse over gaze in terms of ease-of-use and reliability. Therefore, future work of similar topics is expected.

Graether and Mueller [1] provided preliminary insights into the design of a flying robot as a jogging companion. The insights includes four themes: embodiment, control, personality, and communication. Graether and Mueller discovered that the communication between the jogger and the flying robot is best through hand gestures.

Pfeil et al. [3] explored five 3D gesture metaphors for interaction with UAVs. Metaphors of proxy manipulation and seated proxy have the highest rating since the gesture sets are more natural, intuitive. Based on the results of their work, participants using gestural control were capable of completing the trial in less time compared to that using smartphone control. Such result is significant and tremendous for the field of human-computer interaction. However, the gesture sets defined by the author instead of users may not be the most intuitive and natural. We believe that user-defined gestures for drones should do even more improvements on the user experience.

USER STUDY

Developing Gesture Set

Our user study was mainly divided into two parts, namely, Study 1 and Study 2. For Study 1, 16 participants were recruited and asked to think aloud and perform in-air gestures for each basic control (e.g. moving forward, backward, left, right, and so forth). After the gestures being developed, we then classified the gesture mappings along several dimensions in reference to the related works. For Study 2, another 16 participants were told to pick their favorite gesture and rate every gesture from each category of basic controls where all the gesture sets were created by the participants from Study 1.

Referents and Signs

Based on the hardware design of Parrot AR Drone 2.0, the 10 basic movement controls included forward, backward, left, and right translation, rotation, ascend, descend, takeoff, landing, hovering. There were also 2 basic camera controls including the activation of picture taking and video shooting. Thus, there was a total of 12 basic controls on the Drone.

Apparatus

Parrot AR.Drone 2.0, a quadrotor vehicle, equipped with a camera facing was chosen as the flying object since it is popular, programmable, and affordable. FreeFlight app is the official smartphone app allowing joystick-like control made by Parrot SA. We made use of an open source Processing project that can control the Drone with keyboard presses. The Processing application ran on a Macbook with OSX 10.9.4, 8 GB RAM, a 2.6 GHz Intel i5 processor.

Participants

32 paid participants with age ranging from 19 to 25 volunteered for the study. None of the participants had experience with piloting a flying object. All participants were right-handed. All were college students recruited from various fields including financial management, law school, chemical engineering, medicine school, and so forth. Half of the participants consisted of 6 females and 10 males attended Study 1 for gesture set development. The other half consisted of 6 females and 10 males took part in Study 2 for gesture ratings.

Procedure

In Study 1, an approximate 90-second short video of the introduction and motivation of this project was played in the beginning. Second, another 60-second short video showed how the drone moves. Then, the participant was asked to stand up and start to come up with 12 gestures for all 12 basic controls in random order. In addition, there was one advance movement control, point-to-specified-destination control, we thought this could be exceptionally useful and powerful relative to basic movement control for the Drone to reach a desired destination. Therefore, the participant had total of 13

Taxonomy of Game Controls		
Form	Static	no motion occurs in G.
	Large-scale	substantial motion occurs in G.
	Small-scale	little motion occurs in G.
Nature	Symbolic	visually depicts a sign
	Pantomimic	imitates real meaningful actions.
	Pointing	points to a specific direction.
	Manipulative	directly manipulates the object.
	Abstract	mapping is arbitrary.
Flow	Discrete	Action is performed after G.
	Continuous	Action is performed during G.
Handedness	Dominant	G. is performed by the user's dominant hand/arm
	Bi-manual	G. is performed using both hands/arms
Hand Shape		free hand, ball holding, bent hand, fist,pinch, open hand index finger, thumb finger, flat hand, ASL-F, ASL-C, ASL-L, ASL-O, ASL-V, ILY
Trace	Straight line	Moving parts trace a straight line.
	Fan shape	Moving parts trace a fan shape.
	Cone	Moving parts trace a cone.
	Expansion/ Contraction	Hand posture changes with expansion or contraction.
	Traceless	G. is static.

Table 1. "G." means "Gesture"

gestures after Study 1. We allowed the participant to change any previously defined gesture if a new idea came to mind. After all gestures being developed, we employed a Wizard of Oz approach to give the participant an idea of how the gestures they created worked.

After Study 1, we categorized the user-defined gesture sets. Since the gesture for left and right were "symmetric", and so were that for ascend and descend, we grouped them into left-right and ascend-descend gesture sets, respectively, leading to 11 gesture sets. All the actions of all 11 gesture sets were recorded in a video which was played for all participants to pick their favorite gesture for each gesture set in Study 2. Then, the favorite gestures along with Wizard of Oz approach were used in the trial. At last, all participants in Study 2 were asked to watch the same video with all gesture sets and rate for every gesture according to their preference using a 5-point Likert scale.

RESULT & DISCUSSION

Classification of 3D Gestures

Gesture taxonomy has been used in several works involving speech gestures, sign language, touch screen gestures, and so forth. However, works involving gestures for flying object seemed rare lately. We managed to explore the most intuitive gestures for UAVs and provide future developers insights for better control methods.

In reference to related works, we manually classified each gesture along seven dimensions: form, nature, flow, dimensionality, handedness, hand shape, and trace. All the gestures

Referents	Mean	Std.
Ascend and Descend	1.00	0
Forward	1.25	0.50
Left and Right	1.50	1.00
Backward	1.50	0.58
Take pictures	2.25	0.50
Land	2.75	0.96
Take off	3.00	0.82
Record videos	3.00	0.82
Hover	3.25	1.26
Point to specified destination	3.50	1.73
Rotate	3.50	0.58
Mean	**2.41**	**0.79**

Table 2. The 11 commands for which participants chose gestures. Each commands conceptual complexity was rated by the 4 authors (1=simple, 5=complex). During the study, each command was presented with an animation and recorded verbal description.

were performed by participants from Study 1. Each dimension contains multiple categories, shown in Table 2.

The dimension of form describes the motions of all gestures. The categories of this dimension includes static, large-scale motion, and small-scale motion. It is important to find out how many body parts are involved so that future developers will choose appropriate sensor devices to capture the motion.

In the nature dimension, symbolic gestures are visual depictions with no simple equivalent in the world of objects and movements. For example, a time out gesture symbolizes the hovering action. Pantomimic gestures imitate a real meaningful action. For example, a participant would have the arms straight with clenched fist hand shape on both hands and one hand above the other as if turning a steering wheel. A pointing gesture points out the direction where the participant intends the drone to fly. Manipulative gestures have direct relation with real world object interaction. For example, moving the hands forward indicates pushing away the drone. Finally, abstract gestures are with arbitrary mappings. For example, putting a hand on the head to take off the drone would be an abstract gesture.

In the flow dimension, if a gesture is discrete, the gesture must be completed first and subsequently bring out the desired task. For instance, the drone starts to move right after the user's right arm is straightened towards right. A gesture would be continuous if the task is performed during gesticulation. The drone rotates around the user as soon as the user faces at a different direction as an example.

In the dimensionality dimension, gestures are classified according to the number of axes involved. The categories include single-axis, double-axis, tri-axis, and six-axis. The descriptions of each category are listed in TABLE X.

Agreement

After all 16 participants had provided gestures for each referent for Study 1, we grouped the gestures within each referent such that each group held identical gestures. Group size was then used to compute an agreement score A that reflects, in a single number, the degree of consensus among participants.

$$A = \frac{\sum_{r \epsilon R} \sum_{P_i \subseteq P_r} \left(\frac{|P_i|}{|P_r|}\right)^2}{|R|} \tag{1}$$

In Eq. 1, r is a referent in the set of all referents R, Pr is the set of proposed gestures for referent r, and Pi is a subset of identical gestures from Pr. The range for A is $\left[|P_r|^{-1}, 1\right]$. As an example, consider agreement for forward in study 1 and study 2. Both had five groups of identical gestures. The former had groups of size 1, 10, 2, 1, and 2; the latter of size 0, 11, 3, 1, and 1. For study 1, we compute

$$A_{forward} = 2\left(\frac{1}{16}\right)^2 + \left(\frac{10}{16}\right)^2 + 2\left(\frac{2}{16}\right)^2 = 0.48 \tag{2}$$

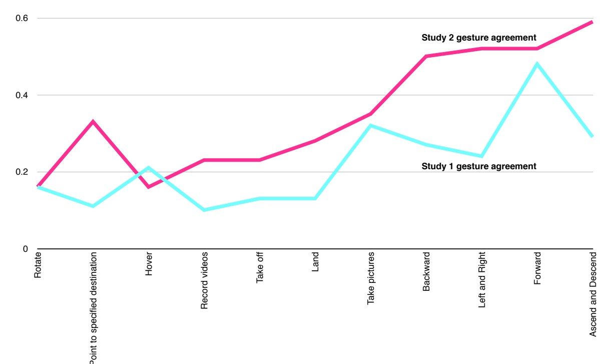

Figure 2. With Caption Below, be sure to have a good resolution image (see item D within the preparation instructions).

CONCLUSION

We have presented a study of interaction between flying object and human gesture, that is leading to a user-defined gesture set based on participants' agreement over 176 gestures. It can reflect the user behavior and this has properties that make it a good candidate for deployment in gesture recognition system. In our opinion, this user-defined gesture can become a good enough reference for developers who study UAV gesture control. We also have presented a taxonomy of gesture for flying object useful for analyzing and characterizing gestures in UAV control. In capturing gestures for this study, we have gained insight into the mental models of non-UAV-control-experience users and have translated theses into implications for technology and design. This work represents a necessary step in bringing interactive flying objects closer to the hands and minds of UAV users. Besides, it also helps designers and developers create gesture interacting to UAV for better user' experience.

ACKNOWLEDGMENTS

We thank CHI, PDC and CSCW volunteers, and all publications support and staff, who wrote and provided helpful comments on previous versions of this document. Some of the

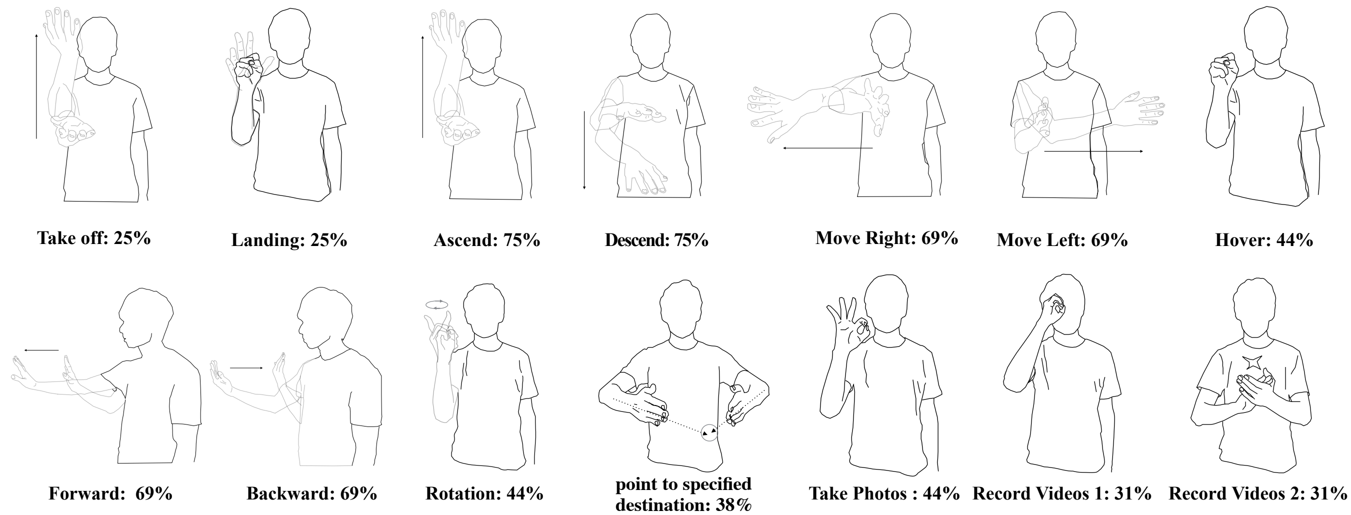

Figure 3. The user-defined gesture of rank 1 in user study 2.

references cited in this paper are included for illustrative purposes only. **Don't forget to acknowledge funding sources as well**, so you don't wind up having to correct it later.

FUTURE WORK

The study of this work mainly focused on gesture itself, and several most favorable gestures in each gesture sets were discovered. We plan on implementing a system that can take gestures as input and output basic control commands to a UAV. The input devices may involve Microsoft Kinect for Windows v2 or wearable devices such as a smart watch, a smart band, or smart glasses for motion capture. After the system being built, we also want to compare the completion time of the trial designed in Study 2 between gesture control and smartphone app control. We are also interested in developing features of the camera in the future so that drones can soon play a role in photography in our daily lives.

REFERENCES

1. Graether, E., and Mueller, F. Joggobot: A flying robot as jogging companion. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '12, ACM (New York, NY, USA, 2012), 1063–1066.
2. Hansen, J. P., Alapetite, A., MacKenzie, I. S., and Møllenbach, E. The use of gaze to control drones. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '14, ACM (New York, NY, USA, 2014), 27–34.
3. Pfeil, K., Koh, S. L., and LaViola, J. Exploring 3d gesture metaphors for interaction with unmanned aerial vehicles. In *Proceedings of the 2013 International Conference on Intelligent User Interfaces*, IUI '13, ACM (New York, NY, USA, 2013), 257–266.
4. Vafaei, F. Taxonomy of gestures in human computer interaction (2013).
5. Wobbrock, J. O., Morris, M. R., and Wilson, A. D. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, ACM (New York, NY, USA, 2009), 1083–1092.