POS359: Dynamic Area of Interest (AOI) Matching in Simulated Environments via a Direct Coordinate Transform Approach

Research Assistants: Michael Williamson Tabango (will2689@purdue.edu), Sibibalan Jeevanandam (sjeevana@purdue.edu)
Principal Investigator: Dr. Neera Jain (neerajain@purdue.edu)

Motivation

- Eye-tracking is a widely used behavioral sensing modality in human factors research
- Dynamic AOIs are important in applications such as human interactions with conditionally automated (SAE Level 3) vehicles
- Dynamic AOI matching (mapping eye-gaze measures to non-stationary stimuli) is challenging to automate
- **Contribution:** A fully automated and novel method for dynamic AOI matching well-suited for simulated environments in which the locations and orientation of objects of interest are known.

Existing Methods

- **Video Annotation [1]:** A human technician defines the AOIs at keyframes which are interpolated to frames that are not directly annotated (prohibitively time consuming for many applications).
- **Draw and Track Algorithms [2]:** Incorporate the use of edge detectors to follow an AOI that a human technician only initially defines (time consuming and fairly intermittent).
- **Active Video Methods [2,3]:** Involve algorithms that only work with very specific and known models of the camera's motion through space.
- **Image Segmentation [2,4]:** Involve deep learning techniques to segment images into individual objects (high computational time that scales with resolution).

Approach

- Starting with a region in 3D space, the steps shown in the flow diagram below (**Fig. 1**) transform the vertex locations to pixel locations on the screen.

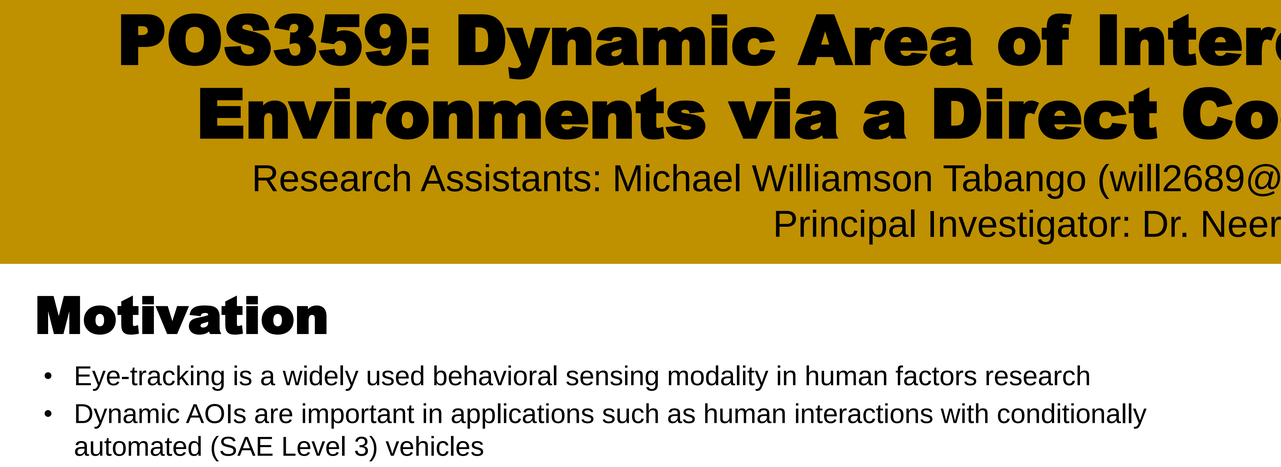

Fig. 1: Flow diagram for the Direct Coordinate Transform method.

This method involves:

1. Defining polyhedrons in 3D simulation coordinates that enclose the object of interest in local coordinate frames
2. Transforming the vertices of these polyhedrons into the camera frame
3. Projecting these transformed vertices into the screen pixel space
4. Computing the convex hull of the projected points to obtain a 2-dimensional polygon that defines the area of interest on a screen (**Fig. 2**)

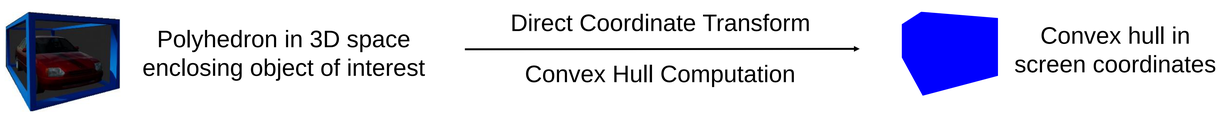

Fig. 2: An illustration of the method.

Advantages

- Fully-automated and does not rely on machine learning algorithms.
- Improvement in speed and robustness to sources of error such as lighting.
- Speed of this method is invariant to the size of the display, unlike image segmentation methods.

Results

- We demonstrate the use of this method for objects simulated in Unreal Engine 5 displayed on 3 screens with a combined resolution of 5760x1080 pixels (**Fig. 3**)

Fig. 3: Typical frame containing objects of interest (stop sign, pedestrian, and other vehicles)

- Camera calibration is achieved by inserting a black and white checkerboard pattern in various locations and orientations in the simulated environment (**Fig. 4**)
- Camera intrinsics are learned using MATLAB's Camera Calibration application

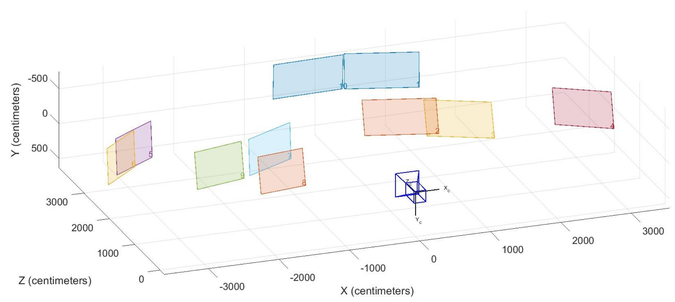

Fig. 4: Camera centric visualization of calibration extrinsics

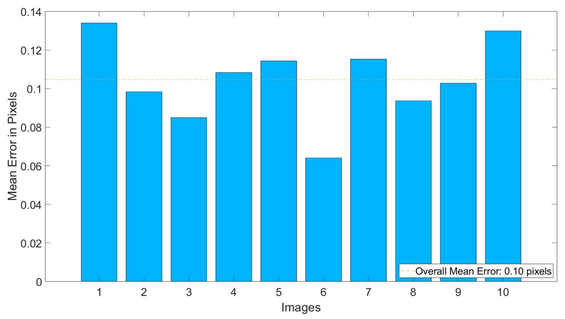

Fig. 5: Mean re-projection error by calibration image

- Unreal Engine uses a camera highly suitable for calibration, resulting in re-projection errors on the order of 0.1 pixels (**Fig. 5**)
  - An order of magnitude more precise than image segmentation for polyhedron objects
- Position and orientation logs at 2Hz do not introduce a significant error due to motion

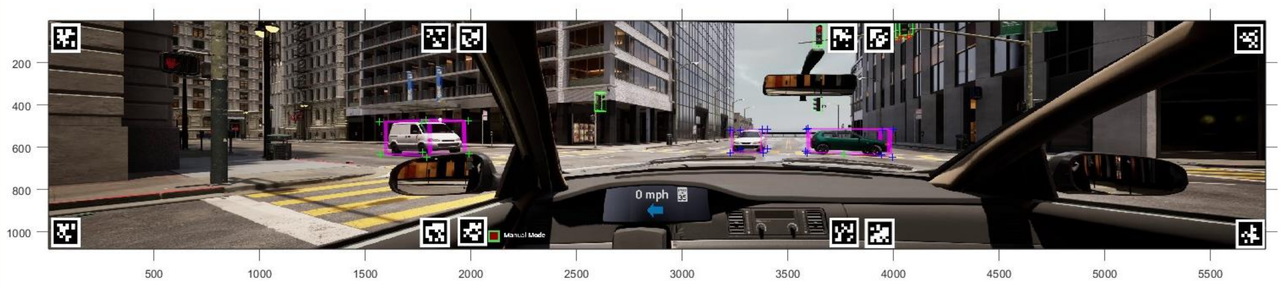

Fig. 6: An illustration of re-projection for cars

Conclusions

- A fully-automated dynamic AOI matching algorithm is developed that requires little specialized knowledge to implement
- This novel method has potential to enable real-time dynamic AOI matching that can be used to measure fixations and saccades online
  - Such measurements can be used as feedback variables in human-automation interaction scenarios

References

[1] Holmqvist, Kenneth, Richard Andersson, Richard Dewhurst, Halszka Jarodzka, and Joost Van de Weijer. Eye tracking: A comprehensive guide to methods and measures. oup Oxford, 2011.
[2] E. A. D. Lagmay and M. M. T. Rodrigo, "Enhanced Automatic Areas of Interest (AOI) Bounding Boxes Estimation Algorithm for Dynamic Eye-Tracking Stimuli," APSIPA Transactions on Signal and Information Processing, vol. 11, no. 1, 2022.
[3] J. Huang and Z.-N. Li, "Automatic Detection of Object of Interest and Tracking in Active Video," Journal of Signal Processing Systems, vol. 65, no. 1, pp. 49–62, Oct. 2011.
[4] K. Panetta, Q. Wan, S. Rajeev, A. Kaszowska, A. L. Gardony, K. Naranjo, H. A. Taylor, and S. Agaian, "ISeeColor: Method for Advanced Visual Analytics of Eye Tracking Data," IEEE Access, vol. 8, pp. 52 278–52 287, 2020.

Acknowledgements

This work was supported by the National Science Foundation under Award No. 2145827.
Any opinions, findings, and material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

PURDUE UNIVERSITY® | JAIN RESEARCH LAB