Predictive Power of Pupil Dynamics in a Team Based Virtual Reality Task

Yinuo Qin* Weijia Zhang* Richard Lee* Xiaoxiao Sun* Paul Sajda*

Columbia University

Abstract

Assessing and tracking physiological and cognitive states of multiple individuals interacting in virtual environments is of increasing interest to the virtual reality (VR) community. In this paper, we describe a team-based VR task termed the Apollo Distributed Control Task (ADCT), where individuals, via the single independent degree-of-freedom control and limited environmental views, must work together to guide a virtual spacecraft back to Earth. Novel to the experiment is that 1) we simultaneously collect multiple physiological measures including electroencephalography (EEG), pupillometry, speech signals, and individual's actions, 2) we regulate the the difficulty of the task and the type of communication between the teammates. Focusing on the analysis of pupil dynamics, which have been linked to a number of cognitive and physiological processes such as arousal, cognitive control, and working memory, we find that pupil diameter changes are predictive of multiple task-related dimensions, including the difficulty of the task, the role of the team member, and the type of communication.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality;

1 Introduction

In the 1995 motion picture *Apollo 13* there is a scene [4] that shows the three astronauts working together to execute a manual burn of the engines to adjust their re-entry angle. In this paper, we describe the VR platform to recreate the *Apollo 13* re-entry maneuver. While a triad team perform the re-entry task, we simultaneously record physiological and behavioral data, including electroencephalography (EEG), pupillometry, eye gaze, controller input and speech. In this paper we focus on the analysis of pupil dynamics, as it has been directly linked to arousal [8]. The state of arousal can significantly affect a person's ability to make optimal decisions, judgments, and actions in real-world dynamic environments [3]. In this paper we investigate how pupil dynamics, as an index of arousal, are predictive of task difficulty, role in task, and type of communication.

2 Experiment

We built a virtual environment, which we term the Apollo-Distributed Control Task (ADCT). In the experiment, three participants, with different partial view of the environment, are asked to coordinate, collaborate and cooperate with each other to navigate a spacecraft to reach a final destination via a re-entry path marked by transparent red rings (Figure 1(**A**)) under a time-constraint. All experiments were approved by the Columbia University's Institutional Review Board (IRB) for human subjects research.

Setup As shown in Figure 1(**B**) for each triad team, the three roles are *yaw*, *pitch*, *thrust*, with each role having different views of the environment as well as controlling different degrees of freedom of the spacecraft. Additionally, there are three modes of communication among team members, which are no communication, command word communication, and free communication. The task performances are categorized as pass or failed to pass a ring.

*e-mail: {yinuo.qin, wz2540, rtl2118, xs2362, ps629}@columbia.edu

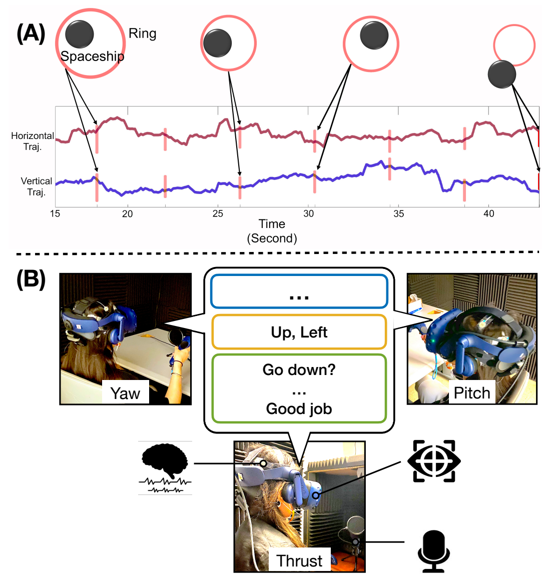

Figure 1: **(A)**An example trial of spacecraft trajectory. The curves are the horizontal and vertical trajectories of the spacecraft. Each black circle and red ring on top indicates the relative position and the size of the spacecraft to a ring (easy, medium, medium, hard). The team failed to pass the last ring. **(B)**A triad team performing the experiment under three different communication modes. Each team member was instrumented with EEG, eye-tracker, and microphone.

Participants There were in total 63 participants (35 males, 28 females) between 19 to 37 years old ($M = 23.79$, $SD = 3.22$) in our study. Participants were split into groups of three, and each group participated in the experiment three times on three distinct days. All participants for this experiment were recruited through voluntary participation forms. Three groups could not complete all sessions due to COVID-19, so their data was omitted from the final dataset.

Apparatus Figure 1(**B**) shows our setup. Each participant wore an HTC VIVE Pro Eye system with a built-in Tobii eye-tracking system. Participants communicated with one other only through microphones and *Mumble* software. EEG signals were recorded for each subject using a B-Alert X24 system with 20 Ag/AgCl electrodes. We connected the HMDs to three different client desktops with an Intel Core i9 CPU and an NVIDIA RTX 2070 Super GPU, and used a server desktop with the same CUP and an NVIDIA RTX 2080 Super GPU. All desktops were connected by a local WiFi network with a 2.6 Gbps router. We synchronized all streams of data with Lab Streaming Layer [7].

Data Processing After each session, we loaded and processed the data using *Matlab* by first removing artifacts. We then interpolated missing data in the time series using a cubic spline. We then computed the Z-score and epoched the data. Each epoch included data from 2 seconds before pass the ring to 2 seconds after pass the ring. All data were then resampled to 100 Hz.

Data Analysis We used a generalized additive model (GAM) to model the changes of participants' pupil sizes across time, under different difficulty levels, roles, communication modes, and task

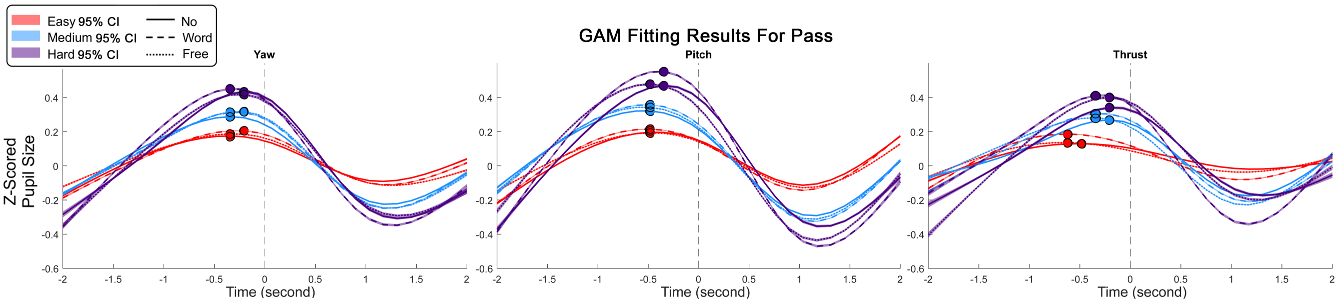

Figure 2: The predicted Z-scored pupil sizes (with 95% *CI*) from the results of all the GAM analyses. Results are plotted as a function of role. Within each role, curves having different colors indicate different difficulty levels with different line styles indicating different communication modes. The solid filled dot on each fitting curve marks the maximum of that curve.

performances. We hypothesized that the Z-scored pupil size would change across time differently as a function of task difficulty levels, subjects' roles, and communication modes. Our GAM analysis is computed using the *R* package *mgcv* [9] in *R studio* (version 4.0.6).

3 Results

Pupil Changes as a Function of Task Difficulty We first evaluated pupil dynamics as a function of task difficulty. As shown in Figure 2, we quantified the changes using the model with independent variable *difficulty* and modeled in both intercept and the smooth term to compare the differences between different task difficulty levels. These curves as significantly different ($p < 10^{-10}$) with the amplitude for the hard tasks being greater than the medium task, which in turn is greater than the easy task.

Pupil Changes as a Function of Role Figure 2 also indicates that there are differences in pupil dynamics between roles of the team members in the task. Using the GAM model with *role* as an independent variable, we find significant differences in Z-scored pupil sizes with different roles ($p < 10^{-5}$). In general, the *thrust* player has the smallest amplitude changes, indicating the *yaw* players, and *pitch* player would have the highest arousal level.

Pupil Changes as a Function of Communication Mode Similar to different difficulty levels and roles, the GAM we used to test the differences of communication modes has *communication* as independent variable. We find that the Z-scored pupil sizes under three different communication modes follow different trends over time. From Figure 2, we can see that these differences among communication mode are significant ($p < 10^{-5}$).

4 Discussion

The Effect of Task Difficulty on Pupil Dynamics Under the hypothesis that pupil size is indexing arousal, greater arousal dynamics in more difficult tasks indicate that greater increases in arousal when approaching a harder ring and greater recovery when successfully passing the ring. These results point to the pupil as a "window" into internal brain states that relate to arousal and cognitive load.

The Effect of Role on Pupil Dynamics Moreover, our findings suggest that the average pupil size is significantly less for *thrust* relative to the other two roles. One potential explanation is that success in the task depends less on the speed control (thrust) than on the position control (yaw and pitch).

The Effect of Communication Mode on Pupil Dynamics From the GAM analysis in Figure 2, we observed a small but significant difference in the pupil dynamics as a function of the mode of communication. Analysis of the other data streams, including the EEG and speech, may yield additional insight on whether there are cognitive signatures that are predictive of the best communication mode for a given team.

Pupil Size and the Affect of Illumination and Accommodation Previous work has shown that luminance changes in a VE strongly influence pupil sizes [1, 2, 6], which increasing luminance lead to pupil constricts. In our VE, the luminance gradually increases as participants approach to a ring. However, we observed that participants' pupils dilate when approaching a ring. In terms of the effect of accommodation on pupil size, previous studies [5] have shown that pupil sizes decrease when focusing on objects moving toward the subject. In our experiment, the rings and Earth continuously move toward participants and we do not see any systematic decrease of pupil size from the beginning of the trial to the end. Thus, there is little evidence that luminance or accommodation had a significant effect on the pupil dynamics we observed, and we therefore conclude that factors such as cognitive arousal are driving the dynamics.

5 Conclusion

The Apollo-Distributed Control Task (ADCT) task we constructed was designed to study neural and physiological signatures of team dynamics during an immersive and stressful cooperative task. The system design and engineering of the VE was non-trivial given it was a multi-person experiment built in VR and allowing for simultaneous recording of multiple data streams including EEG, pupil size, gaze, speech, and participants' behavioral input via controllers. This multi-dimensional dataset provides the potential to better understand team behavior and how it relates to physiological and neurological variables.

Acknowledgments

This work was supported by funding from the Army Research Laboratory's STRONG Program (W911NF-19-2-0139, W911NF-19-2-0135, W911NF-21-2-0125) the National Science Foundation (IIS-1816363, OIA-1934968) and a Vannevar Bush Faculty Fellowship from the US Department of Defense (N00014-20-1-2027).

References

[1] Y.-G. Cherng, T. Baird, J.-T. Chen, and C.-A. Wang. Background luminance effects on pupil size associated with emotion and saccade preparation. *Scientific reports*, 10(1):1–11, 2020.

[2] C. Ellis. The pupillary light reflex in normal subjects. *British Journal of Ophthalmology*, 65(11):754–759, 1981.

[3] J. Faller, J. Cummings, S. Saproo, and P. Sajda. Regulation of arousal via online neurofeedback improves human performance in a demanding sensory-motor task. *Proceedings of the National Academy of Sciences*, 116(13):6482–6490, 2019.

[4] B. Grazer. *Apollo 13*. Universal City Studios, Inc., 1995.

[5] S. Kasthurirangan and A. Glasser. Characteristics of pupil responses during far-to-near and near-to-far accommodation. *Ophthalmic and Physiological Optics*, 25(4):328–339, 2005.

[6] S. Koorathota, K. Thakoor, L. Hong, Y. Mao, P. Adelman, and P. Sajda. A recurrent neural network for attenuating non-cognitive components of pupil dynamics. *Frontiers in Psychology*, 12:12, 2021.

[7] C. Kothe. Lab streaming layer (lsl), 2014.

[8] J. Rajkowski. Correlations between locus coeruleus (lc) neural activity, pupil diameter and behavior in monkey support a role of lc in attention. *Soc. Neurosc., Abstract, Washington, DC, 1993*, 1993.

[9] S. N. Wood. *Generalized additive models: an introduction with R*. CRC press, 2017.