

# INSIDE IMAGE

## TECHNICAL NOTES FOR VIRTUAL STORYTELLING

**Daniele Rossi**  
University of Camerino  
School of Architecture and Design  
[daniele.rossi@unicam.it](mailto:daniele.rossi@unicam.it)

## ESSAY 23/01

VIRTUAL REALITY  
VIRTUAL STORYTELLING  
IMMERSION  
PRESENCE  
EMBODIMENT

The use of virtual reality (VR) to tell stories, and more specifically, the possibilities of constructing 3D spaces to be experienced through an Head Mounted Display (HMD) allows empathy levels that are difficult to reach through other mediums. These levels are concretized in some characteristic VR attributes so they can define different layers of immersion that can be funda-

mental in the narrative setting. Knowing how to distinguish these attributes by recognizing their peculiarities is the starting point for anyone who wants to experiment with VR design. Immersion, presence, embodiment, first-person shot, point of interest and continuity are some of the terms that need to be understood to approach new forms of VR storytelling.

## INTRODUCTION

With the commercialization of new and more effective head mounted displays (HMDs), the potential market for Virtual Reality (VR) has expanded massively. This expansion invites storytellers from traditional genres such as television, film, marketing and, of course, video games to take advantage of the new and effective storytelling techniques offered by VR.

One of the challenges that VR designers will face is the need to strike a balance between the user's agency and the narrative model adopted.

The most immersive VR experiences allow viewers to interact with environments and acquire a feeling of autonomy while making sense of them.

The traditional cinematic lexicon does not contain terms such as interactivity and agency. For this reason, it is necessary to introduce new metrics capable of understanding them. Additionally, storytelling in general needs to be revisited because the immersive experiences of VR have greatly expanded the narrative potential.

The use of virtual reality to tell stories, and more specifically, the possibilities for constructing 3D spaces to be experienced through HMDs, allows empathy levels that are difficult to reach by other mediums.

These levels are concretized in some characteristic VR attributes to define different layers of immersion, and they can be fundamental in the narrative setting. Knowing how to distinguish these attributes by recognizing their peculiarities is the starting point for anyone who wants to experiment with VR design.

## IMMERSION, PRESENCE AND EMBODIMENT

Immersion, presence and embodiment aren't new terms, and their definitions have been extensively discussed and de-

bated over the last 30 years (Shin, 2018; Cavazza et al., 2017; Slater, 2009)

Immersion is an objective condition determined in particular by the technological apparatus that is being used for the VR experience. It therefore depends on the ability of the VR system to trick our senses through the quality of visual, auditory or haptic stimulation. The stimuli provided in this way must deceive our cognitive system by “technically” transporting us to another place. With a high level of immersion, the sensory information provided in a VR system is almost identical to that we have in real life.

Presence instead refers to a subjective condition conformed on the basis of how much the user feels involved around the virtual environment. If the narrative structure involves it, if the modes of interaction are natural then the presence will be high. Presence does not depend directly on the level of Immersion and is fully pursued when technology completely disappears, and the user reacts to virtual stimuli as if he/she were in the non-virtual world. Finally, embodiment refers to the perception of physically interacting with a virtual environment. In traditional video games, interactive animation and three-dimensional navigation immerse a player in the 3D world represented. In this case, however, presence and immersion are not determined by stimuli to the visual and perceptual system but by cognitive and emotional elaboration processes that keep the user involved throughout the duration of the game experience.

When immersion involves a purely representative simulation of a 3D space, the perceptive experience obtained through movements in such spaces becomes an integral and essential part of gaining knowledge about the three-dimensional attributes of the model. In fact, there is no doubt that the perception of a single artifact or a built environment is resolved by visual perception, and therefore, the visual apparatus is involved only up to an extension of scale that allows the user to understand the environment at a glance (Antinucci, 2014).

When the dimensions are set to force the user to move his/her body, even if limited to the rotation of the head alone, then

the perceptive modes related to the sense of sight are integrated with those of the human body and its neuro-brain systems.

To obtain embodiment in a VR experience, a sensory-motor integration of sight and movement is necessary. Without this, the cognitive experience would be compromised. It follows the need to combine motion perception and wayfinding with the sense of sight because the ability to move and navigate are integral parts of direct space perception.

Therefore, presence, immersion and embodiment are opportunities that designers of virtual environments must understand in order to use a narrative model specific for the new medium.

#### WHERE AM I?

Anyone who has ever experienced an immersive virtual environment would have spent the first seconds frantically looking around, trying to figure out where they have been transported. In a VR experience, the fact that the question of “where” precedes the questions of “how” and “why” is one of the main elements that differentiates virtual reality and cinematographic fiction.

In virtual reality, location—whether hyper-realistic or imaginative—is of primary importance. From a narrative point of view, the VR space we experience is able to convey information the way a virtual character or an off-screen voice would. In 2013, interactive illustrator Daniel Ernst defined diorama as a type of VR experience based exclusively on visual discovery. He said, “I wanted to create a type of experience that would convey the sense of wonder the diorama in the Efteling did so perfectly well, and in the end, the design choices supported the name as well,” and “Each diorama is a fantastic hand-painted environment in which interaction is used to tell a story and convey a sense of wonder.” (Ernst, 2014). In one of these dioramas, Ernst draws an environment that reproduces a study by the Russian scientist, Alexey Pajitnov,

**Fig. 1** Screenshot from shoebox diorama *Blocked In* by Daniel Ernst.

creator of the puzzle game Tetris. The richness of the details along with the presence of the typical Tetris blocks that perceptibly fall into the external environment are the fundamental elements of a screenplay in which the story and the way of telling it converge in the setting (Figure 1).



About 200 years earlier, the pictorial panorama was conceived for image construction alongside which, a way of visually presenting such an image was also developed (Bordini, 1984). In fact, well before the development of viewers for virtual reality, the panorama and diorama were the first scenic artifices capable of representing the viewer being inside a simulated environment. Technically, these were large canvases painted and set up on a circular wooden wall so as to completely surround the viewer. The viewer would be at the middle of a cylinder, its inner face primarily depicting landscapes or scenes of war. The resulting effect was a com-

**Fig. 2** The Go-Pro camera rig used to shoot the first person perspective of the movie "Hardcore Henry" directed by Ilya Naishuller (2015). Credit: STX Entertainment.



plete enveloping of vision. Observing a panorama meant immersing oneself inside the painting and being isolated from the outside world because it was no longer possible to find the limits or frame of that representation. Similarly, the first nineteenth-century dioramas were able to introduce sophis-

ticated movements and plays of light to add dynamism and changes within the scene.

In Tricart (2017), David Liu, the former director of virtual reality at Viacom NEXT states that virtual reality is a spatial medium. He refers to virtual content creators as narrative architects whose role is not to geometrically shape the environment but to craft a world of narrative potential, enabling the environment to tell a story that would be interpreted by the user.

If we want to follow the cinematographic lexicon's indications, the use of VR for narrative purposes implies a definite contraction of the traditional compositional vocabulary.

Unlike the normal frame, the spherical space has no off-screen images. All the aspects become visible at the same time, and the viewer has to personally choose which portion of the field to frame each time with his/her gaze.

To understand virtual spaces, it may be more productive to introduce a new definition of off-screen, referring not to the shooting field, but to the viewer's field of view (FOV). It is clear that the centripetal action of the virtual image is not characterized by the traditional out-of-field as much as the virtual out-of-field, which must be identified in real time when exploring VR. An out-of-field image that is always potentially visible to users threatens the foundation of classical cinematographic scripts.

The virtual reality must therefore be exclusively expressed through full shots. Although the spherical space of VR is different from the role of traditional cinema, they share descriptive characters and the function of representing an entire internal or external space.

The technological peculiarities of an immersive virtual environment makes us reflect on some specific features of the narrative model to be adopted.

For example, determining factors such as hyperrealism derived from the sensitive enrichment of the viewer's sensory experience or the ultra-spectacularity in overcoming the limits of the traditional screen have become the main



constituent elements of new experiential VR modes (Rossi, 2017). These factors were fueled by certain contingencies such as the commercial interest in wearable HMDs, the miniaturization of shooting technologies and action cams, the worldwide success of some FPS video games and the release of movies visually based on first-person shots and single long takes (Figure 2).

Another feature of narration in VR is the typically paratactic trend. Each environment is directly dependent on the FOV made available by HMDs and each slice of the immersive sphere is independent from each other.

Unable to handle a traditional narrative formula, VR uses cinematic language to show a scene rather than telling the user about it. Thus, the storytelling in VR comes closer to a stunning graphic edit than a traditionally articulated sequence of events and scenes.

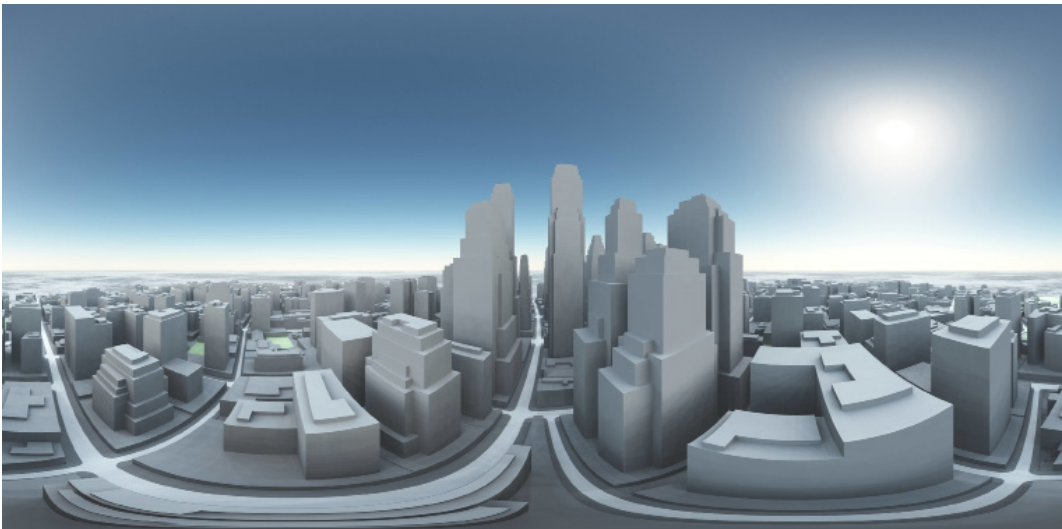
#### WHO AM I?

After wondering, “Where am I?” the user usually questions his/her role within the virtual space and whether he/she is an active part of the story or a silent viewer.

Therefore, after a quick initial exploration, the users tend to look for something to focus on. These are the points of interest (POI). Points of interest are elements that attract a visitor’s attention within a VR experience. Some are extremely easy to find, such as colored geometric primitives on a completely black graphic field, while others are a lot more difficult to find due to the absence of preponderant graphic elements with noticeable shapes, sizes or chromatic contrasts (Figure 3).

During a VR experience, the visitor’s attention shifts from one POI to another in a chain of eye or head movements depending on the virtual content designer’s placement of POIs within the scene. This guides the viewer’s attention and consequently silently directs the visual storytelling. When the POIs are not graphically evident, clues are introduced to help trace them. Such clues can be small animations or color changes as well as audio effects or music tracks.

The designer's skill lies in the ability to hide such clues. If they are too evident, they would break the user's illusion of being alone in an unknown environment. On one hand, disseminating cryptic clues would make it more difficult to construct a sense of visual storytelling; on the other hand, expending greater cognitive effort would enhance the user's virtual immersion (Brillhart, 2016).



**Fig. 3** In this equirectangular 3D illustration the group of skyscrapers in the center represents the only POI. Credit Stanislav Rykunov @ Dreamstime.com.

When more than one POI is present in the virtual environment, being prompted to make an effort could lead to the user being overly engaged in interpreting clues due to the fear of missing out (FOMO) on something important in the story rather than enjoying the virtual experience (Figure 4).

FOMO can be a dangerous tool that's capable of expelling users from the immersive experience because they would find themselves "leaving" the narrative environment to observe their actions. It can also be an excellent narrative stimulation device to create anxiety and stress in the user or to force the user to repeat an experience several times to complete every part of it (Tricart 2017).

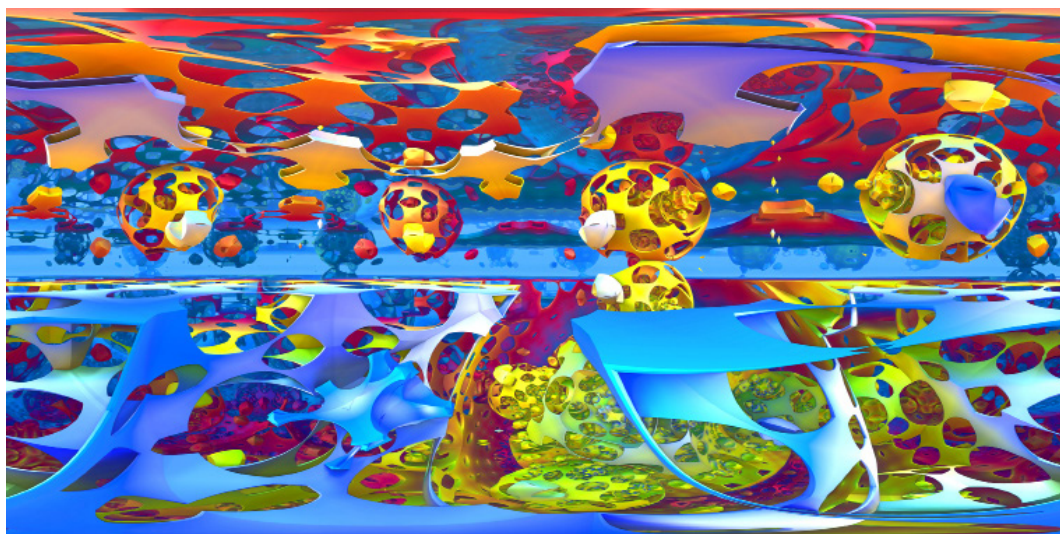
Therefore, FOMO is the basis of a non-linear narrative mode where all the POIs on the scene are consumed not in a precise order but according to the sensitivity and experience of each user.

## AESTHETIC OF CONTINUITY

The remediation process triggered by the digital revolution is characterized by attributes such as fluidity and dynamism. From the logic of postmodern pastiche that dominated artistic and cultural production in the 1980s, we moved on to the multimedia paradigm in which codes from a multitude of media objects overlap in a continuous flow of information.

Another founding characteristic of VR storytelling was derived from this remediation: continuity. It was developed when the artistic production in the early nineties' digital revolution began to transform the world of communication and define new aesthetic and cultural canons. The aesthetic category of continuity is dominant in the three-dimensional drawing practices of interactive virtual spaces. In fact, the exploration of these spaces is articulated in an uninterrupted manner and without editing cuts. Lev Manovich (2002) identifies this lack of editing as the primary linguistic convention that best suits the first-person narrative modalities of the immersive virtual space experience.

**Fig. 4** In this equirectangular 3D illustration the many shapes and bright colors can generate a fear of missing out (FOMO) in the viewer due to the richness of detail and potential POIs. Credit: Pitris @ Dreamstime.com



According to him, the digital paradigm shifts from the tradition of representation, qualified by the presence of a frame/screen, to the tradition of simulation, characterized by the absence of a reduction scale; this causes the physical space to coincide with the virtual one.

The lack of editing becomes a precise compositional strategy wherein first-person navigation prevails, and it also suggests the use of a shooting technique borrowed from film grammar: the long shot, which is a shot that performs the functions of a sequence or scene on its own. In other words, a single shot completes an entire narrative sequence without the need for editing.

Continuity can therefore be considered as another intrinsic attribute of the virtual image, characterized by the high degree of subjectivity of the shots.

The size and shape of the HMD screen strategically activates the viewer's peripheral vision and strongly intensifies the illusion of reality. The VR visitor becomes actively immersed in the simulation of a "real" space rather than passively enjoying an audiovisual document.

Ruggero Eugeni (2015) defines the first-person shot as a symbolic form and more precisely, "a ubiquitous and almost omnipresent figure within the intermediate and post cinematic galaxy that characterizes contemporaneity". The POV shot, which represents a well-defined shooting technique in film grammar, is combined with domains and experiences from the world of FPS video games and has become a symbol of a visual culture characterized by perceptive habits dominated by first-person experiences.

Therefore, the first-person shot contains enormous potential for different ways of unveiling 3D environments. Exploration with the personal point of view improves the geometric perception of three-dimensional models. The stereoscopic visualization of the latest HMD generation, combined with the ability to move freely with this point of view, improves the user's ability to understand the spatial quality of the places in which he/she is immersed.

Any environment can therefore be experienced as if

the user is actually inside it. The user can virtually move around in the 3D scene by approaching, moving away or turning their head. In other words, one can benefit from a system of simulation of possible movements in space through a natural intuitive interface such as the human body.

## REFERENCES

- Antinucci, F. (2014). *Comunicare nel museo*. Roma, IT: Laterza.
- Bordini, S. (1984). *Storia del Panorama. La visione totale nella pittura del XIX secolo*. Roma, IT: Officina.
- Brillhart, J. (2016). *In the Blink of a Mind*. Retrieved from <https://medium.com/the-language-of-vr/>
- Cavazza, M., & Young, R. M. (2017). Introduction to interactive storytelling. In R. Nakatsu, M. Rauterberg, P. Ciancarini (Eds.). *Handbook of Digital Games and Entertainment Technologies*, 377-392. New York, NY: Springer.
- Ernst, D. (2014). *Why nothing is not the answer to: "what do I do?"*. Retrieved from [theshoeboxdiorama.blogspot.com/](http://theshoeboxdiorama.blogspot.com/)
- Eugeni, R. (2015). *La condizione postmediale*. Brescia, IT: La Scuola
- Rossi, D. (2017). Shaping Virtual Image. In A. Luigini, D. Basso, S. Brusaporci, E. Cicalò, M. M. Moretti... D. Villa, *Proceedings of International and Interdisciplinary Conference IMMAGINI? Image and Imagination between Representation, Communication, Education and Psychology*. Basilea, CH: MDPI.
- Shin, D. (2018). Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? *Computers in Human Behavior*, 78, 64-73.
- Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535), 3549-3557.
- Manovich, L. (2001). *Il Linguaggio dei Nuovi Media*. Milan, IT: Olivares.
- Tricart, C. (2017). *Virtual reality filmmaking: Techniques & best practices for VR filmmakers*. Abingdon, UK: Taylor & Francis.

**Article available at**

DOI: 10.6092/issn.2724-2463/11077

**How to cite**

**as article**

Rossi, D. (2019). Inside Image: Technical Notes for Virtual Storytelling. *img journal*, 1(1), 298-311.

**as contribution in book**

Rossi, D. (2019). Inside Image: Technical Notes for Virtual Storytelling. In A. Luigini, C. Panciroli (Eds.) *img journal 01/2019 Manifesto*, 298-311. ISBN 9788899586096



© 2019 by the authors. Licensee PUBLICA, Alghero, Italy. Open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC-BY-NC-ND 4.0) license.