RAE-IC
Revista de la Asociación Española de Investigación de la Comunicación

RAE-IC, Revista de la Asociación Española de Investigación de la Comunicación
vol. 11, núm. Especial (2024), raeic11e01
ISSN 2341-2690
DOI: https://doi.org/10.24137/raeic.11.e.1

Recibido el 15 de mayo de 2024
Aceptado el 31 de mayo de 2024

Intelligent Immersion: The current landscape of AI tools in the Mixed Reality development pipeline for creative experiences

Dima, Mariza
Brunel University London (BUL)
mariza.dima@brunel.ac.uk

Daylamani-Zad, Damon
Brunel University London (BUL)
damon.daylamani-zad@brunel.ac.uk

Forma de citar este artículo:
Dima, M., & Daylamani-Zad, D. (2024). Intelligent Immersion: The current landscape of AI tools in the Mixed Reality development pipeline for creative experiences. *RAE-IC, Revista de la Asociación Española de Investigación de la Comunicación*, *11*(Especial), raeic11e01. https://doi.org/10.24137/raeic.11.e.1

Abstract:

Among the technologies that hold immense potential to revolutionize how we interact with information and digital content, Mixed Reality (MR) offers unique immersive experiences that seamlessly integrate virtual objects into the user's physical environment. This groundbreaking fusion of the physical and digital worlds has a complex development process which is a fertile ground for applications of Artificial Intelligence (AI). This article aims to provide a comprehensive overview of AI tools and

their applications, in all stages of the development of MR experiences for the creative sector. It also discusses the challenges and opportunities of incorporating them in the development pipeline and offer some use guidelines as a compass to navigate this rapidly changing landscape.

Keywords: Artificial Intelligence, Mixed Reality, Tools, Deep Learning, Machine Learning, Creativity, Development pipeline, Experience Design.

1. INTRODUCTION

At the back of the ultimate goal to create machines that exhibit human-like intelligent behaviour, the field of Artificial Intelligence (AI) has, through many breakthroughs, developed architectures and techniques to address complex processes that continue to shape and reshape work methods, making tasks easier and faster, and offer tools that redefine what is possible. The application of AI techniques is nowadays increasingly ubiquitous, from simple tasks to more advanced applications, including methods to enable creative explorations, which have found a prominent place in the creative sector.

The creative sector requires a different set of cognitive processes, skills, and approaches to produce creative outputs compared to those required by more routine and repetitive tasks often found in productivity settings. Creative tasks involve generating new ideas, solutions, or content that are original, innovative, and novel. These tasks often require divergent thinking, experimentation, and imagination to explore multiple possibilities and perspectives which are often difficult to address by constrained learning systems found in AI techniques (Anantrasirichai & Bull, 2022). The nature of the integration of AI tools in creative and artistic work has been an on-going discourse. In principle, artists and other creatives use AI to assist the creative process as another tool in their toolset that can support inspiration and experimentation rather than substitute and take the human out of the loop.

The support that AI tools offer in machine vision, NLP, and content creation has made a huge impact in the development of experiences in the domain of spatial computing,

where Mixed Reality (MR) is situated. MR combines Augmented (AR) and Virtual Reality (VR) to create a hybrid environment featuring holographic material that is usually, but not exclusively, viewed through a pair of smart glasses (Head Mounted Displays-HMD). The users can interact with the virtual content in the physical space, and with use of appropriate sensors, the users' movement through space can affect the superimposed content. MR has opened many creative possibilities in the creative and entertainment sectors such as gaming, and virtual production. Creating content and interactions in cyber-physical spaces requires a complex development pipeline, and cross disciplinary work.

In this paper we review how AI is currently used in MR development for creative experiences. We provide an overview of AI techniques for creative processes that pertain to MR development, followed by an in-depth review of tools for each different stage of the MR development pipeline. Finally, we discuss challenges and opportunities of integrating AI tools in the development process presenting a set of use considerations that can guide designers and developers.

2. LITERATURE REVIEW

Machine Learning (ML) algorithms have paved the way for groundbreaking applications of AI across diverse domains such as healthcare (Ali et al., 2023), manufacturing (Liu et al., 2020), business (Bharadiya et al., 2023), education (Zawacki-Richter et al., 2019), and the creative industries (Anantrasirichai et al., 2022). A large focus of ML research is on advancing Deep Learning algorithms (DL), a method that is based on neural networks and hierarchical representation of data (LeCun, 2015). The DL method is very useful for MR-related applications as it provides accuracy and efficiency in the development of content and interactions.

The integration of AI and MR technologies is emerging as a powerful tool that has started revolutionising both research domains as well as domains where these technologies are applied. MR technologies have grown rapidly in recent years due to advances in hardware, large data storage options, and the advent of Industry 4.0 developments. An increasing number of MR applications are found in diverse domains and verticals from heritage to healthcare for purposes such as training, education, and

entertainment. Under the experience economy (Pine & Gilmore, 2011), where metaphors from theatre are employed to craft immersive interactions that foster deep emotional responses and fuelled by the building of the Metaverse (Thien-Hu, 2023), MR is widely researched and used to create immersive, affective journeys for audiences and consumers.

AI-driven MR tools provide solutions to several robotics applications (Bassyouni & Elhajj, 2021), medical training (Bissonnette et al. 2019), military training (Israelsen et al., 2018), industry (Ming-Tang et al., 2023; Rüßmann et al., 2015), retail (Rathore, 2021), aviation (Jiang et al., 2023), and education (Hensen 2023). In a many of these applications the focus of AI is on improving the 3D visualisations used in MR to ease understanding of complex structures and systems, and supporting the interactions designed within the system to enhance learning and increase efficiency and productivity. AI extends and enriches MR experiences, and vice versa (Reiners et al., 2021).

In the development of creative 3D MR experiences, AI can facilitate every step of the creation process aiming to produce seamless hybrid experiences. Mood boards, storyboards, and mock-up prototyping are useful tools in the conceptualisation stage and can be populated through text description which is an AI driven process based on Natural Language Processing (NLP) and Natural Language Understanding (NLU) - both using ML and DL algorithms. NLP and NLU recognise not only words but also the structure of a language and many tools leverage it for text to image with AI Image Synthesis Models such as Open AI's DALL-E[1], and Stability AI's Stable Diffusion[2], text to 3D model generation (Poole et al., 2022; Wang et al., 2023) such as Meshy[3], and text to video (Blattman et al., 2023; Hong et al., 2022; Uriel et al., 2022; Wu et al., 2023) such as Synthesia[4] and Stable Video Diffusion[5].

In MR experiences, it is important to have smooth interactions between virtual content, user movements, and physical space. Therefore, space mapping and environment

[1] https://openai.com/index/dall-e-2/
[2] https://stability.ai/stable-image
[3] https://app.meshy.ai/discover
[4] https://www.synthesia.io/features/text-to-video
[5] https://stability.ai/stable-video

scanning, movement recognition, and computation of real time light conditions considering illumination and reflections, occlusion and depth of field are important parameters to compute correctly. Advanced image processing (Yeh et al., 2019), depth sensing (Lai et al., 2019), object recognition and retrieval (Liu et al., 2020; Zhu et al., 2018), and gesture/gaze estimation, recognition and tracking tasks (Anantrasirichai et al., 2016; Huynh-The et al., 2020; Hu et al., 2020; Soccini, 2017; Ojeda-Castello et al., 2022) are assisted by DL algorithms. DL architectures are widely used in automatically generated animations (Khatri, 2021). Neural Radiance Fields (NeRF) are neural networks that can reconstruct complex three-dimensional scenes from a partial set of two-dimensional images and are used for the reconstruction of dynamic 3D environments. They are commonly used to reconstruct the appearance of human face and its dynamics such as talking and facial expressions (Gafni et al., 2021). Outside computer vision, music and sound are important parts of MR experiences. There are many applications of AI for music generation such as JukeBox (Dhariwal et al., 2020) and NSynth (Engel et al., 2017).

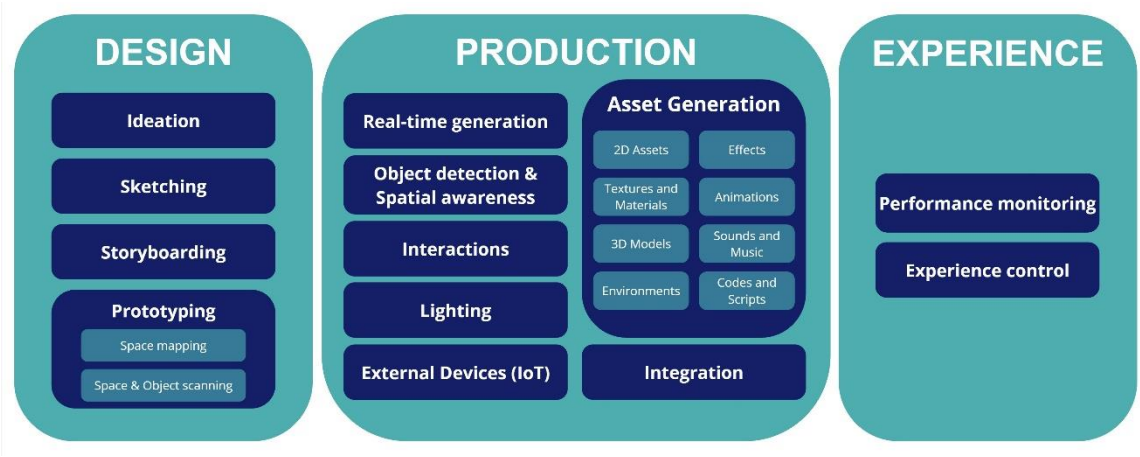
Considering the current possibilities and applications provided in the state-of-the-art, there is real potential with many opportunities to use AI-driven solutions in MR design and development. The next section discusses where and how AI solutions present opportunities in the MR development pipeline with some examples of applications in past MR experiences.

3. AI TOOLS IN MIXED REALITY PIPELINE

The potentials of AI solutions are vast and advancing quickly. Hence, it is not possible to cover all available solutions. The AI solutions could range from Generative AIs (Gen AI) which create new assets based on training data, stochastic methods and pre-trained models, to Deterministic AIs which always have the same output for the same input. Gen AIs are considered mostly non-deterministic, yet with higher levels of training can lead to almost deterministic behaviour. Concurrently, deterministic AI approaches can be generative or non-generative (analytical, such as classifiers and regressors).

With the current advances of AI tools, especially generative AI, there are various areas of development for MR which can support and enhance MR experience development. We consider the possibilities and opportunities for the stages of MR development pipeline: design, production, release. **Figure 1** presents the standard pipeline of MR experience development with a breakdown at each stage of areas that present opportunities for AI tools.

Figure 1. MR experience creation with AI pipeline stages



It is important to note that the use of AI should not be a goal in itself, but rather means and methods for facilitating creativity, while improving and optimising the design and development of the experience. Therefore, whilst the current and state-of-the-art AI tools can be used in various stages, designers and developers should choose the right tools for each stage of the pipeline that suits them.

3.1. DESIGN STAGE

The incorporation of Gen AI tools in the ideation phase is now a popular practise among many designers. Tools such as ChatGPT[6], Midjourney[7], Adobe AI Assistant [8]and Adobe Sensei[9], and Creative AI Lab[10] provide the ability to create and generate visuals and mock-ups using textual and visual inputs. MR developers can harness these tools for

[6] https://chat.openai.com/
[7] https://www.midjourney.com/
[8] https://www.adobe.com/
[9] https://www.adobe.com/uk/sensei/generative-ai.html
[10] https://creative-ai.org/

their ideations, and conceptualisation of their ideas including mock-ups and high-fidelity sketches from low-fidelity hand drawn or digital sketches.

AI tools for storyboarding is a growing field as the practice is used across different creative industries. These range from script generation using generative text tools such as ChatGPT and Google Gemini[11] to tools that generate visual storyboards such as storyboarder.ai[12], Boords[13], Storyboard Hero[14], and Elai.io[15]. Each of these provide features and facilities that allow rapid generation of storyboards to explore the possible creative avenues for an MR experience and communicate the concept to all stakeholders.

A major requirement in MR experiences is the interaction between physical space and real-time 3D content. Spatial understanding is at the core of MR experiences and is what sets them apart from other extended reality experiences. Examples of AI tools that help with the mapping of spaces using cloud points, wireframes, and textured 3D models are Scaniverse[16], Agisoft Metashape[17] and Azure AI Vision[18] which accurately (depending on technology and tools used) scan objects and environments. Having a high-quality accurate scan of the physical environment allows for rapid prototyping of the experience which can be very useful for understanding the space and the feel for the interactions. **Figure 2** presents an actual room and its prototype 3D scans created as part of an Immersive Mixed Reality experience 'Sancho's Journey' at Marble Hill House, a London 18th century stately house (September 2023).

[11] https://gemini.google.com/
[12] https://storyboarder.ai/
[13] https://boords.com/
[14] https://www.storyboardhero.ai/
[15] https://elai.io/
[16] https://scaniverse.com/
[17] https://www.agisoft.com/
[18] https://learn.microsoft.com/en-us/azure/ai-services/computer-vision/concept-object-detection

Figure 2. Hallway and mahogany staircase: Photos (top) Prototype 3D scans of the room (bottom) – design stage, Marble Hill, London, UK 2023.

3.2. PRODUCTION STAGE

The power of AI tools and Gen AI is much more pronounced when it gets to production stage. Traditionally all assets (2D and 3D models, 3D scenes, animations, videos, music, sounds) and code were produced and developed by designers and developers. This process meant higher costs in terms of budgets and time. The various types of assets used in MR projects require a wide verity of expertise. This also creates a barrier for creation and curation of content by non-experts in MR, which may well be experts in the specific application area of MR, e.g. education, healthcare, heritage. Real-time asset

generation from text removes this barrier and the methods offered by tools such as Scenario[19] and Layer AI[20] offer creative control and allow for unique and personalised outputs. For example, Layer AI creates assets in the creator's personalised artistic direction. It is worth noting that whilst these tools can be used for real-time generation, they also can be used for offline asset generation as well.

The arguments for offline asset generation are very similar to the real-time generation. There are a variety of tools for various types of assets. Adobe Sensei, Midjourney and Scenario are used to create 2D assets in a variety of styles. There are also features within the existing tools which provide AI powered features for generating textures, maps and materials. Substance 3D[21] can procedurally generate textures and materials that transcend the traditional approach and assist with creativity and uniqueness in material generation. Very useful is the ability to create 3D models of objects and environments. There are many tools for procedural generation of vegetations, foliage, full environments, as well as creating 3D models based on 2D images. Tools such as Didimo[22] and Kaedim3D[23] allow users to upload 2D images and generate rigged 3D models from them within minutes. These are processes that would otherwise have taken days if not weeks. There are also AI tools that use scans of objects and people to create 3D models, including AgiSoft Metashape, iGoodie[24] and many others. **Figure 3** presents a portrait used in 'Sancho's Journey' to create a rigged and animated 3D model.

[19] https:// scenario.com/
[20] https://layer.ai/
[21] https://www.adobe.com/creativecloud/3d-ar.html
[22] https://www.didimo.co/
[23] https://www.kaedim3d.com/
[24] https://www.igoodi.eu/

Figure 3. Portrait of King George II at Marble Hill House (left) and 3D generated model using AI tools (right)

Auto riggers and animators such as Mixamo[25] have been around for some time and are constantly improving. They allow developers to upload 3D models and have them rigged and animated in under a minute. Whilst they are not great for bespoke characters, they work exceptionally well with bipeds. RADiCAL[26], Move AI[27] and the works of AI4Animation[28] on deep learning for animation provide new ways of animating characters (bipeds and quadrupeds) quickly and accurately (Starke et al., 2021; Zhang et al., 2018).

The concept of AI sound and music generation has been the topic of research for many years (Yin et al., 2023; Daylamani-Zad et al., 2006). The current tools are reaching new heights with adaptive music and sound effects that change as the experience unfolds. Jukebox (Dhariwal et al., 2020) and MuseNet[29] by OpenAI[30] provide excellent

[25] https://www.mixamo.com/
[26] https://radicalmotion.com
[27] https://move.ai/
[28] https://github.com/sebastianstarke/AI4Animation
[29] https://openai.com/index/musenet/
[30] https://openai.com/

groundings, whilst tools such as Soundraw[31], Mubert[32] and Loudly[33] enable developers to reduce the time and cost of music generation. Tools such as ElevenLabs[34] can generate a wide variety of sound effects that can facilitate the development process.

Object detection and spatial awareness is crucial to MR experiences. They bring a new dimension to spatial computing which allows for higher levels of realistic interaction between users, holographic and physical objects. With most current devices (smartphones and HMDs) benefitting from high quality cameras and depth detection, it is becoming an expected feature of any MR experience. Object detection and spatial awareness tools such as VisionLib[35], Azure AI Vision and Open CV[36] have paved the way to achieving real-time spatial awareness being able to simulate the behaviour of virtual models as if they were real, for example implementing occlusion. Whilst these do have a setup process and require some learning, the benefits are undeniable. **Figure 5** presents an example of spatial awareness implementing occlusion for the 3D holograms which are displayed behind the furniture as part of an Immersive Mixed Reality experience in a 18th century house, 2023, at Historic Deerfield, Massachusetts [further info redacted]. The image is taken as a snapshot from a user's MR HMD.

Interactions and coding can also benefit from AI. Tools such as Modl AI[37] , Github Co-pilot[38] and ChatGPT generate code after a prompt and facilitate many scripting challenges, especially for the less experienced developers. They also include several other features that streamline the coding processes.

Lighting in MR is an aspect that has not yet had the same level of support from AI-powered tools as other areas. There are obvious visual challenges that hinder this. However, techniques for adjusting lightning offline and in real time in the relevant development environments using features such as raytracing, real-time lighting

[31] https://soundraw.io/
[32] https://mubert.com/
[33] https://loudly.com/ai-music-generator
[34] https://elevenlabs.io/
[35] https://visionlib.com/
[36] https://opencv.org/
[37] https://modl.ai/
[38] https://github.com/features/copilot

management and light baking use AI-powered tools under the hood which have helped enhance the visual quality of MR artefacts.

Figure 4. Spatial awareness for 3D holograms demonstrating occlusion at Historic Deerfield, MA.

Figure 5. Spatial awareness for 3D holograms demonstrating occlusion at Historic Deerfield, MA.

Connecting and integrating external devices such as IoT sensors and actuators is an aspect of MR development that offers exciting possibilities for making the blending between the physical and virtual seamless. NLP and NLU approaches can help provide better interfaces between IoT devices and MR artefacts, leading to more human-centred integration for MR experiences. Platforms such as Azure AI, IBM Watson[39] and Google's Cloud IoT Core[40] provide facilities for seamless integration of IoT devices to a range of artefacts and experiences.

[39] https://www.ibm.com/watson
[40] https://cloud.google.com/iot-core

3.3. EXPERIENCE STAGE

The MR experience development pipeline does not finish when the final touch on the experience is put. There is much to be considered during the run of the experience as the users go through it. AI-enhanced motion and gaze tracking can help finetune the users' walkthrough. It is possible to create deeper engagement by having an accurate sense of the behaviour and level of engagement of the user and respond to it accurately. Whilst motion tracking and gaze tracking are very accessible and relatively easy to integrate for MR experiences, AI pipelines such as Azure AI, Google's Cloud AI and IBM Watson create quick and easy solutions to make sense of the data, automate responses and adjustments to the experience in real-time. The monitoring and control of the experience through AI-enhanced features paves the way for adjusting the position, rotation, and colour of assets, adjusting the volume and positioning of sound sources, and adapting the virtual content to the different lighting levels of the real environment.

4. GUIDELINES FOR CHOOSING AI TOOLS OR SOLUTIONS

Considering the possibilities provided by current state-of the-art AI tools, it becomes clear that there are many aspects of design and development that can be improved or optimised. The challenge however, it how to decide when to use them and how to choose the most suitable tool for each task to accomplish an immersive experience. This challenge is further highlighted by the notion that depending on the task, the goal, and the vision, AI solutions may not always be the best approach. Rather at times, the "traditional" approach may produce a better result, faster. Taking an example where an artefact requires a mythical/historical character with no existing images or examples in a very specific art style, it would be very difficult to produce a 3D model for this using AI tools. However, the same can be achieved much easier and at a high quality by 3D artists. Therefore, it is important to make the decisions with trepidation and considering all the factors that affect the decision. **Figure 6** presents a five factor guide towards approaching these decisions to assist designers and developers with making the right decisions for their project.

Figure 6. Five factors to consider as a guideline for deciding to use AI tools and how to choose between them.

These factors should be considered as a whole and not necessarily steps. Whilst availability is the first factor to consider, the rest may have advantages and disadvantages. Depending on the vision, the task, available team, timeframe and budget, the order and importance of the factors may change. Hence, the factors should be considered as a holistic guideline which should be tailored, prioritised and ordered based on the project, its audience, gaols and requirements.

4.1. AVAILABILITY

When considering a task to be achieved for the design or development of the MR artefact, the first step is to consider if there are any AI tools or solutions for the task and if so, how easily are they available. There are tools that are available as plug and play with extensive documentation, yet there are also libraries and tools that may need

customisation, integration, re-compiling and even training. It is important to consider these before embarking on adopting these approaches.

4.2. APPROPRIATENESS

Once it is established that there are available AI tools or solutions for the task at hand, it is prudent to consider if using the tool would facilitate the creativity or hinder it. The decision comes down to if it would be easier or faster to achieve the task without engaging the AI tools. If so, then the use of AI tools is acting as a hinderance to achieving the task. For example, using a tool such as DALL-E to create concepts of an environment, that there is a very clear vision of, may prove to be less effective than conceptualising it the "traditional" way. The process of describing the vision accurately and clearly enough for DALL-E for it to understand, and guiding it to get to the vision in mind may prove to be much more time consuming than conceptualising it in the first place. Therefore, it is crucial to consider if this tool is an appropriate means to achieving the task at hand.

4.3. LEARNING CURVE AND USABILITY

Many AI tools and solutions come with a learning curve. They may have their own taxonomy and terminology. They may require understating of some level of programming to make them work as desired. They may also require inputs such as an input dataset to train on. These are not necessarily extreme challenges. Most tools and solutions have comprehensive tutorials and documentations accompanied by extensive samples and support from the community of users. However, when deciding, it should be considered if the learning curve is affordable within the timeframe and budget of the project. A learning curve of two weeks to use a 3D model generator may not be reasonable in a five-week timeline.

4.4. AFFORDABILITY

Cost is always an important factor. Many cloud solutions such as Azure AI and Google Cloud services can come with various costs which may not be apparent up front. Many cloud-services charge per interaction which can round up to staggering costs. Other tools may offer initial free/affordable instances at the start but by the time the learning

curve has passed, and tests are done, the actual artefact requires licensing or purchases which may not fit in the budget. The risk is that time and staff have been invested in the learning and testing, therefore, by then the project may be committed to the costly path with no time to seek alternative solutions. Some AI solutions require high power computing to produce the desired results which is also a considerable cost and may not be easily available.

4.5. QUALITY

The quality of the assets is one the most important factors in an MR artefact. Small details, inconsistencies or defects can break the user immersion. Hence, it is paramount to consider, compare, and judge the quality of assets produced using AI tools. Looking at previous examples, seeking feedback from users can really help to gauge the quality of the assets produced. The quality may at times be the format or type of textures they produce. If the output format requires further adjustments or conversions, it may be an indication to consider the suitability of the solution. Any adjustment or conversion may result in loss of quality, compromises and, undoubtedly, time.

5. CONCLUSIONS

In this article we have provided a review of the AI tools available to MR designers and developers, their potential, and the challenges and opportunities related to their integration in the MR development pipeline for creative experiences. We outlined seminal work on AI architectures that pertain to MR development such as ML and DL algorithms for machine vision, media generation, and networking. After an in-depth introduction of the several AI tools available for every stage of MR development, the study concludes with a set of use guidelines for how to approach and integrate AI in projects.

As the discussion evolves around creative processes the relationship between human creativity and technology takes centre stage. Despite recent advances, AI cannot replicate the process of human creativity, a fuzzy concept in itself. There are significant challenges to overcome such as training in large sets of unbiased and diverse data, and

work with ill-defined problems, to name just the main ones. Instead, the biggest threat currently is the use of AI to create something for the sake of creating it instead of using it as a tool within a wider creative vision. The current synergy between human and AI in the creative sector is about using AI tools to assist and augment the imagination, problem solving capacity, divergent thinking, and innovative concept exploration, and speed up the delivery of the creative output. AI tools such as the ones presented in this article free creatives from repetitive technical tasks and challenges and allow them to focus on the creative part, while also democratise the creative processes for junior creatives. This synergy will be further explored and change as advances in AI will offer more variety and power that will highly influence the creative vision and outcome. Coupled with advances in MR, and the gradual shaping of the Metaverse, the potential impact on human creativity can be immense. Besides, human creativity applies also to the way that creatives will use AI tools even if these reach the holy grail of being truly creative machines. In the years to come, this will be an interesting space to watch.

6. ACKNOWLEDGEMENTS

The design and development of the two Immersive Mixed Reality Performances referenced in the article was supported by a joint Arts and Humanities Research Council (UK, AH/W005530/1) and National Endowment For the Humanities (US, HND-284975-22) research grant under the theme New Direction for Digital Scholarship in Cultural Institutions.

7. REFERENCES

Ali, O., Abdelbaki, W., Shrestha, A., Elbasi, E., Abdallah, M., Alryalat, A., & Dwivedi, K.Y. (2023). A systematic literature review of artificial intelligence in the healthcare sector: Benefits, challenges, methodologies, and functionalities. *Journal of Innovation & Knowledge*, 8(1).

Anantrasirichai, N., & Bull, D. (2022). Artificial intelligence in the creative industries: a review. *Artificial Intelligence Review, 55*, 589–656. https://doi.org/10.1007/s10462-021-10039-7

Anantrasirichai, N., Gilchrist, I.D., & Bull, D. (2016). Fixation identification for low-sample-rate mobile eye trackers. In: *IEEE international conference on image processing* (ICIP) (pp. 3126–3130). https://doi.org/10.1109/ICIP.2016.7532935

Bassyouni, Z., & Elhajj, I. (2021). Augmented Reality Meets Artificial Intelligence in Robotics: A Systematic Review. *Frontiers in Robotics and AI*, *8*, 724798. https://doi.org/10.3389/frobt.2021.724798

Bharadiya, J., & Thomas, R., & Ahmed, F. (2023). Rise of Artificial Intelligence in Business and Industry. *Journal of Engineering Research and Reports*, *25*, 100807. https://doi.org/10.9734/JERR/2023/v25i3893

Bissonnette, V., Mirchi, N., Ledwos, N., Alsidieri, G., Winkler-Schwartz, A., & Del Maestro, R. F. (2019). Artificial Intelligence Distinguishes Surgical Training Levels in a Virtual Reality Spinal Task. *The Journal of Bone & Joint Surgery, 101*(23), e127. https://doi.org/10.2106/JBJS.18.01197

Blattmann, A., Dockhorn, T., Kulal, S., Mendelevitch, D., Kilian M., & Lorenz, D. (2023). Stable Video Diffusion: Scaling Latent Video Diffusion Models to Large Datasets. *ArXiv* abs/2311.15127. https://doi.org/10.48550/arXiv.2311.15127

Daylamani Zad, D., Araabi, B. N., & Lucas, C. (2006). A novel approach to automatic music composing: Using genetic algorithm. In *International Computer Music Conference Proceedings*. MPublishing, University of Michigan Library, Vol. 2006, 551-555.

Dhariwal, P., Jun, H., Payne, C., Kim, J.W., Radford, A., & Sutskever, I. (2020). Jukebox: A Generative Model for Music. ArXiv, abs/2005.00341. https://doi.org/10.48550/arXiv.2005.00341

Engel, J., Resnick, C., Roberts, A., Dieleman, S., Norouzi, M., Eck, D., & Simonyan, K. (2017). Neural Audio Synthesis of Musical Notes with WaveNet Autoencoders. ArXiv, abs/1704.01279. https://doi.org/10.48550/arXiv.1704.01279

Gafni, G., Thies, J., Zollhofer, M., and Nießner, M. (2021). Dynamic neural radiance fields for monocular 4d facial avatar reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 8649–8658.

Hensen, B. (2023). A Systematic Literature Review of Mixed Reality Learning Approaches. In: De Paolis, L.T., Arpaia, P., Sacco, M. (eds) *Extended Reality. XR Salento 2023. Lecture Notes in Computer Science*, vol 14219. Springer, Cham. https://doi.org/10.1007/978-3-031-43404-4_2

Hong, W., Ding, M., Zheng, W., Liu, X. and Tang, J. (2022). CogVideo: Large-scale Pretraining for Text-to-Video Generation via Transformers. ArXiv:2205.15868. https://doi.org/10.48550/arXiv.2205.15868

Hu, Z., Li, S., Zhang, C., Yi, K., Wang, G., & Manocha, D. (2020). Dgaze: Cnn-based gaze prediction in dynamic scenes. *IEEE transactions on visualization and computer graphics*, *26*(5), 1902-1911.

Huynh-The, T., Hua, C. H., Tu, N. A., & Kim, D. S. (2020). Learning 3D spatiotemporal gait feature by convolutional network for person identification. *Neurocomputing*, *397*, 192-202.

Israelsen, B., Ahmed, N., Center, K., Green, R., and Bennett, W. (2018). Adaptive Simulation-Based Training of Artificial-Intelligence Decision Makers Using Bayesian Optimization. *Journal of Aerospace Information Systems, 15*, 38–56. https://doi.org/10.2514/1.I010553

Jiang, Y., Hieu Tran, T., & Williams, L. (2023). Machine learning and mixed reality for smart aviation: Applications and challenges. *Journal of Air Transport Management*, *111*, 102437. https://doi.org/10.1016/j.jairtraman.2023.102437

Khatri, P. (2021). The Future of Automatically Generated Animation with AI, In *Deep Learning in Gaming and Animations*. CRC Press.

Lai, P. K., Xie, S., Lang, J., & Laganière, R. (2019). Real-time panoramic depth maps from omni-directional stereo images for 6 dof videos in virtual reality. In *IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 405-412.

LeCun Y., Bengio Y., Hinton G. (2015). Deep learning. *Nature*, *521*(7553), 436-444.

Liu Y., Rai R., Purwar A., He B., Mani M. (2020). Machine learning applications in manufacturing. *Journal of Computing and Information Science in Engineering*, *20*(2), Article 020301.

Ming Tang, Y., Ting Kuo, W., Lee, C.K.M. (2023). Real-time Mixed Reality (MR) and Artificial Intelligence (AI) object recognition integration for digital twin in Industry 4.0. *Internet of Things*, *23,* 100753. https://doi.org/10.1016/j.iot.2023.100753

Ojeda-Castelo, J. J., Capobianco-Uriarte, M. D. L. M., Piedra-Fernandez, J. A., & Ayala, R. (2022). A survey on intelligent gesture recognition techniques. *IEEE Access*, 10, 87135-87156.

Pine, J., & Gilmore, J. H. (2011). *The experience economy*. Harvard Business Press.

Poole, B., Jain, A., Barron, T.J., & Mildenhall, B. (2022). Dreamfusion: Text-to-3d using 2d diffusion. arXiv:2209.14988. https://doi.org/10.48550/arXiv.2209.14988

Rathore B. (2023). Digital Transformation 4.0: Integration of Artificial Intelligence & Metaverse in Marketing. Eduzone: *International Peer Reviewed/Refereed Multidisciplinary Journal*, *12*(1), 42–48.

Reiners, D., Davahli, M. Reza, Karwowski, W., & Cruz-Neira, C. (2021). The Combination of Artificial Intelligence and Extended Reality: A Systematic Review. *Frontiers in Virtual Reality*. 2. 721933. https://doi.org/10.3389/frvir.2021.721933

Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M. Justus, J., Engel, P., & Harnisch, M. (2015). Industry 4.0: The future of productivity and growth in manufacturing industries. *Boston consulting group, 9*, 1, 54-89.

Soccini AM (2017). Gaze estimation based on head movements in virtual reality applications using deep learning. In: *IEEE virtual reality (VR),* 413–414.

Starke, S., Zhao, Y., Zinno, F., & Komura, T., 2021. Neural animation layering for synthesizing martial arts movements. *ACM Transactions on Graphics (TOG), 40*(4), 1-16.

Uriel, S., Polyak, A., Hayes, T., Yin, X., An, J., Zhang, S., Hu, Q., Yang, H., Ashual, O., Gafni, O., Parikh, D., Gupta, S., & Taigman, Y. (2022). Make-A-Video: Text-to-Video Generation without Text-Video Data. *ArXiv* abs/2209.14792. https://doi.org/10.48550/arXiv.2209.14792

Wang, Z., Lu, C., Wang, Y., Bao, F., Li, C., Su, H., and Zhu, J. (2023). Prolific dreamer: High fidelity and diverse text-to-3d generation with variational score distillation. arXiv:2305.16213. https://doi.org/10.48550/arXiv.2305.16213

Wu, J.Z., J., Ge, Y., Wang, X., Lei, W.S., Gu, Y., Shi, Y., Hsu, W., Shan, Y., Qie, X., & Shou, Z.M. (2023). Tune-A-Video: One-Shot Tuning of Image Diffusion Models for Text-to-Video Generation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 7623-7633.

Yeh, C. H., Huang, C. H., & Kang, L. W. (2019). Multi-scale deep residual learning-based single image haze removal via image decomposition. *IEEE Transactions on Image Processing, 29*, 3153-3167.

Yin, Z., Reuben, F., Stepney, S., & Collins, T. (2023). Deep learning's shallow gains: A comparative evaluation of algorithms for automatic music generation. *Machine Learning, 112*(5), 1785-1822.

Zawacki-Richter, O., Marín, V.I., and Bond, M. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, *39.* https://doi.org/10.1186/s41239-019-0171-0

Zhang, H., Starke, S., Komura, T., & Saito, J. (2018). Mode-adaptive neural networks for quadruped motion control. *ACM Transactions on Graphics (TOG), 37*(4), 1-11.

Zhu, Y., Zhao, C., Guo, H., Wang, J., Zhao, X., & Lu, H. (2018). Attention CoupleNet: Fully convolutional attention coupling network for object detection. *IEEE Transactions on Image Processing, 28*(1), 113-126.