Augmented Reality Markers: CAD Renderings vs Photos

Problem Statement
The goal of this project is to design and build an interactive augmented reality work instruction interface. We compare the 2D renderings of 3D CAD models vs. photos of the actual objects as markers in terms of:
- Setup difficulty
- Feature point extraction
- Orientation detection

Background
Existing methods for detecting 3D objects:
- 2D image fiducials
  Cons: textured, physical augmentation required
- 3D object geometry
  Cons: real-time mesh capture of workspace needed, expensive computation
- 2D photographs
  Cons: physical staging setup, lighting variation

Proposed Solution

Results
- Color rendering takes avg 23 ms
- Depth rendering takes avg 29 ms + 21 ms normalize
- Multi-resolution feature extraction from 4 to 72 dpi takes >15 sec per marker

Conclusion
- More feature points for more robust detection
- Flat shaded areas do not provide enough feature points for color CAD renderings
- Background color must match for detection
- Match extraction and capture resolutions for best performance
- For photos and renderings, choose likely poses
- Photos are more robust to marker/model mismatches than CAD renderings
- Overall recommendation:
  - Photos have a less technically involved setup
  - Photos have the most feature points
  - Photos result in the best detection

Limitations
- Mismatched CAD model
- Depth camera resolution requires larger objects and marker poses from farther away
- 2D markers limit the user’s manipulation range
- Slow feature point extraction limits real-time generation
- Complex depth feed capture process

Future Work
- Add NFT to JSARToolKit5 for turnkey integration
- Add poses/logic to expand manipulation range
- Evaluate correlation of Sobel, Canny, LoG edges

References
Arpit Gupta, Dieter Fox, Brian Curless, Michael Cohen, DuploTrack: a real-time system for authoring and guiding duplo block assembly, Proceedings of the 25th annual ACM symposium on User interface software and technology, October 07-10, 2012, Cambridge, Massachusetts, USA

Tools
- Objects: Duplo bricks
- CAD model: OBJ file
- Marker rendering: Three.js
- Extraction & Detection: ARToolKit.org Natural Feature Tracking
- Color camera: iSight 1920x1080 (Recorded 29 fps)
- Depth camera: Structure Sensor 640x480 (Recorded from NiViewer: 680x360, 15 fps)
- Overlay rendering: Unity

For more information, please contact Michelle Norris mln7740@cs.rit.edu

By: Michelle Norris mln7740@cs.rit.edu
Rochester Institute of Technology – Computer Science
Advisors: Dr. Reynold Bailey & Dr. Daniel Ashbrook