Conceptualising Cinematic Augmented Reality

Jason S. Kao
Inland Norway University of Applied Sciences, Norway
Ole Christoffer Haga
Inland Norway University of Applied Sciences, Norway
Meisam Taheri
Inland Norway University of Applied Sciences, Norway

Abstract

Augmented Reality (AR) is considered as an innovative medium that presents novel experiences by overlaying virtual content over a physical space. Current AR narrative applications are mainly location-based and function-orieanted for users to experience the storytelling. The cinema-centred AR for group viewing has remained undiscovered. This work conceptualises Cinematic Augmented Reality (CAR) through discussing the potential implementation of AR imaging over a viewing screen to form a believable cinematic composition. We discussed the advantages of applying CAR as the next generation filmmaking from the aspects of cinematic realism, field of view and the potential interactivity. The proposed approach, Augmented Dimension (AD) led to the identifications and alternations of scene design techniques, including advanced storyboarding, adaptable field of view, and a hybrid production pipeline. The result presents insights into the effective integration of AR with conventional cinematic techniques, and proposing a fundamental framework for future explorations in AR-based cinema.

Keywords: Augmented reality, Extended reality, Cinematic virtual reality, Cinematic augmented reality, Augmented dimension.

Introduction

The advancement of Augmented reality (AR) is allowing wider audiences to experience immersive and interactive storytelling by superimposing virtual elements in physical space. Current AR narrative practices mainly focus in location-based applications such as guide tours at specific locations, museum experience enhancements or portal spaces for entertainment. Shilkrot et al. described three major types of AR narratives as Situated, Location-based and World-level through the scale of space. While Situated requires the user to explore the AR storytelling at specific installations, such as art gallery or museum, and Location-based encourages users to explore the AR storytelling in certain locations such as parks, campuses or historical sites. World-level narrative invites users for mass-participatory interactive narratives to accomplish tasks by roaming the streets which led users to drive the progress of the storytelling (Shilkrot, Montfort, and Maes 2014). Shin et al. reviewed 64 AR storytelling works and categorized four types of virtual-real connections in augmented narrative space: Pairing, Registration, Function and Placement rule. The study expanded the types into 11 patterns to define the approaches and purposes of the visual narrative space (CHI). These works indicate current implementations of AR as a storytelling medium are varied and mainly serving as function-oriented applications. For cinematic experience developers, the utilization of AR as a medium for cinema has remained unclear. In contrast, Cinematic Virtual Reality (CVR) as a term to define VR as the medium for the cinematic experience has been widely adopted and developed by practitioners. The viewer watches omnidirectional movies through VR devices and determines preferred viewing directions (Rothe, Buschek, and Hußmann 2019). However, CVR is criticized for its limitations for cinematic experience such as missing points of interest (POI) and the restriction of individual viewing. Kao indicated certain attention guidance is needed in CVR due to critical diegetic elements can be missed due to the freedom of the omnidirectional viewing. In the aspect of cinematic experience creation, AR has potential advantages in increasing viewers narrative immersion, engagement and group viewing (Kao and Kao 2022). AR allows the virtual content to be extended into the user's physical environment. Unlike virtual reality (VR) in which the user is immersed in a virtual world without seeing the real world, AR merges CGI and animations with the real world and allows tangible interaction with virtual objects. This study conceptualised Cinematic Augmented Reality (CAR) by proposing the potential production process of Augmented Dimension (AD) to initialise a more cinema-centered perspective of implementing AR for cinematic experience creation.

Cinematic Realism and Augmented Dimension

Over a century, audiences watch motion pictures on a rectangle screen that the fiction world is presented in. The cinematic realism is presented by various aspects such as storytelling techniques, acting, fictional world buildings and the theatrical advancements such as large-format projection, high resolution and high frame rate filming, digital sound system and stereoscopic third dimension. The enhanced cinematic realism unveils a self-enclosed fictional world to audiences and strengthens a more absorbing and powerful sensory viewing experiences. The Stereoscopic three-dimensional (S3D) film was a breakthrough attempt to bring cinematic realism to a different level. However, S3D did not fulfil audiences' expectations that content can be perceived as so-called

three-dimensional as objects cognized in physical reality. Such expectation of the ideal three-dimensional effect is far beyond that photorealism can achieve. The realism of S3D can be considered an extension of photorealism. It is more an alteration of photorealism that presents a photographic-based illusion (Prince 1996).

A New Level of Realism: Perceptual Realism

With the thriving of extended reality (XR), a realism revolution in the entertainment industry has been initiated by practitioners. The potential of perceptual realism brought by XR is conceivable. Digital imaging theorists and artists deemed perceptual realism an effect that can be reconstructed with digital tools (Prince 1996). However, cognitive scientists and philosophers define perceptual realism as a sense view that physical objects exist independently to the perceivers. From the aspect of philosophy of perception, it is similar to naïve realism that the senses provide us with direct awareness of objects as they really are (Le Morvan 2018). In other words, perceptual realism is a realism that is based on what we see, smell, hear, taste and touch in real life, while most digital imagery is the photorealism based on what the camera sees (Turnock 2012). Thus, the realism in XR can be considered as a different level perception compared with the conventional cinema as it creates the illusion that the perceived virtual CGI objects are frameless-base, and they are relatively independent to the viewer in the same reality. CVR is gaining attention in the filmmaking industry as a new medium for storytelling. Issues such as attention guidance and motion sickness in CVR has been widely discussed to reveal a further clear potential. Prior research indicates that AR creates a unique experience given by the viewer's respective physical reality and thus increased the sensation of involvement (Shin and Shin 2011). The borderless feature in AR stimulates the viewer's psychological and emotional responses to immerse in both the physical reality and the virtual content (Dumic, Grgic, and Grgic 2010). That is, AR's feature of placing virtual content in physical reality creates the foundation of perceiving perceptual realism particularly in the condition of utilising AR glasses instead of mobile screen-based devices. A clear vision in the CGI industries is to create realistic rendered objects that audiences cannot distinguish with real objects (Lestari et al. 2022). The high-resolution realistic rendering in AR devices can be expected in near future.

Cinematic XR

With the declining popularity of S3D films and the growing formation of adapting game engine for filmmaking, the new cinematic realism is in time to be explored to push the boundary of cinematic immersion. Several CVR productions have been published and new theories and approaches are studied to investigate the compositions, point of interest and the attention guidance issues under CVR productions. The series film Invisible directed by Doug Liman, explores several traditional filming techniques in the CVR space, such as parallel editing, camera movement, mask application and the blocking of actors (Liman 2016). One critical advantage of Cinematic Augmented Reality (CAR) compared with CVR is that the audience still sees the physical reality. This bridges conventional viewing tendency with the new technology and allows group viewing while CVR is based on the sole viewer. Meanwhile, Metaverse has been introduced by Meta in early 2020s which indicates the potential mass adoption on Mixed Reality (MR) headset that may change how people work, communicate, and entertain. Based on these trends, this paper discusses the potentials of CAR through proposed approach, pipeline and evaluation.

Augmented Dimension

AR's feature of presenting virtual imaging in the physical world founds the ground to build an extended dimension over a conventional viewing screen. The extended dimension associated to the content showing in the viewing screen that could be a cinema screen, a desktop display or a smart mobile display. This extended dimension associated to the viewing screen, this work gives it a term, Augmented Dimension (AD). AD allows augmented content correlated to the content in the viewing screen to be placed over or beyond the viewing screen to bring visual impact to the viewer, and meanwhile support the storytelling with potential interactivity. In other words, AD requires a viewing display such as cinema screen, TV, or a mobile device to associate with the AR content. The viewer sees through AR glasses to perceive the composition formed by the viewing screen and the AR content. Imagine if a character or an object can be outside of the screen frame and be close to the viewer, such as a football is kicked to outside of the screen and dropped on the floor in the cinema and eventually bounced back into the screen world, or the devil in a horror movie can crawl out from the screen and walk toward to the audience. The AR content can also be designed to cover the whole physical environment. For example, a spaceship's interior (AR content) fulfils the viewer's physical room and leaves only the viewing screen as the frame of canopy. These scenarios can by no means happen in S3D films. The following section discusses the potential workflow of the approach to integrate AR content over conventional screen content to demonstrate a believable visual composition and moving images. The goal is to identify the potential of this approach as the next generation cinematic content creation.

The Hybrid Pipeline

The development of AD integrates the processes of filmmaking, visual effects (VFX) in the game engine and the AR application development. The proposed pipeline is based on the workflow of game production including pre-production, production, calibration, and testing. Calibration is a specific phase added in the pipeline for the accurate alignments between AR content, screen, and targeted viewing environments. See Image 1.
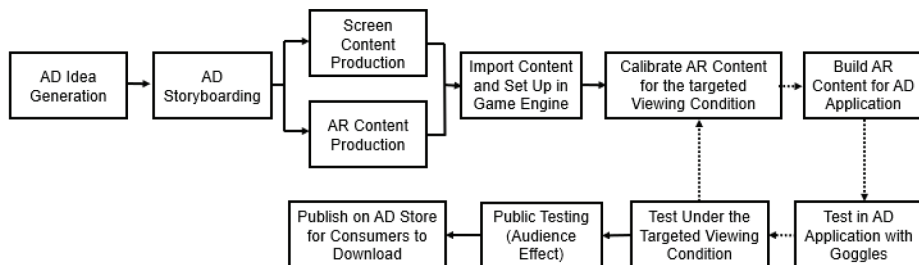
Image 1 – The potential pipeline of AD

Beginning with the idea generation, AD developers brainstorm creative visual compositions to increase the impact of storytelling. The storyboarding for AD should express the association between the AR content, the screen content, as well the viewing environment to visualize what the viewer will see through the AR glasses. Therefore, a new storyboard layout will be developed to express the spatial motion design. In the production phase, regular filmmaking process will be implemented to create the on-screen content while AR content is produced as CGI assets for VFX compositions. Both on-screen and AR content then be imported into the game engine for the scene setup. AD artists must consider a layout that is based on the targeted viewing environment as a performance stage, and the screen will be the major focal center for the viewers. The scale and the motion of AR objects, distances to the screen and the viewers will be precisely planned and calibrated for the best visual impact and comfort. The calibration process requires testing with application platform and the AR glasses. The final build will be shipped to relevant online stores for viewers to download. See Figure 2 for the proposed pipeline for AD production. Viewers need to wear AR glasses with AD applications installed to watch the associated on-screen content. The AD application will sync the on-screen playtime and trigger the AR content in the AR glasses in due course.

The Fourth Wall and Field of View

Breaking the Fourth Wall

A major strength of AD in relation to inducing immersive experiences is the ability for its content to (perceivably) transcend the cinema screen. In other words, breaking the fourth wall. Conventionally, three boundaries form the performing area, one behind the set and two on the sides, while the fourth wall is the imaginary one that performers face toward to the audience or cameras. If actors address or pay attention to the audience, it is considered the break of the fourth wall that separates the fictional and the real space. The playwright Brecht used this to provoke reflection and critical engagement in the audience by disrupting the immersion in fiction (Silberman, Giles, and Kuhn 2015). In addition to the Brechtian distancing effect, breaking the fourth wall results "closeness effect"

described by Wijer (Wijers 2018). The audience can experience united with the fictional world while cognise that it is mediated. She argues that active participation stimulates an intellectual reflection on the fictional content where the viewer, along with an awareness of the medium, feels empathy and immersion with what is presented. For example, when a fictional character addresses the audience it can, as much as an alienating effect, also be experienced as a personal connection with the character. This might correspond to Murray's concept of transformation as how new media can allow users to interact with stories rather than to merely witness them (Murray 1997). This refers to the malleability provided by real-time media that can adapt what is being presented to user actions. However, in media such as VR, it enhances users' sense of presence by acknowledging through diegetic characters. Bucher compares VR to immersive theater, a form of performance where actors leave the traditional stage, move out among the audience, and often interact with them (Bucher 2018).

Position AR Content

AR works differently than VR since digital content does not cover the user's vision entirely but is perceived as superimposed on top of the physical world. The medium can still contribute to a form of immersion by augmenting the physical world with digital content. AR not only blends the virtual with the real world, but it also often adapts to it as well. Commonly, something visible in the physical world is recognized by an application as an AR marker triggering digital content to be spawned. Additionally, the positioning of the digital content corresponds in real-time to the field of view presented in the appliance used through head-tracking (Wereszczynski, Cyrana, and Nwobodo 2023). The accuracy of the perpetual linkage between what is seen through the camera, and the representation of the digital content, strengthens the illusion of augmenting the user's physical reality.

As noted by Liestøl, AR provides situated simulations where the representation of digital content aligns with the physical space (Liestøl 2018). When users turn their head wearing AR-glasses, the perceived distance and perspective of the superimposed graphics changes accordingly to nurturing the illusion of being part of the real world. Regarding space, this is achieved by

technology measuring the physical environment optically and inertially. Liestøl demonstrates the challenges around coordinating real-world time and "scenic" time. The latter implies that time represented on film equals the duration of the sequence of fictional events, for example, being a one-take shot without cutting, or manipulating footage speed (Genette 1980; Liestøl 2018). In AD, it is critical to consider this regarding storytelling. In cinema, montage can for example condense time so a time span of years can be told within minutes. For AD cinema there needs to be a synchronization of time represented and the users' real-world time. If digital content at one moment 'leaves' the screen and is perceived to be a part of the users' physical environment, there will most likely be inappropriate to cut to another shot before the content is brought back to the screen. Otherwise, the users might perceive something being abruptly removed from their physical reality, which would interfere with the illusion. But there might of course be occurrences where digital content could leave the screen and return into the next shot after cutting. The most important here in relation to AD, is that the triggering of AR content should be timed properly according to screen content, such as at a specific time code in the film, and always be adapted to a real-time experience for the user.

Another aspect to consider is how much head movements to expect from the users, considering them being seated together in a cinema auditorium.

According to Lim & Lee, motion sickness in VR increases with a higher Field of View (FOV), such as the angle of what is observable from a particular vantage point, measured in degrees (Lim and Lee 2023). A restricted FOV offers a more comfortable experience and McCurley suggests 154 degrees when allowing head turn (McCurley 2016). For a viewing experience seated in a movie theater, we assume this angle would be appropriate for AD. See Image 2.
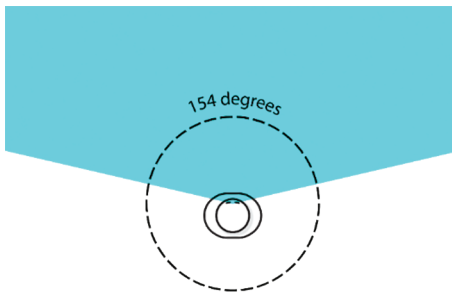

Image 2 – The comfortable FOV of AD

Innovative Storyboarding

With its legacy in cinema, AD can draw on many of the same development techniques, but with adjustments to accommodate the expanded dimensions. Storyboarding is a common activity when planning and conceptualizing movies. In XR, however, users can perpetually change their FOV, meaning that conventional rectangular sketches can become insufficient. Isometric 3D renders and/or overhead sketches can provide a more complete visualization in this regard since they can indicate both the spatial aspects and the user's perceived vantage point (Haga 2024; Vindenes and Nyre 2023). Recently, tools have also emerged where developers can create spatial storyboards inside virtual environments. One example is ShapesXR, an XR-based application for XR experience developments, where users outline immersive narratives or experiences within XR, either by drawing freehand or placing and manipulating 3D objects or other digital content. The occurrence of the assets can be linked to a timeline or user interactivity that makes it easier to predict how a finalized experience will unfold. Image 3 shows a way of pre-visualisation an AD cinematic experience, showing both the delimited cinema screen and the AR content, as well as indicating the audience vantage point.
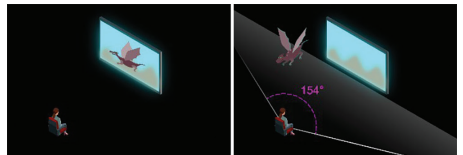

Image 3 – Pre-visualisation of AD

The concept of AD is based on diegetic existents being perceived to be moving out of the cinema screen. This implies that parts of the viewer's physical space will contain cinematic elements, which need to be conceptualized. For AD, storyboarding could draw from the preproduction pipeline of VR. For example, Vindenes and Nyre suggest spherical images providing an overhead visualization of the virtual environment (Vindenes and Nyre, 2023). To complement, they also propose stretching the field of view to equirectangular images, as common when visualizing the rounded Earth as a map. Compared with VR, AD does not require a 360 degrees viewing angle. If we adapt McCurley's suggestion for comfortable FOV, AD storyboarding could utilize isometric or top-down images, only reducing VR's 360-degrees FOV to 154-degrees.

Production Platforms and Interactivity Model

The concept of XR is to engage viewers in an environment where they can sense the presence of an actor/3D model and potentially to interact with it. This actor, represented by a 3D model, seemingly emerges from the screen, creating an immersive experience propelled by technical components. In this study, we aim to leverage the capabilities exclusive to mixed reality, which include the fluid integration of actual and virtual environments, and the real-time interaction with both tangible and digital objects.

Our proposed concept of CAR aims to create an extended dimension and enable viewers to explore a novel dimension of the extended reality. Described in our

proposed pipeline, the workflow integrates filmmaking and AR application process. The screen content and AR content will be calibrated in the game engine and result a convincing composition to the viewers, further enhancing their immersive experience. By seamlessly blending the physical and virtual worlds, we hope to revolutionize the viewing experience. This project attempts to push the boundaries of what is currently achievable in the world of XR, potentially setting a new standard for future cinematic experiences.

Unreal Engine 5 (UE5) is a powerful tool for creating immersive XR experiences and provides sufficient tools for real-time film productions. UE5 is a game development engine used widely for creating interactive 3D, 2D, VR, and AR experiences allowing users to integrate 3D models, animations, sounds, and lighting effects and giving the ability to control every aspect of the VR environment. In the context of AR or XR, these platforms can be used to create AR scenes with and without markers. When a particular visual marker—like a QR code—is observed through a camera, marker-based augmented reality (AR) is activated. Conversely, marker-less augmented reality leverages a device's sensors, accelerometers, and gyroscopes to deliver an AR experience that is contingent upon the user's movement or position.

A key component of XR experiences is interaction, which determines how users will interact with the 3D representations. Depending on the intended level of immersion and the available technology, this might require a range of inputs. For instance, gestures can be utilised with portable controllers that capture hand and finger motions or in virtual reality settings where users wear gloves. The direction of the user's gaze affects how they interact with the 3D models in both VR and AR experiences. This technology is known as gaze input. When a person or item approaches a particular 3D model, proximity sensors may recognize that proximity and initiate actions. Additional input methods may include physical buttons on portable controllers, voice instructions, or, in more sophisticated configurations, brain-computer interfaces. Ultimately, the choice of interactivity inputs will depend on the desired user experience and the technical feasibility of implementing these inputs in the XR application. Image 4 demonstrates how sensory feedback enhances the immersive experience and seamlessly integrates different scenes. Proximity sensors and location-based triggers allow interactions between the viewer and the actors in the film, activating various sensors based on the film's actions. This innovative approach ensures that the sensory stimuli are precisely synchronized with the narrative, providing a highly engaging and realistic experience. The fluid integration of these elements allows viewers to feel as though they are an integral part of the story, bridging the gap between passive observation and active participation.
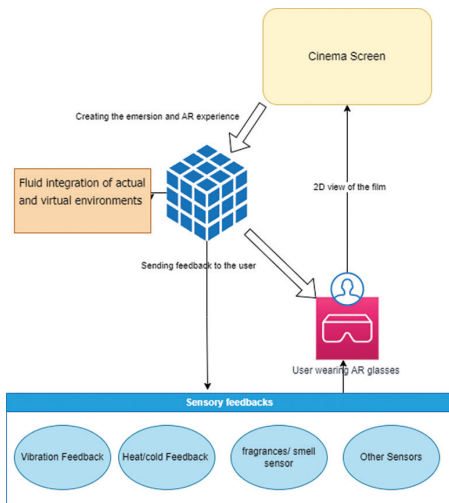

Image 4 – Potential interactivity model in CAR

Conclusion

The emergence of extended reality brought potentials for filmmakers to break through a different level of realism. The potential of AR to be applied in cinema is greater than VR since its holistic feature that audiences still see the physical world. Comparing with S3D, the three-dimensional perception in AR can be overwhelming and the dimension outside of the screen frame can be limitless. The new dimension of storytelling is revealed and the extended reality as a critical role for the next generation cinematic realism would be inevitable.

References

Bucher, John. 2018. *Storytelling for Virtual Reality: Methods and Principles for Crafting Immersive Narratives* (Routledge: New York).

Dumic, Emil, Sonja Grgic, and Mislav Grgic. 2010. 'Comparison of HDTV formats using objective video quality measures', *Multimedia Tools and Applications*, 49: 409-24.

Genette, Gérard. 1980. *Narrative discourse* (Cornell University Press: Ithaca, New York).

Haga, Ole Christoffer. 2024. 'Shifting Diegetic Boundaries.' in Kath Dooley and Alex Munt (eds.), *Screenwriting for Virtual Reality: Story, Space and Experience* (Palgrave Macmillan).

Kao, J. S., and K. Kao. 2022. "The Effect of Attention Guidance and the Potential of Cinmatic Augmented Reality in Narrative Immersion." In *2022 IEEE Games,Entertainment,MediaConference(GEM)*,1- 5.

LeMorvan, Pierre. 2018. "Chapter Nine - Perceptual Realism's Fundamental Forms." in John Smythies and Robert French (eds.), *Direct versus Indirect Realism* (Academic Press).

Lestari, Budianto, Setiawan Slamet, Retnaningdyah Pratiwi, Barus Pijar Krupskaya, Ningsih Bilqis Aurell Widya, and Amelia Diah Riska. 2022. 'The Power of The Computer-Generated Imagery (CGI) in Avengers Endgame Movie: Hyperreality Perspective', *Ethical Lingua: Journal of Language Teaching and Literature*, 9.

Liestøl, Gunnar. 2018. "Story & Storage – Narrative Theory as a Tool for Creativity in Augmented Reality Storytelling.", *Virtual Creativity*, 8:75-89.

Lim, Chae Heon, and Seul Chan Lee. 2023. 'The Effects of Degrees of Freedom and Field of View on Motion Sickness in a Virtual Reality Context', *International Journal of Human-Computer Interaction*: 1-13.

Liman, Dong. 2016. "Invincible." In. USA: Jaunt.

McCurley, Vincent. 2016. "Storyboarding in Virtual Reality." In *Virtual Reality Pop (Medium)*.

Murray, Janet Horowitz. 1997. *Hamlet on the holodeck: the future of narrative in cyberspace* (New York : Free Press, [1997] © 1997).

Prince, S. 1996. 'True Lies - Perceptual Realism, Digital Images, and Film Theory', *Film Quarterly*, 49: 27-37.

Reipschläger, Patrick, Severin Engert, and Raimund Dachselt. 2020. "Augmented Displays : Seamlessly Extending Interactive Surfaces With Head-Mounted Augmented Reality." In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–4., Honolulu, USA: Association for Computing Machinery.

Rothe, Sylvia, Daniel Buschek, and Heinrich Hußmann. 2019. 'Guidance in Cinematic Virtual Reality-Taxonomy, Research Status and Challenges', *Multimodal Technologies and Interaction*,3.

Shilkrot, Roy, Nick Montfort, and Pattie Maes. 2014. "nARratives of augmented worlds." In *2014 IEEE International Symposium on Mixed and Augmented Reality - Media, Art, Social Science, Humanities and Design (IMSAR-MASH'D)*, 35-42.

Shin, Dong-Hee, and Youn-Joo Shin. 2011. 'Why do people play social network games?', *Computers in Human Behavior*, 27: 852-61.

Silberman, Marc, Steve Giles, and Tom Kuhn. 2015. *Brechton theatre* (Bloomsbury Publishing:London).

Turnock, Julie. 2012. 'The ILM Version: Recent Digital Effects and the Aesthetics of 1970s Cinematography', *Film History*, 24: 158-68.

Vindenes, Joakim, and Lars Nyre. 2023. 'Prototyping first- person viewer positions for VR narratives with storyboards and pilot productions', *Journal of Screenwriting*, 14: 251–69.

Wereszczynski, Kamil, Krzysztof Cyrana, and Onyeka J. Nwobodo. 2023. 'A review on tracking head movement in augmented reality systems', *Procedia Computer Science*, 225: 4344–53.

Wijers, Eva. 2018. 'Emersive storytelling: An exploration of animation and the fourth wall as a tool for critical thinking', *Animation Practice, Process & Production*, 7: 41-65.