AI-POWERED AUGMENTED REALITY FOR ADVANCING HUMAN-COMPUTER INTERFACE IN ASSEMBLY TASKS

innovationX

Powered by: Texas A&M School of Innovation

Background:

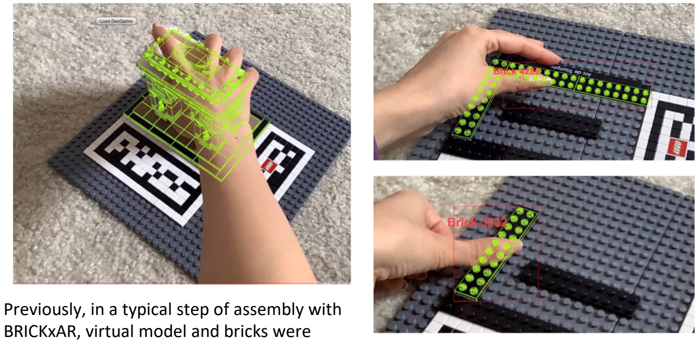
Major assembly, manufacturing, construction, and maintenance projects are increasingly complex. Automation and robotics are replacing many manual, repetitive, and standard tasks. Still, more than 95% of jobs consist of activities that need human labors. Assembly tasks can take advantages of human-machine interfaces (HMIs) that allow human operators to collect data, and monitor, program, and control the system. However, traditional HMI cannot effectively contextualize and interact with future workflows that now include physical and digital work information. As a major field of HMI, Human-Computer Interface (HCI) has started to integrate Augmented Reality (AR) in the workflow in assembly tasks. AR superimposes digital images on the real-world view of human users, putting the answers right where the questions are, and may greatly benefit manufacturing, building construction, and part assembly by human workers. Artificial Intelligence (AI) has the potential to significantly empower AR and advance HCI in assembly tasks. The proposed project aims to research and develop innovative AI-Powered AR for advancing HCI in assembly tasks.

Project Development:

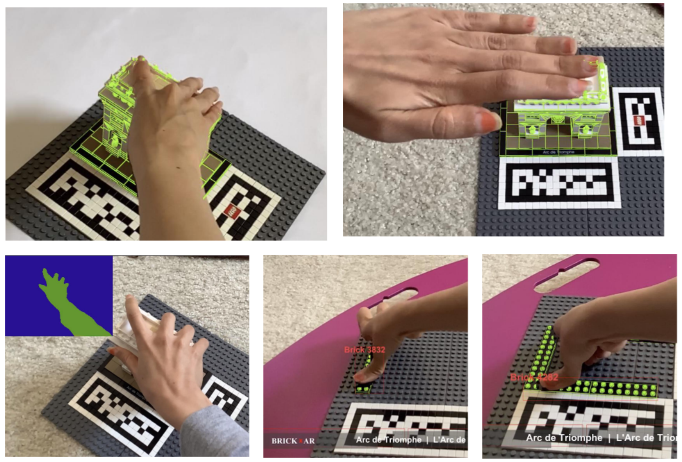
The project builds AI components into our prior work BRICKxAR - AR-aided instruction for construction, using virtual bricks to guide physical assembly (e.g. LEGO) step by step, through precise AR model registration.
(Demo video: https://youtu.be/7JDW_IDv7FU)

AI for hand detection and occlusion in AR:

To enhance hand-eye coordination and realistic immersive AR experience of users, the project has previously tested "grasping virtual objects with real hands", through hand detection and hand-brick occlusion using Computer Vision techniques. The latest development is made by upgrading BRICKxAR and using the AI feature of body occlusion supported by Apple's ARKit and Unity's AR Foundation. Using the body occlusion technique, the virtual model occlusion is enabled in real-time through the steps of the assembly process. In this way, the virtual model is occluded by a user's physical hands/body based on the virtual model's relative depth compared to the hands/body in the AR camera view. More precise depth detection is possible through the state-of-the-art Lidar technology which is in the future plan of the team.



Previously, in a typical step of assembly with BRICKxAR, virtual model and bricks were rendered in front of a user's hand



BRICKxAR uses the AI feature supported by Apple's ARKit and Unity's AR Foundation to detect hands/body to enable virtual models occluded by the user's hands.

TEXAS A&M UNIVERSITY College of Architecture

TEXAS A&M UNIVERSITY Engineering

AI-POWERED AUGMENTED REALITY FOR ADVANCING HUMAN-COMPUTER INTERFACE IN ASSEMBLY TASKS

innovationX

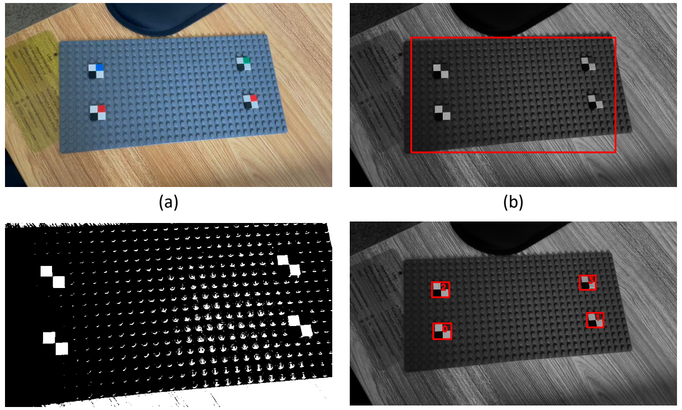
Powered by: Texas A&M School of Innovation

AI for AR model registration:

By using predefined artificial AR markers, we can detect the AR markers from the camera or video images and estimate the anchor to localize the AR model. Previously, the AR markers in BRICKxAR used two large AprilTag-like images, which were composed by many LEGO bricks. To simplify the design of AR markers, we utilize four 2-by-2 square checkerboards to mark the region where the AR model is placed.

We utilize the precise corners formed by the strong black and white squares to compensate the reduction of the number of features available in the new AR markers. Besides, the difficulty in the pattern detection raises when we reduce the size of the new AR markers. The small makers make the original image tracking method invalid. We start to apply Deep Learning methods to the AR registration problem. However, it is impossible to deploy the Deep Learning method with complex graph model in the pattern detection due to the power and computation constraints on the mobile devices. To achieve the real-time requirement, we choose SSD-MobileNetV2 as our deep learning model since SSD is a much simpler graph structure and MobileNet is light-weighted, which can perform the inference.

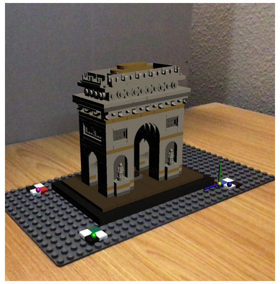
To tackle the small pattern detection problem, we convert the image into a binary image to produce much stronger signal of black-white pattern. We also guide the user to align the physical marker patterns to increase the success rate in the initialization step by restricting the region of interest (ROI) of the pattern detection. After the initialization is complete, we track both ROI and the detected patterns to maintain the registration.



(a) (b) (c) (d)

(a) New checkerboard-inspired AR markers, (b) the guiding border when performing the initialization, (c) the binary image used for pattern detection, and (d) the result of pattern association.

We use Harris corner detection to detect the corner formed by the strong black and white squares. We apply the geometric constrains such as that the corner must be close to the center of the pattern to filter out the false corners. We associate the patterns by using both the rectangular shape and the color. After we receive at least 3 associated corners, we perform Perspective-n-Point algorithm to estimate the camera pose. By combining the result provided by visual-inertial algorithm embedded in ARKit, we then are able to localize the AR model in the AR scene.



AR model deployment using YOLOv3 model during experiments

Project Team

TEXAS A&M UNIVERSITY College of Architecture

TEXAS A&M UNIVERSITY Engineering