Motor Exploration Is Key to Decoding Perceptual Primitives

ARMADILLO 2010

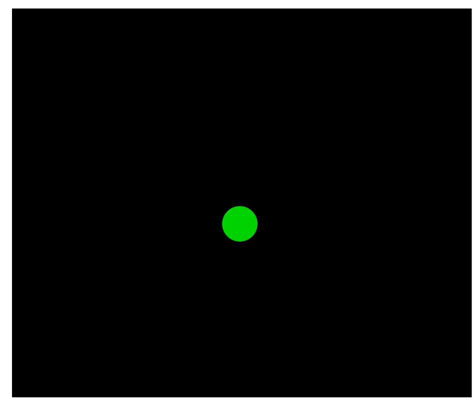
October 22, 2010

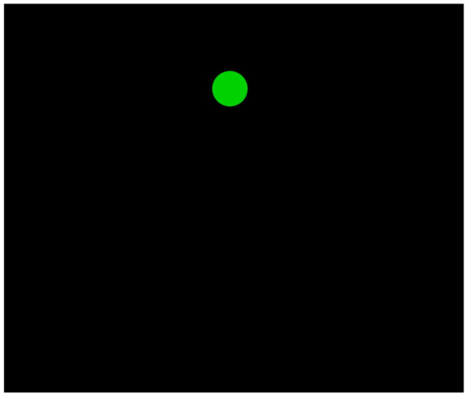
Yoonsuck Choe, Ph.D.
Department of Computer Science & Engineering
Texas A&M University

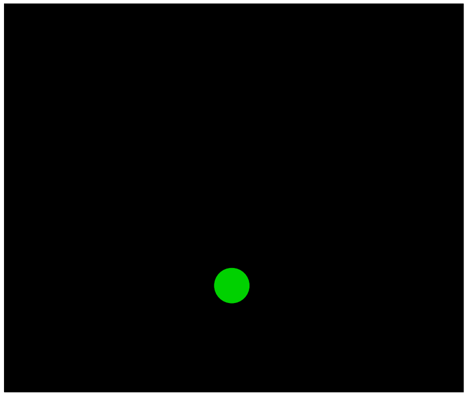
With Noah Smith, Huei-Fang Yang, and Navendu Misra.

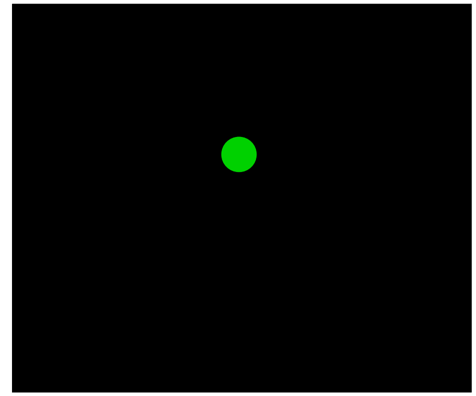
What Is the Meaning of ...

Words
↓
Representations
↓
Brain Activation Patterns
↓
Single Neuronal Spikes ⟶ Let's focus on **perceptual primitives**, like the visual cortical edge detectors.





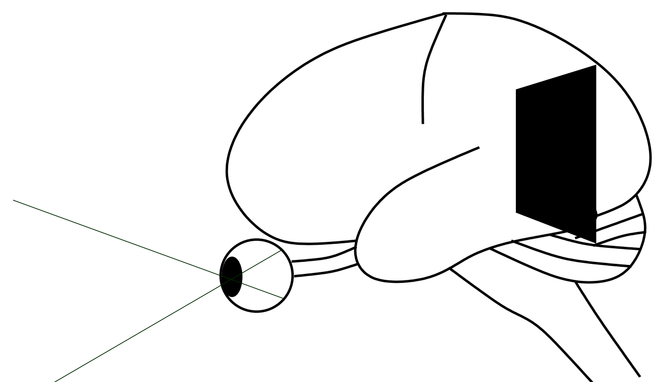




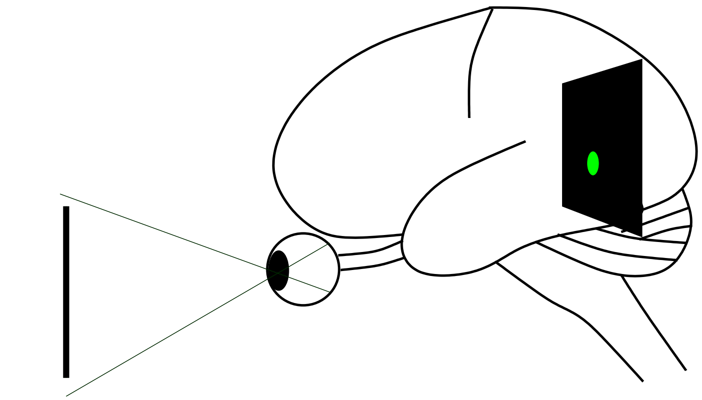
What Is the Meaning of the Green Lights?

- It is hard to get any idea at all.
- If these are neuronal spikes, there's no hope in understanding the meaning of these!

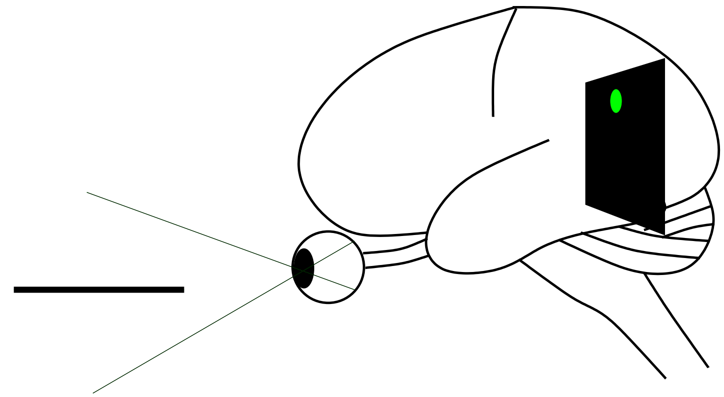
They Are Visual Cortical Responses to Oriented Lines



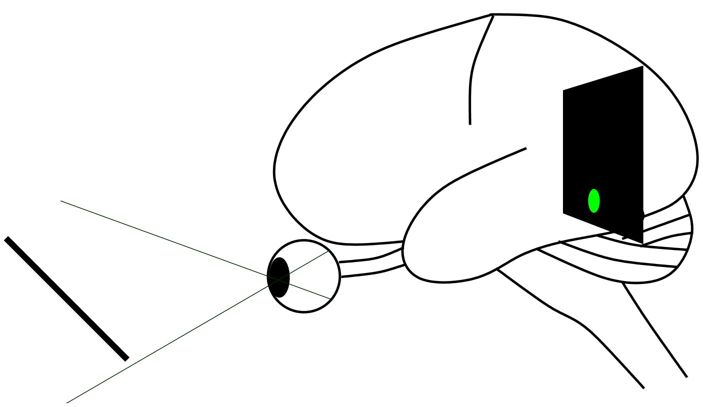
They Are Visual Cortical Responses to Oriented Lines



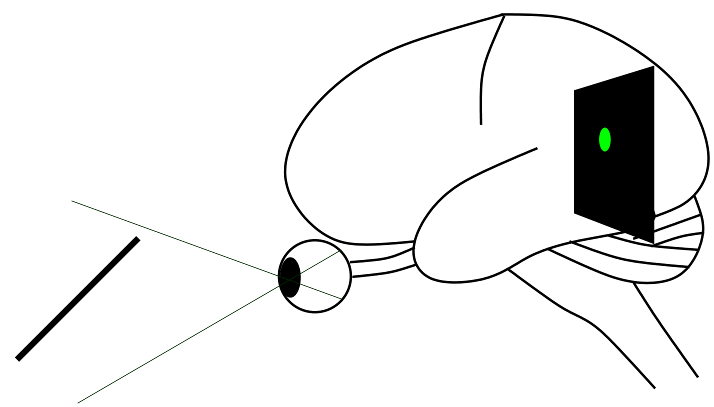
They Are Visual Cortical Responses to Oriented Lines



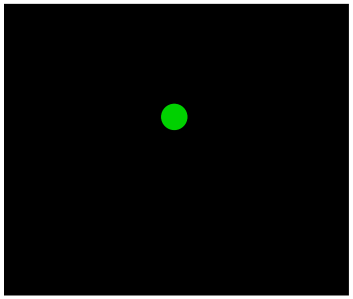
They Are Visual Cortical Responses to Oriented Lines

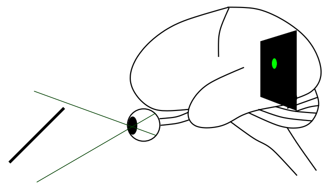


They Are Visual Cortical Responses to Oriented Lines



What Happened Here?





Input not known

↓

Clueless

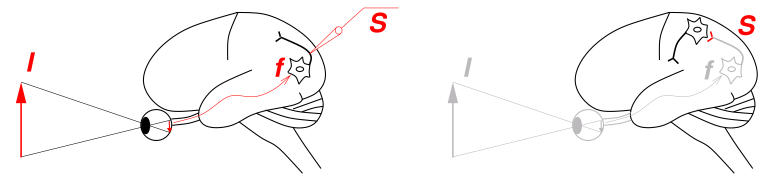
Input known

↓

Meaning clear

Do we need the input then, to understand the meaning?

Two Approaches to Meaning



(a) External observer   (b) Internal observer

- Fom the outside – seems straightforward but artificial.

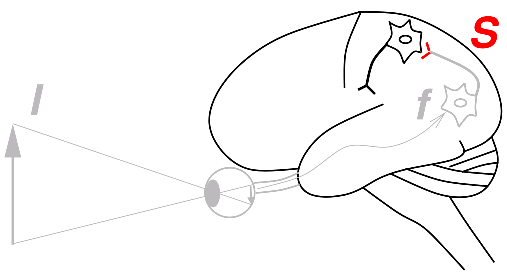
  $\longrightarrow$ Neuroscientist's approach. 3rd person.

- From the inside – seems impossible but natural.

  $\longrightarrow$ The brain's approach. 1st person.

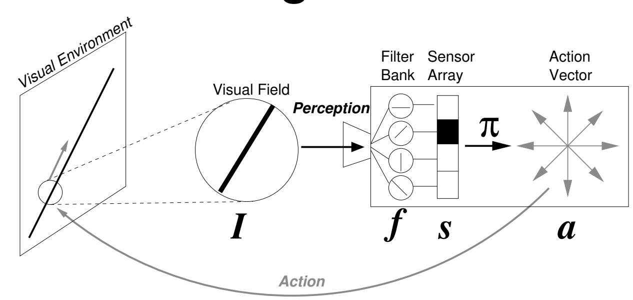
Why does the natural seem more impossible?

Possible Solution: Allow Action



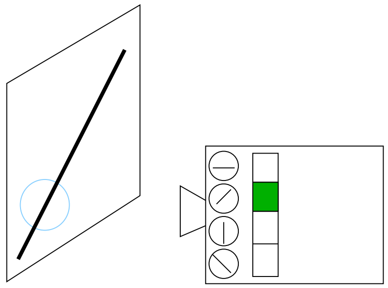
- A major problem in the picture is the **passiveness** of the whole situation.
- Adding action **can help solve** the problem.
- But **why** and **how**?

Approach: Semantic Grounding Through Action



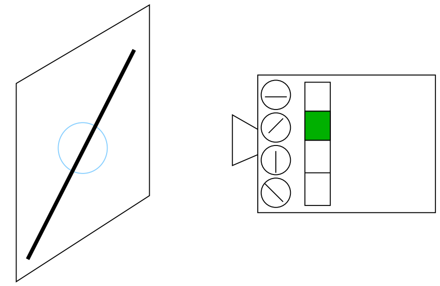
- Direct access only to **encoded internal state**.
- Action: **can move the gaze**.
- How does this solve the grounding problem?

Action for Unchanging Internal State



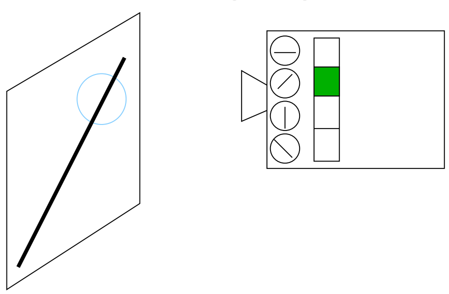
- Diagonal motion causes the *internal state* to **remain unchanging** over time.
- Property of such a movement **exactly reflects** the property of the input I: Semantics figured out through action.

Action for Unchanging Internal State



- Diagonal motion causes the *internal state* to **remain unchanging** over time.
- Property of such a movement **exactly reflects** the property of the input I: Semantics figured out through action.

Action for Unchanging Internal State

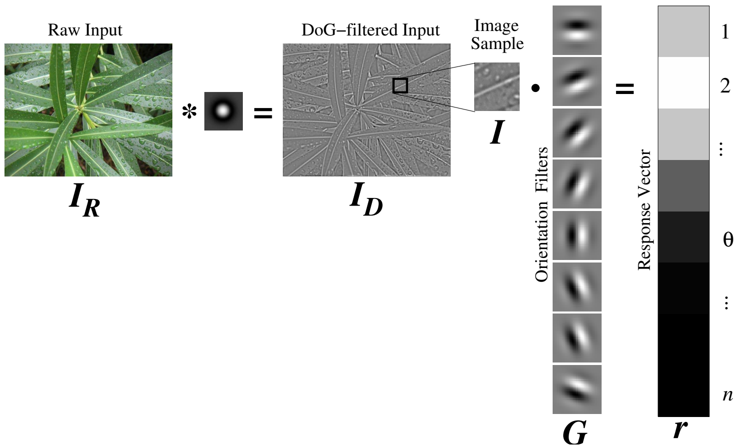


- Diagonal motion causes the *internal state* to **remain unchanging** over time.
- Property of such a movement **exactly reflects** the property of the input I: Semantics figured out through action.

Task

- Given an encoded perceptual signal s, we want to learn action a that **maximizes the invariance** in the internal state over time.
- The learned action a will give **meaning** to s.
- This is basically a **reinforcement learning** task.

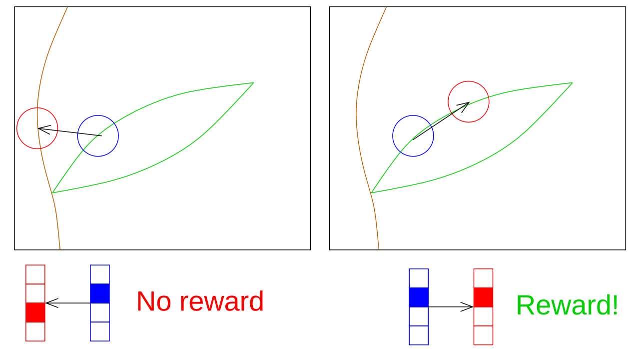
Methods: Orientation Response



Perceptual state:

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

Methods: Reinforcement Learning



$R(s, a)$: How desirable is action a in state s?

- $R(s, a)$ increased if action a in state s leads to unchanged internal state.
- $R(s, a)$ decreased otherwise.

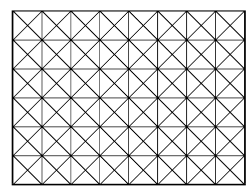
Reward Probability Table

A: direction of motion

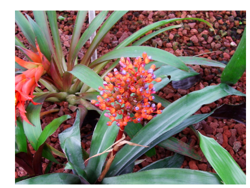
S: sensory state (orientation)	→	↗	↑	↖	←	↙	↓	↘
—	0.5	0	0	0	0.5	0	0	0
/	0	0.5	0	0	0	0.5	0	0
|	0	0	R(s ,a)	0	0	0	0.5	0
\	0	0	0	0.5	0	0	0	0.5

- 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).

Input Images



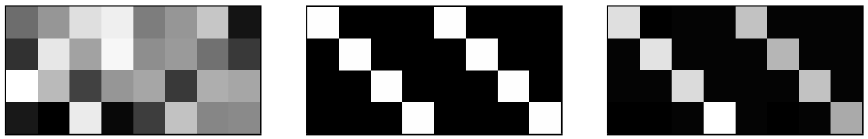
Synthetic



Natural (plant)

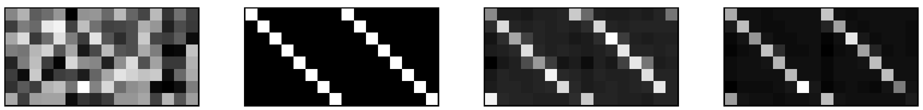
Natural (oleander)

Results: Learned $R(s, a)$



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

Synthetic image

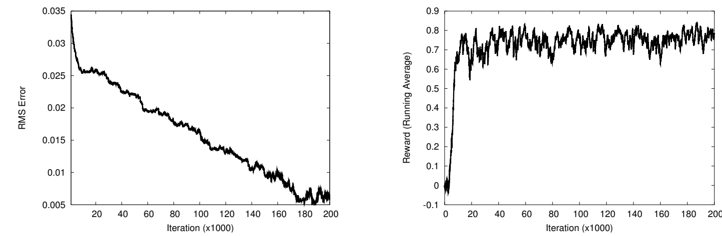


(a) Initial (b) Ideal (c) Plant (d) Oleander

Natural images

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

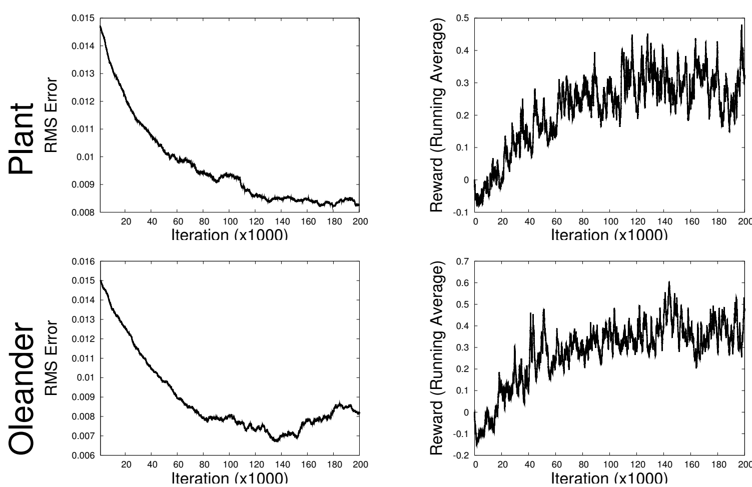
Results: Error in R and Average ρ



Synthetic Input

- Left: Root-mean-squared error in $R(s, a)$ compared to the ideal case.
- Right: running average of immediate reward ρ.

Results: Error in R and Average ρ

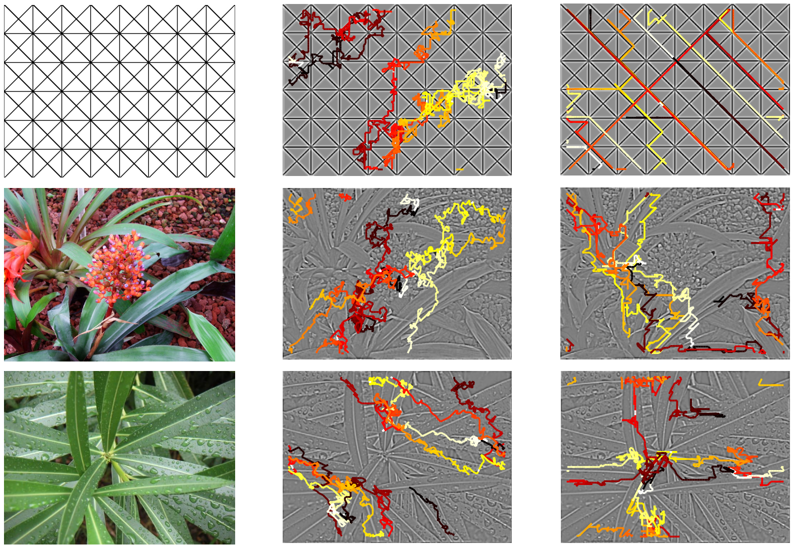


(a) Error (b) Reward

Natural Input

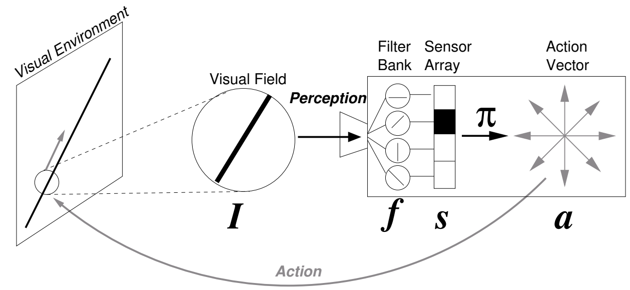
Results: Demo

Results: Gaze Trajectory



(a) Input (b) Initial (c) Final

Summary

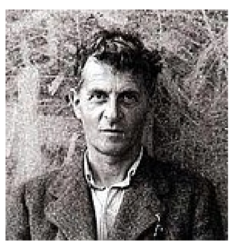


- (1) Using **invariance** as the only criterion, (2) particular **action pattern** was learned, (3) that has the **same property** as the input that triggered the sensors.

Related Work

images from wikipedia

Kant



Wittgenstein

- Kant: bottom-up + prior knowledge: 3rd person
- Wittgenstein: meaning in language use: 1st person
- Perceptual grounding (Barsalou et al. 2003)
- Sensorimotor grounding (**?**)
- Sensormotor account of vision and consciousness (O'Regan and Noë 2001).

Discussion

- Main contribution: Discovery of the invariance criterion for sensorimotor-based semantic grounding.
- Importance of self-generated action in autonomous understanding.
- Richer motor primitive repertoire can lead to richer understanding.
- Tool use can dramatically augment motor primitive repertoire, and thus understanding.

Conclusions

We must ask how the brain understands itself.

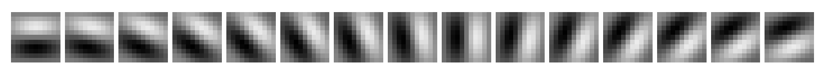
- Action is important for understanding/grounding.
- Simple criterion (state invariance) can help link perceptual coding with meaningful action.

Credits

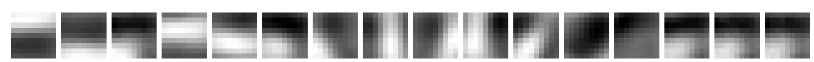
- Contributors: Kuncara A. Suksadadi, S. Kumar Bhamidipati, Noah Smith, Stu Heinrich, Navendu Misra, Huei-Fang Yang, Daniel C.-Y. Eng
- Choe et al. (2008, 2007); Choe and Smith (2006); Choe and Bhamidipati (2004)

More Advanced Results

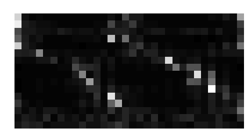
Simultaneous Learning of Perceptual Primitives and $R(s, a)$



Reference RFs



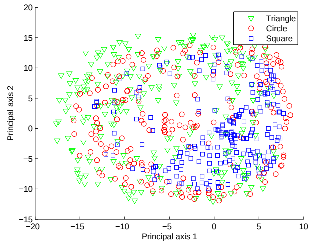
Reordered final RFs



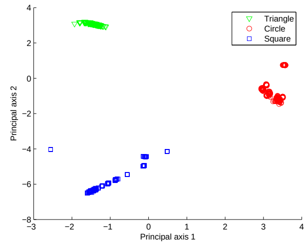
Reordered final $R(s, a)$

- Both can be learned at the same time!

Motor vs. Sensory Representation Compared



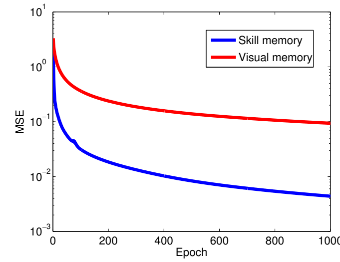
(a) Visual Memory



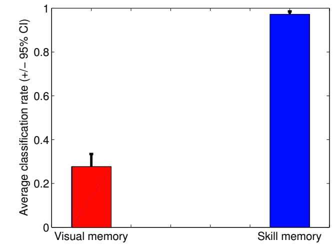
(b) Motor Memory

- 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 in Motor vs. Sensory Representation



(a) Training Speed



(b) Generalization Accuracy

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

References

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.

Choe, Y., and Bhamidipati, S. K. (2004). Autonomous acquisition of the meaning of sensory states through sensory-invariance driven action. In Ijspeert, A. J., Murata, M., and Wakamiya, N., editors, *Biologically Inspired Approaches to Advanced Information Technology*, Lecture Notes in Computer Science 3141, 176–188. Berlin: Springer.

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(AAAI 2006)*. 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.

Choe, Y., Yang, H.-F., and Misra, N. (2008). Motor system's role in grounding, receptive field development, and shape recognition. In *Proceedings of the Seventh International Conference on Development and Learning*, 67–72. IEEE.

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