

# Motor System's Role in Grounding, Receptive Field Development, and Shape Recognition

Yoonsuck Choe, Huei-Fang Yang, and Navendu Misra

Department of Computer Science

Texas A&M University

College Station, TX 77843-3112

Email: choe@tamu.edu, hfyang@cs.tamu.edu, navendu.misra@gmail.com

**Abstract**—Vision is basically a sensory modality, so it is no surprise that the investigation into the brain's visual functions has been focused on its sensory aspect. Thus, questions like (1) how can external geometric properties represented in internal states of the visual system be grounded, (2) how do the visual cortical receptive fields (RFs) form, and (3) how can visual shapes be recognized have all been addressed within the framework of sensory information processing. However, this view is being challenged on multiple fronts, with an increasing emphasis on the motor aspect of visual function. In this paper, we will review works that implicate the important role of motor function in vision, and discuss our latest results touching upon the issues of grounding, RF development, and shape recognition. Our main findings are that (1) motor primitives play a fundamental role in grounding, (2) RF learning can be biased and enhanced by the motor system, and (3) shape recognition is easier with motor-based representations than with sensor-based representations. The insights we gained here will help us better understand visual cortical function. Also, we expect the motor-oriented view of visual cortical function to be generalizable to other sensory cortices such as somatosensory and auditory cortices.

## I. INTRODUCTION

The primary visual cortex (V1), a major visual processing area in the brain, has been extensively studied through neurophysiological and theoretical/computational methods. In the past few decades, researchers investigated various issues including what kind of stimulus properties are encoded in V1 neuron's spikes, how the receptive fields (RFs) of V1 neurons develop over time and how they are influenced by natural stimulus statistics, what kind of topographic organization exists in V1, etc. (see [1] for a review). Since vision is basically a sensory modality, almost all of these investigations have focused on the sensory aspect of visual information processing.

However, there is a strong indication that visual function is inseparable from motor function and only through this perspective we can reveal how the visual cortex really works. Earlier theoretical investigations by Dewey [2], Bergson [3], and Wittgenstein [4], emphasized the importance of action in (visual) perception. More recent works came to similar conclusions, e.g., Varela et al.'s enactive cognition [5], Humphrey's work on evolution of consciousness [6], O'Regan and Noe's sensorimotor theory of perception [7], Llinás's virtual action patterns [8], Schank's conceptual dependence theory based on motor primitives [9], Von Foerster's eigenbehavior [10], [11], and Freeman's pragmatism-view of brain function [12].

Concrete experimental works also point in the same direction, i.e., motor function is inseparable from perceptual function. For example, the following works show the tight integration of motor function with visual perceptual function: Milner and Goodale's work on the ventral and dorsal visual pathways [13], Rizzolatti et al.'s work on the mirror neurons (neurons in the prefrontal cortex responding both to visually perceived gesture and to enacted gesture of the same type) [14], Salinas's work on motoric requirements influencing RF form [15], etc. There is also a large number of engineering approaches on enactive perception/cognition [16]–[18], behavioral robotics [19], [20], and autonomous mental development [21] that emphasize the sensorimotor link.

In this paper, we will review our prior works on how such a motor perspective can be applied to the modeling and understanding of visual cortical development and function. More specifically, we will show how (1) grounding [22], (2) RF development [23], and (3) shape recognition [24] can be influenced and aided by motor function.

The remainder of the paper is organized as follows. We will first discuss the three topics above in individual sections, each section containing background and our experimental results. Finally, we will discuss the implications of these results on understanding visual cortical function, and provide a summary in the conclusion.

## II. GROUNDING OF INTERNAL VISUAL REPRESENTATIONS

How does an agent, which cannot directly access the external world but only its internal state (i.e., brain state) learn the basic properties of the external world [25]? That is, how can the agent decode (or *ground*) its internal state (e.g., the activity patterns in V1), based on the internal state only? For example, consider Fig. 1. External observers (e.g., scientists) have access to both the external input and the internal states of the brain under observation (Fig. 1(a)), so we can infer what properties are encoded by the internal state. However the brain can only access the internal state of itself (Fig. 1(b)), so the approach taken by the external observer is not applicable. We showed that this problem can be solved through the use of action (Fig. 2) that fulfills a simple criterion of internal state invariance (Fig. 3) [22]. Note that by "invariance", we simply mean that the activity pattern does not change over time: We

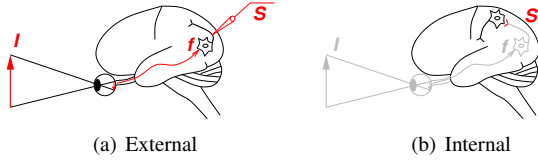


Fig. 1. **External vs. Internal Observer.** (a) External observers (e.g., scientists observing a brain) have full access to both the environmental stimulus and the spikes. By correlating the spikes to the input, the stimulus properties carried by the spikes can be inferred. (b) Internal observers (e.g., downstream neurons) have access to the upstream neuron’s spikes only, making such an inference seem impossible. Adapted from [22], [26].

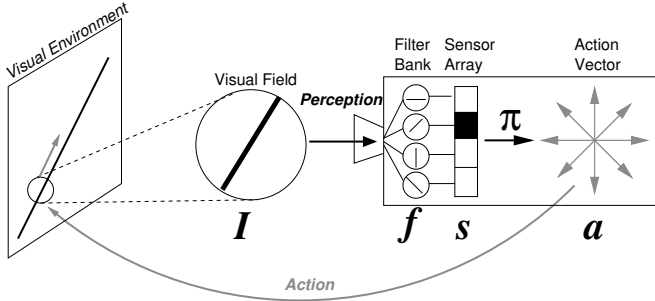


Fig. 2. **A Visuomotor Agent.** The agent has a limited field view and a set of sensory primitives  $f$ . Sensory primitives  $f$  receive input and generate an activity in the sensory array  $s$ . Adapted from [22], [26].

are *not* referring to perceptual invariances such as translation, rotation, and scaling invariance.

Consider a simple agent shown in Fig. 2, with limited visual field view and a set of sensory receptive fields. Neurons with sensory receptive fields (shown as oriented lines) receive input and generate activities in the sensory array  $s$  (modeling the visual cortex). Based on the activity of the sensory array, motor response (eye movement) can be generated. This can be modeled as a mapping  $\pi$ . How can this added capability to move the gaze help in decoding (or grounding, [27]) the internal state (the spikes in the visual cortex)? Our main observation was that the meaning of the spikes can be recovered through a specific kind of eye movement. Take the state of the agent shown in Fig. 3 for example. At time  $t = 1$ , the input activates the neuron tuned to  $45^\circ$  input. (Note that the agent has access to the neuron’s spikes.) Suppose the agent happened to move in the  $45^\circ$  direction and back ( $135^\circ$ ). The result would be that the internal state does not change over time ( $t = 1, t = 2, t = 3$ , on the right) even though the agent is generating motion. The key observation here is that (1) the property of such a motion and (2) the stimulus property represented by the spike are identical. In converse, if the agent had the goal of moving while maintaining its internal state invariant over time, it will end up generating action that reflects the property of the encoded spike. This way, the internal state can be decoded (grounded), all without direct access to the external world. Thus, internal state invariance can serve as a strong criterion for grounding internal representations (which naturally includes visual cortical representations).

In this following, we will describe how to process the input

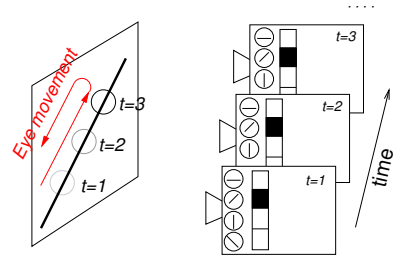


Fig. 3. **Invariance in Internal State during Motion.** The agent moving in diagonal direction ( $45^\circ$  or  $135^\circ$  in this example) will keep the internal state unchanged. By generating a behavior congruent with input stimulus, the agent can infer external world from internal state. Adapted from [22], [26].

image, generate sensory activities, learn the mapping from the sensory states to motor primitives.

#### A. Initial input processing

Below, we follow the steps described in our earlier work [22], so we will only provide a sketch of the overall process. Please refer to [22] for details. The whole process is summarized in Fig. 4, from left to right. First, the raw input image  $I_R$  ( $640 \times 480$ ) is convolved by a Difference of Gaussian (DoG). The DoG filter was  $15 \times 15$  in size, with the width parameter  $\sigma$  set to  $15/4$ . The resulting image was normalized to have zero mean and the max range scaled down between -1.0 and 1.0. This resulted in the convolved image  $I_D$ . From this image a small  $9 \times 9$  area was sampled, resulting in the input  $I$ .

#### B. Response generation

The cortical response generation process is outlined in Fig. 4. Given a set of receptive fields the vectorized dotproduct between the input sample  $I$  and the receptive fields  $G_i$  are calculated, resulting in the response vector  $r$ . (In the figure, eight oriented Gabor filters are shown, but these can be an arbitrary pattern, as it may have to be adapted throughout development.) The vector  $r$  is then normalized by its  $l_2$ -norm  $\|r\|$ . The current state index  $s$  is determined by:

$$s = \arg \max_{\theta=1..n} r_i, \quad (1)$$

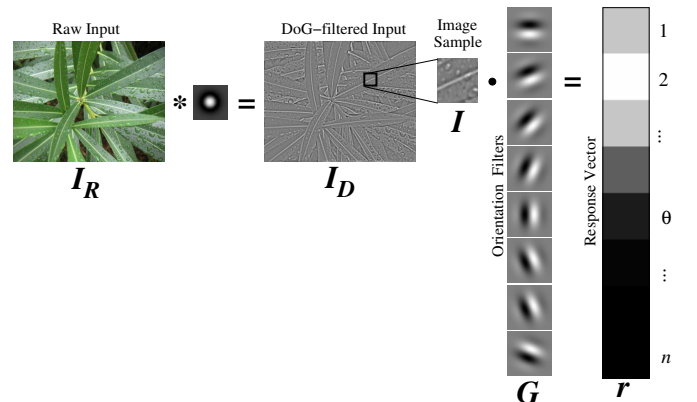


Fig. 4. **Response Generation.** An overview of the response generation process is shown. See text for details.

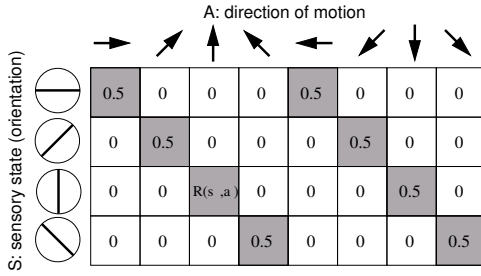


Fig. 5. **Reinforcement Learning of Internal State to Action Mapping.** The rows represent the internal sensory state (orientation in this case), and the columns the direction of motion (gaze). For each orientation, in an ideal case, there are two optimal directions to move to maintain invariance in the internal state. For example, for  $0^\circ$  orientation, one can either move left or right. Thus, the reward table  $R(s, a)$  has a diagonal structure as shown above, in an ideal case. Adapted from [22], [26].

where  $\theta$  is the index of the receptive field, and  $n$  the number of receptive fields.

### C. Learning the Motor Policy

Learning in the agent model occurs in two different parts: (1) internal state to action mapping  $\pi$ , and (2) receptive field structure. Learning of  $\pi$  follows closely our previous work [22], so again, we will just provide a brief overview rather than going into details. The agent’s gaze is controlled by a stochastic process based on the conditional probability  $P(a|s)$  where  $a \in A$  is the gaze direction and  $s \in S$  is the internal state indicating the maximally responding unit in the sensory array. Given a specific current state  $s$ , action  $a$  is chosen with the probability  $P(a|s)$ . The task is to learn this conditional probability so that when the above scheme is employed, maximum state-to-state invariance is achieved over time. For convenience, let us write  $R(s, a)$  (and call it the reward table) instead of  $P(a|s)$ , in the following (see Fig. 5).

$R(s, a)$  is basically trained using a reinforcement learning algorithm. The degree of invariance serves as the reward ( $\rho$ ), and it is simply defined as the dotproduct of successive response vectors at time  $t - 1$  and  $t$ :  $\rho_t = \mathbf{r}_t \cdot \mathbf{r}_{t-1}$ .  $R(s, a)$  is updated as:

$$R_t(s_{t-1}, a_{t-1}) = R_{t-1}(s_{t-1}, a_{t-1}) + \alpha \rho_t \quad (2)$$

where  $R_t(\cdot, \cdot)$  is the reward table at time  $t$  and  $\alpha$  the learning rate (set to 0.002 typically). Then,  $R_t(s_{t-1}, a)$  values are normalized by their sum for all  $a \in A$ .

### D. Results

After training the algorithm for 200,000 iterations on a natural image (Fig. 4), the reward table  $R(s, a)$  converged to a nearly ideal pattern (Fig. 6(c)). The gaze trajectory before learning is erratic and random-walk-like (Fig. 6(d)), but it becomes more organized and reveals the underlying image structure (Fig. 6(e)). See [22] for full results.

## III. RF DEVELOPMENT AND MOTION

The discussion in the preceding section did not take into account the fact that receptive fields are not fixed and given

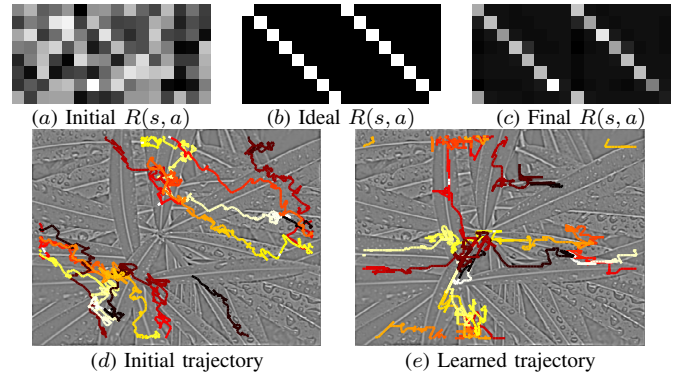


Fig. 6. **Learned  $R(s, a)$  and Behavioral Results.** The (a) initial, (b) ideal, and (c) final learned  $R(s, a)$ , and the (d) initial and (d) final gaze trajectories are shown. (a) Initial reward table is random. (c) The learned reward table shows the signature diagonal pattern of the ideal one (b). (d) Initial gaze trajectory is erratic, similar to a random walk. The repeating color coding from black to red to yellow to white represent time. Note that the DoG-filtered image is shown in the background just as a reference, and is not directly used in gaze generation. (e) Gaze trajectory after learning shows more structure, revealing the underlying image structure. Adapted from [22].

from the beginning. Further, for an agent residing in a specific environment, its sensory receptive fields should reflect the statistics of the external stimulus space. Below, we will investigate if receptive field development and grounding proceed simultaneously.

Neurons in the primary visual cortex respond to specific patterns of visual input. These patterns define the receptive field of the neurons (see [1] for a review). Thus, in some sense, the neurons “encode” visual features resembling their receptive fields.

How these encodings are learned has been intensively studied, both in neuroscience [28], [29] and in computational models [1], [30]–[35]. The focus of these earlier studies has been on the sensory aspect only (which was their prime limitation), i.e., the representational properties relating to natural image statistics and information theoretic concerns.

Only recently researchers started to raise questions about how subsequent stages of visual processing can utilize the response of the primary visual cortical neurons. For example, Salinas showed that functional requirements of downstream (motor) neurons determine the response property of early sensory neurons [15], and Sejnowski proposed that we should look at the “projective fields” as well as receptive fields [36].

In our previous work reviewed in Sec. II, we proposed a model based on reinforcement learning to allow subsequent stages discover the encoded feature properties in visual cortical spikes (i.e., “decode” the spikes, or semantically ground the representations), through the use of motor primitives [22], [24], [26], [37]. However, in that model, the receptive fields had fixed oriented Gabor patterns. In this section, we will extend our previous model to include receptive field development. Unlike approaches based on statistical properties of images alone, we propose that receptive field development should take into account the motor component.

There are other works where learning of receptive fields

are modulated by the motor system [17], [18], but the role (or goal) of the motor component was unclear. It is apparent that the involvement of the motor system will inevitably bias the statistical properties of the received input [38], and that will lead to differentiation in structure and function of the visual cortex [1], [30].

The real question however is, is there anything beyond that? In our case, the goal of such motor engagement is clear: autonomous semantic grounding of internal brain states. We present a model which uses (1) competitive learning to develop receptive fields and (2) reinforcement learning to link the visual cortical spikes to meaningful motor primitives. Our results suggest that receptive field development is strongly influenced by the motor system (confirming results by other researchers), and, more importantly, that visual receptive fields and the sensorimotor mapping that helps decode the spikes can develop simultaneously.

#### A. Receptive Field Learning

The agent not only learns the mapping from receptive fields to motor primitives but also adapts its receptive fields at the same time. We used competitive learning to adapt the RFs:

$$g_{ij} = \frac{g_{ij} + \alpha(I_{ij} - g_{ij})}{\sum_{mn} g_{mn} + \alpha(I_{mn} - g_{mn})}, \quad (3)$$

where  $g_{ij}$  is a weight of the RF to location  $(i, j)$  in the input, and  $I_{ij}$  the input at location  $(i, j)$ , and  $\alpha$  the learning rate. The  $\alpha$  value was positive for the RF with the highest response and negative for the other RFs. The magnitude of  $\alpha$  was initially 0.02 and 0.002 for the positive and the negative case respectively, and both decayed exponentially over time ( $\alpha_t = \alpha_{t-1} \times 0.99998$  where  $t$  is time step). Furthermore, receptive fields were trained only when the correlation in previous and current response vectors were above a certain level (typically 0.2), taking into account the invariance criterion.

#### B. Experiments and Results

In order to assess the influence of motor factors in receptive field learning, and to investigate whether receptive field development and internal-state-to-action mapping can be learned simultaneously, we conducted two experiments.

In the first experiment, we compared the development of receptive fields under two different, fixed internal-state-to-action mappings: (1) random reward table (Fig. 6(a)), and (2) ideal reward table (Fig. 6(b)). Again, these reward tables were fixed throughout the learning trial. With this, we could measure the effects of motor policy on receptive field learning. In this experiment, we trained 8 receptive fields on the natural image shown in Fig. 4. The agent was trained for 70,000 iterations. Fig. 7 shows that with random motor policy, the receptive fields are formed slowly and malformed (second and the last column), and the order is random whereas with ideal  $R(s, a)$  motor policy, the receptive fields are formed quickly and more importantly the order reflect the motor primitives. Thus, the property of the motor primitives (“downstream requirement”) dictates the receptive field property (cf. [15]).

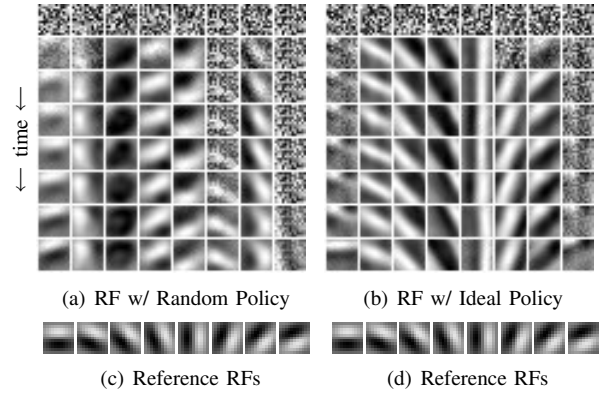


Fig. 7. **Learned receptive fields (RFs) using fixed policies.** (a, b) The fixed reward tables; (c, d) the learned receptive fields over time, from top to bottom; and (e, f) reference receptive fields (plain Gabor filters) are shown. In all plots, black represents min and white max. The receptive fields trained with a random policy has malformed receptive fields (second and last column), which is not the case for those trained with the ideal policy. Adapted from [23].

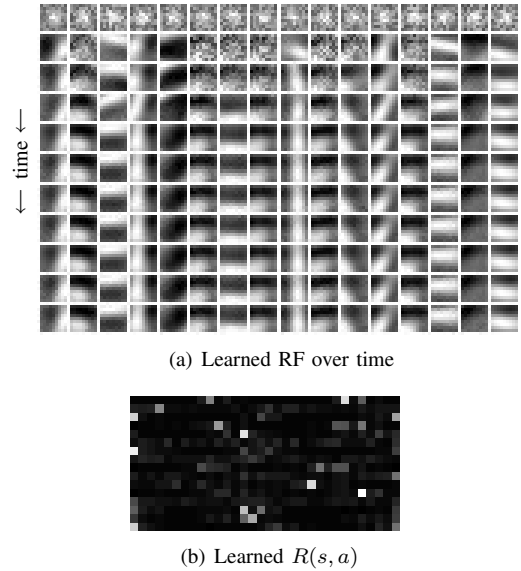


Fig. 8. **Learned Receptive Fields and  $R(s, a)$ .** (a) The learned RFs over time and (b) the final  $R(s, a)$  values are shown. Adapted from [23].

In the second experiment, we allowed both the receptive fields and the reward table to learn at the same time. The agent was given 16 sensory receptive fields (for finer representation of orientation) and thus had 32 corresponding motor primitives, and trained for 100,000 iterations, with other conditions identical to those in Sec. II. Fig. 8 shows the learned receptive fields over time from top to bottom, and the final  $R(s, a)$  values. The receptive fields learned reasonably well, but the reward table  $R(s, a)$  looks totally disorganized. However, this is simply because the ordering of receptive fields do not follow that of the standard, gradual change shown in Fig. 7(c). Reordering the learned receptive fields (their column location) gives Fig. 9(a). To reflect the new ordering, the reward table’s rows also need to be rearranged, which gives the reordered reward table Fig. 9(b), which shows the diagonal

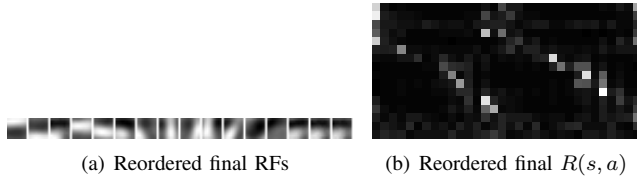


Fig. 9. **Reordered Learned Receptive Fields and  $R(s, a)$ .** (a) The receptive fields in Fig. 8 (the columns) were reordered to have a similar ordering as Fig. 7(c). (b) The rows in Fig. 8(b) were reordered accordingly, resulting in a reordered  $R(s, a)$ . Once reordered, the diagonal structure typical of an ideal reward table becomes apparent. Adapted from [23].

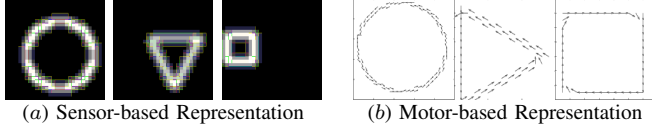


Fig. 10. **Sensory vs. Motor Representation of Shapes.** The (a) gray-scale bitmap-based sensory representation and (b) vector-based motor representation are compared. Adapted from [24].

pattern typical of an ideal case. Certain orientations are over expressed, such as the horizontal one, and as a result, there are more rows in the reward table associated with horizontal movement (see the first five rows in the reordered  $R(s, a)$ ).

#### IV. SHAPE RECOGNITION BASED ON MOTOR REPRESENTATIONS

The final topic of this paper is about visual shape recognition [24]. The question we wish to address here is how can the motor-based approach in the two sections above be extended so that it can handle more complex visual tasks such as shape recognition. To be more specific, what kind of internal representations of visual shapes enable invariant recognition (using “invariance” in the usual sense)? We believe Lashley’s concept of “motor equivalence” [39] is a good starting ground (see similar discussions in [8]). Motor equivalence is the ability to generate basically the same kind of motor behavior using different effectors (e.g., making a circular trajectory with the eyes, head, finger, arm, etc.). Lashley’s main idea is that there should be one shared motor representation that express itself in these different behavioral output.

Here, we assess the relative merit of motor-based representations compared to sensor-based representations. Given a particular shape as in Fig. 10(a), processes similar to the sensorimotor mapping method we described above would result in a series of motor vectors as shown in Fig. 10(b). The question is which one of these representations are easier to learn and to generalize?

We generated one thousand random grayscale images ( $30 \times 30$ ) containing circles, triangles, and squares of different sizes at different locations (Fig. 10(a)), and traced these shapes to generate an equivalent 900-dimensional vector containing a sequence of gaze directions (Fig. 10(b)). (Note that the motor-based representations are essentially normalized for translation and scale, but not for rotation.) The inputs were fed through a multilayer perceptron, trained using backpropagation (see [24] for details). 75% of the input was used for training and

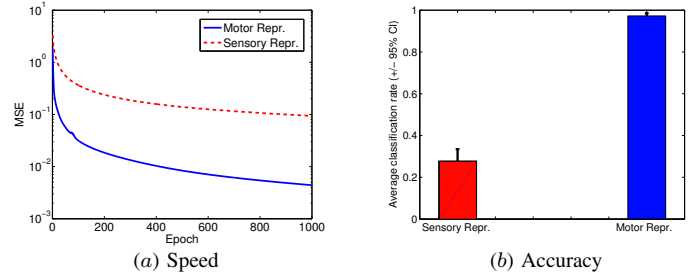


Fig. 11. **Speed and Accuracy of Learning of Shape Representations.** A comparison of (a) learning speed and (b) accuracy of sensory vs. motor representations are shown. (a) The mean squared error (MSE) during training shows that motor representations reach lower levels of error faster. (b) The generalization performance (classification rate on novel test inputs) is significantly higher ( $p = 0, n = 10$ ) for the motor representations (97%) than the sensory representations (28%). Adapted from [24].

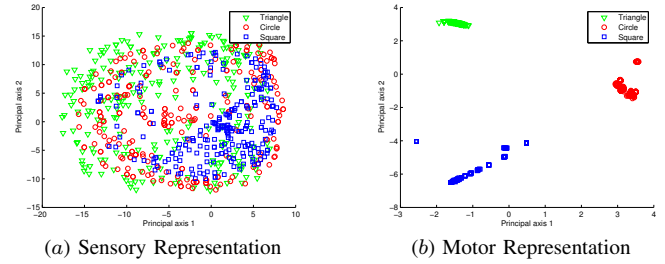


Fig. 12. **Principal Components Analysis of Sensory vs. Motor Representation of Shapes.** The projection of the data points (all three shape categories) onto the 1st and the 2nd PCA axes are shown for the sensory vs. motor representations. (a) The projections of the three classes are inseparable for the sensory representation, while (b) they are clearly separable for the motor representation. Adapted from [24].

25% for testing. The results are reported in Fig. 11. Both the training time and generalization performance on the test set was much better for the motor-based representation.

A subsequent analysis using principal components analysis (PCA) explains the superior learnability of the motor representations (Fig. 12). The sensory representations of the shape classes are broadly scattered and overlap with each other, whereas the motor representations are concentrated into clusters and are separated.

These results show that motor-based internal representations could be a good underlying mechanism for invariant shape recognition. The motor-based representations have built-in invariances, but it could be argued that extracting such invariances is easier through the mobilization of the motor system than through static visual analysis.

#### V. DISCUSSION AND CONCLUSION

The main contribution of our work summarized in this paper is to have shown the critical role played by the motor system in grounding, RF learning, and shape recognition. Our investigation is strongly grounded in neuroscience research, which allows us to bring in theoretical frameworks into neuroscience investigations. The framework that we presented here is currently focused on the visual cortex, but it is quite general enough so that it can be generalized to other cortical regions



such as the auditory cortex or the somatosensory cortex. Whether the same is true for more qualitative modalities such as color or olfaction is yet an open question.

One important implication of our work is that an organism's understanding of the external world is fundamentally bounded by the types of motor primitives available to the organism. Recent results by Hatsopoulos et al. [40] showed how these motor primitives (what they call "pathlets") are encoded and laid out in the motor cortex. Linking such an organization to the sensory counterpart could lead to important discoveries.

In summary, we have shown how the motor system plays a key role in visual function, and we expect our approach to help us better understand the development and function of the visual cortex. In the long run, similar approaches will allow us to build more autonomous agents that have inherent understanding of itself and the environment.

#### ACKNOWLEDGMENT

This overview paper is based on [22]–[24], [41], [42].

#### REFERENCES

- [1] R. Miikkulainen, J. A. Bednar, Y. Choe, and J. Sirosh, *Computational Maps in the Visual Cortex*. Berlin: Springer, 2005, uRL: <http://www.computationalmaps.org>.
- [2] J. Dewey, "The reflex arc in psychology," *Psychological Review*, vol. 3, pp. 357–370, 1896, reprinted in [43], pp.136–148.
- [3] H. Bergson, *Matter and Memory*. New York, NY: Zone Books, 1988, translated by Nancy Margaret Paul and W. Scott Palmer.
- [4] L. Wittgenstein, *Philosophical Investigations*. Oxford: Blackwell, 1953, G.E.M. Anscombe and R. Rhees (eds.), G.E.M. Anscombe (trans.).
- [5] F. J. Varela, E. Thompson, and E. Rosch, *The Embodied Mind: Cognitive Science and Human Experience*, 1993.
- [6] N. Humphrey, *A History of the Mind*. New York: HarperCollins, 1992.
- [7] J. K. O'Regan and A. Noë, "A sensorimotor account of vision and visual consciousness," *Behavioral and Brain Sciences*, vol. 24(5), pp. 883–917, 2001.
- [8] R. R. Llinás, *I of the Vortex*. Cambridge, MA: MIT Press, 2001.
- [9] R. C. Schank and R. P. Abelson, *Scripts, Plans, Goals, and Understanding: An Inquiry Into Human Knowledge Structures*, 1977.
- [10] H. von Foerster, *Understanding Understanding*. New York: Springer, 2003.
- [11] L. M. Rocha, "Eigenbehavior and symbols," *Systems Research*, vol. 13, pp. 371–384, 1996.
- [12] W. J. Freeman, *How Brains Make Up Their Minds*. London, UK: Wiedenfeld and Nicolson Ltd., 1999, reprinted by Columbia University Press (2001).
- [13] A. D. Milner and M. A. Goodale, "Visual pathways to perception and action," *Progress in Brain Research*, vol. 95, pp. 317–337, 1993.
- [14] G. Rizzolatti, L. Fadiga, V. Gallese, and L. Fogassi, "Premotor cortex and the recognition of motor neurons," *Cognitive Brain Research*, vol. 3, pp. 131–141, 1996.
- [15] E. Salinas, "How behavioral constraints may determine optimal sensory representations," *PLoS Biology*, vol. 4, pp. 2383–2392, 2006.
- [16] R. Pfeifer and C. Scheier, *Understanding Intelligence*. Cambridge, MA: MIT Press, 1999.
- [17] D. Floreano, M. Suzuki, and C. Mattiussi, "Active vision and receptive field development in evolutionary robots," *Evolutionary Computation*, vol. 13, pp. 527–544, 2005.
- [18] R. Wyss, P. König, and P. F. M. J. Verschure, "A model of the ventral visual system based on temporal stability and local memory," *PLoS Biology*, vol. 4, p. e120, 2006.
- [19] R. A. Brooks, "Intelligence without representation," *Artificial Intelligence*, vol. 47, pp. 139–159, 1991.
- [20] C. Breazeal and B. Scassellati, "Robots that imitate humans," *Trends in Cognitive Sciences*, vol. 6, pp. 481–487, 2002.
- [21] J. Weng, J. L. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen, "Autonomous mental development by robots and animals," *Science*, vol. 291, no. 5504, pp. 599–600, 2001.
- [22] Y. Choe, H.-F. Yang, and D. C.-Y. Eng, "Autonomous learning of the semantics of internal sensory states based on motor exploration," *International Journal of Humanoid Robotics*, vol. 4, pp. 211–243, 2007, <http://dx.doi.org/10.1142/S0219843607001102>.
- [23] H.-F. Yang and Y. Choe, "Co-development of visual receptive fields and their motor-primitive-based decoding scheme," in *Proceedings of the International Joint Conference on Neural Networks 2007 Post conference Workshop on Biologically-inspired Computational Vision (BCV) 2007*, 2007, [Online].
- [24] N. Misra and Y. Choe, "Shape recognition through dynamic motor representations," in *Neurodynamics of Higher-Level Cognition and Consciousness*, R. Kozma and L. Perlovsky, Eds. Berlin: Springer, 2007, pp. 185–210.
- [25] D. Philipona, J. K. O'Regan, and J.-P. Nadal, "Is there something out there? Inferring space from sensorimotor dependencies," *Neural Computation*, vol. 15, pp. 2029–2050, 2003.
- [26] Y. Choe and N. H. Smith, "Motion-based autonomous grounding: Inferring external world properties from internal sensory states alone," Y. Gil and R. Mooney, Eds., 2006, 936–941.
- [27] S. Harnad, "The symbol grounding problem," *Physica D*, vol. 42, pp. 335–346, 1990.
- [28] B. Chapman, M. P. Stryker, and T. Bonhoeffer, "Development of orientation preference maps in ferret primary visual cortex," vol. 16, no. 20, pp. 6443–6453, 1996.
- [29] L. C. Katz and C. J. Shatz, "Synaptic activity and the construction of cortical circuits," *Science*, vol. 274, pp. 1133–1138, 1996.
- [30] Y. Choe and R. Miikkulainen, "Contour integration and segmentation in a self-organizing map of spiking neurons," *Biological Cybernetics*, vol. 90, pp. 75–88, 2004.
- [31] B. A. Olshausen and D. J. Field, "Sparse coding with an overcomplete basis set: A strategy employed by v1?" *Vision Research*, vol. 37, pp. 3311–3325, 1997.
- [32] A. J. Bell and T. J. Sejnowski, "The "independent components" of natural scenes are edge filters," *Vision Research*, vol. 37, p. 3327, 1997.
- [33] T.-W. Lee, T. Wachtler, and T. J. Sejnowski, "The spectral independent components of natural scenes," in *Lecture Notes in Computer Science, V.1811*, S.-W. Lee, H. H. Bülthoff, and T. Poggio, Eds. Springer, 2000, pp. 527–534.
- [34] S. Grossberg, "On the development of feature detectors in the visual cortex with applications to learning and reaction-diffusion systems," *Biological Cybernetics*, vol. 21, pp. 145–159, 1976.
- [35] C. von der Malsburg, "Self-organization of orientation-sensitive cells in the striate cortex," *Kybernetik*, vol. 15, pp. 85–100, 1973, reprinted in [44].
- [36] T. J. Sejnowski, "What are the projective fields of cortical neurons?" in *Twenty Three Problems in Systems Neuroscience*, L. J. van Hemmen and T. J. Sejnowski, Eds. Oxford, UK: Oxford University Press, 2006, pp. 394–405.
- [37] Y. Choe and S. K. Bhamidipati, "Autonomous acquisition of the meaning of sensory states through sensory-invariance driven action," in *Biologically Inspired Approaches to Advanced Information Technology*, ser. 3141, A. J. Ijspeert, M. Murata, and N. Wakamiya, Eds. Berlin: Springer, 2004, pp. 176–188.
- [38] P. Reinagel and A. M. Zador, "Natural scene statistics at the center of gaze," *Network: Computation in Neural Systems*, vol. 10, pp. 341–350, 1999.
- [39] K. S. Lashley, "The problem of serial order in behavior," in *Cerebral Mechanisms in Behavior*, L. A. Jeffress, Ed. New York: Wiley, 1951, pp. 112–146.
- [40] N. G. Hatsopoulos, Q. Xu, and Y. Amit, "Encoding of movement fragments in the motor cortex," *Journal of Neuroscience*, vol. 27, pp. 5105–5114, 2007.
- [41] Y. Choe and H.-F. Yang, "Co-development of visual receptive fields and their motor primitive-based decoding scheme," in *Society for Neuroscience Abstracts*. Washington, DC: Society for Neuroscience, 2006, program No. 734.6. Online.
- [42] N. Misra, "Comparison of motor-based versus visual representations in object recognition tasks," Master's thesis, Department of Computer Science, Texas A&M University, College Station, Texas, 2005.
- [43] J. J. McDermott, Ed., *The Philosophy of John Dewey*. Chicago, IL: University of Chicago Press, 1981.
- [44] J. A. Anderson and E. Rosenfeld, Eds., *Neurocomputing: Foundations of Research*, 1988.