Opportunities and Challenges in Immersive Entertainment

Wallace S. Lages
Virginia Tech
School of Visual Arts
Blacksburg, USA
wlages@vt.edu

***Abstract*—The next major shift in digital entertainment will be the wide adoption of mixed reality platforms. Although owners of mixed reality devices are still a minority today, immersive systems are quickly getting better and cheaper. We identity eight research topics on immersive entertainment and discuss possible research directions that can lead to advances in the field.**

***Index Terms*—mixed reality, virtual reality, storytelling, games**

I. Introduction

The digital entertainment industry saw its latest great shift around mid-2000s with the growth of mobile gaming. Smartphones enabled a completely new ecosystem in which novel designs appears and new technical solutions were developed. Mobile gaming broke many previous assumptions of the field, including the typical profile of players and how gaming hardware should look like. In less than 10 years, the mobile games sector revenue surpassed PC and consoles [1]. Although smartphones were not designed as primary gaming devices, their wide availability, new interaction methods, and ubiquitous connectivity also untapped new areas for research.

We are now witnessing a similar shift, with the second generation of affordable and high-quality virtual reality (VR) headsets reaching the market. The consumer hardware evolved quickly from simple smartphone-based VR to standalone headsets with six degree of freedom tracking. Applications also followed accordingly from pure cinematic experiences to highly interactive and fast-paced games. New applications pushed the boundaries of storytelling and theatrical performance. Due to its more complex requirements, augmented reality devices have not yet reached the same point, although they are likely to follow suit in the next few years.

How will immersive entertainment look like 10-15 years from now? If the trajectory of mobile gaming serves of indication, it is safe to assume that systems will become cheaper, better, and ubiquitous. Display, optics and processing power will improve, along with sensors and networking. We will also likely encounter true mixed reality devices, which will be able to operate in VR, AR, or blended modes. Improved quality and form factor will enable people to use these devices for many hours at a time. Social experiences will be commonplace, either physically co-located or along with other users around the world.

In this paper, we look even further ahead to identity challenges in design and enabling technology required to transition to the next level of immersive experiences.

II. Immersive Entertainment

We will use the term "immersive entertainment" to include any entertainment application designed for immersive systems, for example, games, narrative experiences, and performance work. Although there might be a considerable overlap with work, therapeutic, training, or educational applications, our concern here is primarily entertainment.

Today these systems are characterized by egocentric, head-tracked, stereoscopic rendering. For example, in the form-factor of head-worn virtual reality displays, CAVE projection systems, or augmented-reality glasses. These systems also provide some type of spatial input by means of six degree-of-freedom tracked controllers or hand-tracking. These systems also provide a 360 degrees field of regard, allowing the user to be completely surrounded by visual imagery (in contrast to a single screen display).

Advances in immersive entertainment have happened so far in the interdisciplinary fields of play, games, and human-computer interaction, fields which are themselves supported by efforts in many other disciplines including artificial intelligence, computer graphics, communication, theater, computer engineering, sociology, storytelling, and psychology. For this discussion, we have selected eight topics which are likely to continue open in the next decade and in which advances could significantly transform immersive entertainment:

A) Understanding presence and user experience
B) Creating believable characters and engaging storytelling
C) Design of novel interaction techniques
D) Improving rendering, tracking, and space understanding
E) Exploring uses for physiological sensing and biofeedback
F) Creating better social experiences
G) Safeguarding users and promoting responsible design
H) Improving and democratizing content generation

The topics are presented in no particular order, and we make no claim of completeness. They are only meant to give an overview of the challenges of this area. For example, haptics, optics, and input devices, are all active research areas on immersive technology but they will not be discussed here.

A. Understanding presence and user experience

What distinguish immersive systems apart from other technologies is the realness of the experience, also known as *presence*. Presence is often discussed in terms of separate constructs: *Place Illusion*, *Plausibility Illusion*, and *Social Presence Illusion* [2], [3]. Place Illusion refers to the feeling of existing or being in a place. It is the sensation of being somewhere else, even when the user knows that is not the case. Plausibility illusion is the acceptance that the events depicted are really happening. Social Presence is used to refer to the feeling of being together with another sentient being.

Supporting these illusions is desirable, since it elicits realistic behavioral, emotional, and psychological responses which makes immersive entertainment so compelling. If we want a solid basis upon which experiences can be created, it is uttermost importance to understand the mechanisms behind these constructs. The understanding of how system's software and hardware characteristics affect presence and other aspects of the user experience is complex, given the multiple parameters, individual differences, and fuzzy and interconnected nature of these illusions. They remain a challenge for the next decade and beyond. Some broad questions include:

- How can we reliably measure different aspects of presence?
- How different characteristics of a system affect presence, immersion and user experience?
- Is there an optimal amount of presence?
- How much does presence impact the actual user engagement and experience?

B. Creating believable characters and engaging storytelling

Storytelling is a central part of games and other narrative experiences. An immersive story can range from a completely fixed narrative (e.g., a journalistic piece) to an open-ended, AI-managed experience. To create engaging story worlds and make the most of the user's agency inside the virtual environment, it will be necessary to combine computational approaches with the creativity of human writers. Interactive storytelling and computational narratives [4] are, thus, an integral part of the future of immersive entertainment.

Compelling intelligent characters will be necessary not only for story purposes, but also for generating plausibility. Storytelling in virtual environments is spatial, with the narrative events being experienced through the user actions and existence in the environment [5]. Characters need to be aware of space, how human express themselves through space, and behave accordingly. Natural interaction with non-playable characters (NPCs) will also involve techniques beyond point-and-click, such as dialog generation, speech synthesis, and procedural animation. We may even see NPCs designed to replicate the personality of an individual during a reenactment, which will bring them closer to an artificial life.

- How can we create intelligent agents which can connect emotionally with users?
- How can we adapt narrative trajectories to provide replayability?
- How can we guide story generation towards the storyteller vision?
- How can we understand and analyze spatial information in virtual environments?
- How can we convey narrative events outside the user's field of view?

C. Design of novel interaction techniques

Early research on interaction techniques for virtual environments have focused on quantifying the properties of fundamental interaction problems. However, entertainment applications tend cover a wide variety of themes. This opens the opportunity to explore highly specific interaction techniques, designed for a specific setting and which contributes towards a specific aesthetic goal. For example, using storytelling elements, such as a bookshelf or a bird to achieve narrative-consistent redirection in VR [6] or using a magic carpet to fly [7].

Advances in hardware will also provide new opportunities for research. New sensors will enable completely new techniques. Some promising avenues include using radio frequency signals to detect touch or recognize gestures [8], [9]. Interaction design research will continue to be one of the pillars supporting good user experiences once devices are small enough for practical use.

- What are good travel techniques for different entertainment applications?
- How do we best use natural walking and bare-hand interaction?
- What is the influence of interaction techniques on presence?
- How can we improve input devices for general and domain-specific input?
- How can we design practical techniques for biometric sensors (eye-based, brain computer interfaces)?

D. Improving rendering, tracking, and space understanding

Immersive applications already have to balance several *synthesis challenges*: they must render stereoscopic views at low latency, provide enough visual quality to achieve creator's aesthetic goals, and run on standard consumer hardware. In addition, most future systems will also have significant *reconstruction challenges*: map and locate the position of the user in the world (for inside-out systems), create a model of the physical surroundings (geometry of surfaces, materials, light sources), and create semantic models of the world (identify floors, objects, furniture, people).

Rendering in immersive systems share most of the traditional real-time rendering challenges: how to achieve fast and realistic rendering of illumination, shadows, surfaces, etc. Even though immersive systems have stricter constraints on latency, they also offer some unique opportunities. For example, images don't need to be accurate, only *perceptually* accurate. Techniques like foveated rendering can make the use of integrated headset sensors to reduce computational load by using gaze direction to adjust graphics to match the acuity of different parts of the visual field [10]. Advances will improve on existing vision models to include eye movement, brightness, etc.

In computer vision, one important problem is the development of inverse models, that can be used to match rendered graphics with real scenes. Fast methods of acquisition and relighting [11] will be of increasingly importance in AR. Advanced scene mapping and understanding will allow more realistic integration between physical and virtual objects in mixed reality [12]. Advances in this area will required increased robustness to different materials, textures, and shapes that still elude current systems. Some interesting problems include:

- How can we recognize objects with headset embedded sensors?
- How can we use the limits of human perception to achieve more with less?
- How can we adapt AR experience to the user's physical location?

E. Exploring uses for physiological sensing and biofeedback

Immersive systems will contain more physiological sensors than previous technologies. Current headsets already include sensors for tracking different facial features (e.g., pupil, eyebrow, mouth) as a way to drive avatar animation. Headsets under development are investigating to incorporate other sensors, such as electroencephalogram (EEG), electrooculography (EOG), electromyography (EMG), electrodermal activity (EDA), and photoplethysmography (PPG) [13].

Physiological sensors will enable the creation of more immersive and dynamic game experiences by capturing information on mental and emotional states. Emotional states can also be used to detect the players attitudes towards NPCs, adjust game difficulty based on frustration or excitement levels, or indicate player intention for an AI system. When integrated with adaptive gameplay, they can be used to close a emotional biofeedback loop [14].

Physiological signals can also be used directly to control user actions [15] or other parameters. For example, PPG could be change effort levels during physically active games and EDA to automatically trigger VR comfort options on cybersickness onset [16]. Questions of relevance include:

- How can we model valence and arousal?
- How do we remove artifacts from sensor data?
- How to integrate player's emotional state into AI systems?
- How can infer intention from sensors?
- How do we use physiological signals to drive an optimal experience?

F. Creating better social experiences

A significant part of future immersive entertainment will be social or have elements for sharing the experience with others. Some of these experiences might be co-located and others will be remote. Since interaction on immersive environments are mediated by the interface, designing new mechanisms for people to interact with each other and express themselves within virtual environments will be fundamental.

Promoting positive experiences is likely to remain a fruitful area of research, as most online communities today are afflicted by toxic behavior that discourages the participation of several groups. Solutions will need to be based on both technical development and deep understanding of the social dynamics in virtual worlds.

Additional challenges include the design of networked (possibly distributed) system architectures to support persistent spaces, avatar customization systems, and new ways of self-expression. Among the questions around social experiences, we can list:

- How can we design stories for multiple users where each participant is relevant?
- How can we design distributed and persistent virtual worlds?
- How can we design mechanisms for rich expression in virtual environments?
- How can we promote positive social experiences?
- How are social relationships in immersive worlds?

G. Safeguarding users and promoting responsible design

Immersive applications can be extremely compelling and believable, potentially leaving post-experience effects [17]. These effects can used to achieve positive results, for example, treating patients through exposure therapy and promote behavioral changes [18], [19]. However, they could also lead users to perform immoral, cruel, or psychologically harmful acts.

Immersive systems will be able to gather personal information about users and their surroundings. Tracking of the head and body parts are an integral part of their operation. In addition, many devices will include cameras, microphones, and sensors to scan and map the environment. It has been shown that data from body motion can be used to identify users and even diagnose certain medical conditions [20], [21]. Research will be required on anonymization techniques and/or data security.

Questions also exist on the representational aspects of the experiences. Unlike a conventional documentary, an immersive piece lacks the clear presence of the director, inducing the user to think the event depicted is real and that the user is part of it [22]. Researchers can help identify and navigate the ethical questions around immersive content, including accessibility, algorithmic bias, speech, virtual resurrections, etc.

- How can we secure biometrical data acquired by immersive systems?
- How can we make content accessible to all users?
- How do we filter disrupting behavior while allowing free expression?
- How do we design systems that collect minimal information?
- Which principles should we follow for content design?
- Which policies or guidelines should be institute regarding privacy?

H. Improving and democratizing content generation

One final important enabling aspect of immersive entertainment is the availability of content creation tools. Good tools can reduce the cost of production and enable non-specialists access and voice in the space. For realistic experiences or depicting real places, tools for quickly capturing geometry and material information from real scenes will be essential. That will require a leap from structured lightstages to single-cameras or portable setups. Deep learning methods have recently allowed good quality 3D reconstruction of faces from single images [23]. Similar technology would need to be developed for full body reconstruction, including challenging parts like hair, eyes, and lips.

Other areas that would benefit from tooling are the design of spatial AI systems and narrative systems. Those are likely to be powered by sophisticated machine learning but that could be trained to achieve the needs of a specific creator. Computational narratives spaces can also get complex quickly [24], so tools could incorporate visual models to facilitate the understanding of branches and story outcomes. Questions for exploration include:

- How can we allow quick capture of existing spaces?
- How can we reduce the time required to create new characters?
- How can we design and visualize complex narrative spaces?
- How can we capture the performance of actors without expensive infrastructure?

III. Conclusion

Immersive entertainment will grow to become a major category of digital entertainment in the next decade and will be the source of challenging problems in several computer science areas. In this paper, we have identified eight topics of relevance: understanding presence and user experience; creating believable characters and engaging storytelling; designing of novel interaction techniques; improving rendering, tracking, and space; exploring uses for physiological sensing and biofeedback; creating better social experiences; safeguarding users and promoting responsible design; and improving and democratizing content generation. Advances in these areas are likely to impact related immersive applications, such as education and training.

References

[1] "The global games market reaches $99.6 billion in 2016, mobile generating 37%," https://newzoo.com/insights/articles/global-games-market-reaches-99-6-billion-2016-mobile-generating-37, accessed: 2021-09-01.

[2] M. Slater, "Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1535, pp. 3549–3557, 2009.

[3] R. Skarbez, F. P. Brooks, Jr, and M. C. Whitton, "A survey of presence and related concepts," *ACM Computing Surveys (CSUR)*, vol. 50, no. 6, pp. 1–39, 2017.

[4] A. Alabdulkarim, S. Li, and X. Peng, "Automatic story generation: Challenges and attempts," *arXiv preprint arXiv:2102.12634*, 2021.

[5] A. Hameed and A. Perkis, "Spatial storytelling: Finding interdisciplinary immersion," in *International Conference on Interactive Digital Storytelling*. Springer, 2018, pp. 323–332.

[6] R. Yu, W. S. Lages, M. Nabiyouni, B. Ray, N. Kondur, V. Chandrashekar, and D. A. Bowman, "Bookshelf and bird: Enabling real walking in large vr spaces through cell-based redirection," in *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 2017, pp. 116–119.

[7] D. Medeiros, M. Sousa, A. Raposo, and J. Jorge, "Magic carpet: Interaction fidelity for flying in vr," *IEEE transactions on visualization and computer graphics*, vol. 26, no. 9, pp. 2793–2804, 2019.

[8] D. Kim, K. Park, and G. Lee, "Atatouch: Robust finger pinch detection for a vr controller using rf return loss," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–9.

[9] M. Scherer, M. Magno, J. Erb, P. Mayer, M. Eggimann, and L. Benini, "Tinyradarnn: Combining spatial and temporal convolutional neural networks for embedded gesture recognition with short range radars," *IEEE Internet of Things Journal*, 2021.

[10] R. K. Mantiuk, G. Denes, A. Chapiro, A. Kaplanyan, G. Rufo, R. Bachy, T. Lian, and A. Patney, "Fovvideovdp: A visible difference predictor for wide field-of-view video," *ACM Transactions on Graphics (TOG)*, vol. 40, no. 4, pp. 1–19, 2021.

[11] J. Park, H. Park, S.-E. Yoon, and W. Woo, "Physically-inspired deep light estimation from a homogeneous-material object for mixed reality lighting," *IEEE transactions on visualization and computer graphics*, vol. 26, no. 5, pp. 2002–2011, 2020.

[12] X. Zhao, Y. Pang, J. Yang, L. Zhang, and H. Lu, "Multi-source fusion and automatic predictor selection for zero-shot video object segmentation," *arXiv preprint arXiv:2108.05076*, 2021.

[13] "Tobii, valve & openbci collaborate on 'galea' vr brain-computer interface," https://www.vrfocus.com/2021/02/tobii-valve-openbci-collaborate-on-galea-vr-brain-computer-interface/, accessed: 2021-09-01.

[14] S. Houzangbe, O. Christmann, G. Gorisse, and S. Richir, "Fear as a biofeedback game mechanic in virtual reality: effects on engagement and perceived usability," in *Proceedings of the 13th International Conference on the Foundations of Digital Games*, 2018, pp. 1–6.

[15] L. E. Nacke, M. Kalyn, C. Lough, and R. L. Mandryk, "Biofeedback game design: using direct and indirect physiological control to enhance game interaction," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2011, pp. 103–112.

[16] N. Martin, N. Mathieu, N. Pallamin, M. Ragot, and J.-M. Diverrez, "Virtual reality sickness detection: An approach based on physiological signals and machine learning," in *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 2020, pp. 387–399.

[17] K. Y. Segovia and J. N. Bailenson, "Virtually true: Children's acquisition of false memories in virtual reality," *Media Psychology*, vol. 12, no. 4, pp. 371–393, 2009.

[18] Y. S. Schmitt, H. G. Hoffman, D. K. Blough, D. R. Patterson, M. P. Jensen, M. Soltani, G. J. Carrougher, D. Nakamura, and S. R. Sharar, "A randomized, controlled trial of immersive virtual reality analgesia, during physical therapy for pediatric burns," *Burns*, vol. 37, no. 1, pp. 61–68, 2011.

[19] A. Garcia-Palacios, H. Hoffman, A. Carlin, T. A. Furness III, and C. Botella, "Virtual reality in the treatment of spider phobia: a controlled study," *Behaviour research and therapy*, vol. 40, no. 9, pp. 983–993, 2002.

[20] M. R. Miller, F. Herrera, H. Jun, J. A. Landay, and J. N. Bailenson, "Personal identifiability of user tracking data during observation of 360-degree vr video," *Scientific Reports*, vol. 10, no. 1, pp. 1–10, 2020.

[21] M. Alcaniz Raya, J. Marín-Morales, M. E. Minissi, G. Teruel Garcia, L. Abad, and I. A. Chicchi Giglioli, "Machine learning and virtual reality on body movements' behaviors to classify children with autism spectrum disorder," *Journal of clinical medicine*, vol. 9, no. 5, p. 1260, 2020.

[22] H. Kool, "The ethics of immersive journalism: A rhetorical analysis of news storytelling with virtual reality technology," *Intersect: The Stanford journal of science, technology, and society*, vol. 9, no. 3, 2016.

[23] A. Lattas, S. Moschoglou, B. Gecer, S. Ploumpis, V. Triantafyllou, A. Ghosh, and S. Zafeiriou, "Avatarme: realistically renderable 3d facial reconstruction "in-the-wild". 2020 ieee," in *CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[24] C. Siler and S. Ware, "A good story is one in a million: solution density in narrative generation problems," in *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, vol. 16, no. 1, 2020, pp. 123–129.