Learning What the Internal State Means, Through Action

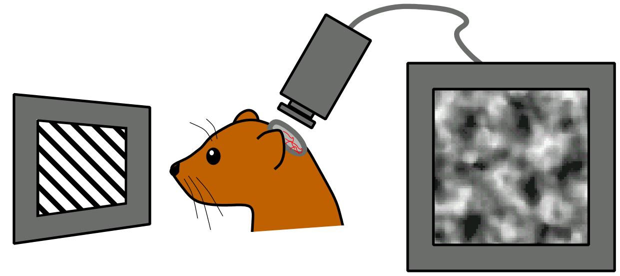
Yoonsuck Choe

Department of Computer Science

Texas A&M University

† Joint work with S. Kumar Bhamidipati, Daniel Eng, Navendu Misra, Stuart B. Heinrich, Noah H. Smith, and Huei-Fang Yang

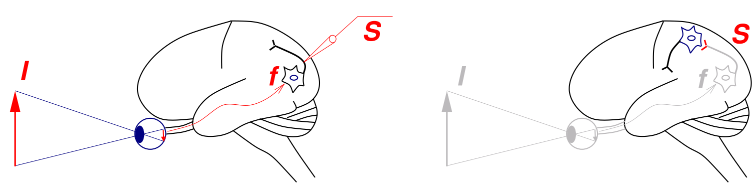
The Main Question



How can we understand what the **pattern of activity** in the brain **means**? (cf. Freeman 2003)

1. How can **scientists** understand the pattern?
2. How does **the brain itself** make sense of its **own activity**?

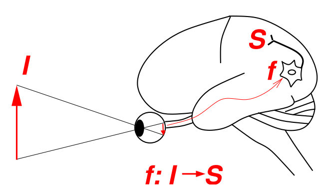
Scientist vs. the Brain



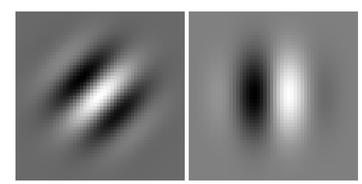
(a) External observer (b) Internal observer

- External observer (e.g., a neuroscientist) **can figure out** how S relates to I (transformation $f : I \rightarrow S$).
- Internal observer **cannot**: But the brain does this all the time, so this does not seem right!

Example: The Visual Cortex



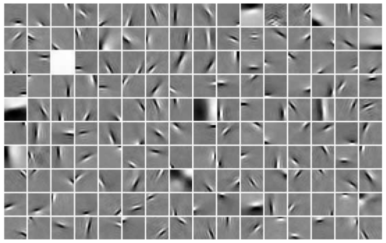
V1 Response to Input



Gabor-like RFs

- With access to both I and S, Hubel and Wiesel (1959) figured out $f : I \rightarrow S$ in V1 (oriented Gabor-like receptive fields Jones and Palmer 1987).
- But even before that, and with access to only S, humans had no problem perceiving orientation.

Theories on RF Formation



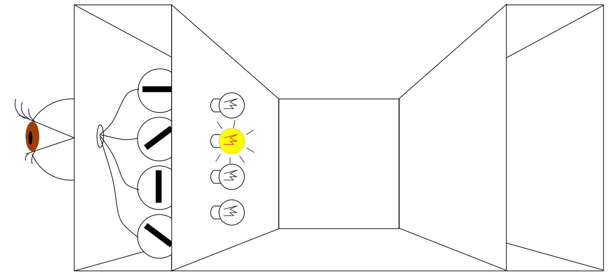
Hoyer and Hyvärinen (2000)

Well-developed understanding on how RFs form:

- Olshausen and Field (1997): Sparse coding; Barlow (1994): Redundancy reduction; Bell and Sejnowski (1997): Information maximization; Miikkulainen et al. (2005): Self-organization through Hebbian learning.

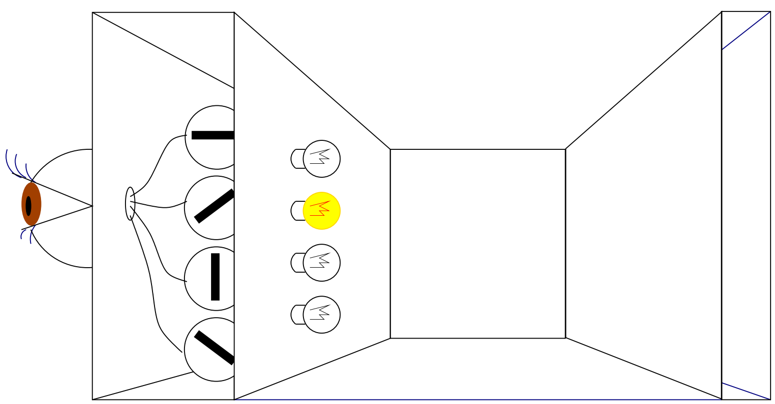
However, how is the resulting code to be used remains a question.

A Metaphor of the Problem



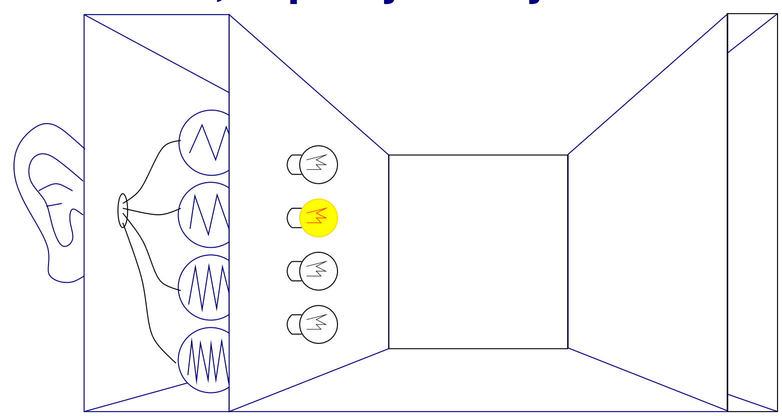
- Imagine sitting in a room, looking at blinking lights, without knowledge of the sensors nor the RFs.
- The lights may be due to any other sensory modality (as in vision-audition rewiring Sur et al. 1999).
- Similar to the **Chinese Room** (Searle 1980): Problem of "**Symbol Grounding**" (Harnad 1990).

The Sensory Organ Can (Possibly) Give Us a Clue



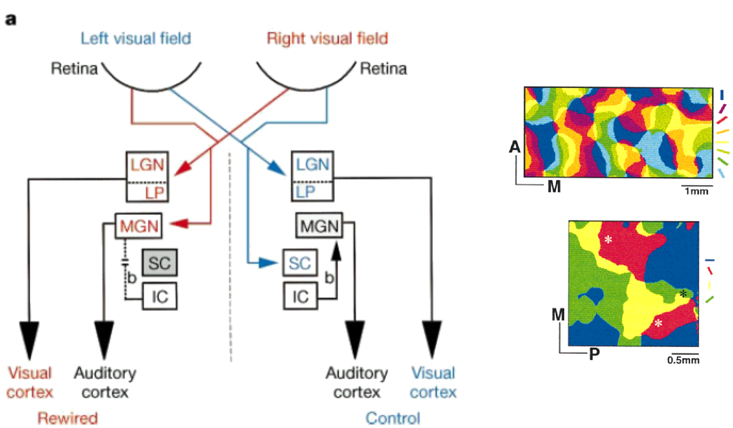
- It could have been caused by a **visual input**.

But, Equally Likely Is ...



- It could have been caused by an **auditory input**.
- Sur et al., Rewiring cortex, *Journal of Physiology*, 41:33–43, 1999

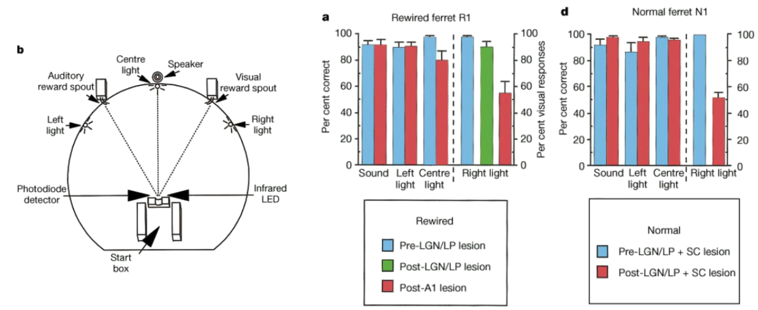
Rewiring Vision to Auditory Area



von Melchner et al. (2000); Sharma et al. (2000); Sur et al. (1999)

- Rewired auditory cortex develops **visual cortex-like** organization.
- Question: does it **see** light or **hear** light?

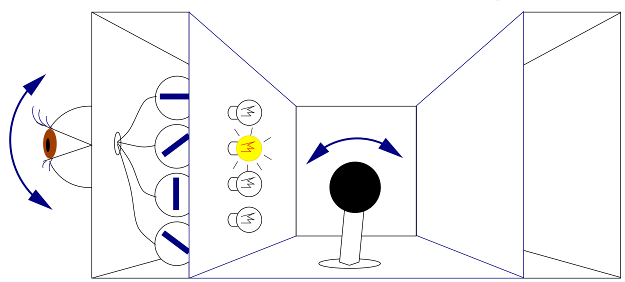
Rewiring: Behavioral Results



von Melchner et al. (2000); Sharma et al. (2000)

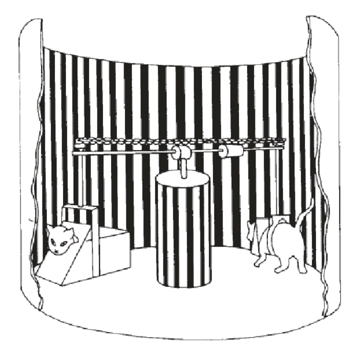
- Ferret **trained to behave differently** for visual vs. auditory stimuli: Behavior *suggests* that the ferret is actually **seeing** light **with** its **auditory** cortex!

Possible Solution: Through Action



- A major problem in the metaphor is the **passiveness** of the whole situation.
- Adding action **can help solve** the problem.
- But **why** and **how**?

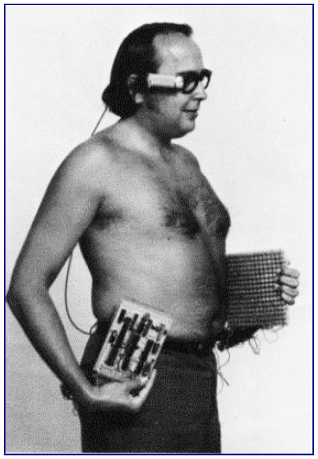
Experimental Evidence



Held and Hein (1963)

- Active animal developed normal vision.
- Passive animal did not.
- Suggests the importance of action in vision.

Experimental Evidence



Bach y Rita (1972; 1983)

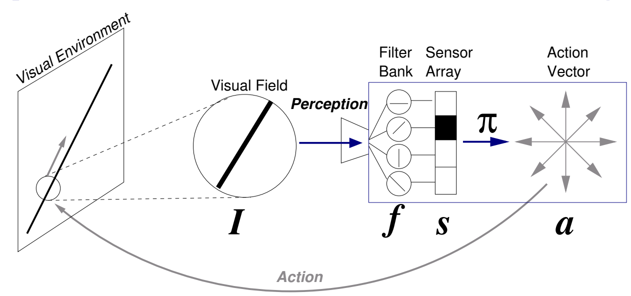
- Vibrotactile array linked to a video camera.
- Passive viewing results in **tactile** sensation.
- Moving the camera results in a **vision-like** sensation.
- Sensation as related to **voluntary/intentional action** may be the key!

Theoretical Insights

- Philipona et al. (2003) showed that properties of ambient space (such as the dimensionality) can be inferred based on internal sensory input alone.
- The key concept is about the compensability between ego-motion and the change in the environmental input conveyed to exteroceptors.

Part I: Learning the Meaning of Internal State

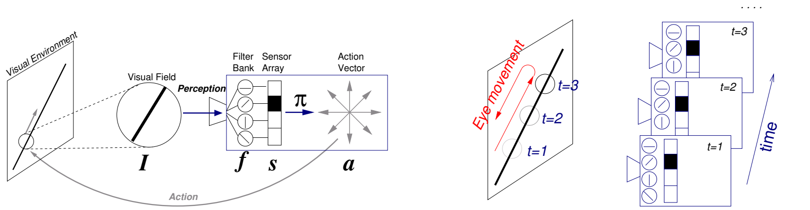
Approach: A Sensorimotor Agent



Choe and Bhamidipati (2003)

- A simple **visuomotor** agent.
- **How** can it **learn** about the visual world?
- What should be the **objective (or goal)** of learning?

Action for Internal Invariance



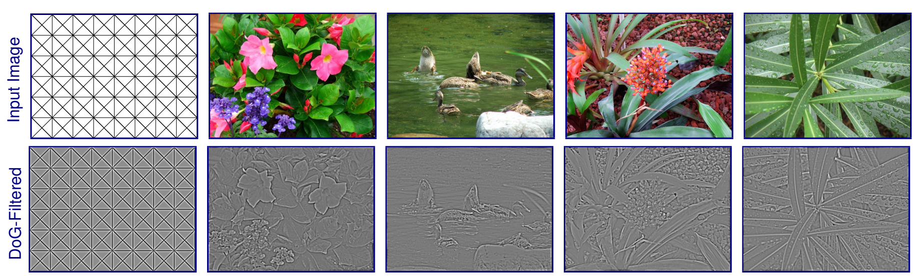
(a) Sensorimotor Agent (b) Sensory Invariance during Motion

- Agent can **move** its visual field.
- Movement in a certain direction (diagonal) causes the *sensory array* to **stay invariant** over time.
- Property of such a movement **exactly reflects** the property of the input I.

Outline of Experimental Methods

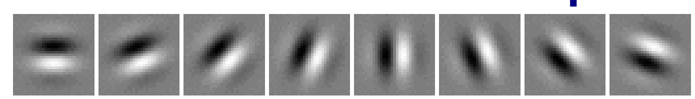
- Input preparation.
- Orientation response calculation.
- Learning algorithm and policy generation.

Methods: Input Preparation



- Convolve with Difference-of-Gaussian (DoG) filter (15×15).
- Then, sample a 31×31 region.

Methods: Orientation Response



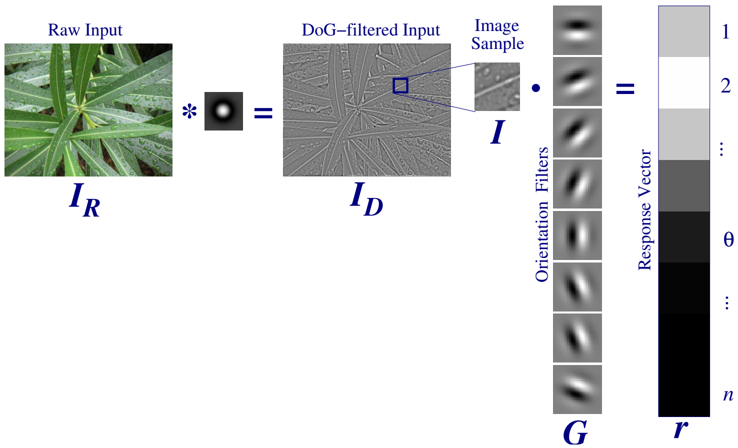
- Find the vectorized dot product of the 31×31 input I and the n Gabor filters G_i ($i = 1..n$, $\theta = \lfloor (i-1)\pi/n \rfloor$):

$$r_i = \sum_{x,y} G_i(x,y) I(x,y).$$

- The above results in a response vector $\mathbf{r}$, and the orientation response s:

$$s = \arg\max_{i=1..n} r_i$$

Orientation Response



Sensory state:

$$s = \arg\max_{1 \le \theta \le n} r_\theta.$$

Methods: Reinforcement Learning (Reward)

- Immediate reward is measured as the dot product of current and previous response vectors:

$$\rho_{t+1} = \mathbf{r}_t \cdot \mathbf{r}_{t+1}$$

- The task the agent is to learn a state-to-action mapping so that it maximizes the reward ρ.

Methods: Policy π

Suppose we know the probability $P(a|s)$ (let us call this $R(s,a)$), where stochastically generating action given the state s with this probability maximizes the reward.

1. Given the current state $s_t \in S$, randomly pick action $a_t \in A$.
2. If a_t equals $\arg\max_{a \in A} R(s_t, a)$,
   (a) then perform action a_t,
   (b) else perform action a_t with probability $R(s_t, a_t)$.
3. Repeat steps 1 to 3 until exactly one action is performed.

In practice, momentum was added so that $a_{t+1} = a_t$ with a 30% chance, and in step 2, if a random draw from $[0..1]$ was less than $cR(s_t, a_t)$, then the action was accepted.

Methods: Learning $R(s,a)$

- A simple update rule was used:

$$R_{t+1}(s_t, a_t) = R_t(s_t, a_t) + \alpha\, \rho_{t+1},$$

  where $\alpha = 0.002$ is the learning rate, and ρ_{t+1} the immediate reward.

- $R_{t+1}(s_t, a)$ was then normalized by:

$$R_{t+1}(s_t, a) := \frac{R_{t+1}(s_t, a)}{\sum_{a' \in A} R_{t+1}(s_t, a')}, \text{ for all } a.$$

Reward Probability Table

A: direction of motion

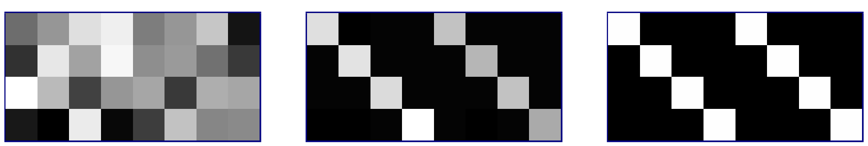
S: sensory state (orientation)	→	↗	↑	↖	←	↙	↓	↘
—	0.5	0	0	0	0.5	0	0	0
/	0	0.5	0	0	0	0.5	0	0
|	0	0	R(s ,a)	0	0	0	0.5	0
\	0	0	0	0.5	0	0	0	0.5

- Reward probability $R(s,a)$ can be tabulated.
- In an ideal case (world consists of straight lines only), we expect to see two diagonal matrices (shaded gray, above).

Results: Overview

1. Synthetic input and natural image input.
2. Learned $R(s,a)$.
3. Error in $R(s,a)$ and average reward ρ over time.
4. Distribution of reward ρ.
5. Gaze trajectory.

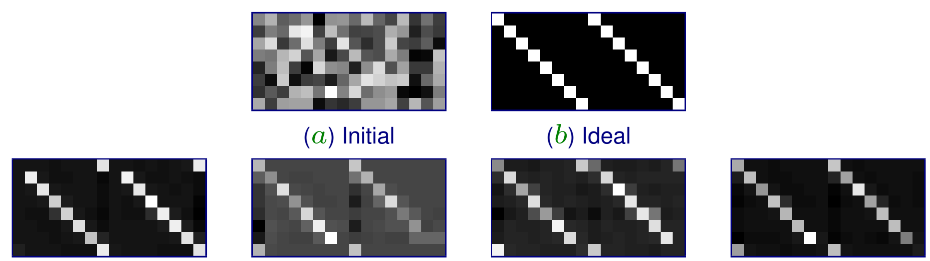
Results: Learned $R(s,a)$ for Synthetic Input



(a) Initial (b) Final (c) Ideal

- Learned $R(s,a)$ close to ideal.

Results: Learned $R(s,a)$ for Natural Images

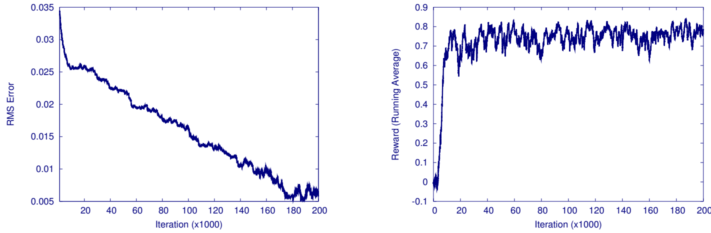


(a) Initial (b) Ideal

(c) Flowers (d) Ducks (e) Plant (f) Oleander

- Learned $R(s,a)$ close to ideal even for natural image inputs.

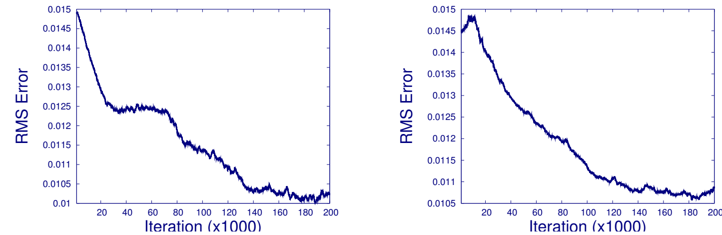
Results: Error in R and Average ρ



Synthetic Input

- Left: Root-mean-squared error in $R(s,a)$ compared to the ideal case.
- Right: running average of immediate reward ρ: $\mu_t = (1-\alpha)r_t + \alpha\,\mu_{t-1}$, ($\mu_1 = \rho_1$, $\alpha = 0.999$).

Results: Error in $R(s,a)$

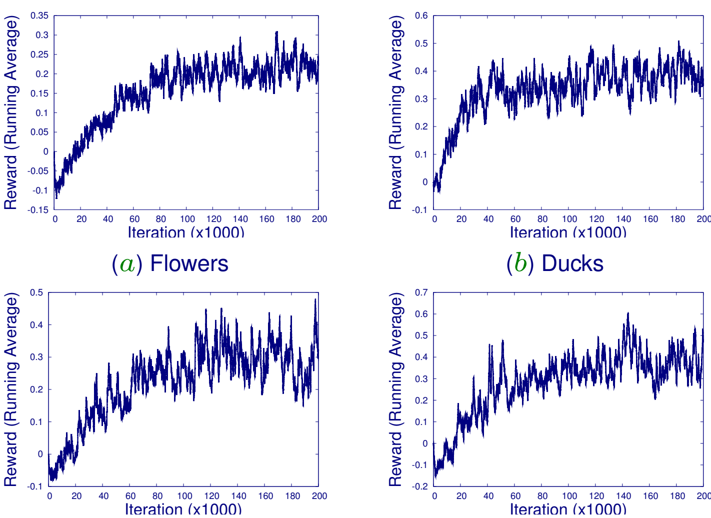


(a) Flowers (b) Ducks

(c) Plant (d) Oleander

Natural Input

Results: Average ρ

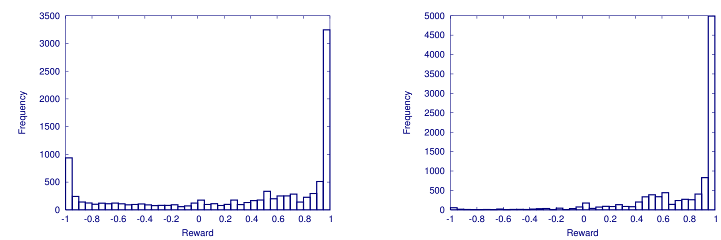


(a) Flowers (b) Ducks

(c) Plant (d) Oleander

Natural Input

Results: Distribution of ρ

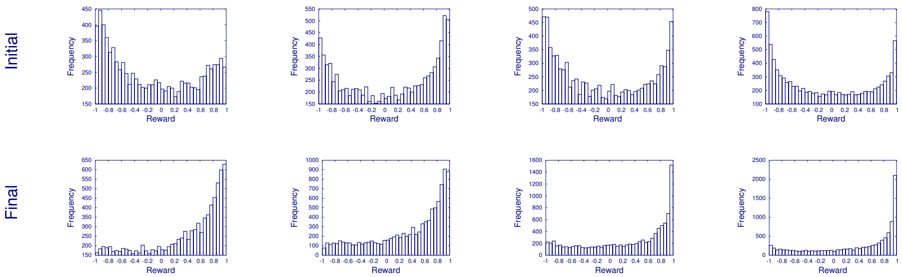


(a) Initial (b) Final

Synthetic Input

- Initially, two peaks: near negative min and positive max ρ.
- Near the end, only one peak: near positive max ρ.

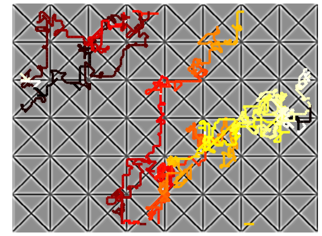
Results: Distribution of ρ



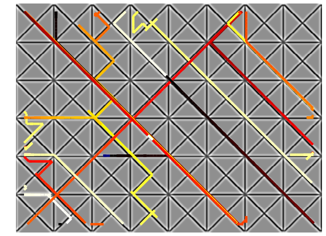
(a) Flowers (b) Ducks (c) Plant (d) Oleander

Natural Input

Results: Gaze Traj. for Synth. Input



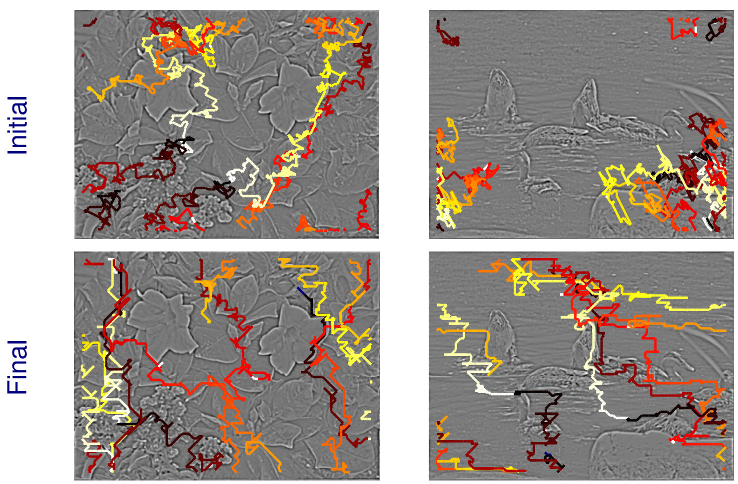
(a) Initial



(b) Final

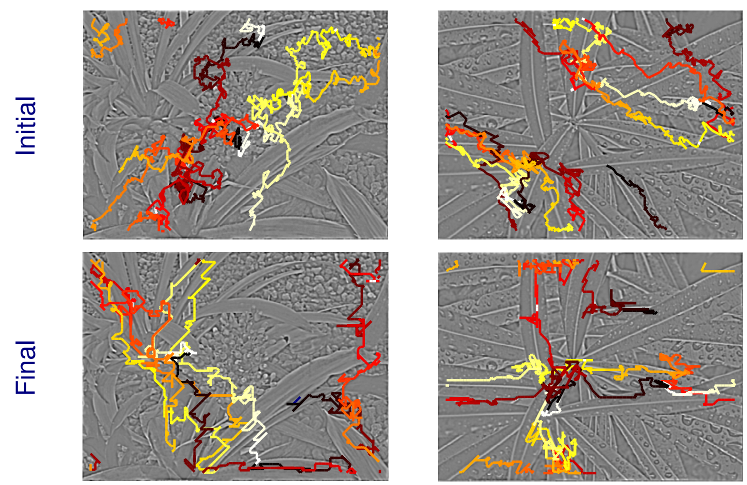
- Gaze trajectory reflects orientation represented by internal state.

Results: Gaze Traj. for Nat. Input



(a) Flowers (b) Ducks

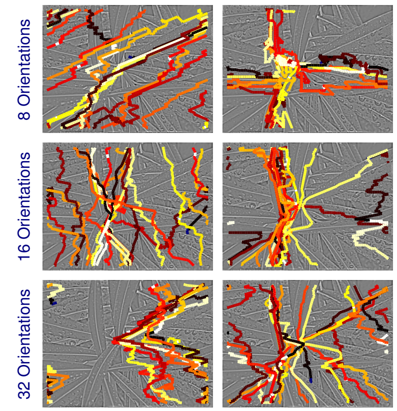
Results: Gaze Traj. for Nat. Input



(c) Plant (d) Oleander

Results: Demo

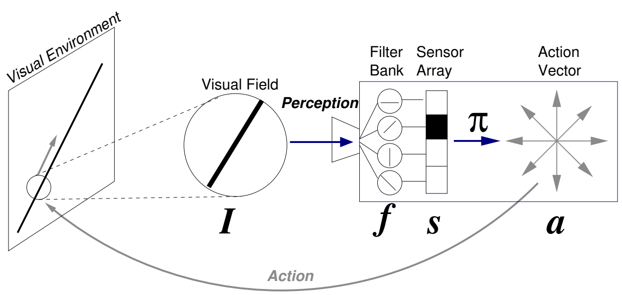
Work in Progress: Q-Learning



(a) Initial (b) Learned

Trajectories from Q-Learning sessions (Choe and Smith 2006).

Interpretation of the Results



- Using **invariance** as the only criterion, particular **action pattern** that has the **same property** as the input that triggered the sensors was learned.
- Question: Can this approach be extended to learning complex stimulus concepts?

Part II: RF Learning

Learning RFs along with Their Grounding (Decoding)

- Grounding (decoding): Same as Part I.
- RFs develops through normalized Hebbian learning:

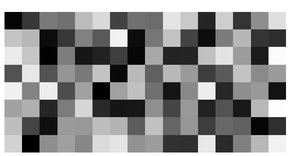
$$g_{ij} = \frac{g_{ij} + \alpha(I_{ij} - g_{ij})}{\sum_{mn} g_{mn} + \alpha(I_{mn} - g_{mn})},$$

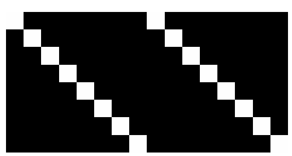
where g_{ij} is the afferent connection weight and I_{ij} the input pixel value.

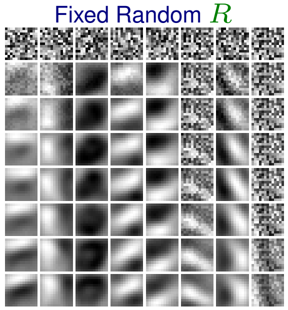
Experiments

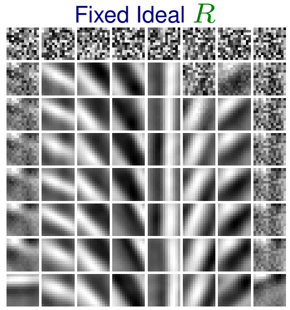
- Effects of different action policy on RF learning.
  - Random $R(s, a)$
  - Ideal $R(s, a)$
- Simultaneous learning of RF and action policy.
  - RF learning through normalized Hebbian learning
  - Reinforcement learning of $R(s, a)$ based on internal-state invariance

Effects of $R(s, a)$ on RF Learning


Fixed Random R


Fixed Ideal R

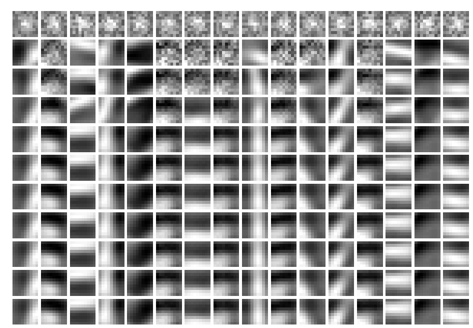

RF w/ Random Policy

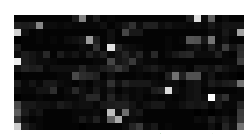

RF w/ Ideal Policy


Reference RFs


Reference RFs

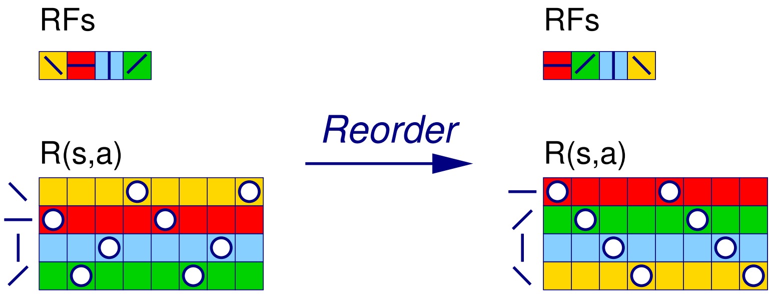
Simul. Learning of RFs & $R(s, a)$


Learned RFs


Learned $R(s, a)$

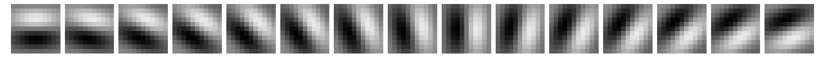
- Seemingly unordered RFs and $R(s, a)$ results.

Reordering RFs

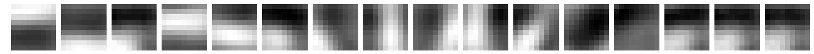


- The $R(s,a)$ result looks bad because each row's corresponding RF orientation is not ordered.
- Reordering RF orientation reorders the rows in $R(s,a)$.

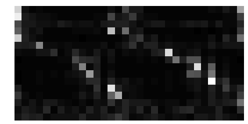
Reordered RFs and $R(s,a)$



Reference RFs



Reordered final RFs



Reordered final $R(s,a)$

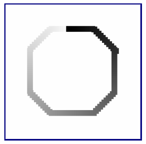
- However, reordering the RFs and their corresponding $R(s,a)$ rows shows the true underlying structure! (Not perfect, but a good start!)

Part II: Summary

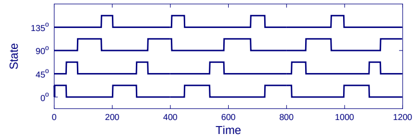
- Action policy strongly influences RF properties, by altering the input statistics.
- Certain action policies may give better RFs, faster.
- Receptive fields and action policy can be learned simultaneously, from scratch, thus allowing encoding/decoding to evolve together.

Part III: Learning Complex Object Concepts

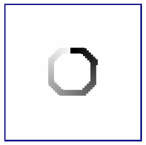
Learning About Complex Objects



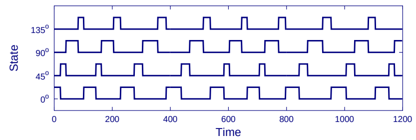
(a) Eye position (large input)



(b) Internal state (large input)



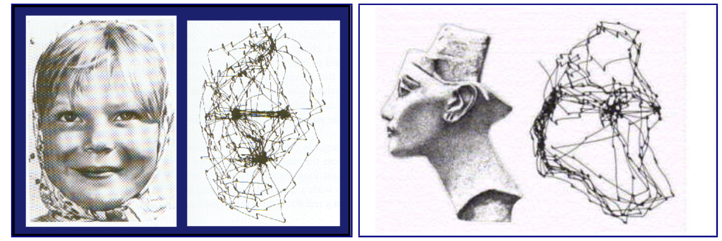
(c) Eye position (small input)



(d) Internal state (small input)

- For complex objects, a history of sensory activity may be needed (i.e., some form of memory).
- Invariance can be detected in the spatiotemporal pattern of sensor activity.

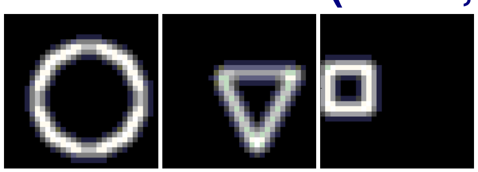
Supporting Evidence?



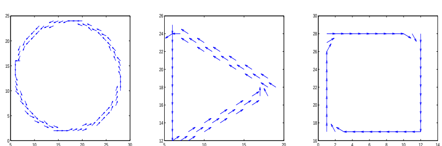
Yarbus (1967)

- When we look at objects, our gaze wanders around.
- Could such an interaction be necessary for object recognition?

Advantage of Motor-Based Memory (Habit, or Skill)



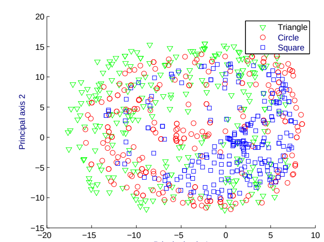
(a) Sensor-based Representation



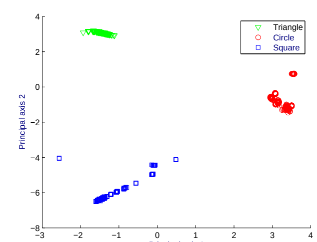
(b) Motor-based Representation

- Sensor-based representations may be hard to learn and inefficient.
- Motor-based approaches may generalize better.
- Comparison: Make both into a 900-D vector and compare backpropagation learning performance.

Class Separability



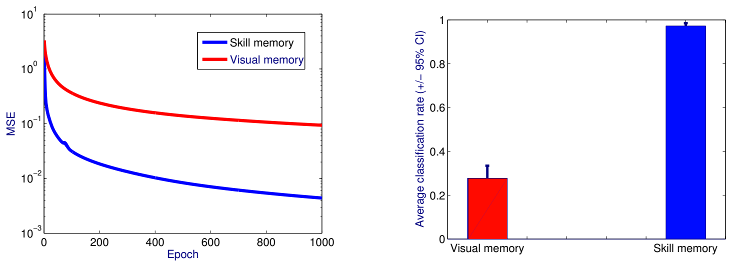
(a) Visual Memory



(b) Motor Memory

- Comparison of PCA projection of 1,000 data points in the visual and motor memory representations.
- Motor memory is clearly separable.

Speed and Accuracy of Learning



(a) Training Speed (b) Generalization Accuracy

- Motor-based memory resulted in faster and more accurate learning (10 trials).

Summary

- Internal observer can learn about the properties of the external environment – through **action** maximizing **invariance** in neural activity.
- Such actions **closely reflect** the property of the stimulus that triggered the sensory neuron to fire: **Meaning** of the spike recovered (through action)!
- **Main contribution**: The **invariance criterion** for autonomously learning the **meaning** of neural states.

Related Work (Selected)

- Piaget (1952): Sensorimotor period in child development
- Freeman (1999): Brain creates meaning through action and choices. Also see Kozma and Freeman (2003) for a KIV model of the emergence of goal-directed, intentional behavior.
- O'Regan and Noë (2001): Sensorimotor contingency theory
- Philipona et al. (2003): Inferring space through sensorimotor interaction
- Rizzolatti et al. (2001): Mirror neurons
- Gibson (1950): Direct perception of invariance and affordance
- Harnad (1990): Symbol grounding on robotic capabilities.
- Taylor (1999): Corollary discharge and awareness of attention movement prior to sensory awareness.

Discussion

- Why is **knowing** ones own **action** any easier than **perceptual interpretation**?: Knowledge of own action may be more **immediate** than perception (cf. Moore 1996, citing Bergson).
- What gives rise to **voluntary, intentional action** and why is it special? (Freeman 1999; Kozma and Freeman 2003; Taylor 1999).
- A different view of **invariance**: Not (only) something to be detected in the environment (cf. Gibson 1950), but something that we **actively seek within**.

Discussion (Cont'd)

- Why not just **analyze the input directly?**: The raw input is only available at the immediate sensory surface.
- What about **other sensory modalities** (such as touch, olfaction, or audition)?
- The learning scheme **depends** on **structure** in the environment: If the environment didn't have structure, the agent can never learn.

Discussion (Cont'd)

- Relation to **mirror neurons** (Rizzolatti et al. 2001)?
- Role of **attention** (e.g. Rensink et al. 1997; Taylor 1999)?: Attention may be needed when ambiguities are present.
- Do **motor primitives** restrict the kind of sensory property that can be learned? What kinds of motor primitive do we have?

Discussion (Cont'd)

- What about meaning **other than sensorimotor-like**, such as reinforcement signals (Rolls 2001) or "feeling" (Harnad 2001)?
- **Grounding on perception alone** may not be sufficient: cf. Perceptual symbol system (Barsalou et al. 2003).
- What to make of the segregation in the **dorsal–ventral pathway?** (Goodale and Milner 1992).

Predictions

- Perceived orientation of a line can be altered by eye movement in the direction of incompatible orientation.
- Motor structures (cerebellum, basal ganglia) may be intimately involved in semantics.
- Geometrical understanding may be limited by the motor primitive repertoire.

Future Work (and Work in Progress)

- Lateral inhibition in sensory array.
- Crossmodal association through sensory invariance.
- Extending to more complex concepts.

Conclusions

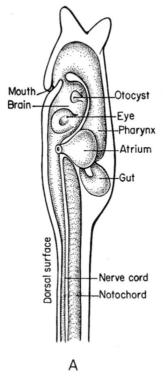
- We must ask how the brain understands itself.
- Autonomous understanding of own internal state is non-trivial without direct access to the stimulus.
- Action can help solve the conundrum.
- Action that maintains invariance in internal state can recover meaning (the property of the stimulus).

Credits

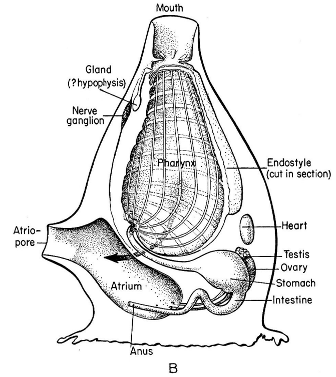
- Kuncara A. Suksadadi helped in the early stages of the idea's development.
- Thanks to Ricardo Gutierrez-Osuna, Ronnie Ward, Stevan Harnad, James Clark, and Ben Kuipers for helpful discussions.
- Thanks to Texas A&M Cognoscenti and NIL members for insightful comments.
- Partially supported by Texas Higher Education Coordinating Board (ATP 000512-0217-2001).

Why Do We Have a Brain?

Tree
(no Brain)



Tunicate
Free-floating
(w/ Brain)



Tunicate
Settled
(w/o Brain)

- Brain vs. no brain

Sources: http://homepages.inf.ed.ac.uk/jbednar/ and http://bill.srnr.arizona.edu/classes/182/Lecture-9.htm

References

Bach y Rita, P. (1972). *Brain Mechanisms in Sensory Substitution*. New York: Academic Press.

Bach y Rita, P. (1983). Tactile vision substitution: Past and future. *International Journal of Neuroscience*, 19:29–36.

Barlow, H. (1994). What is the computational goal of the neocortex? In Koch, C., and Davis, J. L., editors, *Large Scale Neuronal Theories of the Brain*, 1–22. Cambridge, MA: MIT Press.

Barsalou, L. W., Simmons, W. K., Barbey, A. K., and Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, 7:84–91.

Bell, A. J., and Sejnowski, T. J. (1997). The “independent components” of natural scenes are edge filters. *Vision Research*, 37:3327.

Choe, Y., and Bhamidipati, S. K. (2003). Learning the meaning of neural spikes through sensory-invariance driven action. Technical Report 2003-8-3, Department of Computer Science, Texas A&M University.

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

Choe, Y., Yang, H.-F., and Eng, D. C.-Y. (2007). Autonomous learning of the semantics of internal sensory states based on motor exploration. *International Journal of Humanoid Robotics*, 4:211–243. http://dx.doi.org/10.1142/S0219843607001102.

Freeman, W. J. (1999). *How Brains Make Up Their Minds*. London, UK: Wiedenfeld and Nicolson Ltd. Reprinted by Columbia University Press (2001).

Freeman, W. J. (2003). A neurobiological theory of meaning in perception. In *Proceedings of the International Joint Conference on Neural N etworks*, 1373–1378. Piscataway, NJ: IEEE Press.

Gibson, J. J. (1950). *The Perception of the Visual World*. Boston: Houghton Mifflin.

Goodale, M. A., and Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15:20–25.

Harnad, S. (1990). The symbol grounding problem. *Physica D*, 42:335–346.

Harnad, S. (2001). TTT guarantees only grounding: But Meaning = Grounding + Feeling. *Think*, 12(045).

Held, R., and Hein, A. (1963). Movement-produced stimulation in the development of visually guided behavior. *Journal of Comparative and Physiological Psychology*, 56:872–876.

Hoyer, P. O., and Hyvärinen, A. (2000). Independent component analysis applied to feature extraction from colour and stereo images. *Network: Computation in Neural Systems*, 11:191–210.

Hubel, D. H., and Wiesel, T. N. (1959). Receptive fields of single neurons in the cat’s striate cortex. *Journal of Physiology*, 148:574–591.

Jones, J. P., and Palmer, L. A. (1987). An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58(6):1233–1258.

Kozma, R., and Freeman, W. J. (2003). Basic principles of the KIV model and its application to the navigation problem. *Journal of Integrative Neuroscience*, 2:125–145.

Miikkulainen, R., Bednar, J. A., Choe, Y., and Sirosh, J. (2005). *Computational Maps in the Visual Cortex*. Berlin: Springer. URL: http://www.computationalmaps.org.

Moore, F. C. T. (1996). *Bergson: Thinking Backwards*. Cambridge, UK: Cambridge University Press.

Olshausen, B. A., and Field, D. J. (1997). Sparse coding with an overcomplete basis set: A strategy employed by v1? *Vision Research*, 37:3311–3325.

O’Regan, J. K., and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5):883–917.

Philipona, D., O’Regan, J. K., and Nadal, J.-P. (2003). Is there something out there? Inferring space from sensorimotor dependencies. *Neural Computation*, 15:2029–2050.

Piaget, J. (1952). *The Origins of Intelligence in Children*. New York: Norton.

Rensink, R. A., O’Regan, J. K., and Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8:368–373.

Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, 2:661–670.

Rolls, E. T. (2001). Representation in the brain. *Synthese*, 129:153–171.

Searle, J. R. (1980). Minds, brains and programs. *Behavioral and Brain Sciences*, 3.

Sharma, J., Angelucci, A., and Sur, M. (2000). Induction of visual orientation modules in auditory cortex. *Nature*, 404:841–847.

Sur, M., Angelucci, A., and Sharma, J. (1999). Rewiring cortex: The role of patterned activity in development and plasticity of neocortical circuits. *Journal of Neurobiology*, 41:33–43.

Taylor, J. G. (1999). *The Race for Consciousness*. Cambridge, MA: MIT Press.

von Melchner, L., Pallas, S. L., and Sur, M. (2000). Visual behaviour mediated by retinal projections directed to the auditory pathway. *Nature*, 404(6780):871–876.