Sensory-Invariance Driven Action as a Method for Autonomous Grounding in a Pan-Tilt Camera

ERICA BOLAN, DEPARTMENT OF COMPUTER SCIENCE, CARLETON COLLEGE

FACULTY MENTOR: DR. YOONSUCK CHOE, DEPARMENT OF COMPUTER SCIENCE, TEXAS A&M UNIVERSITY

bolane@carleton.edu,choe@tamu.edu

Abstract

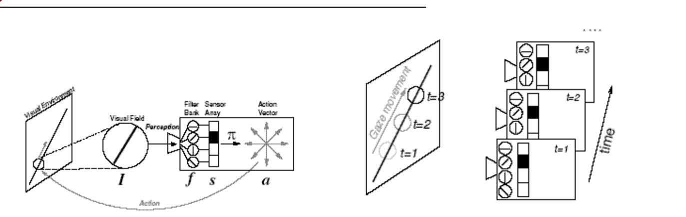
The creation of truly autonomous agents is one of the primary goals of Artificial Intelligence. A key characteristic of such an autonomous agent is the ability to represent its environment internally and to extract meanings and rules based solely on that representation. Internal representation of the environment is relatively straightforward, but deriving meaning from that representation in a truly autonomous fashion has proven difficult. One approach to this setback that has met relative succes is the Sensory-Invariance Driven Action (SIDA) algorithm (Choe & Bhamidipati (2004)). SIDA proposes that an agent can extract visual meaning from an internal representation of its environment, using sensory invariance as a criterion for directing the trajectory of its gaze. Previous work has demonstrated that this is a plausible and effective solution; however, it has been limited to synthetic and stationary images. Our work here integrates SIDA with a web-cam to demonstrate the algorithm's ability to provide agents with a method to interact with and extract meaning from a natural environment that is being continuously updated. The main contribution of this work is the integration of SIDA with an autonomous agent that acts on a natural environment in real time. This work also shows that the agent is capable of learning a set of rules based on a static, synthetic environment, and successfully applying those rules to a dynamic, natural environment.

Introduction

The Physical Agent

The main contribution of this research centers around the camera shown above, the Logitech OrbitView webcam. The camera has pan and tilt capabilities that allow it to be directed to shift its gaze around the environment. The SIDA algorithm was implemented using this camera as its autonomous agent.

The Agent and its Environment



(a) a simple agent with sensory filters and action primitives (b) an SIDA-driven agent interacting with a stimulus

The agent does not have direct access to the external stimuli and so must determine the structure of its environment based on its own internal representations of that environment. SIDA proposes that sensory invariance is a key step in this process. The agent determines its sensory state and acts in a manner intended to maintain that sensory state, allowing it to to construct an internal representation whose structure closely represents the structure of the environment.

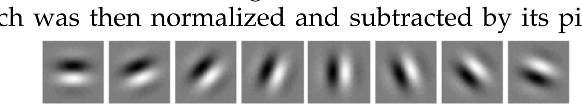
SIDA Algorithm

- Agent is assumed to have a given number n (8, 16, or 32) of sensory primitives and $2n$ corresponding motor primitives
- Agent is trained on a single natural image, which is preprocessed via Gaussian convulution
- Agent then uses a stochastic Q-learning policy to train, resulting in a reward table mapping sensory states to actions
- After training, the agent uses the reward table to explore and extract meaning from natural stimuli

Training

Preprocessing

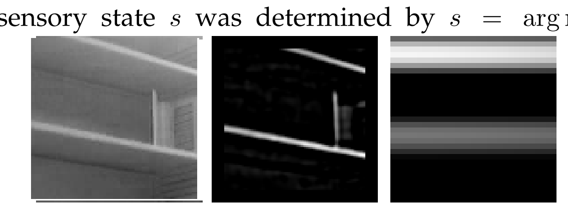
The following closely follows Choe & Smith (2006). Training input was a raw image of a natural scene, 320×240 pixels. This image was convolved with a Difference-of-Gaussian filter $D(x, y) = g_{(\sigma/2)}(x, y) - g_{(\sigma)}(x, y)$, where g_b is a Gaussian filter of width b. Filters were of size 15×15 for all experiments. The initial image I_R was convolved with $D(\cdot)$ to generate I_D, which was then normalized and subtracted by its pixelwise mean.



The agent's n sensory states were represented by a series of Gabor filters G_i of size $m \times m$), defined as:

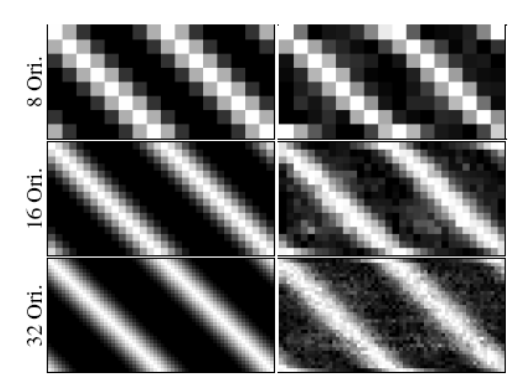
$$G_{\theta,\phi,\sigma,\omega} = \exp - \frac{x'^2 + y'^2}{\sigma^2} \cos(2\pi\omega x' + \phi), \tag{1}$$

with $\sigma = m/2$, $\phi = -\pi/2$, and $\omega = 2/m$, $x' = x\cos(\theta) + y\sin(\theta)$, and $y' = -x\sin(\theta) + y\cos(\theta)$. Values of θ were determined by $\theta_i = (i-1)\pi/n$, where i ranged from 1 to n. Filter response was a normalized vector s' containing the dot-products of the input block I and the Gabor filters G_i. Given a block of convolved input, the current sensory state s was determined by $s = \arg\max_{\theta_i, i=1..n} s'_i$.



The image block on the left (part of a bookshelf) yielded the DoG-filtered image seen in the middle, and matched most closely to the horizontally oriented Gabor filter.

Q Learning



(a) Ideal (b) Learned

(a) Choe & Smith (2006)

The reward table above was generated during offline training by allowing the agent to interact freely with its environment, recording the reward for each action taken. Reward for performing action a_t is $r_{t+1} = s'_t \cdot s'_{t+1}$, where s_t is the filter response vector. Complete invariance between two states results in maximum reward, or $r_{t+1} = 1$.

1. Given current sensory state s_t, randomly choose an action a_t.
2. If a_t equals $\arg\max_{a \in A} Q(s_t, a)$,
   (a) then perform action a_t,
   (b) else perform action a_t with probability proportional to $Q(s_t, a_t)$.
3. Repeat until an action has been performed.

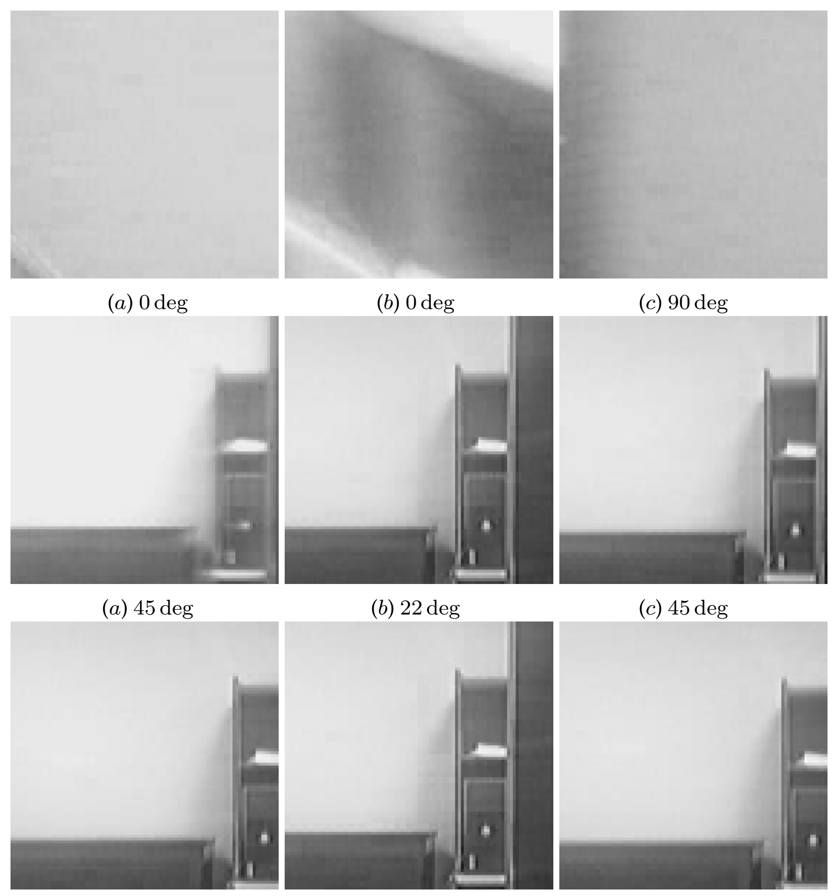
At each timestep t, given a current state (s, a), the agent chose an action a' and updated the table at $Q(s, a)$:

$$Q_t(s,a) := (1-\alpha_t)Q_{t-1}(s,a) + \alpha_t\left(r_t + .85\max_{a_i \in A} Q_{t-1}(s', a_i)\right), \tag{2}$$

where s' was the new sensory state reached via action a', r_t was the reward acheived via a' and $\alpha_t = \frac{1}{1+v_t(s,a)}$, with $v_t(s, a)$ representing the number of visits to $Q_{s,a}$.

Results

Image Blocks



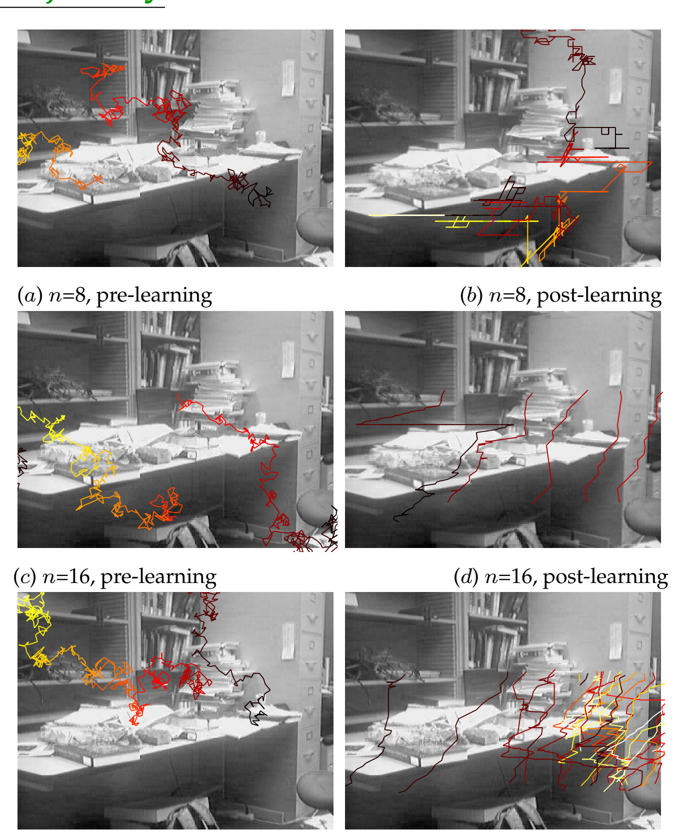
(a) 0 deg (b) 0 deg (c) 90 deg

(a) 45 deg (b) 22 deg (c) 45 deg

(a) 68 deg (b) 68 deg (c) 90 deg

Above are nine 93×93 image blocks, sampled in consecutive timesteps and classified using n=8 Gabor filter orientations. Below each image sample is the orientation classification assigned it by the learning algorithm. Very slight changes in input can trigger dissimilar filters.

Gaze Trajectory



(a) n=8, pre-learning (b) n=8, post-learning

(c) n=16, pre-learning (d) n=16, post-learning

(e) n=32, pre-learning (f) n=32, post-learning

The gaze trajectories after learning demonstrate more precise patterns and travel in the same direction for longer periods of time, whereas the trajectories before learning travel in random paths, appearing to follow no guidelines. Because the nature of the agent allows it to roam freely, the images on which the trajectories are overlaid are estimates, sample halfway through the exploration process.

Conclusion

Discussion

The reward and sensory state data gathered show that the algorithm allows the agent to maximize its reward by attempting to maintain its sensory state as it moves from gaze location to gaze location. Immediate reward from timestep to timestep is consistently high. The gaze trajectory data gathered shows that after the learning process, the agent is able to direct its gaze in longer, more consistent strokes that reflect the underlying stimuli. Adjusting as the input changes, the agent is able to autonomously choose its next action based only on its internal representation of the environment. Having no direct access to the external stimuli, the agent's actions correlate closely to the structure of the environment.

Contributions

Previous work has demonstrated SIDA's effectiveness in directing gaze trajectory in a static environment, based on a synthetic image, and provided a persuasive case for SIDA as a plausible method of autonomous grounding. The main contribution of this paper builds on the results of that work by integrating SIDA with an everyday real-time agent and providing concrete examples of its ability to glean structural knowledge from natural, dynamic input stimuli.

Acknowledgements

This work draws from previous work (Choe & Smith (2006),Choe & Bhamidipati (2004)) on the design, development, and testing of SIDA. The images of the agent model, raw and DoG-filtered leaves, Gabor filters, and reward tables were taken from Choe & Smith (2006). The new results presented in this poster are based on the reward tables generated by their work.

Resources

For more information, please visit:

- http://faculty.cs.tamu.edu/choe or
- http://people.carleton.edu/˜bolane

References

Choe, Y., and Bhamidipati, S. 2004. Autonomous acquisition of the meaning of sensory states through sensory-invariance driven action. In *Biologically Inspired Approaches to Advanced Information Technology*. Berlin: Springer. 176–188.

Choe, Y., and Smith, N. 2006. Motion-based autonomous grounding: Inferring external world properties from internal sensory states alone. In *Proceedings of the 21st National Conference on Artificial Intelligence*, 936–941.