

Bringing Reality Back to Virtual Reality - A Collaborative Tool for Multidisciplinary Teams

Sampsa Kohtala¹, Thomas Kaland², Live Jacobsen², Pasi Aalto³, Martin Steinert¹

¹*Department of Mechanical and Industrial Engineering, NTNU Trondheim
sampsakohtala@ntnu.no, martin.steinert@ntnu.no*

²*Department of Design, NTNU Trondheim
thomaskaland@gmail.com, livejac@gmail.com*

³*Department of Architecture and Technology, NTNU Trondheim
pasi.aalto@ntnu.no*

Abstract

The use of Virtual Reality (VR) supports larger strategic goals in the industry, including the geographically diverse working markets, large-scale projects and increased collaboration, both locally and globally. At the same time, VR is becoming increasingly popular as the technology becomes cheaper, more accessible, and better in terms of processing power and ease-of-use. The technology offers new possibilities for play, work and collaboration in various areas, from ordinary consumers to the industry. However, there are still many challenges before VR becomes a seamless experience for the user and can even cause restrictions in collaborative and multidisciplinary projects in some scenarios. One of the challenges with VR is the lack of a natural interface for interacting with the digital environment together with multiple users.

In this paper we propose a new system combining a physical interface and a virtual environment that can be accessed through VR. Our aim is making collaboration in multidisciplinary teams in large-scale projects more productive through the joint effects of immersion and presence through VR and the tactility and intuitiveness of a physical interface. Our proof-of-concept prototype is presented, consisting of a physical 2D map and 3D printed objects that are being tracked with object detection and computer vision, which are linked to a digital twin. The digital twin can be accessed in VR in real time. We also explore solutions that can improve upon the prototype, such as producing an accurate physical 3D model of a real landscape to increase the resolution and authenticity of the physical model.

A pilot experiment with 14 participants using our prototype has been conducted. Results of the study shows potential benefits of the system, such as the simplicity of a physical model promoting engagement, and the in-depth exploratory capabilities of the VR application enabling more solutions to be found. We also present limitations with potential solutions.

Keywords: *Collaboration, virtual reality (VR), multi-/cross-/trans-disciplinary approaches, prototyping*

1 Introduction

The Norwegian government emphasizes the role of well-functioning regions and rural districts as the basis for value creation, as well as the necessity of providing them with forward thinking employment opportunities (Perspektivmeldingen 2017b, p. 40). Digitalization is key in these new employment opportunities. In consumer behavior, where transportation is now the second largest cost for consumers, more consumption is moving to digital platforms, and smart products for homes and health are becoming more prevalent (Framtidas forbruker - grønn, smart og digital 2019b, pp. 18-19). The Norwegian government is focusing on the potential of using Virtual Reality (VR) technology as well as satellite data globally to reduce cost in international collaboration, especially in developing nations (Digital transformasjon og utviklingspolitikken, 2019a, pp. 33-34). Reduced transportation needs and specifically the «polluter must pay»-principle connect emissions to company cost (Klimastrategi 2030, 2017a, p. 41), something also reflected in how the government is organizing its own public services (Digital agenda for Norge, 2016, Chapter 6). In summary, the use of VR supports larger strategic goals in Norway, including bettering working conditions in geographically diverse working markets, reduction of transport related CO2 emissions and increased collaboration, both locally and globally.

At the same time, VR is becoming increasingly popular as the technology becomes cheaper, more accessible and better (Craig & Sherman, 2003; Wraga, Creem-Regehr, & Proffitt, 2004). The development of VR applications is also becoming easier and more available through cross-platform game engines and editors, such as Unity and Unreal Engine, providing tools and tutorials to facilitate the development process. The technology offers new possibilities for play, work and collaboration in various areas, from ordinary consumers to the industry. While common consumers mostly use VR for entertainment, the industry has found several use-cases including training, visualization, and education. It has for example been shown that moving through a virtual environment can give enough knowledge to the subject to be used as a training environment when the real-life environment is unavailable, such as training for an emergency situation (Witmer, Bailey, Knerr, & Parsons, 1996).

However, there are still many challenges before experiences in VR are becoming more seamless (Bastug, Bennis, Médard, & Debbah, 2017), and it has been shown that the technology can even cause a reduced level of learning in students while increasing the cognitive load (Makransky, Terkildsen, & Mayer, 2019). The threshold for starting to use and benefit from VR can be high, especially for industries that are not yet familiar with the technology. There is a certain level of separation from the present surroundings when a user is immersed in the VR environment, which can make the use of this technology difficult in collocated collaboration. Awareness of each other's position and emotions can be a challenge, although different viewpoints and capabilities can enhance collaboration in virtual tasks (Spante, Axelsson, & Schroeder, 2006). Alatta and Freewan (2017) found several benefits of using VR in the early stage of the design process in architecture and design but mentions that virtual environments are rarely used for collaboration due to the lack of support of such activities. Attempts have been made to develop physical interfaces or haptic feedback in VR to increase immersion and collaborative capabilities. He, Zhu, and Perlin (2017) proposed a method for enabling collaborators to share the same physical feedback using actuated robots. Piumsomboon, Lee, Lee, and Billinghamurst (2017) combined both VR and AR (Augmented Reality) in remote collaboration to enhance the experience and awareness between users by simulating a collocated environment and utilizing natural communication cues to further improve the interaction.

In this paper we investigate how combining VR with a physical interface can improve collaboration. We believe a system comprising a separate physical interface connected to a virtual environment can exploit the benefits from both worlds. Our overarching goal is making collaboration in multidisciplinary teams in large-scale projects more productive through the joint effects of immersion and presence through VR and the tactility and intuitiveness of a physical interface. We hypothesize that a physical interface will promote ease-of-use, thereby lowering the threshold for engagement by different disciplines in co-located collaboration. Meanwhile a digital twin of the physical interface provides a means for inspection, observation, and conceptualization in real scale through VR, which can also invite remote collaborators. Our focus is on the industrial park and real-estate planning industry, although different application areas are being investigated, such as planning larger operations, workflows, logistics, risk management and marketing strategies. A proof of concept (POC) prototype has been developed, consisting of a physical 2D map and 3D printed objects that are being tracked with object detection and computer vision (CV), which are then linked to a digital twin. The digital twin can be accessed in VR in real time. The system is developed with the intention of making it scalable and simple to implement, so that separate projects can use the same setup with different components. We also discuss how different parts of the system can be improved, such as how the realism of the map can be increased by acquiring heightmap data to create an accurate physical and digital 3D model of the area. A pilot experiment has been conducted to investigate how users interact with the proposed system, and to find its requirements and limitations.

2 System Overview

There are many parts to the system, and considerations must be made for each subsystem in order to make it scalable and simple to implement for different use-cases. For example, the same system should be applicable for planning a new a shipyard, manufacturing plant or residential area, each of which will require different landscapes and constructions (objects) in different scales. In this section we go through the subsystems of the latest POC prototype and how they can be solved or improved in the future, including some preliminary testing to determine the feasibility of the different solutions. We begin with a general overview of the POC prototype.

2.1 POC Prototype

2.1.1 Physical Setup

The physical setup consists of a scaled version of a landscape, fitted to a small conference table. Currently we have used a printed 2D map with contour lines constructed from the terrain of the landscape, but we have also looked into possibilities of generating a true 3D representation of the landscape in Section 2.2. Corners of the map are marked with QR codes to define the reference plane between the physical and digital model. Buildings (objects) to be placed are 3D printed in the same scale as the map. A camera is placed above the table at a slight angle to fit the whole map within the capturing frame. The camera captures a live feed of the table, where each frame is processed for locating the map and tracking different objects relative to the map. The coordinates for each tracked object are constantly updated in a text file on a local server.

2.1.2 Virtual Environment

The virtual environment is made with Unity using the SteamVR plugin, which supports HTC Vive and other VR hardware. The digitally reconstructed 3D landscape corresponds to the physical 2D map and contains the same objects. The digital objects can have more details than their 3D printed counterparts, such as texture and color, but are still usually a simplified version

of the original 3D model for better real-time performance. The x-y coordinates for each object are read from the local server so they can be placed on the corresponding location on the digital landscape, where the z-location (height) is determined by the ground level.

2.1.3 Intended Usage and Interaction

The physical setup provides a holistic overview of the planned area. Users can interact with the physical setup by moving objects (e.g. buildings) and try different configurations. At any point in time, a user can view the same configuration in more details and in real scale through the VR application, where they can teleport to any location within the virtual area using one of the hand-controllers. While the physical setup gives an overview and a simple way of interacting with the model, the VR application provides a channel for closer examination and immersion in real scale. We hypothesize that this simple and intuitive approach will lower the threshold of engagement by different collaborators and disciplines, thereby increasing both the overall efficiency of the collaboration and the holistic understanding of the project, while additionally providing more detail-oriented inspection through the immersive digital twin.

2.2 Physical Scale Model

A coherence between the physical landscape model and virtual landscape is necessary. Introduction of translation errors that result in differences between the physical and virtual landscape models could potentially lead to insecurity as to which model is correct in the use scenario. To ensure a viable pathway to a physical model, a 1:1000 scale model suitable for the proposed team interaction was produced.

The scale model uses data from the Norwegian Detailed Heightmodel (Nasjonal Detaljert Høydemodell, NDH) from the Norwegian Map Authority. A DTM1 tile of 1538x1538 pixels is downloaded in GeoTIFF format, corresponding to 1538x1538 meters in-situ and 1538x1538 mm in the corresponding CNC-machined physical landscape model in 1:1000 scale. The file is processed in QGIS to render it to a TIFF image that can be used as a heightfield in several 3D applications, including Unity and Rhinoceros 3D, allowing further processing. In addition, the Z height of the model is recorded.

To verify a future capacity to CNC machined 3D models, the image was imported to Rhinoceros 3D, using a heightfield object of corresponding 1538x1538 mm. The resulting surface was used to calculate parallel finishing type CNC paths for a 3-axis router using RhinoCAM with typical settings after material choice and milling strategy. A first prototype was machined to verify that a valid workflow to a 3D machined landscape exists.

2.3 Object Tracking

2.3.1 Overview of Tracking Methods and Technologies

There are several methods that can be used to track physical objects in 3D-space. For a VR-setup to function, the controllers and the VR-headset must be accurately tracked in real time. Consumer VR gear has managed to do this reliably and with high precision for several years. These methods can be repurposed for tracking other objects as well. In order to find a tracking solution to fit our needs we have reviewed existing tracking methods. A summary of the findings is presented in Table 1.

Two of the biggest actors in consumer VR gear today is Oculus and Vive, both of which produce several different VR headsets. Across these two companies and their VR headsets there are three major types of tracking used; outside-in, inside-out and standalone. Both outside-in and inside-out systems require manual setup of base stations, where outside-in refers to the cameras (sensors) being fixed at the base stations with transmitters (usually non-visible light) on the moving objects, with inside-out being the opposite. Standalone systems (Oculus Quest) have the camera on the VR headset itself and does not require a setup of external hardware. Outside-in systems (Oculus Rift) typically use stationary infrared cameras with infrared emitting diodes on the moving objects. The unique pattern of infrared light on each object is then captured by the cameras to calculate its location and orientation. A typical inside-out system (Lighthouse used by Vive) uses base stations to emit patterns of non-visible light, which are received by photosensors on the moving objects to calculate location. Honnet and Lopes (2019) has developed a miniature version of the HTC Vive tracker, called HiveTracker, that can potentially be used to track smaller objects without having to extensively alter their original shape.

Other tracking methods that utilize already present signals, such as WiFi-based positioning (Kotaru, Joshi, Bharadia, & Katti, 2015), have come a long way. Reaching decimeter level accuracy, this type of indoor localization has many applications such as health monitoring, indoor navigation and warehouse automation. However, this technology is still considered too inaccurate for most VR applications.

Camera (RGB) based tracking has gained popularity with the improvement of hardware, and the development of CV and machine learning algorithms, making it possible to achieve real-time performance within many applications. Methods range from simple edge detection algorithms (e.g. Canny edge detection) that can be used to for example derive the location and orientation of predefined markers, to complicated convolutional neural networks that require a lot of data and the consequent training of models that can then be used to locate and classify different objects in images or video frames.

Table 1. Pros and cons of different tracking methods and technologies.

Tracking method	Pros	Cons
Infrared light and camera (Oculus)	Low-cost hardware High update frequency Precise	Unique light pattern for each object Advanced recognition algorithms
Lighthouse (Vive sensors)	Precise Designed to be scalable	Photosensors in each object Requires at least two base stations
IMU (orientation and acceleration)	Low-cost hardware High update frequency On-board processing	Drift (accumulated error) Power consumption
Positioning from signal strength	Low-cost hardware Ubiquitous technology	Large margin of error Require at least two receivers Signals easily obstructed
Camera (object detection)	Low-cost hardware Ubiquitous technology No hardware inside objects	Sensitive to environmental variation (light, material, resolution etc.) Need data and training 3D tracking difficult

Although there are many pros and cons of each method mentioned, as shown in Table 1, our aim has been to build a system that is scalable, simple to assemble, and adaptable to different

projects. We therefore mostly weighted the ability to track 3D printed objects without having to insert any extra hardware. The fact that we track objects on a surface that is basically flat makes the use of a camera more feasible, in addition to making the setup simple to assemble. The main reason for going forward with object detection, as opposed to for example attaching QR-codes on the models and using conventional CV algorithms, is to preserve the realism of the physical model. Ideally, the 3D printed buildings would have even higher levels of detail, such as true colors, which would increase the realism as well as increasing their uniqueness, which would also improve the object detection accuracy. An object detection model can also be expanded when introducing new objects to be tracked, with no extra cost except for the training process. Furthermore, other functionality can be incorporated to the system by for example training the model to also recognize hand gestures.

2.3.2 *Object Detection Concept*

To track objects using only a camera, we have used the open source neural network framework by Redmon and Farhadi (2018), which is made for both utilizing and training object detection models. Their method is known for its speed, reaching real-time performance with GPU-accelerated computing, and match other state-of-the-art object detection systems in terms of accuracy (Zhao, Zheng, Xu, & Wu, 2019). Through transfer learning we have repurposed the darknet53 model (pre-trained convolutional weights) on our own training data to speed up training and increase accuracy.

To verify the feasibility of this method, we have trained several classifiers with different 3D printed buildings. Training a classifier has been done by first capturing a video of the scene where we manually move the objects around randomly, then extract relevant frames, usually ending up with around 90 training images and 30 test images for each classifier. Each of the images are then manually labeled, i.e. specifying the bounding boxes for each object. For most of the classifiers the loss function has converged after only a few hours of training on a local machine, and resulting in an accuracy of around 90% mAP (mean average precision) with the IoU (intersection over union) threshold at 0.5, or approximately 85% mAP at IoU 0.75. The relatively high accuracy can be explained from the test set being taken with the same settings as the training data, such as image quality, distance and angle of the camera to the table, and room lighting, which can cause overfitting. By applying the classifiers in similar settings as the training data we can to some extent exploit overfitting during the prototyping and testing phase, effectively reducing resources and training time since we don't need to generalize the data for the model to work properly.

To get meaningful data (input for the digital scene) from the object detection output, we must convert the detected bounding box coordinates in the pixel space into x and y coordinates relative to the 2D plane of interest. This is accomplished by first detecting the plane through three QR markers using Open Source Computer Vision Library (OpenCV), as shown in Figure 1, which is done automatically to make the camera positioning simple and trivial. The four corner points are then mapped to a square with side lengths of one to get the perspective transformation matrix M , which is calculated using Gaussian elimination through a function in OpenCV. We then multiply M with the pixel coordinates of every detected object center to get the Homogeneous coordinates relative to the plane, which are converted to Cartesian coordinates. Finally, the coordinates with corresponding object IDs are sent to a local server which the Unity program has access to.

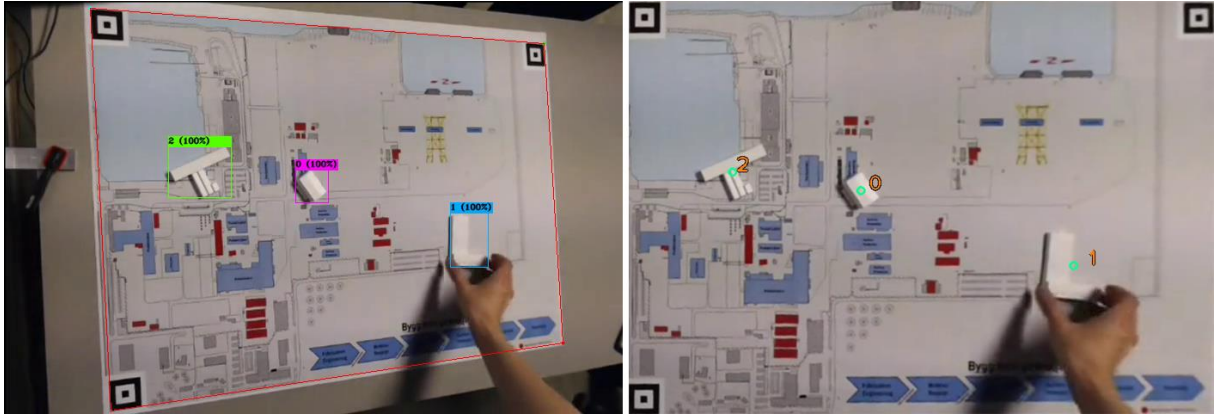


Figure 1. Detecting three different 3D printed buildings and the QR-defined plane from a live video stream on the left, with the corresponding transformed area with object centers on the right.

2.4 Virtual Environment and VR

The virtual environment is developed using Unity, a cross-platform game engine supporting the creation of games, simulations and other experiences that can be deployed on devices such as desktops, smartphones and multiple VR headsets. Unity projects are mainly developed using a scripting API in C#, with an extensive asset store supporting all aspects of the development including 3D models, environments and software development kits. This makes the development process efficient and accessible, which has allowed us to rapidly test our concept with both the HTC Vive (used in the pilot experiment) and Oculus Quest headsets.

The Vive headset uses base stations for tracking and is connected by cables to a computer for processing. While this can enable a virtual environment with higher resolution, given a powerful computer, the cables and base stations adds a constraint to the user. With Oculus Quest being a standalone system and doing all the processing within the headset itself, it is not constrained by the cables and is thus the preferred platform.

3 Pilot Experiment

In order to test the concept of combining immersive virtual reality with a physical interface in collaboration tasks, we conducted a simple experiment to see how people would interact with the system and with each other. At this point in time, the technology to track physical objects in real time was not yet implemented. Therefore, the experiment consisted of a “wizard of oz” test, where we would manually simulate the functionality of the system by having a third person manually move the digital models. When an object was moved on the physical model by a participant, the third person (experimenter) would manually move the corresponding object in VR. For the test participants the system appeared to be fully automatic.

The scene contained a simple white terrain with contour lines in similar positions to the physical map, without details such as grass or plants. Other features, such as the seafront, road and lighthouse, was placed to see if the participants would use them as reference points for orientation. The physical and digital setups are shown in Figure 2.

The experiment consisted of 14 participants divided into pairs, and the test was conducted seven times. In each test, the participants were given the task to place houses in a fictional residential area, as shown in Figure 2. The participants were given three goal criteria; (1) the houses should be placed in such a way that the fictional residents would get a great view of the nature, (2) the

houses should be secluded from view from other houses and (3) the houses should be placed in a relatively flat terrain. The participants chose themselves how much they would consider each of the criteria and their task was completed with little input and feedback from the test conductors. While observing the tests, we focused on how the system affected the participants in terms of presence, i.e. the ability to cooperate and their use of language, and immersion with focus on their spatial awareness and the general ease-of-use.

In two of the tests, we replaced the physical model of the houses and the residential area with a digital model on a regular computer screen. This was done to investigate if there are differences between using a physical scale model versus using a digital scale model in terms of how participants collaborate and interact with each other.

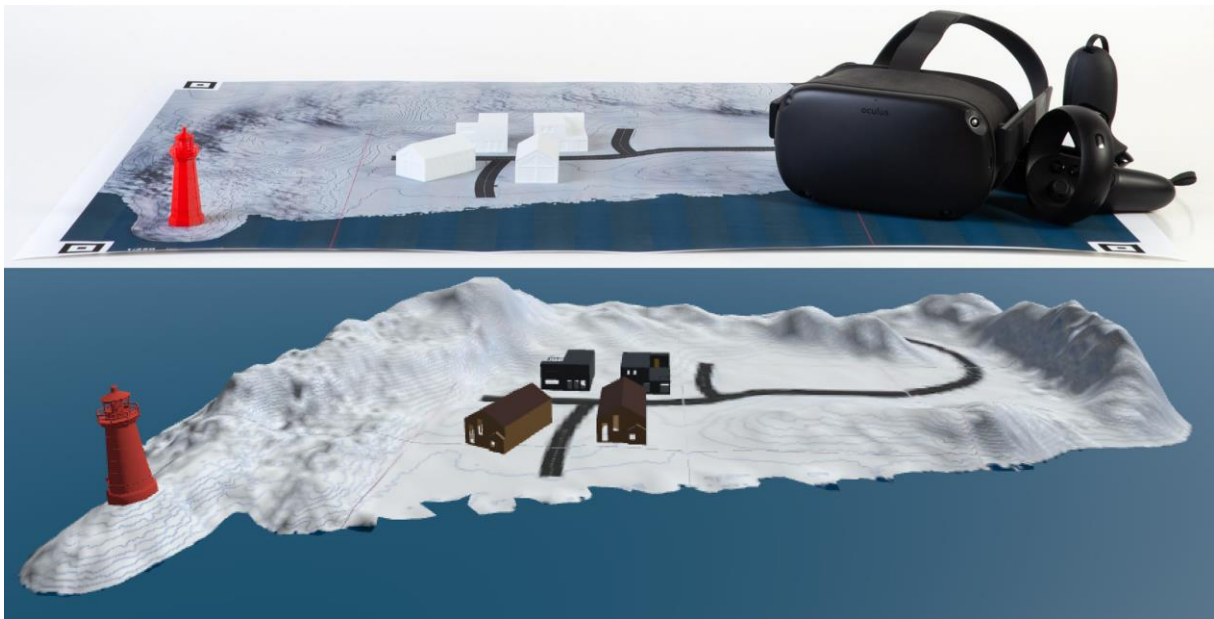


Figure 2. Experiment scene with the physical part shown at the top and the digital VR-scene on the bottom.

4 Results

During the pilot experiment it became obvious that communication between the non-VR user and the VR user was the biggest issue, especially when it came to communicating positions. Here, many participants in the physical space started to point and gesture towards the physical model, even though their partner were inside VR and unable to see what their partner was pointing at. Likewise, the participants inside VR would often struggle to communicate their position, which led to them to pointing and gesturing out in thin air. This habit of pointing in VR only worked when the person outside VR focused on the screen showing the point of view the person in VR, but they were often fixated on the physical scale model instead.

A word search was done on the transcripts from the pilot experiment, with the occurrence of positional words shown in Figure 3. The words “here”, “there” and “that” occurred significantly more times than the other positional words. These three words are all subjective positional words and describes position relative to the person using them. This creates communication problems when the two persons are not in the same relative environment, as is the case if one person is in VR and the other is not.

The participants seemed to develop more empathy for the residents of the houses when they started to actively use VR. They started to imagine the everyday life of the residents and imagined use cases for the areas between the houses. Some participants also expressed having a better understanding of the terrain when immersed in the VR scene, leading to a more feasible solution.

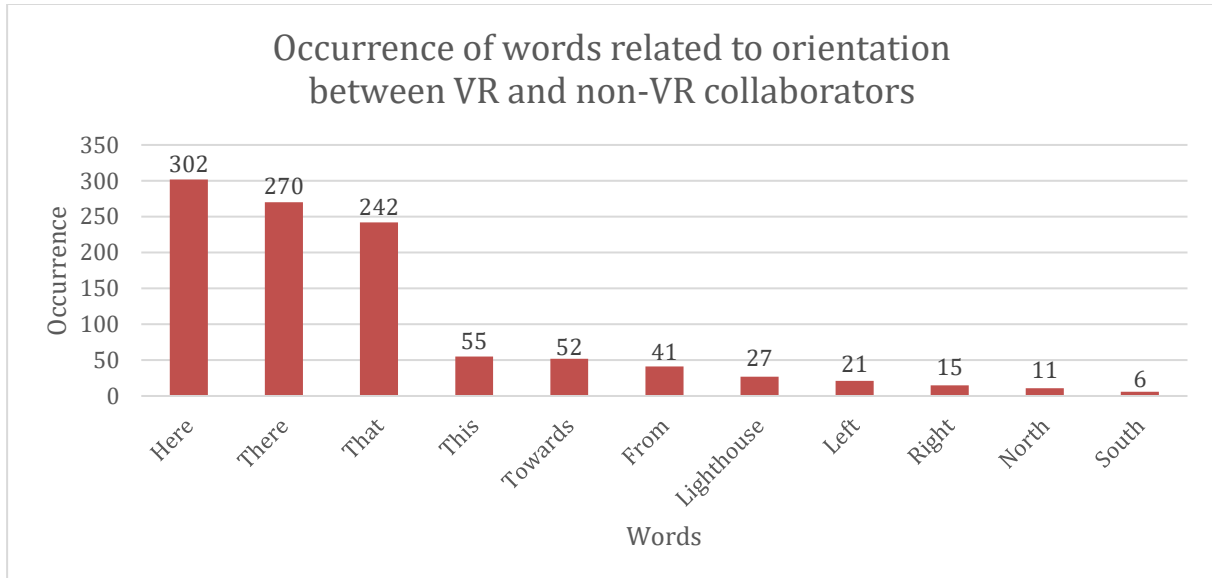


Figure 3. Occurrence of words related to communicating positions between a person in VR and a person outside, both working with the same model in different mediums and scales.

The separation of roles, i.e. when each participant focused only on using either VR or moving houses, appeared stronger in the tests using a screen compared to the physical interface. The digital interface groups also spent more time placing houses in the beginning, compared to a more direct and playful approach from the physical interface groups.

5 Discussion and Future Application

5.1 System Usage and Interaction

The pilot experiment indicates that the system has potential benefits in collaborative tasks involving planning of large-scale projects, although the study is limited by the low number of participants. One of the key observations is the perceived ease-of-use by participants, being quickly able to start testing their ideas on the physical setup. Furthermore, the immersion of VR resulted in an increase in empathy, enabling the users to further improve their solution by seeing and understanding the task from different perspectives. This result is analogous to what Alatta and Freewan (2017) found in their study of using VR in architecture and design for human experience, where they observed many benefits including an increase in spatial awareness and improvement of critical thinking. From Figure 3 we can also argue that the participants felt a strong presence within the space, since they could relate to the physical and digital scenes and consequently use their natural language instead of referring to specific terms or abstract interpretations.

One of the limitations discovered was the separation, in terms of each other’s spatial awareness, between a user inside the virtual world wearing the VR headset and the user who focused on the physical model. A future requirement of the system should therefore account for better communication between the real and virtual world. There are many potential solutions to this

problem, for example projecting the location and viewing direction of the user immersed in VR on the physical map and recognizing hand gestures from the physical scene to approximate and project the area that is pointed at to the virtual scene.

To answer our hypothesis, a more comprehensive experimental design or real-world case study including more participants and disciplines should be done in the future.

5.2 Scalability

As mentioned in the early sections, it is important that the system is simple to implement and adaptable to both small changes and completely different scenarios. Our solution attempts to solve this by keeping the general setup simple using a scaled map, 3D printed objects and camera-based tracking, all of which are achieved with technologies considered ubiquitous today.

We have also explored ways of increasing the resolution of the physical model by using available heightmap data from the real world, and how the data can be processed to end up with a realistic CNC-machined physical model of the scene.

There are however still challenges to make it a truly scalable and simple system. Object detection methods can detect many different objects, where new objects to the scene can be added by training the model further with new data. Acquiring and preparing training data can be a cumbersome and time-consuming task. We are currently looking into possibilities for generating synthetic training data, since we can assume the digital object already exists, which we can leverage to make the entire data acquisition and training process completely automatic. We are currently unable to extract the orientation of the tracked objects. Since this is a regression problem, making it more difficult compared to simple classification, the ability to generate large amounts of training data including orientation can help to develop a method for extracting the orientation of each detected object as well.

Since Unity applications are built with scripting languages, there is also a potential to make some parts of the program general purpose or fully automated, such as loading 3D models and different landscapes. Connecting the corresponding digital and physical parts can potentially be solved with a database management system for improved overview and control. The ability to incorporate other relevant parts of an organization into the database, such as engineers working with computer aided design, could potentially make the approach seamless and more useful. There is also a need to explore other functions that the system can support, such as planning operations, workflows, maintenance, and marketing.

6 Conclusion

A new collaborative tool concept has been demonstrated through a proof-of-concept prototype, aimed at advancing collaboration in large-scale projects with multidisciplinary teams. The system consists of a physical scale model of a landscape with 3D printed objects that are tracked using object detection and computer vision, and a digital twin of the physical model that can be accessed through virtual reality in real time and real scale. The physical model serves as a simple interface to engage collaborators, while the virtual environment enables a more in-depth experience of the scene.

Our pilot experiment suggests that the proposed system is simple to use, thereby lowering the threshold for engagement by its users, while providing a channel for more in-depth exploration

that can reveal otherwise unseen problems and solutions. Communication between a VR and non-VR user of the system was discovered as a limiting factor, causing frustration in each other's spatial awareness. A proposed solution is projecting the VR-users position and orientation on the physical model and recognizing hand gestures to reflect the non-VR persons intentions in the digital scene.

A more comprehensive real-world case study is planned in the future to gain better insights of the potential of this system. More requirement elicitation is needed to improve the system, such as finding the required accuracy of object tracking. We hope others will find our concept intriguing and help us bring reality back to virtual reality.

References

- Alatta, R. A., & Freewan, A. (2017). INVESTIGATING THE EFFECT OF EMPLOYING IMMERSIVE VIRTUAL ENVIRONMENT ON ENHANCING SPATIAL PERCEPTION WITHIN DESIGN PROCESS. *ArchNet-IJAR: International Journal of Architectural Research*, 11(2), 219.
- Bastug, E., Bennis, M., Médard, M., & Debbah, M. (2017). Toward interconnected virtual reality: Opportunities, challenges, and enablers. *IEEE Communications Magazine*, 55(6), 110-117.
- Craig, A. B., & Sherman, W. R. (2003). Understanding Virtual Reality: Interface, Application, and Design. In: San Francisco, CA: Morgan Kaufmann Publishers.
- He, Z., Zhu, F., & Perlin, K. (2017). *Physhare: Sharing physical interaction in virtual reality*. Paper presented at the Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology.
- Honnet, C., & Lopes, G. (2019). *HiveTracker: 3D positioning for ubiquitous embedded systems*. Paper presented at the Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers.
- Kotaru, M., Joshi, K., Bharadia, D., & Katti, S. (2015). *Spotfi: Decimeter level localization using wifi*. Paper presented at the Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication.
- Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, 60, 225-236.
- Piumsombon, T., Lee, Y., Lee, G., & Billinghurst, M. (2017). CoVAR: a collaborative virtual and augmented reality system for remote collaboration. In *SIGGRAPH Asia 2017 Emerging Technologies* (pp. 1-2).
- Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.
- Spante, M., Axelsson, A.-S., & Schroeder, R. (2006). The good inequality: Supporting group-work in shared virtual environments. In *Avatars at work and play* (pp. 151-166): Springer.
- The Norwegian Government. (2016). *Digital agenda for Norge*.
- The Norwegian Government. (2017a). *Klimastrategi 2030*.
- The Norwegian Government. (2017b). *Perspektivmeldingen*.
- The Norwegian Government. (2019a). *Digital transformasjon og utviklingspolitikken*.
- The Norwegian Government. (2019b). *Framtidas forbruker - grønn, smart og digital*.

- Witmer, B. G., Bailey, J. H., Knerr, B. W., & Parsons, K. C. (1996). Virtual spaces and real world places: transfer of route knowledge. *International journal of human-computer studies*, 45(4), 413-428.
- Wraga, M., Creem-Regehr, S. H., & Proffitt, D. R. (2004). Spatial updating of virtual displays. *Memory & cognition*, 32(3), 399-415.
- Zhao, Z.-Q., Zheng, P., Xu, S.-t., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE transactions on neural networks and learning systems*, 30(11), 3212-3232.