



Identification of Exploration and Exploitation Balance in the Silkmoth Olfactory Search Behavior by Information-Theoretic Modeling

Cesar A. Hernandez-Reyes^{1*}, Shumpei Fukushima^{1,2}, Shunsuke Shigaki³,
Daisuke Kurabayashi¹, Takeshi Sakurai⁴, Ryohei Kanzaki⁵ and Hideki Sezutsu⁶

¹ Department of Systems and Control Engineering, Tokyo Institute of Technology, Tokyo, Japan, ² MHPS Ltd., Takasago, Japan, ³ Department of Systems Innovation, Osaka University, Osaka, Japan, ⁴ Department of Agricultural Innovation for Sustainability, Tokyo University of Agriculture, Atsugi, Japan, ⁵ Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan, ⁶ Transgenic Silkworm Research Unit, National Agriculture and Food Research Organization, Tsukuba, Japan

OPEN ACCESS

Edited by:

Dai Owaki,
Tohoku University, Japan

Reviewed by:

Hitoshi Aonuma,
Hokkaido University, Japan
Dominique Martinez,
UMR7503 Laboratoire Lorrain de
Recherche en Informatique et ses
Applications (LORIA), France

*Correspondence:

Cesar A. Hernandez-Reyes
hernandez.cesar@irs.sc.e.titech.ac.jp

Received: 14 November 2020

Accepted: 08 January 2021

Published: 01 February 2021

Citation:

Hernandez-Reyes CA, Fukushima S,
Shigaki S, Kurabayashi D, Sakurai T,
Kanzaki R and Sezutsu H (2021)
Identification of Exploration and
Exploitation Balance in the Silkmoth
Olfactory Search Behavior by
Information-Theoretic Modeling.
Front. Comput. Neurosci. 15:629380.
doi: 10.3389/fncom.2021.629380

Insects search for and find odor sources as their basic behaviors, such as when looking for food or a mate. This has motivated research to describe how they achieve such behavior under turbulent odor plumes with a small number of neurons. Among different insects, the silk moth has been studied owing to its clear motor response to olfactory input. In past studies, the “programmed behavior” of the silk moth has been modeled as the average duration of a sequence of maneuvers based on the duration of periods without odor hits. However, this model does not fully represent the fine variations in their behavior. In this study, we used silk moth olfactory search trajectories from an experimental virtual reality device. We achieved an accurate input by using optogenetic silk moths that react to blue light. We then modeled such trajectories as a probabilistic learning agent with a belief of possible source locations. We found that maneuvers mismatching the programmed behavior are related to larger entropy decrease, that is, they are more likely to increase the certainty of the belief. This implies that silkmoths include some stochasticity in their search policy to balance the exploration and exploitation of olfactory information by matching or mismatching the programmed behavior model. We believe that this information-theoretic representation of insect behavior is important for the future implementation of olfactory searches in artificial agents such as robots.

Keywords: *Bombyx mori*, infotaxis, olfaction, ethology, adaptive-behavior, exploration-exploitation

1. INTRODUCTION

Odor source localization is a search problem that requires fast decision-making based on sporadic and stochastic detection of chemical particles. Despite the challenge of turbulent and dilute plumes that often have a complex spatio-temporal structure (Mafra-Neto and Cardé, 1994; Celani et al., 2014), insects such as the fruit fly (van Breugel and Dickinson, 2014) and various species of moths (Vickers, 2005) rely on olfactory searches to conduct essential behaviors such as searching for food or potential mates. The high performance that insects show on such a complex search problem

despite their simple brain motivates researchers to further analyze and understand the decision processes that these insects execute when conducting olfactory searches (Baker et al., 2018).

With this motivation, our research group has analyzed the olfactory behavior of the male silk moth *Bombyx mori* (Lepidoptera: Bombycidae). Despite having wings, this insect is unable to fly, and has a body that is on average 30 mm long and 10 mm wide. It has two antennae of approximately 6 mm in length on its head. This insect has been widely employed to analyze olfactory behavior because it exhibits only one action: It walks only when it detects a pheromone (*Bombykol*) released by its female counterpart (Obara, 1979). Such behavior consists of a series of maneuvers called a “surge,” “zigzag,” and “loop.” This sequence of maneuvers has been approximated to a mean-response model denoted as “programmed behavior” (Kanzaki et al., 1992).

Based on the mean durations of the surge, zigzag, and loop maneuvers, the programmed behavior has been algorithmically defined as follows: first, immediately after a pheromone stimulus, the moth advances in a straightforward manner through a *surge* motion. Then, if there is an absence of pheromone detections, the moth moves on a *zig-zag* pattern, trying to detect pheromones again. Finally, if the pheromone remains undetected, the moth transitions into a *loop* motion until the next detection. A diagram of the programmed behavior is shown in **Figure 1**. Because the silk moth is motionless by default and only elicits its programmed behavior after the first pheromone hit, this search strategy has been labeled as “reactive” by Voges et al. (2014). Despite the simplicity of this sequential pattern, the male silk moth can effectively locate females with remarkable efficiency.

However, this model does not reflect how the motions of the moth vary in response to fine spatio-temporal fluctuations of the odor plume and individual differences among specimens. In previous studies, such variability was investigated by identifying maneuver transitions with machine learning (Shigaki et al., 2018b) and fuzzy logic (Shigaki et al., 2019b). Although these studies succeeded in identifying deviations from the programmed behavior, they relied on data from electrophysiological signals obtained from implanting electrodes in the wing muscles or brain of the silk moth; however, electrode implantation is technically challenging and risks degrading the tissues of the moth. Therefore, an analysis method that allows modeling adaptive olfactory behavior from non-intrusive experimental measurements is necessary.

To identify adaptive olfactory behavior, recent studies have used the information-theoretic framework of infotaxis, which was first proposed by Vergassola et al. (2007). A recent study by Pang et al. (2018) investigated the features of odor encounters that modulate the intensity of upwind turns in the fruit fly *Drosophila melanogaster* and the mosquito *Aedes aegypti*. The authors found through simulations that, compared to a *centerline inferring* odor source search algorithm, infotaxis produced trajectories that were more similar to those of the actual animals, in the sense that they exhibit weaker upwind turns later in a sequence of odor encounters. Similarly, Calhoun et al. (2014) recently demonstrated the possibility of using infotaxis to model the multi-stage foraging behavior of the nematode

Caenorhabditis elegans. In their paper, the authors showed that infotaxis-like search strategies, which minimize the entropy of the probability distribution of odor source locations, reflects both the “local” and “global” stages of the *C. elegans* foraging behavior.

In this paper, we investigate the potential causes of variability in the behavioral maneuvers of the silk moth *B. mori* by using a non-invasive experimental method and an infotaxis-based model similar to those described in recent studies. We measured the silk moth trajectories and input stimuli data with a tether, a two-dimensional treadmill, and a virtual odor plume. To ensure accurate and reproducible stimuli, we used optogenetic silk moths that react to the impulses of blue light in the same way as with pheromone particles. We modeled the trajectories and stimuli measurements as infotaxis agents and found that; maneuvers that mismatch the *programmed behavior* model correspond to higher expected information rewards regarding the location of the source. In summary, we believe that this paper demonstrates the possibility of using non-invasive experimental measurements and infotaxis-based modeling to identify variability in the olfactory behaviors of the male silk moths.

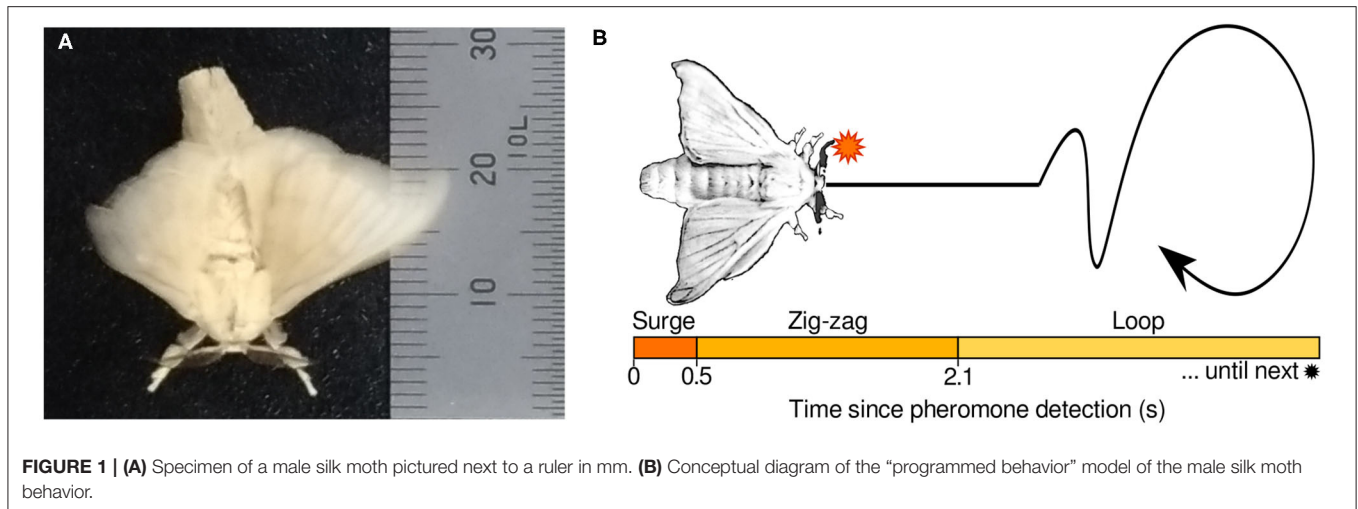
This paper is structured as follows: section 2 states the research questions of this paper. Section 3 describes the usage of optogenetic silk moths, the experimental virtual reality system to measure their behavior, and how to model it as infotaxis agents. Section 4 shows the results of the behavior measurement experiments and calculations of the information entropy of infotaxis-modeled silk moths. Section 5 discusses the contributions of this study and possible future areas of research.

2. PROBLEM STATEMENT

In this paper, we look for possible causes of adaptive mechanisms in the olfactory behavior of the silk moth, which are not represented in the programmed behavior model. Specifically, we investigate the following two hypotheses:

- Are deviations from the programmed behavior motivated by higher information gains?
- Can a probabilistic framework such as infotaxis explain how the male silk moth balances exploration and exploitation of olfactory information?

To test the first hypothesis, we need to measure the behavior of the silk moth in an olfactory environment that can be accurately reproduced in each experimental run. Therefore, in this paper we utilize a “virtual reality” behavioral measurement system in which we can subject moths to virtual odor plumes and measure their motor response to odor stimuli. However, such a system faces the challenge of an accurate stimulation of the moth antennae. In other words, stimulating the antennae with gaseous pheromone particles results in uncertain stimulation because such particles diffuse in the air; hence, they do not produce stimuli with the same intensity or duration each time. To overcome this, we employed genetically modified silkmooths that elicit their normal olfactory behavior response when subjected



to a blue light stimulus at their antennae; thus, we can present reproducible olfactory inputs.

To test the second hypothesis we modeled the trajectories of silkmoths as an agent that minimizes the information entropy of its probabilistic belief of the location of an odor source. Such a maximally informative agent is based in the infotaxis algorithm (Vergassola et al., 2007). We related the decrease in entropy of the infotaxis-modeled moth to the time steps in which the moth behavior matched or mismatched the programmed behavior model. Finally we determined whether infotaxis can explain the exploration-exploitation strategy of the silkmoth behavior by evaluating the distribution of entropy reductions by either matching or mismatching behaviors.

3. MATERIALS AND METHODS

Here, we describe our methodology for conducting olfactory search experiments with optogenetic male silkmoths and a non-invasive behavior measurement system. We also describe the method we used to represent the silkmoth trajectories as those of an infotaxis agent. The silkmoth experiments in this study were examined and approved by the Tokyo Institute of Technology Gene Recombination Experiments Safety Management Committee.

3.1. Virtual Reality System for Measurement of Moth Behavior

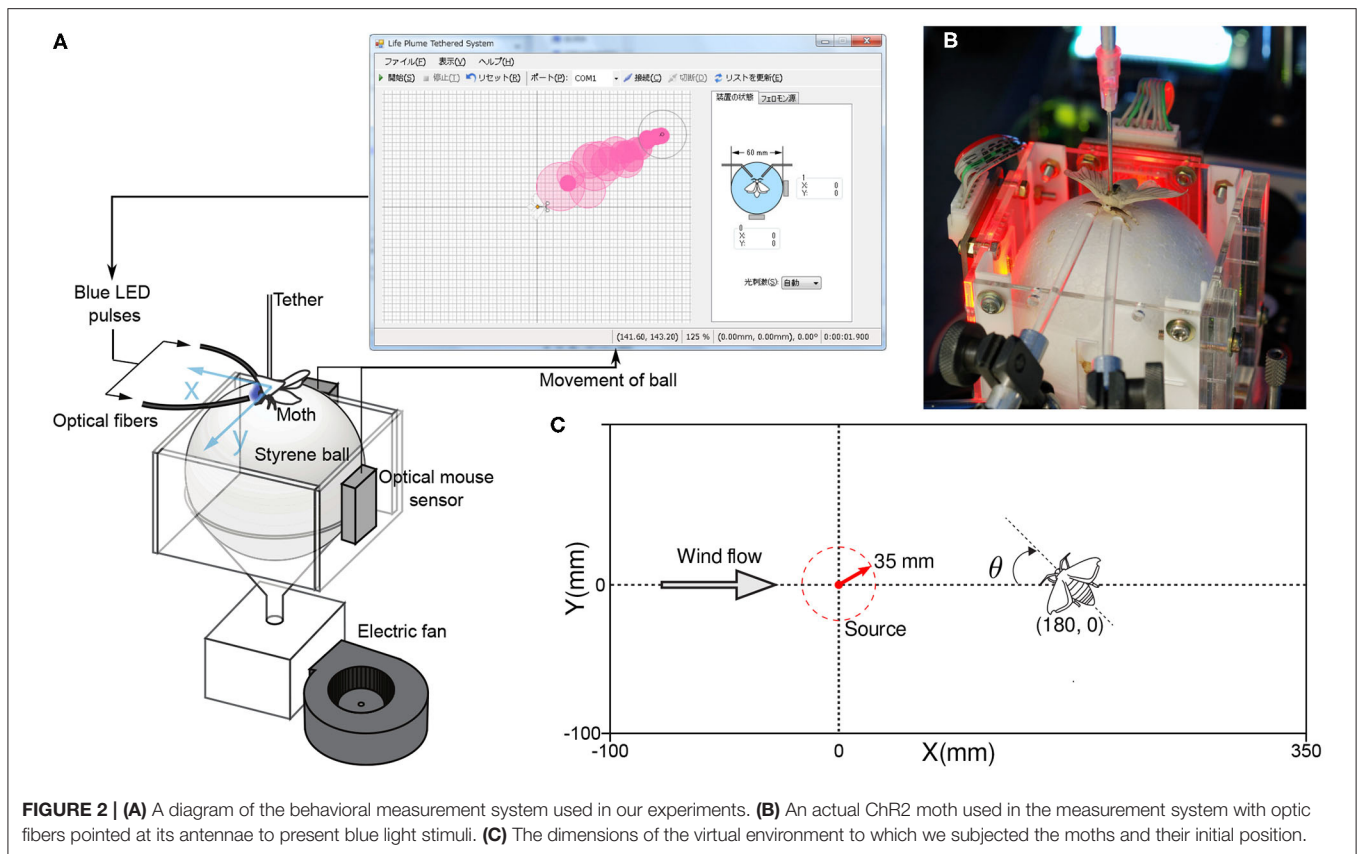
We conducted non-intrusive behavioral measurements on tethered male silkmoths. Although similar systems to measure the olfactory behavior of insects have been used in the past (Shigaki et al., 2018a, 2019b), in this study we ensure that odor stimuli are accurately presented by using optogenetic silkmoths. Using genetically modified specimens that react to blue light stimuli in the same way as normal specimens react to the pheromone bombykol, allowed us to present stimuli accurately and with reproducibility. This is because gaseous pheromones diffuse in the air; therefore, not all stimuli present the same amount of pheromone molecules to the antennae

of the moth. Furthermore, in this case, the response of the antennae is measured using an electroantennogram (EAG), which is technically challenging and subjected to electrical noise; in addition, damage to the antennae may occur. Our non-invasive behavior measurement system for the silkmoth is shown in **Figure 2**, and fulfills the following purposes:

- Measuring the pose (x, y, θ) of the moths.
- Accurately presenting light stimuli to the antennae of the moth.
- Subjecting moths to a virtual odor plume to which we can alter the emission rate, wind speed, and other parameters.

To measure the pose of the silkmoth, we fixed its back to a thin aluminum rod (\varnothing 2 mm; length 150 mm) with glue (G17 Bond, Konishi K.K., Osaka, Japan) and placed it on a polystyrene sphere (\varnothing 60 mm), which served as a two-dimensional treadmill. When the moth walked, the sphere moved in response because it was being levitated by the flow of wind from a small fan (FW1251-1051C2ALARX, ARX, Wanchai, Hong Kong). The movements of the sphere were detected using two optical sensors, such as those found in a computer mouse (ADNS-5030, Avago Technologies, California, USA), at a sampling rate of 20 Hz. They were then translated into translational and rotational movements of the moth, that is, the pose.

We developed a virtual representation of an odor plume by modeling the dispersion of white smoke in a wind tunnel. First, we recorded videos of the dispersion of smoke. We also calculated the statistics of the position and intensity of the pixels in the smoke video. Based on these statistics, we programmed a random process that generates virtual circular puffs that match the intensity and transit the positions of the real smoke puffs in the video. An example of a virtual plume is shown in **Figure 2A**. In addition to the virtual representation of the odor plume, we also programmed a virtual representation of a silkmoth. As in the real world, the virtual moth reacts to the virtual plume and travels toward its source. By using a virtual odor plume environment, we can tune parameters such as wind speed, emission rate, and particle lifetime. Tuning such parameters is



particularly useful in infotaxis-based behavior modeling because it allows for faster testing of various plume structures and higher reproducibility; compared with real plume experiments. In summary, the following process describes the operation of our experimental device:

1. The moth in the virtual world encounters a puff of pheromone.
2. Blue light is shown to the real moth depending on which antenna of the virtual moth reacted.
3. The real moth moves after receiving the stimulus.
4. The movement of the real moth is sent to the virtual world.
5. The virtual moth reflects the movement of the real moth.
6. The loop is repeated until either the moth reaches the virtual source or until a predetermined time limit is passed.

3.2. Use of Optogenetic Moths for Accurate Antennae Stimulation

The presentation of accurate stimuli is important for the applied infotaxis-based analysis because updating the probability distribution of the source position; as well as the calculation of the expected entropy decrease, are directly affected by whether the agent experiences a hit or not at a given time step. In addition, reproducible odor stimuli are an overall useful property for an olfactory behavior measurement system because their duration and frequency can be finely tuned. Both properties have been reported to directly influence the olfactory behavior of moths (Celani et al., 2014) and other animals (Ache et al., 2016). To

present olfactory stimuli to the moth, previous studies have presented pheromones from glass tubes placed directly in front of the antennae of the moth. However, the amount of pheromone particles that effectively reach the antenna varies owing to their gaseous nature.

To ensure that each stimulus has the same intensity and is accurately sensed by the antennae, we utilized genetically modified moths. These BmOR1-GAL4/UAS-ChR2 silk moths (ChR2 hereinafter); express channelrhodopsin-2 in their olfactory receptor neurons. As a result, they execute their olfactory search behavior when their antennae encounter blue light, rather than pheromone particles. This property has been used in previous studies to ensure that all stimuli are reproducible with the same intensity and duration (Shigaki et al., 2018b, 2019a). To activate channelrhodopsin-2, i.e., blue light sensitivity in these moths, we injected all-trans retinal (ATR) into their abdomen on the day before the experiments; because insects do not intrinsically possess ATR. All behavior measurement experiments were conducted from 9:00 to 17:00 to reduce circadian effects (Tomioka et al., 1993). It is reported that brain serotonin level increases in the daytime and that serotonin enhances pheromones sensitivities in the silk moth (Gatellier et al., 2004).

We generated stimuli for the ChR2 silk moths with LEDs (LBW5AP-JYKY-35-Z; Osram Opto Semiconductors), which produced blue light with a 470 nm wavelength and a light intensity of more than 1.6 mW/mm². Such values of wavelength

and light intensity have been reported to reliably produce olfactory search responses in ChR2 moths (Tabuchi et al., 2013). On each LED, we attached optical fibers of 3 mm in diameter to ensure that blue light was directed only to each antenna, as seen in **Figure 2B**. In addition, moths are unable to make yaw turns because their back is glued to an aluminum rod. The only rotation they are able to make is on their neck (see **Supplementary Video**). However, this neck rotation is very small and it does not decrease the sensibility or the amount of stimulation to the antennae.

3.3. Modeling the Silk Moth as an Infotaxis Agent

Infotaxis was first proposed by Vergassola et al. (2007) as an odor source search algorithm for turbulent environments. In this algorithm, a point-mass agent is located at a position \mathbf{r} and searches for an odor source by iteratively reducing its uncertainty about the distribution of possible source locations \mathbf{r}_{src} . The agent has knowledge of its trajectory, \mathcal{T}_t , which contains its sequence of positions as well as the odor “hits” it has experienced throughout the search. The agent also maintains a probability map $P(\mathbf{r}_{src}|\mathcal{T}_t)$ or “belief” (Thrun et al., 2005) about the location of the source. This belief spans all possible locations of the source \mathbf{r}_{src} that in both the original infotaxis study and the present paper, consist of a two-dimensional lattice of discrete locations. The certainty of the belief $P(\mathbf{r}_{src}|\mathcal{T}_t)$ is represented by Shannon’s entropy as in Equation (1):

$$S_t = S[P(\mathbf{r}_{src}|\mathcal{T}_t)] = - \sum_{\mathbf{r}_{src}} P(\mathbf{r}_{src}|\mathcal{T}_t) \ln(P(\mathbf{r}_{src}|\mathcal{T}_t)) \quad (1)$$

The goal of infotaxis is to minimize the entropy of the belief $P(\mathbf{r}_{src}|\mathcal{T}_t)$; therefore, at every time step, the agent calculates the expected change of entropy by moving from its current position \mathbf{r}_t to a future position \mathbf{r}' as defined in Equation (2).

$$E[\Delta S(\mathbf{r}_t \mapsto \mathbf{r}')] = p^* \Delta S^* + (1 - p^*) \Delta S \quad (2)$$

Where p^* is the probability of finding the source at \mathbf{r}' , and ΔS^* and ΔS are the change in entropy if the source is found or not found at \mathbf{r}' , respectively. The agent then executes the move $\mathbf{r}_t \mapsto \mathbf{r}'$ with the largest negative value of $E[\Delta S]$, or; in other words, the move that causes the greatest reduction of uncertainty in the agent’s probability map of the possible source locations. **Figures 3A,B** show conceptual representations of the agent’s belief as well as the effect of odor detections on such belief. Detailed derivations of the infotaxis formulae are presented in **Appendix A** of this paper.

We modeled the body of the silk moth as a point agent with a radius of 10 mm (half of its average body length). We reduced the three degrees of freedom of the moth to (x, y) coordinates because an infotaxis agent moves in a two-dimensional grid ignoring the orientation. Furthermore, we considered as odor hits only those that occurred when the moth was facing upwind, that is, when $\cos(\pi - \theta + \theta_{src}) > 0$ (see **Figure 3C**), where θ and θ_{src} are the angle of the moth and the plume’s centerline, respectively. We considered this capture region because real moths limit the

odor hits to those coming from the front by flapping their wings (Loudon and Koehl, 2000).

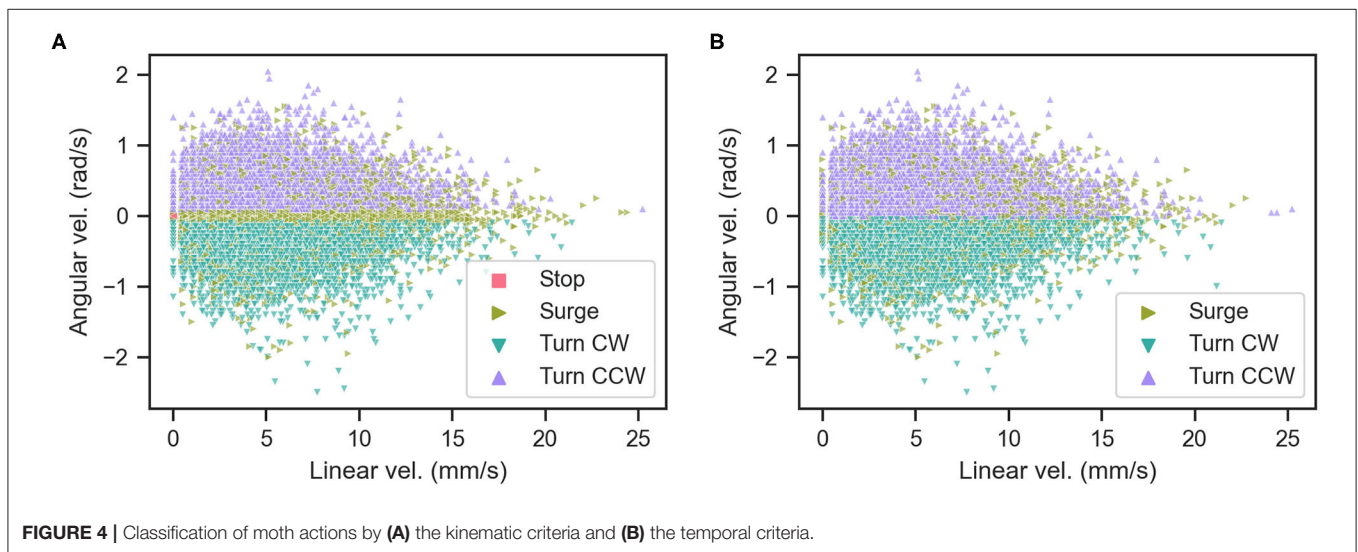
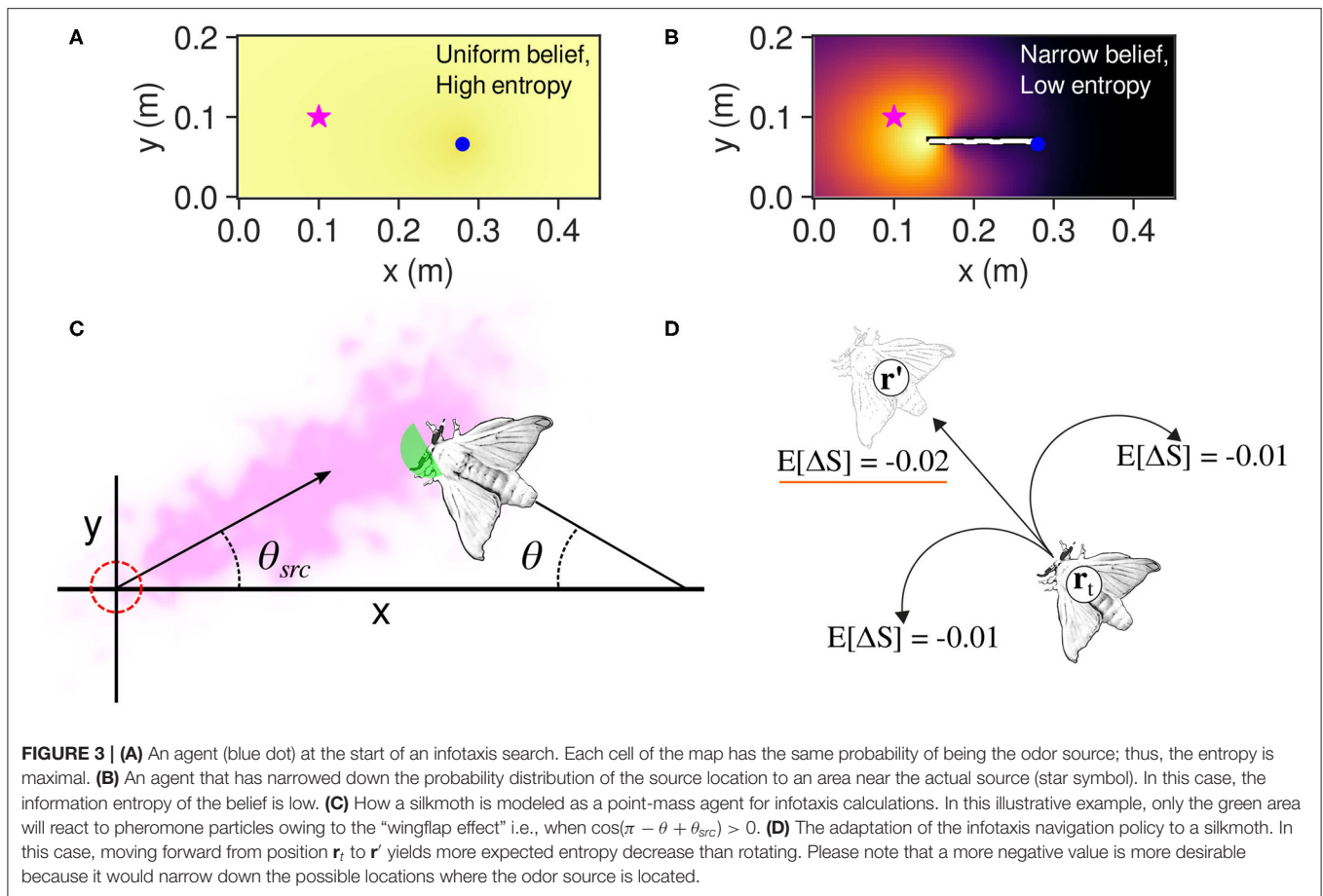
3.4. Classification of Variability in the Moth Behavior

We determined whether the behavior of the silk moth matches the definition of the programmed behavior (Kanzaki et al., 1992) by comparing it to the definition of Minegishi et al. (2012). Accordingly, we classified the maneuvers of the silk moth by simply considering the time elapsed since the last odor hit, which we call “blank duration” τ_b as in Celani et al. (2014). We also classified maneuvers according to both τ_b and the moth’s linear and angular velocities (v and ω , respectively) based on Minegishi et al. (2012). We denote the first and second classification as “temporal” and “kinematic,” respectively. **Table 1** shows a comparison of both schemes used to classify maneuvers and **Figure 4** shows the result of using each scheme. The blank duration threshold of 500 ms in the “temporal” classification of **Table 1** was selected because this is the average duration of surge motions after an odor hit as reported in Kanzaki et al. (1992). Throughout all olfactory search experiments, we classified the moth maneuvers by both schemes and labeled the state of the moth at each time step as “matching” if it matches the criteria of both schemes and “mismatching” if it only matches the “kinematic” criteria.

To determine whether “mismatching” behaviors are motivated by higher information gains, we analyze the value of the entropy change ΔS and the expected entropy change $E[\Delta S]$ regarding the rate of odor hits and the cumulative odor hits experienced by moths over a search. We are particularly interested in these variables because recent studies identified that they influence the decision-process of olfactory behaviors (Celani et al., 2014; Pang et al., 2018). We also evaluate whether the distribution of ΔS is different for “matching” and “mismatching” behaviors with a two-sample Kolmogorov-Smirnov test and by comparing their histograms. In addition, we calculate the cumulative density function (CDF) of ΔS and $E[\Delta S]$ to specifically determine whether “mismatching” behaviors have a higher probability of obtaining larger negative values of those variables, that is, greater information gains. Finally, we calculate the root mean squared error (RMSE) between the values of ΔS and $E[\Delta S]$ to determine what type of behavior is more similar to infotaxis, regarding the rate of odor hits and the cumulative sum of hits, which are our variables of interest. The following section presents the results of the calculations of ΔS and $E[\Delta S]$ regarding hit rate and cumulative hits, the histograms and CDFs, and the RMSE of “matching” and “mismatching” behaviors.

4. RESULTS

Here, we present the results of the VR odor source search experiments using optogenetic silkmoths. First, we present the trajectories of the moths as well as their information entropy. We then show the statistics of the matching versus mismatching states, followed by the relationship between those two states and the expected decrease in information entropy for each.



4.1. VR Olfactory Search Experiments

We subjected ChR2 silkmoths to olfactory search experiments. We conducted 20 trials in which the moth searched for a pheromone source in a 350 mm long by 200 mm wide virtual environment where the wind was blowing in the positive

x -direction at a mean speed of 0.1 m/s. The initial position of the moth in the virtual environment was $(x, y, \theta) = (180, 0, -\pi/6)$, where θ is in radians. Moths searched for a source located at $(x, y) = (0, 0)$ by entering a radius of 35 mm around it under a time limit of 180 s. The mean \pm std. dev. of the time required to reach

TABLE 1 | Definitions for the maneuvers of silk moths when classified by either a temporal or a kinematic state.

	Temporal	Kinematic
Surge	$\tau_b \leq 500$ ms	$\tau_b \leq 500$ ms and $v > 0$ or $\tau_b > 200$ ms and $ \omega < 5$ deg/s
Rotate	$\tau_b > 500$ ms	$\tau_b > 500$ ms and $ \omega > 0$
Stop	Otherwise	Otherwise

the source was 73.92 ± 46.5 s. **Figure 5A** shows the information entropy for the experiments where moths found the pheromone source. The solid line represents the average value, the shaded range represents the standard deviation, and the gray lines show the value for each trial. **Figure 5B** shows the moth trajectories of these successful trials. The color gradient represents the value of the information entropy. **Table 2** shows the statistics of the matching and mismatching moth states. Surge (temporal) and Rotate (kinematic) represent the proportion of time taken when the silk moths exhibited a mismatching state over the entire duration of the search experiments. In total we conducted 20 experiments with 10 specimens. Out of these, 12 trials from six specimens successfully found the odor source under the time limit; thus achieving a success rate of 60.0%. We considered only the data from the successful trials for the classification of matching and mismatching behaviors.

4.2. Relationship Between Behavior Variability and Information Gains

We investigated whether there is a relationship between mismatching maneuvers and a higher expected decrease in entropy $E[\Delta S(\mathbf{r}_t \mapsto \mathbf{r}')]$. **Figure 6** shows the actual rewards ΔS and expected rewards $E[\Delta S]$ of the match and mismatch behaviors. As can be seen in **Figures 6A,C**, matching and mismatching behaviors generate large decreases in entropy at low or high hit rates, respectively. In addition, the matching behaviors generated penalties (entropy increase) at high numbers of accumulated hits. Please note that entropy is non-monotonous (Hajieghrary et al., 2016; Rodriguez et al., 2017) and can increase on detection to non-detection sequences since the agent's belief is narrowed by the detection but broadens again at the non-detection. **Figures 6B,D** show that the expected rewards are greater at low or high hit rates for mismatching and matching behaviors, respectively. **Figures 7A,C** show histograms of the actual and expected rewards, respectively. We validated the statistical difference in the distributions of the matching and mismatching states (Kolmogorov-Smirnov test $p < 0.01$).

Figure 7B shows the cumulative density function of the actual rewards ΔS for matching and mismatching behaviors. As shown in the figure, mismatching behaviors have a higher probability of greater entropy reductions (particularly values of approximately 10^{-4} and 10^{-1}). Mismatching behaviors also have a higher probability of a larger decrease in entropy (values of approximately -4×10^{-3}) as shown in **Figure 7D**. **Figures 8A,B** show the cumulative odor hits and hit rate against the root mean squared error between the actual

ΔS and the expected reward $E[\Delta S]$. This was calculated as shown in Equation (3), where N is 20 because the sampling frequency of the behavioral measurement system is 20 Hz.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\Delta S_{i+1} - E[\Delta S_i])^2} \quad (3)$$

5. DISCUSSION

In this study, we investigated the possible causes of variability in the programmed behavior model of the male silk moth. Specifically, we asked whether such variability leads to higher information gains; in other words, if it minimizes the information entropy of the probability distribution of the moth regarding the location of an odor source. We also investigated whether the probabilistic framework of infotaxis can explain how the male silk moth selects maneuvers to balance the exploration and exploitation of the expected rewards.

5.1. Relationship Between Behavioral Variability and Information Rewards

In a recent study, Shigaki et al. (2019b) simultaneously measured the odor search behavior of male silkmoths and the neural activity from their lateral accessory lobe (LAL). The LAL generates motor commands in response to odor stimuli. That study found that silkmoths are less likely to “surge” (move forward) as the frequency of odor hits increases. In terms of infotaxis, this can be interpreted as moths preferring rotations (exploration) because, at high odor encounter rates, the expected decrease in entropy is less than at low rates. Our results found that matching and mismatching behaviors generate rewards at high and low hit rates, respectively (**Figure 6**). Thus, this leads us to believe that at high hit rates, silk moths prefer reactive or more exploitative behaviors, and at low rates, they prefer more stochastic or explorative behaviors such as rotations instead of straight forward moves. Furthermore, this tendency was observed on all specimens that reached the odor source.

An interesting interpretation of these results can also be made from the viewpoint of reinforcement learning (RL). In this field, an agent learns to behave according to an optimal policy with the highest expected accumulated reward over a time horizon. Nonetheless, many RL algorithms face the exploration and exploitation dilemma in which greedily selecting the actions with the highest reward can lead to suboptimal policies stuck in the local maxima. A common way to avoid this is to add stochasticity in the selection of actions; thus balancing exploration and exploitation, using methods such as ϵ -greedy algorithms (Sutton and Barto, 2018). An analogy can be made to the behavior of the silkmoth in the sense that some randomness in the selection of the “surge” maneuver leads to higher information gains and possibly a better odor source search performance. This can be clearly seen in **Figures 7B,D**, where

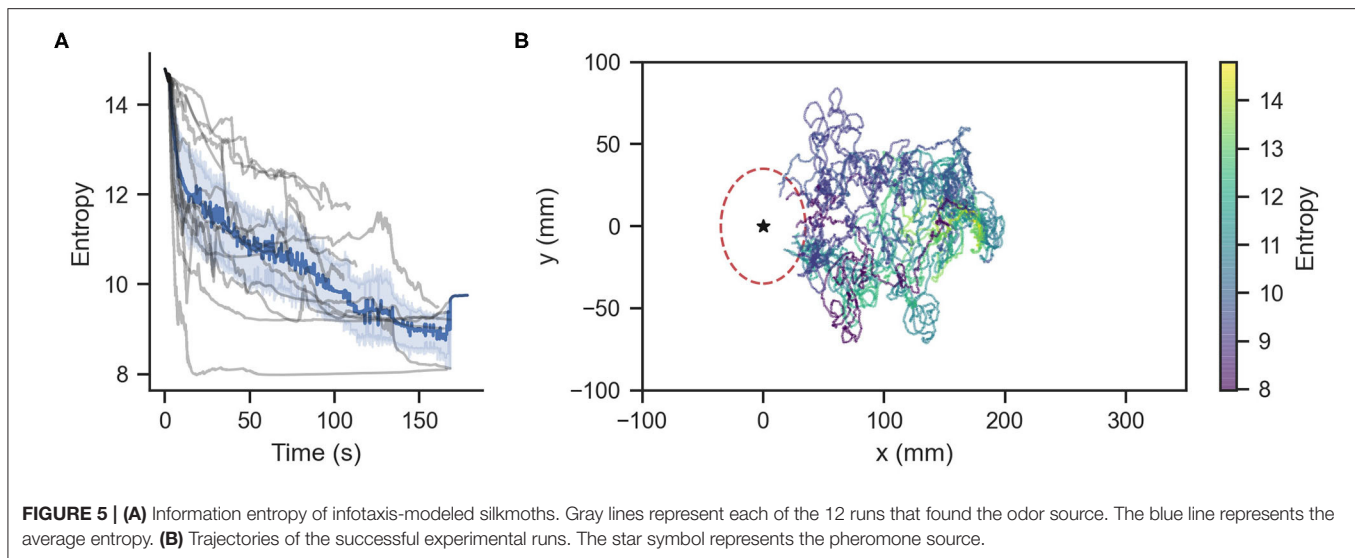


TABLE 2 | Normalized counts of each maneuver taken by the moths.

Kinematic	Temporal	
	Surge	Rotate
Surge	0.1597 ± 0.07	0
Rotate	*0.1939 ± 0.11	0.6464 ± 0.19

The “kinematic” classification scheme is based on the linear and angular velocities of the moths. The “temporal” scheme is based on the time since the last odor hit. The values with the asterisks indicate “mismatching” behaviors.

the probability of obtaining better rewards is higher for the mismatching behaviors.

5.2. Exploration and Exploitation in Silk Moth Behavior

We found that maneuvers that deviate from the programmed behavior model correspond to a larger expected decrease in entropy, that is, a higher expected reward in the terminology of reinforcement learning. Therefore, we demonstrated the capability of the infotaxis strategy to quantitatively express maneuvers that deviate from the programmed behavior as explorative and those that match it as exploitative.

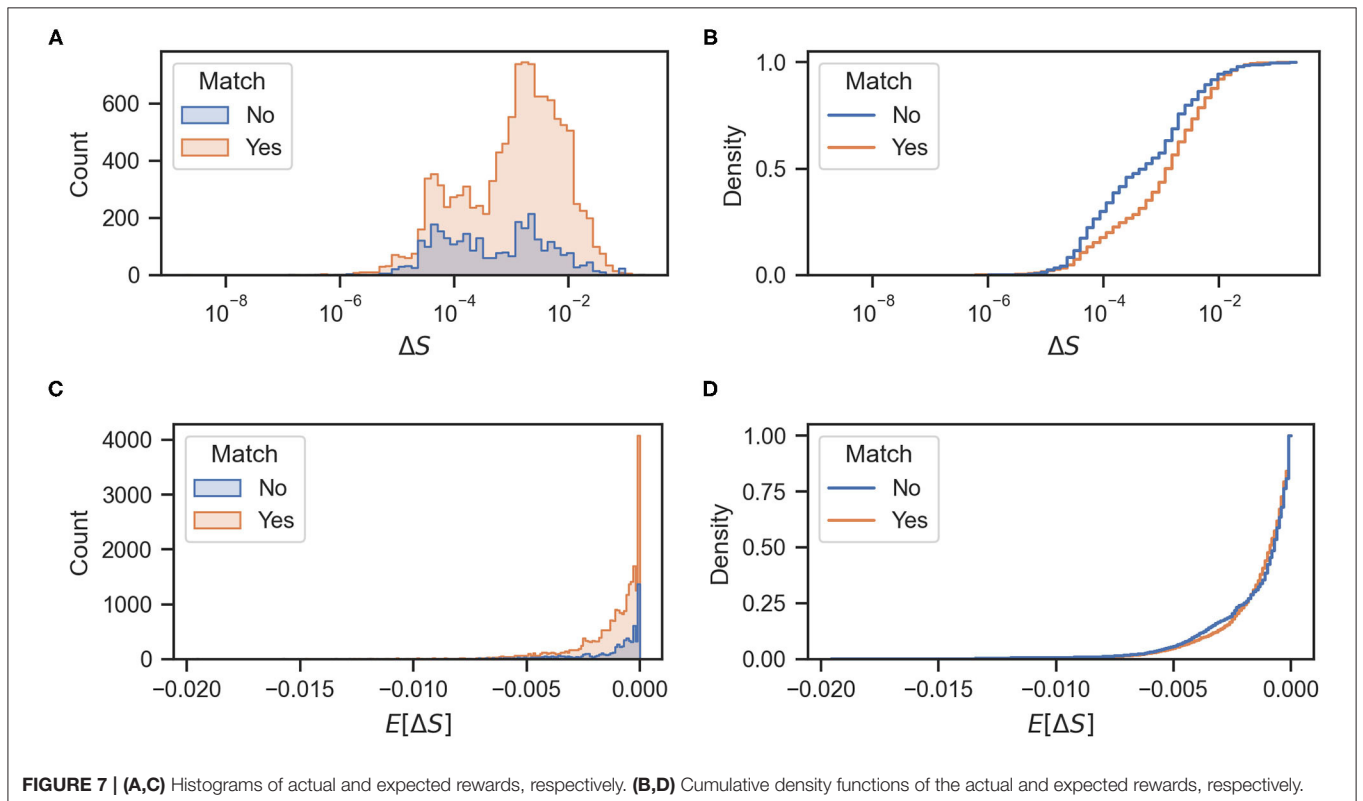
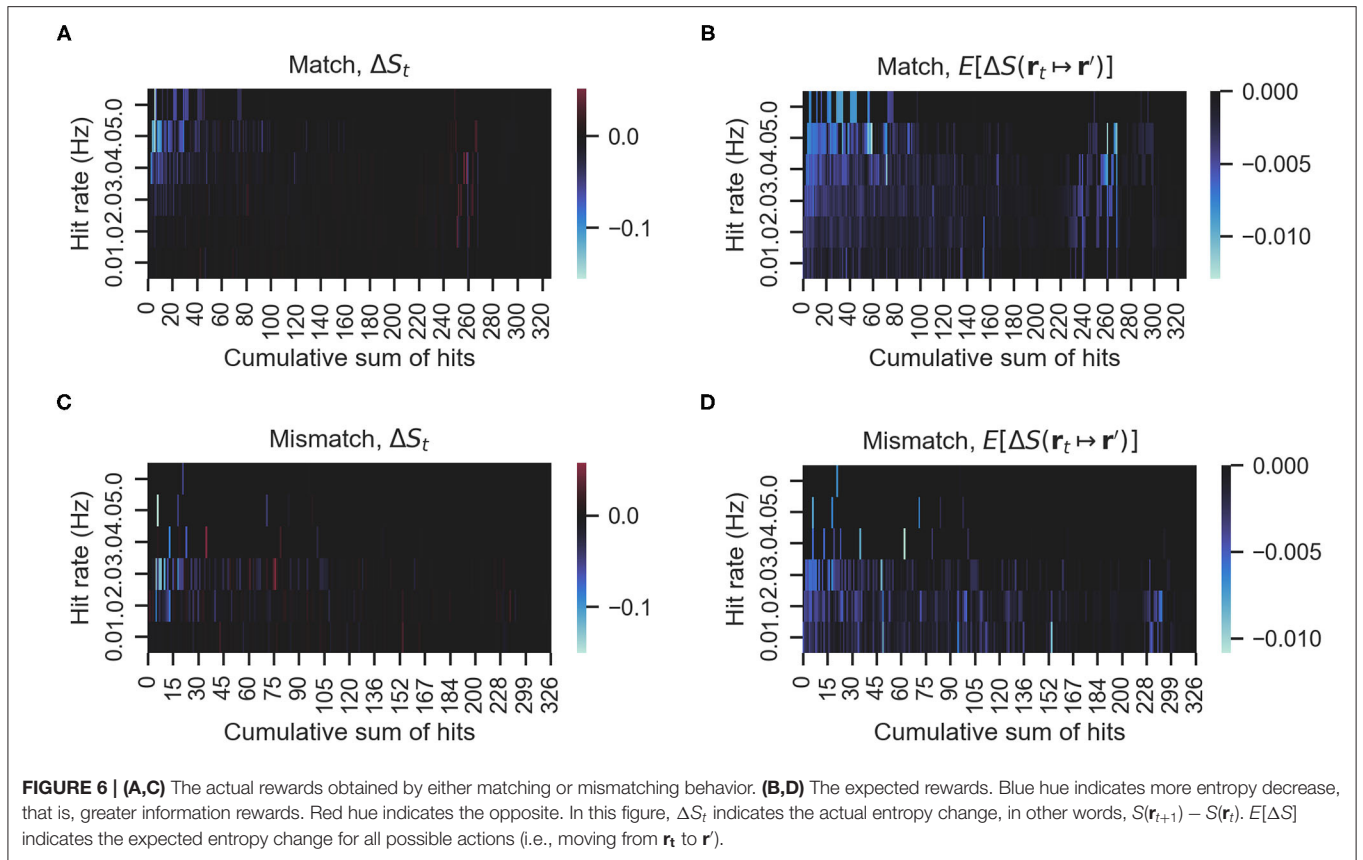
Another interesting point to note is the relationship between matching and mismatching behaviors with the root mean squared error (RMSE) of the real vs. expected rewards. As shown in **Figure 8A**, the error decreases proportionally to the accumulation of odor hits. This is relatively intuitive because more detections narrow down the belief of the source location. However, more RMSE occurs between real and expected rewards at times of high hit rates. Furthermore, the matching behaviors have a lower error than the mismatching behaviors. One possible interpretation for this is that matching behaviors are more

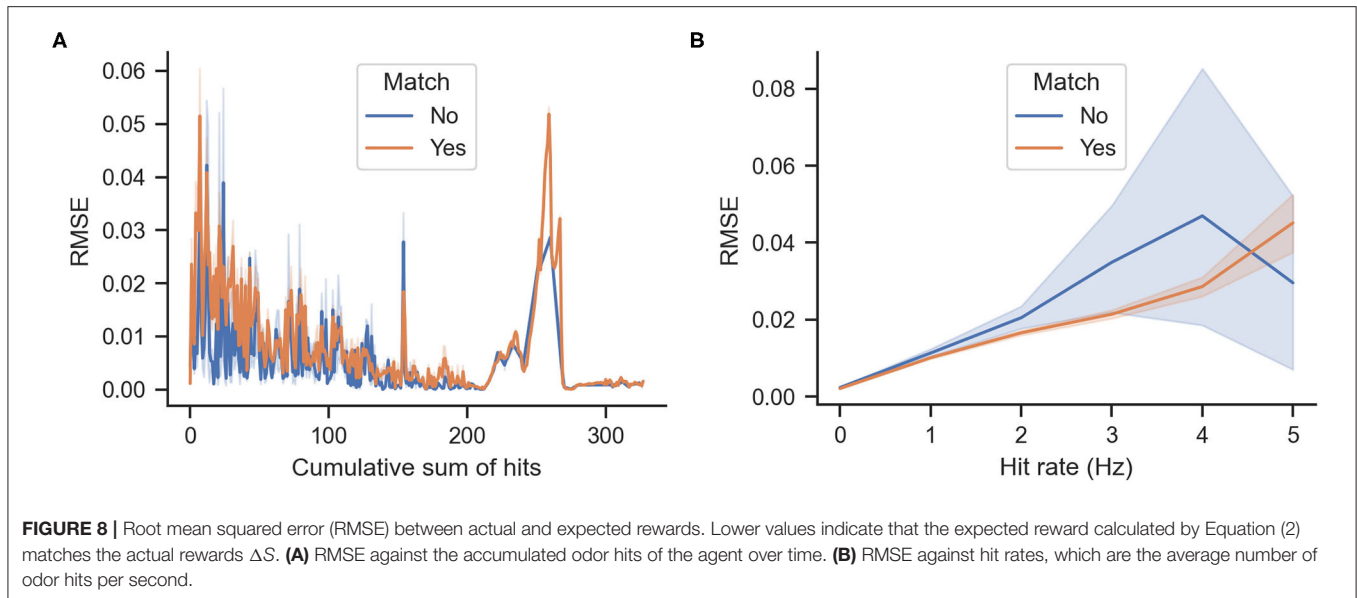
exploitative; thus they are more similar to the greedy infotaxis policy, whereas the mismatching behaviors are more explorative; hence, they differ from the expected reward of the infotaxis strategy.

We believe that being able to represent animal olfactory behavior through a method such as infotaxis is an important contribution to the fields of ethology and robotics because having a representation of the decision process of animals in terms of probabilistic beliefs and expected rewards facilitates the algorithmic implementation of these processes in robots. Furthermore, it allows for the refinement of these decision processes using tools such as machine and reinforcement learning.

6. CONCLUSION

In this study, we measured the behavior of moths using a virtual reality system that presents accurate and reproducible odor stimuli by using blue light and optogenetic moths. We then took trajectories from these measurements and modeled them as an infotaxis (Vergassola et al., 2007) strategy. We used infotaxis-based modeling to determine if variability in the silkmoth behavior is related to higher gains in information regarding the probabilistic distribution of the source location. We found that variations have a higher probability of obtaining larger information gains than “programmed behaviors” (i.e., reactive, exploitative behaviors). This suggests that silkmoths incorporate some stochasticity into their behavior to balance the exploration and exploitation of information gains. Future studies should be conducted to develop ways to extract decision-making mechanisms from free-running silkmoths. In this study, we used tethered moths walking on a treadmill, and, although such a device imposes minimal disturbances on the moth behavior, we believe it is necessary to study whether models from free-running experiments will differ from those in this specific study. It would also be useful to develop an olfactory search algorithm based on





the silkmoth exploration/exploitation mechanisms elucidated in this paper and then implement such an algorithm on a robot to test whether the search performance is improved compared with either the programmed behavior or the infotaxis strategy.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Files**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

CH-R, SF, SS, and DK contributed conception and design of the study. SF conducted the virtual reality silkmoth experiments. CH-R performed the numerical analyses and wrote

the manuscript. TS, RK, and HS provided genetically modified silkmoths. All authors read and approved the submitted version.

FUNDING

CH-R acknowledges funding from Instituto de Innovación y Transferencia de Tecnología (I2T2) and Consejo Nacional de Ciencia y Tecnología (CONACYT). This work was also partially supported by JSPS KAKENHI under Grant JP19H02104, JP19H04930, and JP19K14943.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncom.2021.629380/full#supplementary-material>

REFERENCES

- Ache, B. W., Hein, A. M., Bobkov, Y. V., and Principe, J. C. (2016). Smelling time: a neural basis for olfactory scene analysis. *Trends Neurosci.* 39, 649–655. doi: 10.1016/j.tins.2016.08.002
- Baker, K. L., Dickinson, M., Findley, T. M., Gire, D. H., Louis, M., Suver, M. P., et al. (2018). Algorithms for olfactory search across species. *J. Neurosci.* 38, 9383–9389. doi: 10.1523/JNEUROSCI.1668-18.2018
- Calhoun, A. J., Chalasani, S. H., and Sharpee, T. O. (2014). Maximally informative foraging by *Caenorhabditis elegans*. *Elife* 3:e04220. doi: 10.7554/eLife.04220
- Celani, A., Villermaux, E., and Vergassola, M. (2014). Odor landscapes in turbulent environments. *Phys. Rev. X* 4:041015. doi: 10.1103/PhysRevX.4.041015
- Gatellier, L., Nagao, T., and Kanzaki, R. (2004). Serotonin modifies the sensitivity of the male silkmoth to pheromone. *J. Exp. Biol.* 207, 2487–2496. doi: 10.1242/jeb.01035
- Hajieghrary, H., Hsieh, M. A., and Schwartz, I. B. (2016). Multi-agent search for source localization in a turbulent medium. *Phys. Lett. A* 380, 1698–1705. doi: 10.1016/j.physleta.2016.03.013
- Kanzaki, R., Sugi, N., and Shibuya, T. (1992). Self-generated zigzag turning of *bombyx mori* males during pheromone-mediated upwind walking (physiology). *Zool. Sci.* 9, 515–527.
- Loudon, C., and Koehl, M. (2000). Sniffing by a silkworm moth: wing fanning enhances air penetration through and pheromone interception by antennae. *J. Exp. Biol.* 203, 2977–2990.
- Mafra-Neto, A., and Cardé, R. T. (1994). Fine-scale structure of pheromone plumes modulates upwind orientation of flying moths. *Nature* 369:142. doi: 10.1038/369142a0
- Minegishi, R., Takashima, A., Kurabayashi, D., and Kanzaki, R. (2012). Construction of a brain-machine hybrid system to evaluate adaptability of an insect. *Robot. Auton. Syst.* 60, 692–699. doi: 10.1016/j.robot.2011.06.012
- Obara, Y. (1979). *Bombyx mori* mating dance: an essential in locating the female. *Appl. Entomol. Zool.* 14, 130–132. doi: 10.1303/aez.14.130
- Pang, R. (2018). *Infotaxis Summary*. Available online at: https://nbviewer.jupyter.org/github/rkp8000/infotaxis/blob/master/test_infotaxis.ipynb
- Pang, R., van Breugel, F., Dickinson, M., Riffell, J. A., and Fairhall, A. (2018). History dependence in insect flight decisions during odor tracking. *PLoS Comput. Biol.* 14:e1005969. doi: 10.1371/journal.pcbi.1005969

- Rodríguez, J. D., Gómez-Ullate, D., and Mejía-Monasterio, C. (2017). On the performance of blind-infotaxis under inaccurate modeling of the environment. *Eur. Phys. J. Spec. Top.* 226, 2407–2420. doi: 10.1140/epjst/e2017-70067-1
- Shigaki, S., Fikri, M. R., Hernandez Reyes, C., Sakurai, T., Ando, N., Kurabayashi, D., et al. (2018a). Animal-in-the-loop system to investigate adaptive behavior. *Adv. Robot.* 32, 945–953. doi: 10.1080/01691864.2018.1511473
- Shigaki, S., Haigo, S., Reyes, C. H., Sakurai, T., Kanzaki, R., Kurabayashi, D., et al. (2019a). Analysis of the role of wind information for efficient chemical plume tracing based on optogenetic silkworm moth behavior. *Bioinspir. Biomimet.* 14:046006. doi: 10.1088/1748-3190/ab1d34
- Shigaki, S., Sakurai, T., Ando, N., Kurabayashi, D., and Kanzaki, R. (2018b). Time-varying moth-inspired algorithm for chemical plume tracing in turbulent environment. *IEEE Robot. Autom. Lett.* 3, 76–83. doi: 10.1109/LRA.2017.2730361
- Shigaki, S., Shiota, Y., Kurabayashi, D., and Kanzaki, R. (2019b). Modeling of adaptive chemical plume tracing algorithm of insect using fuzzy inference. *IEEE Trans. Fuzzy Syst.* 28, 72–84. doi: 10.1109/TFUZZ.2019.2915187
- Sutton, R. S., and Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT press.
- Tabuchi, M., Sakurai, T., Mitsuno, H., Namiki, S., Minegishi, R., Shiotsuki, T., et al. (2013). Pheromone responsiveness threshold depends on temporal integration by antennal lobe projection neurons. *Proc. Natl. Acad. Sci. U.S.A.* 2013:201313707. doi: 10.1073/pnas.1313707110
- Thrun, S., Burgard, W., and Fox, D. (2005). *Probabilistic Robotics*. Cambridge, MA: MIT Press.
- Tomioka, K., Ikeda, M., Nagao, T., and Tamotsu, S. (1993). Involvement of serotonin in the circadian rhythm of an insect visual system. *Naturwissenschaften* 80, 137–139. doi: 10.1007/BF01131019
- van Breugel, F., and Dickinson, M. H. (2014). Plume-tracking behavior of flying drosophila emerges from a set of distinct sensory-motor reflexes. *Curr. Biol.* 24, 274–286. doi: 10.1016/j.cub.2013.12.023
- Vergassola, M., Villermaux, E., and Shraiman, B. I. (2007). “Infotaxis” as a strategy for searching without gradients. *Nature* 445:406. doi: 10.1038/nature05464
- Vickers, N. J. (2005). Winging it: moth flight behavior and responses of olfactory neurons are shaped by pheromone plume dynamics. *Chem. Senses* 31, 155–166. doi: 10.1093/chemse/bjj011
- Voges, N., Chaffiol, A., Lucas, P., and Martinez, D. (2014). Reactive searching and infotaxis in odor source localization. *PLoS Comput. Biol.* 10:e1003861. doi: 10.1371/journal.pcbi.1003861

Conflict of Interest: SF was employed by MHPS Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Hernandez-Reyes, Fukushima, Shigaki, Kurabayashi, Sakurai, Kanzaki and Sezutsu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX A: INFOTAXIS STRATEGY

Herein, we provide a more detailed explanation of the derivation of the infotaxis formulae. We based this explanation on the work of Pang (2018) and the original infotaxis strategy developed by Vergassola et al. (2007). The agent's belief in the source location $P(\mathbf{r}_{src}|\mathcal{T}_t)$ can be written using Bayes' theorem, as indicated in Equation (A1).

$$P(\mathbf{r}_{src}|\mathcal{T}_t) = \frac{P(\mathcal{T}_t|\mathbf{r}_{src})P(\mathbf{r}_{src})}{P(\mathcal{T}_t)} \propto P(\mathcal{T}_t|\mathbf{r}_{src})P(\mathbf{r}_{src}) \quad (A1)$$

where $P(\mathcal{T}_t|\mathbf{r}_{src})$ is the likelihood of the source position and $P(\mathbf{r}_{src})$ is the prior distribution of the source. Infotaxis assumes that odor hits and misses are independent of one another and the likelihood of the source position takes the following form:

$$P(\mathcal{T}_t|\mathbf{r}_{src}) = \prod_t P(h(\mathbf{r}_t)|\mathbf{r}_{src}) \quad (A2)$$

where $h(\mathbf{r}_t)$ is 1 if the agent detects an odor hit at time t and 0 if it detects a miss. The infotaxis strategy considers that the number of hits follows a Poisson distribution; hence, the probability of a hit or miss becomes the following:

$$P(h(\mathbf{r}_t) = 0|\mathbf{r}_{src}) = \exp[-R(\mathbf{r}_t|\mathbf{r}_{src})\Delta t] \quad (A3)$$

$$P(h(\mathbf{r}_t) > 0|\mathbf{r}_{src}) = 1 - P(h(\mathbf{r}_t) = 0|\mathbf{r}_{src}) \quad (A4)$$

where $R(\mathbf{r}_t|\mathbf{r}_{src})\Delta t$ is the mean rate of hits the agent expects at \mathbf{r}_t , during a time period of Δt , given a source position \mathbf{r}_{src} . The **Supplementary Material** of the original paper on infotaxis indicate that the hit rate is derived from the advection-diffusion equation of a turbulent plume and define it as follows:

$$R(\mathbf{r}|\mathbf{r}_{src}) = \frac{E}{\ln\left(\frac{\lambda}{\alpha}\right)} e^{\frac{(x_{src}-x)V}{2D}} K_0\left(\frac{|\mathbf{r}-\mathbf{r}_{src}|}{\lambda}\right) \quad (A5)$$

$$\lambda = \sqrt{\frac{D\tau}{1 + V^2\tau/(4D)}} \quad (A6)$$

where E is the emission rate of odor particles, which have an effective diffusivity D and, a finite lifetime τ , and are advected by a wind with mean velocity V that blows in the positive x -direction. K_0 is the modified Bessel function of order zero, and α is the radius of a round-shaped agent. In our calculations of the information entropy of the silkmoth we used the following parameters into Equation A6: $\alpha=10$ mm, $E=1$, $\tau=6.3$ s, $D=0.012$, and $V=0.1$ m/s to match the wind speed in the moth experiments. The range of possible values for the source location was a 1730×770 lattice; i.e. the size of each cell was 0.26 mm. For Equation A4 we set Δt to 50 ms; which is the same as the sampling period of the treadmill described in section 3.1. At each time step, the belief

of the source position distribution $P(\mathcal{T}_t|\mathbf{r}_{src})$ can be recursively updated as follows:

$$P(\mathbf{r}_{src}|\mathcal{T}_{t+1}) = P(h(\mathbf{r}_{t+1})|\mathbf{r}_{src})P(\mathbf{r}_{src}|\mathcal{T}_t) \quad (A7)$$

At each time step, the agent considers five possible actions: moving forward, backward, left, right, or waiting. For each possible action, it calculates the probability p_* that the action will result in finding the source:

$$p_* = \sum_{\mathbf{r}_{src}} P(\mathbf{r}_{src}|\mathcal{T}_t)P(|\mathbf{r}' - \mathbf{r}_{src}| \approx 0) \quad (A8)$$

Consequently, the probability of not finding the source is $1 - p_*$. If the source is found, then the entropy of the belief will become zero, that is, $\Delta S^*=(0 - S_t)=-S_t$. To balance the exploration and exploitation, the agent also considers the case in which it does not find the source after taking an action. In such case, it would sample from the environment either a miss with a probability p_m or a hit with a probability $p_h=1 - p_m$. The probability of sampling a miss is the average of the miss probability over the range of possible source locations:

$$p_m = \sum_{\mathbf{r}_{src}} P_m(\mathbf{r}_{src}|\mathcal{T}_t) = \sum_{\mathbf{r}_{src}} P(h(\mathbf{r}') = 0|\mathbf{r}_{src})P(\mathbf{r}_{src}|\mathcal{T}_t) \quad (A9)$$

where \mathbf{r}' is the future position of the agent after taking an action. The agent also estimates how its source position belief, as well as its entropy, would change after moving. The change in entropy after sampling a miss or a hit at \mathbf{r}' would be the following:

$$\Delta S_m = - \sum_{\mathbf{r}_{src}} P_m(\mathbf{r}_{src}|\mathcal{T}_t) \ln(P_m(\mathbf{r}_{src}|\mathcal{T}_t)) - S_t \quad (A10)$$

$$\Delta S_h = - \sum_{\mathbf{r}_{src}} P_h(\mathbf{r}_{src}|\mathcal{T}_t) \ln(P_h(\mathbf{r}_{src}|\mathcal{T}_t)) - S_t \quad (A11)$$

Overall, the agent calculates the expected change of entropy by moving from \mathbf{r}_t to \mathbf{r}' as follows:

$$E[\Delta S(\mathbf{r}_t \mapsto \mathbf{r}')] = p^*(0 - S_t) + (1 - p^*)(p_m \Delta S_m + p_h \Delta S_h) \quad (A12)$$

where the terms on the left and right sides of the sum are the change in entropy if the source is found or not found at \mathbf{r}' , respectively. Finally, the agent chooses the action a with the largest expected decrease in entropy (see **Figure 3D**) as:

$$a = \arg \min_{\mathbf{r}'} (E[\Delta S(\mathbf{r}_t \mapsto \mathbf{r}')]) \quad (A13)$$

After making a move, the agent encounters either a miss or a hit from the odor plume and updates the probability distribution of the source location. The agent then repeats the navigation policy process iteratively until it finds the source.