Speculations on the representation of architecture in virtual reality

Hermund, Anders; Klint, Lars; Bundgård, Ture Slot

Publication date:
2017

Document Version:
Peer reviewed version

M. Aurel (eds.), *Back to the future: The next 50 years, 51st International Conference of the Architectural Science Association 2017,* pp. 1–10. ©2017, The Architectural Science Association and Victoria University of Wellington.

Speculations on the representation of architecture in virtual reality: How can we (continue to) simulate the unseen?

Anders Hermund[1], Ture Slot Bundgaard[2] and Lars Simon Klint[3]
[1,2,3] The Royal Danish Academy of Fine Arts, Schools of Architecture, Design and Conservation
School of Architecture, Copenhagen, Denmark
anders.hermund@kadk.dk, ture.bundgaard@kadk.dk, lklin@kadk.dk

Abstract: This paper discusses the present and future possibilities of representation models of architecture in new media such as virtual reality, seen in the broader context of tradition, perception, and neurology. Through comparative studies of real and virtual scenarios using eye tracking, the paper discusses if the constantly evolving toolset for architectural representation has in itself changed the core values of architecture, or if it is rather the level of skilful application of technology that can inflict on architecture and its quality. It is easy to contemplate virtual reality as an extension to the visual field of perception. However, this should not necessarily imply an acceptance of the dominance of vision over the other senses, and the much-criticized retinal architecture with its inherent loss of plasticity. Recent neurology studies indicate that 3D representation models in virtual reality are less demanding on the brain's working memory than 3D models seen on flat two-dimensional screens. This paper suggests that virtual reality representational architectural models can, if used correctly, significantly improve the imaginative role of architectural representation.

Keywords: virtual reality; representation; perception; neurology.

1. Architectural representation

Architecture and forms of representation have always been inseparable. It is not possible to build without having some sort of representation, of at least the idea, of the building. This representation can exist on many levels of abstraction and can be created in many ways, with the 3D scale model being one of the most obvious and widely used. It can be a simple model made of wood, or a wildly advanced digital model created with the latest print, milling, or simulation technologies. An important and interesting feature about the nature of the representation is that it is not in itself the final product, but only an essential means of achieving the realization of the full-scale architectural idea. One can add yet another layer to this representational hierarchy by insisting that even the building is not the final product, but only a framing or articulation of the interaction taking place within it.

> 'A building is not an end in itself; it frames, articulates, structures, gives significance, relates, separates and unites, facilitates and prohibits. Consequently, basic architectural experiences have a verb form rather than being nouns.' (Pallasmaa, 2005)

Thus, architectural representations serve at least two different purposes as a medium to convey both the rational concise organization of matter, but also as a framework for the imagination of the activities that make the building an architectural experience and a meaningful space for humans to occupy. This purpose, has not fundamentally changed, even though the digital tools and technologies are rapidly evolving and changing the possibilities of creating digital architectural representations on new levels of immersion. Focusing on architectural representation, the imaginative model, virtual reality (VR) becomes an interesting means of communicating and interacting through exactly those layers of experience Pallasmaa addresses.

2. A comparative study

To investigate the perception of space and potential use of VR models in architecture, an experiment was set up to compare a real physical space to its virtual counterpart.

2.1 Experiment setup

The experiment was set up using a medium size auditorium as the real life scenario and a BIM model of the same auditorium as the virtual scenario. In the real life scenario, we define the "normal condition" as a situation where a test person is experiencing a specified existing physical architectural space.

In the virtual scenario, the same architectural space, is presented to the test person through the HMD VR technology (a modified Oculus Rift SDK 2 fitted with eye tracking equipment), using a 3D digital architectural building information model. We used 60 students for the test: 30 for each scenario, respectively the real life scenario and the virtual scenario.

We used both eye tracking and a quantitative/qualitative interview matrix to assess to what extent the perception of space through VR technology holds similarities to the experience under normal, or close to normal, circumstances. In both scenarios we used the same two specific locations: the first being just inside the entrance to the space, and the second approximately in the middle of the room. We applied each test person's correct eye-height in the viewer. In addition to a silent ten-seconds-long eye tracking logging at the beginning of each location, a series of questions were posed to the test subjects, and the answers, along with the eye tracking data, were logged. The questions addressed four different areas of architectural perception: the space itself, lighting, sensation and estimation, and materials. The matrix questionnaire draws inspiration from C. Melhuish's thoughts on corpothetics (Rattenbury 2002) and S. Zeki's perception model involving microconsciousnesses (Mallgrave, 2010).

2.2 Reflective and immediate perception

While the abovementioned interview matrix data addresses a conscious reflection about perception of space, interesting on its own account, this paper introduces results from the first rough comparison analysis of the eye tracking data accompanying the questions. The collected eye tracking data related to each question, address a layer of behavioral perception less consciously controllable by the test subjects. Answering a question about the room involves both a conscious reflection on answering and a less conscious eye movement behavior. Thus, we could log the immediate perception results from eye tracking along the verbal reflective perception from the answers to the questions.

By tracking where the test subjects are actually looking, in respectively a real and a virtual scenario, we can compare the results and estimate how similar the eye movement behavior are in the two scenarios, when answering the same question from the same location.

While the eye tracking, once the equipment is set, is collecting data quite automatically, the subsequent analysis can be a very time-consuming process that requires the development of a consistent way of picking the right ranges for analysis. We picked eight instances of each participant's response (Table 1) representing both a free perception (no questions asked) and a more guided perception (by posing questions).

Table 1: Instances and questions selected for eye tracking

#	Description of instance or question
1	Instance (no specific question): First 10 seconds after entering the room
2	Question - The room: The want to go deeper into the room
3	Question - The lighting: Is the room sufficiently lit
4	Instance (no specific question): First 10 seconds at second location in the room
5	Question - Sensation and estimation: Estimation of the height of the room
6	Question - Sensation and estimation: Estimation of the width of the room
7	Question - Sensation and estimation: What row to sit at for a lecture
8	Question - Materials: Sensation of the quality of materials

For the analysis, we extracted 10 seconds of eye tracking data at the beginning of each of the two locations and subsequently the 10 seconds immediately after a question had been posed.

The resulting maps from the eye tracking contain a lot of different information, such as Time To First Fixation (TTFF), Time Spent in the different areas of interest, and the number of fixations. A fixation is a cluster of gaze points in the same spot. Below is an example (Figure 1) of the visualization of eye tracking data in respectively an aggregated heat map (A) and an opacity map (B). The Opacity map has been converted to grayscale, and using the histogram produced by Adobe Photoshop (C) it is possible to measure the average intensity value in the image to measure the balance between the number of black and white pixels in the grayscale image.

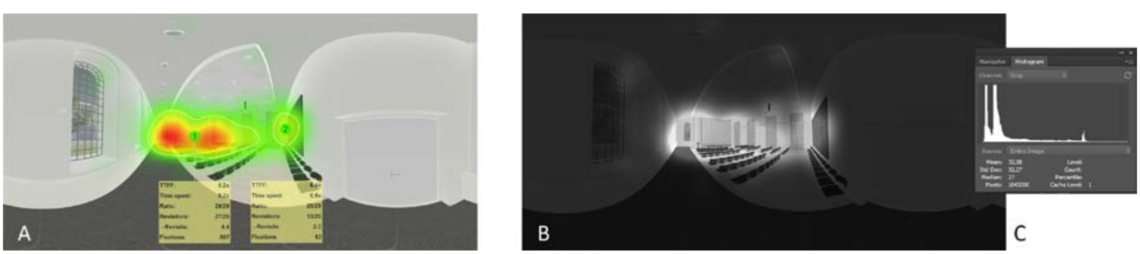

Figure 1: (A) The heatmap on a gradient scale from red to green showing eye tracking data related to a set of answers to a specific question. (B) An opacity map of the same data without the colours but with a gradient intensity. (C) The Histogram used to calculate the average intensity value in the image (Mean).

Combining the opacity map feature with the Adobe Photoshop *histogram* and *difference filter* provided a precise and relatively fast way of comparing the overall areas of interest (AOIs) in the various images. This comparison is designed to give a general indication of the trends between the real and virtual environment. In this way, we managed to cross-reference the AOIs of all the images as described in the following method of comparing.

2.3 Method of comparing

Comparing the images from the eye tracking using the opacity maps has an obvious advantage in relation to the amount of time needed to perform the analysis. Using the visual output and layering it in Photoshop using the difference filter gives a possibility to see and measure the difference of fixations.

Figure 2: Instance 1 real life (A) compared to instance 1 in virtual reality (B) combined into the resulting difference map (C). (source: the author, 2017)

In the example (Figure 2) the comparison between eye movements related to the location in respectively the real life scenario (A) and the virtual reality scenario (B) is created by finding all the non-corresponding areas, the difference, of the two images (C). This resulting image of the differences between the two previous (A & B) can then be measured by using the histogram mean value.

Figure 3: Question 2 in RL (A) & VR (B) and differences (C). (source: the author, 2017)

The example (Figure 3) show the same procedure with a comparison between two scenarios, this time with the same question in the real life, *RL,* scenario (A) as in the virtual reality*, VR*, scenario (B) with the resulting difference (C).

In the third example (Figure 4), the comparison of eye tracking in the same location, but with two different questions, results in more white pixels in the difference image (C) signifying a higher amount of difference than in the previous two examples. By measuring all the images against each other, we should be able to spot where the highest amount of difference exists in relation to fixations in the various scenarios.

Figure 4: Question 5 RL (A) compared with question 7 VR (B) and the resulting differences (C). (source: the author, 2017)

2.4 Differentiation

The average intensity value in the grayscale pictures are represented by the 'Mean' value, measured using the Adobe Photoshop software. The results from applying this method on all the combinations of the images are collected in the following tables.

Table 2: The difference between same instances and questions

Same instances and questions		Difference Filter (Mean)	%	Average Mean	Average %	Total average %	Of full results%
RL 1	VR 1	4,24	1,66	3,0333333	1,19	**2,04**	**37,09**
RL 2	VR 2	2,18	0,85				
RL 3	VR 3	2,68	1,05				
RL 4	VR 4	9,60	3,76	5,19125	2,54		
RL 5	VR 5	6,64	2,60				
RL 6	VR 6	2,88	1,13				
RL 7	VR 7	3,70	1,45				
RL 8	VR 8	9,61	3,77				

The first measurements (Table 2) show the difference between the same instances and same questions in respectively real life and virtual reality, RL and VR. The number after RL or VR corresponds to the number of instance or question (Table 1). The horizontal separation line in the middle of the table is between the two different locations. RL & VR 1-3 are the first location in the scenario and the RL & VR 4-8 are the second location, a bit deeper in the room.

In the next table (Table 3), the measurements are cross-referenced comparing the eye movements between different questions, while still naturally comparing only within the same location, since a change between locations would make the incongruent images useless for the purpose of this investigation. Obviously, there are more combinations of non-matching instances and questions than of matching, which is why the percentage of the full results, both matching and non-matching instances and questions, is calculated in the last column.

Table 3: The difference between mixed instances and questions

Mixed instances and questions		Difference Filter (Mean)	%	Average Mean	Average %	Total average %	Of full results%
RL 1	VR 2	4,64	1,82	4,1244444	1,62	**3,45**	**62,91**
RL 1	VR 3	3,68	1,44				
RL 1	RL 2	5,02	1,97				
RL 1	RL 3	3,50	1,37				
RL 2	VR 1	6,32	2,48				
RL 2	VR 3	4,09	1,60				
RL 2	RL 3	2,31	0,91				
RL 3	VR 1	5,42	2,13				
RL 3	VR 2	2,14	0,84				
RL 4	VR 5	9,30	3,65	8,815	4,00		
RL 4	VR 6	8,96	3,51				
RL 4	VR 7	9,87	3,87				
RL 4	VR 8	12,65	4,96				
RL 4	RL 5	9,03	3,54				
RL 4	RL 6	8,46	3,32				
RL 4	RL 7	10,42	4,09				
RL 4	RL 8	7,37	2,89				
RL 5	VR 4	15,91	6,24				
RL 5	VR 6	13,35	5,24				
RL 5	VR 7	14,74	5,78				
RL 5	VR 8	16,44	6,45				
RL 5	RL 6	13,03	5,11				
RL 5	RL 7	13,96	5,47				
RL 5	RL 8	11,58	4,54				
RL 6	VR 4	13,03	5,11				
RL 6	VR 5	11,21	4,40				
RL 6	VR 7	3,70	1,45				
RL 6	VR 8	9,94	3,90				
RL 6	RL 7	4,25	1,67				
RL 6	RL 8	4,19	1,64				
RL 7	VR 4	15,13	5,93				
RL 7	VR 5	11,19	4,39				
RL 7	VR 6	4,39	1,72				
RL 7	VR 8	12,56	4,93				
RL 7	RL 8	6,41	2,51				
RL 8	VR 4	12,95	5,08				
RL 8	VR 5	10,27	4,03				
RL 8	VR 6	5,66	2,22				
RL 8	VR 7	6,38	2,50				

2.5 Findings

Comparing the result from the investigation, we can see that the correlation between same questions in the two different scenarios is higher than the correlation between different questions (Table 4).

The difference between that matching instances and questions in RL and VR amounts to 37.09 % of the full result while the difference between non-matching questions amounts to 62.91 % of the full result. The divergence between the two results are thus 25.83 %.

Table 4: Divergence in percent

Same instances and questions	Mixed instances and questions	Resulting Divergence
37,09	62,91	25,83

This shows that a there is actually a better correspondence in eye movements between a scenario in real life and its counterpart in virtual reality, in the specific situations, than between eye movements in general when observing the space. In other words, the 25.83 % divergence shows clearly more consistency between the RL and VR scenarios than what can be attributed to coincidence.

Table 5: Results of mixed instances and questions (1, 4 & 8 screened out)

Mixed instances and question		Difference Filter (Mean)	Average Mean	Average %	Of full results %
RL 2	VR 3	4,09	8,20	3,21	69,39
RL 2	RL 3	2,31			
RL 3	VR 2	2,14			
RL 5	VR 6	13,35			
RL 5	VR 7	14,74			
RL 5	RL 6	13,03			
RL 5	RL 7	13,96			
RL 6	VR 5	11,21			
RL 6	VR 7	3,7			
RL 6	RL 7	4,25			
RL 7	VR 5	11,19			
RL 7	VR 6	4,39			

It could be argued, that a screening out of some of the instances and questions could by useful for the clarity of the investigation. Especially instances 1 and 4, which are not accompanied by questions can be difficult to compare in this way, since a free look around can vary a lot with no direct link to perception of a specific sensation or quality of the space. In addition, question 8, dealing with the quality of materials, could be screened out as well due to both a slight fuzziness in the articulating of the question, but also due to an obvious (though intended) lack of material quality in the VR model, which was not constructive in this specific case.

If these premises are accepted, the resulting measurements (Table 5, Table 6) shows the same trend with even more clarity.

Table 6: Results of same instances and questions (1, 4 & 8 screened out)

Same instances and questions		Difference Filter (Mean)	Average Mean	Average %	Of full results %
RL 2	VR 2	2,18	3,62	1,42	30,61
RL 3	VR 3	2,68			
RL 5	VR 5	6,64			
RL 6	VR 6	2,88			
RL 7	VR 7	3,7			

The calculation (Table 7) now shows a resulting divergence of 38.78 % between the matching and the non-matching instances and questions.

Table 7: Divergence in percent (1, 4 & 8 screened out)

Same instances and questions	Mixed instances and questions	Resulting Divergence
30,61	69,39	38,78

3. Perceiving space

Perception of space has always been a crucial area of interest to architects. The relation of body, consciousness, and perception has been the focus of much consideration in attempts to define a phenomenological structure of perception (Merleau-Ponty, 1962) that has continued to intrigue us.

Recently the field of perceiving aesthetics is being addressed, by neurology, with studies of the visual cortex and the biological foundation of vision. For instance, it can be discerned that colour is perceived before form, which is again perceived before motion (Zeki, 1999). Architecture as environmental neuroaesthetics as summerized by N. Rostrup (Lauring, 2014) activates many different areas of the brain, that is the perception of a building façade may be mentally processed both as a background and as a configuration of objects, involving different areas of the brain.

Even though the limited space of this paper does not allow much elaboration, such studies of the brain must be addressed to assist the understanding of perception, and should be embraced by architects as a new source of information to the intricate connections between perception of space, aesthetics and behaviour.

3.1 Perception in VR

The phenomenon of perception in VR has been described by Jonathan Steuer (1992) as a combination of the variables ‘vividness’ and ‘interactivity’ together forming the level of sensation of telepresence.

> Vividness means the representational richness of a mediated environment as defined by its formal features, that is, the way in which an environment presents information to the senses. […] Interactivity is defined as the extent to which users can participate in modifying the form and content of a mediated environment in real time.

It is noteworthy that interactivity seems to be the most important factor of the two, overruling the level of realism in the VR scene (Slater, 2009). Research that is more recent shows that e.g. distance perception is not decreasing significantly even with a severe degradation of visual realism in a scene (Vaziri, 2017). An argument is that the sense of being present or immersed in a virtual environment highly depends on the ability to act and receive direct feedback from one's actions in the virtual environment (Sanchez-Vives, 2005). This is sustained by the suggestion that the reflections on experiences of architectural space in VR can be simulated corresponding to the same scenario in real life, in many aspects (Hermund, 2016). Comparisons between answers in respectively a real environment and a virtual reality model, indicates a high similarity in the answers given about the perception of space.

3.2 VR perception and neurology

Neurology studies indicate that the brain uses less working memory when perceiving representation models in VR compared to flat two-dimensional representation models due to reduced task complexity and an eliminated need for cross-dimensionality in the translation of geometry from 2D to 3D (Neubauer, 2010). A recent test (DR2, 2017), involving assessment of a small 3D game in VR and on a traditional screen while measuring the electroencephalogram (EEG), showed that the VR version did actually induce less cognitive load on the brain than the same task performed on a flat screen. This is in accordance with results from testing 2D & 3D non-immersive and 3D immersive scenarios (Kozhevnikov, 2012) concluding that the desktop graphics are ineffective or even counterproductive in relation to 3D immersive environments when aiming for tasks in simulated real world scenarios from an egocentric frame of reference. VR, in this case, creates a more natural environment for the brain, than 2D screens.

4. Discussion and Conclusion – simulating the unseen

It has become apparent that a virtual environment through the visual dimension of perception, in combination with interactivity feedback, can be helpful in generating immersiveness, the sensation of being present, even while one is maintaining a conscious knowledge that it is only an illusion – a model.

The model, and digital models, are still a central means of creating architecture. Until recently, the digital regime, although fundamentally a 3D regime, all happened on flat screens. However, VR can perhaps be a step towards a more natural way of representing and dealing with digital architecture. The investigation presented in this paper suggests that VR can indeed simulate a physical scenario to a degree where human behaviour shows correspondences. Furthermore, a virtual scenario contains the possibility to incorporate elements of interaction that cannot be provided to the same extent, using traditional drawings or even non-immersive 2D and 3D. That is exactly what makes a difference and where VR has a potential to surpass a static retinal architecture. In this case, the visual world is not a grand illusion (Noë, 2002), but the visual representations can combined with our sensorimotor skills provide a usable experience of perception. This sort of movement is exactly what can be supported by the clever use of VR simulations. Instead of striving for a photorealistic virtual paradise we suggest that the virtual reality simulations are opened to imagination by actively removing some of the visual effects and adding more possibilities for interaction. In this way, interaction becomes crucial in simulating the unseen cohesion of architecture.

That our brains are combining many different areas to create an experience does not necessarily mean that we need much more than a few indications to imagine a scenario. The crucial part is to find those indications which are needed, and those which are just noise in the model. This is presumably highly dependent on the profession of the people who experience the simulation. We need to establish a sufficiently solid general foundation for assessing where a VR representation, on a specific level of detail, can convey a realistic, and still imaginative, scenario for a future architectural vision. Then the next step is to look closer at the different professions in the construction sector and establish a parametric adjustable VR scenario responder system, which can be applied and adjusted for the viewer, when needed throughout the whole lifecycle of a project.

We have seen that architectural virtual reality exists, and believe that a collaboration between architecture, neurology, and technology can be a promising path towards better understanding of perception of architectural space. That is why we find it appropriate to suggests that virtual reality representational architectural models can significantly improve the imaginative role of architectural representation, and thus eventually a better architectural quality.

5. References

DR2 (2017). *So ein Ding: De nye digitale virkeligheder,* 08.05.2017, DR2, Danish Broadcasting Corporation 2017.

Hermund, A., & Klint, L. (2016). *Virtual and Physical Architectural Atmosphere*, NZAAR International Event Series on Natural and Built Environment, Cities, Sustainability and Advanced Engineering, New Zealand 2016.

Kozhevnikov, M. & Dhond, R. P. (2012). *Understanding immersivity: image generation and transformation processes in 3D immersive environments*, Frontiers in Psychology, Perception Science, August 2012, volume 3, article 284

Lauring, J. O. (ed.). (2014). *An Introduction to Neuroaesthetics: The Neuroscientific Approach to Aesthetic Experience, Artistic Creativity, and Arts Appreciation*, Museum Tusculanum Press, Copenhagen, Denmark.

Mallgrave, H, M. (2010). *The Architect's Brain: Neuroscience, Creativity, and Architecture*, Wiley-Blackwell.

Merleau-Ponty, M. (1962). Phenomenology of Perception, Routeledge & Kegan Paul, London.

Neubauer, A. C. et al. (2010). *Two- vs. three-dimensional presentation of mental rotation tasks*, Intelligence, Volume 38, Issue 5, September–October 2010, Pages 529-539.

Noë, A. (2002). *Is the Visual World a Grand Illusion?* Journal of Consciousness Studies, 9, no 5-6, 2202, pp.1-12

Pallasmaa, J. (2005). *The eyes of the skin: Architecture and the senses*, Chichester, Wiley-Academy.

Rattenbury, K. (ed.). (2002). *This Is Not Architecture: Media constructions,* Routledge.

Sanchez-Vives, M.V. & Slater, M. (2005). *From presence to consciousness in virtual reality*, Nature Reviews Neuroscience, 6, 332-339.

Slater, M. et al. (2009). *How we experience immersive virtual environments: the concept of presence and its measurement*, Anuario de Psicología 2009, vol. 40, nº 2, 193-210, Facultat de Psicologia Universitat de Barcelona.

Steuer, J. (1992). *Defining Virtual Reality: Dimensions Determining Telepresence*, Journal of Communication, 42(4) (Autumn, 1992), 73-93.

Vaziri, K. et al. (2017). *Impact of Visual and Experiential Realism on Distance Perception in VR using a Custom Video See-Through System*. In Proceedings of SAP ’17, Cottbus, Germany, September 16-17, 2017.

Zeki, S. (1999). *Inner Vision: An Exploration of Art and the Brain*, University College London, Oxford University Press,
Inc., New York.