Motor-Based Autonomous Grounding in a Model of the Fly Optic Flow System

Amey Parulkar
Department of Computer Science
Texas A&M University
College Station, TX 77843-3112
Email: ameyp20191@tamu.edu

Yoonsuck Choe
Department of Computer Science
Texas A&M University
College Station, TX 77843-3112
Email: choe@tamu.edu

***Abstract*—The fly visual system, although tiny when compared to the mammalian visual system, can still perform highly sophisticated spatial tasks like collision avoidance, landing on objects, pursuit of prey, etc. Flies outperform human-made autonomous flying systems in solving such spatial tasks by a long way. This is partly due to their ability to perceive and respond to optical flow generated by motion in the environment. They are also known to take actions that actively shape the image flow on their eyes. Higher level neurons in the fly visual system respond to different types of complex optical flows due to rotation and translation, by pooling information from local motion detectors called the elementary motion detectors (EMDs) in the lower level. In this sense, neuronal responses (spikes) from these optical flow detectors in the fly carry highly encoded signals: a single spike can represent a complex dynamical pattern of movement in the visual field. In this paper, we investigate how such highly encoded signals can be interpreted and utilized in the fly's brain, while solely operating on the internal spike patterns within their brain and no direct external sensory information, i.e. a form of grounding. With a computational model of the optical flow detectors based on those in the fly, we show that action (or coordinated motor output) is the only way that a fly can decode its internal spikes and generate meaningful, relevant behavior based on that.**

I. Introduction

The fly visual system can perform highly sophisticated functions such as detecting optical flow, to help the animal navigate through the environment (see e.g. [1]). The behavior and anatomy of the optical flow system in the fly have been extensively studied [2], [3], [4], [5], [6] and modeled computationally [6], [7], [8]. Perceiving and responding to the optical flow is critical to the animal's survival and success, since based on this they perform rapid and complex maneuvers such as stabilization and pursuit.

Higher level neurons in the fly visual system respond to different types of optical flow due to rotation and translation of the animal, by pooling information from elementary motion detectors (EMDs) in the lower level [9], [10], [11]. In this sense, neuronal responses (spikes) from these optical flow detectors in the fly carry highly encoded signals.

In this paper, we investigate how such highly encoded optical flow signals can be interpreted and utilized within the fly's brain, while solely operating on the internal spike patterns within its brain. For example, just observing incoming spikes might look like the situation shown in Fig. 1. (Can you tell what these activity patterns represent?)

This is basically a problem of "grounding" [12], [13], [14], i.e., trying to figure out the meaning of internal representations in the brain. Strangely, the grounding problem seems trivial from an external observer's point of view, with full access to the external stimulus (Fig. 2, Fig. 3b), while it seems intractable from an internal observer's view (the brain's native view; Fig. 1, Fig. 3a).

Fig. 1: **A Clueless Situation.** What do these light spots (or activity patterns) represent? [15]

In our previous work on grounding visual cortical orientation detector responses, we showed that action is critical in making the internal observer case (Fig. 3a) feasible [16], [17], [15]. See the "Autonomous Grounding ..." section for details.

The main contributions of this paper are:

1) The main goal of our paper is to introduce the concept of grounding (in AI) to neuroscience, by using the fly as an example and showing how the fly can operate on internal encoded neural spikes alone within its brain to generate ecologically meaningful action.
2) With a computational model of the optical flow detectors (neurophysiology) in fly, we show that action (or coordinated motor output) is the only way that the fly can decode its internal spikes and generate meaningful, relevant behavior, all just based on its internal spike patterns.
3) We use learning algorithm based on reinforcement learning (AI) as a tool to show a sensory state to action mapping.

II. Background: The Fly Visual System

The fly's visual system is briefly described below (closely following the description in [5]). The fly has 2 compound eyes in the anterior region of its body giving it a wide panoramic

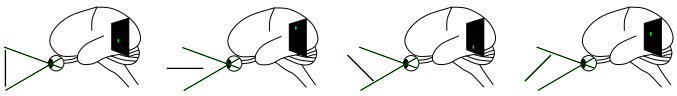
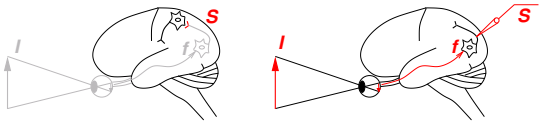


Fig. 2: **The Clueless Situation, Explained.** The four light spots are neural spikes in the visual cortex representing four different orientations in the visual field. Adapted from [18].



(a) Internal observer (b) External observer

Fig. 3: **Decoding neural spikes.** Decoding the meaning of neural spikes from (a) an internal observer's perspective (the brain itself) where only the spikes S are available, compared to (b) an external observer's perspective where both the spikes S and the environmental stimulus I are available to the observer. In (a), the grayed-out part is unknown to the neuron receiving the spike. Adapted from [19].

field of view. Each compound eye consists of tens to hundreds of thousands of ommatidia that form a hexagonal lattice. They are involved in sensing directional motion. The fly optical flow system consists of three neuropils and retina, which is connected to the hexagonal eye lattice of the compound eye. The retina houses the photoreceptors that respond to light stimuli while the three neuropils - lamina, medulla, and lobula plate are involved in processing motion information. As a system that uses visually detected motion information, it has two important properties: (1) spatially distributed local directional motion sensor and (2) mechanism to integrate this information to create a global picture.

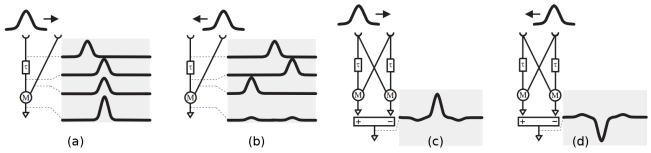


Fig. 4: **Elementary Motion Detectors (EMD)**. EMDs detect local directional information by spatio-temporal correlation of light intensity of neighboring location on the retina. (a) Half-Detector, produces output by multiplying delayed input from point 1 to input from point 2, depending on the direction of point's alignment (preferred direction), the output can be positive or (b) negative for null direction. (c) and (d) are fully directional-selective EMDs, outputs of two mirror-symmetrical half-detectors are subtracted from each other, giving positive response to motion in preferred direction while negative for null direction motion. Adapted from [20].

The fly has an array of local motion detectors called Elementary motion detectors (EMDs) spanning the visual field. EMD is the theoretical model of neural mechanism present between lamina and medulla that process the changes in the light intensity as sensed by photoreceptor cells to generate local directional motion information. The EMD are thought to be based on Reichardt Detector model (see reviews [10], [11], [20]) as shown in Fig. 4. Their response depends directly on the velocity and direction of visual motion. In computer vision, similar spatio-temporal correlation based methods are available for optical flow generation. We use Lucas-Kanade method [21] and Horn-Schunck methods [22] for our agent design.

Finally, integration of local motion information is done by 60 Lobula Plate Tangential cells (LPTC), present in the lobula plate neuropil, to produce response commensurate with global optical flow. Out of these, output LPTCs - 3 horizontal systems cells (HS) and ten vertical systems cells (VS) are the most important for processing wide-field optical flow generated by fly self-motion. The spikes generated by these neurons represent a complex dynamical pattern of movement in the visual field. Horizontally arranged HS cells respond mainly to horizontal motion while vertically arranged VS cells respond to vertical motion (Fig. 5). VS cells are modeled on the idea of matched filters [23] which says that VS cells respond to certain definite motions like roll and pitch instead of simple response to plain vertical motion. The premotor descending neurons, postsynaptic to LPTCs decide the motor response based on LPTC output and accordingly direct the motor neurons. More complex models of fly optical flow system exists like communication among VS cells model [4] but are of little importance to our purpose.

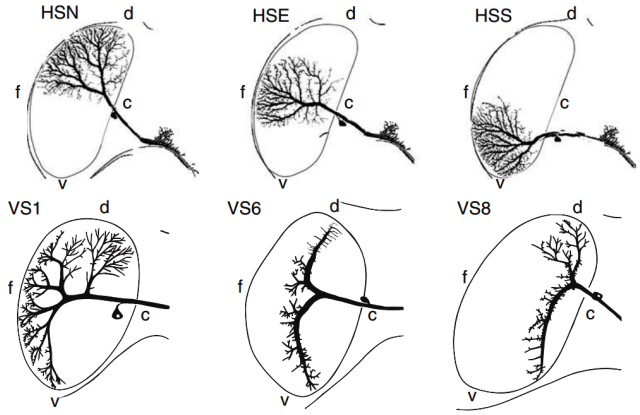


Fig. 5: **Distribution of dendrites of the fly's right eye Lobula Plate Tangential Cells (LPTCs).** HS dendrites (top row) span across the frontal(f)-caudal(c) direction responding best to front-back motion. VS dendrites (bottom row) span across the dorsal(d)-ventral(v) direction responding best to up-down motion.The sensitivity of neuron's response across the field of view depends on density of dendrites. Right eye: Elevation Range-[v,d]=$[-45°, 45°]$, Azimuth Range-[f,c]=$[-30°, 90°]$. Left eye: Elevation Range-[v,d]=$[-45°, 45°]$, Azimuth Range-[f,c]=$[-90°, 30°]$. Adapted from [5].

III. Autonomous Grounding Through Sensorimotor Interaction

Consider a simple sensorimotor agent that is capable of sensing motion information in the form of optical flow and can perform a certain set of actions. The agent receives motion information from a moving scene which is processed by motion detector filters (modeled on the fly's optical flow processing neurons) to produce spike pattern in the sensory array. The design of these filters is based on the structure, arrangement and sensitivity of neurons responsible for processing optical flow in the fly's visual system. The sensory array indicates the sensory state s which is used by the agent to infer the semantics of the motion information and how it is related to the external environmental changes. The agent does not have direct access to external input I nor understands the functional properties of filters f. This situation is similar to Fig. 3a, where external world properties need to be inferred based on internal spikes only. It is not possible for the agent to associate motion information with the sensory spikes by passive observation of these spikes. However, a solution is possible when the agent uses motor primitives to actively shape the external input (it is known that flys actively shape their visual input through action: [1]), which in turn changes the internal state.

Now assume that the agent receives an input in the form of optical flow generated by relative motion between the agent and the environment. The motion detector filters process this information and trigger a specific sensory state corresponding to the perceived optical flow. Based on this encoded state, how can the agent understand what the encoding means? A critical insight in our previous work [15], [16], [17] was that a specific pattern of action that maintains the internal state invariant over time is bound to embody the same stimulus property encoded by that internal state (also see [24], [25], [26]). For example, if the perceived optical flow generates "rotate clockwise" state, the agent rotating counterclockwise will keep the "rotates clockwise" state invariant over time (the *invariance criterion*). Thus, by seeking action that maintains invariance in its internal sensory state representation, the agent can understand what these encoded internal states mean, in terms of its own actions. Even with no prior knowledge, no direct access to external world, the agent can still learn the meaning of motion in the environment via its internal sensory states. Below, we elaborately describe the model of the agent, propose a sensorimotor learning algorithm based on reinforcement learning that is able to learn sensory state to action mapping with sensory invariance as the reward signal.

IV. Computational Model of the Fly Visual System

In this section, we describe the computational model of the agent's optical flow detection system and the sensory state to action mapping algorithm. We start with input processing, response generation and finally the learning rule that enables agent to learn the sensory state to motor response mapping.

(a) Synthetic (b) Natural 1 (c) Natural 2

Fig. 6: **Images of input backgrounds used for training.**

A. Input Processing

The agent observes a background (Fig. 6) that is moving in 3D at a constant speed in a particular manner. For training, the background is moved to generate pitch, roll, yaw, and zooming (radiating) optical flows, in different instances with respect to the agent (Fig. 7). These optical flows can be generated by rotating or translating the agent along certain axes and directions with respect to the background. The agent observes these optical flows as projected on the plane of its visual sensors. The agent's local motion detectors based on the same principles as those of the fly's EMDs detect this motion and produce a signal corresponding to the direction and velocity of the local motion. These detectors together span the entire visual field.

The following equations govern the optical flow calculation:

$$I(x, y, z) = I(x + \Delta x, y + \Delta y, t + \Delta t) \tag{1}$$

where intensity $I(x, y, t)$ at point (x, y, t) is moved by Δx,Δy over time Δt. It is assumed that the intensity does not change over small magnitude of motion. Using Taylor series expansion, we get, $\frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t = 0$ which results in, $\frac{\partial I}{\partial x}V_x + \frac{\partial I}{\partial y}V_y + \frac{\partial I}{\partial t} = 0$ where $\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial I}{\partial t}$ are derivatives of image at (x, y, t) in corresponding directions. This equation is the standard equation for optical flow calculation. This equation has two unknowns (velocity), V_x and V_y. It can only be solved by having another assumption. We use Lucas-Kanade method [21] to implement the agent's local motion detectors. It assumes that the flow is constant in the local neighbourhood of the pixel (x, y) to solve this ambiguity. This differential method gives response that is linearly dependent on velocity and direction of motion. Finally, the optical flow is normalized over the entire field of view using their l_2-norm.

B. Sensory primitives

The sensory state is generated by the modeled HS and VS cells using the optical flow generated by local motion detectors. The agent has two eyes that receive input from different regions in the visual field. Each eye, left and right, has its own set of mHS (Modeled HS) and mVS (Modeled VS) neurons. However, the overall sensory state is the output of all the neurons of both eyes, as they all act as inputs for the premotor decision system. To keep the design simple, we consider that the agent has 3 mHS cells and 5 mVS neurons corresponding to each eye. These cells respond to inputs from their corresponding regions of the visual field. The visual field is 90° wide in elevation and 180° wide in azimuth. The sensitivities of HS and VS cells to motion depend upon the the

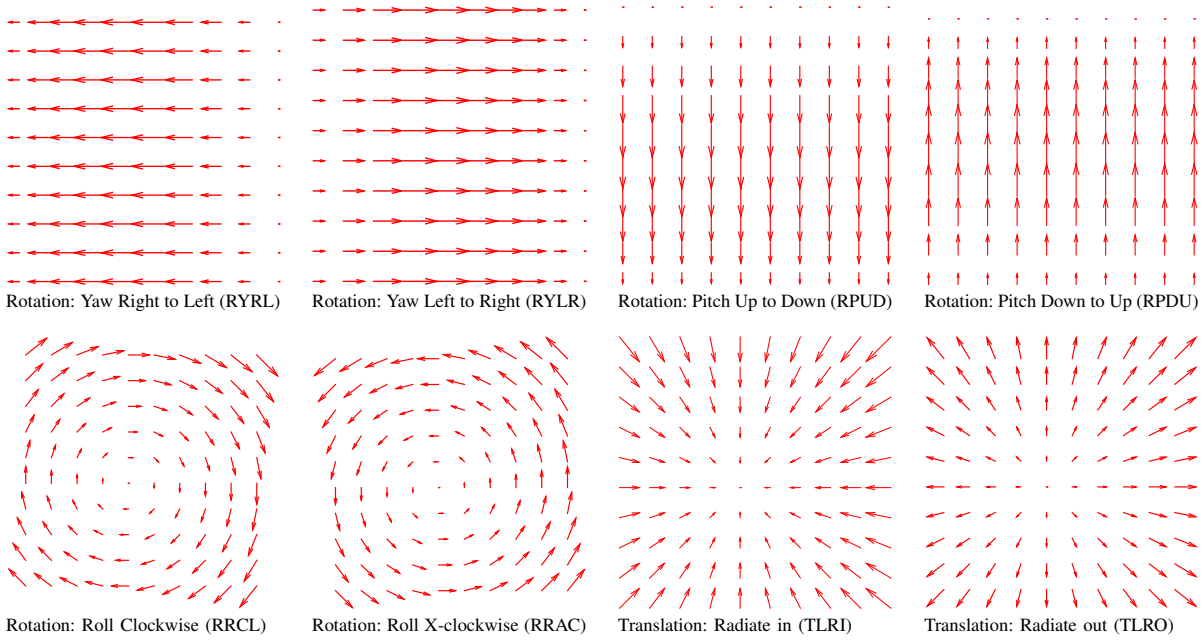


Fig. 7: **Optical Flows (Ideal).** Optical flows (as observed by the agent) induced by rotation (yaw, pitch, roll), and translation (forward and backward only) are shown. These are standard optical flow patterns (see e.g., [27]).

mHS	**Elev. Range**	**Azim. Range**	(ϕ_m, θ_m)	σ_θ	σ_{ϕ^+}	σ_{ϕ^-}
R mHSN	$[9°, 45°]$	$[-30°, 90°]$	$(11.25°, 27°)$	$15°$	$75°$	$33.75°$
L mHSN	$[9°, 45°]$	$[-90°, 30°]$	$(-11.25°, 27°)$	$15°$	$33.75°$	$75°$
R mHSE	$[-18°, 18°]$	$[-30°, 90°]$	$(11.25°, 0°)$	$15°$	$75°$	$33.75°$
L mHSE	$[-18°, 18°]$	$[-90°, 30°]$	$(-11.25°, 0°)$	$15°$	$33.75°$	$75°$
R mHSS	$[-45°, -9°]$	$[-30°, 90°]$	$(11.25°, -27°)$	$15°$	$75°$	$33.75°$
L mHSS	$[-45°, -9°]$	$[-90°, 30°]$	$(-11.25°, -27°)$	$15°$	$33.75°$	$75°$

(a) mHS

mVS	(ϕ_c, θ_c)	**mVS**	(ϕ_c, θ_c)
R mVS1	$(90°, 15°)$	L mVS3	$(0°, 0°)$
L mVS1	$(-90°, 15°)$	R mVS4	$(22.5°, -2°)$
R mVS2	$(90°, -15°)$	L mVS4	$(-22.5°, -2°)$
L mVS2	$(-90°, -15°)$	R mVS5	$(45°, -4°)$
R mVS3	$(0°, 0°)$	L mVS5	$(-45°, -4°)$

(b) mVS (matched filter)

TABLE I: **Parameter values of modeled HS cells (mHS) and modeled VS cells (mVS, matched filter).**

spanning direction and density of the dendritic ramifications across the field (Fig. 5). The weighted response to optical flow by mHS and mVS neurons is given as a two dimensional Gaussian function (see [7], [8]). We model the mHS neurons as,

$$w(\phi,\theta) = \exp\left(-\left(\frac{\theta-\theta_m}{\sigma_\theta}\right)^2 - \left(\frac{\phi-\phi_m}{\sigma_{\phi^+}}\right)^2\right), \text{if } \phi \geq \phi_m$$
$$= \exp\left(-\left(\frac{\theta-\theta_m}{\sigma_\theta}\right)^2 - \left(\frac{\phi-\phi_m}{\sigma_{\phi^-}}\right)^2\right), \text{otherwise,} \tag{2}$$

where θ is the elevation angle, ϕ is the azimuth angle, and (ϕ_m, θ_m) is the center of receptive field for that neuron (Fig. 8). The angular width for elevation is σ_θ while for azimuth it is σ_{ϕ^+} and σ_{ϕ^-}. The parameter values for all 6 mHS neurons are shown in Table Ia.

For mVS neurons (Fig. 8),

$$w(\phi,\theta) = \exp\left(-\left(\frac{\theta-\theta_m}{\sigma_{\theta^+}}\right)^2 - \left(\frac{\phi-\phi_m}{\sigma_\phi}\right)^2\right), \text{if } \theta \geq \theta_m$$
$$= \exp\left(-\left(\frac{\theta-\theta_m}{\sigma_{\theta^-}}\right)^2 - \left(\frac{\phi-\phi_m}{\sigma_\phi}\right)^2\right), \text{otherwise,} \tag{3}$$

where θ is the elevation angle, ϕ is the azimuth angle, and (ϕ_m, θ_m) is the center of receptive field for that neuron. The angular width for azimuth is σ_ϕ while for elevation it is σ_{θ^+} and σ_{θ^-}. The parameter values for all 10 mHS neurons are shown in Table II.

Finally, we normalize the weighted response by

$$w(\phi,\theta) = \frac{w(\phi,\theta)}{\sum_{\phi,\theta} w(\phi,\theta)}. \tag{4}$$

Each model neuron integrates local optical flow information from a region that they respond (Tables Ia and II) to generate a response commensurate with global optical flow. These neurons give positive response in their preferred direction and negative response in the opposite direction. The preferred direction of mHS neurons is from center to side of the field

while that of mVS neurons is from up to down. Also, mVS cells are matched to respond to rotational fields about the axes (as shown in Table Ib) in their region of response. The response of all neurons form a column vector r with each element r_i corresponding to an individual neuron's response. r_i is the weighted sum of net motion in the neuron's preferred direction, given as $r_i = v_{i^+} - v_{i^-}$ where,

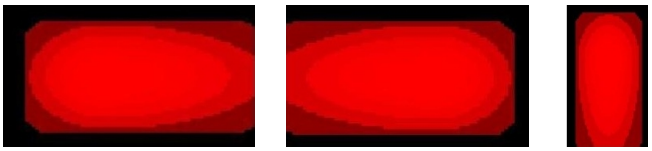
$$\begin{aligned} v_{i^d} &= \sum_{\phi,\theta} w_i(\phi,\theta) m_d(\phi,\theta), \text{ for mHS neurons} \\ &= \sum_{\phi,\theta} w_i(\phi,\theta)(F(\phi,\theta) \cdot \mathrm{MF}_{i,d}(\phi_c,\theta_c)), \text{ for mVS neurons.} \end{aligned} \tag{5}$$

Here $i = 16$ is the total number of neurons, $d = (+/-)$, $w_i(\phi,\theta)$ is the weight distribution of i^{th} neuron, $m_d(\phi,\theta)$ is the normalized optical flow in neuron's preferred direction ($d = +$) or null direction ($d = -$). $\mathrm{MF}_i(\phi_c,\theta_c)$ is the matched filter mask (rotation along center (ϕ_c,θ_c) as shown in Table Ib) of i^{th} neuron, mVS neuron, in its preferred direction ($d = +$) or null direction ($d = -$). $F(\phi,\theta)$ is the normalized global optical flow.

The sensory response vector r is normalized using the l_2-norm: $r = \frac{r}{|r|}$. The agent then compares this generated response to reference responses r_{ref} that are the ideal standard optical flows (Fig. 7)- yaw, pitch, roll, zoom (radiate): $s = \underset{\psi_i, i=1\ldots n}{\arg\max}(r \cdot r_{\mathrm{ref}(i)})$ where s is the sensory state that corresponds to $i = 8$ types of optical flows. This sensory state is what the agent is trying to figure out the meaning of.

mVS	Azimuth Range	(ϕ_m, θ_m)
R mVS1	$[-30°, 0.4°]$	$(-14.8°, 15°)$
L mVS1	$[-0.4°, 30°]$	$(14.8°, 15°)$
R mVS2	$[-7.6°, 22.8°]$	$(7.6°, 15°)$
L mVS2	$[-22.8°, 7.6°]$	$(-7.6°, 15°)$
R mVS3	$[14.8°, 45.2°]$	$(30°, 15°)$
L mVS3	$[-45.2°, -14.8°]$	$(-30°, 15°)$
R mVS4	$[37.2°, 67.6°]$	$(52.4°, 15°)$
L mVS4	$[-67.6°, -37.2°]$	$(-52.4°, 15°)$
R mVS5	$[59.6°, 90°]$	$(74.8°, 15°)$
L mVS5	$[-90°, -59.6°]$	$(-74.8°, 15°)$

TABLE II: **Parameter values of modeled VS cells (mVS).** Some parameters that are common among all mVS are: Elevation Range=$[-45°, 45°]$, $\sigma_\phi = 24°$, $\sigma_{\theta^+} = 24°$ and $\sigma_{\theta^-} = 37.5°$.



(a) mHS (Right) (b) mHS (Left) (c) mVS

Fig. 8: **Weight distribution of mVS and mHS neurons.** The 2-D Gaussian plots show weight distribution of mVS and mHS neurons across the visual field based on dendrite distribution and density pattern (Fig. 5).

C. Learning Algorithm

The agent uses reinforcement learning algorithm to learn the sensory state to action mapping [16][17]. The learning happens postsynaptic to the mHS and mVS neurons. The output of the learning mechanism controls the motor actions. Consider that the sensory state is s_t at time t. The agent performs an action a_t, which leads to change in sensory state to s_{t+1} at time $t+1$. This transition of state really depends upon change in flow patterns as observed by the agent due to active shaping of motion input through action a_t. As discussed, the agent wants to minimize the variation in state (invariance criterion) while performing actions. This means that the reward that the agent gets should be inversely proportional to variation in current state s_t. The immediate reward ρ_{t+1} is calculated as, $\rho_{t+1} = 1/\sqrt{\sum_i (r_{t+1,i} - r_{t,i})^2}$ where $r_{t,i}$ is the i^{th} element of the sensory response vector r_t at time t. Here the reward is calculated as the inverse of the Euclidean distance between the current state response vector r_t and the next state response vector r_{t+1}, as opposed to using simple binary values of variant or invariant. Since the denominator in the equation can approach zero, leading to very high ρ_{t+1}, an upper bound of $\rho_{max} = 100$ is enforced.

Now, the agent has to learn a state-action mapping such that the reward ρ_t is maximized at time t. The agent has to learn to take action a_t such that the variation in current state (s_t) to next state (s_{t+1}) is minimized. Since, the state transition is probabilistic (to avoid greedy exploitation) so the problem is to determine $P(a_t|s_t)$, which is the conditional probability of taking action a_t at state s_t resulting in s_{t+1} that is highly likely to be the same as s_t. Let us call this the reward probability function $R(s_t, a_t)$ (Fig. 9). The learning algorithm is executed in 2 phases, after initializing the $R(s,a)$ table randomly.

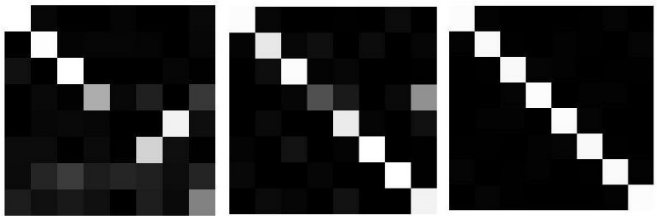
1) **Phase I:** Given the sensory state is $s_t \in S$. For 200 iterations,
   a) Randomly pick action $a_t \in A$.
   b) Perform action a_t.
2) **Phase II:** Given the sensory state is $s_t \in S$.
   a) Randomly pick action $a_t \in A$.
   b) If a_t is $\underset{a \in A}{\arg\max} R(s_t, a)$
      i) Then perform action a_t
      ii) else perform action a_t with probability $R(s_t, a_t)$.
   c) Repeat (a) and (b) until exactly one action is performed.

During Phase I, the learning algorithm selects the actions uniformly in random fashion to make sure that no action gets preference due to randomly generated initial values of reward probability function table. Enough iterations are allowed to sufficiently nullify this initial value bias. The $R(s,a)$ entries (Fig. 9) are updated after performing selected actions every iteration. This is done as follows:

$$R_{t+1}(s_t, a_t) = R_t(s_t, a_t) + \alpha \rho_{t+1} \tag{6}$$

$R(s,a)$	Direction of Motion (a)							
Optical Flow States (s)	→	←	↑	↓	↺	↻	⊙	⊕
RYRL	1	0	0	0	0	0	0	0
RYLR	0	1	0	0	0	0	0	0
RPUD	0	0	1	0	0	0	0	0
RPDU	0	0	0	1	0	0	0	0
RRCL	0	0	0	0	1	0	0	0
RRAC	0	0	0	0	0	1	0	0
TLRI	0	0	0	0	0	0	1	0
TLRO	0	0	0	0	0	0	0	1

Fig. 9: **Reward table $R(s,a)$ with ideal reward values.** Actions: 1=Yaw Left to Right, 2=Yaw Right to Left, 3=Pitch Down to Up, 4=Pitch Up to Down, 5=Roll Anti-Clockwise, 6=Roll Clockwise, 7=Translation Backward (Zoom In), 8=Translation Forward (Zoom Out).



(a) Synthetic (b) Natural 1 (c) Natural 2

Fig. 10: **Results: Learned reward tables.** Final R-tables obtained after training on input images in Fig. 6 in the same order. White=1.0, black=0.0.

where $R_t(.,.)$ is the reward probability function at time t, and α is the learning rate ($\alpha = 0.001$). Finally, the reward table is normalized (similar to Eq. 4),

$$R_{t+1}(s_t, a) = \frac{R_{t+1}(s_t, a)}{\sum_{a_i \in A} R_{t+1}(s_t, a_i)}. \quad (7)$$

V. Experiments and Results

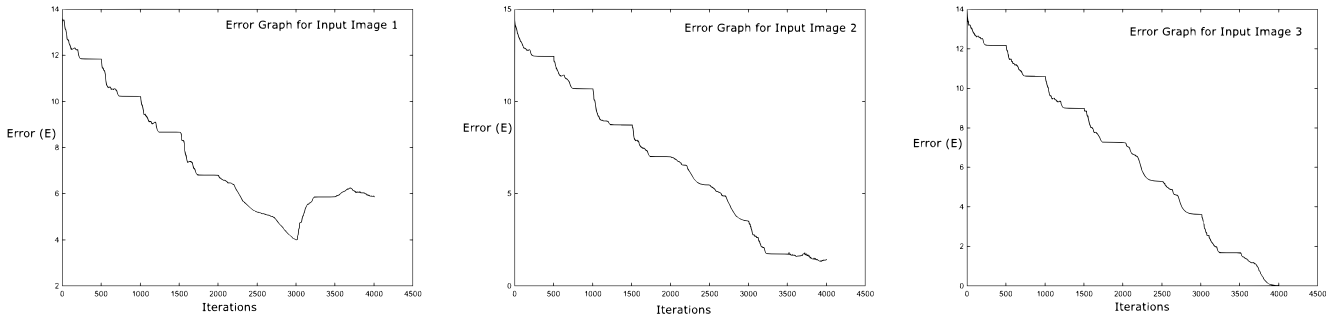
In order to test the effectiveness of the learning algorithm as mentioned in the previous section, we conducted experiments to study the nature of state-action mapping obtained using the given models. The experiments were conducted with different sets of parameter values in the model over various backgrounds - synthetic and natural scenes. We experimented with Lucas-Kanade method and Horn-Schunck methods for optical flow computation, where both gave similar results with slight variation based on input image's characteristics. Different values for learning rate (α), modeled neuron's parameters ($\sigma_\theta, \sigma_\phi$) were experimented. We also tried various techniques for generating reward (ρ), e.g., dot product vs. inverse of Euclidean distance.

Here we show results for 3 input background images as shown in Fig. 6. The first one is a high contrast-low density texture synthetic scene, the second a low contrast-high density texture natural scene, and the third a high contrast-high density texture natural scene. The agents were trained using 8 sensory primitives and 8 motor primitives (Fig. 9). The training was done for 500 iterations per sensory primitive, where the input background was moved with respect to the agent in a particular fashion to generate that specific sensory primitive. The learning algorithm was executed to learn a state-action mapping for this state. This procedure was repeated for all 8 sensory primitives. Therefore, overall the training involved a minimum of 4,000 iterations to get good results.

The reward table $R(s,a)$ was initially randomly initialized before training. The ideal R-table is an identity matrix of size 8. The learned R-table becomes quite similar to the ideal R-table when the learning algorithm is executed for a large number of iterations. The algorithm performs very well for images that have high contrast and dense texture. This is expected as the local motion detectors give smooth optical flow pattern when the image has lot of variation in intensity across the scene, as can be seen in Fig. 10a-c, which shows the learned R-tables for input images in Fig. 6a-c, respectively. It is quite clear from this result that the algorithm performs best for the background in Fig. 6c as it has dense texture and high contrast. Synthetic background (Fig. 6a) performs sub-optimally due to sparse texture (perhaps due to the strong aliasing effect), while the Fig. 6b input gives better results (but not perfect) due to low contrast.

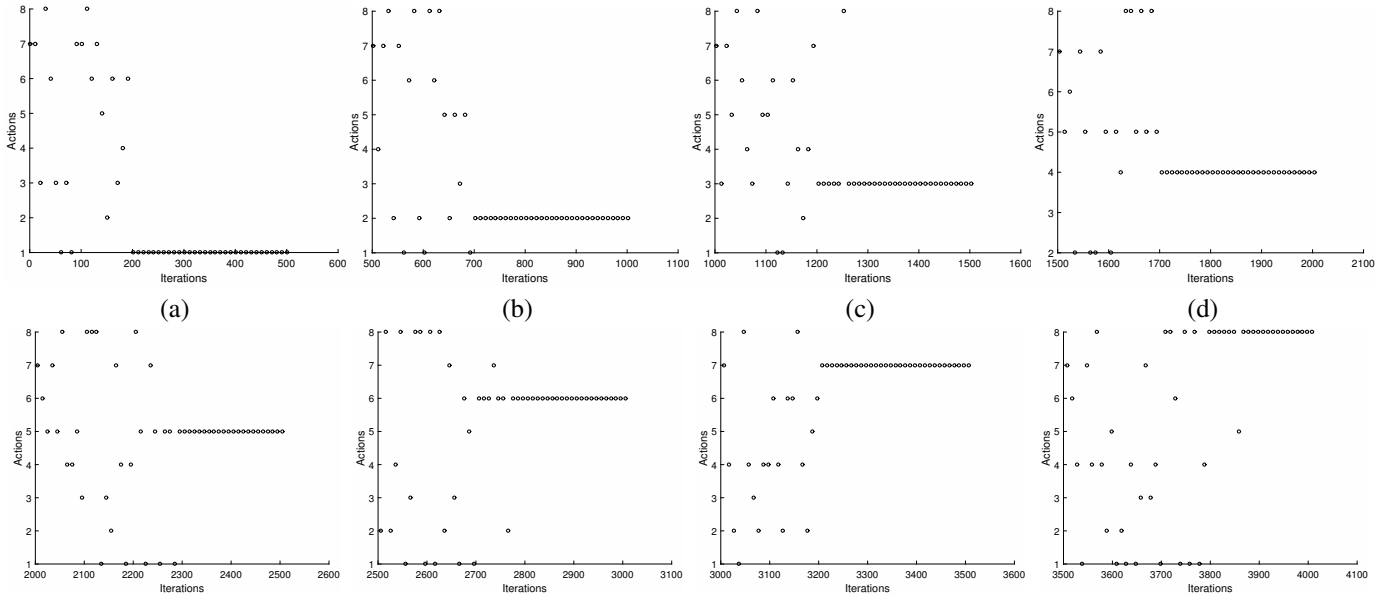
To quantitatively measure the performance of the algorithm, we compared the Ideal R-table ($R_I(s,a)$) and the learned R-table ($R_L(s,a)$) over the training iterations. This was computed as $E = \sum_{s,a} |R_L(s,a) - R_I(s,a)|$, where $|..|$ is the absolute value. The error (E) represents the discrepancy between the ideal R-table and the learned R-table. This error is calculated after every iteration in the learning algorithm and is shown to decrease to near zero over the iterations, for all three training images (Fig. 11). The error at the end of the training process approaches zero, meaning that the learning algorithm works as expected. Fig. 12 shows learned action choices over training iteration for each input state. We can see that within 200 to 300 training iterations after the input state changes, the correct action (see the ideal actions in Fig. 9) is learned and maintained.

The results show that the proposed agent model and learning algorithm is able to give promising results for the motion-based grounding problem in synthetic and natural scenarios. The agent learns a unique relationship between each of its sensory states and actions. Fig. 10 shows that actions that are able to maintain and reinforce the sensory states are mapped to those very states in the learning process. For example, consider the case where sensory state is roll optical flow in clockwise sense (RRCL). Now, the action roll in counter clockwise sense (counter clockwise arrow) will generate the same optical flow as observed by the agent. Thus, this action reinforces the current sensory state while other actions would change the state, making the the learning algorithm map roll counter clockwise action to generate roll clockwise optical flow. A similar relationship can be observed for all other corresponding state-action pairs.



(a) Synthetic (b) Natural 1 (c) Natural 2

Fig. 11: **Results: Learning error over training iteration.** Error in the reward table shown as a function of training iteration, based on the three input images (Fig. 6), in the same order. The x axis represents training iteration and the y axis error.



(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 12: **Results: Actions chosen over training iterations.** The progress graphs shows the actions (y axis: see Fig. 9 for action-number mapping) chosen over training iterations (x axis) for Natural 2 image in Fig. 6. Training with initial state: (a) RYRL (action a=1), (b) RYLR (a=2), (c) RPUD (a=3), (d) RPDU (a=4), (e) RRCL (a=5), (f) RRAC (a=6), (g) TLRI (a=7), (h) TLRO (a=8).

VI. Conclusion and Future Work

In this paper, we showed that an optical flow model based on the fly's visual system can be used by an agent to learn a sensory-motor mapping that relates (1) the agent's internal sensory states that encode optical flow information to (2) simple motions like rotation, translation, etc. that are congruent with the encoded sensory states. We expect our computational study to help better understand how the real fly visual system decodes its own spikes, thus achieving autonomous grounding. In future work, we will investigate compositional optical flow, for example, optical flow induced by combined roll and thrust (forward movement), etc. Furthermore, our results offer a strong prediction that can be tested in the experimental lab: electrical stimulation of the optical flow detectors in the fly brain would elicit motor behavior that reinforce the detected optical flow. For example, stimulating RRCL detector will result in counter clockwise rotation.

Acknowledgments

This paper is largely based on the first author's master's thesis [28]. The first author is now at FactSet, Austin, Texas.

References

[1] M. Egelhaaf, N. Boeddeker, R. Kern, R. Kurtz, and J. P. Lindemann, "Spatial vision in insects is facilitated by shaping the dynamics of visual input through behavioral action," *Frontiers in neural circuits*, vol. 6, 2012.

[2] S.-y. Takemura, A. Bharioke, Z. Lu, A. Nern, S. Vitaladevuni, P. K. Rivlin, W. T. Katz, D. J. Olbris, S. M. Plaza, P. Winston, T. Zhao, J. A. Horne, R. D. Fetter, S. Takemura, K. Blazek, L. Chang, O. Ogundeyi, M. A. Saunders, V. Shapiro, C. Sigmund, G. M. Rubin, L. K. Scheffer, I. A. Meinertzhagen, and D. B. Chklovskii, "A visual motion detection circuit suggested by drosophila connectomics," *Nature*, vol. 500, pp. 175–181, 2013.

[3] H. G. Krapp, B. Hengstenberg, and R. Hengstenberg, "Dendritic structure and receptive-field organization of optic flow processing interneurons in the fly," *Journal of Neurophysiology*, vol. 79, no. 4, pp. 1902–1917, 1998.

[4] H. Cuntz, J. Haag, and A. Borst, "Modelling the cellular mechanisms of fly optic flow processing," in *The Computing Dendrite*. Springer, 2014, pp. 259–275.

[5] G. K. Taylor and H. G. Krapp, "Sensory systems and flight stability: what do insects measure and why?" *Advances in insect physiology*, vol. 34, pp. 231–316, 2007.

[6] M. Egelhaaf, R. Kern, H. G. Krapp, J. Kretzberg, R. Kurtz, and A.-K. Warzecha, "Neural encoding of behaviourally relevant visual-motion information in the fly," *Trends in neurosciences*, vol. 25, no. 2, pp. 96–102, 2002.

[7] J. Lindemann, R. Kern, J. van Hateren, H. Ritter, and M. Egelhaaf, "On the computations analyzing natural optic flow: quantitative model analysis of the blowfly motion vision pathway," *The Journal of neuroscience*, vol. 25, no. 27, pp. 6435–6448, 2005.

[8] H. G. Meyer, J. P. Lindemann, and M. Egelhaaf, "Pattern-dependent response modulations in motion-sensitive visual interneurons–a model study," *PLoS One*, vol. 6, no. 7, p. e21488, 2011.

[9] H. Eichner, M. Joesch, B. Schnell, D. F. Reiff, and A. Borst, "Internal structure of the fly elementary motion detector," *Neuron*, vol. 70, no. 6, pp. 1155–1164, 2011.

[10] W. Reichardt and W. Rosenblith, "Autocorrelation, a principle for evaluation of sensory information by the central nervous system," in *Symposium on Principles of Sensory Communication 1959*. MIT Press, 1961, pp. 303–317.

[11] W. Reichardt, "Evaluation of optical motion information by movement detectors," *Journal of Comparative Physiology A*, vol. 161, no. 4, pp. 533–547, 1987.

[12] S. Harnad, "The symbol grounding problem," *Physica D*, vol. 42, pp. 335–346, 1990. [Online]. Available: http://www.ecs.soton.ac.uk/~harnad/Papers/Harnad/harnad90.sgproblem.html

[13] J. Modayil and B. Kuipers, "Autonomous development of a grounded object ontology by a learning robot," in *Proceedings of the 22nd National Conference on Artificial Intelligence*, 2007, pp. 1095–1101.

[14] D. M. Pierce and B. J. Kuipers, "Map learning with uninterpreted sensors and effectors," *Artificial Intelligence*, vol. 92, pp. 162–227, 1997.

[15] Y. Choe, "Action-based autonomous grounding," in *Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence*, 2011.

[16] Y. Choe and N. H. Smith, "Motion-based autonomous grounding: Inferring external world properties from encoded internal sensory states alone," in *PROCEEDINGS OF THE NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE*, vol. 21, no. 1. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2006, p. 936.

[17] 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, no. 02, pp. 211–243, 2007.

[18] Y. Choe, "Action-based autonomous grounding," in *AAAI-11 Workshop on Lifelong Learning from Sensorimotor Experience*, J. Modayil, D. Precup, and S. Singh, Eds. Palo Alto, CA: AAAI Press, 2011, pp. 56–57, aAAI Workshop Technical Report WS-11-15. [Online]. Available: http://www.aaai.org/ocs/index.php/WS/AAAIW11/paper/view/3798

[19] Y. Choe and N. H. Smith, "Motion-based autonomous grounding: Inferring external world properties from internal sensory states alone," in *Proceedings of the 21st National Conference on Artificial Intelligence(AAAI 2006)*, Y. Gil and R. Mooney, Eds., 2006, pp. 936–941. [Online]. Available: http://faculty.cs.tamu.edu/choe/ftp/publications/choe.aaai06.pdf

[20] A. Borst and M. Egelhaaf, "Principles of visual motion detection," *Trends in neurosciences*, vol. 12, no. 8, pp. 297–306, 1989.

[21] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision." in *IJCAI*, vol. 81, 1981, pp. 674–679.

[22] B. K. Horn and B. G. Schunck, "Determining optical flow," in *1981 Technical Symposium East*. International Society for Optics and Photonics, 1981, pp. 319–331.

[23] H. G. Krapp, "Neuronal matched filters for optic flow processing in flying insects," *International review of neurobiology*, vol. 44, pp. 93–120, 2000.

[24] 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. [Online]. Available: http://www.bbsonline.org/Preprints/ORegan/

[25] D. Philipona, J. K. O'Regan, J.-P. Nadal, and O. J.-M. D. Coenen, "Perception of the structure of the physical world using unknown multimodal sensors and effectors," in *Advances in Neural Information Processing Systems 16*, S. Thrun, L. Saul, and B. Schölkopf, Eds. Cambridge, MA: MIT Press, 2004, pp. 945–952.

[26] 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. [Online]. Available: http://nivea.psycho.univ-paris5.fr/Philipona/space.pdf

[27] T. Dijkstra, P. Snoeren, and C. Gielen, "Extraction of three-dimensional shape from optic flow: A geometric approach," *JOSA A*, vol. 11, no. 8, pp. 2184–2196, 1994.

[28] A. Parulkar, "Autonomous grounding of the optical flow detectors in a simulated visuomotor system of the fly using behaviorally meaningful actions," Master's thesis, Department of Computer Science, Texas A&M University, 2015.