The interplay between affective and cognitive factors in shaping early proficiency in mathematics

Published Version:
Cargnelutti, E., Tomasetto, C., Passolunghi, M.C. (2017). The interplay between affective and cognitive factors in shaping early proficiency in mathematics. TRENDS IN NEUROSCIENCE AND EDUCATION, 8-9, 28-36 [10.1016/j.tine.2017.10.002].

Availability:
This version is available at: https://hdl.handle.net/11585/614734 since: 2022-02-23

Published:
DOI: http://doi.org/10.1016/j.tine.2017.10.002

Investigating Depth Perception in Immersive Hypothetical Reconstructions: 1816 Canova's Exhibition in Spirito Santo Church in Bologna

Fabrizio Ivan Apollonio[1][0000-0001-5186-1378], Federico Fallavollita[1][0000-0003-2916-0262], and Riccardo Foschi[1*] [0000-0001-6828-8133]

[1] Department of Architecture, Alma Mater Studiorum Università di Bologna, Italy,
`fabrizio.apollonio@unibo.it`, `federico.fallavollita@unibo.it`,
[1∗corresponding] `riccardo.foschi2@unibo.it`

Keywords: VR, spatial visualization, real-time interactive 3D, 360° omnidirectional stereoscopic spherical panorama, perspective projection.

1 Introduction

Digital hypothetical 3D reconstructions are often presented through precomputed visualizations, nevertheless, new Extended Reality (XR) technologies allow the users to experience the same architectonic spaces in a much more effective and immersive way. Different technologies provide different levels of immersion and quality, and if not properly used sometimes they could distort the perception of such spaces rather than improve it. Virtual spaces captured as spherical panoramas and visualized through a headset, sometimes are perceived as "too big" or "too far" but feel "of the right size" when viewed through a fully explorable Virtual Reality (VR) real-time rendered interactive experience. This observation leads to interesting hypotheses related to how different immersive visualization technologies influence our perception of depth and the interpretation of the size and shape of virtual architectonic spaces. The case study of the hypothetical virtual 3D reconstruction of Canova's exhibition, held in 1816 in the Spirito Santo Church in Bologna (presented at the *Notte Europea dei Ricercatori* – Society held in Cesena), is a valuable opportunity to investigate the criticalities and potentialities of advanced immersive visualization technologies in the architectural field, with a particular focus on the perception of depth and size of objects and architectonic spaces. Thus, this research aims to investigate the state of the art of the most popular available immersive and non-immersive visualization modes and technologies (e.g., mono and stereo still images and animated videos, spherical 360° panoramas, VR interactive experiences) and investigate how much, and under which circumstances, such new technologies can improve or undermine the perception of architectonic spaces.

In traditional architectural representations, many aspects influence the perception of objects and spaces [1]. For example, it is known that in architectural perspective views, a wide field of view (short focal length) empathizes the apparent perspective aberration and gives the impression of expanding space along the camera axis [2] the apparent distortion is particularly evident closer to the edges of the frame. A higher point of view contributes to shrinking the surroundings of the observer while a lower point of view expands it. The orientation of the camera (the angle of its projection plane) might also produce apparent distortions to shapes and proportions of the objects in relation to how fast lines that are parallel in 3D space converge to their relative vanishing points in the 2D projection.

Mastering these parameters is something that architects and photographers are trained to do, however, the new media, frontier of architectural visualization, such as 360° spherical panoramic images and immersive Virtual Reality (VR) explorable scenes, present new challenges. Some of these new media differ from monocular still images by the fact that they add an improved perception of depth, which is achieved by differentiating what the right and left eyes see (mimicking the stereoscopic vision), or by compensating for the lack of binocular disparity by conveying the depth information through interactive motion. However, this newly added dimension brings into play further optical and psychological factors that might distort the correct perception of these virtual spaces [3, 4].

In this research, through the case study of the hypothetical reconstruction of Canova's exhibition in Spirito Santo Church in Bologna in 1816 [5], we will compare several static, animated, and interactive visualization techniques, such as omnidirectional or unidirectional 360° spherical panoramas (monocular or stereoscopic), real-time rendered explorable interactive virtual scenes viewed through a screen or a VR headset, and we will analyze and discuss the criticalities and potentialities of each method in comparison with the more traditional visualization techniques, to help the scholars and professionals performing informed choices according to their needs.

2 Visualization modes and technologies

Architectural representation has a centuries-old history. In the past, the buildings were represented only through static drawings and physical maquettes. The digital revolution allowed the representation techniques to evolve and new technologies and methodologies were discovered up to our age where we are not only bound to static and physical representations anymore.

Architectural visualization nowadays can be passive (e.g., images, precomputed animations) or interactive (e.g., active exploration of a virtual scene with a headset or through a traditional display with mouse and keyboard). Both passive and interactive visualizations can be experienced through traditional bidimensional displays or VR headsets (e.g., as planar projections, spherical/cubical projections, or completely explorable interactive 3D scenes). Depending on the technology used, either the passive and the interactive architectural visualizations can be presented in monocular mode (there is no distinction between what the right and left eyes see) or stereoscopic mode (the right and left eyes perceive two different images).

Several visualization modes can derive from the combination of all these variables:

- Monocular still images (e.g., traditional raster images);

- Stereoscopic still images (e.g., anaglyphic images, autostereograms, polarized images);
- Monocular 360° spherical panoramas (e.g., cubic projections, equirectangular projections)
- Stereoscopic 360° omnidirectional spherical panoramas (e.g., over-under or side-by-side equirectangular stereo projection, stereo cubic projection);
- Monocular pre-computed/pre-recorded animations (e.g., traditional videos and movies)
- Stereoscopic pre-computed/pre-recorded animations (e.g., the so-called 3D movies)
- Monocular pre-computed/pre-recorded spherical panoramic animations (e.g., the so-called 360° panoramic videos)
- Stereoscopic pre-computed/pre-recorded omnidirectional spherical panoramic animations (e.g., the so-called 360° panoramic 3D videos)
- Parallax pre-computed animations (e.g., the so-called 2.5D animations)
- Interactive fully explorable/immersive stereoscopic VR experiences (e.g., game-like experience)
- ...

Not only the visualization mode but also the technological device used for the visualization plays an important role in the correct interpretation of depth and shape in architectural visualization. Nowadays, three-dimensional data can be visualized in various forms and with various technological devices:

- Physical 2D prints/drawings (e.g., hand-drawn, printed on a sheet of paper)
- Physical 3D models (e.g., 3D printed, handcrafted, CNC carved)
- Digital 2D monocular displays (e.g., computer display, smartphone screen)
- Digital 3D stereoscopic displays (e.g., stereoscopic and autostereoscopic displays, with or without active/passive glasses)
- Digital 3D spatial displays (e.g. Asus spatial vision, Sony spatial reality display)
- 3-DoF (Degrees of Freedom) electronic VR headsets (e.g., smartphone + Google Cardboard, only capable of tracking rotations)
- 6-DoF electronic VR/AR headsets (e.g., Meta Quest 3, Microsoft Hololens, Apple Vision Pro, capable of tracking rotations and translations)
- …

Different visualization modes and technological devices can convey different depth cues, superimpose different apparent perspective aberrations, or generate optical illusions, due to the type of projection used, the position of the observer, and other boundary conditions. Given these assumptions, knowing how to handle such complexity is of fundamental importance for whoever operates in the field of architectural visualization, where the correct communication and interpretation of shape and size are crucial aspects. In the next section, the most popular visualization modes experienced through various technological visualization devices will be investigated and compared.

3 Faithful Representation of Architectonic Spaces

The problem of representing architectonic spaces faithfully is something that architects, artists, and photographers investigated for centuries. It is known that a point of view close to the ground gives the impression of an expanded space, and vice versa. In photographic terms, a wide Angle of Vision (AoV), or a short Focal Length (FL), gives the impression of a space expanded along the camera axis (Fig. 1) [2].

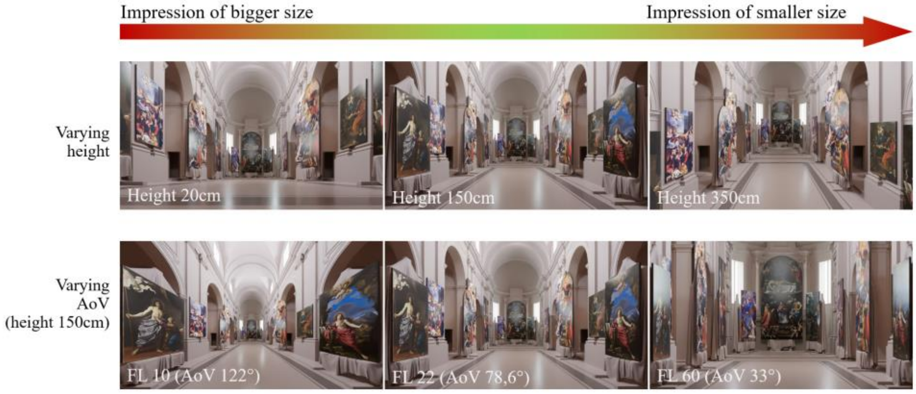

Fig. 1. Varying the height and the Angle of vision (AoV) change the impression of the size of architectonic spaces in monocular still images; (Author: Riccardo Foschi).

Other camera parameters can influence the perception of depth and size such as the angle of the projection plane, and the amount of blurriness of the foreground/background. In Fig. 2 extreme settings of the camera (height, angle, aperture, and focal length) were used to produce relevant upscaling and miniaturization effects. Architects and photographers use human-height viewpoints to represent spaces faithfully and higher/lower viewpoints to convey particular effects. Frank Lloyd Wright for example, in some of his indoor drawn perspectives [6–9], used to choose a point of view at the seated height (about 120/130 cm) giving the impression of a bigger space compared to when observed from a standing position (150/170cm).

Fig. 2. The same architectonic space captured with extreme camera settings to produce an upscaling (left) or miniaturization (right) effects; (Author: Riccardo Foschi).

3.1 Monocular and Stereoscopic Depth Cues

The interpretation of depth in the examples shown in Fig. 1 and 2 was conveyed exclusively thanks to monocular depth cues [10 pp. 155–158] (i.e., which can be perceived only with one eye/image). However, in the real world, the perception of depth is also influenced by stereoscopic depth cues [10 pp. 158–161] (i.e., that can be perceived only with two eyes). Monocular depth cues are many more than stereo depth cues, and after a certain distance stereo cues start to be less and less relevant, this is why the size of an architectonic space can roughly be interpreted by just looking at its 2D projection. Despite that, VR experiences do not only add stereo cues but also motion parallax which greatly contributes to improving the correct perception of depth. However, the added third dimension, and the possibility to dynamically move the point of view, bring into play further optical and psychological factors that, if not properly calibrated, might undermine the correct perception of the architectonic spaces, rather than improve it. The main monocular and stereoscopic depth cues are listed in Table 1.

Table 1. Main monocular and stereoscopic depth cues [10 pp. 155–158].

Monocular depth cues	Stereoscopic depth cues
Retinal image size	Ocular vergence
Height in the visual field	Binocular disparity
Accommodation	
Motion Parallax	
Shadows	
Interposition	
Image blurriness/sharpness	
Atmospheric haze	
…	

Monocular depth cues. The retinal image size is responsible for giving the impression that bigger figures are closer and smaller ones are farther. The height in the visual field causes the elements of the image that are closer to the horizon line to be interpreted as farther. Not only the depth but also the size of the objects can be influenced by their height in the visual field, in fact, the famous moon illusion [11] which causes the moon to appear larger when it is near the horizon could be related to this effect. Accommodation is the process by which our eyes are capable of changing their optical power to focus on close or far objects, our brain is capable of evaluating the change in the eye curvature and uses this information to contribute to the interpretation of depth. Motion parallax cue kicks in when the viewer is in motion relative to the observed objects, thanks to motion parallax the brain is capable of interpreting depth by comparing the speed of the retinal images of far and near objects (farther objects move slower, closer objects move faster). The interpretation of the shadows projected by an object onto another can help evaluate which of them is closer or farther. The interposition of figures helps recognize which of them is in front or behind.

The amount of blurriness, the amount of atmospheric haze, and other monocular cues are also relevant factors that can help interpret the depth of a scene.

Stereoscopic depth cues. Ocular vergence refers to the simultaneous movement of both eyes inward (convergence) or outward (divergence) in order to point to a target object preventing double vision. The brain is capable of evaluating depth by interpreting the change in the tension of the muscles of the eyes responsible for their convergence on a near or far object. Lastly, binocular disparity refers to the ability of the brain to evaluate depth by interpreting the differences between the right and left retinal images of the same subject. For more about depth cues and human vision mechanisms, refer to S. M. La Valle's book: Virtual Reality [10].

3.2 Analysis of monocular and stereo depth cues in different visualization modes

Each of the visualization modes listed in Section 2 relies on different monocular or stereoscopic depth cues. The correct interpretation of the depth and shape of the architectonic spaces strictly relates to the different depth cues that each mode is capable of conveying. A synthesis of the various depth cues conveyed by the different visualization modes is presented in Table 2.

Table 2. Monocular and stereoscopic depth cues analysis in different visualization modes.

Visualization modes	**Mono cues**					**Stereo cues**	
	Retinal image size	Height in the visual field	Accommodation	Motion parallax	Shadows Interposition Blur Haze (…)	Ocular vergence	Binocular disparity
Monocular still images	✓	✓			✓		
Stereoscopic still images	✓	✓			✓	✓	✓
Monocular 360° spherical panoramas (3-DoF)	✓	✓			✓		
Stereoscopic 360° omnidirectional spherical panoramas (3-DoF)	✓	✓			✓	✓	✓
Monocular pre-computed animations	✓	✓		✓ (non-interactive)	✓		
Stereoscopic pre-computed animations	✓	✓		✓ (non-interactive)	✓	✓	✓
Monocular pre-computed spherical panoramic animations	✓	✓		✓ (non-interactive)	✓		
Stereoscopic pre-computed omnidirectional spherical panoramic animations	✓	✓		✓ (non-interactive)	✓	✓	✓

Parallax pre-computed animations (2.5D)	✓	✓		✓ (faked)	✓		
Interactive stereoscopic immersive VR experiences (6-DoF)	✓	✓		✓ (interactive)	✓	✓	✓

Monocular still images (planar perspectives). They can only convey some monocular depth cues (except accommodation and motion parallax). Refer to section 3 for more about depth perception in still images.

Stereoscopic still images. They have the same monocular depth cues as monocular still images plus the binocular disparity and ocular vergence. Thanks to the added stereo cues, those who have stereoscopic vision can perceive a surprising three-dimensionality effect from the image. However, to have correct depth perception, the size and position of the image should be calibrated in relation to the viewpoint of the observer; and the Inter-Pupillary Distance (IPD) [12] should be calibrated based on the specific observer, which is rarely the case since these images are usually produced by referring to an average IPD (i.e., 6.5cm) and are viewed at different distances depending on how the viewer approaches the image.

Monocular 360° spherical panoramas. Unintuitively, can convey the same depth cues of monocular still images, even when experienced through a headset. This happens because monocular spherical panoramas, even when viewed from a headset, provide the same image to both eyes. Furthermore, the spherical image was generated by projecting the scene onto a sphere with a predetermined fixed center which causes the headset to react only to rotations and not translations of the head, and rotation alone cannot provide any parallax effect (as shown in Fig. 3). It must be noted that even 6-DoF headsets behave like 3-DoF headsets when visualizing 360° spherical panoramas.

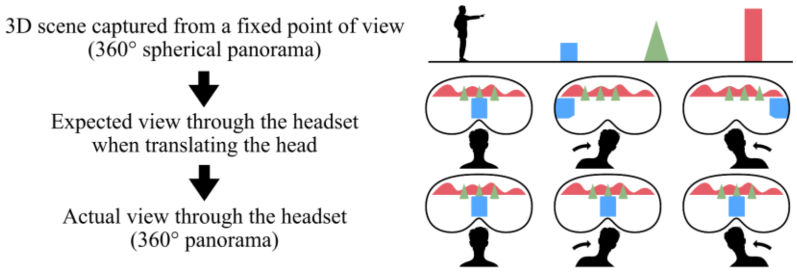

Fig. 3. Expectations versus the actual perceived scene captured as a 360° spherical panorama viewed through a headset (3-DoF); (Author: Riccardo Foschi).

Despite no additional depth cues being added compared to still images, the level of immersivity provided by 360° spherical panoramas experienced through headsets is surely improved because the user feels inside the environment. However, immersivity does not automatically improve the perception of depth, in fact in this case it is the opposite. A 360° spherical panorama viewed through a headset can produce various unwanted illusions, e.g. the illusion of flying, bigger spaces, sliding floors and walls. The reason for that deals with the missing of any motion parallax effect.

In our all-day experience, when we move (translate) our head relative to some objects, the retinal images of the objects closer to us move faster than the retinal images of farther objects. In 360° spherical panoramas, we never experience any perceivable shift between objects in the foreground and background. In reality, this is only possible for objects very far away from us. Our experience also teaches us that architectonic spaces are of relatively small dimensions. Thus when we observe architectonic spaces in the form of 360° spherical panoramas through headsets the interpretation that our brain tries to give (objects without any parallax shifting = very far objects) conflicts with our expectations of that space (the walls aren't farther than a few meters), and this creates confusion in the interpretation of depth, size and shape. Fig. 4 and 5 graphically synthesize the most common illusions that 360° spherical panoramas viewed through a headset can produce: bigger spaces, sliding floors and walls, and the illusion of flying.

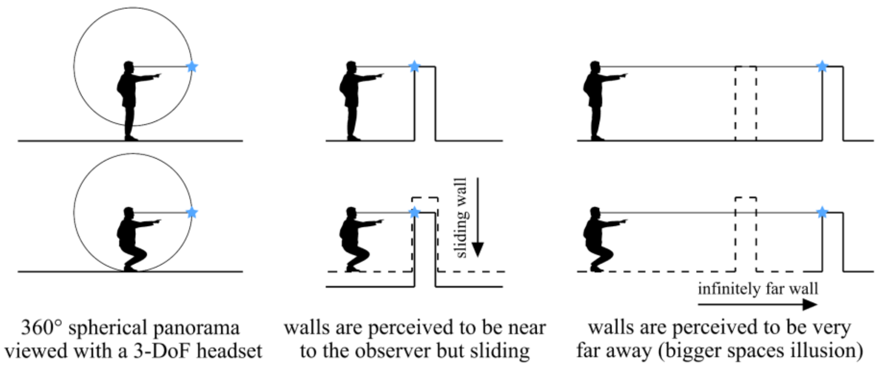

Fig. 4. Graphical explanation of sliding walls / bigger spaces illusion; (Author: Riccardo Foschi).

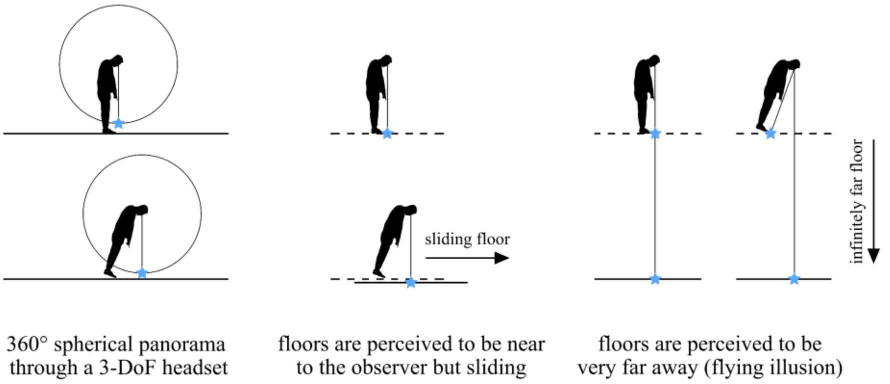

Fig. 5. Graphical explanation of sliding floors / flying illusion; (Author: Riccardo Foschi).

The fact that the predetermined projection center of a 360° spherical panorama is always aligned with the observer's viewpoint (when viewed through a headset) does not only have drawbacks. In fact, viewing any type of projective image from the correct viewpoint prevents anamorphic distortions and gives an improved perception of depth. This point can be better understood by referring to the traditional problem of the *Veduta Vincolata* [1]. The *Trompe l’Oeil* technique, for example, provides the best illusion if the observer is in the precise spot where the artist placed the center of projection, the effect is lost if the observer moves. So, when viewing a 360° spherical panorama through a headset to enjoy the benefits while limiting as much as possible unwanted illusions, a good approach would be always closing the eyes before rotating the head and looking at the panorama only without moving or rotating the head.

Stereoscopic 360° omnidirectional spherical panoramas. When adding the binocular disparity and ocular vergence cues to 360° spherical panoramas the perception of depth is improved, however, the missing motion parallax cue with all its consequences still persists. It must be noted that the creation of plausible stereoscopic 360° spherical panoramas is not as simple as producing two spherical projections, one for the right eye and one for the left eye, with their centers at interocular distance. In fact, in this way, the binocular disparity would be correct only when looking toward a specific direction (Fig. 6).

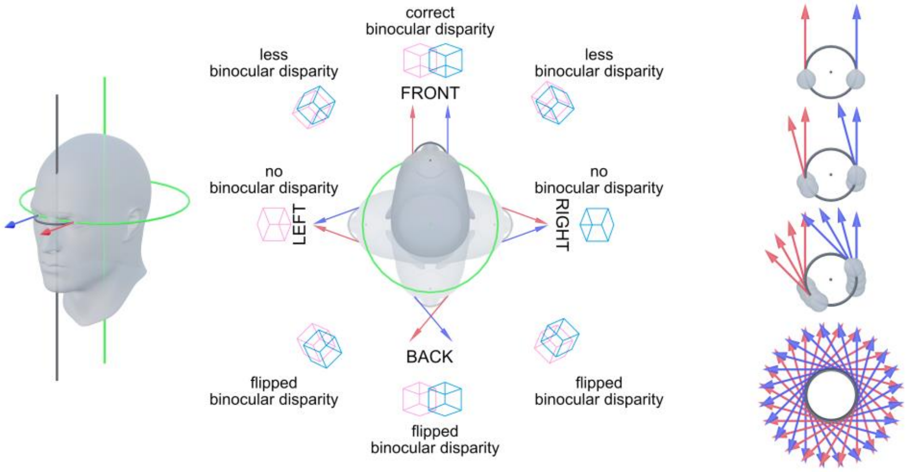

Fig. 6. Left: view directions centered in the eyes (blue and red arrows), axis passing through the neck (green), axis passing through the eyes midpoint (grey). Centre: the binocular disparity problem in unidirectional 360° stereoscopic spherical panoramas. Right: improved omnidirectional distribution of view directions; (Author: Riccardo Foschi).

So, to produce an acceptable approximation of the binocular disparity from all directions, it is important to use a special arrangement of the right and left eye projection rays known as omnidirectional stereo projection [13–15]. Most 3D rendering applications already provide special virtual cameras capable of producing such type of stereoscopic omnidirectional projection. In case the application of use only provides monocular spherical cameras, a good approximation of an omnidirectional 360° stereoscopic spherical panorama can be created by the method explained by Bourke in 2010 [16].

Pre-computed animations. Both monocular and stereoscopic pre-computed animations, spherical or not, share the same depth cues as their still counterparts, with an added non-interactive motion parallax cue. The motion parallax cue is present thanks to the predetermined motion of the camera, however, it is not interactive because it does not respond to the observer's movements.

Parallax (2.5D) pre-computed animations. In these particular types of animations the motion parallax is predetermined as in the previous case, but it’s faked manually by starting from a monocular still image and animating it by sliding the cutouts of the objects in the foreground and background at different speeds based on the intended distance from the observer. This type of 2.5D animation is mainly for artistic purposes and is not suitable for conveying accurate depth in architectural visualization.

Interactive stereoscopic immersive VR experiences. They are usually navigated through electronic headsets and provide all the depth cues present in the previous visualization modes including binocular disparity and ocular vergence, but are the first of the visualization modes analyzed that convey an accurate and interactive motion parallax. The binocular disparity and ocular vergence depth cues are possible because two different images are sent to the right-eye and left-eye displays, and the motion parallax cue is possible because the user’s motion is tracked via sensors placed in the environment around the user or on the headset itself. VR experiences, despite being very advanced ways to experience virtual architectonic spaces, are still not capable of conveying accommodation. None of the technologies analyzed so far is capable of reproducing accurate accommodation, at present only prototype technology can provide accurate accommodation (e.g., light-field displays). The lack of accommodation does not cause major problems in the interpretation of depth and shape of architectonic spaces, however, the accommodation-vergence conflict [17] can cause fatigue or motion sickness for some subjects.

VR experiences are capable of conveying the most amount of monocular and stereoscopic depth cues compared to the other analyzed visualization modes, however, other factors need to be taken into account to guarantee the best possible experience. For example:

- the angle of vision of the virtual camera must match the headset angle of vision for the correct retinal image size of the objects;
- the tracking of the headset must be correctly calibrated both for angles and translation to avoid motion sickness;
- the Inter-Pupillary Distance (IPD) must be as close as possible to the one of the user to avoid the illusion of bigger or smaller spaces (virtual eyes too close make the scene look bigger and vice versa);

- the height of the virtual point of view with respect to the ground floor must match exactly the user's height;
- other factors such as display refresh rate, lens aberrations, color depth, brightness, pixel density, etc. also play a role in the correct perception and interpretation of a three-dimensional virtual scene, however, these aspects are not controllable by who designs the 3D scene or the app, but depends on the hardware.

4 Case Study: 1816 Canova's Exhibition in Spirito Santo Church in Bologna

Given the studies presented in the previous sections, VR visualization mode is indeed the solution, among the ones analyzed, which provides the best interpretation of architectonic virtual spaces. Thus, for the case study of the hypothetical virtual 3D reconstruction of Canova's 1816 exhibition held in Spirito Santo church in Bologna (Fig. 7), we opted for an immersive experience viewed through a VR headset. The immersive VR experience was developed for the *Notte Europea dei Ricercatori – Society*, held in Cesena in September 2023. The 3D model was based on a previous work developed for the exhibition *Antonio Canova e Bologna, alle Origini della Pinacoteca* [18] and was developed by a multidisciplinary team of architects and historians. In this section, we will present the workflow that we followed to prepare a 3D model suitable for VR applications in the context of scientific dissemination.

Fig. 7. Hypothetical 3D reconstruction of Spirito Santo church exhibition (frame from the animation presented at *Antonio Canova e Bologna, alle Origini della Pinacoteca* [18]); (Author: Riccardo Foschi).

4.1 Historical Context and Gathering of the Sources

During the Napoleonic looting of art (1797-1815 after the Treaty of Tolentino), several artistic masterpieces were seized by the French army from the conquered territories, including Italy, and brought to France. In 1815 the Congress of Vienna ordered the restitution of the works and Papa Pio VII sent Antonio Canova as an official envoy to retrieve these artistic treasures. Canova was able to retrieve several masterpieces and his effort culminated in a significant public exhibition of paintings housed in Santo Spirito church in Bologna. Among the witnesses of that event was Marquess Antonio Bolognini Amorini. His precious testimony about the exhibition was published presumably in 1818 [19]. Thanks to this direct textual source, we know now, with a certain degree of reliability, which paintings were exhibited and their disposition in the church (Fig. 8). All the original Paintings are now housed in the *Pinacoteca Nazionale di Bologna* and the *Pinacoteca Civica di Cento*. Other historical sources [20] highlighted that the church was renovated in 1788 by the architect Giuseppe Jarmorini (1732–1816), and after that, the building didn't receive any major change, until 1943 when its roof was demolished due to the bombardments of World War II, and then restored in the '80s as it was before [21].

Fig. 8. Layout of the paintings in the Spirito Santo Church retrieved from Antonio Bolognini Amorini's text [19]; (Author: Riccardo Foschi).

4.2 Constructing the 3D Model

Spirito Santo church still exists today in a good state of preservation and the gathered sources provide information about its configuration and ornaments which are probably very close to how they were during Canova's exhibition in 1816. Thus, the laser scanning acquisition was the starting point for the 3D modeling phase. The present state was then compared to the available historical evidence to digitally reconstruct the historical configuration as a NURBS mathematical 3D model (Fig. 9).

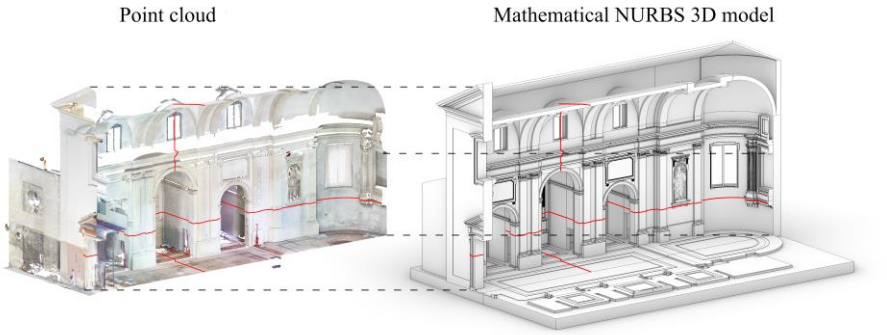

Fig. 9. Reconstruction process of the 3D model of Spirito Santo church. Left: laser scanned point cloud; right: NURBS 3D model; (Author: Riccardo Foschi).

The model was semantically segmented (Fig. 10) to simplify the 3D modeling process, improve the future analysis possibilities, foster an eventual future upload on a shared public scientific repository, and speed up the NURBS-to-mesh conversion process which requires the meshes to be tessellated differently based on their level of detail and geometric complexity.

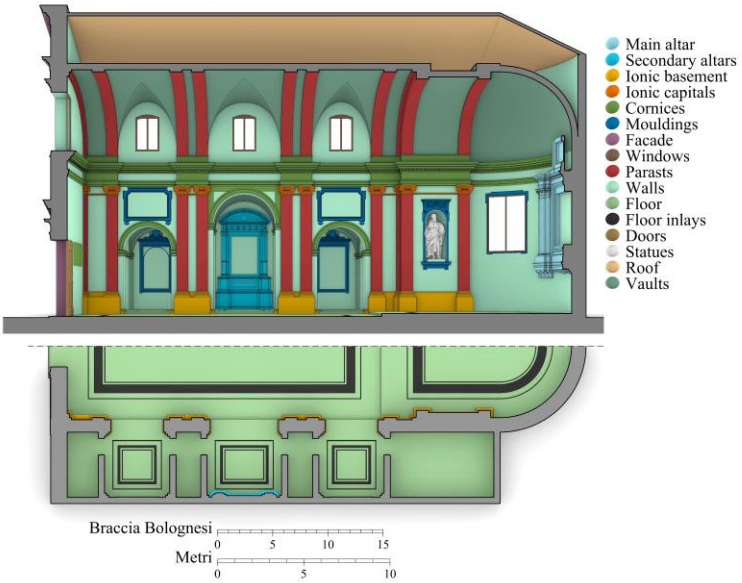

Fig. 10. Segmented NURBS 3D model of Spirito Santo church. Top: perspective section; bottom: orthographic plan; (Author: Riccardo Foschi).

4.3 Preparation of the Model for VR

The 3D model of the church was modeled in McNeel Rhinoceros, and the software that was chosen for implementing the VR experience was Unreal Engine. To transfer the model from Rhinoceros to Unreal it was converted into a mesh model, exported as OBJ, and processed into Blender before importing it into Unreal. Blender was used to optimize the mesh geometry, unwrap it, set up the lights, and bake the lights and shadows into the textures of the model (Fig. 11). Blender was also used to model the painting stands. In the original reconstruction prepared for the exhibition at the Pinacoteca of Bologna [21] human characters roaming around the main nave were also 3D modeled, however, to avoid distracting too much from the scope of the VR experience and to maximize the frame rate the characters were removed.

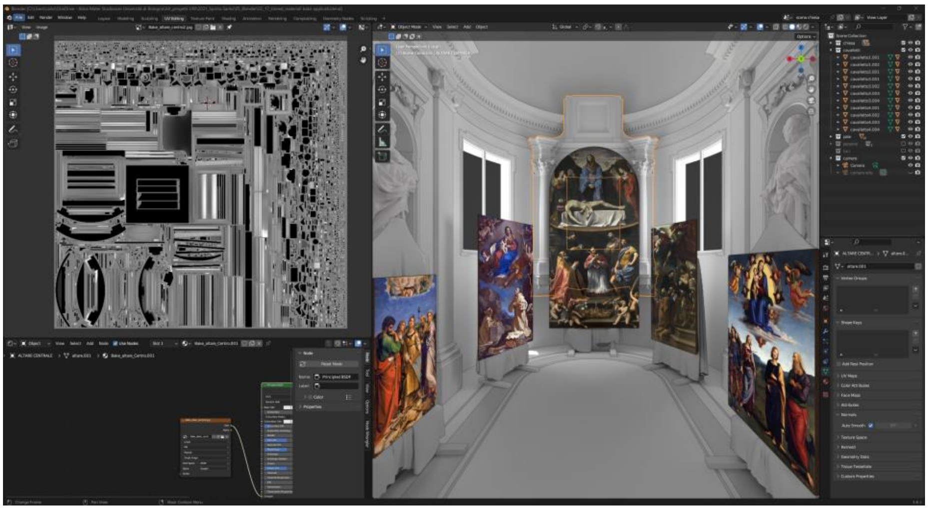

Fig. 11. Model optimization in Blender. Top left: unwrap and light baking of the main altar. Right: the main nave of the Spirito Santo church; (Author: Riccardo Foschi).

The textured model was then imported into Unreal where it was elaborated to make it explorable. Since the lights were already baked into the imported model they were deactivated completely in Unreal. The collision meshes were set up in order to avoid intersections between the player and the 3D elements, and the navigation domain was set up to allow player teleportation only on certain areas (main nave) of the church (Fig. 12). The teleportation was chosen as the preferred locomotion method because it demonstrated to be more tolerable by people not used to move in VR. Floating texts with the main info about the exhibited paintings (title, author, year,) were added next to each painting and they were set up to make them disappear and reappear at the click of a button on the VR controller or by proximity with the player (Fig. 13). Lastly, an additional player was set up to enable an external view of the VR player from the PC screen, this was useful to allow other visitors to see what the player saw and to guide the new players through oral command based on their observed behavior.

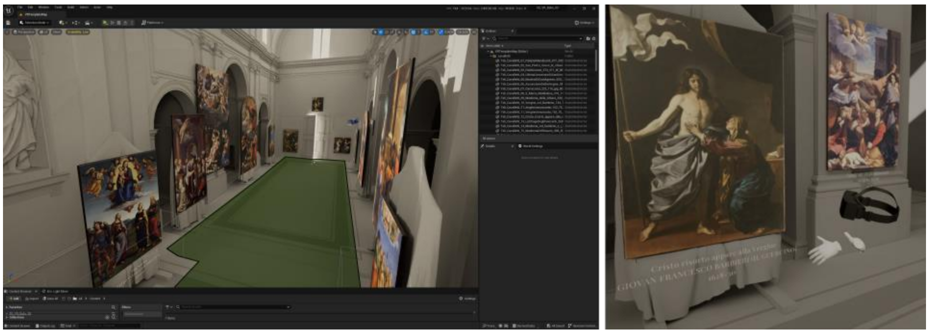

Fig. 12. Left: Setup of the scene in Unreal, the green area represents the area where the player can teleport which is automatically calculated based on the collision meshes and the navigation domain. Right: the VR player approaching a painting observed by the third-person player; (Author: Riccardo Foschi).

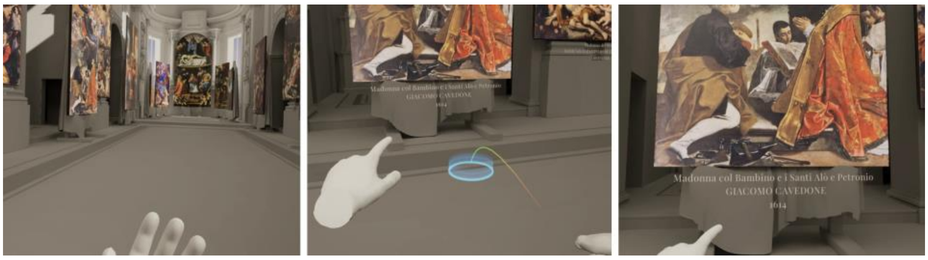

Fig. 13. First-person immersive experience through a stereoscopic 6-DoF VR headset (the light blue circle in the middle is the teleportation target location); (Author: Riccardo Foschi).

The hardware used to run the interactive experience was a VR-ready laptop connected via cable or wireless to a VR all-in-one headset. The full specifications of the hardware used are in Table 3. The application was also tested in the stand-alone mode (without the need for any PC connected) however due to hardware limitations of the headset and the high amount of polygons in the scene we opted to stream the VR experience from a more powerful laptop to achieve better overall video quality.

Table 3. Hardware used.

Laptop configuration	**VR headset**
CPU: i7-10750H 2.60GHz	Model: Meta Quest 2 All-in-One
GPU: NVIDIA GeForce RTX 2060 (6GB)	DoF: 6
RAM: 32 GB	IPD: Adjustable with 3 Settings
Operating system: Windows 64-bit	External sensors: not necessary
	Controllers: Two Touch Controllers
	Resolution: 1832 x 1920 Resolution Per Eye
	Refresh rate: 60Hz to 90Hz

The experience was presented and tested extensively at the *Notte Europea dei Ricercatori – Society* (Fig. 14) in September 2023, in Cesena. After navigating the scene in VR for 2 to 10 minutes, all the users reported no motion sickness nor major fatigue; the perceived scale of the space matched their expectations; and they didn't experience any of the unwanted illusions typical of 360° panoramas mentioned in Section 3.2.

Fig. 14. VR experience presented at the *Notte Europea dei Ricercatori – Society*, on September 2023, in Cesena; (Photo: Federico Fallvollita, Luciana Aloisio Delgado e Lucrezia Dell'amore).

5 Conclusions

The experience conducted allowed us to investigate the problem of depth perception in the available visualization technologies popularly used in the architectural field. Spherical stereoscopic and monocular 360° panoramas experienced with headsets, can improve immersivity while visualizing architectonic spaces, however, both stereoscopic and monocular 360° spherical panoramas viewed through both 3-DoF and 6-DoF headsets, often produce unwanted illusions (e.g. bigger spaces and sliding floors or walls) which cause a wrong perception of depth and shape. On the contrary, VR immersive real-time-rendered experience viewed through 6-DoF headsets is, among the ones analyzed, the visualization mode that guarantees the most reliable perception of both size and shape of architectonic spaces, thanks to its capability to reproduce a higher number of monocular and stereoscopic depth cues reliably. Furthermore, interactive VR experiences can give an improved perception of depth also to those individuals who have no stereo vision, because they can compensate for their missing binocular disparity with the interactive motion parallax monocular cue by moving their head in the 3D space. The case study of the hypothetical virtual 3D reconstruction of Canova's 1816 exhibition held in Spirito Santo church in Bologna confirmed the expectations and demonstrated that the VR real-time rendered interactive experience can be an effective and engaging tool to experience architectonic spaces faithfully in the context of museums and popular science.

Acknowledgements. We want to thank Alessio Costarelli who gathered the historical sources and organized the first exhibition, without whom this research would have never been possible; Gian Piero Cammarota whose help was para-

mount in the interpretation of some of the sources; Vittoria Barbiero, who contributed to the photographic survey of the church. All authors are equally responsible for the survey of the church and the study of the historical sources. All authors agree with the content of the manuscript and are responsible for the revision of the draft. Riccardo Foschi is responsible for the research on perception in VR, the writing of the initial draft, and the development of the images, 3D models, and the VR experience. Federico Fallavollita is responsible for organizing and conducting the on-field experimentation. Fabrizio Ivan Apollonio and Federico Fallavollita are responsible for the acquisition of the funds and the hardware technologies that made this research possible.

References

1. Migliari, R., Romor, J.: Perspective: Theories and Experiments on the "Veduta Vincolata". Journal for Geometry and Graphics, 19(1), pp. 57–77 (2015).
2. Karelin, D. A., Karelina, M. A.: Methods of reconstructions' presentation and the peculiarities of human perception. In Der Modelle Tugend 2.0, pp. 187–201 (2019).
3. Wann, J. P., Rushton, S., Mon-Williams, M.: Natural problems for stereoscopic depth perception in virtual environments. Vision Research, 35(19), pp. 2731–2736 (1995).
4. Jin, X., Meneely, J., Park, N. K.: Virtual Reality versus Real–World Space: Comparing Perceptions of Brightness, Glare, Spaciousness, and Visual Acuity. Journal of Interior Design, 47(2), pp. 31–50 (2022).
5. Apollonio, F. I., Fallavollita, F., Foschi, R.: The Critical Digital Model and Two Case Studies: the Churches of Santa Margherita and Santo Spirito in Bologna. Nexus Network Journal, pp. 1–8 (2023).
6. Wright, F. L.: Herbert F. Johnson house (Wing Point, Wisconsin), n. 3703.151. Avery Architectural & Fine Arts Library: Frank Lloyd Wright Drawings (Black & White Reference Photographs). https://library.artstor.org/asset/28515279 last accessed 2024/01/08.
7. Wright, F. L.: Rhododendron Chapel for Edgar J. Kaufmann, Jr. (Mill Run, Pennsylvania), n. 5308.002. Avery Architectural & Fine Arts Library: Frank Lloyd Wright Drawings (Black & White Reference Photographs), https://library.artstor.org/asset/28600684 last accessed 2024/01/08.
8. Wright, F. L.: Hoffman Jaguar Showroom (New York, New York), n. 5622.003. Avery Architectural & Fine Arts Library: Frank Lloyd Wright Drawings (Black & White Reference Photographs), https://library.artstor.org/asset/28602675 last accessed 2024/01/08.
9. Wright, F. L.: Phi Gamma Delta fraternity house (Madison, Wisconsin), n. 2504.002. Avery Architectural & Fine Arts Library: Frank Lloyd Wright Drawings (Black & White Reference Photographs), https://library.artstor.org/asset/28512735 last accessed 2024/01/08.
10. LaValle, S. M.: Virtual Reality. Cambridge University Press (2023).
11. Moon Illusion, https://en.m.wikipedia.org/wiki/Moon_illusion, last accessed 2024/01/08.
12. Pupillary distance, https://en.wikipedia.org/wiki/Pupillary_distance, last accessed 2024/01/08.
13. Marrinan, T., Papka, M. E.: Real-time omnidirectional stereo rendering: generating 360 surround-view panoramic images for comfortable immersive viewing. IEEE Transactions on Visualization and Computer Graphics, 27(5), pp. 2587–2596 (2021).
14. Peleg, S., Ben-Ezra, M., Pritch, Y.: Omnistereo: Panoramic stereo imaging. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(3), pp. 279–290 (2001).

15. Rendering Omni-directional Stereo Content, https://developers.google.com/static/vr/jump/rendering-ods-content.pdf, last accessed 2024/01/08.
16. Bourke, P.: Capturing omni-directional stereoscopic spherical projections with a single camera. In 16th International Conference on Virtual Systems and Multimedia, pp. 179–183. IEEE (2010).
17. Vergence-accommodation conflict, https://en.wikipedia.org/wiki/Vergence-accommodation_conflict, last accessed 2024/01/08.
18. Antonio Canova e Bologna. Alle origini della Pinacoteca, https://www.pinacotecabologna.beniculturali.it/en/home/2-non-categorizzato/2976-antonio-canova-e-bologna-alle-origini-della-pinacoteca, last accessed 2024/01/08.
19. Bolognini Amorini A.: Descrizione de' quadri restituiti a Bologna: i quali da' Francesi, che occuparono l'Italia nel MDCCXCVI. erano stati trasportati in Francia. Tipografia de' Franceschi alla Colomba, Bologna (1818).
20. Fanti, M.: Il restauro della chiesa dello Spirito Santo già Santa Maria dei Celestini. Lit. Bodoniana, Bologna (1965).
21. Apollonio, F. I., Fallavollita, F., Foschi, R.: La ricostruzione digitale della mostra allo spirito santo. In, Costarelli A. (ed.) A.B.C. Antonio Canova e Bologna, alle origini della pinacoteca di Bologna, pp. 104–113. Electa, Milano (2021).