Active Visual Search in a 3D Real World Environment

Tiffany Wu (tiffwu@yorku.ca), John K. Tsotsos (tsotsos@yorku.ca)
Department of Electrical Engineering and Computer Science, York University, Toronto, Canada

Introduction

Active visual search in the real world has not been investigated in great detail.

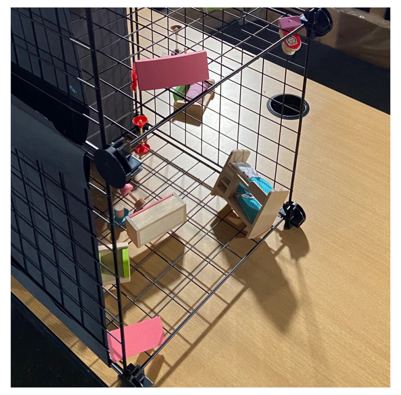
- Classic visual search paradigm focuses on the 2D, passive version of visual search
- Existing psychophysics studies on active search are rare, despite its importance
- Development of computational models are similarly lacking
- Active search has been more extensively researched in robotics, but with little relation to human search behavior

Investigating human strategies allows a better understanding of mechanisms behind human active search, with potential application in robotics.

Motivation

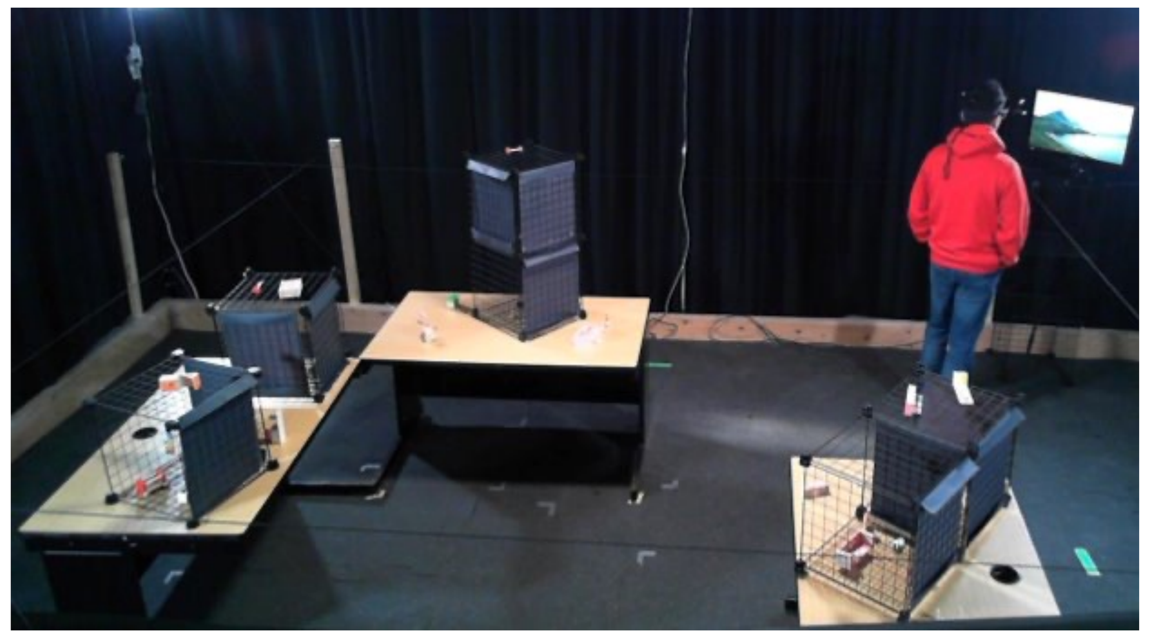
- Most psychophysics research has involved *passive search* conducted on *2D computer screens* with artificial stimuli [1,2]
- No computational models of human behavior exist for active search in a 3D environment (though models in robotics exist , e.g.[5,6])
- Active vision is a key component for performing search in the real world [3] to deal with problems of:
  - Viewpoint selection
  - Resolution
  - Visibility
  - Occlusion


Target


Occluded view


Not occluded view

- Few human active search experiments exist
  - Tracking eye gaze is insufficient to fully understand behavior


Experiment environment

Research Questions

How do eye and head movements during active search vary by target presence, visibility, difficulty, and set size?

Are there eye and head movement patterns that can be uncovered corresponding to different search strategies?

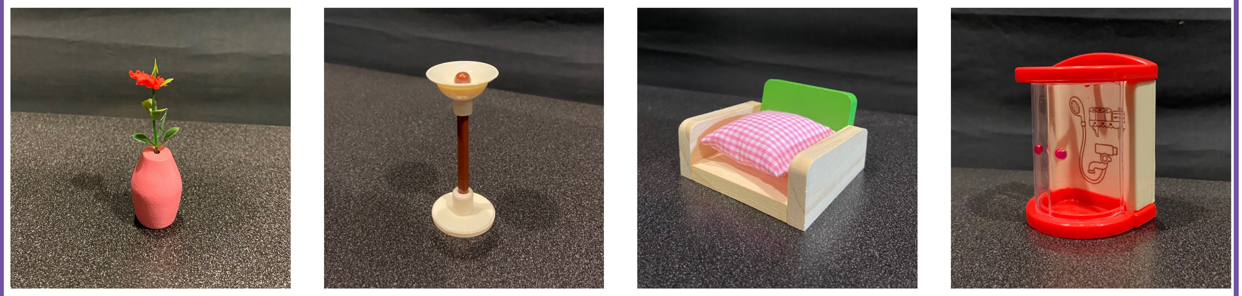
Methodology

Active visual search task

- Determine if specified target is present or not
  - Canonical image of target presented at start of each trial for 5 seconds
- Subjects can move around to search
- Respond by verbalizing ("I found it!" / "It's not here!")

Stimuli: dollhouse furniture of everyday objects

- Varying in color, material, size, shape



Experiment environment

- PESAO – Psychophysical Experiment Setup for Active Observers [4]
  - 6 Optitrack cameras for target and head movement tracking
  - Tobii Pro Glasses 2 for eye tracking
  - Software precisely synchronizes trackers
  - Subjects untethered and free
- Tables and cages arranged in pre-determined configurations
- Objects placed in pre-determined locations

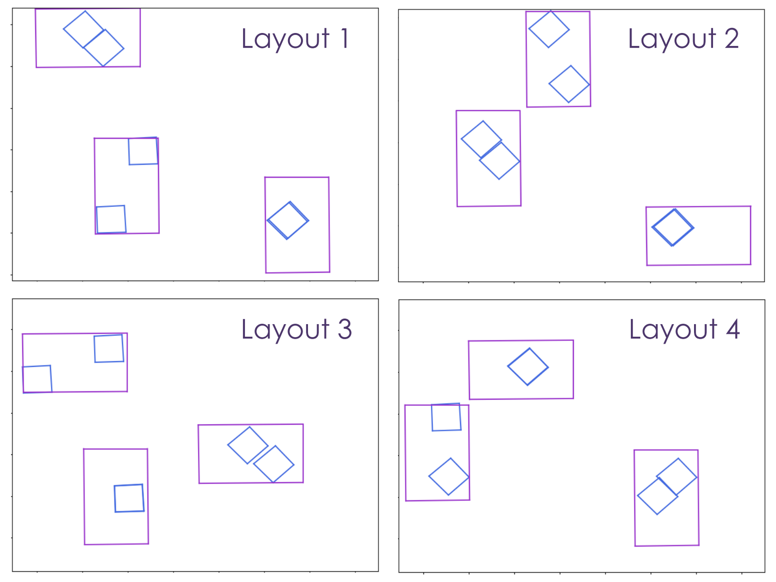
Independent variables:

- Target: presence (y/n), visibility (from starting location), orientation
- Trial difficulty (3 levels)
- Set size (4 levels)
- Layout : table and cage configurations
- Trials: objects + placements

Dependent variables:

- Response time
- Accuracy
- Eye and head movement data

4 layouts used in experiment:

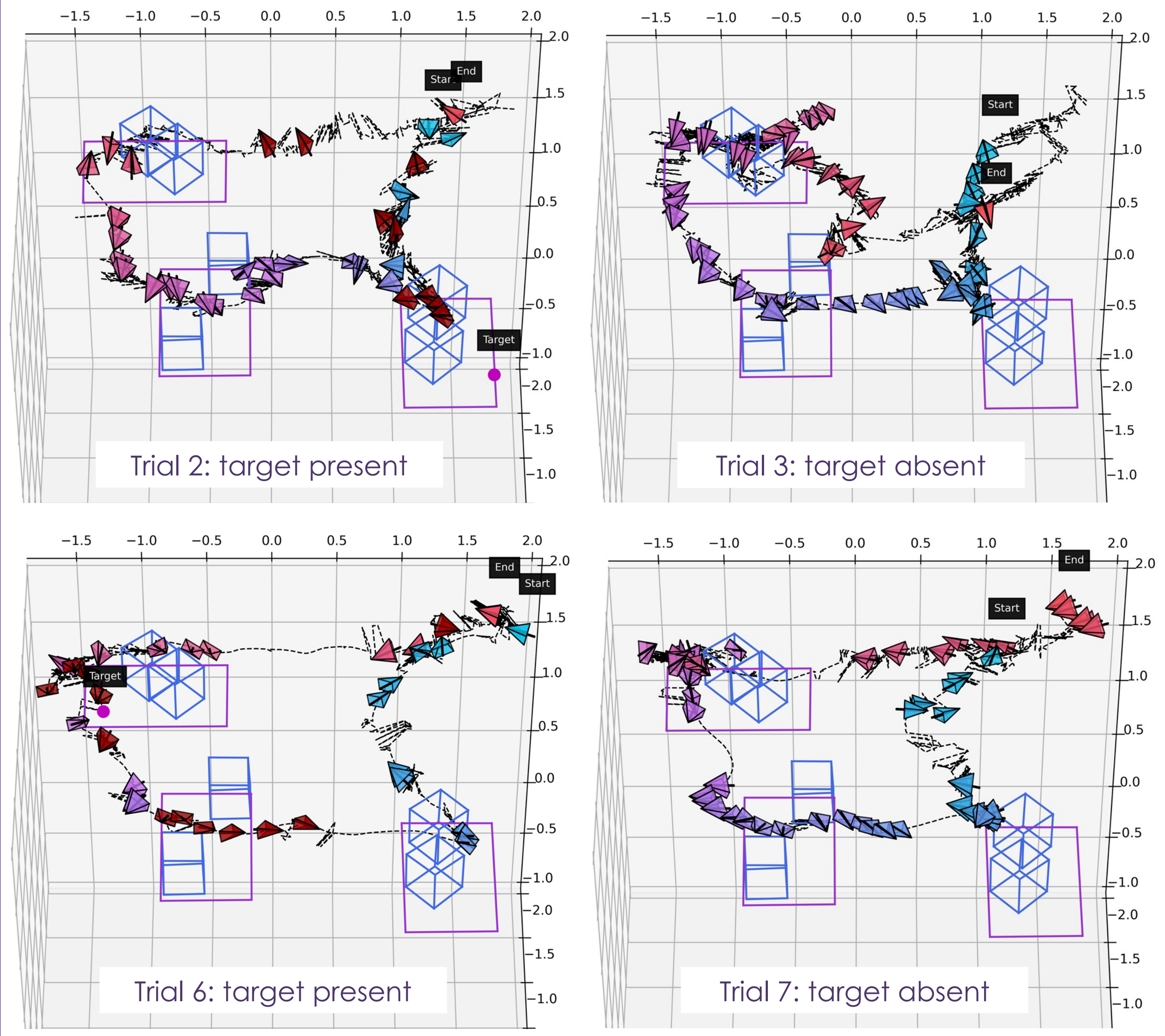


Results

Overall performance

N=48

- Counterbalancing trial conditions (4) by layouts (4)
- 3 subjects in each condition



Example of 1 subject's trial scanpaths; general trajectory does not change regardless of previous trial's target location

		RT (s)		# fixations		Accuracy (%)	
Target presence		Present	Absent	Present	Absent	Present	Absent
Correct	Mean	21.2	48.2	42.1	102.6	89.9 SD=13	96.5 SD=6.8
	SD	13.4	26.4	28.2	61.3		
	Count	259	278	259	278		
Incorrect	Mean	41.5	37	85.5	72.4		
	SD	17.3	7.8	39.3	75		
	Count	29	10	29	10		

Descriptive stats summary table

Target present trials

- Subjects take around 8 target-facing fixations (up to 30° eccentricity) before confirming target

"Miss" trials

- Trials which subjects incorrectly reported as absent had similar scanpaths and behavior to target absent trials

Strategy analysis

Metrics extracted from eye and head movement data:

- Fixations and saccades
- Crouching, head tilts
- Distance travelled, distance from target
- Fixation "look-at" locations
  - Revisits
  - Cage, table, wide scans

Initial strategy

Look around first, then focus in on 1 location

- At first location, on average:
  - 4.3 fixations
  - 2 locations looked at
- After 1 step:
  - 62% of fixations looked at a previously fixated location
- Subjects tended to use same initial strategy across trials (66.8%) - including where they looked first

Overall strategy

- Subjects tended to pick a path to investigate each location in order, consistently across trials

Implications

- New dataset of human active search behavior in an untethered 3D environment, with eye and head movement data
- 3D real world search shares some similarities with 2D passive search, but the difference in task and method make it hard to compare
- Insights into active search behavior can be uncovered from eye and head movement data
- Human active search strategies involve starting with a quick scan of the environment, before going on a path traversing possible target locations that remain consistent for the duration of the experiment

Next steps

- Different layouts: different number of tables and cages
- Change concentration of objects in tables and cages
  - Manipulate occlusion of objects
- Change initial view of trial:
  - Change starting point of subject
  - Change distance of tables and cages
- Show target object prompts in 3D instead of images

References

[1] Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, *12*(1), 97-136.

[2] Wolfe, J. M. (1998). What can 1 million trials tell us about visual search?. *Psychological Science*, *9*(1), 33-39.

[3] Bajcsy, R., Aloimonos, Y., & Tsotsos, J. K. (2018). Revisiting active perception. *Autonomous Robots*, *42*, 177-196.

[4] Solbach, M. D., & Tsotsos, J. K. (2020). PESAO: Psychophysical Experimental Setup for Active Observers. *arXiv preprint arXiv:2009.09933*.

[5] Shubina, K., & Tsotsos, J. K. (2010). Visual search for an object in a 3D environment using a mobile robot. *Computer Vision and Image Understanding*, *114*(5), 535-547.

[6] Rasouli, A., Lanillos, P., Cheng, G., & Tsotsos, J. K. (2020). Attention-based active visual search for mobile robots. *Autonomous Robots*, *44*(2), 131-146.