

Learning the Sensorimotor Structure of the Foveated Retina

Jeremy Stober and Lewis Fishgold
Department of Computer Sciences
The University of Texas at Austin
{stober,lewfish}@cs.utexas.edu

Benjamin Kuipers
Computer Science and Engineering
University of Michigan
kuipers@umich.edu

Abstract

We identify two properties of the human vision system, the foveated retina, and the ability to saccade, and show how these two properties are sufficient to simultaneously learn the structure of receptive fields in the retina and a saccade policy that centers the fovea on points of interest in a scene.

We consider a novel learning algorithm under this model, *sensorimotor embedding*, which we evaluate using a simulated roving eye robot on synthetic and natural scenes, and physical pan/tilt camera. In each case we compare learned geometry to actual geometry, as well as the learned motor policy to the optimal motor policy. In both the simulated roving eye experiments and the physical pan/tilt camera, our algorithm is able to learn both an approximate sensor map and an effective saccade policy.

The developmental nature of sensorimotor embedding allows an agent to simultaneously adapt both geometry and policy to changes in the physical model and motor properties of the retina. We demonstrate adaptation in the case of retinal lesioning and motor map reversal.

1. Introduction

In the human eye, the retina is a non-uniform array of photoreceptive rod and cone cells. The human retina has a foveal pit, a single region of maximum density of cone photoreceptors. In addition, a human can change the location of the retina relative to a scene through ballistic actions known as saccades (Palmer, 1999). The combination of a small, high-resolution fovea with the ability to saccade to regions of interest is an economical strategy for both humans and robots to achieve high-resolution vision across large fields of view.

Gathering and interpreting visual information requires a *motor map* and a *sensor map* of the retina. The motor map encodes the motor commands necessary to move the eye to new locations in the visual scene and is used in generating saccades. The sensor map represents the geometric structure of the retina, specifically the positions of

sense elements within the sensor array, and can be used to perform geometric operations on the visual signal such as edge detection. We show how, by exploiting the relationship between motor commands and sensor geometry, an autonomous agent with foveated vision can *simultaneously* learn both the motor and sensor maps.

For simple sensors, these maps can be manually specified, but as sensors become more complex and adaptive, learning approaches such as ours are of increasing value to robotics. In addition, as lifetimes of autonomous robots increase, the robust nature of this developmental approach will allow robots to adapt to changing sensors and motors.

2. Related Work

2.1 Learning Motor Maps

In previous work on learning motor maps for saccades, the learning was driven by the two-dimensional difference between the pre-saccadic and post-saccadic position of a target on the retina. These models assume that the structure of the retina is known when learning the motor map, allowing calculation of the distance between a target and the fovea.

In (Pagel et al., 1998) the authors use learning to improve upon rough predictions made by first-principle geometric calculations. They represented the motor map using growing neural gas. Using a training scheme that involves corrective saccades, the agent experiences more training examples in the foveal region, causing an increase in the density of units in the region of the motor map that represents the fovea.

In (Rao and Ballard, 1995) the authors also used a strategy based on corrective saccades. They relied on the ability to locate a point of interest in the post-saccadic image using multiscale spatial filters, though the ability to locate interest points using this method may be too strong an assumption for a young infant with an immature visual cortex (Slater, 1999).

In (Shibata et al., 2001), the authors use fifth order splines and saliency maps (Itti and Koch, 2001) to generate realistic saccade trajectories and that closely resemble human motion. In this work, we opt for a simpler saccade

model that allows us to learn both sensor and motor maps simultaneously.

The model used in (Weber and Triesch, 2006) is one of the most recently published models and is the most similar to ours. Like us and unlike previous work, they use an error signal based on total retinal activation, exploiting cases where the total activation of a foveated retina is proportional to the degree of success of a saccade. Their model treats learning the horizontal and vertical components of saccades separately in accord with the experimental results of (Noto and Robinson, 2001).

2.2 Learning Sensor Maps

In previous work on learning sensor maps, (Pierce and Kuipers, 1997) demonstrated how sensor maps for a mobile robot can be discovered from uninterpreted high-dimensional sensor streams while motor babbling, and (Olsson et al., 2006) later extended these results to physical robots with visual perception. These studies generate sensor maps using dimensionality reduction algorithms that discover low-dimensional sensor arrangements that approximate distances between sensor trace histories. Two sensors are close in the sensor map if their corresponding sense histories are highly correlated.

In this work, we take a complementary but related approach and exploit some additional available structure, namely the availability of motor commands. We base our embedding, which we call *sensorimotor embedding*, on the motor system’s ability to change the sensory signal.

The algorithm we present here utilizes the relationship between sense and action to *simultaneously* extract useful geometric features (i.e. sensor position) along with primitive animate vision behaviors. Our method is appropriate for cases with an easily identifiable reward signal (e.g. activation), linear ballistic motor commands, and a high number of sense elements. We exploit the structure of the sensorimotor domain to produce an explicit mapping between motor commands and sensor features. This map has two interpretations, one as a primitive behavior that maximizes reward (the policy interpretation), and another as a structure for the sensor array (the geometric interpretation).

3. A Foveated Retina

3.1 Model

Our abstract model of the foveated retina is inspired by the anatomy of the human retina. In our model, a retina is a collection of receptive fields, or sense elements, with fixed geometry arrayed across a two dimensional surface. Each receptive field responds to sensory input from a portion of an image or scene according to its own activation function. Our learning rule requires that the distribution of activations across the retina be non-uniform and achieve a single maximum at the fovea. In addition, un-

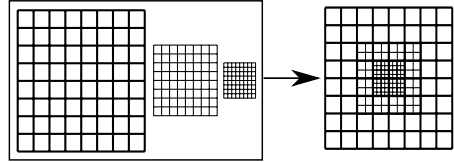


Figure 1: Our implementation of the fovea consists of overlapping layers of receptive fields. As the layer resolution increases, the extent of each receptive field decreases, and the number of bits necessary to describe the layer state remains constant.

der our model, ballistic motions instantaneously change the location of the retina in an image or scene.

Many implementations of a foveated retina satisfy this model. In biological systems, receptive fields are often distributed according to a log-polar distribution (Schwartz, 1977) and many computational models of saccade generation build upon this model of foveation (Weber and Triesch, 2006, Rao and Ballard, 1995). For this work, we view the specific distribution of receptive fields as an implementation issue, and expect that any distribution that satisfies the modeling assumptions above will behave similarly to our implementation.

3.2 Implementation

In our implementation, the learning agent has a foveated retina with N layers of receptive fields (Figure 1). Each layer has receptive fields of uniform extent and resolution. Layers with higher resolution and smaller extent overlap layers with lower resolution in the center of the retinal field of view. The fovea is the region with the highest concentration of overlapping receptive fields, and is also the region of maximal activation, so this implementation satisfies the model assumptions specified above. We stress that alternative implementations satisfying the model assumptions should behave similarly.

The implementation of each individual receptive field may also vary. In this case, each receptive field must map a patch of underlying pixel or sensor values to an activation level. Let I_k denote the image patch that affects the state of the k^{th} receptive field. Let \mathcal{I} denote the set of all such patches.

In addition to the image patch associated with each receptive field, the activation depends on the global state of the entire retina. In the case of a pan/tilt camera, we can describe the retina state using the horizontal and vertical angle of the camera lens (θ, ϕ) . In the case of the roving eye, we can describe the state of the retina in terms of the horizontal and vertical offsets (u, v) that describe the position of the retina in the larger image. However the state space is parametrized, we denote the set of all states by \mathcal{S} .

We require that the receptive field implement an activation function $\delta : \mathcal{I} \times \mathcal{S} \rightarrow [0, 1]$. In our implementation, $\delta(I_k, s)$ is the total activation of the pixels in the image patch I_k given the current retina state s , normalized to

$[0, 1]$ as a fraction of the maximum possible activation.

The activation over the entire retina is the sum of the activations for each receptive field for the current retina state,

$$R_{\mathcal{I}}(s) = \sum_{I_k \in \mathcal{I}} \delta(I_k, s) \quad (1)$$

4. Reinforcement Learning Problem

In our computational model, saccades result in 2D displacements of the image on the retina or pan/tilt changes for a physical camera. Each action or saccade $a : \mathcal{S} \rightarrow \mathcal{S}$ is described by two-element vector denoting horizontal and vertical motion and results in a single globally rigid transformation of the image or scene.

If the receptive fields in the retina are of uniform size and distribution, and they are exposed to input consisting of a small spot of light against a uniform background, then $R_{\mathcal{I}}(s)$ would be approximately constant for all retinal states s , regardless of where the spot of light falls. However, with a *foveated* retina, $R_{\mathcal{I}}(s)$ will have a dramatic maximum for retina states that cause the spot of light to fall on the fovea, due to the larger density of receptive fields there.

Using the total activation of all the receptive fields for the current retina state, $R_{\mathcal{I}}(s)$ in Equation 1 as the reward, combined with saccade actions, we can define a simple reinforcement learning problem, the goal of which is to find a policy, or choice of action, that maximizes retinal activation.

We factor the global learning problem into an individual learning problem for each receptive field. The goal of each receptive field is to learn a policy that greedily maximizes the total retinal activation $R_{\mathcal{I}}(s)$,

$$\pi_k(s) = \arg_a \max R_{\mathcal{I}}(a(s)) \quad (2)$$

The problem is episodic and spans a pre- and post-saccade state. The collective policy π^* for the entire retina is the weighted average of the actions preferred by the individual receptive fields,

$$\pi^*(s) = \frac{1}{R_{\mathcal{I}}(s)} \sum_{I_k \in \mathcal{I}} \delta(I_k, s) \cdot \pi_k(s) \quad (3)$$

In this factored learning problem, the only information a receptive field has about the state of the retina is the intensity level for that receptive field's visible patch I_k . If the intensity is high ($\delta(I_k, s)$ is close to 1), then the policy $\pi_k(s)$ will have a large impact on the global policy calculated in Equation 3. In this case, we want the policy to suggest an action $\pi_k(s) = a$ that maximizes the reward $R_{\mathcal{I}}(a(s))$. The action that accomplishes this takes the activation that the current receptive field sees and shifts it to the fovea, where the density of receptive fields is higher.

If the intensity is low, then the policy for that receptive field will have little impact on the policy for the entire retina since $\delta(I_k, s)$ is close to zero. As a consequence,

we can treat $\pi_k(s)$ as a constant. So in the factored problem, each receptive field only needs to estimate the optimal action and observe its own intensity level.

We predict that (after sufficient training), the action specified by π_k will approximate the saccade that moves an image-point from receptive field k directly to the fovea. Consider the inverse $-\pi_k$ of the policy estimate for each receptive field. This is the action that would move an image-point from the fovea to the receptive field k . In other words, the inverse of the policy is a position for the receptive field relative to the fovea. We expect that physically proximate receptive fields will have similar saccade policies, and hence similar learned positions. Note that we have not used any knowledge of the location of receptive fields within the fovea. In fact, that knowledge has been learned by the training process, and is encoded in the policy π_k . Spatial knowledge that was implicit in the anatomical structure of the retina becomes explicit in the policy.

The reinforcement learning problem described above has two unusual properties that constrain the choice of learning algorithm. First, the action space is continuous (as opposed to small and discrete). Second, the problem is episodic, and each episode spans only one choice of action.

During learning, each receptive field maintains an estimate for π_k , the current best action, and R_k , the current maximum estimated reward after performing the current best action. Initially, each π_k is set to a random action, and the reward estimate is initialized to zero.

At the beginning of each iteration or training, we randomly reposition the retina. For exploration, some noise ϵ is added to the current greedy policy. The retina agent executes $\pi^*(s) + \epsilon$, and measures the reward (R). Each individual receptive field's reward estimate and current policy are updated proportional to its state activation prior to the saccade ($\delta_k = \delta(I_k, s)$) since the optimal policy π^* is weighted according to those activations. We use a moving average learning rule to update both the reward estimate and current policy. For each receptive field k , we update the reward as follows

$$R_k^{new} = \begin{cases} R_k^{old} + \delta_k \cdot \alpha \cdot (R - R_k^{old}) & \text{if } R > R_k^{old} \\ R_k^{old} & \text{otherwise} \end{cases} \quad (4)$$

If the reward received, R , is greater than our current reward estimate, we move the current policy π_k for that receptive field closer to the global policy responsible for the increased reward

$$\pi_k^{new} = \pi_k^{old} + \delta_k \cdot \alpha \cdot (\pi^* - \pi_k^{old}) \quad (5)$$

By varying the learning rate α , we can change how much recent experience affects both the estimate of reward (R_k) and the estimate of the optimal saccade (π_k) itself. We discuss cases where R_k may decrease in Sections 5.2 and 5.3.

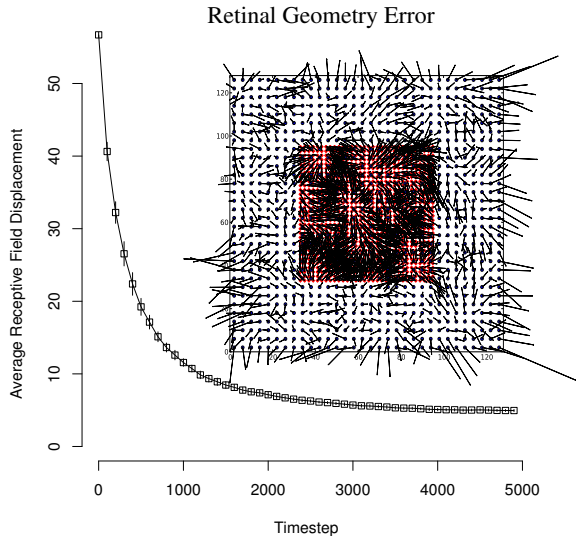


Figure 2: This figure plots the mean geometric error as a function of training time. The mean and standard errors are shown for ten independent training runs using a single dot image. The subfigure shows the result of interpreting learned receptive field policies as positions. Each line represents the error between the true position and learned position — the head (dot or diamond depending on the layer) is the true location of the field. The tail is the learned position. For clarity, only two layers are shown.

5. Experimental Evaluation

5.1 Simulated Saliency

We trained a simulated foveated retina with four layers of receptive fields on an image with a single white spot on a black background, meant to simulate the result of a saliency map. Each retina layer contained 32x32 receptive fields. The extent of each receptive field varied by layer, with the largest layer having receptive fields of size 4x4 (for a total retinal pixel area of 128x128). Actions corresponded to horizontal and vertical translations of the retina across the image.

We randomly initialized the policy for each receptive field and used a training rate $\alpha = 0.5$. ϵ was normally distributed with a mean of 0 and a standard deviation of 10 pixels.

We use two criteria to measure the success of our learning algorithm. The first computes the mean of the Euclidean distances between the learned position (interpreted as the additive inverse of the policy) and the true position $pos(I_k)$ of all receptive fields (Equation 6).¹ The results of training are shown in Figure 2.

$$E_{geometry} = \frac{1}{N} \sum_{k=1}^N \| -\pi_k - pos(I_k) \|_2 \quad (6)$$

For the second criterion, we compare the accuracy of the learned saccade against the optimal saccade, which

¹This analysis compares pixel positions to action space positions. This is only possible since translations of the roving eye retina are specified in pixels. In experiments using a pan/tilt camera, we do not have the same access to error free ground truth actions.

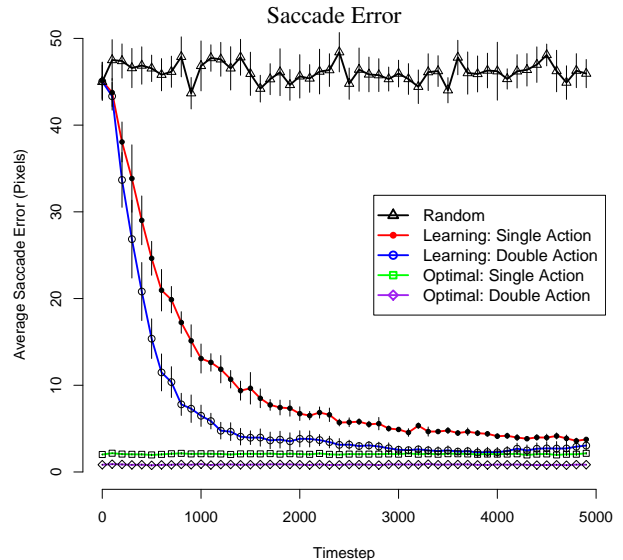


Figure 3: The saccade error as a function of the number of training iterations using the learning algorithm of Section 4. The saccade error is computed over thirty random repositions every 100 timesteps for ten independent trials. Note that even with an optimal policy, saccades are not entirely accurate because of low resolution in the periphery of the retina.

would center the retina on the area of high activation. We also test two-saccade accuracy, where the retina makes a second saccade after the first during testing but not training.

During the training process, every 100 training steps, we stop training and test saccade and two saccade accuracy for 30 random repositions. The average and standard errors of the accuracies over ten training trials are shown in Figure 3, which also includes comparisons with a randomly initialized policy and an optimal policy (where each policy is initialized to the inverse of that receptive field's position).

The learning algorithm achieves near-optimal saccade accuracy after 5000 training steps. Comparing Figures 2 and 3, we see that the geometric error decreases as accuracy increases, though the final sensor map only approximates the true positions of the receptive fields. Our algorithms final saccade error of 5 pixels is less than that of (Pagel et al., 1998) and requires only a quarter of the number of training steps.

5.2 Lesioning

In natural scenes, or in cases where the number of receptive fields in the fovea changes as with macular degeneration, the maximum achievable reward changes. In these cases, the maximum achievable reward may decrease to a level below the current reward estimate for each receptive field, $R < R_k^{old}$ and so no updates will take place. To account for this kind of variation over time, we can change the learning rule to maintain a recency-weighted average estimated reward, instead of maintaining an estimate of maximum reward.

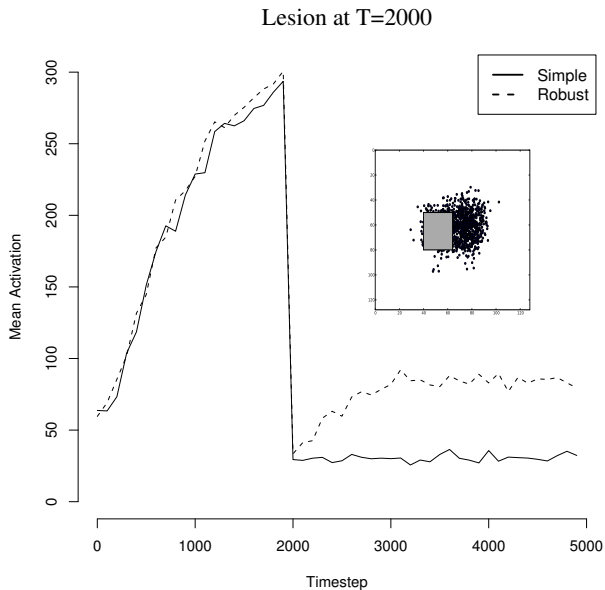


Figure 4: As a result of lesioning, a retina, with a robust learning rule as described in this section, adapts its policy to favor saccades to regions just outside the damaged region (see subfigure), providing higher post-saccadic activation in the case of lesioning than the previous optimal saccades directly to the fovea. We note that this increases the position error relative to the ground truth, but provides a coordinate system consistent with the sensorimotor properties of the damaged retina. The basic learning rule from Section 4 fails to adapt following a lesioning event.

This learning rule would require that the reward estimate be updated each timestep

$$R_k^{new} = R_k^{old} + \delta_k \cdot \alpha \cdot (R - R_k^{old}) \quad (7)$$

instead of only updating during timesteps where $R > R_k^{old}$.

We tested the ability of this modified algorithm to adapt to lesioning a small off-center part of the foveal region of the retina after 2000 steps of normal training. The mean post-saccade activation increases after lesioning when the agent uses the the robust learning rule (Figure 4). The basic learning rule, however, does not adapt to the lesioning event.

5.3 Motor Map Reversal

The modified algorithm presented above to deal with lesioning may require very high sample complexity to properly adapt to large changes in the motor model of the foveated retina.

Even though the reward estimates for each receptive field would adjust downward after a large change in the semantics of the motor commands, exploration still depends on adding noise to the previous policy estimate for each receptive field. In cases where the motor model changes radically, this exploratory bias may handicap any attempt to learn an alternative motor map.

Humans have shown some capacity for adapting to drastic changes in sensorimotor experience. For exam-

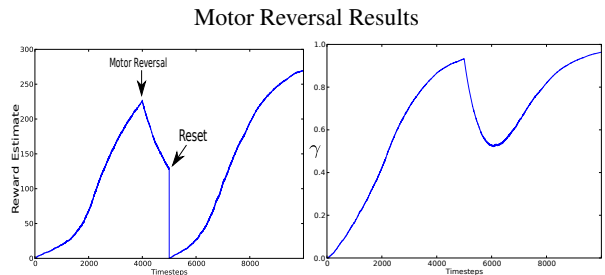


Figure 5: The left figure shows the moving average estimate of rewards experienced during training. A reversal in the motor map occurs after 4000 timesteps results in a decrease in the moving average reward estimate. After decreasing over 1000 timesteps, the retina resets the rewards estimate and the estimates for each receptive field and begins adapting to the new motor model. This results in a decrease in γ and an increase in exploration as shown on the right.

ple, in a self study using prismatic inverting eye-wear (Dolezal, 1982), Dolezal reports both initial difficulty in simple reaching tasks followed later by comfortable mastery.

In Dolezal’s inverted perceptual world, pointing up results in the visual perception of pointing down. By reversing the result of a motor command along one axis, we can simulate a similar (but less complex) change in the relationship between the motor actions and perceptual response. Though our experiment does not capture the full range of altered sensorimotor contingencies presented in (Dolezal, 1982), this experiment illustrates the need for a different kind of adaption in the face of significant changes in sensorimotor contingencies.

In this modification, each receptive field maintains an estimate of the optimal reward and policy as before. The retina also maintains an estimate of the maximum observed reward, a moving average of all the observed rewards, along with the reward estimates associated with each receptive field.

The exploration/exploitation trade-off is driven by a parameter, γ , that is meant to measure the extent to which the learned policy for currently active receptive fields will be able to achieve the maximum observable reward as estimated by the retina as a whole.

For a given pre-saccade retina state s , we compute both the current action estimate a and the reward estimate r_a . γ is then the ratio of r_a to r_{max} , the maximum observed reward for the entire retina. Intuitively, if r_a is close to r_{max} then the action a is likely close to optimal, and so little exploration is necessary. Similarly, if r_a is less than r_{max} , the action a is likely suboptimal, and so more exploration is required. The actual action taken is

$$\gamma a + (1 - \gamma) a_{exp}$$

where a_{exp} is a random saccade.

We use a large negative change in the moving average of all the rewards as an indicator of a major change to the retina motor or sensor map (Figure 5). When detecting

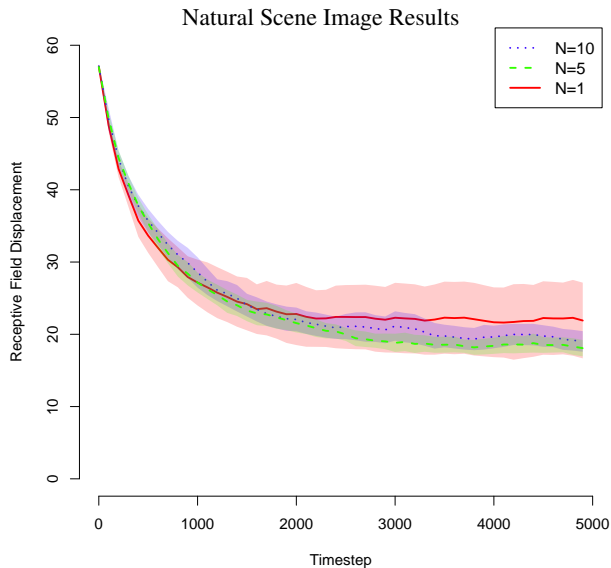


Figure 6: For this experiment, subsets of natural scene images were chosen randomly. This graph shows the mean and variance of ten runs for each subset size and is best viewed in color. Training across sets of images results in more consistent learning curves than training over single images, since the variance is smaller for training that takes place across subsets. Even in the single image case (where each run drew training examples from a single image) the mean learning curve was qualitatively similar to the others, but the high variance suggests that some images are “bad” sources of training examples.

this kind of change, the retina resets the reward estimates of all the receptive fields to their original values. This significantly decreases γ , triggering an increase in exploration and decreasing the contribution of the previously learned policy.

5.4 Natural Scenes

To recapture the features of the single spot case in natural scenes, we construct a *proto-saliency* map from natural scenes by first blurring the image under the retina using a Gaussian blur with a 5×5 filter size², then thresholding the image and taking pixels that fall into the top one percent brightness level in the region under the retina. If the number of active pixels is less than 500 pixels, we proceed to train on that portion of the image, otherwise the agent performs a new random saccade without training. This is to avoid training in situations of homogeneous brightness that wash out any existing progress on learning the optimal policy.

We note that humans tend to avoid saccading to areas of high luminance at low spatial scales (e.g. sky, solid colors) (Tatler et al., 2005). By avoiding training when the number of active pixels after thresholding is too high, we avoid training on precisely these kinds of high-

²Blurring is incompatible with the assumption that geometric information is not available. However, this blurring step is meant to simulate the optical characteristics of infants during early development (Slater, 1999), not infant visual processing.

luminance inputs.

Due to the variation in learning performance across images, we examine how the learning process behaves when trained over subsets of images randomly chosen from the Berkeley segmentation dataset (Martin et al., 2001). For each run, we select a set of images ($N=1, 5$ or 10) to train over. We cycle through the images, training 19 times over each image before moving to the next image in the cycle to continue training. As before, we evaluate the learning performance by measuring geometric errors every 100 timesteps of training. The results are shown in Figure 6.

Even though the final error rates are higher than when trained with the synthetic scene (Section 5.1), we note that the fixed point behavior of the policy (allowing repeated corrective saccades) does result in accuracy comparable to what training achieves on an ideal version of a saliency map after a similar number of training steps. The following table shows the accuracy after one and two saccades, as well as after the number needed to reach a fixed point (or in rare cases, a cycle – in which case the closest cycle point is counted).

1 Saccade	2 Saccades	Fixed Point
20.4	12.5	7.6

5.5 Pan/Tilt Camera

For the physical pan/tilt experimental setup, we used a Logitech QuickCam Orbit AF placed 15 feet from a single light source. To reduce training time, we modified the exploration policy to search randomly for a bright light. The agent performs a random saccade away from the light source. During training the agent then performs the opposite saccade back towards the light source, and uses the resulting retinal activations to learn a function from field activation to optimal saccades using the algorithm described in Section 4 with the proto-saliency method as described in Section 5.4. Unlike a learned policy, this open-loop training policy cannot account for relocation of the salient light source.

Figure 7 shows the decrease in saccade error and the increase in post-saccade reward (or activation) after intervals of 100 training steps. Each data point is the mean of 10 test trials. Each trial randomly saccades away from the light source, then computes the return saccade as the activation weighted average of the learned receptive field policies. For a trained retina, the post-saccade reward is independent of the initial random saccade, since the state of highest reward is reachable from any random starting position.

In our simulation experiments, the learned policies correspond to ground truth pixel geometry, since actions for the simulated roving eye camera are pixel unit translations over an image. The action space of the pan/tilt camera, however, is not represented in pixel unit shifts. The motor commands represent control signals sent directly to the piezoelectric motors in the camera appara-

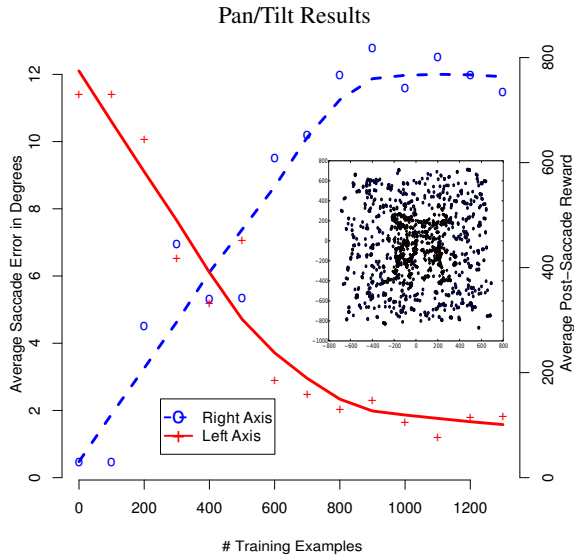


Figure 7: Every 100 training timesteps, we perform 10 test trials with the pan/tilt camera, randomly saccading away from the light source, then using the learned saccade policy to attempt to recenter on the light source (as opposed to using the inverse of the random saccade as in training). As training progresses, each receptive field learns a policy that centers local activation at the fovea resulting in greater post-saccade reward (dashed line) and lower saccade error (solid line). The subfigure shows the corresponding action space coordinates of each receptive field for two different layers of receptive fields after training.

tus. Camera geometry, along with irregularities in camera control, make the correspondence between motor signals and pixel shifts in the field of view necessarily inexact. We made no attempts to improve the correspondence through any alternative method of system identification beyond running our algorithm.

As a result of the learning process, for each region of interest we have access to the motor coordinates that center the camera on the region of interest. The geometry of these action space coordinates approximates (up to a scale factor) the ground truth geometry of the receptive fields in pixels.

Our approach is not limited to finding a sensor map in the coordinate system of the action space. With access to the ground truth pixel geometry for each receptive field, we can also construct a map from ground truth pixel coordinates to the corresponding action space coordinates, providing the ability to switch between pixel and motor geometry as a method of controlling the pan tilt camera. Selecting pixel coordinates (and activating the corresponding receptive fields) for a region of interest is sufficient to generate the corresponding motor mapping that brings those pixels to the center of the field of view. In other words, the learning algorithm autonomously provides a method for going from pan/tilt (or joystick) control, to point and click control in the view frame.

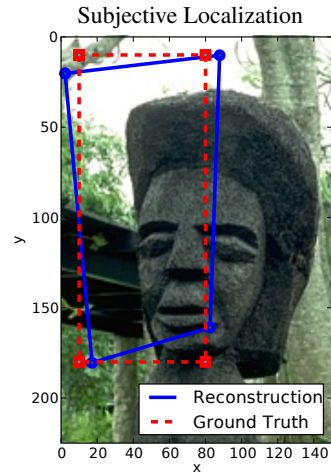


Figure 8: The results of localization in a roving eye domain. A roving eye was able to use features and their associated policies learned through sensorimotor embedding to reconstruct a visual path.

6. Future Directions

Sensorimotor embedding can be applied to other types of structure discovery problems. As an example, an agent can use sensorimotor embedding to visually localize by associating sensor inputs with ballistic actions that bring about desired changes in sensor state. This provides an alternative to *action respecting embedding* (Bowling et al., 2007) in continuous action spaces.

We applied sensorimotor embedding to the “roving eye” domain by first generating a set of 50 principle component basis vectors using random samples of a scene. We then formed a feature set consisting of principle projections of random samples onto these principle components. Associated with each feature is a reward and ballistic policy estimate just like the receptive fields described above.

During training, the projection of each eye image is compared to each feature. The winning feature determines the next (noisy) action. After each action, the reward is the least of the inverse of the distance to a predefined point in the scene or one. Updates to reward and policy estimates are the same as in Section 4. Once trained, a sequence of images can be embedded directly in the learned motor space by comparing each images projection with the feature set. An example embedding for a visual path of a roving eye is shown in Figure 8.

7. Discussion

Our experimental results confirm that, under simple assumptions, an agent can simultaneously discover motor and sensor maps for a foveated retina. Like Weber and Triesch, we use total activation as a reward signal to learn the motor map; however, we demonstrate the ability to learn without prior knowledge of the sensor map. To do so, we generate a proto-saliency map directly from natural scenes in a geometry-free way. After learning

the motor map, we generate the sensor map by exploiting the relationship between sensor geometry and motor commands. Previous approaches to sensor map construction use dimensionality reduction techniques and do not exploit additional available domain structure, namely access to motor commands.

Representing the sensor map in motor units may appear to be a limitation of the approach. However, in the absence of some external system identification, we would expect that a developmental agent would have difficulty discovering sensor geometry in units *other* than those which correspond in some way to motor semantics.

Our method is appropriate for cases with an easily identifiable reward signal (e.g. activation), linear ballistic motor commands, and a high number of sense elements. We exploit the structure of the sensorimotor domain to produce an explicit mapping between motor commands and sensor features. This map has two interpretations, one as a primitive behavior that maximizes reward (the policy or motor map interpretation), and another as a structure for the sensor array (the geometry or sensor map interpretation).

The *sensorimotor embedding* algorithm we present above, and the general approach of utilizing action spaces to better understand sensor spaces represents a fundamental first step in building a computational model of vision that follows the “seeing is acting” paradigm (O’Regan and Noë, 2001).

Any developmental process or autonomous robot depends on robust sensorimotor primitives that can adapt to changes over time. We demonstrate the robustness of our learning process under both lesioning and motor map reversal. We believe that focusing on associating structure with motor commands that bring about desirable changes in perceptual state, as in foveated retina and localization, will result in precisely the kind of robust sensorimotor primitives required for autonomous developmental robots.

Acknowledgments

This work has taken place in the Intelligent Robotics Lab at the Artificial Intelligence Laboratory, The University of Texas at Austin. Research of the Intelligent Robotics lab is supported in part by grants from the Texas Advanced Research Program (3658-0170-2007), from the National Science Foundation (IIS-0413257, IIS-0713150, and IIS-0750011), and from the National Institutes of Health (EY016089).

References

Bowling, M., Wilkinson, D., Ghodsi, A., and Milstein, A. (2007). Subjective localization with action respecting embedding. *Robotics Research*, 28:190–202.

Dolezal, H. (1982). *Living in a World Transformed: Per-*

ceptual and Performatory Adaptation to Visual Distortion. Academic Press.

Itti, L. and Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2(3):194–203.

Martin, D., Fowlkes, C., Tal, D., and Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proc. 8th Int’l Conf. Computer Vision*, volume 2, pages 416–423.

Noto, C. and Robinson, F. (2001). Visual error is the stimulus for saccade gain adaptation. *Cognitive Brain Research*, 12(2):301–305.

Olsson, L. A., Nehaniv, C. L., and Polani, D. (2006). From unknown sensors and actuators to actions grounded in sensorimotor perceptions. *Connection Science*, 18(2):121–144.

O’Regan, J. and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(05):939–973.

Pagel, M., Maël, E., and von der Malsburg, C. (1998). Self calibration of the fixation movement of a stereo camera head. *Autonomous Robots*, 5(3):355–367.

Palmer, S. (1999). *Vision science: photons to phenomenology*. MIT Press, Cambridge, MA.

Pierce, D. and Kuipers, B. J. (1997). Map learning with uninterpreted sensors and effectors. *Artificial Intelligence*, 92(1-2):169–227.

Rao, R. P. N. and Ballard, D. H. (1995). Learning saccadic eye movements using multiscale spatial filters. *Advances in Neural Information Processing Systems*, 7:893–900.

Schwartz, E. (1977). Spatial mapping in the primate sensory projection: Analytic structure and relevance to perception. *Biological Cybernetics*, 25(4):181–194.

Shibata, T., Vijayakumar, S., Conradt, J., and Schaal, S. (2001). Biomimetic oculomotor control. *Adaptive Behavior*, 9(3-4):189–208.

Slater, A. (1999). *Perceptual Development: Visual, Auditory and Speech Perception in Infancy*. Psychology Press.

Tatler, B., Baddeley, R., and Gilchrist, I. (2005). Visual correlates of fixation selection: effects of scale and time. *Vision Research*, 45(5):643–659.

Weber, C. and Triesch, J. (2006). A possible representation of reward in the learning of saccades. *Proceedings of the Sixth International Workshop on Epigenetic Robotics*, pages 153–60.