Machine Learning for Interaction Analysis in Virtual Reality

YUJIE HUANG, Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, F-44000 Nantes, France

ALEXANDRE BRUCKERT, Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, F-44000 Nantes, France

PATRICK LE CALLET, Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, F-44000 Nantes; Institut universitaire de France (IUF), France

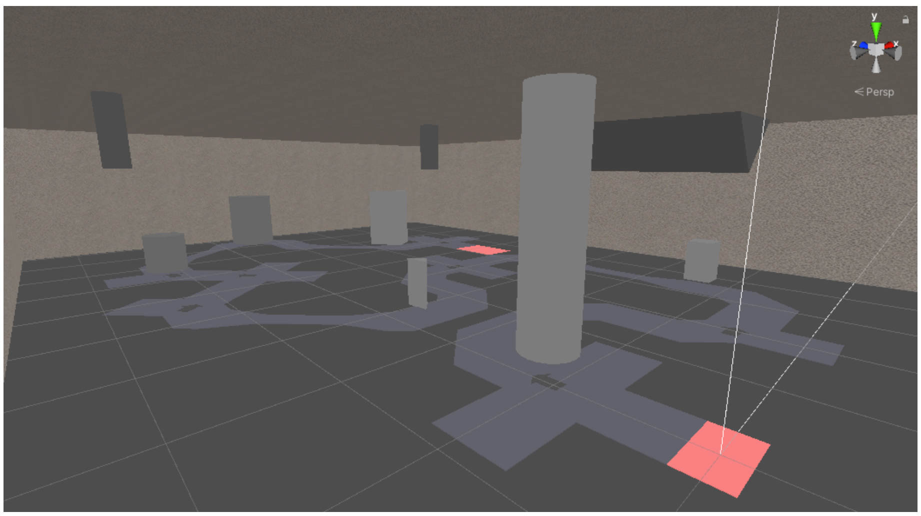

Fig. 1. Test trials are conducted using a virtual labyrinth with obstacles

Functional vision refers to the ability to effectively use visual information to perform everyday tasks such as reading, writing, navigating, or recognizing objects and faces. However, to assess the severity of visual impairments, ophthalmologists mostly rely on tests designed to evaluate the visual function, which refers to a wider range of visual abilities (visual acuity, color vision, contrast sensitivity…) without necessarily considering their practical applications. Thus, there is a crucial need for designing protocols dedicated to evaluating the functional vision for patients with visual impairments, developing appropriate therapies, as well as monitoring the evolution of the patient's quality of life. In this work, we will explore the use of virtual reality (VR) and data mining methods to objectively and subtly assess the impact of treatments on patients' quality of life. By modeling interactions in ecological virtual environments and leveraging advanced data analysis methods, such as stochastic models and graph-based methods, we aim to extract underlying information that can yield robust evaluation results. We expect such evaluations to be an additional tool for physicians to inform the monitoring of their patients, at a relatively low cost.

CCS Concepts: • **Human-centered computing → Virtual reality**; • **Applied computing → Health care information systems**.

Manuscript submitted to ACM

Additional Key Words and Phrases: Functional Vision, Virtual Reality, Data Mining, Orientation and Mobility

ACM Reference Format:
Yujie Huang, Alexandre Bruckert, and Patrick Le Callet. 2018. Machine Learning for Interaction Analysis in Virtual Reality. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 6 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Health-related Quality of Life (QoL) [20] is a widely discussed measure of an individual's well-being in medical research and therapy development. Visual impairment is a significant public health issue that impacts the QoL of individuals across all age groups [6], posing challenges in medical and social care. In France, it is estimated that the number of people suffering from visual impairments exceeds 2% of the total population and is constantly increasing with the aging population. Faced with this reality, substantial investments are made not only in public research but also by the socio-economic sector to develop remediation devices, prosthetics, and particularly promising new gene therapies [14].

However, objectively evaluating the clinical effectiveness of these devices and treatments remains a challenge. Several programs have been discontinued despite patients reporting an improvement in their quality of life. The lack of tools for objectively measuring functional vision may prevent the adoption of vision biotherapies. Unlike visual functions, such as visual acuity and contrast sensitivity, which is related to organ systems [9], functional vision is the ability to perform visual tasks in daily life, and its assessment should be discussed in real-world scenarios [5]. Several physical courses have been developed to assess the mobility of visually impaired individuals [6, 17]. However, there are some shortcomings with these courses, such as the difficulty of modifying visual cues [11], setting up the physical course is time-consuming [17] and costly. Virtual reality (VR) can overcome these limitations and reproduce ecological situations, and the head-mounted display (HMD) can capture real-time kinematic data, making it a suitable choice for assessing functional vision [2, 7].

Going beyond quantitative measures such as time or performance, this thesis aims to model interactions and behaviors in such situations. Combined with advanced algorithmic analysis methods, this approach enables the development of a quantitative assessment of functional vision for longitudinal clinical effectiveness monitoring.

2 RELATED WORK

It is necessary to consider both visual function and functional vision to regulate the rating system for the recovery effect of ophthalmic treatment [8]. Functional vision has quality-of-life-related characteristics that are not completely correlated with visual function. Bennett et al. [5] discussed the disconnection between the two terms, illustrating the importance of assessing both. The evaluation of visual function can be divided into assessing factors such as visual acuity using clinical methods. However, for functional vision, the assessment should focus on real-world tasks [5].

Functional vision is commonly measured using subjective and objective methods. Subjective methods mostly rely on patients providing self-reported assessments by the use of questionnaires or scoring scales [13]. Objective methods are mostly based on the Orientation and Mobility (O&M) test [2]. Geruschat et al. [12] used time to complete the course and the number of errors as metrics to compare the improvement of functional vision. To overcome the variability in different subjects' natural walking speed, the percentage of preferred walking speed (PPWS) was adapted in [17]. To represent the subject's performance, errors were used to assign penalties for both accuracy and speed, an accuracy score and a time score were calculated [2, 6].

The physical mobility O&M courses have several limitations due to environmental conditions and difficulty in reproduction. Virtual reality enables the manipulation of the environment and has been widely used in diagnosing spatial navigation tasks [7]. Compared to traditional paper-and-pencil tests, VR tests offer greater ecological validity [10], despite participants' performances not being entirely the same as real-world behavior [23]. Nonetheless, the validity of using VR tests for diagnostic purposes has been confirmed. Aleman et al. [2] reproduced the idea of [6] in virtual reality and provided proof-of-concept data in support of the usability of virtual reality O&M (VR-O&M); Bennett et al. [4] used an eye tracker integrated virtual environment to analyze the behavior of patients with different type visual impairment. This tool enables researchers to improve scientific monitoring and create a safer and more reproducible experimental environment.

Data mining aims to discover patterns and information from a given set of data. Stochastic models have the potential to extract features or underlying patterns. For instance, Shaily and Mangat [21] proposed using the Hidden Markov Model (HMM) on sequential data for human activity recognition; Ben-Gal et al. [3] used the Markov Chain for mobility modeling. Graphs also prove to be a powerful tool for representing data. Rossi et al. [19] used graph-based method to cluster subjects with the same navigation pattern in VR. With different modeling strategies, related learning methods can be further explored.

3 RESEARCH QUESTIONS

To address problems and challenges in this field, we have defined four research questions (RQs) to investigate:

- RQ1: How can individual participant behavior be characterized when facing the same task?

Facing the same task, each individual may exhibit different behaviors. Characterizing these differences and extracting the implicit information will be the first research question.

- RQ2: How do environmental factors influence participants' behaviors?

Environmental factors, such as luminosity and contrast, can have an impact on QoL [6]. The objective is to determine the factor(s) that influence(s) participants' behavior(s) and how they do so. This information can also serve as indicators for future virtual environment design.

- RQ3: Which behavior(s) represent(s) functional vision?

Assessment methods based solely on the time taken and number of errors in the O&M course are not accurate or subtle enough [12]. To provide a more objective assessment of a patient's functional vision, we're looking to incorporate participants' behaviors. This will improve the reliability and accuracy of the tools. In this stage, we will use the clustering method combined with behavioral information to provide a more robust assessment of the patient's vision.

- RQ4: How to objectively assess functional vision?

This final research question aims to define a learning model using the clues extracted from the previous stages, to evaluate the functional vision improvement under different environmental factors. During this stage, we will try to explore the potential of Graph Neural Network, combined with graph signal processing method [22], to create ultimately an objective functional vision assessment tool.

4 METHODOLOGY

To answer these questions, we have divided the entire research objective into 3 sub-tasks: data collection, behavior modeling, and objective assessment. This section details the planned methodology for each.

4.1 Data collection method

Data collection provides the foundation for this thesis. To complete this sub-task, we are taking advantage of virtual reality to create a safe, flexible, and well-monitored system for conducting the test trials. The initial idea for our virtual environment was inspired by the multi-luminance mobility test (MLMT) [6] and a virtual reality orientation and mobility (VR-O&M) protocol [2]. With the HMD and the virtual environment created by the Unity software, a replicated O&M test can be conducted in our laboratory. We can easily adjust the environmental factors and get richer information during the test.

4.2 Behavior modeling method

To better understand human behavior, high-level processing is also involved [1]. Since the data collected from the virtual environment are time series, each subject has a variable test duration, and a dynamic presentation method is considered. Dynamic Time Warping (DTW) [15] and Hidden Markov Model (HMM) [18] are both suitable methods for sequence representation and comparison. The emergence of graphs as a powerful way to model human behavior data and interactions [16, 22] will also guide the way that we address this modelization.

4.3 Objective assessment method

The objective measurements were typically based on low-level metrics such as time and number of errors [2, 6, 12, 17]. For a more subtle measurement, we propose to use clustering approaches to differentiate between individual behavior models [3]. Graph neural networks are also another tool that we will examine to this end, considering their convincing performances [24].

5 WORK IN PROGRESS AND FUTURE PLANS

In a preliminary phase, data collection was performed in a virtual environment. For a brief introduction of the data collection, all data were collected using HTC Vive Pro Eye. A group of healthy subjects and a group of patients were recruited to participate with informed consent, and the data will be validated by the ethical committee. Two training sessions were conducted before data collection. We kept the main point of the O&M test: navigation and obstacle course, learning from the previous environment design [17], and integrated them into virtual labyrinths with multiple objects placed in different positions. During the test, the subject remained seated and controlled the direction of the avatar by rotating the chair as shown in figure 2, which was a safer and equivalent way to move the avatar. Subjects followed the labyrinth and attempted to destroy each object that they discovered. Both time-series and contextual data were recorded after each trial, providing a rich set of scene and behavioral information.

After the data were collected, some preliminary analyses were performed and the validity of our designed test was verified. We found that between healthy subjects, as the level of luminosity decreased, both the duration of the trial and the number of errors increased; and on average, patients spent more time to complete the required tasks than healthy subjects. In addition, for a better understanding of the scene, some visualization tools for trajectories during the test were developed, as shown in figure 3: the figure on the left represents the elements of the complete scene; the figure on the right shows the direction of the head and the potential visibility of obstacles.

All of these preliminary analyses lay the groundwork for our future research. According to the result of the current stage, we aim to dig deeper into the collected data. The current work is based on the metrics and scoring system proposed by the previous research. In the next working stage, we plan to add more information such as head and hand

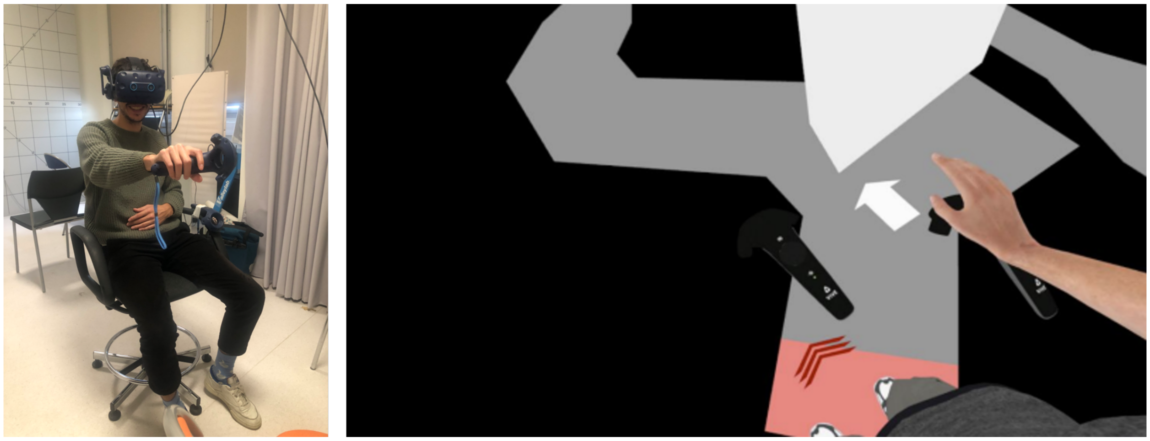

Fig. 2. A subject undergoing a trial and the subject's view in virtual environment

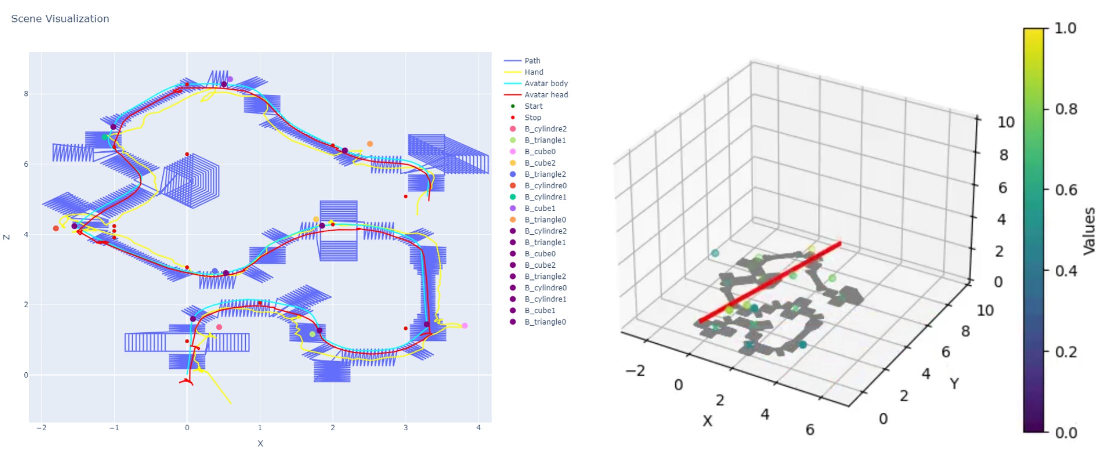

Fig. 3. Scene visualization tools developed during the preliminary phase

movement to improve the performance of the objective assessment tool. Focusing on the collected data, especially on the behavioral time-series, we aim to explore the potential of stochastic models and graph-based models in order to extract the high-level information, which is hypothesized to be better representative of patients' health status.

6 CONCLUSION

In conclusion, this paper briefly introduces my thesis, entitled "Machine Learning for Interaction Analysis in Virtual Reality". It highlights the importance of overcoming objective and subtle functional vision assessment challenges. By leveraging virtual reality and data mining methods, we aim to develop a more nuanced assessment tool to provide a more accurate evaluation result. This thesis' outcomes will improve the reliability of clinical diagnosis and ultimately the quality of life for patients with visual impairment.

REFERENCES

[1] Jake K Aggarwal and Sangho Park. 2004. Human motion: Modeling and recognition of actions and interactions. In *Proceedings. 2nd International Symposium on 3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004.* IEEE, 640–647.

[2] Tomas S Aleman, Alexander J Miller, Katherine H Maguire, Elena M Aleman, Leona W Serrano, Keli B O'Connor, Emma C Bedoukian, Bart P Leroy, Albert M Maguire, and Jean Bennett. 2021. A virtual reality orientation and mobility test for inherited retinal degenerations: testing a proof-of-concept after gene therapy. *Clinical Ophthalmology* (2021), 939–952.

[3] Irad Ben-Gal, Shahar Weinstock, Gonen Singer, and Nicholas Bambos. 2019. Clustering users by their mobility behavioral patterns. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 13, 4 (2019), 1–28.

[4] Christopher R Bennett, Emma S Bailin, Timothy K Gottlieb, Corinna M Bauer, Peter J Bex, and Lotfi B Merabet. 2018. Virtual reality based assessment of static object visual search in ocular compared to cerebral visual impairment. In *Universal Access in Human-Computer Interaction. Virtual, Augmented, and Intelligent Environments: 12th International Conference, UAHCI 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15-20, 2018, Proceedings, Part II 12.* Springer, 28–38.

[5] Christopher R Bennett, Peter J Bex, Corinna M Bauer, and Lotfi B Merabet. 2019. The assessment of visual function and functional vision. In *Seminars in pediatric neurology*, Vol. 31. Elsevier, 30–40.

[6] Daniel C Chung, Sarah McCague, Zi-Fan Yu, Satha Thill, Julie DiStefano-Pappas, Jean Bennett, Dominique Cross, Kathleen Marshall, Jennifer Wellman, and Katherine A High. 2018. Novel mobility test to assess functional vision in patients with inherited retinal dystrophies. *Clinical & experimental ophthalmology* 46, 3 (2018), 247–259.

[7] Michel Cogné, Mathieu Taillade, Bernard N'Kaoua, Annabelle Tarruella, Evelyne Klinger, Florian Larrue, Hélène Sauzéon, P-A Joseph, and Eric Sorita. 2017. The contribution of virtual reality to the diagnosis of spatial navigation disorders and to the study of the role of navigational aids: A systematic literature review. *Annals of physical and rehabilitation medicine* 60, 3 (2017), 164–176.

[8] August Colenbrander. 2005. Visual functions and functional vision. In *International Congress Series*, Vol. 1282. Elsevier, 482–486.

[9] August Colenbrander. 2010. Assessment of functional vision and its rehabilitation. *Acta ophthalmologica* 88, 2 (2010), 163–173.

[10] Raquel Quimas Molina da Costa, José Eduardo Pompeu, Emerson Moretto, Juliana Magalhães Silva, Michelle Didone Dos Santos, Ricardo Nitrini, and Sonia Maria Dozzi Brucki. 2022. Two immersive virtual reality tasks for the assessment of spatial orientation in older adults with and without cognitive impairment: concurrent validity, group comparison, and accuracy results. *Journal of the International Neuropsychological Society* 28, 5 (2022), 460–472.

[11] Fábio B Daga, Eduardo Macagno, Cory Stevenson, Ahmed Elhosseiny, Alberto Diniz-Filho, Erwin R Boer, Jürgen Schulze, and Felipe A Medeiros. 2017. Wayfinding and glaucoma: a virtual reality experiment. *Investigative ophthalmology & visual science* 58, 9 (2017), 3343–3349.

[12] Duane R Geruschat, Ava K Bittner, and Gislin Dagnelie. 2012. Orientation and mobility assessment in retinal prosthetic clinical trials. *Optometry and Vision Science* 89, 9 (2012), 1308–1315.

[13] Vijaya K Gothwal, Jan E Lovie-Kitchin, and Rishita Nutheti. 2003. The development of the LV Prasad-Functional Vision Questionnaire: a measure of functional vision performance of visually impaired children. *Investigative ophthalmology & visual science* 44, 9 (2003), 4131–4139.

[14] Albert M Maguire, Katherine A High, Alberto Auricchio, J Fraser Wright, Eric A Pierce, Francesco Testa, Federico Mingozzi, Jeannette L Bennicelli, Gui-shuang Ying, Settimio Rossi, et al. 2009. Age-dependent effects of RPE65 gene therapy for Leber's congenital amaurosis: a phase 1 dose-escalation trial. *The Lancet* 374, 9701 (2009), 1597–1605.

[15] Meinard Müller. 2007. Dynamic time warping. *Information retrieval for music and motion* (2007), 69–84.

[16] Antonio Ortega, Pascal Frossard, Jelena Kovačević, José MF Moura, and Pierre Vandergheynst. 2018. Graph signal processing: Overview, challenges, and applications. *Proc. IEEE* 106, 5 (2018), 808–828.

[17] Shrinivas Pundlik, Matteo Tomasi, and Gang Luo. 2015. Evaluation of a portable collision warning device for patients with peripheral vision loss in an obstacle course. *Investigative ophthalmology & visual science* 56, 4 (2015), 2571–2579.

[18] Lawrence R Rabiner. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* 77, 2 (1989), 257–286.

[19] Silvia Rossi, Francesca De Simone, Pascal Frossard, and Laura Toni. 2019. Spherical clustering of users navigating 360 content. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 4020–4024.

[20] Leopold Schmetterer, Hendrik Scholl, Gerhard Garhöfer, Lucas Janeschitz-Kriegl, Federico Corvi, SriniVas R Sadda, and Felipe A Medeiros. 2023. Endpoints for clinical trials in ophthalmology. *Progress in Retinal and Eye Research* 97 (2023), 101160.

[21] Shagun Shaily and Veenu Mangat. 2015. The hidden Markov model and its application to human activity recognition. In *2015 2nd International Conference on Recent Advances in Engineering & Computational Sciences (RAECS)*. IEEE, 1–4.

[22] David I Shuman, Sunil K Narang, Pascal Frossard, Antonio Ortega, and Pierre Vandergheynst. 2013. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE signal processing magazine* 30, 3 (2013), 83–98.

[23] Ineke JM van der Ham, Annemarie ME Faber, Matthijs Venselaar, Marc J van Kreveld, and Maarten Löffler. 2015. Ecological validity of virtual environments to assess human navigation ability. *Frontiers in psychology* 6 (2015), 637.

[24] Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. 2020. Graph neural networks: A review of methods and applications. *AI open* 1 (2020), 57–81.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009