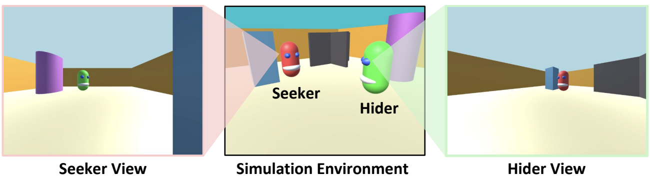
# Visual Hide and Seek

Boyuan Chen, Hod Lipson, Carl Vondrick
Columbia University
http://www.cs.columbia.edu/~bchen/visualhideseek/

## Visual Hide and Seek




We propose *Visual Hide and Seek*, to study the relationship between agents and their environment. In this game, a "prey" must run to avoid capture from a "predator," only relying on first-person visual observations. Our game enables agents: 1) to learn representations of other agents; 2) to leverage occlusions to solve their tasks.

## Methodology

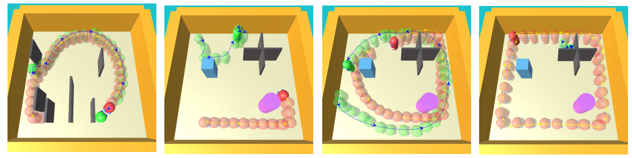
(A) Our agent is trained with Reinforcement Learning (PPO) to learn to hide. The hider gets a positive reward when it successfully survives, and a negative reward when it is captured. The objective of our algorithm is to the expected sum of future reward:

$$\max_{\theta_f \theta_p} \mathbb{E}_{\pi(\phi(o_t;\theta_f);\theta_p)}[\Sigma_t r_t]$$

(B) We perturb the environmental factors to systematically study how the behaviors evolve with different environment setup. We study why certain behaviors emerge and how to quantify behaviors.

## Learning to Hide

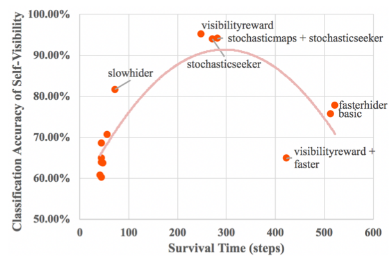
Agents learn very different hiding policies under various environmental conditions



Some agents learn to better infer the visibility of others, and its self-visibility to other agents

| Environment | Seeker Recognition ( /random init policy) | Awareness of Self-Visibility ( /random init polcy) |
|---|---|---|
| basic | $91.03 \pm 0.21$ / $79.25 \pm 2.42$ | $75.84 \pm 0.66$ / $63.92 \pm 0.50$ |
| fasterhider | $90.01 \pm 0.34$ / $74.84 \pm 1.91$ | $77.88 \pm 0.05$ / $60.30 \pm 0.13$ |
| visibilityreward + fasterhider | $79.19 \pm 1.11$ / $74.18 \pm 1.95$ | $65.05 \pm 1.88$ / $60.84 \pm 0.84$ |
| stochasticseeker | $96.17 \pm 0.25$ / $77.54 \pm 1.29$ | $94.00 \pm 0.58$ / $65.05 \pm 0.13$ |
| stochasticmaps + stochasticseeker | $95.55 \pm 0.38$ / $80.75 \pm 1.67$ | $94.25 \pm 0.75$ / $70.80 \pm 0.38$ |
| visibilityreward | $96.21 \pm 0.46$ / $77.65 \pm 0.15$ | $95.30 \pm 0.38$ / $63.88 \pm 2.30$ |
| slowerhider | $83.96 \pm 0.46$ / $77.79 \pm 0.54$ | $81.71 \pm 0.63$ / $68.59 \pm 0.34$ |

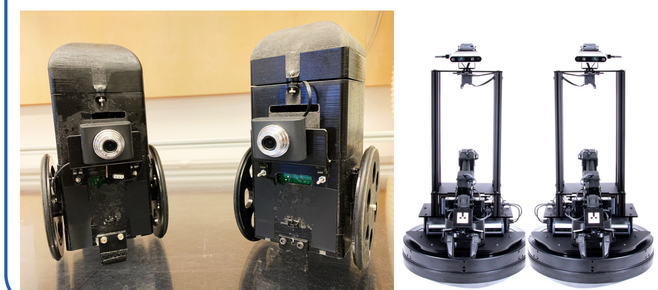
## The Cause of Emergent Features




We show that the best representations arise when the agent has intermediate disadvantages. When the agent has a weakness, the agent learned to overcome it by instead learning better features.

## Learning in the Wild

We plan to ground the visual hide and seek task onto real-world physical robots. Real-world environment allows us to incorporate physics into our system. It also requires investigating more-efficient learning algorithms.



## Potential Impact

Our work has the potential to create robots that ecologically leverage their environments to solve tasks without additional supervision. We have shown how occlusions and the agent's movement abilities change the learned representations of the agent itself and other agents. In the future, we plan to explore other incidental clues, such as sound. We also plan to explore applications with dynamic environments and deformable objects.

2020 National Robotics Initiative (NRI) Principal Investigators' Meeting
FEBRUARY 27 - 28, 2020 | ARLINGTON, VIRGINIA
Award ID#: 1925157