2019

Traditional vs Gesture Based UAV Control

Sam J. Holley
Embry-Riddle Aeronautical University, holle710@erau.edu

Brian Sanders
Embry-Riddle Aeronautical University, sanderb7@erau.edu

Dennis Vincenzi
Embry-Riddle Aeronautical University, vincendzd@erau.edu

Yuzhong Shen
Old Dominion University, yshen@odu.edu

Scholarly Commons Citation
J. Chen (Ed.): AHFE 2018, AISC 784, pp. 15–23, 2019. https://doi.org/10.1007/978-3-319-94346-6_2

Traditional Vs Gesture Based UAV Control

Brian Sanders[1(✉)], Dennis Vincenzi[2], Sam Holley[2], and Yuzhong Shen[3]

[1] Department of Engineering and Technology, Embry-Riddle Aeronautical University, Worldwide, Daytona Beach, USA
sanderb7@erau.edu

[2] Department of Aeronautics, Graduate Studies, Embry-Riddle Aeronautical University, Worldwide, Daytona Beach, USA
{vincenzd,holle710}@erau.edu

[3] Department of Modeling, Simulation, and Visualization Engineering, Old Dominion University, Norfolk, VA, USA
YShen@odu.edu

Abstract. The purpose of this investigation was to assess user preferences for controlling an autonomous system. A comparison using a virtual environment (VE) was made between a joystick based, game controller and a gesture-based system using the leap motion controller. Command functions included basic flight maneuvers and switching between the operator and drone view. Comparisons were made between the control approaches using a representative quadcopter drone. The VE was designed to minimize the cognitive loading and focus on the flight control. It is a physics-based flight simulator built in Unity3D. Participants first spend time familiarizing themselves with the basic controls and vehicle response to command inputs. They then engaged in search missions. Data was gathered on time spent performing tasks, and post test interviews were conducted to uncover user preferences. Results indicate that while the gesture-based system has some benefits the joystick control is still preferred.

Keywords: Autonomous systems · Leap Motion Controller
Gesture-based interface

1 Introduction

The autonomous systems market continues to grow with systems (aerial and ground) ranging from those that can be held in the palm of your hand to larger, extremely capable military systems. For recreational use, and an emerging business market, the FAA requires an unenhanced line of sight to the vehicle be maintained. However, new technology, in the form of mixed reality headsets are emerging and accompanied by a plethora of gesture-based systems (see for examples, LMC, Oculus with Touch) that may enable a future change in policy and regulations. These head mounted display systems combined with gesture-based command control interfaces offer a new approach to vehicle control that is unencumbered by traditional hand held devices, and offer unique capabilities for mission planning and vehicle, even multivehicle, control.

J. Chen (Ed.): AHFE 2018, AISC 784, pp. 15–23, 2019.
https://doi.org/10.1007/978-3-319-94346-6_2

The design of these systems will require careful investigation of human factors issues to populate gesture libraries that are natural and intuitive, as well as cognitive loading considerations due to the easy availability of a vast amount of visual information. This combination of gesture libraries and cognitive loading is the focus of this research.

1.1 New Technology for Displays and Controls

The concept of a reality-virtuality continuum was first introduced by Paul Milgram in 1994. In that paper, Milgram and Kishino [1] discuss the concept of a reality-virtuality continuum with the real-world environment on one end of the continuum and a totally virtual computer-generated environment on the other end of the continuum. Between the two ends of the continuum is a wide range of mixed reality variations between total real environment and total virtual environment. Most advanced interfaces today fall somewhere in the mixed reality section of the reality-virtuality continuum.

State of the art technology began revolutionizing the Human-Machine Interface with the movement away from legacy control devices such as joysticks and other physical controls, toward more innovative interface technologies such as touchscreens, virtual reality displays (VR), augmented reality displays (AR), and mixed reality displays (MR). VR displays or VR-like displays are now affordable and commonplace, and are regularly used as the display of choice when immersion into a 3D environment is preferred. See-through displays offer the ability to create AR or MR displays by allowing the real world to be viewed through a see-through display that can be used to augment the real world with additional information. High resolution displays on phones or inexpensive Helmet Mounted Displays (HMDs) such as the Oculus Rift have begun to replace and augment visual systems to develop environments that serve as displays for vehicle parameters as well as provide an egocentric view from the UAS camera [2, 3].

The creation of a MR display that integrates both the real world and computer-generated world into a display that uses relevant parts of both environments is meant to create a display that is more effective than either one by itself. The goal of this MR type display would be to integrate relevant portions of both the real world and virtual world to produce a display that is efficient, functional, user friendly, and (hopefully) intuitive in design. The capability exists for technology to provide more information with realistic visual perspectives similar to looking through a Heads-Up Display (HUD) on a manned aircraft. Utilization of these types of technologies, if designed correctly, can result in a more realistic visual display that provides the information needed for successful operation with minimal training requirements [2].

Although VR and MR type displays have been available since the 1980s and 1990s, MR interfaces that include control components have not. Interactive, MR display and control interfaces have only recently appeared on the consumer market in a usable and affordable form. Typically, a combination of technology can be integrated and utilized to create an inclusive human-machine interface that can be used to both display information in a VR environment while designing a control interface which can be used to manipulate objects in the real or virtual worlds. This combination of technology provides the means to design a MR display and a VR control interface for use in a real or virtual world. Inexpensive hardware used in this experiment include the Leap Motion

Controller (LMC) and Oculus Rift Goggles. The LMC generates the operational gesture recognition environment while the Oculus Rift Goggles provides an immersive or see-through information display environment.

1.2 Cognitive Loading

During the early stages of evolution for gesture-based control applications, cognitive processing and loading issues [4] are introduced as factors that will be incorporated gradually as the study advances. For the present effort, conditions aligned with aerodynamic factors and perception in contrasting environments will be observed. Initially, attention is directed to features of direct viewing in natural settings compared with views from a tablet displaying virtual images and information. This is particularly apparent when switching views between operator real-world view and a virtual framework. Recent evidence indicates that very different brain processes are involved in comprehending meaning from these sources [5], in particular competition among select hippocampal neurons, suggesting accelerated depletion of neural resources under high task loadings [6]. As stated earlier, anecdotal observations will accompany the tasks performed in the current trials to identify particular variables for study later.

As elements and workload increase in number and complexity, task loading will follow. The influences of channelized attention, sensory cues, and loss of energy state awareness [7] are potential targets for further study. In turn, influences on cognitive mapping [8] will be investigated to elaborate effects on sustained attention, conflict resolution, and rapid updating of working memory [9]. Another element for later focus in the current line of research is to evaluate perceptual and cognitive processing features when using a three-dimensional virtual image with an environment that entails moving units or elements. As might be expected, cognitive loading considerations will likely increase [10] in presence and influence. In the intermediate stages of this research trajectory, when operators will use a heads-up display, combined effects of real-world and virtual cognitive processing will be an intense area of inquiry, including consequences of refractory periods and protein cycling limits in memory buffering [11] as they affect vigilance and judgment.

The above discussion highlights some of the opportunities and considerations for the control of autonomous vehicles using head mounted and gesture-based systems. Previous work by the authors involved the development of a gesture-based library for control of a drone [2]. In that paper, 11 basic control functions were identified along with the associated hand gestures to achieve the desired response. This work builds upon that by examining user's ability to control a representative drone in a virtual environment (VE).

2 Virtual Environment and Controllers

The control approaches were compared using a virtual environment (VE) created in Unity3D (™). A partial shot of the VE is shown in Fig. 1 below. Its overall purpose was to enable a scenario that included some depth and scale perception. The drone is a generic representation of a recreational quadcopter drone. It models a 1 kg drone with nominal dimensions of 30 cm × 30 cm × 10 cm and has red lights indicating the forward part of the drone and blinking green lights in the rear of the unit. The VE is of a generic rolling hills environment that contains some natural environment features such as trees, a lake, and several man-made features such as a jeep, tents, and recreational vehicles. The latter were used in the second part of the test when the participant was asked to locate and proceed to one of these targets.

Fig. 1. Virtual environment (probably need one with the Jeep in it or some other target)

The screen contained features to inform the participant about the status of the vehicle. The red letters are a Heads Up Display that provides information on the altitude, speed, and distance to the vehicle. At this point the test is done in screen space mode, but as the research continues and VR and Mixed Reality Goggles are incorporated into the configuration, these features should transition nicely. There is also a directional arrow in the left-hand corner that indicates the direction of the vehicle. It points in the forward direction (i.e., red lights). The green box is an indicator for when the vehicle responds to the virtual hands. Outside of this area other commands can be given, such as turning the vehicle camera on and off.

The commands are given by either a traditional joystick-based controller, or gestures captured by a leap motion controller. The joystick is shown in Fig. 2 and is a typical Xbox 360 controller. The left stick controls the rotation of the vehicle and the fore/aft translation (Fore/aft translation referring to direction of the lights). The right stick controls the altitude and the left/right translation. Finally, the A button changes the perspective from the operator to a view from the drone camera and the B button resets the vehicle to the starting position.

The gesture commands were captured using a leap motion controller (LMC). Figure 3 shows the leap motion sensor and its coordinate system. It is a right-handed coordinate system, as opposed to Unity3D which uses a left-handed coordinate system. The LMC is able to capture hand motions with a sub-millimeter accuracy (Weichert et al.) [12]).

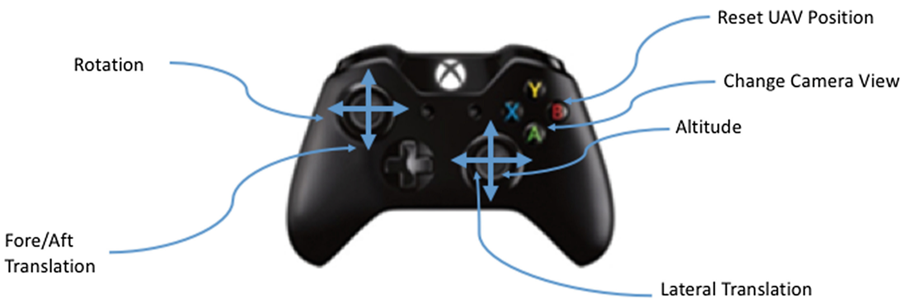

Fig. 2. Joystick controller

Multiple hand models come with the LMC. These range from basic hand models shown in Fig. 4 below to robotic looking hands to humanoid looking hands. For this investigation the user could roll the left hand about the z-axis for the vehicle to yaw and pitch it about the x-axis for a forward and aft motion. A rotation of the right hand around the z-axis results in a left/right translation of the vehicle while a right hand rotation around the x-axis results in a change in altitude. One final capability was via the use of the LMC based user interface to control the camera. It is a slick capability available in the Orion Version of the API [13]. In this case a UI is attached to the left hand and is visible when that hand is rotated toward the user as shown in Fig. 5 below. In this case the interface enabled the view selection from either the view of the operator or the view from the camera. The view from the camera assists in identifying targets.

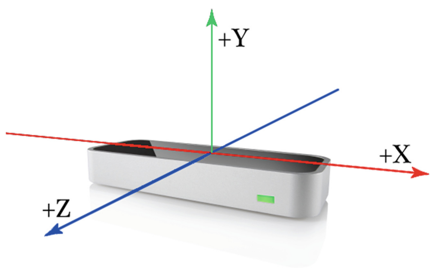

Fig. 3. Leap motion controller and coordinate system [13]

3 Methodology

The intent of this phase of the research was to determine a participant's control approach preference and what led to that decision. To help the user make an informed decision, each participant was exposed to a play environment and an operational environment. No tasks were assigned to the user in the play environment. The play environment was meant to allow the participant to become familiar with the basic controls and sensitivity of the vehicle while using either the joystick or gesture-based command inputs. While it did contain several of the elements to help with scale and

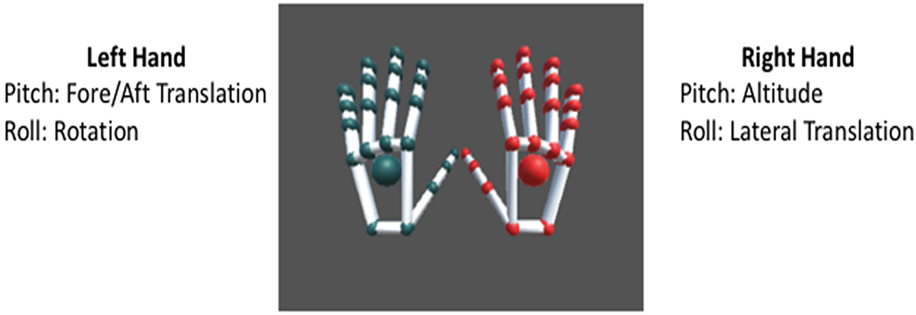

Fig. 4. Leap motion UI for controlling camera view

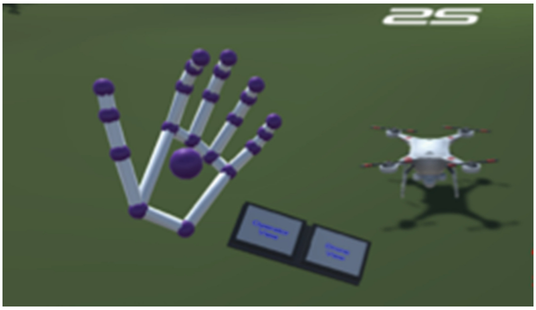

Fig. 5. Leap motion UI for controlling camera view

depth perception, it did not contain any of the targets mentioned above. In the operational environment users were tasked to find a target (i.e., ground of tents), proceed to that target, and then land in an identified landing zone. The landing zone had a 10 m radius, so it was quite large compared to the drone.

This approach provided some insight into how well a user could perform a task without being overly burdensome by asking them to fly specific flight plans. The flying of precision flight paths will be reserved for future investigations where assessment of accuracy and performance are needed. The qualitative data gathering from this was the amount of time spent in play and the amount of time to complete the assigned task in the operational environment. Post test interviews consisted of a series of six questions measured on a Likert scale and constructed to uncover the perceived suitability and advantages of the gesture-based control approach, as well as other features the simulation should include to enhance usability and functionality of the interface.

4 Results and Discussion

Two sets of tests were conducted to date. The first set of tests included two participants that were recreational drone pilots. The purpose was to provide feedback on the basic virtual environment requirements. For this early work it was desirable to keep the cognitive loading to a minimum, so the operator could focus on the control system, and the feedback indicated the fundamental data related to vehicle orientation and scaling references that need to be included. The two participants help guide what basic information needed to be included which was: textural information related to altitude, speed, and slant range between the operator and the drone. They also suggested that some objects be placed in the scene to help process scaling issues related to distance. The final scene represented a camping area at a lake with recreational vehicles at one location and a group of tents at the other.

The second set of tests included four participants. The purpose of these tests was to provide feedback on the two control system approaches. As mentioned above each participant engaged in four scenarios. Two in play (one with the Joystick and one with the LMC system) and similarly two in a search mission, such as looking for and traveling to the RV park. On average twice as much time (11 min vs 22 min) was spent in play mode with the gesture-based system. Users were able to quickly feel comfortable with the joystick approach. On the other hand, while the controllability significantly improved from the start the users still did not feel as comfortable at the end of the play session with LMC system as compared to the joystick. Likewise, mission times for the joystick were on the order of 3 min while the missions for the LMC were rarely completed due to fatigue and frustration with the system. This suggests that a change in test approach may be required. For example, as opposed to running a participant through all four scenarios maybe just have them focus on one or the other control system.

Total test time and post interview per participant took just under an hour, and the participants reported feeling fatigued at the end, mostly due to using the LMC system. It is thought that this is due to the hand gestures requiring more energy compared to the finger motion that can be used with the joystick. Even though it was more energy the users like that it made them feel more connected to the vehicle response. Also, use of hand gestures is a newer approach so the cognitive load was most likely higher since they were processing more (i.e., the visuals of the hands) visual information and correlating the vehicle response to their inputs. The users also commented that they preferred the joystick for making small command inputs. This notion is consistent with what is reported by Weichert et al. [12] when they compared the accuracy of the mouse to that of the LMC. The LMC is highly accurate when it comes to detecting the hand placement and motion, but the challenge is transforming that information into precise control such as the vehicle in this case. Achieving the same control precision with hand motions may require some additional filtering of the hand movements translated to command actions. The addition of filtering algorithms, automation, or some form of artificial intelligence to dampen, regulate, adjust, or interpret gestures intended to control a vehicle is something to be considered for future iterations.

For the most part the visual content was satisfactory for the participants. The location of the textural information was enough, and the users' responses did not indicate they were overly taxing of information processing. In fact, they were typically so focused on the vehicle that they needed to be told this information was available. On the other hand, the virtual hands were distracting. This concern was alleviated by making them smaller. It still provided a point of reference, but it could be accessed when needed rather than constantly in the visual processing path. Thus, the larger hands were a distraction and loaded up the visual processing apparatus. While the smaller hands still provided an orientation but did not seem to distract attention.

5 Summary and Conclusions

The promise of emerging technology for the design of new control environments for autonomous systems is alluring. These new systems will enable a departure from using highly accurate (i.e., joystick), low information visual systems to less accurate but more flexible gesture-based systems accompanied by a visual system that can rapidly provide vast amounts of information. In the present case, the LMC itself is highly accurate, but this particular application design has not transferred that accuracy to a system control approach yet.

References

1. Milgram, P., Kishino, F.A.: Taxonomy of mixed reality visual displays. IEICE Trans. Inf. Syst. **E77-D**(12), 1321–1329 (1994)
2. Sanders, B., Vincenzi, D., Shen, Y.: Gesture based UAV control. In: Proceeding from the 8th International Conference on Applied Human Factors and Ergonomics, Los Angeles, CA, 17–21 July 2017
3. Chandarana, M., Meszaros, E., Trujillo, A., Allen, B.: Challenges of using Mulitmodal HMI for unmanned mission planning. In: Proceeding from the 8th International Conference on Applied Human Factors and Ergonomics, Los Angeles, CA, 17–21 July 2017
4. Antonenko, P., Paas, F., Grabner, R., van Gog, T.: Using electroencephalography to measure cognitive load. Ed. Psychol. Rev. **22**(4), 425–438 (2010)
5. Ravassard, P., Kees, A., Willers, B., Ho, D., Aharoni, D., Cushman, J., Aghajan, Z., Mehta, M.: Multisensory control of hippocampal spatiotemporal selectivity. Science **340**(6138), 1342–1346 (2013)
6. Ungar, M.: The social ecology of resilience: addressing contextual and cultural ambiguity of a nascent construct. Am. J. Orthopsychiatr. **81**(1), 1–17 (2011)
7. Dodd, S., Lancaster, J., Miranda, A., Grothe, S., DeMers, B., Rogers, B.: Touch screens on the flight deck: the impact of touch target size, spacing, touch technology and turbulence on pilot performance. In: Proceedings of 58th Annual Meeting of the Human Factors and Ergonomics Society, pp. 6–10, Chicago (2014)
8. Gevins, A., Smith, M.E., Leong, H., McEvoy, L., Whitfield, S., Du, R., Rush, G.: Monitoring working memory load during computer-based tasks with eeg pattern recognition methods. Hum. Factors **40**(1), 79–91 (1998)
9. Bowers, M., Christensen, J., Eggemeier, F.: The effects of workload transitions in a multitasking environment. In: Proceedings of 58th Annual Meeting of the Human Factors and Ergonomics Society, pp. 220–228, Chicago (2014)

10. Rorie, R., Fern, L.: UAS measured response: the effect of GCS control mode interfaces on pilot ability to comply with ATC slearances. In: Proceedings of 58th Annual Meeting of the Human Factors and Ergonomics Society, pp. 64–68, Chicago (2014)
11. Cottrell, J., Levenson, J., Kim, S., Gibson, H., Richardson, K., Sivula, M., Li, B., Ashford, C., Heindl, K., Babcock, R., Rose, D., Hempel, C., Wiig, K., Laeng, P., Levin, M., Ryan, T., Gerber, D.: Working memory impairment in calcineurin knock-out mice is associated with alterations in synaptic vesicle cycling and disruption of high-frequency synaptic and network activity in prefrontal cortex. J. Neurosci. **33**(27), 10938–10949 (2013)
12. Weichert, F., Bachmann, D., Rudak, B., Fisseler, D.: Analysis of the accuracy and robustness of the leap motion controller. Sensors **13**(5), 6380–6393 (2013). https://doi.org/10.3390/s130506380
13. Leap motion developer: https://developer.leapmotion.com/. Accessed 2 March 2017