GAZE, STEERING, AND ACTIVE VISION DURING SKILLED QUADCOPTER FLIGHT

Nathaniel V Powell

Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Approved by:
Dr. Brett Fajen, Chair
Dr. Chris Sims
Dr. Wayne Gray
Dr. Oliver Layton

Department of Cognitive Science
Rensselaer Polytechnic Institute
Troy, New York

[August 2022]
Submitted July 2022
# TABLE OF CONTENTS

LIST OF FIGURES .............................................................................................................................. v
ABSTRACT........................................................................................................................................ viii

1. VISION-BASED CONTROL, GAZE, AND QUADCOPTER FLIGHT .............................................. 1
   1.1 CHAPTER OVERVIEW .................................................................................................................. 1
   1.2 SOURCES OF VISUAL INFORMATION, GAZE, AND GUIDING LOCUMOTION .................................. 3
      1.2.1 OPTIC FLOW GUIDES LOCOMOTION ................................................................................... 4
      1.2.2 DRIVERS USE SPECIFIC SOURCES OF VISUAL INFORMATION TO STEER ...................... 8
      1.2.3 THE ROLE OF TASK, CONTEXT, AND UNCERTAINTY ON HUMAN GAZE ............................... 12
   1.3 GAZE AND STEERING DURING QUADCOPTER FLIGHT ....................................................... 15
      1.3.1 DIFFERENCES BETWEEN CONTROLLING AUTOMOBILES AND QUADCOPTERS ............. 19
   1.4 THESIS OVERVIEW .................................................................................................................. 22

2. A PARADIGM FOR STUDYING THE GAZE AND STEERING BEHAVIOR OF QUADCOPTER PILOTS ............................................................................................................................... 23
   2.1 HARDWARE ................................................................................................................................ 24
   2.2 VIRTUAL ENVIRONMENT .............................................................................................................. 26
   2.3 PERFORMANCE AND OPTIMIZATION ......................................................................................... 27
   2.4 EYE TRACKING QUALITY CONTROL ............................................................................................. 28
   2.5 DATA RECORDING AND POST-PROCESSING ............................................................................... 29
   2.6 FUTURE USE OF THIS EXPERIMENTAL PARADIGM .............................................................. 31

3. EXPERIMENT 1: COORDINATION OF GAZE AND STEERING DURING QUADCOPTER FLIGHT ................................................................................................................................. 33
   3.1 METHODS ................................................................................................................................... 36
      3.1.1 PARTICIPANTS .......................................................................................................................... 36
      3.1.2 APPARATUS AND VIRTUAL ENVIRONMENT .......................................................................... 38
      3.1.3 PROCEDURE AND DESIGN ................................................................................................... 39
      3.1.4 DATA POST-PROCESSING AND ANALYSES ........................................................................ 41
         3.1.4.1 EYE TRACKING DATA .................................................................................................... 41
         3.1.4.2 GAZE AND STEERING TIME-SERIES ............................................................................ 42
   3.2 RESULTS AND DISCUSSION ......................................................................................................... 44
      3.2.1 GENERAL MEASURES OF PERFORMANCE ............................................................................. 44
      3.2.2 GAZE BEHAVIOR .................................................................................................................... 47
      3.2.3 DO PILOTS LOOK IN THE DIRECTION TOWARD WHICH THEY INTEND TO STEER? .......... 53
         3.2.1 EVIDENCE SUGGESTS ANTICIPATORY EYE MOVEMENTS AND STEERING .................. 55
LIST OF FIGURES

Figure 1: Example of a quadcopter flight path (main panel). Quadcopter point of view showing fixation on future trajectory cannot be maintained (subpanel). .............................. 17

Figure 2: (A) Empty ground template with path painted over natural features. (B) Template with random tree placement. .......................................................... 25

Figure 3: Visualization and post-processing tool. (A) drone/quadcopter data can be viewed in playback (B) The visualization can be adjusted by changing camera point of view (C) The data being explored can easily be changed along with the experimental condition cooresponding to that data (D) The gaze location in screen coordinates can be visualized. Each crosshair represents the left eye (blue) the right eye (yellow) and the cyclopian eye (green) (E) Lower-level data about gaze behavior can be viewed in playback (F) A image exportation tool that will genereate a series of images that can be used to create detailed videos of the quadcopter flight and gaze behavior. .......................... 29

Figure 4: The left panel shows the first-person point of view seen by the pilot with gaze overlay and object highlighted. The right panel shows the top-down view with gaze vector cast into world space. ................................................................. 30

Figure 5: (A) Contains the setup of the experiment with user controling the drone and eye-tracking software running. (B) Shows the mounted pupil-labs eye-tracker in a vive pro HMD. (C) Shows the taranis drone controller used. (D) Shows the path outline that participants needed to negotiate through the racecourse. .................................................. 37

Figure 6: The experimental design. Each block contained two laps of each condition. subsequent blocks alternated the ordering of the conditions, with the path and hoops condition always being second. ................................................................. 39

Figure 7: An example of each of the experimental conditions. ................................................................. 40

Figure 8: Top-down point-of-view of the angles used for cross correlation analysis. Green is the angle between gaze vector and heading vector, purple is the angle between the trust vector and heading vector, and blue is the angle between a vector representing the orientation of the forward-facing camera and the heading vector............................. 43

Figure 9: General performance measures across conditions and blocks. ................................. 44

Figure 10: Distribution of gaze categories by condition and block. All proportions were calculated by lap and then averaged across participants. ................................. 46

Figure 11: (Top Panel) The overall distribution across conditions. The yellow line is the mean and the red lines are one standard deviation. (Bottom Row) Probability Distribution of gaze distance by condition. ................................................................. 48
Figure 12: Measures of gaze behavior relative to hoops across hoop-containing conditions. ..... 49

Figure 13: The ratio between how often hoop centers and edges fell under the gaze crosshair and the number of possible instances where gaze could have fell towards those objects. The black line shows the average between the four subjects, and the bottom graph shows the density of inter-hoop segment lengths mirrored over 0. The hoop combined ratio plot is the sum of the two plots above it. The vertical black line is the mean and the vertical red line is the median. Four subjects were chosen for this analysis due to the accuracy of their eye tracking data. The proportions don’t add up to one right at the hoop due to instances where participants missed the hoop or had the vehicle not oriented directly through the hoop........................................................................................................ 50

Figure 14: Cross-correlation of thrust and gaze angles broken down by lag and correlation. The blue star corresponds to the results found by Pfieffer & Scaramuzza (2021). The top and side panels illustrate the correlation densities of the lag and correlation coefficients. ........................................................................................................ 53

Figure 15: Diagram of the variables that make up the linear models. The relationship of hoop N and N+1 predicts how pilots fly through hoop N. The relationship between flying through N-1 relative to hoop N were also calculated as covariates. ν is the angular offset between hoops N and N+1, and γ is the relative orientation between hoops N and N+1. ........................................................................................................ 56

Figure 16: (Top) Partial r-squared for coefficients predicting approach angle at hoop N. (Middle) partial r-squared for coefficients predicting camera angle at hoop N. (Bottom) partial r-squared for coefficients predicting thrust angle at hoop N. Each point is the partial r-squared for a single linear model (a total of 30 models). ........................................ 58

Figure 17: Distribution of the angle between camera and heading across participants and broken down by condition. ........................................................................................................ 60

Figure 18: Example images from each of the conditions. ......................................................... 66

Figure 19: Block design for experiment 2. A total of 5 blocks existed, each containing an instance of 2 laps in each condition........................................................................................................ 67

Figure 20: Performance measures across conditions (Left) and blocks (Right). ....................... 69

Figure 21: Average gaze distributions to different object categories across conditions............ 70

Figure 22: (Left) Overall gaze distance probability densities. (Middle and Right) Gaze distance densities across both conditions. Black lines represent densities for one block from each participant. The yellow line is the overall mean and the red lines are one standard deviation........................................................................................................ 72
Figure 23: (Top) Average gaze distance to points along the path. (Bottom) Average time to reach points gazed along the path.
ABSTRACT

Previous accounts of how humans locomote have focused on movements and perceptual judgments along the 2D ground plane, such as when driving automobiles or walking over various types of terrain. Humans, however, can also locomote in very different conditions as evidenced by the flight of experienced quadcopter pilots as they fly through cluttered three-dimensional environments. In this thesis, I investigated how quadcopter pilots coordinate gaze with steering in a variety of environmental conditions. Chapter 1 reviews previous accounts of visually guided locomotion in humans and discusses whether such accounts may be able to explain the gaze and steering behavior of quadcopter pilots. The second chapter outlines a novel experimental paradigm designed to investigate how quadcopter pilots coordinate gaze with steering in a virtual reality (VR) environment. Chapters 3 and 4 detail the results from two experiments. There are three main takeaways from the present set of experiments. First, coordination of gaze and steering during quadcopter flight share some similarities with gaze during other forms of locomotion. During a purely path following task, participants spent a large portion of time orienting gaze towards the ground over which they would soon pass. However, in Experiment 1, the presence of hoops resulted in gaze predominantly being oriented through the center of those hoops. Second, pilots adapt their current trajectory in accordance with upcoming environmental conditions, suggesting anticipatory steering through the nearest hoop to align themselves with the subsequent. Lastly, Experiment 2 showed that even when obstacles lie directly on or close to the path participant gaze was still oriented towards the path outline or the terrain near the path. The findings from both studies provide evidence in support of a predominant gaze strategy where participants look towards their desired direction of steering, with little time spent orienting gaze.
towards obstacles or other objects. The last chapter is a general discussion, outlining the limitations and future directions of the current study.
1. VISION-BASED CONTROL, GAZE, AND QUADCOPTER FLIGHT

1.1 CHAPTER OVERVIEW

Humans rely on their visual capabilities to guide themselves safely through the world. When driving a car down a busy highway or when hiking over rocky, uneven mountain trails, humans must sample the available visual information using well-placed eye movements to select appropriate actions for locomotion. The use of vision to control action has been studied in a wide variety of domains, from expert performance in sports (Land & McLeod, 2000; Williams, Davids, Burwitz, & Williams, 1994; Hayhoe, McKinney, Chajka, & Pelz, 2012) to the regulation of flight patterns in flying insects (Srinivasan, Zhang, Lehrer & Collett, 1996; Hateren & Schilstra 1999). Studying the ways that humans use vision to guide their actions is important not only for improving our understanding of human perception and motor behavior, but also for improving algorithms for the control of autonomous vehicles and robotics. The creation of such algorithms makes it possible for researchers and engineers to develop systems that behave like real biological systems who have evolved efficient solutions to vision-based control problems.

In this chapter, I review background literature relevant to a series of experiments designed to expand our understanding of how humans coordinate their gaze and steering while navigating through cluttered environments. The experimental setup involves a novel virtual reality eye-tracking paradigm that records the actions and gaze behavior of human participants as they control a small quadcopter through a visually realistic virtual environment. At the time of this writing, there have been few previous attempts to explain the ways in which humans’ sample and use visual information to guide their movements within three-dimensional spaces. Rather, the
past research on gaze behavior in pilots has focused primarily on human factors issues in aviation, such as how eye movements can be a predictor of pilot fatigue (Peißl, Wickens, & Baruah, 2018). Questions about perception and control while flying multirotor-based aerial vehicles have largely been neglected, only recently having been explored by researchers (Pfeiffer & Scaramuzza, 2021).

This chapter is broken up into two main sections (Sections 1.2 and 1.3) that are followed by a brief section (Section 1.4) previewing the remaining chapters of this thesis. The next section (1.2) first reviews the past research on how humans actively generate eye movements to sample visual information needed to guide locomotion and steer along a 2D ground plane. Previous accounts of visually guided steering predict that specific sources of visual information are sampled via eye movements so that an actor may produce appropriate steering adjustments. For example, Lappi and Mole (2018) provide a theoretical account of why an actor makes eye movements to points along their future path to guide steering. An actor’s future path is defined as the trajectory along which an actor will traverse if their current path curvature remains unchanged. A future path sampling account of steering may not generalize to control under conditions in the 3D domain. The questions then emerge: How are humans able to control their actions primarily using vision when movement is not restricted to a 2D ground plane? What visual information allows them to efficiently guide their movement? What kinds of gaze and control strategies are they using to perform this task? Many previous investigations of gaze and steering have focused on controlling an automobile. Therefore, quadcopter flight highlights a potential gap in our understanding of visually controlled locomotion as it is very different from automobile driving and offers a novel task space by which researchers can investigate how humans use visual information to guide movement in cluttered environments.
In Section 1.3, I qualitatively describe quadcopter flight and the visual information that may be available while steering a vehicle through three dimensional spaces. Based on the description of flight I explore how current theoretical accounts of visually guided locomotion may inform our understanding of how quadcopter pilots use gaze to sample sources of visual information to guide steering. Specifically, I discuss the ways in which quadcopter flight is both similar and different to the visual conditions experienced when driving automobiles.

**1.2 SOURCES OF VISUAL INFORMATION, GAZE, AND GUIDING LOCOMOTION**

To complete tasks as seemingly simple as walking down an empty sidewalk or as complicated as driving down a highway at rush hour, humans must select and execute actions by moving through their environment in accordance with the available visual information. The visual information that is available during such tasks, however, is not constant. Instead, humans must strategically place eye movements that sample task-relevant information when and where it is available. Researchers have found evidence of several qualities of human eye movements as they complete sensorimotor tasks. Lappi (2016) reviews these qualities and summarizes that human gaze is often repeatable, focused on task-relevant information, intermittent, based on spatial memory, and is always coordinated with actions of the body.

One theoretical framework that has been developed as an attempt to unify our understanding of how eye movements shape human behavior and perception is active vision (Bajcsy, 1988; Bajcsy, Aloimonos, & Tsotsos, 2018). Proponents of active vision (or more broadly speaking, active perception) argue that eye, head, and body movements can all be critical for generating information that is needed to maintain robust perceptual contact with the environment. The active generation of visual information through strategically placed eye movements has been
explored in many different task spaces, such as walking over complex terrain (Matthis, Muller, Bonnen, & Hayhoe, 2022) and driving automobiles (Lappi & Mole, 2018). However, our understanding of how humans coordinate their gaze with locomotion, what visual information is sampled, and when it is sampled is still an ongoing topic of investigation.

1.2.1 OPTIC FLOW GUIDES LOCOMOTION

Optic flow is one source of visual information that humans can use to perceive their self-motion and guide themselves through their environment. It is defined as the changing pattern of motion generated by movement through space or by externally moving objects (Gibson, 1950). For example, while driving an automobile forward down an empty street a pattern of radially expanding motion is generated and can be observed by the forward-facing driver. Since Gibson’s initial hypothesis that humans use optic flow to control their actions, there have been numerous studies on how optic flow can be used to guide actors through the world. In one such study, Warren, et al., (2001) provided evidence in support of Gibson’s hypothesis and found that humans use optic flow to steer their body towards goal locations. They predicted that if optic flow is being used to control movement, then participants will align the focus of expansion (FOE) with their goal location by adjusting their movement to cancel the error between the perceived direction of heading and the goal location. This strategy contrasts with an egocentric-direction strategy, which predicts that an observer will align their locomotor axis with the perceived direction of goal location (Rushton, Lloyd, Harris, & Wann, 1998). Warren’s experiment tested sensitivity to optic flow by using a virtual reality manipulation that displaced the FOE 10 degrees to the left or right of the direction a participant was walking. In conditions where optic flow was not available to participants, the observed behavior was consistent with the egocentric direction hypothesis. However, as more structure was added to the scene, and hence,
more optic flow made available, participants acted more in accordance with an optic flow-based guidance strategy. This suggests that humans rely on optic flow to steer towards goal locations when it is available (Warren et al, 2001).

The optic flow field is often made more complex by rotations from the body, head, or eyes, which add rotational components to the translational flow vectors. In the case of rotational components produced by eye movements, the optic flow being sensed in the eye-centered frame of reference is called retinal flow. During many types of motion, the introduction of rotational components to the optic array makes perceiving the direction of self-motion more complicated than determining the direction of the FOE. For example, during movement along a curved path there is no FOE defining the actor’s instantaneous heading.

Predicting the type of body movement that produced certain patterns of flow is not always a trivial task. For example, the instantaneous retinal flow generated from moving along a straight path while rotating about the vertical axis through the eyes and the retinal flow from moving along a curved path can be identical (Warren et al., 1991; Royden, 1994). Due to the potential ambiguity in the flow field, some researchers have argued that the visual system must remove or account for these added rotational components to recover the components due to pure translation and maintain an accurate perception of self-motion. Banks, Ehrlich, Backus, and Crowell (1996) tested the influence of real versus simulated rotation resulting from eye movements on an actor’s ability to judge their instantaneous heading, which they defined as the direction of translation. They found that heading estimates were robust to the added optic flow components due to rotation if the eye movements were real rather than simulated (Banks et al, 1996).

An actor’s ability, or lack-thereof, to accurately perceive his or her direction of self-motion during the simulated eye rotation condition in Banks et al. (1996) may suggest a role for extra-
retinal and non-visual information when perceiving heading. For example, if extra-retinal information is necessary for the perception of heading, then during real eye movements the motor signal provided by the rotation of the eye can be informative about the observer’s rotation and motion. When the rotation is simulated, however, the fixed position of the eye relative to the motion of the observer will not generate any signal about rotation. Li and Warren (2000) tested this hypothesis further and showed that during simulated rotation an actor’s judgement of heading remains unbiased if the environment contains dense texture and enough depth variation to generate motion parallax. The findings from this study showed that heading can be perceived during simulated rotation relying only on vision in natural conditions.

Evidence has also shown that humans can account for the added rotational components in the optic flow field through active gaze strategies, simplifying retinal flow instead of making it more complex. This view contends that in some situations the brain does not need to remove the added rotational components in the optic flow field to accurately perceive self-motion and other aspects of the environment. Glennerster, Hansard, and Fitzgibbon (2001) showed that the retinal flow that follows from a series of fixations helps simplify the information that is available and can be used to recover the direction of heading. Similarly, Matthis, et al., (2022) showed that while locomoting over complex terrain, humans make eye movements that stabilizes retinal flow. The pendular motion from walking generates rotational components along the horizontal and vertical axes in the optic flow field that seems to complicate the perception of heading if one assumes that heading perception is based on the head-centered global radial optic flow pattern. Their findings suggest that human gaze generates a rich source of visual information in the retinal flow field that can aid natural locomotion. One source of visual information that Matthis and colleagues argue in favor of is the curl of flow at the fovea. Curl of the optic flow is defined as
how spiral the flow is relative to the point of gaze. If an observer fixates a point along their current movement trajectory, then the retinal flow will have no rotation. Maintaining a fixation off the current movement trajectory, however, will result in foveal curl that observers could use to determine their trajectory relative to that point of fixation. For example, while fixating a specific reference object, an actor could then turn their body until there is zero foveal curl, orienting them towards some desirable location. The foveal curl, therefore, behaves systematically with relation to an observer’s current movement trajectory and gaze position. The curl of optic flow in the fovea is therefore another potential source of visual information that humans could use to guide their locomotion while walking. However, more systematic manipulation on the curl of flow while walking is necessary to determine how and if it influences locomotor behavior.

Human gaze has also been shown to be correlated with the direction towards which an actor is steering (Wann & Swapp, 2000; Wilkie & Wann, 2003). Wann and Swapp (2000) provided an account of visually guided locomotion that shows how maintaining fixation on a point in the environment and following the optic flow components along the ground plane that define future path can lead an actor to their desired point relative to the fixated reference point. Their account states that an actor can steer towards a target using retinal flow by fixating a feature in the environment close to their future path, after which they can reorient their trajectory by steering so that the flow lines become straight. Additionally, Wilkie, Kountouriotis, Merat, and Wann (2010) provide further evidence that human gaze is predominantly oriented towards their desired path. They showed that driving along a curved path at high speeds requires fixations in the direction that an actor wants to steer rather than where an actor is currently steering.
Contrary to the studies on the importance of optic flow specifying future path, there is also evidence that perceiving instantaneous heading based on the expanding radial pattern of optic flow, and not future path, is more important for steering towards a goal (Li & Cheng, 2011). In one study, participants were presented with visual information where egocentric direction was unavailable for steering, forcing them to rely on optic flow (Li & Cheng, 2011). The authors found that participants steered to align their instantaneous heading with their current goal and found no evidence that information specifying future path was used to steer. Similarly, Li, Sweet, and Stone (2006) found that humans can perceive their heading when future path information is not directly accessible during a steering adjustment task. Evidence also shows that in the presence of a rich optic flow field future path information does not influence of the perception of heading. However, in the presence of a sparse optic flow field the perception of heading is aided by future path information. These findings show that humans can perceive their instantaneous heading even when future path information is unavailable (Li, Chen, & Peng, 2009). Whether humans actively use path information for perceiving heading and other visual variables needed for steering in natural conditions is still debated.

1.2.2 DRIVERS USE SPECIFIC SOURCES OF VISUAL INFORMATION TO STEER

Decades of research on steering automobiles has resulted in two dominant accounts of how drivers coordinate gaze and steering: the travel-point account and waypoint account. Both accounts predict drivers will produce gaze patterns to specific sources of visual information to make current steering corrections and plan for upcoming controls. Travel-point accounts predict that driver fixations are made to points in the environment that remain fixed in the egocentric (i.e., human-centered) frame of reference (Lappi, 2016; Tuhkanen et al., 2019; Lappi et al.,
For example, one travel-point account, the tangent point hypothesis, predicts that drivers will fixate the tangent point of a curve, or the point at which curve reverses direction in the visual field (Land & Lee, 1994). Land and Lee (1994) showed that curvature can be determined from information specifying the visual direction of the tangent point and the distance from the inside lane edge, and that such information can be used to make steering adjustments along curved paths. Further work showed that similar behavior could be produced by maintaining a constant visual direction of the tangent point while steering (Wann & Land, 2000).

Land and Horwood (1995) found further evidence for the use of travel-point information, and other researchers have developed computational models that rely on near- and far-point information to predict driver behavior. For example, the two-point model (Salvucci & Gray, 2004) showed that near and far point information is needed to maintain stable steering under normal driving conditions. The near point information is needed to stay within a lane and maintain smooth driving, while far point information is needed to plan for upcoming road conditions. Their computational model showed how different sources of visual information can lead to stable steering trajectories like those observed in human drivers.

Early research provided evidence in support of the tangent point hypothesis (Land & Lee, 1994; Land & Tatler, 2001; Kandil, Rotter, & Lappe, 2010; Land & Horwood, 1995), but more recently, experimental manipulations of the available visual information have provided evidence that drivers may instead be using a strategy where they fixate waypoints along their future path through which they plan to steer (Lappi, Lehtonen, Pekkanen, & Itkonen, 2013; Lappi & Mole, 2018; Lappi et al, 2020). The gaze behavior of automobile drivers while steering has indicated sensitivity to information that specifies the future-path along which they will soon move. Lappi and Mole (2018) outline an integrated theoretical account of the active nature of
gaze as it relates to fixating waypoints along a driver’s future path and how it is coordinated with steering.

In contrast to the gaze strategy predicted by travel-point accounts, the waypoint fixation account predicts that eye movements are oriented towards locations that are fixed in the world frame of reference (Lappi et al, 2020). There are several components that make up the waypoint account: (1) predictive eye movements are made to waypoints along the predicted future path, (2) eye movements compensate for actor rotation, and (3) there is a functional role for eye movements directed to points in the environment approximately two seconds or further into the future, otherwise known as look-ahead fixations (LAF – see below).

Lappi et al (2020) investigated whether gaze compensates for actor rotation while steering along a curved path. They found that the horizontal velocity of driver eye movements was negatively correlated with vehicle yaw rate, suggesting that optokinetic eye movements were actively being generated to account for driver rotation in the opposite direction as they move along curved paths. The authors argued that this evidence is more consistent with a waypoint fixation gaze strategy because the smooth pursuit eye movements remain fixed in the world frame of reference, which is predicted by a waypoint tracking strategy. Another study, from Tuhkanen, Pekkanen, Lehtonen, and Lappi (2019), provided evidence for the existence of predictive eye movements to waypoints along the future path during real and simulated curved path driving. The experiment asked drivers to steer along curved paths where the waypoints were visually outlined at regular intervals along the path. On 25% of potential waypoint instances, however, the waypoint was not shown. Even on these trials participants still made saccades to the location on the ground plane that the waypoint would appear. This finding suggests that eye
movements are generated in anticipation of expected waypoints that are used to path plan and steer.

Other research has focused on the relationship of gaze switching and steering behavior. Gaze switching has been described as the patterns of changing eye movements made between different points in the visual scene during locomotion. For example, how a driver may intermittently shift gaze from a leading car to a parked car on the side of the until the parked car is passed safely. Wilkie, Wann, and Allison (2008) showed that while steering a virtual bicycle sequentially through a series of gates the timing of participant gaze shifts from the closest gate to the next gate influenced steering error and the smoothness of steering trajectories. Their experiment found that participants preferred to switch gaze from the closest gate to the next approximately 1.5 seconds before reaching the closest gate. Gaze switching earlier than 1.5 seconds before the closest gate was associated with smoother steering trajectories but increased steering error between the actor and the gate. On the other hand, gaze switching later than 1.5 seconds before the closest gate was associated with decreased steering error, but less smooth path trajectories (Wilkie, Wann, & Allison, 2008). This evidence supports the idea that an actor’s pattern of gaze, or in other words what an actor looks at and when, is tightly related to their steering.

The results from Wilkie, Wann, and Allison (2008) also provided evidence that anticipating future actions based on current visual information is important for steering. Upcoming road conditions can be sensed using sweeping eye movements that scan for the information needed to plan safe steering maneuvers. Other researchers have also shown that human eye movements serve an anticipatory function while acting under natural conditions (Lehtonen et al., 2013; Mennie, Hayhoe, & Sullivan, 2007). These kinds of eye movements are
referred to as look-ahead fixations. Drivers intermittently use LAFs to plan their upcoming actions by allowing them to briefly monitor upcoming traffic and road conditions (Lehtonen, Lappi, Kotkanen, & Summala, 2013). It has been argued that anticipatory gaze behavior may be necessary for driving because road conditions often change quickly, creating uncertainty about the future state of the environment for the driver (Mennie, Hayhoe, & Sullivan, 2007). Using LAFs drivers can monitor the area ahead of their current position and anticipate how events may unfold to make appropriate current actions. Therefore, LAFs are potentially used to look towards targets of later possible actions relative to an actor’s current state and the changing environmental conditions (Ballard, Hayhoe, & Pelz, 1995; Mennie, Hayhoe, & Sullivan, 2007). Additionally, Lappi and Mole (2018) argue that LAFs are necessary for a waypoint fixation gaze strategy and are used to anticipate a driver’s future-path relative to visual information two seconds or more into the future.

Many researchers have taken a modeling approach to explain human steering behavior. Fajen and Warren (2003) developed a dynamical model that treats goal locations as attractor points and obstacles as repeller points, which are both weighted by the actor’s distance and visual direction relative to those points. The resulting model was able to predict whether humans would steer to the left or to the right of an obstacle to reach a goal. A similar model proposed by Wilkie and Wann (2003) uses an actor’s gaze location to mark point attractor locations for steering to desired locations.

1.2.3 THE ROLE OF TASK, CONTEXT, AND UNCERTAINTY ON HUMAN GAZE

The placement of eye movements can also be linked to the structure and demands of the task being completed. For instance, gaze often departs from a current fixation to the next once
the current action is close to completed or entirely completed (Johansson, Westling, Backstrom, & Flanagan, 2001). Similar findings have been shown while completing everyday tasks like making sandwiches (Hayhoe, Shrivastrava, Mruczek, & Pelz, 2003). In their experiment, Hayhoe and colleagues outfitted participants with an eye tracker and asked them to make a peanut butter and jelly sandwich. While completing this high-level task, participants often made fixations to visual information relevant to the next subtask they needed to complete in the sandwich making process. This finding suggests that while participants are completing a current subtask (e.g., spreading the peanut butter) they are planning for the next subtask (e.g., fixating the jar of jelly) in the process of completing the overarching task of making a sandwich (Hayhoe et al, 2003).

The context in which actors find themselves has also been shown to be a predictor of where and towards what sources of visual information humans allocate their gaze. Rothkopf, Ballard, and Hayhoe (2007) showed that context matters when determining where an observer will look while performing multiple subtasks as they completed the overarching task of walking down a virtual sidewalk. Participants were asked to pick up targets (i.e., litter along the path) and avoid obstacles while following the sidewalk. The results showed that participants spend more time fixating on objects related to the immediate subgoals that they were asked to complete more often than other extraneous visual information. They also found that the context actors find themselves relative to the current visual scene could account for some of the variance in participant fixations. Their results showed that in the obstacle avoidance condition, participants still fixated on litter on the ground but less often than the information given priority (e.g., the visual information of the obstacles). From their analysis, they concluded that participant fixations to points in the visual field that are related a subtask can be explained by the proportion of visual information in the visual field relevant to that subtask. For example, fixations made to litter can
be explained by the proportion of litter in the visual field to obstacles in the visual field at any given point in the experiment. This showed that high-level goals and the context in which an actor is situated in the visual scene is predictive of where subjects would look, even when exposed to highly salient features in the visual field (Rothkopf, Ballard, & Hayhoe 2007).

The ways that humans orient their gaze have also been shown to be driven by a person’s level of perceptual uncertainty about task relevant stimuli, which often arises because of ambiguity in the sensory information provided by a stimulus. Sullivan et al. (2012) found that while driving in a virtual environment, participants tended to fixate longer on stimuli that related to the task to which they asked to give priority. Subjects were instructed to maintain a constant speed and follow a leading car but asked to pay more attention to performing one of those goals over the other (i.e., the speedometer or the leading car). Additionally, in one condition noise was added to the gas pedal signal or the visual information of the lead car, making it more difficult to obtain an accurate perception of the driver’s vehicle velocity or position of the vehicle relative to the lead car. They found that increasing the noise, or uncertainty, of the stimulus increased the amount of time fixating on that stimulus, but only when that stimulus related to the goal given priority (Sullivan et al. 2012). This finding was echoed further by Tong, Zohar, and Hayhoe (2017), who also showed that both task and uncertainty are important when avoiding obstacles and intercepting target objects while locomoting through a virtual environment. Objects that were more relevant to the task and moving (i.e., of which they were more uncertain) were attended to more frequently and for longer periods of time to resolve uncertainty about that object's motion. Another study, conducted by Dominguez-Zamora, Gunn, and Marigold (2018), found that gaze is allocated to reduce uncertainty about target location while stepping towards target points. Furthermore, this finding showed that more time was spent fixating target steps
when more precision was needed to hit that target. These studies suggest that perceptual uncertainty and task relevance are major predictors of human gaze during visually guided tasks.

The previous studies just discussed above provide us with several principles for understanding the role of visual information and gaze for the locomotor capabilities of humans. First, the timing and position of gaze is directly related to an actor’s ability to efficiently complete locomotor tasks (Lappi & Mole, 2018; Wilkie, Wann, & Allison, 2008). Second, actors often generate eye movements to sample task-relevant information that helps them maintain a robust perception of the environment. Such eye movements have been shown to reduce perceptual uncertainty (Sullivan et al, 2012) or generate useful information via retinal flow (Matthis et al, 2022). These principles are potentially useful for understanding how humans coordinate gaze and steering while acting in all three spatial dimensions, such as while flying quadcopters. In the next section, I provide a qualitative description of quadcopter flight and the resulting visual information. In the section that follows the next I describe how the patterns of gaze and steering may be both similar and different from other forms of locomotion, such as driving automobiles.

1.3 GAZE AND STEERING DURING QUADCOPTER FLIGHT

Skilled quadcopter pilots are capable of quickly flying their vehicles through cluttered 3D spaces. Videos\(^1\) of the complicated maneuvers performed by a skilled quadcopter pilot provide evidence that humans can navigate cluttered 3D spaces relying primarily on vision. It remains unclear whether their gaze and steering can be accounted for by current theoretical accounts of

---

\(^1\) This is a link to an example of complex quadcopter flight from an experienced pilot in the Rensselaer Polytechnic Institute Drone Club: https://www.youtube.com/watch?v=6ZjLPWFLix0
vision-based locomotion in humans, such as the waypoint fixating hypothesis and the role of task-related uncertainty. The patterns of gaze and steering of these pilots, therefore, provide researchers with a rich and largely untapped source of data that can help advance our understanding of visually guided locomotion in humans.

While flying a quadcopter, pilots are not constrained to actions along a 2D ground plane like when driving a car, instead being able to act freely along all axes of orientation. Evidence from studies on visually guided navigation of flying insects, unsurprisingly, shows that optic flow is important for controlling actions unbound within 3D space. Srinivasan (2011) provided evidence that bumblebees maintain constant image velocity in their ventral field for smooth landings. Bees are also sensitive to changes in optic flow speed in their periphery (i.e., asymmetrical optic flow) while maintaining constant flight speed and distance from walls (Srinivasan, 2011). Like bees, birds have also been shown to have visual systems that are sensitive to optic flow (Frost, 2010). Results from Schiffner and Srinivasan (2016) showed that birds, specifically Budgerigars, use optic flow in their periphery to maintain speed while moving down textured corridors. However, studies designed to investigate the ways in which humans perceive and act in similar situations have only recently begun in earnest.
While racing, quadcopter pilots need to quickly move from one goal location to another, such as when navigating from one hoop or flag to another as they move through a racecourse (see Figure 1). Pilots must simultaneously avoid obstacles like flags and the edges of the hoops that make up the racecourse. In these conditions it is important to select immediate actions that also plan for upcoming actions. To optimize the selection of immediate actions, it is likely necessary to visually search the visual field for relevant information pertaining to planning and anticipating those actions, such as using gaze polling strategies like those shown in Wilkie, Wann, and Allison (2008) and Lappi and Mole, (2018). Figure 1 shows the outline of a section of racecourse where the goal is to sequentially move through each of the hoops as quickly as possible. Pilots can accomplish this by strategically allocating their attentional resources to help plan and adapt their steering accordingly for immediate and future actions that are tuned to the

Figure 1: Example of a quadcopter flight path (main panel). Quadcopter point of view showing fixation on future trajectory cannot be maintained (subpanel).
dynamics of a vehicle moving through space (i.e., through experience pilots learn the physics of quadcopters, like their speed, acceleration, and momentum).

Depending on the context in which a pilot is flying, planning upcoming actions becomes more important for efficiently controlling locomotion. For instance, while flying through a racecourse with a hoop placed close to another and offset by some distance, it may be necessary to move along a nonoptimal path through the nearest hoop so that the quadcopter’s orientation as it passes through the hoop is optimal for moving to the next hoop. Path planning of this nature may be reflected by anticipatory gaze and steering behavior, such as look-ahead fixations to the future hoop. This LAF could precede a steering adjustment through the nearest hoop or goal location towards which the pilot desires to move that leaves the vehicle in an optimal location and orientation for locomotion to next goal. It is also possible that pilots are using peripheral visual information to track the future path, which would mean that look-ahead fixations may not be seen in the patterns of gaze exhibited pilots. However, it may still be reflected in their steering behavior. Additionally, it may be the case that pilots rely on their memory of the spatial layout of the racecourse to anticipate the upcoming hoop positions. Experimental manipulations would need to be created to investigate these possibilities.

At the time of this writing, and to this author’s knowledge, there has only been one previous study of how the gaze behavior of quadcopter pilots. Pfeiffer and Scaramuzza (2021) used a virtual flight simulator and gaze tracking paradigm to explore the ways quadcopter pilots coordinate gaze and steering while flying through a series of gates setup in a standard figure-eight style racecourse. The pilots flew a quadcopter through a simulated environment that was displayed on a computer monitor, and the gaze position on the monitor was recorded as they flew through the course. They conducted a series of cross-correlation analyses and found that the
angle between the direction of thrust applied to the quadcopter and its instantaneous speed was highly correlated ($r = 0.83$) with the angle between gaze and instantaneous speed with a lag of approximately 223ms. This means that the pilots made eye movements ahead of their current direction of self-motion towards where they are going to move in-the-near-future. This suggests that pilots are looking towards where they are going to move, consistent with past studies on visually guided locomotion. Their findings also showed that gaze tended to be directed towards the next gate approximately 1.5s and 16m before reaching it. They also speculate that there is an image stabilizing role for how gaze is used to guide flight. The ways in which quadcopter pilot steering is influenced by eye movements, specific sources of visual information, and task have not yet been investigated.

1.3.1 DIFFERENCES BETWEEN CONTROLLING AUTOMOBILES AND QUADCOPTERS

As noted at the beginning of this chapter, studies on the control of quadcopter flight in the field of sensorimotor control have largely been unexplored except for the recent Pfeiffer and Scaramuzza (2021) described above. That being the case, the most similar task that researchers have abundantly investigated is automobile driving. The available visual information while driving, however, may be very different from that resulting from quadcopter flight due to changes in height above the ground plane and rotations along different axes. These differences may have an influence on the gaze and steering strategies that are employed under these conditions. The prominent waypoint fixation account of gaze and steering (Lappi & Mole, 2018), for example, may not extend to movements in 3-dimensions. The waypoint fixation account would predict that an actor would make eye movements to the point(s) in the environment through which they will pass. However, while flying quadcopters, pilots might instead fixate on points that they need to
avoid, such as the edges of hoops and other obstacles. The waypoint fixation account does not make any predictions about gaze and steering behavior in the presence of obstacles or when future-path information may be unavailable. This shortcoming highlights one potential gap in our understanding of the capabilities of human visual control.

While driving an automobile there is almost always an explicit predefined path outlined in the environment, like a road or trail to follow. For instance, while driving down a highway a driver’s path is outlined by the lane along which they are driving and the side of the road. Lanes have highly salient features, such as the actual markings that define the lane and points between the markings that denote the path the driver is following. As discussed above, the visual information provided by the ground is sampled to guide locomotion during tasks like driving and walking. This is very different from flying quadcopters, as they fly through space instead of directly on the ground plane. Instead, there are often a series of obstacles or goal locations above the ground plane that they must avoid or intercept. Having access to a predefined path while driving has been shown to provide salient and useful information about an actor’s self-motion, future path, and spatial features that allow for correcting steering errors. Quadcopter pilots, however, do not have access to visual information that specifies waypoints directly along their future path because their future trajectory is not along the ground plane (see Figure 1). The inability to sample information specifying an actor’s future path may mean that a different gaze strategy is needed to efficiently steer in three dimensional spaces above the ground plane.

While flying a quadcopter, pilots must be able to produce actions that take the vehicles translational drift due to inertia and the effects of gravity into account to steer efficiently. This is not (usually) necessary while driving an automobile due to friction between the vehicle and the ground. Additionally, while driving an automobile the only types of motion that can be produced
are translating forward or backwards and curvilinear motion to the left or right. The differences between the physics needed to control quadcopters versus automobiles may result in very different steering strategies.

Another difference between driving automobiles and flying quadcopters is that while driving the main contribution of rotational components added to the optic flow field are from curvilinear motion and lateral head/eye movements. This differs from quadcopter flying where rotational components can be added to optic flow from curvilinear motion, rotation along the X, Y, or Z axis, eye movements, or any combination of these types of motion. The results from Lappi et al., (2020) showed that during curved path driving humans make smooth pursuit eye movements anchored to waypoints along their future path. These smooth pursuit eye movements simplify retinal flow by accounting for the lateral rotational component due to curvilinear motion. This simplification stabilizes the retinal flow about the waypoint through which they will pass. It is possible that during nonlinear quadcopter flight, similar smooth pursuit eye movements are generated which remove the rotational components due to nonlinear movement, but how the eye movements will deal with rotations due to roll remains unclear and must be evaluated.

These differences are by no means an exhaustive list of differences between controlling automobiles and quadcopters. It is possible, and likely, that these differences influence the gaze and steering behavior in ways that cannot be accounted for by models of steering automobiles. Chapters 3 and 4 discuss the findings from a pair of studies designed to explore the coordination of gaze and steering of quadcopter pilots as they fly through a series of racecourses that were created by manipulating the available visual information as it relates to future-path, waypoints, and obstacles. Additionally, this experiment investigated how the gaze and steering strategies of
quadcopter pilots is different from what would be predicted by various theoretical accounts, such as the waypoint fixation account.

1.4 THESIS OVERVIEW

Chapter 2 describes the setup of the novel eye tracking and quadcopter flight simulator. Chapters 3 and 4 focus on the results from a pair of experiments designed to investigate the ways quadcopter pilots use visual information to guide their actions. In both experiments, I examined the eye movements pilots made while flying a simulated quadcopter through virtual environments. The experiments used a novel eye-tracking paradigm implemented in virtual reality by means of a head-mounted display (HMD) to give quadcopter pilots a first-person view from the perspective of a simulated quadcopter’s onboard camera. The setup simultaneously recorded the actions executed by participants synced with an in-HMD eye-tracker which tracked their gaze location as they move through cluttered environments, such as through a dense forest, along forest paths, and through obstacle courses. The first experiment (Chapter 3) explored the role of path information on guiding locomotion through a racecourse environment. The second experiment (Chapter 4) examined how avoiding obstacles near an explicit path influences gaze and steering. The fifth and final chapter is a general discussion focused on limitations and future-directions.
2. A PARADIGM FOR STUDYING THE GAZE AND STEERING BEHAVIOR OF QUADCOPTER PILOTS

First-person view (FPV) quadcopter racing offers researchers a novel context in which to explore how humans actively generate eye movements to control their movement. The task of piloting a quadcopter through a cluttered environment is well-suited for advancing our understanding of vision-based control because pilots rely almost entirely on vision, move at high speeds, steer through waypoints, avoid obstacles, and move in 3D space within visually naturalistic environments. However, studying the eye movements of FPV drone pilots in the real world poses some significant challenges, such as accurately tracking drone movement and position, estimating gaze location relative to objects in world space, the ability to control the environment and its layout, and the safety of the pilot. In addition to these issues, exploring this task in the real world would require synchronizing an eye tracker with the video feed onboard a quadcopter, which poses significant technical challenges.

To create a tool for investigating quadcopter flight, I developed an experimental setup\(^2\) that can be used to explore gaze and steering during quadcopter flight in virtual reality using Unity and a Pupil Labs eye tracking system. The result is a visually realistic virtual environment and setup that allows for full control over the features/textures of the environment, the physics of the quadcopter, and eye tracking. The paradigm also includes a data post-processing pipeline to

\(^2\) The experimental paradigm (and the data visualization tool described below) was created with significant contributions from an experienced undergraduate Unity programmer, Xavier Marshall.
extract useful sources of information for analyses and visualization. This chapter starts by discussing the hardware and software used to develop the paradigm. Next, I describe aspects of the virtual environment and the procedure for creating it. I end the chapter by detailing how the setup can be used to extract data for analyses and visualizations.

The paradigm was developed through the work of three individuals. The bulk of the low-level Unity/C# code was written by Xavier Marshall, an undergraduate research assistant. His work also covered parts of the data post-processing that involved the Unity environment (i.e., extraction of instantaneous heading). I, along with my advisor Brett Fajen, provided guidance for the implementation of the setup by describing the necessary data to record, aspects of the task (i.e., how the quadcopter behaved, how the world looked, etc.), and the experimental procedure. I also conducted extensive testing of the setup throughout the development process to determine where changes needed to be made.

2.1 HARDWARE

The VR environment and eye tracking system were developed on a Windows PC equipped with an Intel i9 (11 series) ten-core processor, NVIDIA GeForce RTX3090 24GB graphics processor, and 32GB of 3200MHz DDR4 RAM. The environment was viewed in an HTC Vive Pro head-mounted display (HMD) with head tracking turned off to simulate realistic quadcopter HMD conditions. The HTC Vive Pro has a total resolution of 2160x1200 (1080x1200 per eye), a refresh rate of 90Hz, and 110-degree horizontal and vertical fields of view. A Pupil-Labs VR/AR extension (up to 200Hz refresh rate) eye tracker (Kassner, Patera, & Bulling, 2014) was mounted inside of the HMD. The eye tracker was setup so that each eye recorded at 120Hz with a resolution of 196x196. An HTC Vive or Vive Pro is necessary for using this model of eye tracker.
The simulated quadcopter is controlled using a Taranis Q X7 RC controller with micro-USB connection. This controller is a popular choice among pilots who race quadcopters. However, any controller that can be set up as a Unity game controller can be used with this paradigm (though some adjustments to input control may be necessary).

Figure 2: (A) Empty ground template with path painted over natural features. (B) Template with random tree placement.
2.2 VIRTUAL ENVIRONMENT

The Unity game/graphics engine (version 2019.3.14f) was used to create the environments. It consisted of a visually realistic ground terrain, types of trees, and lighting. A curved path was placed along the ground that maintained natural textures, and in some instances, hoops were placed along the path to fly through. A template virtual environment (see Figure 2A) that would be later modified to produce specific experimental conditions was created using the following processes. First, the ground was created using a procedural terrain generator (Gaia Pro asset from Unity store) in Unity that produces realistic and natural ground topologies. The scale, variability, and textures of the terrain can be manipulated using this asset to produce realistic conditions. The second step was to create the path. A curved path loop was then hand-crafted over the terrain using a Path Painting Asset available in Gaia Pro, which places a path along the generated terrain while also maintaining a natural visual appearance and realistic textures. The path varied in width between 1 and 2 meters. The path painting tool works clicking points along the ground terrain to set waypoints along the ground through which the software will paint a path. The terrain was then discretized into a grid-like pattern and various types of visually realistic trees were placed at random positions within each section of the grid. If trees were in undesirable locations, such as wide trees being placed directly on the center of the path or trees placed too close to one another, they were moved by hand in the Unity editor. The difference between the empty template and the random tree placement can be seen in Figure 2.

The quadcopter’s physics relied on an asset from the Unity store and mirrored a typical photography drone with the maximum speed set to approximately 17 meters per second. The left joystick controls the altitude (Y axis) and the yaw (X axis), while the right joystick controls the forward/backward (Y axis) and sideways (X axis) trust/translations. Moving joystick positions
back to their origins, or the center position, results in a stabilized quadcopter that hovers in place with no additional forces. The quadcopter controller model in Unity neglects the force of gravity as the vehicle is continuously in hover mode and air resistance was minimal. Collisions were not a major detriment to flying. If the vehicle collides with a tree, hoop, or the ground, then it simply bounces off and flight can continue. Some types of trees can cause the vehicle to get “stuck” in the branches, and if this happens the current lap around the racecourse can be restarted with a keyboard command.

This template environment was manipulated to create the experimental conditions used in Experiments 1 and 2 by changing aspects of the scene, such as placing hoops along the path, removing the visual outline of the path, and changing the relative proximities of the trees to one another (See Chapters 3 and 4). Each of these changes corresponds to the different experimental conditions found in those experiments. UXF, an open-source Unity-based experimental interface tool, was used to move participants through each phase of the experiment.

2.3 PERFORMANCE AND OPTIMIZATION

This setup requires hardware that can produce high resolution graphics rendering at high frame rates. The PC used for development was among the best commercially available hardware for graphics performance. During data collection, the eye tracker and VR environment must run simultaneously which puts a significant load on the computer and can make achieving the necessary frame rates difficult due to the resolution of the environment in the HMD and the resolution of the eyes in the eye tracker. Maintaining a stable frame rate of 90Hz for the HMD, 60Hz for quadcopter position recording, and 120Hz for eye tracking (per eye) was necessary for sufficient data recording (see below). The high level of detail of objects in the virtual
environment resulted in difficulty maintaining a stable 90 frames per second in the HMD, however using the methods below we were able to achieve such performance.

To overcome these challenges, we optimized the data recording and graphics settings so that many aspects of the environment were prerendered by Unity. Additionally, we modified the level of detail for the objects in the world so that it balanced visual realism with performance. This was particularly important for rendering the lighting sources in the environment. This was possible in the current state of the paradigm because the environment is static. If future studies require moving objects, overcoming the computational challenges may be more difficult because prerendering objects will not be possible as they need to be updated in real-time. Additional benchmarking will be necessary in future iterations of this paradigm as more sophisticated optimization is required.

2.4 EYE TRACKING QUALITY CONTROL

Eye movements are recorded, classified, and post-processed using the Pupil-Labs Core software. Eye tracker calibration was performed using an 18-point depth mapped calibration routine provided by Pupil-Lab open-source software. In addition to calibration, an assessment routine was created that used nine points that cover the visual field to determine the accuracy and quality of calibration. After calibration assessment, the average eye-tracking error in visual degrees is displayed on the screen. The experimenter can then choose to move the participant in
the experiment or have them recalibrate. If recalibration is chosen, the previous calibration recording is overwritten. This allows for some control of the quality of recorded eye tracking data.

2.5 DATA RECORDING AND POST-PROCESSING

While flying through the VR environment, the quadcopter’s position, orientation, and collisions were recorded at 60Hz and the eye tracking system recorded gaze data from each eye at 120Hz. There were two types of gaze data that were recorded. The first was made up of the normalized X and Y gaze positions on the image plane at each frame. The second is the 3D mapped point in world space of participant gaze relative to the camera. This is provided by casting a ray in the VR environment and finding the closest point that the ray collided with to

Figure 3: Visualization and post-processing tool. (A) Drone/quadcopter data can be viewed in playback (B) The visualization can be adjusted by changing camera point of view (C) The data being explored can easily be changed along with the experimental condition corresponding to that data (D) The gaze location in screen coordinates can be visualized. Each crosshair represents the left eye (blue) the right eye (yellow) and the cyclopian eye (green) (E) Lower-level data about gaze behavior can be viewed in playback (F) A image exportation tool that will generate a series of images that can be used to create detailed videos of the quadcopter flight and gaze behavior.
determine the position in world space that gaze is fixated and with what object that ray collides. This results in a recording of the objects gaze is oriented towards at a rate of 120Hz. The eye tracking data and quadcopter position data are collected on the same PC, and therefore synchronization of recording is essentially a nonissue. Some latency between the two data streams can occur, but those should be quite small. The quadcopter position data and the gaze data are combined by matching Unix timestamps. The data from Pupil-Labs was recorded at a higher frequency than the vehicle’s position and orientation. Therefore, after combination the gaze data was averaged within a frame to down sample it to the same frequency as the position data.

During an experimental session, the vehicle’s position and orientation are recorded, as well as the gaze location in world coordinates. This data can then be imported into a data visualization/extraction tool (see Figure 3) to make visualizations of the data and export further sources of information, such as gaze target, instantaneous heading in screen coordinates, and the object aligned with instantaneous heading. The data tool’s interface shown in Figure 3 has a variety of features. Any dataset flown through the experiment can be easily imported into the tool (Figure 3C) and the information presented on the screen can be easily changed (Figure 3B.

Figure 4: The left panel shows the first-person point of view seen by the pilot with gaze overlay and object highlighted. The right panel shows the top-down view with gaze vector cast into world space.
and F). The tool can then be used to inspect the data, such as the gaze location estimated from either of the eyes and what object is aligned with gaze (Figure 3A, D, and E).

The recorded positions and orientations of the quadcopter were used to recreate the trajectory flown by participants. During trajectory recreation (Figure 3F) the system records the images corresponding to what pilots saw during their laps around the virtual environment, as well as a top-down view of the quadcopter (see Figure 4). Recreating the visual information must be accomplished in post processing because of the hardware demands from rendering graphically realistic virtual environments in real-time simultaneously with the eye-tracking software. The trajectory visualization tool can place several useful markers signifying location of gaze, heading, and gaze target. These markers can be modified for both the first-person view and the top-down view, and are shown in Figure 3F.

Using the recorded quadcopter position, the vehicle’s velocity was calculated and using that the vehicle’s instantaneous heading was calculated at each location to which the vehicle moved. The resulting heading vector relative to the quadcopter was then extended into world space until it collided with an object, like the procedure for determining gaze location in 3D space. The object aligned with heading was then recorded at each frame. This can also be viewed and exported using the data visualization tool.

2.6 FUTURE USE OF THIS EXPERIMENTAL PARADIGM

This setup can be used in the future to further investigate the role of different sources of visual information on steering through cluttered 3D spaces. As described above, this setup can be used to easily manipulate a variety of different sources of information, such as the low-level visual information, the movement of external objects, and/or the physics of the vehicle. The setup’s flexibility paired with its ability to extract useful information and produce visualizations
of gaze and steering make it well suited for researchers investigating vision-based locomotion, eye movements, and/or motor control. In the following two chapters, I describe how I used this setup to explore the relationship of gaze and steering in the presence of an explicit path along the ground plane and in the presence of obstacles that lie near or directly on the path.
3. EXPERIMENT 1: COORDINATION OF GAZE AND STEERING DURING QUADCOPTER FLIGHT

During natural human locomotion, eye movements are used to sample relevant visual information so that actions can be selected and executed to move the body safely through the world. Human gaze during locomotion has been examined in many different task-spaces, from walking over different kinds of terrain (Marigold & Patla, 2007; Matthis et al., 2018; Matthis et al., 2022) to steering automobiles under various conditions (Lappi, 2016). In past studies, researchers have provided evidence for the tight coupling of gaze and motor commands, yet questions surrounding the exact nature of how humans coordinate gaze with locomotion remain.

As explained in Chapter 1, the past research on how humans use gaze to steer has focused almost exclusively on human locomotion along 2D ground planes with access to visual information specifying a path to follow. To recap the most relevant points from that chapter, there is a long history of research being done to investigate how drivers use gaze to sample points along the ground to steer automobiles. The presence of a visually explicit path provides actors with salient features along the ground plane to which a fixation can be made and used for steering adjustments, such as the lines that make up a lane on a road (Land & Lee, 1994; Land & Horwood, 1995; Lappi & Mole, 2018). Humans, however, are also capable of guiding locomotion during tasks that are very different from driving automobiles, such as the task of flying a quadcopter at high speeds through 3D racecourses and forests with densely spaced trees. During quadcopter flight, both the control of steering and the available visual information are
quite different from driving automobiles. Recent research has provided some insight into the patterns of gaze and steering that emerge during high-speed quadcopter flight (Pfeiffer & Scaramuzza, 2021), but little is known about how different sources of visual information guides the coordination of gaze and steering during flight through 3D spaces. The question arises, how do quadcopter pilots use gaze to sample the available visual information to guide their steering? It is possible that during quadcopter flight, as in driving, information provided by the ground plane may be useful for guiding flight. If pilots use information provided by the ground plane to guide locomotion, then they should spend a significant portion of time with gaze oriented towards regions along the ground.

Given the limited amount of previous research on gaze behavior during high-speed steering and obstacle avoidance in 3D environments, the present study is exploratory with the goal of developing an accurate description of gaze and steering in these conditions. This experiment was designed to answer two sets of questions. The first set of questions are oriented towards comparing flight through 3D space to locomotion along the ground. Do quadcopter pilots use gaze and steering strategies in the same way as drivers? What gaze and steering strategies do pilots use in conditions very different from driving a car, such as the presence of waypoints (i.e., hoops) above the ground through which to pass?

To answer this first set of questions, we used the experimental paradigm described in Chapter 2 to investigate how the absence of an explicit path outline along the ground influences the coordination of gaze and steering behavior. The experiment comprised three conditions. The first condition (Path Only) is a path following task, where participants were instructed to fly a quadcopter along a curved path as quickly as possible while minimizing deviations from the center of the path. The curved path contains a visually explicit outline defined by natural
textures. This condition more closely simulates the kind of environment an automobile driver would experience and may be better suited for participants to use a gaze strategy dominated by eye movements to points along the ground over which they will soon move. The second condition (Hoops Only) is similar, but participants were asked to fly through a series of intermittently placed hoops above the ground plane. In this condition, the visually explicit textures that define the path were removed. The hoops were placed at points along the same, now invisible, path from the Path Only condition. The third condition (Path and Hoops) combines the first two conditions, placing hoops above the path. The Hoops Only and Path and Hoops conditions resemble situations very different from driving an automobile because their future trajectory is through space above the ground and there are explicit waypoints above the ground through which to fly.

If pilots use the visual information provided by the ground plane to guide action, then they should spend time gazing at locations along the outlined path when it is present to guide steering over those locations. The inclusion of waypoints above the ground plane should attract attention too, but intermittent time would be spent looking at points along the path they will fly over if that information is being used to guide flight. If little to no time is spent orienting gaze towards the path or the terrain near the path, then it may be that pilots predominantly rely on other sources of information to guide steering (e.g., hoop edges, obstacles, trees in the distance, etc.).

The second set of questions were formulated to further explore and describe the gaze and steering behavior of quadcopter pilots. Does gaze lead steering? Pfeiffer and Scaramuzza (2021) provided evidence for gaze leading steering in the same direction. Does their finding generalize to the current task? Research on driving (Lehtonen, et al., 2013; Wilkie, Wann, & Allison, 2008)
has provided evidence for anticipatory eye movements. In the present task, do pilots make look-ahead eye movements and/or anticipatory steering adjustments? Lastly, quadcopter pilots control their rate of yaw independent of their translation. In Chapter 1, I described how lateral rotation can be problematic for the perception of heading. Do pilots deal with camera rotations? And if so, how? These questions are answered using a set of analyses and allow us to situate quadcopter pilot behavior in this task within the broader scope of visually guided locomotion and task-dependent eye movements.

3.1 METHODS

3.1.1 PARTICIPANTS

Participants were students recruited from Rensselaer Polytechnic Institute (18-29 years old) with valid drivers’ licenses and vision corrected to near normal (contact lenses only). Participants had varying levels of experience flying drones, some having experience flying drones in the real world and others having flown them in simulation. Participants all reported having more than 10 hours of experience flying quadcopters, with the highest amount being 200+ hours. There was a total of six participants (N=6), all of whom were male. Included in the experiment is the non-naive author of this paper who has logged a substantial number of hours flying in simulation. The experiment was approved by the Rensselaer Polytechnic Institute Institutional Review Board and all participants signed written informed consent. Participants were given monetary payment for participating in the experiment.

Both experiments reported in this thesis had small sample sizes: N=6 and N=3, respectively. This was due to several factors, such as quadcopter racing being a niche hobby, data collection during a large COVID-19 outbreak, and the risk of simulator sickness. The problems with participants recruitment were first experienced due to the limited participant pool.
of individuals with an appropriate amount of experience flying quadcopters. This problem was exacerbated by the COVID-19 pandemic, which further reduced the possible pool of participants to individuals with on-campus access to the university. Nevertheless, small sample size studies are not uncommon in vision science, and from each of the participants a large amount of data was logged. While caution must be exercised in generalizing to the population, we can still explain the behavior in the current sample and gain insight into how some quadcopter pilots use vision to control their actions. As mentioned above, one of the participants was the author. In many subfields of psychological and vision science, a non-naive participant could introduce bias. This, however, was not the case as the performance and behavior of the author was consistent.

Figure 5: (A) Contains the setup of the experiment with user controlling the drone and eye-tracking software running. (B) Shows the mounted pupil-labs eye-tracker in a vive pro HMD. (C) Shows the taranis drone controller used. (D) Shows the path outline that participants needed to negotiate through the racecourse.
with the other participants.

### 3.1.2 APPARATUS AND VIRTUAL ENVIRONMENT

Participants were asked to control a simulated quadcopter with realistic flight dynamics through a virtual environment built in Unity. The development of the virtual environment was described in Chapter 2. We modified the template environment to create three experimental conditions for Experiment 1. In the two hoop-containing conditions (Hoop Only and Path and Hoop) hoops were placed at various intervals of distance from each other and at varying heights above the curved path racecourse. The hoops were placed at 3D positions along the varying ground topology 17.2m apart on average with a standard deviation of 9m. A total of 42 hoops were placed along the path and it was verified that each hoop was visible from the previous. In the Hoop Only condition, the path along the ground was made invisible and a green grass texture was placed over it. The Path Only condition was essentially the template environment that was described in Chapter 2, with trees randomly placed and moved by hand to make sure no trees fell directly on the path. Example images from each of the conditions are illustrated in Figure 7. The setup, eye tracker mounted in HMD, and controller can be seen in Figure 5A, B, and C.

Several participants stated a noticeable difference between the implemented control scheme and real-world quadcopter flight. The most notable difference being that in the current experiment the thrust is applied by both axes of the right joystick, which is different from real world racing drones where the thrust is solely applied by the Y-axis of the left joystick. For real world racing drones the rate of change in orientation is controlled by the X-axis of the left joystick (yaw) and both axes of the right joystick (pitch and roll). We chose this control scheme because it reduced the complexity of the degrees of freedom of motion in an already complex
control space and as long as participants were able to perform the task these differences in control were not detrimental to our ability to address the experimental questions.

3.1.3 PROCEDURE AND DESIGN

The experiment used a repeated-measures design the implementation of which was controlled by Unity Experiment Framework (UXF). Figure 6 illustrates the block design. There was a total of five blocks, each of which contained three sub-blocks corresponding to the three experimental conditions. The order of conditions reversed on each subsequent block and each sub-block contained two laps around the racecourse. The Path and Hoops condition always occurred second in the ordering.

Participants first signed written consent after which they listened to brief instructions on the experiment. They were then instructed to place the HMD on their head and given time to adjust the head straps so that they can see the full view of the screen in the HMD and so the eye tracker camera has adequate coverage of their eyes. Participants were taught to calibrate the eye

![Figure 6: The experimental design. Each block contained two laps of each condition. Subsequent blocks alternated the ordering of the conditions, with the path and hoops condition always being second.](image)
tracker, which used a built in VR calibration tool from the Pupil-labs software. While learning the calibration procedure, participants were instructed to pay attention to the eye view that is projected into the HMD, which shows a video of each eye. The experimenter asked for HMD position adjustments as needed to keep the eyes visible within the eye tracking cameras. This facilitated consistent eye coverage for accurate and stable tracking of pupil positions throughout.

Figure 7: An example of each of the experimental conditions.
the experiment. Next, participants learned the physics of the virtual quadcopter and the controls. Participants were given three minutes to fly freely in an open environment to learn the controls and any differences from real-world quadcopter flight. After the initial exploration phase, they were asked if they were sufficiently comfortable with the controls to continue to the testing blocks. If they were not, they were given additional intervals of one minute to continue practice flying until they felt comfortable controlling the quadcopter.

After becoming comfortable with the HMD, drone physics, and eye tracker, participants were asked to complete five blocks of the experiment, each of which comprised of the sub-blocks illustrated in Figure 6. At the start of each sub-block, participants were asked to recalibrate the eye-tracker and after each calibration the quality of the calibration is assessed and recorded. After calibration and assessment, they completed two laps per sub-block around the virtual racecourse as quickly as they could. Participants completed a total of ten laps around the course in each condition. If at any point participants strayed too far off the path (approximately 10 meters in left or right direction and 20 meters in the vertical direction), they were automatically transported to the start of their current lap and the recorded data for that lap was overwritten. In the fifth block, participants started each condition facing the reverse direction and were asked to complete two laps per condition flying in the opposite direction. Each lap took approximately two to four minutes to complete depending on a pilot’s skill level.

3.1.4 DATA POST-PROCESSING AND ANALYSES

3.1.4.1 EYE TRACKING DATA

The eye tracking data was post-processed using Pupil-Labs Pupil Player software version 3.4. Pupil estimation and gaze mapping for each recording was processed by hand in the pupil-labs’ editor and exported. The exported gaze positions were filtered by confidence by discarding
gaze positions with confidence below 0.6. The resulting positions were then combined with the live recording of the quadcopter position by comparing Unix timestamps. The eye tracker recorded at a higher frequency than the position of the vehicle, so the gaze locations were averaged and weighted by confidence level at each vehicle position frame. The combined gaze and quadcopter position dataset was then fed back into Unity to recreate the quadcopter’s trajectory. At each position the gaze location in world space was represented as a ray, and the object with which the ray collided was recorded as the object that the participant was looking at in each frame. This same process was done using the instantaneous velocity vector (i.e., heading) to determine what object participants were heading towards at each frame. Lastly, the gaze data recorded from Pupil-Labs was incorrectly recorded from two instances of Path Only conditions in separate participants. This occurred due to software issues. Those recordings were excluded from the analyses.

3.1.4.2 GAZE AND STEERING TIME-SERIES

To determine the relationship between gaze and steering, the data were decomposed into time-series of angles that describe the time-evolving gaze and steering signals. The participants’ flight trajectories were parsed into between-hoop segments, where the first timestep is where they passed through the previous hoop and the last is when they passed through the next hoop. In each of the segments, the angles describe the direction of thrust, the direction of the forward-facing camera, and the direction of gaze relative to the instantaneous velocity vector (i.e., heading). Figure 8 depicts each of those angles from a top-down perspective. Thrust was approximated as the controller’s right joystick offset from center (i.e., control of lateral and forward translation) and rotated the resulting vector relative to the quadcopter. The angle between the resulting thrust vector and the instantaneous heading relative to the quadcopter was
then calculated. Lastly, the angle between gaze location relative to quadcopter position and heading was calculated.

Additionally, I calculated the angle between the instantaneous heading and a normal vector passing laterally through the hoop, representing the approach angle of the quadcopter as it passes through a hoop. This angle was used in a series of linear mixed effects models described in the following section. All mixed linear models were created in the LME4 package in the R programming language. The degrees of freedom that are seen in each of the linear models are adjusted by default by the LME4 package in R. In all figures, error bars correspond to 95% confidence interval around means with between subject variation removed.

The CROSSCOR function in MATLAB was used for cross-correlation analysis. The time-series from each block were parsed into segments defining the periods between hoops, a total of 42 per lap, or 84 per condition within a block. Cross-correlations were conducted on each segment and then the maximum correlation and corresponding lag were averaged across

Figure 8: Top-down point-of-view of the angles used for cross correlation analysis. Green is the angle between gaze vector and heading vector, purple is the angle between the trust vector and heading vector, and blue is the angle between a vector representing the orientation of the forward-facing camera and the heading vector.
segments. The maximum lag was set to 30 frames, or 0.5 seconds. The between hoop segments that contained more than one collision were excluded from cross-correlation analysis.

3.2 RESULTS AND DISCUSSION

3.2.1 GENERAL MEASURES OF PERFORMANCE

The analyses reported in this section were conducted to examine if participants could perform the task, if performance improved over time, if performance varied across conditions, and if familiarity with the path influenced behavior. In general, participants were able to complete laps around the racecourse, pass through hoops, and follow the path in each of the conditions.

![Figure 9: General performance measures across conditions and blocks.](image-url)
Participants stayed relatively close to the center of the path, deviating overall on average 0.83m per lap. No differences in deviation from the center of the path were found between the Path Only and Path and Hoops conditions ($F(2, 168.01) = 2.99, p=0.053, \eta^2 = 0.03$). In the two hoop containing conditions, participants largely flew through all the hoops, completing nearly 96% of them each lap on average. No differences were found in the percentage of hoops completed between the Hoop only and Path and Hoops conditions ($F(1, 113) = 4.09, p=0.04, \eta^2 = 0.03$). In all conditions there were minimal collisions with objects, happening less than 10 times on average. This number is artificially inflated from the first block where collisions were more common. There were no differences in the number of collisions between the conditions ($F(2, 173) = 1.11, p<0.27, \eta^2 = 0.01$). Participants took approximately 79s to complete a lap on average with a significant difference in lap completion time between conditions ($F(2, 168.01) = 11.05, p<0.001, \eta^2 = 0.12$). Using pairwise comparisons, a significant difference was found between Hoops Only and Path Only conditions ($t(168)=4.089, p=0.0002, \eta^2=0.1$) and between Path And Hoops and Path Only conditions ($t(168)=4.104, p=0.0002, \eta^2=0.09$). On average, participants flew the vehicle approximately 12.1m/s, with a significant difference between conditions ($F(2,168.01) = 15.42, p<0.001, \eta^2 = 0.16$). Using pairwise comparisons, a significant difference was found between Hoops Only and Path Only conditions ($t(168)=-4.927, p<0.0001, \eta^2=0.13$) and between Path and Hoops and Path Only conditions ($t(168)=-4.745, p<0.0001, \eta^2=0.12$). All pairwise comparisons used Tukey HSD method for adjusting p-values.

The left and right columns of Figure 9 depict each of these performance measures broken down by condition and by block respectively. Performance across blocks was largely consistent after the first block in all conditions. This likely was due to participants becoming familiar with the environment and task in the first block. Trends showing improvement in some performance
categories (speed, time to complete lap, path deviation, and number of collisions) can be found in the right column of Figure 9. In all performance measures, the difference between the test block where they flew in the reverse direction and the fourth block was minimal. This suggests that either performance was not affected by familiarity with the course layout or that four blocks was not sufficient to gain such familiarity.

Figure 10: Distribution of gaze categories by condition and block. All proportions were calculated by lap and then averaged across participants.
3.2.2 GAZE BEHAVIOR

The analyses reported in this section focus on where participants looked and address the question of whether quadcopter pilots exhibit gaze behavior that is similar to that of automobile drivers, who tend to fixate points on the ground over which they plan to steer (Lappi & Mole, 2018). I begin with an analysis of the distribution of gaze across five categories of possible fixation targets: the path outline, the terrain other than the path, the trees near the quadcopter, the hoop edges, and through the center of the hoop (see Figure 10). Trees near the quadcopter were considered those within 2 average hoop segment distances from the vehicle (34.4 meters) at each frame. The terrain other than the path includes the ground textures outside of the path outline and trees that are far off in the distance.

A significant difference in the amount of time spent orienting gaze towards the path was found between conditions ($F(1,5) = 15.10, p = 0.012, \eta^2 = 0.75$). No significant difference across conditions was found between time spent looking at the ground terrain excluding the path between conditions ($F(2,10) = 1.37, p = 0.85, \eta^2 = 0.16$). There was also no effect of condition on the amount of time spent orienting gaze towards trees near the quadcopter ($F(2,10) = 2.15, p < 0.166, \eta^2 = 0.3$). Within the two hoop-containing conditions there was no difference in the amount of time spent orienting gaze through hoop centers ($F(1,5) = 0.115, p = 0.14, \eta^2 = 0.02$), or their edges ($F(1,5) = 1.145, p = 0.67, \eta^2 = 0.19$).
Next, I will examine the distribution of distances from the quadcopter to the fixated object. As shown in the top panel of Figure 11, participant’s gaze most often fell on points in the environment between 10 and 20 meters away from the vehicle. This pattern was largely consistent patterns across participants, although there was some variation (see black lines). In the Path Only condition, participants on average looked at points closer in the environment compared to the two hoop containing conditions ($F(2,10)=43.00$, $p<0.001$, $\eta^2=0.90$).
The last analysis in this section focuses on when participants first looked at each hoop in the Hoop Only and Path and Hoop conditions. Initial fixations to a hoop or through its center occurred on average ~17m or ~1.6s before reaching the hoop (Figure 12). No difference in distance or time of initial fixation to hoop was found between conditions (Distance: $F(1,5) = 0.09, p=0.771, \eta^2 = 0.02$; Time: $F(1,5) = 0.160, p = 0.706, \eta^2 = 0.03$). On average, participants made their initial fixation to the subsequent hoop just as they reached the nearest hoop, although

![Distance to hoop at first eye movement](image1)

![Time of first eye movement before reaching hoop](image2)

![Timing of first eye movement to next hoop relative to nearest](image3)

*Figure 12: Measures of gaze behavior relative to hoops across hoop-containing conditions.*
it often happened before and after reaching the nearest hoop as well (See Figure 12). No difference was found in the timing of the initial fixation to the subsequent hoop between conditions ($F(1,5) = 0.003, p = 0.960, \eta^2 < 0.001$). The right column of Figure 12 shows these measures broken down by condition and block.

Next, I will attempt to synthesize these findings into a cohesive account of where participants looked in each condition. In the Path Only condition, the condition most similar to driving an automobile, participants looked at the path more often than any other category of

![Graphs showing gaze ratios and densities](image)

**Figure 13:** The ratio between how often hoop centers and edges fell under the gaze crosshair and the number of possible instances where gaze could have fell towards those objects. The black line shows the average between the four subjects, and the bottom graph shows the density of inter-hoop segment lengths mirrored over 0. The hoop combined ratio plot is the sum of the two plots above it. The vertical black line is the mean and the vertical red line is the median. Four subjects were chosen for this analysis due to the accuracy of their eye tracking data. The proportions don’t add up to one right at the hoop due to instances where participants missed the hoop or had the vehicle not oriented directly through the hoop.
object (34%), but they also spend significant time looking elsewhere. Seeing as participants also spent ~25% of the time looking at the terrain other than the path, one might wonder if most of those fixations were directed to locations near the path. To answer this question, I reanalysed the data to estimate the percentage of time participants spent looking at ground terrain within 4 degrees of visual angle of the path center ($mean=50.9\%, CI=[42.3, 59.7]$). Overall, when the task required following a path, the dominant gaze behavior was looking toward the ground plane at or near the path, which is similar to the gaze behavior of automobile drivers. However, participants also intermittently looked at nearby trees and the surrounding terrain.

In the Hoop Only condition, the dominant gaze behavior was looking through the center of the hoop (34% of the time). Less frequently, participants looked at nearby trees (22%) and the ground terrain or far away trees (29%). Interestingly, participants infrequently looked at the edges of the hoop (14%). As participants approached each hoop, the center of the hoop occupied an increasingly larger percentage of the visual field. Eventually, in the last frames before passing through each hoop, the center took up the entire visual field. As such, one might wonder how often participants looked through the hoop center before it was the only object category at which the participant could look. Figure 13 illustrates the how often participants looked towards the hoop or its edges as a ratio of the number of times participants looked at those categories and the number of times they possibly could have looked towards those categories. This was completed as a function of time. The data from both hoop-containing conditions were used in this analysis. The bottom graph in Figure 13 shows the density of between hoop segment lengths mirrored over zero. The red vertical line marks the median of those values at approximately 1.7 seconds long. This means that 50% of hoop segments were shorter than this. During segments less than 1.7 seconds, the majority of frames were spent looking towards the hoop edge or center shown
by the black line in the Hoop and Hoop edge graph. This means that participants spend a large percentage of the possible time looking through or at the hoop.

The gaze behavior observed in the Path and Hoops condition was very similar to the Hoops Only condition. When considered with the finding that these two conditions were also similar across performance measures (Figure 9) and gaze object category distributions (Figure 10), it appears that the visual information from the path played a negligible role when the task required steering through hoops.

This evidence provided by these analyses suggest that quadcopter pilots look towards where they want to move the vehicle and spend little time looking towards obstacles or information unrelated to their current or upcoming movement. In the Path Only condition, that is reflected by gaze being predominantly oriented towards points along or near the path. In the hoop containing conditions, that is reflected by gaze oriented towards waypoints through which they will soon pass. However, participants did spend some time looking towards other sources of information, such as terrain not near the path. Taken together, these findings suggest a gaze strategy mostly consistent with what has been found in past research on automobile driving, which has provided evidence that people tend to fixate the points over which they will soon steer (Lappi & Mole, 2018; Lappi, 2016).
3.2.3 DO PILOTS LOOK IN THE DIRECTION TOWARD WHICH THEY INTEND TO STEER?

The analyses reported in the previous section suggest that quadcopter pilots often look in the direction of waypoints (i.e., points along the ground over which they intend to steer or through the center of the upcoming hoop). The one previous study of gaze behavior in quadcopter pilots (Pfeiffer & Scaramuzza, 2021) reported a finding that is largely compatible with this strategy of looking towards intended waypoints. Specifically, they found that gaze often preceded a steering command in the same direction approximately 223ms ahead of time. The

Cross-correlation: gaze and thrust angles

Figure 14: Cross-correlation of thrust and gaze angles broken down by lag and correlation. The blue star corresponds to the results found by Pfeiffer & Scaramuzza (2021). The top and side panels illustrate the correlation densities of the lag and correlation coefficients.
present analysis was conducted to determine if their conclusions generalize to the task that participants performed in the present study, which was considerably more complex than the repeated figure-8 path following task used by Pfeiffer & Scaramuzza.

Using cross-correlation analysis, I determined the relationship between gaze and steering and the time-offset between the time-series of gaze angle and thrust angle, where gaze angle is the angle between gaze and instantaneous heading and thrust angle is the angle between the direction of thrust and instantaneous heading (see green and purple angles in top right panel of Figure 14). These angles were calculated at each timestep and were then decomposed into segments where the first timestep is when they passed through the previous hoop and the last timestep is when they passed through the subsequent hoop. This resulted in 42 segments per lap. Cross-correlation between the two time-series were conducted with a maximum time lag of 30 frames, or 0.5 seconds. Figure 14 shows the 2D density of the resulting maximum correlation coefficient and corresponding time lag across all participants in both hoop-containing conditions. In a large portion of between hoop segments a strong correlation ($r=0.70$) was found between the gaze and thrust angles clustered around a lag of 200ms, which is illustrated in Figure 14.

The results from the present analysis, however, had a high degree of variability (See density panels in the top and right side of Figure 14). While a large portion of the cross-correlations were at a maximum at a lag of about 200ms, there were many instances where gaze and thrust were negatively correlated or had a negative lag. There are at least two possible reasons for these differences. The first is that the task in the present experiment was more complex (path following and obstacle avoidance during flight) with slightly different vehicle controls compared to the task in Pfeiffer and Scaramuzza’s study. The second was a sinusoidal pattern that often occurred in the angle of gaze or the angle of thrust, which results in the
relationship between the two time-series becoming ambiguous. In these instances, the two time-series will be correlated mirrored around the 0 lag at a time offset roughly equal to the period of the signal. This can result in a seemingly negative correlation and opposite signed lag, making interpretation of those values difficult.

### 3.2.1 EVIDENCE SUGGESTS ANTICIPATORY EYE MOVEMENTS AND STEERING

Past research has shown that drivers make look-ahead eye movements to points in space that allow them to anticipate upcoming conditions (Lehtonen et. al., 2014). It is possible that quadcopter pilots use these kinds of eye movements to plan for upcoming environmental conditions, such as the curvature of the path or the layout of the hoops. This can aid them in steering the vehicle towards desirable locations in cluttered environments. The following analysis was conducted to examine if participants made look-ahead eye movements, and if those eye movements resulted in steering trajectories that were influenced by the layout of the hoops.

Evidence has suggested anticipatory eye movements that sample information more than 3 seconds into the future during tasks like driving cars (Land & Horwood, 1995; Mennie et. al., 2007; Lehtonen, 2014). The results from the gaze behavior analysis (Figure 12) above demonstrated that participants made eye movements to hoop N+1 before reaching hoop N (~0.05 seconds). This suggests anticipatory eye movements may play a role in how they steer through the nearest hoop. This, however, is a weak effect. One potential reason for this is the possibility of a false negative if looking through the center of hoop N+1 before passing through hoop N, as the recording can only determine when looking through the center of one hoop at a
Further analysis of the gaze data will need to be conducted to determine if stronger evidence in favor of anticipatory gaze patterns is present.

Anticipatory actions may also be present in the ways pilots adapt their current flight trajectory with upcoming environmental conditions. For example, aspects of the environmental layout between the nearest hoop and the subsequent may predict how the quadcopter passes through the nearest hoop. A series of linear mixed-effects models was created to predict the
orientation of the quadcopter along its vertical axis (i.e., yaw) relative to the nearest hoop as it passes through that hoop, the approach angle of the quadcopter relative to the normal vector passing through the center of hoop N, and the angle between the direction of thrust and the normal vector of hoop N (See Figure 15). Each linear model consisted of one of the following two predictor variables; either the relative orientation of sequential hoops along their vertical axis or the angular offset between the hoops (See Figure 15). These covariates were included to account for participant behavior as they passed through hoop N-1. Additionally, each model included covariates that represented the relationship of the vehicle to hoop N as the vehicle passed through hoop N-1. These covariates were the angular offset and relative orientation of the vehicle as it passed through hoop N-1 relative to hoop N (See Figure 15). These models were then used to predict approach angle, camera angle, and thrust angle at different normalized timepoints in the interloop segment. In the x-axis of Figure 16, the left side is the beginning of the normalized hoop segment as they had just passed through hoop N-1 and the right side is the point where they pass through hoop N. The points in the middle of the figure refer to the results of the linear model corresponding to 25%, 50%, and 75% of the hoop segment completed.

This resulted in models of the following general form:

\[ \theta_{yt} \sim Subject + Covariates + \theta_{xt} \]  

(1)

This general model (Equation 1) includes predicted variable \( \theta_{yt} \), which refers to either approach angle, camera angle, and thrust angle at the proportion of the hoop segment \( t \). The model accounts for the random effect of subject. The covariates included in the model were the angular offset and relative orientation of the vehicle relative to hoop N as it passed through hoop N-1. Lastly, \( \theta_{xt} \) is the predictor variable \( x \), either angular offset or relative hoop orientation, at proportion of hoop segment \( t \). Using the output from each model corresponding to each
dependent and independent variable combination, I report the partial r-squared for each independent variable to describe how much variance is explained (See Figure 16).

After just passing through hoop N-1, the angular offset is not predictive of approach angle, relative camera angle, or thrust angles, with partial r-squared values ranging from 0.01 to 0.06. However, as the vehicle moves closer to hoop N, the predictive quality of angular offset
increases (See Figure 16). The results show that for all models the angular offset has a partial r-squared of approximately 0.2-0.3 as the vehicle passes through hoop N. This means that the angular offset between hoops N and N+1 accounts for approximately 20% to 30% of the total model variance in each model as they pass through hoop N, significantly predicting approach angle, relative camera angle, and thrust angle. The relative orientation between hoop N and N+1 was not predictive.

The results from these models provide evidence suggesting pilots adapt their current flight trajectory, vehicle orientation, and motor commands based on the spatial relationship of hoop N and N+1. Both the vehicle’s approach angle to the hoop and relative orientation to a hoop as it passes by can be predicted by the relationship between the nearest and subsequent hoop. This provides evidence that participant approach angle is biased by the relative angular offset of the subsequent hoop, suggesting they are attempting to optimize their path by taking non straight approaches through the nearest hoop to better fly through the subsequent hoop given the physics of the vehicle.

The adaptation of current trajectory is surprising given that previous investigations of adapting steering trajectories in anticipation of upcoming environmental conditions has been weak. Past studies on locomoting through similar slalom courses have only shown weak evidence of aligning the body with the next goal before reaching the nearest goal. For example, the experiment conducted by Wilkie, Wann, and Allison (2008) found weak evidence for such steering trajectories. The present finding appears to be stronger than what they observed, but consistent.

This anticipatory behavior may be due to the influence of inertia on steering. In the present study, a steering command will cause drift in the direction of translation due to the inertia
of the vehicle. In past studies on how people steer, their rate of yaw remains fixed with the instantaneous direction of heading, and the effects of inertia do not play a major role in trajectory planning due to friction between the body/vehicle and the ground. Due to those factors, a more reactive steering strategy may be more appropriate because their steering command does not have a lagged effect on a vehicle’s trajectory. In the present case, this kind of reactive vehicular physics would allow them to make a steering command right at the turn (or hoop in our cases) instead of initiating the turn before reaching the hoop to account for inertia. In fact, if a pilot is sufficiently tuned to the dynamics of a quadcopter, then it is necessary to start a turn well before reaching the future intended inflection point in the anticipated trajectory.

### 3.2.2 DO PILOTS DEAL WITH THE ROTATION ALONG THE VERTICAL AXIS?

To steer towards goal locations, it is reasonable to assume that quadcopter pilots need an accurate estimate of their direction of heading. However, during quadcopter flight, pilots control

![Figure 17: Distribution of the angle between camera and heading across participants and broken down by condition.](image-url)
their vehicle’s yaw independent of translation. In these conditions, it may be necessary to deal with that rotation to perceive their direction of self-motion due to added rotational components in the optic flow array. This has been shown to be a difficult phenomenon to explain and has been studied extensively (Li & Warren, 2000; Li & Warren, 2002; Banks et al., 1996). How, if at all, do pilots accomplish this? One way that they may deal with their changing yaw is to try and minimize the deviation between instantaneous heading and the forward-camera on the vehicle (See Figure 8). This would allow them to keep their direction of self-motion close to the center of the screen in the lateral direction. To investigate this, I calculated the angle between the forward orientation of the vehicle’s camera and instantaneous heading at each frame. Only frames where the quadcopter was moving were considered. The distribution of those angles in each of the conditions can be seen in the upper row of Figure 16. The mean angle between heading and the camera is close to zero in each of the conditions (i.e., the vertical blue lines). This shows that participants keep the angle between camera and heading relatively low, but there is a large amount of variation in these distributions, with a standard deviation close to 20 degrees each in condition. Camera angles of these magnitudes reflect that the camera is often noticeably unaligned with the vehicle’s direction of heading. The bottom row of Figure 16 shows the distribution of the rate of change of the camera angle. It shows that the angle often changes very quickly, though the average is close to zero. These findings are similar across conditions. Overall, these findings suggests that participants are likely not using this kind of control strategy to deal with their rotation about the vertical axis. Future experiments will more systematically investigate the influence of independent control of yaw and translation on gaze and steering strategies.
3.3 SUMMARY OF EXPERIMENT 1 FINDINGS

This experiment was designed to investigate several questions that were outlined in the introduction of this chapter. The results from this study provide further evidence that people look towards where they want to go instead of locations they wish to avoid. Furthermore, the regions of the visual scene that gaze is oriented towards is dependent on aspects of the task. Gaze was oriented towards objects (or lack thereof in the case of flying through a hoop) in the world that participants needed to move towards, and those objects changed depending on the experimental condition. In hoop-containing conditions gaze was oriented towards or through the center of a hoop, and in the path only condition gaze was oriented towards the path or nearby terrain over which they would soon move the vehicle. The replication of Pfeiffer and Scaramuzza’s cross correlation analysis also supports this, and the results from their experiment and this experiment show that gaze often precedes steering in the same direction. The hypothesis that participants look towards where they desire to move is investigated further in the next chapter.

I also provided evidence that suggests pilots use anticipatory gaze strategies and steering control. The orientation of the vehicle as it passes through the closest hoop can be predicted by the relationship between hoop N and hoop N+1. Furthermore, participants sometimes made eye movements to hoop N+1 before passing through hoop N. These findings are consistent with past research on anticipatory eye movements, but past research providing evidence for adapting steering for future goal locations has been weak.

The last analysis investigated how participants deal with their yaw changing independent of their translation. The results showed some weak evidence that participants are attempting to keep the forward-facing camera oriented towards the vehicle’s direction of motion, but the data reflects a large amount of variation. Participants maintained a camera offset relative to direction
of heading close to zero, but that deviated plus or minus 20 degrees. Additionally, the rate of change in the camera offset was on average close to 0, but the vehicle’s rate of yaw was quite sensitive, leading to large variations in the rate of change in camera offset relative to heading. This likely made it difficult for participants to keep the camera facing towards their direction of heading.
4. EXPERIMENT 2: GAZE AND STEERING DURING OBSTACLE AVOIDANCE

As explained in Chapter 1, while driving automobiles, people tend to fixate locations along their future path. However, to drive safely, humans must also avoid obstacles that would otherwise hinder their movement or result in injury. For example, the behavior of a leading car during an overtaking maneuver has been shown to influence the visual search patterns of drivers (Zhang et al., 2016). While overtaking a leading car, increasing the speed of the lead car resulted in more time spent looking towards the destination lane. Overall, their finding showed that 65% of driver gaze time was spent shifting gaze between their current lane and the destination lane (Zhang et al., 2016). In another study, Rothkopf, Ballard, and Hayhoe (2007) had participants complete a series of tasks, such as walking down a sidewalk, picking up litter, and avoiding obstacles. Their results showed that the task being completed influenced the object that gaze was oriented towards and the amount of time spent looking at that object. Furthermore, they found the distribution of gaze position on objects to be different depending on whether the goal was to avoid the object or move towards it (i.e., pick up the litter). When moving towards a target participants looked towards the center of the object, but when avoiding an obstacle, they looked at the object’s edges.

Past research has focused on many aspects of visual controlled locomotion but have rarely explored both obstacle avoidance and path following simultaneously. Where do humans look while performing both tasks? This experiment consisted of two conditions (see Figure 20) to answer this question. The first condition (referred to as Sparse Trees) is similar to the Path Only condition used in Experiment 1, except that the trees were placed further from the path so that
collisions were nearly impossible. The second condition (Dense Trees) included trees as obstacles directly on or close to the path that pilots must avoid to successfully complete laps around the racecourse. Obstacles are made up of varying types of trees that may lie on or close to the path, have overhanging branches, and may occlude a participant’s view of the path.

Experiment 1 provided evidence in favor of a gaze strategy where pilots look where they want to go. The aim of this experiment was to investigate if the findings from the Path Only condition of Experiment 1, where the predominant strategy was to orient gaze towards points along the future path, generalizes to situations where there are obstacles directly on or near the path. Participants could use a gaze strategy that is dominated by eye movements towards obstacles that lie near their future trajectory. This could suggest a gaze strategy that is used to perceive the direction of object motion relative to their movement to make avoidance maneuvers. It is also possible that they use a strategy where they look towards points near where they desire to fly their vehicle (i.e., points along or near the path). This would suggest a gaze strategy similar to what has been found on studies of driving automobiles, such as the waypoint fixating account (Lappi & Mole, 2018).

4.1 METHODS

4.1.1 PARTICIPANTS

Participants were students recruited from Rensselaer Polytechnic Institute (18-29 years old) with valid drivers’ licenses and vision corrected to near normal (contact lenses only). Participants (N=3) were recruited from the pool of participants who had successfully completed Experiment 1. As explained in Chapter 3, The small sample size was a result of a difficult recruiting process due to quadcopter flying being a niche hobby. This issue was compounded further by the COVID-19 pandemic, further limiting the size of the potential participant pool.
This experiment used the same hardware, setup, and method for creating the experimental conditions that were used Experiment 1. I created two conditions, both of which used a variation of the same looped forest path environment from Experiment 1. The first condition was a sparse trees condition, composed of no trees near the path (see the left panel in Figure 18). The trees were placed far from the path (approximately 8 meters on average) so that there was no chance of collisions. The second condition was a dense tree condition, which was made up of a dense forest surrounding the path. In this condition, trees often fell on the edges or directly in the middle of the path (see Figure 18) so that avoidance maneuvers around the trees was necessary to successfully complete laps around the course. To make sure that too many trees did not fall directly on the path or cover the entire path, their placement was modified by hand after the initial random tree position placement. For a description of the process for random tree placement refer to Chapter 2.
4.1.3 PROCEDURE AND DESIGN

The procedure for conducting the experiment was identical to Experiment 1 (see Chapters 2 and 3 for details). Participants completed five blocks of testing after an initial exploration period in an open environment consisting of a forest and a path to become familiar with the controls and flight dynamics (See Figure 18 for examples of the conditions). Within each of the testing blocks the participants completed two laps of the two conditions for a total of 4 laps per block (see Figure 19). Consecutive blocks reversed the order of the conditions to minimize ordering effects. In the last block, the participants completed laps in the same conditions but in the reverse direction through the racecourse. Participants were asked to recalibrate and perform an assessment of each recalibration before completing 2 laps of each condition within a block, resulting in 2 recalibrations per block.
4.1.4 ANALYSES

All data was post-processed using Pupil-Labs editor, C# (Unity), and MATLAB. Data was analyzed using MATLAB and R. The eye tracker (120Hz per eye) provides several gaze locations per eye at each recorded quadcopter position which was logged at 60Hz. Eye tracking data was filtered using the same method used in Experiment 1, by taking only gaze location measures with corresponding pupil labs confidence threshold above 0.6. The gaze location within a frame was then averaged to minimize the influence of gaze estimation error. All mixed linear models were created in the LME4 package in the R programming language.

4.2 RESULTS AND DISCUSSION

4.2.1 GENERAL MEASURES OF PERFORMANCE

The first set of analyses sought to answer the question, how well did participants perform this task? Participants flew the vehicle on average 16m/s per lap, with a higher speed observed in the Sparse Trees condition ($F(1,2)=1.90, p=0.004, \eta^2=0.49$). It took approximately 50 seconds on average to complete a lap, with lower average time to complete a lap in the Sparse Trees ($F(1,2) = 2.517, p=0.011, \eta^2 = 0.56$). Participants were able to stay relatively close to the center of the path with an average absolute deviation of approximately 1m, with an average higher path deviation in the Sparse Trees condition ($F(1,2)=1.90, p<0.001, \eta^2=0.49$). Collisions were quite rare, with less than two collisions per lap occurring on average in both conditions with no
significant difference \(F=(1,2)=1.639, p=0.111, \eta^2 = 0.45\). Figure 20 shows the performance measures across conditions (left) and blocks (right). Performance was largely consistent within a
condition and across blocks.

4.2.2 GAZE BEHAVIOR

The analyses reported in this section were performed to determine where pilots look when following a path while avoiding obstacles and how the presence of dense obstacles (i.e., trees) influenced gaze behavior. Do pilots continue to look towards the path outline, or does gaze often shift to obstacles? The categories of objects toward which participants could orient their gaze in Experiment 2 include trees near the quadcopter, the path, and the terrain. Similar to Experiment 1, trees near the quadcopter are considered those within 34.2m of the vehicle and

![Figure 21](image.png)

Figure 21: Average gaze distributions to different object categories across conditions.
terrain was considered the ground other than the path and trees far from vehicle (i.e., two average hoop lengths from Experiment 1). The majority of time was spent with gaze towards the terrain, with no significant difference between the conditions ($F(1,2) = 8.875, p=0.097, \eta^2=0.82$). Participants spent about a quarter of the time per block looking at the path with no difference between conditions ($F(1,2) = 1.520, p=0.323, \eta^2 = 0.43$). The remaining time was spent looking at trees near the quadcopter with no difference between conditions ($F(1,2) = 15.750, p=0.058, \eta^2 = 0.89$). Figure 21 depicts the average percentages of time spent looking at the different object categories across conditions. No differences were found in the distribution of gaze objects across blocks.

Similar to Experiment 1, I will next integrate these findings into a cohesive account of where participants looked in each condition. First, the pattern of gaze in each of the conditions were largely the same with minor differences. Participants spent a large percentage time looking towards the terrain (~60%), with less time looking towards the path (~25%) and nearby trees (~10%). In Experiment 1, I found that a large portion of time spent looking towards the terrain in the Path Only condition was close to the path (within 4 visual degrees). Recreating that analysis, I found a large percentage of time spent looking at the path or the ground within 4 visual degrees of the path center in both the Dense Trees ($mean=62.5\%$, $CI=[56.3, 68.7]$) and Sparse Trees ($mean=69.8\%$, $CI=[63.4, 76.2]$) conditions. On average, the visual angle between gaze locations on the ground terrain and the center of the path was 2.5 degrees in both conditions. Overall, the dominant gaze behavior was looking toward the ground plane at or near the path, which is similar to the gaze behavior of automobile drivers. However, similar to the Path Only condition from Experiment 1, participants also intermittently looked at nearby trees and the surrounding terrain.
I will next examine the distribution of distances from the quadcopter to the fixated object. Overall, gaze distance followed a similar pattern between conditions, which can be seen in the gaze distance densities detailed in Figure 22. As was discussed in the previous paragraph, the majority of time was spent looking towards or near the path. How far ahead on the path do pilots orient their gaze? When gaze fell on the path, participants looked to points ahead of the vehicle 12m on average across both conditions. It took participants roughly 0.7-0.8 seconds to reach their closest position to that point (See Figure 23). The analyses yielded no significant differences in either distance ($F(1, 2) = 0.251, p = 0.666, \eta^2 = 0.11$) or time ($F(1, 2) = 0.060, p = 0.83$).

![Figure 22](image-url) (Left) Overall gaze distance probability densities. (Middle and Right) Gaze distance densities across both conditions. Black lines represent densities for one block from each participant. The yellow line is the overall mean and the red lines are one standard deviation.
Furthermore, participants passed relatively close by the positions (~3m lateral deviation) along the path towards which they previously oriented their gaze.

4.3 EXPERIMENT SUMMARY

Participant ability to complete the task was consistent with the performance observed in the Path only condition from Experiment 1, except that participant flew faster and completed the blocks slightly faster. This is likely due to their familiarity with the task from completing Experiment 1. Participants expressed comfort with the controls, but like in Experiment 1, some
participants stated noticeable differences between the control scheme implemented on the vehicle and real-world quadcopter flight.

The gaze and steering behavior observed in both conditions in Experiment 2 were also consistent with the gaze and steering observed in the Path Only condition in Experiment 1. This includes the time spent looking towards the terrain close to the path, the time to complete a lap, the average speed, and the average path deviation.

The most interesting result from this experiment is that in the dense tree condition there was no significant difference in the amount of time looking at the path or trees compared to the sparse tree condition. This is surprising given that several of the trees were placed directly in the center of the path in this condition. This finding shows that even in the presence of obstacles, the path took gaze priority and provides further evidence that pilots orient gaze towards the direction they will soon move (i.e., towards points along the path). Another interesting result is that in the dense tree condition the average absolute path deviation from the center was significantly different from the sparse tree condition (see Figure 21). The average deviation was instead less in the dense tree condition. This is likely due to the trees forcing participants to fly in specific locations and providing a sort of artificial boundary to fly through.

Given the lack of time spent gazing towards trees (See Figure 21) in the Dense Trees condition, this suggests that participants either monitored them in their periphery or spent very short gaze periods focusing on them. It is also possible, that like Experiment 1, in this experiment gaze was used to anchor on the path beyond a potential obstacle and then make steering adjustments to pass by the obstacle. This could be analogous to looking through a hoop at distant points and then using that to make steering adjustments through the center of the hoop (See Chapter 3). This may be the subject of analyses in the future. Another possibility worth
noting is that participants may not have felt that the presence of the trees posed any risk. If a tree is collided with, then the vehicle would bounce off and they would be able to continue flying. This may need to be accounted for in future experiments.
5. GENERAL DISCUSSION

5.1 SUMMARY OF MAIN FINDINGS

This dissertation explored how quadcopter pilots use visual information to guide their movements along trajectories through 3D space in various environmental conditions. Studying the behavior of quadcopter pilots required the development of an entirely new experimental paradigm that allowed for a high level of flexibility, had realistic flight dynamics, provided easy manipulation of the environment and visual information, and the ability to log relevant data. An experimental paradigm for quadcopter flight that also had these features did not exist until I developed the one described in Chapter 2. Additionally, in Chapter 2 I described how the paradigm was developed so that future researchers can create new experiments, process data, and create detailed visualizations of gaze and steering behavior from participant data.

Due to the limited amount of past research on this topic, this study was conducted first and foremost to develop a description of how quadcopter pilots coordinate gaze and steering while path following, passing through waypoints, and/or avoiding obstacles. There are three main takeaways from the present set of experiments. The results show that the coordination of gaze and steering during quadcopter flight share some similarities with gaze during other forms of locomotion under some conditions. Specifically, quadcopter pilots use a gaze strategy where they orient gaze towards where they desire to steer the vehicle and not gaze towards locations to avoid (i.e., obstacles). Similarly, the presence of hoops in Experiment 1 produced gaze strategies that were dominated by gaze through the center of the hoop, a waypoint through which they desire to pass. A possible reason for this is that eye movements towards the desired location of vehicle
movement produce more accurate steering compared to a strategy where gaze is oriented towards other objects, such as obstacles (i.e., hoop edges or trees).

The second main takeaway comes from Experiment 2, which allowed for the development of a description of gaze and steering during obstacle avoidance through cluttered 3D spaces. The findings are consistent with the gaze strategy observed in Experiment 1. The results showed that even in the presence of dense obstacles, participant gaze was oriented towards the path and terrain, or in other words the direction they desired to move. Little time was spent looking at obstacles (<20% of the time in both conditions), as shown by the gaze object category distributions. This last finding is interesting because previous research has shown obstacle avoidance while walking leads to gaze towards the edges of those obstacles (Rothkopf, Ballard, & Hayhoe, 2007). The present finding is understandable, however, because looking towards obstacles while moving at high speeds can result in steering towards those locations. Instead, looking towards desirable locations towards which to steer may result in more accurate steering. The time participants spent looking towards obstacles was brief. Therefore, intermittent eye movements towards regions in space to avoid may be all that was necessary for efficient avoidance maneuvers.

The finding that pilots look towards where they desire to move the vehicle is largely consistent with past research on steering automobiles. The waypoint fixating account predicts that drivers will make eye movements to waypoints along their future path (Lappi & Mole, 2018). In the present study, pilots made eye movements towards waypoints (i.e., hoops) above the ground in conditions that contain them and towards the path in the conditions that only have a path (e.g., Path Only, Sparse Trees, and Dense Trees). These locations are analogous to the
waypoints described in the waypoint fixating account, supporting the theory that looking where you move is important for high-speed flight in 3D.

The results from both experiments are also consistent with the idea that people spend more time orienting gaze towards task-relevant information (Rothkopf, Ballard, & Hayhoe, 2007; Ballard & Hayhoe, 2009; Hayhoe & Ballard, 2014). In each of the present experiments, and each of its conditions, the task that pilots needed to complete changed slightly. Experiment 1 consisted of either a path following task or hoops to fly through. In Experiment 2, pilots needed to either follow a path or follow a path while avoiding obstacles. As the task changed, the gaze patterns often changed too. This can be seen in the gaze object category distributions from Experiment 1, which showed gaze oriented towards the path in the Path Only condition and gaze oriented towards hoops in the hoop-containing conditions. In Experiment 2, gaze was consistently oriented towards the path or nearby terrain, even in the presence of obstacles. This may be because the task constraints were not sufficient to shift gaze towards obstacles, or it may be because the participants are considering the path following aspect of the task to take precedence over the obstacle avoidance.

The last main takeaway is that during flight, pilots adapt their current trajectory in accordance with upcoming environmental conditions, suggesting anticipatory steering through the nearest hoop to align themselves with the subsequent hoop. Evidence for this was shown by the linear models in Experiment 1, which were able to predict approach angle and relative orientation of the quadcopter as it passed through a hoop based on the relationship between the nearest and subsequent hoop. Further evidence in support of anticipatory actions was that pilots sometimes made eye movements to hoop N+1 before passing through hoop N. These two findings suggest an overall anticipatory strategy for navigating through complex 3D spaces.
Some of these findings are consistent with numerous studies which have provided evidence for look-ahead eye movements (Mennie et al., 2007; Lehtonen et al., 2013; Wilkie, Wann, & Allison 2008). The evidence I provided in support of anticipatory steering is largely consistent with past research on steering, which also provided weak evidence for anticipatory steering in similar tasks (Wilkie, Wann, & Allison 2008). Though the evidence provided by the current experiment was stronger than what Wilkie, Wann, & Allison found, the effect was still overall relatively weak. In Chapter 3, I discussed that the observed anticipatory steering may be due to the physics of quadcopter flight, which make it necessary to adapt current trajectories due to the influence of inertia on the vehicle.

5.2 LIMITATIONS OF THE PRESENT EXPERIMENTS

The present experiments successfully captured the behavior during the intended task, but they were not without some issues. The quadcopter physics and control scheme were noticeably different compared to real world racing quadcopters, which required participants to adapt their previously learned controls. The biggest difference between them being how thrust is applied to the vehicle as was discussed in Chapter 2 Discussion. Those facts notwithstanding, participants were able to quickly learn the new controls and efficiently fly the vehicle through the virtual environment, but in future experiments researchers using this paradigm may want to use updated flight controls.

Recruiting participants was a major challenge, resulting in a relatively small sample size in the first experiment (N=6), and an even smaller sample size (N=3) in the second experiment. The main reason for our recruitment issue was due to a small participant pool from which we could recruit participants. The experiment required individuals to have experience flying first-person viewed quadcopters (via HMD) either in real world or simulation. This is a niche hobby which
significantly cut down on the available participant pool. This was compounded further due to the COVID-19 pandemic, which only allowed individuals within the Rensselaer Polytechnic Institute pandemic testing pool could participate in the experiment. This meant that drone racing clubs, or hobby groups, that exist outside of the RPI community were unable to participate. Future studies should not have this issue.

Simulator sickness was a minor problem that arose in pilot testing, as several participants with little to no experience flying quadcopters had to end approximately 15 to 20 minutes into the experiment. After observing this effect, subsequent participants were screened for vulnerability to simulator sickness and only participants with 10 hours of flight time experience in simulation were considered for the experiment. Simulator sickness was not an issue for any of the participants in Experiment 1 and 2. In virtual tasks, simulator sickness is a common occurrence due to the visual cue matching conflict between moving in virtual space with the actual movement of the head and body. In the present experiment, this conflict is more apparent because tracking the position of the HMD was turned off to mirror actual quadcopter FPV goggles.

To combat these issues, future experiments will adjust the paradigm to include a wider range of participant skills and largely remove the possibility of simulator sickness by having the task completed on a monitor rather than on a head-mounted display. This will remove the visual cue matching problem for simulator sickness. One of the reasons that experienced quadcopter pilots were necessary for the present experiment is because they should be less likely to experience simulator sickness due to their familiarity with using an HMD during flight. Therefore, using a monitor will allow less experienced participants to complete a similar experiment. Using less experienced pilots can also afford us the opportunity to train them to control the vehicle and investigate how people learn to fly quadcopters.
Other studies have used screen-based eye-tracking paradigms to investigate visually guided locomotion, such as Pfeiffer and Scaramuzza (2021) and Tuhkanen et. al., (2021). Implementing the current experiment using a monitor-based solution should not be overtly difficult, although there are some eye-tracking issues that need to be solved first to allow such a setup to maintain the same level of flexibility as the current HMD-based setup. For example, determining where on a screen a participant is looking, and then translating that point to what object in the virtual world gaze is oriented towards is not straight forward. Past studies have used areas of interest in screen coordinates to run analyses, which offer less precise measures compared to what was offered from the VR setup currently in place. To solve this problem, the position of all the objects in the virtual world need to be known and the gaze location in screen coordinates needs to be very accurate. The gaze position on the screen can be mapped into the virtual world-space, and a vector can be cast to collide with the object that lies beneath gaze, recovering the gaze object and the position of gaze in world coordinates. However, the translation between screen to world coordinates will introduce some amount of error and will be further complicated if a mobile eye tracking device is used due to free movement of the head and any slippage of the device on the head. This issue, however, was not present in the current paradigm because pupil labs’ software maps gaze position directly into 3D world coordinates. This allowed us to use the estimated gaze location in world-space to determine the object in the world towards which gaze was oriented. The 3D point in space was used to cast a vector into the VR environment until it collides with an object, and that object was recorded as the gaze target at each frame during experimentation. This process could be completed during data recording or during post processing.
A fourth potential issue was eye-tracker accuracy. Many of the analyses required high confidence in the eye tracking recording. While the eye tracking recordings appeared accurate, it is a known issue that high accuracy eye tracking can be difficult due to slippage of the HMD on the head, lighting conditions, and other factors. Experiments 1 and 2 took precautions to prevent these issues, but participant accuracy still varied. Furthermore, in the present experiments, participants often made eye movements to points well ahead of their current position in the environment. Therefore, any decrease in accuracy could result in objects or distances being recorded that differ from the actual participant behavior. These are known and common problems that arise in experiments that use eye tracking. Hopefully, as mobile eye tracking systems become even more sophisticated, these will become less of an issue for researchers.

5.3 FUTURE DIRECTIONS

The findings from Experiments 1 and 2 led to an idea that was not considered before running the experiments, namely that the independent control of yaw from translation may lead to differences in gaze and steering strategies used by pilots compared to actors performing other forms of locomotion. Under such conditions it may be necessary to deal with those rotations to maintain stable control of the vehicle. The last analysis from Experiment 1 investigated if pilots try to minimize the offset between the forward-facing camera and instantaneous heading. The evidence showed that participants typically maintain an offset of zero degrees with a standard deviation of -20 to 20 degrees between the camera and heading, and that the rate of rotation can vary widely. This suggests that while pilots are mostly keeping the camera facing their direction of heading, there are often large offsets (i.e., > 10 degrees). This coupled with the fact that the rate of rotation is often nonzero suggests that rotations are very common and frequently not perfectly aligned with the direction of heading.
As discussed in Chapter 1, past research has shown that people can judge their direction of self-motion based on the available pattern of optic flow (Warren et al., 2001) and in the presence of rotations from pursuit eye movements (Warren & Hannon, 1988). Further research has shown that rotation can bias the perception of heading if those rotations were simulated (Banks et al., 1996), although a follow-up study showed that perception of heading remains robust to simulated rotation as long as there is sufficient visual depth, motion parallax, and environmental textures (Li & Warren, 2001). In all these studies, however, participants do not have active control of their movement, instead making judgments about their direction of self-motion. In the present study, participants maintain active control of their vehicle and their control is quite different from other forms of locomotion due to how they control yaw.

Given the results from the last analysis in Experiment 1, questions remain about how, if at all, pilots deal with their changing yaw. One possibility is that pilots make eye movements that produce retinal flow to compensate for the added rotational components in the optic flow array. This would be reflected in pilots making eye movements to points in the environment at sufficient distances so that the rate of rotation of the eye is equal and opposite to the rate of yaw. Theoretically, a fixation to a point at a far enough distance during rotation should completely remove the rotational components in the optic flow array and result in a purely radial pattern. If the participant made these kinds of eye movements, they were for short durations and may depend on a variety of factors, such as proximity to the hoop, previous steering commands, collisions, and speed. To investigate this possibility, another experiment or set of analyses will need to be developed in the future.

Due to the flexibility of the experimental paradigm, researchers can conduct experiments that manipulate any number of aspects of the flight task, from the control of the quadcopter to the
available visual information. For example, a future study could create a scenario where yaw remains oriented towards instantaneous heading to test differences in eye movements compared to when yaw is controlled independently. Furthermore, this paradigm also could be modified to conduct studies on judgments of heading by removing active control. I contend that all these points above demonstrate how the present quadcopter flight experimental paradigm is an excellent candidate task space for investigating the influence of rotation on the perception of heading, the control of steering, and the control of gaze during visually guided locomotion.
REFERENCES


*Proc. IEEE Workshop on Visual Behaviors*, Seattle, WA. Retrieved May 5, 2022, 


