Multi-3D-Models Registration-Based Augmented Reality Instructions for Assembly

Seda Tuzun Canadinc*
Webster University

Wei Yan†
Texas A&M University

Abstract

BRICKxAR (Multi 3D Models/M3D) prototype offers markerless, in-situ, and step-by-step, highly accurate Augmented Reality (AR) assembly instructions for large or small part assembly. The prototype employs multiple assembly phases of deep learning-trained 3D model-based AR registration coupled with a step count. This ensures object recognition and tracking persist while the model updates at each step, even if a part's location is not visible to the AR camera. The use of phases simplifies the complex assembly instructions. The testing and heuristic evaluation findings indicate that BRICKxAR (M3D) provides robust instructions for assembly, promising potential applicability at different scales and scenarios.

Keywords: In-situ AR, AR-Assisted Assembly.

Index Terms: Augmented Reality, Assembly, Human-Computer Interaction.

1 Introduction

Augmented Reality (AR) instructions for parts assembly have broad applications, including manufacturing, building construction, repair and maintenance, and Do-It-Yourself (DIY) furniture. The current state of the art for AR assembly instructions is either based on object recognition or relies on markers. However, object recognition-based AR instructions are often limited to overly simplified assemblies with a small number of parts or work only when the assembly parts are large [1]. Besides parts often have distinct geometries that make them easily detectable [2]. When the geometry of parts is not distinct and is similar to each other, it requires constrained setups [3] Similarly, marker-based applications also necessitate constrained setups, imposing limitations on the interaction with the assembly model due to the use of markers. Thus, these methods do not apply well to complex assemblies. Another challenge is detecting small geometry changes in the assembly especially when the parts become smaller. On the other hand, as the assembly becomes more complicated some parts may become partially invisible to the AR camera during the process due to occlusion by other parts. Moreover, if the user cannot hold by hand and examine the model, the instructions cannot be presented in AR at full potential, such as viewing the assembly model from various angles. Although 3D-model-based AR promises high precision and tangibility [4], there are both errors and tolerance for detecting the 3D model. Furthermore, the model's geometry changes as the assembly proceeds, therefore one model trained by deep learning for AR registration would not work for the entire assembly process.

This research aims to address those research gaps by answering the following questions: How to enable 3D AR assembly instructions while there is a geometry change at each step of assembly within potentially hundreds or more assembly steps; How to improve the design of 3D AR assembly instructions toward optimized AR instructions: What new features should be developed to enhance 3D AR assembly instructions. A mobile assembly instruction application, BRICKxAR (M3D), is developed by employing Multi-3D-Models (M3D) Registration-based AR to explore answers to the research questions. The application is evaluated for its accuracy through testing and usability through heuristic evaluation.

* corresponding author, e-mail: stuzuncanadinc@webster.edu
† e-mail: wyan@tamu.edu

2 Prototype

For the test case, LEGO® Arc de Triomphe set consisting of 386 parts (assembly steps) made up of small LEGO bricks, was chosen. To address the complexity of a large number of steps and occluded part locations, the assembly process was divided into five phases, each containing multiple steps, in order to utilize the tolerance in 3D model-based registration for small geometry changes. These phases include ground plane registration for Phase I, utilizing a flat surface to anchor the virtual assembly at the beginning, and 3D-model-based registration for the rest of the four phases. Because in the beginning, there was no model to be detected, which was necessary to initiate the 3D model-based registration. The user aligns the physical assembly with the virtual model, and the assembly model remains on the flat surface until the end of Phase I. To ensure a robust and accurate 3D-model-based registration for the subsequent phases, it was crucial to achieve distinct edges to facilitate the detection of each phase. Asymmetrical geometry was aimed to provide distinct edges for initiating Phases II, III, IV, and V, while also considering structural stability, usability, and the user experience. Through experiments, the assembly order was reorganized differently than LEGO's original order in the instruction booklet, and optimized to meet these requirements. Subassemblies were eliminated in this respect. The CAD models (.obj) of Phases II, III, IV, and V were used for deep learning training in Vuforia Model Target Generator 9.4.6 for object recognition (Figure 1).

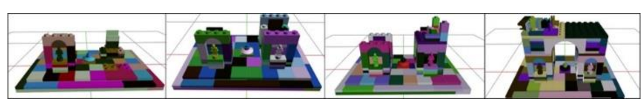

Figure 1: Left to right: BRICKxAR (M3D) assembly Phase II, III, IV, and V models.

In the Unity 3D environment, the deep learning-trained AR registration models for the four phases were combined, and later within the application can be automatically aligned with the physical models of respective phases, phase by phase in AR. The corresponding rendered, wireframe, or transparent virtual models of the assembly were further aligned with the registration phase models. Each of the 386 steps was matched with a corresponding step number based on a step count. The step count, through a user's tapping "Next" or "Previous" buttons, enables and controls the switch between phases and ensures the display of partially-occluded parts. (Figure 2).

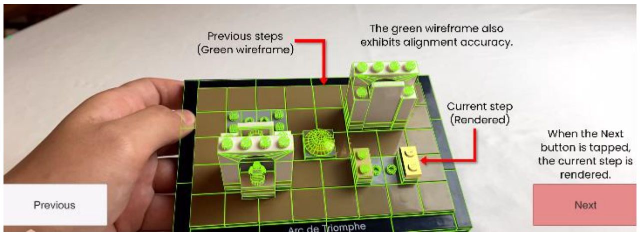

Figure 2: Realistic visualization of assembly steps and accurate registration even if assembly is moved by hand.

3 Testing and Heuristic Evaluation

For prototype testing, the entire assembly process of 386 LEGO bricks was recorded using a screen recorder on a mobile device. An iPhone 11 with a screen resolution of 1792x828 pixels was used. The entire process was observed in terms of registration stability and accuracy, the transition quality between different 3D-model-based AR registration phases, and hand occlusion's impact on the registration. Selected screenshots of different phases were analyzed in Adobe Photoshop. The screenshots were evaluated by drawing perpendicular lines between the edges of the physical LEGO bricks and the AR wireframe lines to measure the errors between the virtual parts and the actual physical model. The gap between these two edges was measured by the distance of the in-between perpendicular line added to the images (Figure 3).

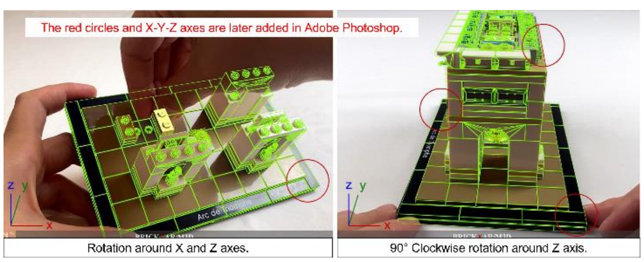

Figure 3: When the model is rotated around both X and Z axes on the left and around only Z axis on the right, the alignment error is less than 1 pixel shown in red circles.

To evaluate the usability and design of the BRICKxAR (M3D) prototype, 30 LEGO builders and 3 experts in Visualization and Human-Computer Interaction were interviewed about assembly, assembly instructions, and AR. They were asked to make a comparison of the prototype with the marker-based version by watching the latter's video [5], and perform a simple assembly task of 10 parts (4 in the front left tower and 6 hidden at the back, challenging visibility when viewed from the front) while guided by the prototype (Figure 4). They were asked to review the prototype and its unique features.

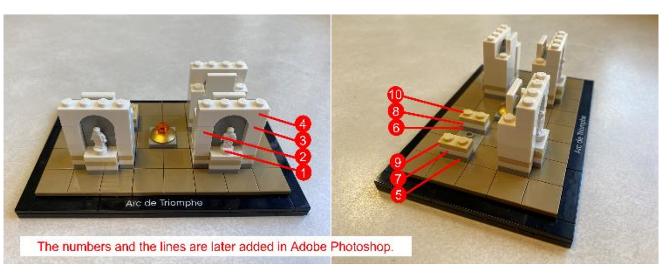

Figure 4: BRICKxAR (M3D) assembly task: the location and order of the parts.

4 Analysis and Results

The findings of the testing show that multi-3D-models-based AR registration in all phases is consistently accurate, even though there is hand occlusion to a greater extent (66% of the model); however, Ground Plane registration varies. Ground Plane relies on screen tapping for scale, making it hard to align digital and physical models. Moreover, the model needs to be at the same location during this phase and requires users always aligning the physical model with AR model on the screen. The transition between different phases of 3D-model registration is seamless, both during assembly and disassembly. The users' review shows that when the assembly step is hidden from sight, users need assistance to point out the assembly location and most of them state that it would be helpful if there is an instruction or a visual clue to guide them to the exact location. Lighting has a significant impact on the efficient working of the prototype. Most users switch their gaze from the screen to the actual assembly while assembling. Most users did not hold the model to see the part locations at the back but rotated the model on the surface instead. Most of the users said eliminating subassemblies or changing the assembly order did not affect them.

5 Observations and Conclusion

The contributions of this research are precise and robust multi-3D models-based, in-situ, step-by-step 3D AR instructions for assembly, even with small parts of a large number, the design of the assembly order, and learning from the heuristic evaluation on how to improve 3D AR instructions to facilitate usability. The mobile phone as AR device was a barrier between the users and assembly, so most users look at the screen for AR instructions and turn their gaze to the physical model outside of the screen for assembly. This limitation could be overcome by using AR headsets. Determining the phase models by a trial-and-error method and preparing instructions for each model individually are not efficient. For future work, assembly assistance with audio and text, assembly order optimization, and instructions automatization will be explored. (The project demo video is available at https://www.youtube.com/watch?v=4MUXR3iRIc8)

Acknowledgments

The research is funded by the National Science Foundation (NSF) Grant No. 2242432 and Grant No. 2119549.

References

[1] Y. Su, J. Rambach, N. Minaskan, P. Lesur, A. Pagani, and D. Stricker, “Deep multi-state object pose estimation for augmented reality assembly,” *Adjun. Proc. 2019 IEEE Int. Symp. Mix. Augment. Reality, ISMAR-Adjunct 2019*.

[2] K. Yang, Y. Guo, P. Tang, H. Zhang, and H. Li, “Object registration using an RGB-D camera for complex product augmented assembly guidance,” *Virtual Real. Intell. Hardw.*, vol. 2(6), pp. 501–517, 2020.

[3] M. Funk, T. Kosch, S. W. Greenwald, and A. Schmidt, “A benchmark for interactive augmented reality instructions for assembly tasks,” *ACM Int. Conf. Proceeding Ser.*, 30-November, pp. 253–257, 2015.

[4] S. Tuzun Canadinc and W. Yan, “3D-Model-Based Augmented Reality for Enhancing Physical Architectural Models,” in *Proceedings of Conf. on Education and Research in Comp. Aided Architectural Design in Europe*, 2022, vol. 2, pp. 495–504.

[5] W. Yan, “Augmented reality instructions for construction toys enabled by accurate model registration and realistic object/hand occlusions,” *Virtual Real.*, vol. 26, no. 2, pp. 465–478, 2022.