VRFromX: From Scanned Reality to Interactive Virtual Experience with Human-in-the-Loop

Ananya Ipsita
aipsita@purdue.edu
C Design Lab, Purdue University
West Lafayette, IN, USA

Hao Li
li3604@purdue.edu
C Design Lab, Purdue University
West Lafayette, IN, USA

Runlin Duan
duan92@purdue.edu
C Design Lab, Purdue University
West Lafayette, IN, USA

Yuanzhi Cao
cao158@purdue.edu
C Design Lab, Purdue University
West Lafayette, IN, USA

Subramanian Chidambaram
schidamb@purdue.edu
C Design Lab, Purdue University
West Lafayette, IN, USA

Min Liu
liu66@purdue.edu
C Design Lab, Purdue University
West Lafayette, IN, USA

Karthik Ramani
ramani@purdue.edu
C Design Lab, Purdue University
West Lafayette, IN, USA

Figure 1: Workflow of VRFromX. User (1) scans a real environment to get raw point cloud using an iPad Pro, (2) creates virtual content by interacting with the point cloud using a Brush Tool and searching in repository through AI assistance, (3) attaches functions to the virtual objects using an Affordance Recommender and, (4) finally interacts with the reconstructed scene in a virtual environment.

ABSTRACT

There is an increasing trend of Virtual-Reality (VR) applications found in education, entertainment, and industry. Many of them utilize real world tools, environments, and interactions as bases for creation. However, creating such applications is tedious, fragmented, and involves expertise in authoring VR using programming and 3D-modelling softwares. This hinders VR adoption by decoupling subject matter experts from the actual process of authoring while increasing cost and time. We present VRFromX, an in-situ Do-It-Yourself (DIY) platform for content creation in VR that allows users to create interactive virtual experiences. Using our system, users can select region(s) of interest (ROI) in scanned point cloud or sketch in mid-air using a brush tool to retrieve virtual models and then attach behavioral propertie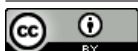s to them. We ran an exploratory study to evaluate usability of VRFromX and the results demonstrate feasibility of the framework as an authoring tool. Finally, we implemented three possible use-cases to showcase potential applications.

CCS CONCEPTS

• **Human-centered computing → Virtual Reality; Graphical user interface;**.

KEYWORDS

Virtual Reality; Embodied interaction; Point cloud interaction; Behavioral modelling; Scene reconstruction; Scene manipulation; Human-centered AI; Graphical user interface

1 INTRODUCTION

We are living in an era where substantial efforts are being made to create seamless pathways between the physical reality and it's digital counterpart. In the past decade, we have witnessed new technologies and systems created to derive the virtual from the reality, or the other way around, aiming to augment each other's capabilities. Due to the visual and spatial nature of reality, it's digital counterpart can be made useful in a variety of applications, such as remote collaboration, hologram conferencing, and simulated training [16] [12] [3] [27].

There are a few different interface platforms for users to interact with the digital version of the real world. In this paper, we focus on VR due to its capability to provide immersive experiences while supporting in-situ spatial and embodied interactions from the users [14]. The choice of 3D platform affords better 3D perception and direct testing of the content within the execution environment that matches the dimensionality of virtual content, without the need of switching from 3D to 2D interfaces. Moreover, 2D interfaces bring interaction issues while adding the extra dimension in unnatural and unconventional ways, and restrict an intuitive "flow" of using hands to work in VR [26].

Because of the state-of-the-art sensing hardware and computer vision algorithms, real world digitization can be easily done by commercially available 3D scanners, RGB [8] and RGB-D cameras [5]. The results from these devices are usually the point cloud or the mesh model reconstruction of the reality which provides a unique benefit of laying out the objects reasonably with realistic dimensions and unskewed spatial relationships. However, they can only serve as a spatially situated visual reference for the user without any interactive functionalities. In order to preserve the interactivity of the scanned real world scene for a desired VR experience, it requires professional developers to manually recreate the virtual scene while using the scanned results as a background modeling reference [7].

To make the scene reconstruction more accessible, prior works have introduced numerous in-situ authoring systems for Augmented Reality (AR) and VR applications [9] [24] [19]. The idea is to provide users with a modeling tool that intuitively uses the reality scene as a modeling reference for derivative scene reconstruction with added interactive functionalities. However, the amount of user interaction and technical modeling knowledge required are still substantial. Recent works in artificial intelligence (AI) have used deep learning to automatically reconstruct the digital scene from 3D scans [23] [2]. However, these automated approaches usually lack the ability to provide user in-the-loop customization. Inspired by previous works, we choose an approach that balances user interaction with AI automation.

We present VRFromX, an end-to-end system framework design that supports the authoring of interactive VR scenes from real world scans. Our workflow is illustrated in Figure 1: user scans a real world scene (a welding workshop in this case) into point cloud using a handheld 3D scanner; then replaces the point cloud objects with corresponding virtual models via embodied interactions with AI-assisted object retrieval; and then attaches functionalities onto the virtual objects and defines logical connections between them; finally the virtual scene is ready for VR interactions (this particular

scene can be used for virtual welding training). The contributions of this paper are listed as follows:

- The system workflow design of VRFromX that enables creation of interactive VR scenes, from scanning real world scenes into point cloud scans, to transforming point cloud objects into virtual 3D models, and to establishing functionalities and logical connections among virtual contents.
- An embodied user interaction design that supports point cloud segmentation and editing, the AI assistant that guides object retrieval and alignment, and the spatial and visual interface for functionality and logic authoring.
- A detailed system development of a welding use-case scenario, evaluation results obtained from a three-session user study, in addition to two example use-case demonstrations.

2 RELATED WORK

2.0.1 Scene Reconstruction and Manipulation in AR/VR. Prior work has deployed 3D reconstruction techniques to automatically reconstruct virtual scenes with scanned RGB-D information. Reality Skins [29] uses consumer-level RGB-D cameras to pre-scan and process the surrounding to translate it into a virtual environment with the best object layout. Oasis [30] procedurally generates an interactive VR experience using pre-scanned data from the physical environment as a template. Dreamwalker [34] leverages RGB-D sensors, Windows Mixed Reality tracking, and GPS to substitute real world walking experience with a dynamic virtual environment. VRoamer [6] scans the physical environment in real-time and dynamically instantiates pre-authored virtual experiences that fit the physical world. These systems mainly focus on employing RGB-D information to ensure safe walking in the physical world and construct pre-authored virtual environments. However, our system considers the RGB-D information as a medium that users can interact with to transform static point clouds into an interactive virtual experience.

Ens et al. [9] build an application to author Internet of Things (IoT) devices in virtual environments with multiple logic nodes. Ng et al.[19] provide a game authoring system that enables users to add virtual contents to the physical world and attach game-logics to virtual objects. Remixed Reality [20] builds a live 3D reconstruction with multiple external RGB-D cameras and provides direct manipulation of the reconstructed virtual scene. AffordIt! [22] offers a solution to assign affordances and constraints to virtual objects which do not have intrinsic behaviors after 3D reconstruction and segmentation. ARchitect [19] reconstructs an interactive virtual scene by scanning the physical environment and aligning virtual proxies with physical objects in consideration of affordances. Inspired by these works, we have devised ways to add logical connections among virtual objects within our authoring tool.

2.0.2 AI integration in AR/VR. Khurana et al. [13] highlighted the importance and breakthroughs that AI and VR can make when combined together and concluded that by the combination of AI, the virtual world will be more than a realistic world. Giunchi et al. [11] proposed a MVCNN-based retrieval method for searching models using free-form sketches within a virtual environment as queries. Zhu et al. [36] proposed a skeleton-based approach to draw 3D models in Head Mounted Displays (HMD) based virtual
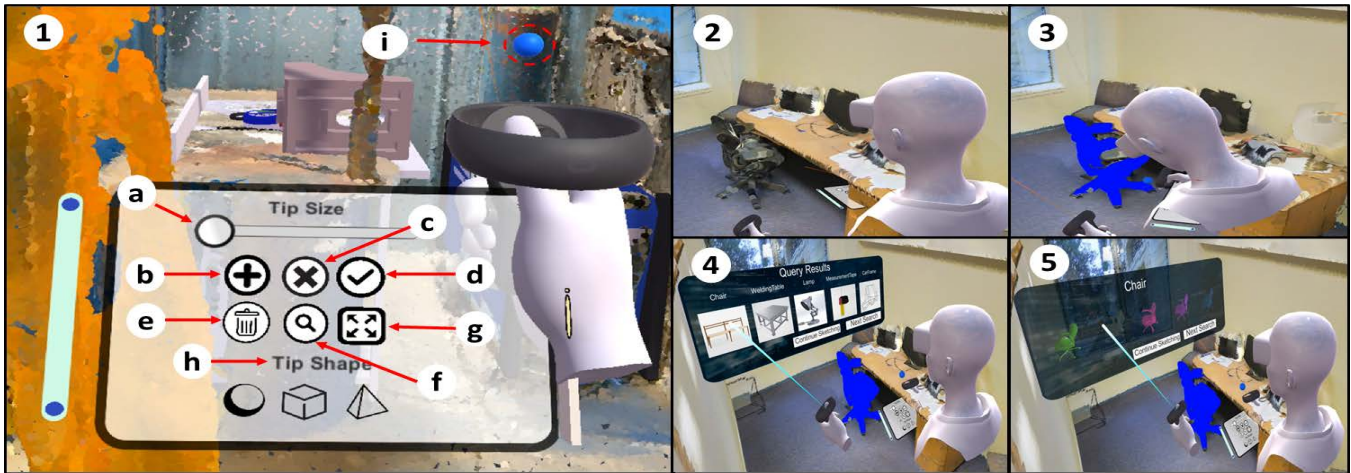
Figure 2: Interface of the Brush Tool: (a) Tip Size, (b) Add, (c) Delete, (d) Select, (e) Clear All, (f) Query, (g) Scale, (h) Tip Shape, (i) Brush Tip. User employs the Brush Tool to retrieve a chair (2-5).

environments using convolution surface-based brushes. Although much progress has been made in the computer vision community [28] [31], few of them have been merged with human interaction workflows to create and make use of 3D context in Mixed Reality applications.

3 THE VRFromX FRAMEWORK

We designed VRFromX as a proof-of-concept prototype that enables users to transfer a scanned point cloud scene into an interactive virtual experience. The major components of our system include: an interactive interface for embodied interaction on point cloud for AI assisted object retrieval, and behavioral modeling for virtual models. Point cloud acquisition of a real environment is performed using an inbuilt LIDAR in iPad Pro. This raw point cloud serves as a medium for user interaction in VR. Spatial and colour information of the point cloud helps in perceiving various objects present inside the scene. The user then employs a brush tool to segment ROI in the point cloud scene. Using the segmented point cloud as the input query, AI algorithms assist user in retrieving corresponding 3D models. Finally the user shifts to the behavioral modelling mode to assign functions to virtual models using an affordance recommender.

3.0.1 Point Cloud Interaction and Interface. In VRFromX, we designed a system for users to interact with scanned point cloud in an embodied manner inside a virtual environment. A brush tool was developed which enables users to select regions on point cloud or sketch in mid-air in a free-form manner. Various features of the brush tool are shown in Figure 2. After having gained a clear perception of the context of scanned scene and objects present inside it, users can exploit the brush tool features for content creation inside the virtual environment. They can manually segment ROI inside the point cloud or modify them (by adding or deleting points) to get corresponding virtual models from the database. The tool also allows users to sketch object shapes in mid-air if they do not find the objects in the partial point cloud. Keeping in consideration the

accessibility of regions inside the scene, the *Scale* feature enables users to manually scale the environment or orient it to a desired setting to suit requirements using simple hand gestures.

3.0.2 Object retrieval and alignment using AI assistance. Neural networks are implemented in VRFromX to assist the user in object retrieval and model alignment in an interactive manner. Object retrieval is performed as a two-step process, object classification followed by similarity check. After user selects a ROI in the point cloud or sketches the object shape, he/she can send the point cloud selection or sketch as a query input to the back-end system. The classification network deployed in back-end outputs top 5 possibilities for object classes present in database. When user selects the desired class, the network performs a similarity check and returns top 5 plausible results of object models that closely resemble the input query. The retrieval process is adaptive to the user's selections, i.e. if the user decides to change the shapes of input query using the brush tool, the AI algorithms adapts to the input and modifies search results accordingly. After the desired model is selected, AI algorithms assist the user by automatically aligning the 3D virtual model with the ROI according to scale, position and orientation. The interaction capabilities provided by VRFromX also enables users to re-purpose the point cloud to retrieve models of interest. Users can also perform simple hand gestures to grab and alter the position, orientation and scale of the virtual models based on their requirements.

3.0.3 Behavioral Modeling Design and Interface. After setting the environment, users can attach affordances to virtual objects to make them interactable. The term affordance is defined as an action that can be performed over an object(s) by an agent in an environment according to Gibson et al. [10]. In our work, we defined such behaviors of virtual objects to be categorized under three categories: *Animations*, *Displays*, and *Controls*. *Animations* can be defined as graphical animations in terms of motion or effects (sound, visual) (Figure 3a). *Displays* are panels used to display data for various attributes of virtual objects (Figure 3b). *Controls* are widgets
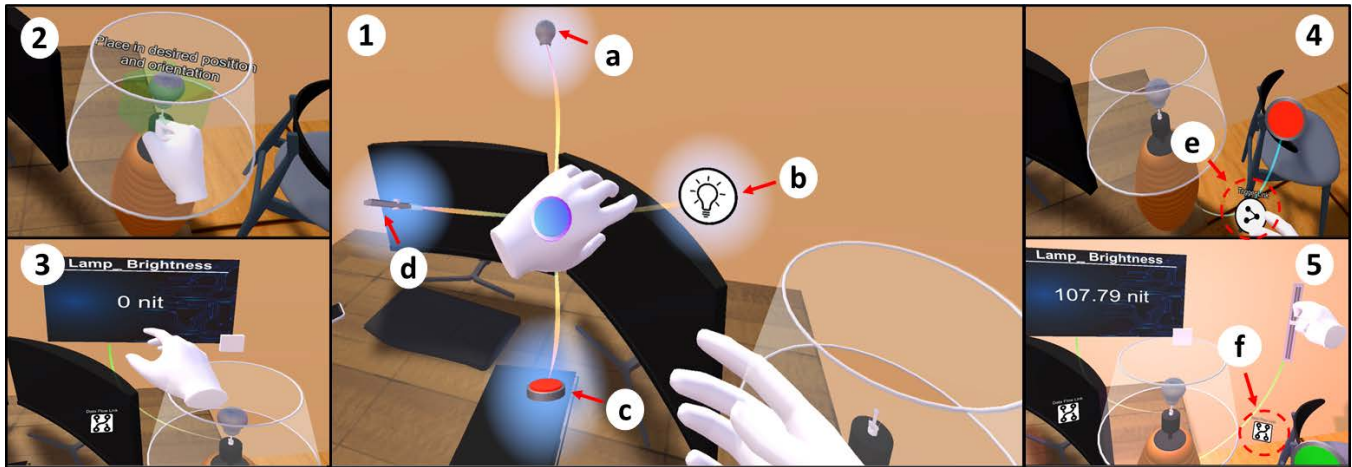
Figure 3: User assigns functions to a virtual lamp with the aid of (1) an Affordance Recommender and attaches (2) *Animations* - Lights (a), (3) *Displays* - Intensity Panel (b), (4) Trigger *Controls* - Switch button (c) with *Trigger Link* (e) and (5) Value Manipulation *Controls* - Slider (d) with *Data Flow Link* (f).

which fall under two categories: Value Manipulation or Trigger. Value manipulation *Controls* (Figure 3d) are used for transferring data from the object to *Displays* and Trigger *Controls* (Figure 3c) are used to trigger *Animations* for an object. Corresponding links are displayed as *Data Flow Link* and *Trigger Link* during the authoring process (Figure 3e, f).

We provide an Affordance Recommender for the user to choose behavioural attributes of a virtual object. Interface design of the affordance recommender is shown in Figure 3-1. Upon grabbing the virtual model of interest, users can enable affordance recommendation by pressing a virtual button that appears on the left hand. Object properties for specific classes are predefined and stored in text format along with 3D models in database. The recommender extracts these properties for target object and provides users with possible options. Users can grab one or more functions from the list of choices displayed by the recommender and attach them to the corresponding virtual object.

4 IMPLEMENTATION

We used Oculus Quest HMD with Oculus Link and Touch controllers to perceive the virtual experience. The system ran on a laptop PC with a Processor Intel® Core™ i7-8700K, NVIDIA GeForce RTX 2070 super graphics card and 64 GB of RAM. Front-end in VR-FromX system was developed within Unity Engine in C# and was connected with the pre-loaded back-end neural networks through a separate service thread in Unity. During the process of user-guided object retrieval, object classification is achieved by a PointNet [25] based classification network that provides the top 5 possible classes for the query input. The classification network is trained with 250 epochs on the ModelNet40 dataset [33] with 40 object classes and fine-tuned with 200 epochs on a welding dataset with 14 classes of welding equipment, each having 50 models. An alignment network based on PointNetLK [1] helps in aligning the selected 3D model to the ROI and is trained on the ModelNet40 dataset with

a data augmentation method that applies to random rotation and variation.

5 USER STUDY

5.0.1 ***Study Setup****.* Our user study was based on a welding use-case scenario. Welding is a hard skill which demands manual dexterity, adept motor skills and attention to numerous details of the process leading to a substantial skills gap and demand for trained welders in industry. Through their research study, Yunus et. al [35] laid out the advantages of using VR simulators in the welding process. From the research analysis, they have showed that respondents agree with the usage of VR welding simulator as a tool to develop basic welding skills to new trainees is necessary. We chose the welding use-case to test our system's capability to create and author a virtual MIG welding simulator. We adjusted our system requirements in terms of object models and their functions in database to enable an experienced welder for authoring the welding simulator. We created a database with model classes belonging to objects that can be typically found in an actual welding environment and trained our back-end neural networks on these objects. Object properties were defined such that the authoring process could involve creating a realistic welding puddle, accurate sounds and movements, simulation of sparks, slag, replicating machine set-up (amperage/voltage) and tracking key weld parameters including work angle, distance and position. In order to obtain the point cloud for the implementation, we scanned the welding station at a local welding-training company. A researcher walked around the scene holding an iPad Pro scanning every possible detail using an application called AppScanner [15]. The process was repeated 5 times each of which took around 10 minutes and the best scan was used for the study.

We ran the exploratory study with 18 participants (4 female, 14 male) with ages between 18-35. 15 users had prior welding experience, 8 were welding experts and 7 had prior authoring experience.

We observed the system's usability while examining how users interacted with VRFromX and what features they expected to achieve using the system. We also tried to evaluate the efficiency of the virtual welding trainer based on the comments from welding experts and our observations, if such simulator can facilitate seamless transition into real world hands-on welding for novices.

5.0.2 **Study Procedure**. At first, users were asked to scan any surrounding region using AppScanner. We included this step to give users an idea about point clouds and the ease of acquisition using a consumer-level product. Next, a virtual tour of the welding station was given in VR by placing virtual markers at appropriate positions in the scan showing pictorial representation of objects from the actual scene. After users had gained a proper understanding of the scene, a researcher explained the instructions verbally about the functions of physical buttons on Oculus controllers and virtual buttons in the system, and objectives of each task during a video demonstration. The entire session was divided into 3 tasks; object retrieval using the brush tool, behavioral modeling of virtual objects, and interaction with the final scene. Users were asked to fill up short questionnaire surveys between the tasks and finished up the study with a conversational interview. The study lasted for approximately 70 minutes and each user was given $20 compensation.

Data was collected for each user in terms of : (i) time required to finish each task, (ii) 7-point Likert Scale based questionnaire to test system's usability, and (iii) qualitative feedback. Each session was recorded from 3 camera views; first-person in VR and third-person in both VR and real environment. We segmented the first-person camera view in VR to record objective results in terms of time needed to perform tasks. After each task, users recorded their experience using the Likert survey. The survey questionnaire design was derived from the standard user experience survey System Usability Scale (SUS) [4]. We audio-recorded all subjective comments and suggestions from users for post-study analysis and summary. We interviewed the users about their experience with the system, comments on improvement and potential applications. The subjective feedback is used to explain the study results and inspire for future design insights.

5.0.3 **Results**. Users were able to complete every task. Survey results were analysed to identify users' scores on each aspect of the system, how usable were the interaction methods, how relatable were the actions to real world entities and how immersive was the experience. All Likert Scale questions are reported with mean (M) and standard deviation (SD).

Task1 : Point Cloud Interaction for Object Retrieval. Users used the brush tool for setting up the virtual welding environment. From video analyses, the average time (in minutes) taken to retrieve a 3D model from database was found to have M=2.3, SD=0.62. From survey results, sketching (M=5.28, SD=1.24) and selection of objects (M=4.89, SD=1.45) were easy inside the point cloud environment. Objects inside the point cloud scene were accessible without much effort (M=5.11, SD=1.24). Users were satisfied with the accuracy of retrieval results (M=4.61, SD=0.89) and time required to fetch objects from database (M=5.17, SD=1.26). Users were comfortable with adjusting scale of point cloud based on their convenience (M=4.78, SD=1.36). Overall, brush tool features were easy to use (M=5.78,

SD=0.97). Users could understand where to pay attention while performing the tasks (M=5.28, SD=1.33). Content created seemed relevant and realistic (M=5.44, SD=1.34). Users had an idea about the content to create in context of the scene (M=5.44, SD=1.34) and were satisfied with the content that they created (M=5.5, SD=1.12). Users were actively engaged during the task (M=6.11, SD=0.99). Oculus controllers felt accurate (M=5.50, SD=1.12) and the physical buttons helped with overall experience (M=5.50, SD=0.96). A low rating was received for weight of controllers affecting the overall process (M=2.17, SD=1.12).

Task2 : Behavioral modeling for virtual objects. Users attached various functions to virtual objects using the affordance recommender to make the virtual scene interactable. From video analyses, the average time (in minutes) taken to assign functions to a virtual model was found to have M=1.6, SD=0.62. From survey results, adding functions to virtual objects was easy (M=5.72, SD=0.99). Users knew where to look and pay attention (M=5.67, SD=1.15) and could assign functions without much effort (M=5.67, SD=1.2). The choices provided by the affordance recommender were meaningful (M=5.67, SD=1.49). Users could relate the functions with real world interactions (M=5.83, SD=1.17) and felt confident about the functions while assigning them (M=5.44, SD=1.30). Users could understand what functions to add in context of the scene (M=5.28, SD=1.15) and were actively engaged in the tasks (M=5.83, SD=1.34). Oculus controllers felt accurate (M=5.61, SD=0.89) and physical buttons helped with overall experience (M=5.00, SD=1.20). A low rating was received for weight of controllers affecting the overall process (M=2.78, SD=1.75).

Task3 : Virtual welding. Users performed virtual welding in the reconstructed scene. From survey results, users liked the welding trainer that they created during the above tasks (M=4.83, SD=1.80). They could relate the virtual interactions to real world interactions (M=5.39, SD=1.06). 13 users tried welding in VR more than once, felt confident (M=4.56, SD=1.50) and performed better with each trial (M=5.28, SD=1.56). Users were actively engaged (M=6.00, SD=1.05) and were willing to use the system again for transforming physical reality to VR (M=5.17, SD=1.64). Oculus controllers felt accurate (M=4.89, SD=1.49) and physical buttons helped with overall experience (M=5.00, SD=1.29). A low rating was received for weight of controllers affecting the overall process (M=2.89, SD=1.59).

6 USE-CASE DEMONSTRATIONS

We implemented 2 use-case scenarios to justify capabilities of our system to perform activities in a Mixed Reality environment. **Remote 3D Printing** use-case demonstrates an example of Mixed Reality fabrication using IoT-enabled devices and was developed to encourage high school students in STEM-based curriculum. Using sketching features, they can retrieve 3D designs for maker-based prototyping and print models remotely without the need of going into the complexity of constraint-based 3D-modelling softwares. After putting the selected models on the plate of the virtual printer, the models can be sent to the physical 3D printer for remote printing. **Robot-IoT Task Planning** use-case demonstrates the applicability of our system in Human-Robot collaboration to interact with IoT-enabled devices. Using the scanned point cloud overlaid in VR,

user specifies the path that the robot has to follow by selecting waypoints and authors the robot tasks at desired waypoints. Using brush tool features, user retrieves the models of IoT-enabled devices that the physical robot has to interact with and assigns functions to them. Finally the physical robot navigates on the path and interacts with the IoT-enabled devices in the real world as authored by the human.

7 DISCUSSION AND FUTURE WORK

Our results show that VRFromX provides users with an interactive platform to create meaningful content inside a virtual environment to transform physical reality into interactive virtual experiences. The interaction procedures are simple, easy to understand and use, and don't demand any specific skill expertise from users. Having been already familiar with the process from before, users are more attached to the authoring process since they can always draw a connection to the corresponding real world process. The system also provides an engaging and immersive experience inside VR through spatial and embodied interactions. From the comments and observations gathered during interview, implications were derived on possible applications in gaming and storytelling, education, training and simulation. *"Comfortable way to teach a soft skill. Before you actually try on a physical set-up"*, *"Escape room-It's a game so that people can escape from room to any mysterious world."*, *"Lab scenario - students can do it at home, can go to lab for data collection, rather than having problems with the set-up"*, *"Training purposes - drill mills, saw, CNCs"*, *"In my old factory, I could use this for setting parameters, training for setting parameters. Training how to do set up parameters for machines Useful when the task might be dangerous. Especially when they are doing some dangerous tasks, VR takes the danger out."*

During the study, some users pointed out to augment the capabilities of the brush tool to provide geometrical features to create simple shapes to assist during sketching. *"Use some features to draw straight lines, circles, rectangles; not draw them ourselves. e.g. in tools like paint, put simple shapes together to build complex shapes"* This can be achieved by adding features in the brush tool to enable creation of constraint based geometric shapes in 3D [42]. We also plan to consider usage of other modalities such as voice input to enhance user interaction. Insights were gained from the study to improve the efficacy of the virtual welding trainer. Adding scope for personalization and integrating other factors such as body size and parameters of biomechanics could be looked into to enhance user interactivity. Another direction to look into is to make use of AI algorithms to assist the users in segmentation of the point cloud [32] [31]. Further collaboration between humans and AI could be targeted by adding features to improve the accuracy and scalability of the retrieval networks through active learning [21]. We plan to explore the potential of VRFromX in reconstructing shapes by making use of shape primitives [17] or geometric features present in the point cloud [18] for unseen classes.

8 CONCLUSION

We have presented the system workfow design of VRFromX that transforms physical reality into interactive virtual experiences inside a VR environment. We have demonstrated how the system

design facilitates an embodied user interaction on scanned point cloud to create virtual content using AI assistance. We also demonstrated the usability of the functionality and logic authoring tool to attach affordances to virtual objects for a meaningful end-user experience. The results of our exploratory study in the welding use-case confirmed the usability, intuitive and engaging aspects of the system and also unveiled the potential directions for future endeavors. Participants commented in affirmation about the ability of the system to create beginner level training experiences for novices. We further implemented 2 more use-cases to demonstrate potential applications of VRFromX.

ACKNOWLEDGMENTS

REFERENCES

[1] Yasuhiro Aoki, Hunter Goforth, Rangaprasad Arun Srivatsan, and Simon Lucey. 2019. PointNetLK: Robust & efficient point cloud registration using PointNet. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Long Beach, CA). IEEE, New York, NY, USA, 7163–7172.

[2] Armen Avetisyan, Manuel Dahnert, Angela Dai, Manolis Savva, Angel X. Chang, and Matthias Nießner. 2019. Scan2CAD: Learning CAD Model Alignment in RGB-D Scans. In *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE* (Long Beach, CA). IEEE, New York, NY, USA, 2609–2618.

[3] Esther Z Barsom, Maurits Graafland, and Marlies P Schijven. 2016. Systematic review on the effectiveness of augmented reality applications in medical training. *Surgical endoscopy* 30, 10 (2016), 4174–4183.

[4] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[5] Kang Chen, Yu-Kun Lai, and Shi-Min Hu. 2015. 3D indoor scene modeling from RGB-D data: a survey. *Computational Visual Media* 1, 4 (2015), 267–278.

[6] Lung-Pan Cheng, Eyal Ofek, Christian Holz, and Andrew D Wilson. 2019. VRoamer: Generating On-The-Fly VR Experiences While Walking inside Large, Unknown Real-World Building Environments. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (Osaka, Japan). IEEE, New York, NY, USA, 359–366.

[7] Brent Cowan and Bill Kapralos. 2014. A survey of frameworks and game engines for serious game development. In *2014 IEEE 14th International Conference on Advanced Learning Technologies* (Athens, Greece). IEEE, New York, NY, USA, 662–664.

[8] Ruofei Du, Eric Turner, Maksym Dzitsiuk, Luca Prasso, Ivo Duarte, Jason Dourgarian, Joao Afonso, Jose Pascoal, Josh Gladstone, Nuno Cruces, et al. 2020. DepthLab: Real-Time 3D Interaction With Depth Maps for Mobile Augmented Reality. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Minneapolis, Minnesota, USA). ACM, New York, NY, USA, 829–843.

[9] Barrett Ens, Fraser Anderson, Tovi Grossman, Michelle Annett, Pourang Irani, and George Fitzmaurice. 2017. Ivy: Exploring spatially situated visual programming for authoring and understanding intelligent environments. In *Proceedings of the 43rd Graphics Interface Conference.* CHCCS/SCDHM, Edmonton, Alberta, 156–162.

[10] James J Gibson. 1977. The theory of affordances. *Hilldale, USA* 1, 2 (1977), 67–82.

[11] Daniele Giunchi, Stuart James, and Anthony Steed. 2018. 3D sketching for interactive model retrieval in virtual reality. In *Proceedings of the Joint Symposium on Computational Aesthetics and Sketch-Based Interfaces and Modeling and Non-Photorealistic Animation and Rendering.* ACM, New York, NY, United States, 1–12.

[12] Hirokazu Kato and Mark Billinghurst. 1999. Marker tracking and HMD calibration for a video-based augmented reality conferencing system. In *Proceedings 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR'99)* (San Francisco, CA, USA). IEEE, New York, NY, USA, 85–94.

[13] Ashish Khurana, Bhanu Prakash Lohani, and Vimal Bibhu. 2019. AI Frame-Worked Virtual World Application-The Ramification of Virtual World on Real World. In *2019 International Conference on Automation, Computational and Technology Management (ICACTM)* (London, UK). IEEE, New York, NY, USA, 582–585.

[14] Konstantina Kilteni, Raphaela Groten, and Mel Slater. 2012. The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments* 21, 4 (2012), 373–387.

[15] Laan Labs. 2020. 3D Scanner App. https://www.3dscannerapp.com/.

[16] Valerie D Lehner and Thomas A DeFanti. 1997. Distributed virtual reality: Supporting remote collaboration in vehicle design. *IEEE Computer Graphics and Applications* 17, 2 (1997), 13–17.

[17] Lingxiao Li, Minhyuk Sung, Anastasia Dubrovina, Li Yi, and Leonidas J Guibas. 2019. Supervised fitting of geometric primitives to 3D point clouds. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Long Beach, CA). IEEE, New York, NY, USA, 2652–2660.

[18] Yangyan Li, Rui Bu, Mingchao Sun, Wei Wu, Xinhan Di, and Baoquan Chen. 2018. PointCNN: Convolution on X-transformed points. In *Advances in neural information processing systems.* NIPS 2018, Montreal, Canada, 820–830.

[19] Chuan-en Lin, Ta Ying Cheng, and Xiaojuan Ma. 2020. ARchitect: Building Interactive Virtual Experiences from Physical Affordances by Bringing Human-in-the-Loop. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI). ACM, New York, NY, USA, 1–13.

[20] David Lindlbauer and Andy D Wilson. 2018. Remixed reality: manipulating space and time in augmented reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal, Canada). ACM, New York, NY, USA, 1–13.

[21] Huan Luo, Cheng Wang, Chenglu Wen, Ziyi Chen, Dawei Zai, Yongtao Yu, and Jonathan Li. 2018. Semantic labeling of mobile LiDAR point clouds via active learning and higher order MRF. *IEEE Transactions on Geoscience and Remote Sensing* 56, 7 (2018), 3631–3644.

[22] Sina Masnadi, Andrés N Vargas González, Brian Williamson, and Joseph J LaViola. 2020. AffordIt!: A Tool for Authoring Object Component Behavior in VR. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (Atlanta, GA, USA). IEEE, New York, NY, USA, 741–742.

[23] Muzammal Naseer, Salman Khan, and Fatih Porikli. 2018. Indoor scene understanding in 2.5/3D for autonomous agents: A survey. *IEEE Access* 7 (2018), 1859–1887.

[24] Gary Ng, Joon Gi Shin, Alexander Plopski, Christian Sandor, and Daniel Saakes. 2018. Situated game level editing in Augmented Reality. In *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction* (Stockholm, Sweden). ACM, New York, NY, USA, 409–418.

[25] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. 2017. PointNet: Deep learning on point sets for 3D classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (Honolulu, HI, USA). IEEE, New York, NY, USA, 652–660.

[26] Udo Schultheis, Jason Jerald, Fernando Toledo, Arun Yoganandan, and Paul Mlyniec. 2012. Comparison of a two-handed interface to a wand interface and a mouse interface for fundamental 3D tasks. In *2012 IEEE Symposium on 3D User Interfaces (3DUI)* (Costa Mesa, CA, USA). IEEE, New York, NY, USA, 117–124.

[27] Will Seidelman, Michael Lee, C Melody Carswell, Travis Kent, Bo Fu, and Ruigang Yang. 2014. User centered design of a hybrid-reality display for weld monitoring. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems.* ACM, New York, NY, USA, 2059–2064.

[28] Tianjia Shao, Weiwei Xu, Kun Zhou, Jingdong Wang, Dongping Li, and Baining Guo. 2012. An interactive approach to semantic modeling of indoor scenes with an RGBD camera. *ACM Transactions on Graphics (TOG)* 31, 6 (2012), 1–11.

[29] Lior Shapira and Daniel Freedman. 2016. Reality Skins: Creating Immersive and Tactile Virtual Environments. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (Merida, Mexico). IEEE, New York, NY, USA, 115–124.

[30] Misha Sra, Sergio Garrido-Jurado, and Pattie Maes. 2017. Oasis: Procedurally Generated Social Virtual Spaces from 3D Scanned Real Spaces. *IEEE transactions on visualization and computer graphics* 24, 12 (2017), 3174–3187.

[31] Harald Steinlechner, Bernhard Rainer, Michael Schwärzler, Georg Haaser, Attila Szabo, Stefan Maierhofer, and Michael Wimmer. 2019. Adaptive point cloud segmentation for assisted interactions. In *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games* (Montreal, Quebec, Canada). ACM, New York, NY, USA, 1–9.

[32] Julien Valentin, Vibhav Vineet, Ming-Ming Cheng, David Kim, Jamie Shotton, Pushmeet Kohli, Matthias Nießner, Antonio Criminisi, Shahram Izadi, and Philip Torr. 2015. Semanticpaint: Interactive 3D labeling and learning at your fingertips. *ACM Transactions on Graphics (TOG)* 34, 5 (2015), 1–17.

[33] Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 2015. 3D ShapeNets: A deep representation for volumetric shapes. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (Boston, Massachusetts). ACM, New York, NY, USA, 1912–1920.

[34] Jackie Yang, Christian Holz, Eyal Ofek, and Andrew D Wilson. 2019. DreamWalker: Substituting real-world walking experiences with a virtual reality. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA). ACM, New York, NY, USA, 1093–1107.

[35] Faizal Amin Nur Yunus, Jamil Abd Baser, Saiful Hadi Masran, Nizamuddin Razali, and Bekri Rahim. 2011. Virtual reality simulator developed welding technology skills. *Journal of Modern Education Review* 1, 1 (2011), 57–62.

[36] Xiaoqiang Zhu, Lei Song, Lihua You, Mengyao Zhu, Xiangyang Wang, and Xiaogang Jin. 2017. Brush2Model: Convolution surface-based brushes for 3D modelling in head-mounted display-based virtual environments. *Computer Animation and Virtual Worlds* 28, 3-4 (2017), e1764.