

New technics of selection, manipulation and placement of objects for editing in Virtual Reality

Dmytro Kadar

Supervisor: Ing. Uršula Žáková

Open Informatics - Computer Games and Graphics

January 2023

Contents

1	Introduction	3
2	Current object manipulation, selection and placement methods in VR	3
3	Proposal	6
4	Implementation	6
5	Testing	8
6	Conclusion	9
7	Links and Sources	9

1 Introduction

As Virtual Reality(VR in later parts of the document) popularity growth, there is more and more developers and designers working with that type of apps. But when it comes to making scenes in VR, we come to problems. We can make scene on a regular computer monitor, but it is very difficult to see it from players perspective and developers must then test it on VR headset. In my project i want to try to solve this problem - what if we already done scene in VR?

In this work i will be focusing on observing current ways of selection, manipulation and placement of objects for editing in VR and implementing my own vision of object selection problem solution.

2 Current object manipulation, selection and placement methods in VR

Effective 3D objects manipulation and scene creation is an old problem. For example "3D Palette: A Virtual Reality Content Creation Tool", a 1997 paper about interface for creating 3D virtual scenes, tried to solve that problem with graphics tablet and voice recognition technology: "3D Palette consists of a Wacom tablet and digitizing pen, both of which are tracked in space using Polhemus Fastrak electromagnetic sensors, a head-mounted microphone and a pair of CrystalEyes stereoscopic shutter glasses. The user holds the tablet in one hand while interacting with the scene using 2D and 3D pen gestures and voice commands"

The closest to my work real use that I found of doing scene in VR is ShapesXR - tool for making 3D models, sounds, scenes in VR, which than can be exported to Unity game engine. There is even a possibility to work as a team on one project, that can help produce content faster. Also worth mentioning AR home designing apps like IKEA Place and PLNAR where you also can manage scenes(in this case your own room or house).

Real world objects are commonly used for 3D object creation as references, but there are technologies to automate 2D photos to 3D object conversion. "We inject the inductive bias of compositionality of the 3D world into BlockGAN in two ways. (1) The generator is designed to first generate 3D features for each object independently, before transforming and combining them into unified scene features, in which objects interact. (2) Unlike other methods that use 2D image patches or layers to represent objects, BlockGAN directly learns from unlabelled images how to generate objects as 3D features. This allows our model to disentangle the scene into separate 3D objects and allows the generator to reason over 3D space, enabling object pose manipulation and appearance

interaction between objects. BlockGAN, therefore, learns to both generate and render the scene features into images that can fool the discriminator.” - Paper about BlockGAN - ” Learning 3D Object-aware Scene Representations from Unlabelled Images”. Also human speech is a great source for inspiration. ”3D SCENE CREATION USING STORY-BASED DESCRIPTIONS” is a paper about ”an approach that enables nonprofessionals to generate a 3D virtual environment (3DVE) by using simplified natural language input.” In my opinion, this project also can help people with vision problems to participate in creation of 3D worlds and express themselves more - ”Through a mixture of user trials and comparison with other similar systems, it was shown that the method presented in this research works successfully in practice through the experiments. While the main usability evaluation was only assessed by a small group of participants, the evaluation showed that the prototype 3DSV system developed is functionally sound and the greatest advantage of the system lies in its ability to evolve tools that can effectively be used by nonprofessional to create a real time interactive virtual environment.”.

VR reality is very close to real world by visuals and sounds, but some techniques are not perfect. For example in a paper ”A visually realistic grasping system for object manipulation and interaction in virtual reality environments” proposed algorithm to make visual grasping not based on premade animations: ”The user is embodied in the simulated world as a virtual human or robot agent, to freely move and interact with the virtual objects using the handheld controllers (e.g. Oculus Touch motion controllers). Unlike most existing VR approaches, based exclusively on interaction with predefined objects through animations, our grasping system is capable of interacting and/or manipulating objects regardless of their geometry. This is because the virtual hand is automatically fitted to the object shape without the need of a different grip animation for each object the user is interacting with. In this way, grasp synthesis is being simplified, since we adapt the virtual hand to multiple object shapes starting from a 6D hand pose which is user predefined.”

Visual programming is a technique, where programming code is replaced with visual nodes. One of the most popular realisations are Scratch(made for beginner programmers who want to understand programming principles) and Unreal Engine Blueprints Visual Scripting(game development, focused more on game designers with little programming experience). Also, exists ”EntangleVR: A Visual Programming Interface for Virtual Reality Interactive Scene Generation”. Paper describes visual programming system based on quantum computing - ”Quantum computing offers many interesting concepts that can be applied to creative computing like interactive scene creation for virtual reality (VR) or game design and other forms of experiential media. Entanglement, a phenomenon that is at the heart of disparity between classical and quantum physics, describes a correlated relationship between the measurement of multiple spatially separated particles [2, 27]. As a unique feature of quantum computing,

it transforms individual object states into non-separable shared states, allowing a single interaction on one object's state to affect all the entangled object states as a whole [22].". Also, technology will be expanded in the future - "We will also expand our system by adding entanglement at different scene creation scales (e.g., procedural scene generation, multi-scene entanglement)."

If we have purposes not to use hand tracking and manipulation while in VR or need additional input method, we can use gaze-supported manipulations. It is a popular technology, that not as popular in VR. However, there are a few reserches and implementations of that. For example: "Gaze-Supported 3D Object Manipulation in Virtual Reality" - "In general, while gaze offers fast and natural pointing, itsuffersfrom the lack of precision and the difculty of confrming a selection. To overcome these challenges, many techniques combine gaze with an additional modality, such as the principle of "gaze select, hands manipulate" [9, 39, 56, 66]. For example, Pfeufer et al. proposed Gaze-touch [39], which enabled usersto control gazelected targets indirectly using multi-touch gestures on interactive surfaces. Another example is the method proposed by Turner et al. [62], which casts the object being looked at by the user to the touch/cursor position to allow further manipulation. In contrast, other approaches [48, 55, 60–63, 67] for content-transfer between diferent displays, have embedded gaze movement into the translation process. These prototypes typically require the use of a hand trigger to "attach" the object to the gaze direction and then release the hand trigger to "drop"it. In a follow up research, Turner et al.[59] pushed this concept further by developing techniques that maintain concurrent rotation and scaling operations when performing translation tasks using gaze and touch.". New Meta's headset - Meta Quest Pro also has abilities to track gaze direction, so this technology might be growing in popularity with game developers and users.

We can use virtual reality to extend our real life. VR can give more space for creativity. In "VRSketchIn: Exploring the Design Space of Pen and Tablet Interaction for 3D Sketching in Virtual Reality" describes a way of creating art in VR using pen and tablet. "Immersive environments can provide a full-body painting experience, which is perceived as very pleasant by users [29]. 3D Palette by Billinghamurst et al. [9] is a system combining pen and tablet with VR to create and draw primitives on a tablet interface and visualize them in 3D space, using widgets for parameter adjustment. They use pen, tablet, and VR as well, but did not focus on sketching or the definition of a design space. Deering [14] investigated pen-based free-floating mid-air VR drawing environments in 1995 but provided no true scale sket-ching at room size. Older work, in general, was implemented with hardware that is quite limited regarding the field of view and display quality compared to current head-mounted displays (HMDs) which can provide a better base for an evaluation."

As for procedural generation, it is very helpful technique that can create many models, big terrains, etc. from small amount of data. It is a little bit difficult to implement algorithm, but it is relatively easy to work with and understand implemented solution. There are even proposal of procedurally generated stories - "The procedural generation of stories has many advantages [142–144]. First of all, creating stories is expensive; many artists, content designers, programmers and audio engineers are usually involved. This effort can be reduced considerably if PCG for stories is employed [145]. Second, when new story variations are created at runtime, this can keep players motivated when replaying the game [144,146]. Third, it becomes possible to generate stories that automatically adapt to the players' skills and preferences [147]".

3 Proposal

I am more interested in object selection techniques, so, in my current work, I will be focused more on the selection. For implementation i will also implement object placement techniques.

As i observed, in my opinion there are those ways of making usable and comfort environment for creating scenes in VR:

- pattern recognition, choosing object by drawing shapes
- voice recognition, choosing object by saying its name
- selecting object from list
- their combinations

My work will be focused on object selection using pattern recognition technology. User will be drawing shapes with his controller motion capture, which then will be translated to text using AI(Artificial Intelligence) recognition algorithms. That text will represent object name, with which we can then add that object to our scene.

For now I will be implementing only Pattern Recognition approach. I was inspired by article on medium.com "Neural Nets + VR = Magic! ".

4 Implementation

I have started my work with creation of pattern recognition algorithm. In Unity Asset store exists extension called "PDollar Point-Cloud Gesture Recognizer" that gathers points from input gesture and, with help of cloud technology, returns guessed gesture name(if guessed) and score(zero if not guessed). The gesture is defined as one stroke, not limited by the time(but, every dot, from which gesture is consisted will disappear after 5 seconds). Also, i used video from Valem's channel called "How to Detect a Movement in VR - Unity VR

Tutorial” parts one and two as reference of PDollar library usage. Here are some examples of gestures sketches and in app representations:

Figure 1: Fabric Sketch



Figure 2: Stool Sketch



Figure 3: Chair Sketch

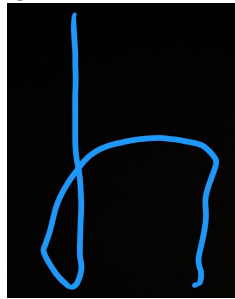


Figure 4: Chair Gesture

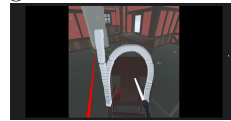


Figure 5: Cabinet Sketch

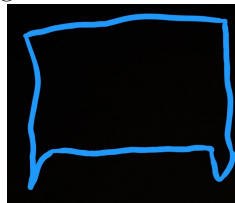
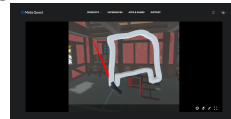


Figure 6: Cabinet Gesture



Movement and object manipulations are done with XR Interaction Toolkit(built in Unity Game Engine package). Test scenes for testing are done with POLYGON - Fantasy Kingdom Pack from Synty Store.

Application was designed with Oculus Quest(first model) in mind. It was used for deployment and testing. Final project is working as standalone application for Oculus Quest, that saves placed game objects in separate save file.

Controls:

- Right Controller:

- Trigger - draw gesture
- Grip button - grab object
- Controll Stick - left
 - right - camera turn, forward
 - backward - adjust object distance.
- Left Controller:
 - Primary button - save created objects
 - Secondary button - change mode(Traing Gesture/ Gesture Recognition)
 - Grip button - delete object
 - Controll Stick - all directions - move camera

Bugs 1. When grabbed, object is move in players direction a little bit (also rotation of the object is changed). I have found video about solving this issue, however despite following all the steps, problem wasn't solved. All the other options also seemed to have no effect.

2. When object is grabbed and very close to camera, it can change players position a little bit. I think that it can be changed by adjusting camera and objects layer in Unity.

3. Camera turns are mapped on control stick in the right controller. Also, selected object rotation is mapped on the same stick. That causes camera left and right movement when trying to rotate selected object in this directions.

5 Testing

For the testing I have created sample scene using Polygon Fantasy Kingdom assets - house with interior and two houses without interior. Then the tester composes scene on regular monitor for first house without interior and composes scene in VR for second house. Each tester tested program separately and didn't know what scene must be composed.

I have tested my implementation on 5 people of different age. Unfortunately, not every person had previous game development experience, but, despite it, they also gave some useful information to improve project.

All of the testers, except one, agreed that scene creation in VR is more intuitive and practical than in 2D monitor space. But, for majority(three people) ordinary object selection(selection of objects from the list) is more preferable. "You can select and spawn more than one object from the list, and then just adjust their individual placement and direction" - one of the testers words.

Some of the testers suggested to add disabling collision by button(for placement time) and object attraction(example: placing object on the same Y axis coordinates as at neighbour objects). Also, they proposed to add 360 degree 3 axis object rotation with Control Stick on controller(there are only one axis rotation in my implementation now).

At least in this implementation of shapes recognition(PDollar) not recognize different hand writings. For example: i have trained model on my handwriting shapes, but some testers had troubles with gesture drawing.

Bug feedback and possible ways to solve it are in the Implementation section.

6 Conclusion

While implementing, I have noticed those disadvantages:

- 3D artists must do additional work, for example write gesture for model, train AI to recognize gestures, add components to move objects in VR
- It is difficult to invent new gestures for objects, the more objects are, the less possibilities for new object gesture you have. Also all the gestures are draw in 2D space and with one move(in my representation)

In my opinion gesture recognition would be more preferable in game play, for entertainment, rather than in games creation. As for VR world creation, development in VR gives more perspectives. Errors are found faster, their fixes are immediate, you don't have to take off your VR headset.

VR world creation needs combination of those methods, would be great to combine all of them(list, gesture recognition, procedural generation, etc.), so game developers can choose more suitable methods for them and their teams purposes.

7 Links and Sources

Whole Unity project is too big, i will upload only part of it on BRUTE, other part will be at my github repository - <https://github.com/dmytrokadar/SemestralVR>
<https://www.shapesxr.com/>

<https://www.citationmachine.net> - citations

<https://unity.com/>

<https://medium.com/hackernoon/neural-nets-vr-magic-3b066538aa5d>

<https://apps.apple.com/us/app/ikea-place/id1279244498>

<https://apps.apple.com/app/plnar/id1282049921>

3D palette: A virtual reality content creation tool - researchgate. (n.d.). Re-

trieved February 3, 2023, from https://www.researchgate.net/publication/221314329_3D_Palette_A_virtual_real

Mengyu Chen Media Arts and Technology, Chen, M., Technology, M. A. and, Marko Peljhan Media Arts and Technology, Peljhan, M., Science, M. S. C., Sra, M., Science, C., University, A. G., University, T., University, O., California, U. of, Tokyo, T. U. of, Canterbury, U. of, & Metrics, O. M. V. A. (2021, December 1). ENTANGLEVR: A visual programming interface for virtual reality interactive scene generation: Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology. ACM Conferences. Retrieved February 3, 2023, from <https://dl.acm.org/doi/10.1145/3489849.3489872>

Author links open overlay panelSergiu Oprea, AbstractInteraction in virtual reality (VR) environments (e.g. grasping and manipulating virtual objects) is essential to ensure a pleasant and immersive experience. In this work, Braido, P., Aydin, Y., Calli, B., Verschoor, M., Garre, C., Li, Y., Goldfeder, C., Ciocarlie, M., Jörg, S., Pollard, N. S., Kry, P. G., Bai, Y., Liu, C. K., Ye, Y., & Feix, T. (2019, July 11). A visually realistic grasping system for object manipulation and interaction in virtual reality environments. Computers & Graphics. Retrieved February 4, 2023, from <https://www.sciencedirect.com/science/article/abs/pii/S0097849319301098>

A survey on the procedural generation of virtual worlds - researchgate. (n.d.). Retrieved February 3, 2023, from https://www.researchgate.net/publication/320722498_A_Survey_on_the_Procedural_Generation_of_Virtual_Worlds

Nguyen-Phuoc, T. H., Richardt, C., Mai, L., Yang, Y., & Mitra, N. (1970, January 1). Blockgan: Learning 3D object-aware scene representations from unlabelled images. Advances in Neural Information Processing Systems. Retrieved February 4, 2023, from <https://proceedings.neurips.cc/paper/2020/hash/4b29fa4efe4fb7bc667c7b301b74d52d-Abstract.html>

Zeng, X., Mehdi, Q., & Gough, N. (1970, January 1). 3D scene creation using story-based descriptions. WIRE HOME. Retrieved February 4, 2023, from <https://wlv.openrepository.com/handle/2436/34232>

Gaze-supported 3D object manipulation in virtual reality - Jorge Goncalves. (n.d.). Retrieved February 4, 2023, from <https://www.jorgegoncalves.com/docs/chi21c.pdf>

<https://www.youtube.com/watch?v=YpobXhmgFQk> - Gaze support Meta Quest Pro

Vrsketchin: Exploring the design space of pen and tablet interaction ... (n.d.). Retrieved February 4, 2023, from https://www.researchgate.net/publication/341692846_VRSketchIn_Exploring_the_Design_Space_of_Pen_and_Tablet_Interaction

<https://assetstore.unity.com/packages/tools/input-management/pdollar-point-cloud-gesture-recognizer-21660>

<https://www.youtube.com/watch?v=GRSOrkmasMM> - video with PDollar usage

https://www.youtube.com/watch?v=kfA_73npjMA - second part of that video

<https://www.youtube.com/watch?v=XRaNYOXU0hw> - video with first bug potential fix

<https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@2.2/manual/index.html>

<https://syntystore.com/products/polygon-fantasy-kingdom>