BILT North America 2018

Union Station Hotel, St Louis
August 9-11, 2018

Session 2.3

Augmented reality and deep learning in the design process

Cesar Escalante, HOK and Alberto Tono HOK

Visualizzazione
- Occlusion
- RECOGNITION
- Tracking

Collaboration
- Pier-to-pier

Contents

Class Description

In this session, we will learn how augmented reality and deep learning in computer vision can be delivered inthe design process.
Today's technologies have limitations and barriers causing friction for designer's adoption. We will talk about the main areas in which we can apply these innovations.
These technologies embed a big "Paradigm Shift" in our industry from a 2D to a real 3D environment in which carries the main value of human interaction and sharing experience, without the focus of the human being.

About the Speaker:

Alberto serves as Research and Development Intern at HOK. Prior to HOK, Alberto worked as a freelancer in Italy providing computational solutions for BIM workflow for large projects. His studies took him to Harbin Institute of Technology (HIT) in China and to Rene van Zuuk Architect Studio in the Netherlands. He is very keen on computational design and machine learning. Learning and this passion inspired him to create the Italian Dynamo User Group and a Deeplearningitalia community.

Cesar Escalante serves as Digital Design Technology Manager at HOK providing training, BIM implementation, and computation design support to the San Francisco and Seattle offices. He is a licensed Architect with 15 years of experience leveraging architecture and design tools as a practitioner, consultant, and educator. He is passionate about computational design, digital prototyping, and a visual programming. He is the founder and coordinator of the San Francisco Dynamo User Group.

Goals

The final goal of this presentation is to defend the theory about why augmented reality will be crucial in the AEC industry, and in particular the design process. We will stress the importance of a **human-centric design**(Kay Sargent definition of the HOK workflow) where values such as **simplicity, collaboration, and gamification** are at the foundation of this process. If we desire to understand the real value we need to dive into its implication with

artificial intelligence and in particular, deep learning and computer vision.This is why we are focused on analysis and product testing based on Apple and LEGO services, companies that embrace these values.

Market Analysis

In order to evaluate a strategic approach to a research and development program that can have an immediate impact in optimizing the current workflow, an accurate market analysis is needed.Thus, using the various reports for augmented reality, a broad vision about the current market can be created. Presently, there is a lack of information regarding the AEC industry. Afterthe report analysis (Digi-capitals, VR/ARA, IEEE,Gartner, ABI, statista) it is possible to understand that there is a considerable shift in theAR market, right now the most amount of fundsare involved in solving technical solutions in the AR space, such as occlusion, persistence, AR cloud so forth.
There are a few parameters to keep in mind with the analysis of the market:

- Market size: 15billion 2017
- Investment: 2.5billion 2017
- Forecast: 210billion for 2022
- Number of HMD: 21 million devices

If we look deeper we can see that 50% of investment went to Magic Leap and Unity but also that the numbers of HMDs are around 2% of the mobile devices.
Unquestionably, it will be a crucial trend in the future, at the moment however,there is a need to overcome a bulk of challenges before we reach a massive and democratic adoption.

A Paradigm Shift

All these innovations such as augmented reality, virtual reality, and immersive experience with AI support, as Thomas Khun described years ago:“will shape a **Paradigm Shift** that can guide investigation in the absence of rules and standards”. This paradigm will drive the shift from a 2D to 3D atmosphere, where all the information is related to a specific point in the real world. This will be based on the AR Cloud, where virtual objects can be placed in their real locations as an overlay.

Figure 1 Thomas Khun

This deals with Augmented Reality because it focuses more on a unique and contextualized experience that can be shared with others in a more social setting.
Cesar Escalantestates, “We realized that the biggest challenge of Building Information Modeling (BIM) is not the hardware,not software either, but it’s real people and people’s mindsetsare our biggest attractors.”
Teamworkand collaboration are the most important soft-skills required in this space, but the challenges are great. The numbers demonstrate interest in overcoming these challenges and bringing AR into reality. If we analyze the number of patents, AR has 78,886 results and the AR cloud has 16,564 results.
AR cloud for AEC can educate everyone about buildings. In this way, we can provide transparency about architecture and thus, information will be displayed as an overlay on top of the real building. The building will be “naked” in front of the future generations. Structural, design, and sustainable analysis can be displayed on top of the building in real-time in results of having a connection between IoT devices and AR capabilities.

Values

It is always crucial to identify a metric of the success in order to prove the value of our research.
Let’s explore this below with some parameters such as time, ergonomics and others.

Time Saving

Based on collecting feedback and data about performance improvement in a scientific approach we have articulated a formula to generate a time saving technique.
The multiusers session in AR allows users to review the model in a quicker way, proportional to the number of users.

$$P_r = n * t * f_a$$

P_r=Performance of revision.
n = number of people.
t = time spent for revision.
f_a = it is the accuracy factor of the revision driven by the analytic data that we can extract form each session.
Before the equation was only related to time, based on the fact that the conversation was guided from a shared screen.

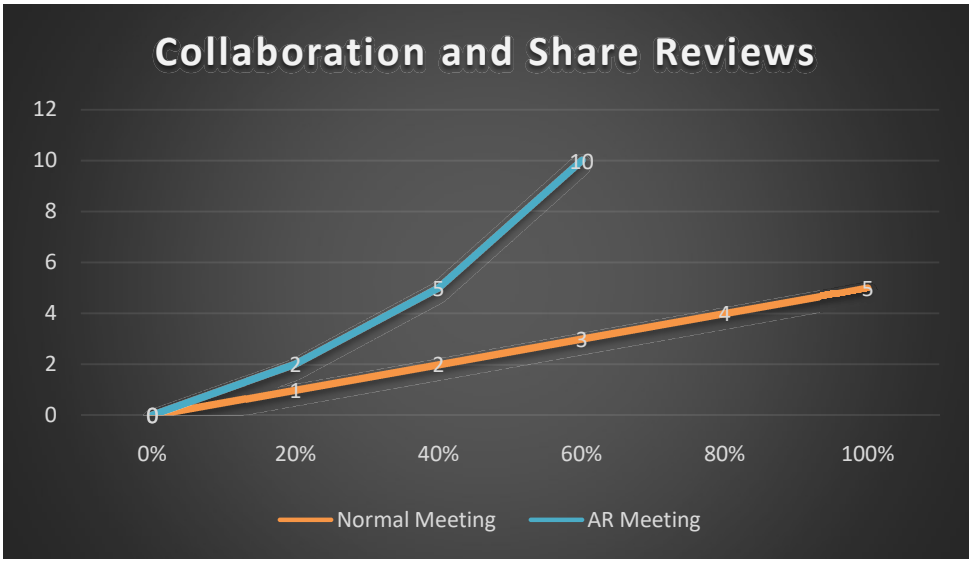

Figure 2 Chart of AR collaboration improvement

The most important achievement provided by augmented reality is its way of drastically decreasing the cognitive workload needed in order to understand the project. The time that you need in order to set up the experience is proportional to your model size doesn't require the manual work.

Work condition

Another relevant research topic is about its improvement for fitting the task to the worker in order to maximize productivity while reducing discomfort, fatigue, and injury.

Ergonomic

Ergonomics also hasahuge impact in the way we work. Research conducted by Human Scale Consulting shows that 53% of workers report discomfort in the work environment, from 5/7 hours lost during the week and 61 Billion $ of annual lost in the US market. We spend more than 60% of our life sitting at work and intransportation etc. This number is increasing exponentially.

The LPL enzyme that facilitates the transportation of fatty acids inside the cells throughout our bodies, such as muscle, adipose, and heart tissue will decrease its level after a certain time of inactivity. So crucial for our bodies to keep and maintain a dynamic life.

AI and AR will affect the way in which we work and interact with the surroundingsin addition to time and how we willdesign directly in the physical space.Potentially in the future, no monitor will be needed. We will attain moreorganizedworkspaces and better work-privacy.

Comfortable device and wireless. Comfortable, wireless device

Nowadays within the workplaces of architecture and design,there arevarious pursuits for computing comfort, but ideally, we can replace our monitors and reimagine the way in which we comfortably use computers. After 48minutes of testing Hololens, it started to raise my level of discomfort. Alternatively,I could work with Meta formore than 3hours with non-existent weight on my nose such as with other HMD systems.Although Meta has a favourable FOV which makes the experience more immersive, Meta presently does not have wireless capabilities.

Capabilities

The proliferation of this digital toolset has given us the ability to transcend traditional architectural practices for the first time in generations. XR tools are fostering innovation and disrupting legacy workflows across the architecture industry.

Our reality becomes immersive at the intersection of digital computerization and real-time computer vision overlay. This presentation won't focus on the right or wrong naming conventionssuch as immersive reality, mixed reality, augmented reality, virtual reality ect. We will be focused on the real values which you can integrate into your firm and how to directly apply in the design process. Some of the applications of computer vision have pursuedthe potential of ARwithin their operation. The range of work includes detailed augmented reality building, virtual 3D data overlays on physical scale models, and real-time object tracking to simplify user manipulation in custom experience. Augmented reality implements a user interaction, giving a first-person manipulation of digital elements and 3D representation of objects around us. As a result, we retain more memory, empathize more with our surroundings, and we are able to gain insight into information at an outstandingly precise real-time depiction.

As architects, we're accustomed to looking at a design be it a rendering, a sketch or a physical model and interpreting how a building or landscape should look and feel spatially as a finished project. The same isn't necessarily true for our clients, who aren't creating and interpreting designs on a daily basis. AR tools give our

designers a powerful aid for explaining ideas and concepts to clients and other stakeholders. It's this use as a presentation or communication tool.

In order to better frame a research programwe can divide it is into 3 macro areas:

1. Visualization
2. Collaboration
3. ARCloud

Each can be analysed with a deeper understanding of its each of them with a deeper understanding of its artificial intelligence implications.

Visualization

The virtual representation of models are getting better and better but at the moment they are not able to replicate manyof natural effects such as real-time iridescence (also known as goniochromism) and others.

Analysis

Reducing the cognitive load of sustainable analysis, such as shadows, daylight, and wind, which can be difficult to convey, can be displayed in real time using AR and thus,having theability to change various outcomes in real time with the use of our hands and gestures will enable us to speed up the decision process at the beginning of the design process.

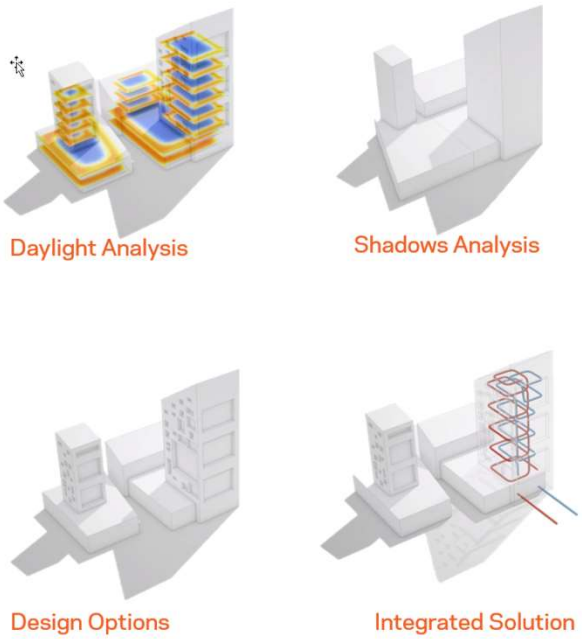

Figure 3 Analysis displayed in AR during the research

One parameter that influences a substantial score in the energy certifications is the ability to bring solutions to focus on the improvement of human life in the space.

Patterns

All the computer vision algorithms aim to understand the world around us to recognize common patterns and simplify or enhance the human capabilities.

Perception can be seen as the capacity to have a sentient experience based on qualia. On the contrary, bits in the memory, electromagnetic energy, or electrical signals in the brain or in a machine do not produce qualia, but they can facilitate and enhance a better and faster understanding of such in a proper way.

On the other hand,**Comprehension** is the capacity to organize, integrate, and understand the information contained in a given experience, and to give it the widest possible meaning within the context of the totality of our experience. As we define these words we can understand how the artificial intelligence can help humans perform this action in a better way, thanks to the knowledge graphs we can have a wide overview of the context around us that can make this experience uniqueevery time. Also thanks to artificial intelligence (AI) we can perceive to discover some hidden pattern in the data that allows us to understand in a flash of insight. As we defined earlier in the presentation, thanks to augmented reality and IoT sensors, we can easily display the colourmapping of some analysis, structure, daylight, humidity, temperature and many of others.Having a camera that can point to the building and provide to us added insight based on our experiences can allow us to examine the encoded memory in nature.

The difference between pattern and code is important in order to understand the potential of AR. While code is a common convention established between human beings, patterns appear to be a natural manifestation.

Actionis usually not driven by consciousness within the machine, but they have been placed in the same way, when an entity can receive information from the physical environment and affect the environment in return.Anexcellent example is provided by Philip Beesley (Figure 4) and his curiosity based machine learning installation.

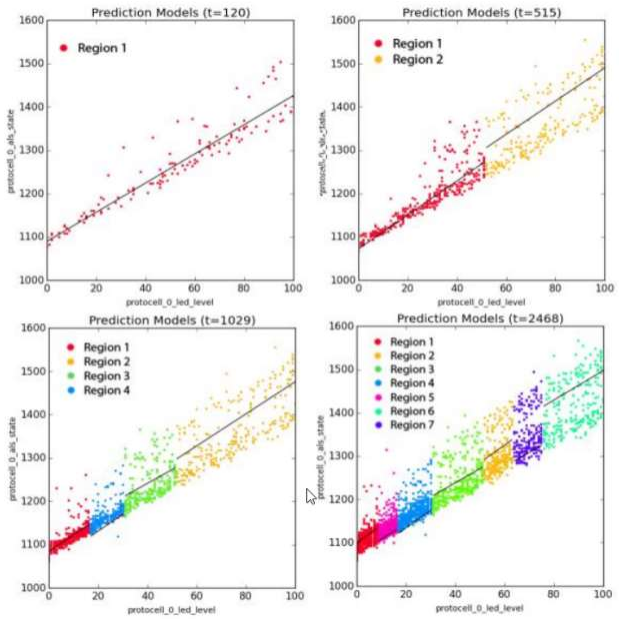

Figure 4 Hybrid Sentient Canopy: Curiosity Based Machine Learning Action based on Knowledge and curiosity, Evolution of the prediction models, each point represents an example of action that the machine can do accordingly with the behaviour of the people in the space. They machine learn from the people behaviour and act in a control way. "Curiosity-Based Learning Algorithm for Distributed Interactive Sculptural System" from Philip Beesley, Matthew T.K. Chan, Rob Gorbet.

Having this in mind, we can let the machine processsome of our actions, sentiments and then perform tasks accordingly. Identity and free will are also important components in order to understand the value of the insights given by the AI. These are key components of a human-centric design based on collaboration and sharing values. In the end, we can process a Re-entrant feedback loop that will allow us to create a better experience and decision thank to the support of AI that is capable to recognize properly different elements around us and interact with them.

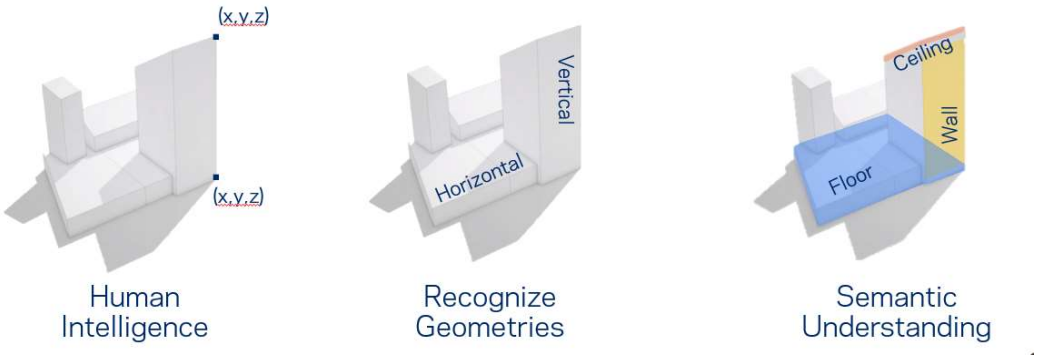

Figure 5 Semantic Understanding of the surroundings.

Accordingly with deep learning capabilities it is possible to have a semantic understanding of the environment around us, the camera can divide the meshes created between floor, wall and ceiling and it will be possible to reconstruct a BIM model in real time only visiting and playing in a specific environment.

Metadata

Another important aspect of the augmented reality experience is the value of metadata, being that all the information that we bring in our BIM models need to be translated as well. Presently, great results are achieved and provided by Forge with its integration with Unity. In this way, it is possible to transfer into the virtual environment and all the information needed.

Collaboration

Multiplayer in AR opens the door for many applications such as collaborative design in 3D. For instance, two design teams can work in a 3D architectural model from different locations by connecting to a shared AR scene. Not only theycan interact and edit the design in real/time, they can also leave design feedback for other team members directly in the 3D environment. Clearly this provides benefits for big design teams. The experience can also be guided from the web or directly in AR while the other players can be immersed in the virtual space.

Figure 6 Multiusers guided-experience in AR and VR

Sharing and Multiusers approach.

Untill now, AR experiences have been mainly focused on asingle-player. In a full multiplayer AR game, users can interact with shared experiences simultaneously. This means that if one user places a virtual object in the scene, other users will be able to see the changes. On the other hand, in a single player scenario the game is run locally, and while this is shared in the cloud, it is not shared with other players. Similarly, the multiplayer games demonstrate a peer-to-peer system where data is not shared on the cloud, thereby avoiding privacy concerns.
Thanks to the new WebXR, APIs will be able to directly manage all the experience throughouta web page without the need to install applications.

Human connection

One of the main topics of this conference is related to the new ways of **connecting people**, in order to improve human connection.Right now space and time are the main barriers in front of human relationships. In bigger enterprises it is always difficult to establish good solutions and strategies to overcome this problem. In this moment,on of the main project categories at HOK is related to aviation and one of the main challenges in this is to control the flow of people within the airport. The software that runs thetraffic analysis is costlyand is based on later design phases. Our solution is to bring this analysis at the beginning in a simplified and gamified way. Using deep reinforcement learning it is possible to use the behaviour of the people interacting with the game in order to collect data useful to the machine to be able to autonomously provide the solutions in optimizing the shape accordingly to the crowd simulations.
The challenge in this case was to build a dynamic navigation mesh on top of our model in order to allow our AI to perform correctly and enable the finding of the shortest path,the algorithmdoing just that. These kinds of operationsinvolving AI are difficult to perform in an AEC environment because most of the time they don’t have the same AI culture as the game engine software.
On the navigation mesh we can simulate the flow of people to perform crowd analysis at the early stage of design.Thus, if we want to optimize the shape of our airport we can implement a deep reinforcement learning in order to let the machine train and test by itself untill it provides us insight.
We can associate each gate to each player in our environment and at the same time provide each player a neural network which learns the main pattern in a different way. The policy will always focus on reaching an optimized solution that allows all people to reach the airport terminal in the shortest amount of time.

Transfer Learning

In the case that we associate each terminal an agent and each agent starts with a random flow that represents the distribution of the passengers in the scene, the learning in each agent also shapes the anticipated learning of the other agents in the environment. With LOLA learning (learning with opponent learning awareness) we can include an additional term which accounts for the impact of one agent's policy on the anticipated parameter update of the other agents. In this way it is possible to converge to the Nash equilibrium (an optimal solution for both agents). The agents learn to cooperate out of self-interest.
The time in which people need to reach the specific location in the airport will be a parameter. However, other parameters will drive the shape of the building in order to optimize the architectural mode for reducing the time in which everyone reaches their final destination.
This learning method for multi-agent settings considers the learning processes of other agents. This is the perfect use case of this specific approach where two different neural networks collaborate in order to find a solution.

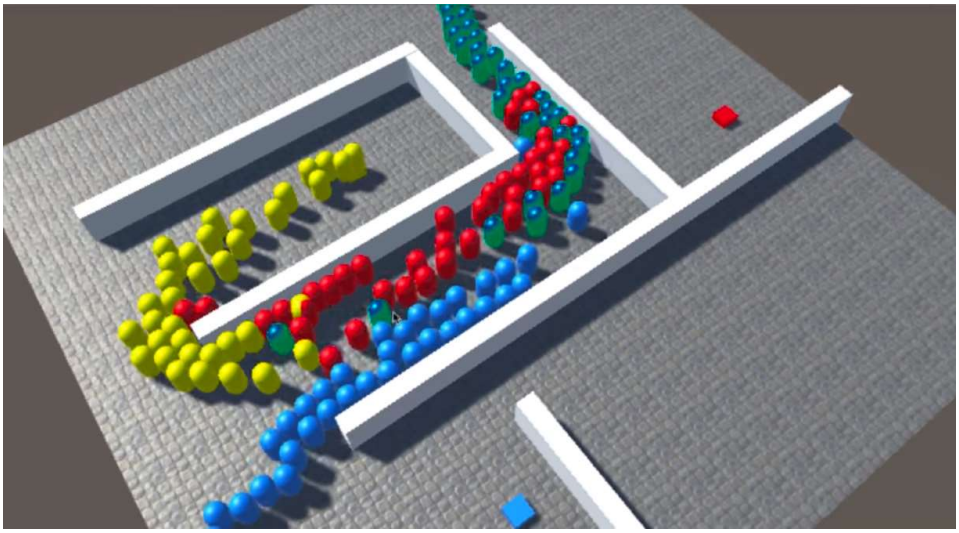

***Figure 7 Unity Crowd analysis and optimization with deep reinforcement learning with AgentML application of LOLA based on* https://arxiv.org/abs/1709.04326 Learning with opponent-learning awareness Jakob N. Foerster, Richard Y. Chen, Maruan Al-Shedivat, Shimon Whiteson, AieterAbbeel, Igor Mordatch**

AR Cloud

The ability to associate information in an accurate position and tracking will start a revolution comparable to the .com one.

Urban spatial intelligence

Understanding the importance of AR cloud applications in the public environment and finding the connection within urban planning will be crucial for the future development of the virtual space that will surround us. One way to supply the limitation of accuracy provided by GPS based solutions is to correlate a satellite image database with real-time streaming from a camera in the physical environment. This will allow a visual position system with great spatial intelligence given by a deep learning approach that allows us to have improved accuracy in order to identify the exact location.

Privacy

One of the most monumental concerns for AR cloud adoption is related to privacy. The main goal of designers isto provide shelter and better places for people, where they can feel safe. The designer will have the duty to manage and control the digital world around us. An example of this can be enablingthe saturation of advertisements reaching into our private virtual spaces (see problems related with Pokémon GO). There are solutions which have beenimplemented, such asblockchain, in the processto provide ownership to the digital space and what is visible to you as a user. Initiatives such as GDPR and other emerging groups, privacy will play an ongoing role in such experiences.

AEC INDUSTRY'S PROBLEMS.

After the analysis of the current position of AR technologies, it is important to have a common vision in order to set the future goals and milestones. The ability to create a digital twin in augmented reality on top of a building in construction will disrupt the entire ecosystem starting from the facility management to the early stage of the design. The digital twin is the virtual representation of a physical object or system across its life-cycle (design, construction, operation) using real-time operational data and other sources to enable understanding, learning, reasoning, and dynamically re-calibrating for improved decision making. More often, we are missing the aspects or comparisons of the product as-designed, as-built, and the communication between product stakeholders. There are engineers involved throughout an LCA, or "circular" workflow that requires a platform to align the definitions of

the virtual object.This process stands to bring operational/physical data into the digital twin, which allows the understanding of how a product is performing compared to its as-designed intent. All this analysis needs to be done before the beginning of construction in order to understand the future impact of the building in that specific area.
In order to transform this vision into reality we need to face the challenges that can be divided into 2 macro areas strongly related to each other:
1) Achieve a hyper-realistic experience.
2) Enable real-time collaboration.
In order to provide a realistic experience to the users, the object itself needs to be designed realistically enough to fool the human eye. It needs to simulate qualities reflection, refraction, real shadows, and ambient light that are partially solved. Another important property is that ofocclusion. This is one of the most difficult challenges of computer vision. The object needs to fit and be scaled around its own physical space and hidden behind real objects.

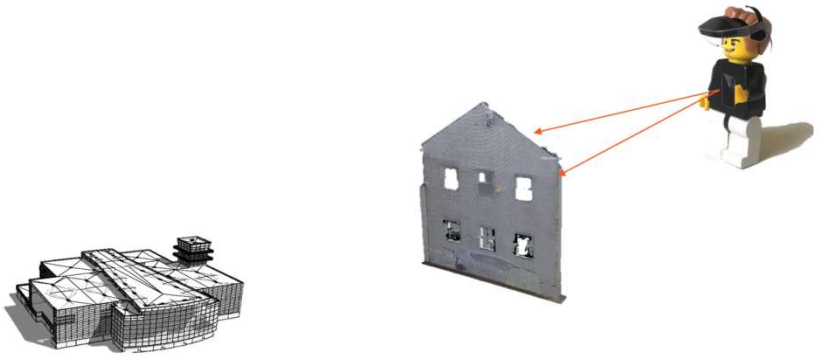

Figure 8 Occlusion, will allow to have a more realistic experience

This problem breaks the immersion in AR, the digital asset unable to integrate with space believably. This surrounding understanding can be solved thanks to the use of deep learning algorithms of computer vision, performing semantic understanding, classification and segmentation.

Conclusion.

Architecture is essentially human. Why move into an isolatedvirtual space when you can share your design with other people and collaborate with more users?

Figure 9 Collaboration with hardware and software agnostic approach

Acknowledgment

Thanks to all the people that helped in the development of this research.
Thanks to Torno Subito and all the people that allow this experience to happen, thanks to our Mars team: Tang Li Qun, Theodore Galanos, Daniele Santucci, Sheridan, Jay, Silvia Croce, Benjamin Taube, Ary Wicaksana, Stefano Schiavon. All the great mentors Claudio Benghi, David Oh, Metaverse aka Damon Hernandez, Paolo Borin, Carlo Zanchetta e Andrea Giordano.
Our HOK Team, Adam Chernick, Konrad Sobon, Hans Palacio, Rotimi Seriki, Mark Perry, Sean Quinn, Niraj, our principal Matt Staublin. But also all the help from our AEC community: Andrea Zani, Nathaniel Jones, Adam Rotman, Ian Keough, Chiara Rizzarda, Antonello di Nunzio, Claudio Vittori Antisari, Hannah Bianchini, IlkerKaradag,MarziaBolpagni, Valentina Sumini, Hannah Bianchini, JoekobFoerester and Federico Faggin. And the biggest thanks goes to Hannah Luxenberg.