Challenges for XR in Games

Esteban W. G. Clua
Universidade Federal Fluminense
Depto. de Ciência da Computação
Niterói, Brazil
esteban@ic.uff.br

Daniela Trevisan
Universidade Federal Fluminense
Depto. de Ciência da Computação
Niterói, Brazil
daniela@ic.uff.br

Thiago Porcino
Dalhousie University
Graphics and Experiential Media Lab
Hallifax, Canadá
tmp1986@gmail.com

Bruno A. D. Marques
Universidade Federal do ABC
Centro de Matemática, Computação e Cognição
Santo André, Brazil
bruno.marques@ufabc.edu.br

Eder Oliveira
Universidade Federal Fluminense
Depto. de Ciência da Computação
Niterói, Brazil
eder.oliveira@ifmt.edu.br

Lucas D Barbosa
Universidade Federal Fluminense
Depto. de Ciência da Computação
Niterói, Brazil
barbosa_lucas@id.uff.br

Thallys Lisboa
Universidade Federal Fluminense
Depto. de Ciência da Computação
Niterói, Brazil
thallys.lisboa15@gmail.com

Victor Ferrari
Universidade Federal Fluminense
Depto. de Ciência da Computação
Niterói, Brazil
victorferrari@id.uff.br

Victor Peres
Universidade Federal Fluminense
Depto. de Ciência da Computação
Niterói, Brazil
victor.peres93@gmail.com

***Abstract*—Extended Reality as a consolidated game platform was always a dream for both final consumers and game producers. If for one side this technology had enchanted and called the attention due its possibilities, for other side many challenges and difficulties had delayed its proliferation and massification. This paper intends to rise and discuss aspects and considerations related to these challenges and solutions. We try to bring the most relevant research topics and try to guess how XR games should look in the near future. We divide the challenges into 7 topics, based on extensive literature reviews: Cybersickness, User Experience, Displays, Rendering, Movements, Body Tracking and External World Information.**

***Index Terms*—Extended Reality, Virtual Reality, Digital Entertainment, Head-mounted Displays, UX**

I. Introduction

Extended Reality (XR) platform can be considered as an increment of Virtual Reality in relation to immersion and interaction aspects. While VR platforms are mostly dedicated to visual issues and AR uses real scenes as the main stage, XR includes more external elements and senses, such as movements, tactile, haptics and the usage of the real environment as the application stage [1]. According to the Milgram Continuum, the virtual immersion is a result that comes not only from accurate visual aspects, but mostly from a precise combination of all human senses, orchestrated in such a way that all of them enhances each other. While many progresses had been achieved in graphics, audio, tracking and interfaces issues, there are still many remaining challenges, mostly related to a correct combination on adaption to recent XR hardware devices. In this paper we propose a division of areas for these challenges. We believe that for a real consolidation for games within this platform it is necessary to have robust solutions in each field. We divide the challenges into 7 topics: Cybersickness (CS), user experience and design guidelines, Display and Fovea, Image quality and rendering, movements and redirect walking, body tracking and finally External world information and acquisition.

II. Cybersickness

Motion sickness (MS) is defined as the discomfort felt during a forced visual movement (without body movement), which typically happens in airplane trips, boats, or land vehicles. Such discomfort is also experienced in virtual environments and is called VIMS (Visually Induced Motion Sickness). MS can be split into two subcategories [2]: transportation sickness, which is tied to the real world and simulator sickness, which is associated to the virtual world and includes CS, as shown in Fig. 1. XR environments that use head-mounted displays (HMDs) are strongly related to common indications of discomfort [3]. Among the potential causes, CS deserves special attention as it is the most common and is usually associated to long exposures to HMDs. Additionally, more than 60% of HMDs usability problems are considerably related to discomfort [3]. The most persistent symptoms caused by CS are general discomfort, headache, stomach awareness, nausea, vomiting, sweating, fatigue, drowsiness, disorientation, and apathy [4]. These symptoms influence the user experience and impact the profit and coverage of XR game manufacturing. In addition, discomfort symptoms can vary over people, where some individuals are more susceptible than others. Several studies have been conducted using deep learning models to predict and mitigate CS, such as convolutional neural network (CNNs) and recurrent neural networks (RNNs) [5], [6]. Although deep learning classifiers are the most suitable approach for CS prediction, deep neural networks are black boxes that are very difficult to grasp. In contrast, a recent approach apply techniques to make deep learning models explainable [7], although the literature is still scarce in the topic. Furthermore, symbolic machine learning algorithms enable a straight understanding of decision paths [8]. Another critical problem in CS researching is associated with data labeling. In general, researchers collect verbal, haptic, or brain signal feedback to construct the ground truth of sickness. While verbal feedback is highly subjective and different from each participant, collecting haptic feedback when participants are under discomfort can often be corrupted by the delay associated with participant feedback. A straightforward challenge is related to gender differences tied to XR tasks. Some works [9], [10] pointed out that specific tasks can produce different results of CS for different user-profiles and groups. Overcoming these issues will help designers to produce better XR content and improves the user experience and retain users for longer XR exposures.

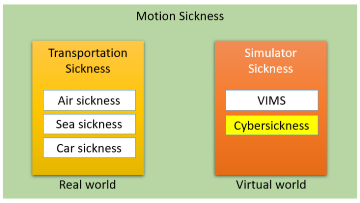

Fig. 1. Motion sickness and its subcategories according to environments and trigger mechanisms.

III. User Experience

Since XR is a new technology and there are many people experiencing it for the first time, it is important that XR designers make their experiences as intuitive and memorable as possible. The game UX accounts for the whole experience players have with a game, from first hearing about it to navigating menus and progressing in the game. The question is: How to make a better game user experience (UX)? Celia Hodent [11] says UX is about understanding the gamer's brain: understanding human capabilities and limitations to anticipate how a game will be perceived, the emotions it will elicit, how players will interact with it, and how engaging the experience will be. As Celia said we believe that UX and cognitive and behavioral psychology can provide very concrete and easy-to-use guidelines to anticipate and even solve design problems. Additionally, techniques from HCI domain such as users interviews, surveys, usability heuristics, analysis of physiological signals, wizardOz [12], among others, need to be properly studied and applied to understand and evaluate the whole UX in the XR context of usage.

IV. Displays, Foveated Rendering

The advent of wider Field of View and higher resolutions HMDs have amplified shading complexities [13]. These features bring a computing power bottleneck, requiring some sort of optimization for keeping target frame rates. Knowing that the human eye has a non regular distribution of cones and rods, some studies have suggested to create non regular pixel distribution (Foveated Rendering). While it is already being explored by rendering engines, there still many challenges for optimizing and customizing it. Foveated Rendering technique exploits human visual system to render the best resolution possible only where the user looks, as in Fig. 2. According to Swafford et al. [14], since the human eye perceives more detail at the center of vision, this uniformity in resolution, regardless of user focus, is a waste of valuable resources such as computing power.

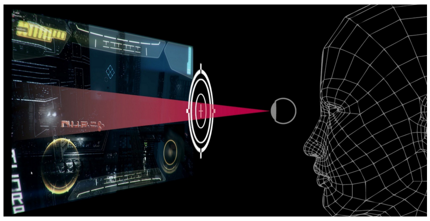

Fig. 2. Foveated Rendering Simulation. Source: [15]

Besides finding a correct balance for the foveated areas, there are still many challenges, such as understanding how this impacts human perception, color distortions and dynamic factors according with the game scene (games with constant colors in large areas naturally requires less pixels to be rendered and enhances the foveated optimizations). It is also important to create robust factors for measuring and better calibrating rendering parameters. Finally, we believe that this concept can also be transposed for different refresh rates for each foveated area, taking into consideration that rods are more dense at the peripheral human vision area.

V. Image Quality and Rendering

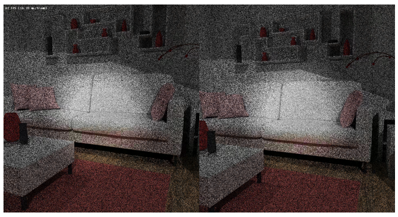

Fig. 3. Path traced image in dual screens using foveation, before denoiser.

Path tracing achieves a higher degree of realism due to the global illumination effects, with a Monte Carlo integration method. Still, it is computationally costly to render a 3D scene at an HMD's resolution. In 2018, with the RTX architecture, access to GPUs capable of optimizing the intersection calculation of a ray with a polygon and thus accelerating realistic rendering became available. With this, some areas of research have been reignited. One of them is the use and optimization of path tracing or hybrid rendering (with rasterization and path tracing) algorithms for virtual reality devices such as HMDs, although they remain changeable. By using the properties of vision such as the concentration of cone distribution in the fovea and devices that allow tracking of the user's gaze, we can avoid rendering parts of the screen with such sharp details or rendering at a reduced spatial sampling frequency. Previous studies have experimented with users to define what an optimal distribution would be, with probabilistic selection of which pixels will be selected by rays, thus decreasing the amount of traced rays and optimizing the algorithm [16]. Similarly, other studies use a fixed texture for ray selection. Reconstruction algorithms called denoisers, which are already commonly used in path tracing rendering, are even more relevant with the fovea distributions, as in Fig.3. The few works with approaches in this regard try to adapt reconstruction to be compatible with log-polar space rendering.

VI. Movements and Redirect Walking

Moving is another form people unconsciously and continuously interact with their surroundings, and developers invented several techniques to move in XR. The problem with this kind of interaction is that while developers can create infinite worlds, users play in confined spaces. Hence, the locomotion, made in real life, can't be mimicked in the virtual world. The literature indicates a growing search to solve this issue [17].

One solution that has the purpose of creating, in the user, the feeling of mimicking his movement by misleading his senses is Redirect Walking [18], which is based on tricking the user's perception and make him feel that is walking forward, but in reality, he is walking in a curved path. The main problem with this is how to shift the virtual environment without triggering the user's perception, which can cause cybersickness and break the immersion. Instead of only divert the player's movement or turn the whole scenery, researchers are using devices, tools, and methods to improve Redirect Walking. Methods such as pointed by Sun et al. [19] recognizes when there are saccade movement of the eye and shifts the scene at the same time. Redirect Walking is the least used method of movement in VR applications mainly because of the necessity of bigger spaces to fully reach its potential of making the user unaware of the reorientation of his movement [?]. Matsumoto et al. [20] measured exprimented that a circular arc of 22m is necessary to avoid perception, but [21] managed to constrain the movements to an area of 6m x 6m.

VII. Body Tracking

XR games must provide users with an immersive experience with a sense of presence and satisfying natural interaction. Body tracking allows reconstruction of the body movement needed to achieve a satisfying natural interaction [22], especially in multiplayerames [23], enabling users to observe other players' movements. A virtual body is crucial for a good level of immersion, and when the user identifies himself with this virtual body, we can see the feeling of presence [24]. Although there are many important works related to the subject [25], most are related to showing only floating hands or VR controllers, due to the lack of movement data. In the application domain, XR in Games, vision-based body tracking remains a challenge because of the sudden change in object motion, cluttered background, partial occlusion, and camera motion. The hands are the most used body parts in XR in games, as they provide a robust form of interaction. Vision-based hand-tracking is a topic of interest of several researchers. Most work on hands-tracking focuses on the use of depth cameras [26] or RGB [12]. Depth-based approaches present results that are superior to RGB-based approaches. A depth camera provides hand geometry in terms of a 2.5D point cloud, and the model-based approaches can reliably fit a hand mesh to the reconstructed point cloud [26]. Using hand tracking input with mobile technology is a problem mainly due to the high energy consumption. Han et al. [27] present a real-time tracking system that uses four egocentric monochrome fisheye cameras to produce 3D hand pose estimates and run not only on PC but also on mobile processors (Fig.4). The system presents failures in hand-hand and hand-object interactions showing that grasping objects and training data generation are still open issues in mobile hand tracking interactions.

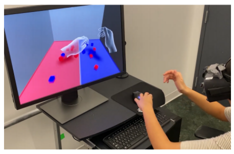

Fig. 4. A real-time hand-tracking system using four monochrome cameras mounted on a VR headset. Source: [27]

VIII. External World Information

Acquiring and processing the external world information is an essential and challenging aspect of XR applications. Real-world data acquisition for XR applications comes from different sources such as motion sensors, cameras, depth sensors, and other hardware. Aggregating and incorporating this data into meaningful information for XR games is not a trivial task.

Fig. 5. Left: rendering of XR scene with inconsistent lighting. Right: relighting of XR scene with consistent environment lighting.

Considering the visual features of XR applications, one crucial aspect is the consistent appearance between virtual and real-world objects. One of the main characteristics that drive the consistent appearance is the lighting between virtual and real-world objects. One possible approach to solve this problem is the relighting of real-world objects into a specific lighting setting [28]. However, relighting the entire scene is still an open challenge. Another possible approach is the relighting of virtual objects (Fig. 5) into a lighting setting that matches the environment lighting [29]. Both of those approaches require the estimation of environmental lighting from the external world. Usually, all the lighting information is estimated from images of the environment, captured in real-time, thus posing as a computer vision problem. With the advances in deep learning, those methods are able to estimate the lighting information and provide a way to improve the XR experience regarding visual fidelity. Extracting and estimating external world information is still considered a difficult task. Developments regarding the representation of the information, including the recent advancements in computer vision methods, can dramatically improve XR environments by allowing new forms of interaction between the virtual and the real world.

IX. Conclusion

Extended Reality as a game platform has an incredible potential, due its high immersive conditions. However, it is a totally new computational and ubiquitous environment and brings many challenges and problems, some of them not trivial. In this workshop we categorize these issues in 7 different topics, although there can be many others. This classification is not exhaustive and there are many other aspects that could also be included, such as audio, new interface devices and collaborative environments.

References

[1] L. Valente, B. Feijó, J. C. S. do Prado Leite, and E. Clua, "A method to assess pervasive qualities in mobile games," *Personal and Ubiquitous Computing*, vol. 22, no. 4, pp. 647–670, 2018.

[2] A. Kemeny, J.-R. Chardonnet, and F. Colombet, *Getting Rid of Cybersickness: In Virtual Reality, Augmented Reality, and Simulators*. Springer Nature, 2020.

[3] E. M. Kolasinski, "Simulator sickness in virtual environments." DTIC Document, Tech. Rep., 1995.

[4] M. S. Dennison and M. D'Zmura, "Cybersickness without the wobble: experimental results speak against postural instability theory," *Applied ergonomics*, vol. 58, pp. 215–223, 2017.

[5] D. Jeong, S. Yoo, and J. Yun, "Cybersickness analysis with EEG using deep learning algorithms," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2019, pp. 827–835.

[6] J. Kim, W. Kim, H. Oh, S. Lee, and S. Lee, "A deep cybersickness predictor based on brain signal analysis for virtual reality contents," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 10 580–10 589.

[7] N. Xie, G. Ras, M. van Gerven, and D. Doran, "Explainable deep learning: A field guide for the uninitiated," *arXiv:2004.14545*, 2020.

[8] T. Porcino, E. O. Rodrigues, A. Silva, E. Clua, and D. Trevisan, "Using the gameplay and user data to predict and identify causes of cybersickness manifestation in virtual reality games," in *2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH)*. IEEE, 2020, pp. 1–8.

[9] S. Grassini and K. Laumann, "Are modern head-mounted displays sexist? a systematic review on gender differences in hmd-mediated virtual reality," *Frontiers in Psychology*, vol. 11, 2020.

[10] C. Curry, R. Li, N. Peterson, and T. A. Stoffregen, "Cybersickness in virtual reality head-mounted displays: Examining the influence of sex differences and vehicle control," *International Journal of Human–Computer Interaction*, pp. 1–7, 2020.

[11] C. Hodent, *The Gamer's Brain: How Neuroscience and UX Can Impact Video Game Design*. CRC Press, 2017. [Online]. Available: https://books.google.com.br/books?id=JzyhDwAAQBAJ

[12] E. de Oliveira, E. W. G. Clua, C. N. Vasconcelos, B. A. D. Marques, D. G. Trevisan, and L. C. de Castro Salgado, "Fpvrgame: Deep learning for hand pose recognition in real-time using low-end hmd," in *Joint International Conference on Entertainment Computing and Serious Games*. Springer, 2019, pp. 70–84.

[13] R. Albert, A. Patney, D. Luebke, and J. Kim, "Latency requirements for foveated rendering in virtual reality," *ACM Transactions on Applied Perception (TAP)*, vol. 14, no. 4, pp. 1–13, 2017.

[14] N. T. Swafford, J. A. Iglesias-Guitian, C. Koniaris, B. Moon, D. Cosker, and K. Mitchell, "User, metric, and computational evaluation of foveated rendering methods," in *Proceedings of the ACM Symposium on Applied Perception*, 2016, pp. 7–14.

[15] "Fove's $250,000 kickstarter campaign wants to bring eye-tracking control to virtual reality," Apr 2021, accessed: 2021-07-28. [Online]. Available: https://thenextweb.com/news/foves-250000-kickstarter-campaign-wants-to-bring-eye-tracking-control-to-virtual-reality

[16] M. Weier, T. Roth, E. Kruijff, A. Hinkenjann, A. Pérard-Gayot, P. Slusallek, and Y. Li, "Foveated real-time ray tracing for head-mounted displays," in *Computer Graphics Forum*, vol. 35, no. 7. Wiley Online Library, 2016, pp. 289–298.

[17] A. Cannavò, D. Calandra, F. G. Pratticò, V. Gatteschi, and F. Lamberti, "An evaluation testbed for locomotion in virtual reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 3, pp. 1871–1889, 2021.

[18] E. Jonathan, C. Roberts, S. Presentation, S. Razzaque, Z. Kohn, and M. Whitton, "Redirected walking," *Proc. of Eurographics*, 08 2001.

[19] Q. Sun, A. Patney, L.-Y. Wei, O. Shapira, J. Lu, P. Asente, S. Zhu, M. McGuire, D. Luebke, and A. Kaufman, "Towards virtual reality infinite walking: dynamic saccadic redirection," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 4, p. 67, 2018.

[20] K. Matsumoto, Y. Ban, T. Narumi, Y. Yanase, T. Tanikawa, and M. Hirose, "Unlimited corridor: Redirected walking techniques using visuo haptic interaction," in *ACM SIGGRAPH 2016 Emerging Technologies*, ser. SIGGRAPH '16. New York, NY, USA: Association for Computing Machinery, 2016. [Online]. Available: https://doi.org/10.1145/2929464.2929482

[21] M. Rietzler, J. Gugenheimer, T. Hirzle, M. Deubzer, E. Langbehn, and E. Rukzio, "Rethinking redirected walking: On the use of curvature gains beyond perceptual limitations and revisiting bending gains," 10 2018, pp. 115–122.

[22] S. Kasahara, K. Konno, R. Owaki, T. Nishi, A. Takeshita, T. Ito, S. Kasuga, and J. Ushiba, "Malleable embodiment: changing sense of embodiment by spatial-temporal deformation of virtual human body," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 6438–6448.

[23] F. Jiang, X. Yang, and L. Feng, "Real-time full-body motion reconstruction and recognition for off-the-shelf vr devices," in *Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry-Volume 1*, 2016, pp. 309–318.

[24] M. Slater and S. Wilbur, "A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments," *Presence: Teleoperators & Virtual Environments*, vol. 6, no. 6, pp. 603–616, 1997.

[25] P. Caserman, A. Garcia-Agundez, and S. Göbel, "A survey of full-body motion reconstruction in immersive virtual reality applications," *IEEE transactions on visualization and computer graphics*, vol. 26, no. 10, pp. 3089–3108, 2019.

[26] J. Taylor, V. Tankovich, D. Tang, C. Keskin, D. Kim, P. Davidson, A. Kowdle, and S. Izadi, "Articulated distance fields for ultra-fast tracking of hands interacting," *ACM Transactions on Graphics (TOG)*, vol. 36, no. 6, pp. 1–12, 2017.

[27] S. Han, B. Liu, R. Cabezas, C. D. Twigg, P. Zhang, J. Petkau, T.-H. Yu, C.-J. Tai, M. Akbay, Z. Wang *et al.*, "Megatrack: monochrome egocentric articulated hand-tracking for virtual reality." *ACM Trans. Graph.*, vol. 39, no. 4, p. 87, 2020.

[28] R. Pandey, S. Orts-Escolano, C. LeGendre, C. Haene, S. Bouaziz, C. Rhemann, P. Debevec, and S. Fanello, "Total relighting: Learning to relight portraits for background replacement," vol. 40, no. 4, 2021.

[29] B. A. D. Marques, E. W. G. Clua, and C. N. Vasconcelos, "Deep spherical harmonics light probe estimator for mixed reality games," *Computers Graphics*, vol. 76, pp. 96–106, 2018.