A neuromorphic model of active vision shows how spatiotemporal encoding in lobula neurons can aid pattern recognition in bees

HaDi MaBouDi[1,2,3,4*], Mark Roper[4,5], Marie-Genevieve Guiraud[4,6], Mikko Juusola[2,3], Lars Chittka[4], James A.R. Marshall[1,3]

[1]Department of Computer Science, University of Sheffield, Sheffield, UK
[2]School of Biosciences, University of Sheffield, Sheffield, UK
[3]Neuroscience Institute, University of Sheffield, Sheffield, UK
[4]School of Biological and Behavioural Sciences, Queen Mary University of London, London, UK
[5]Drone Development Lab, Ben Thorns Ltd, Colchester, Essex, UK
[6]School of Natural Sciences, Macquarie University, North Ryde, NSW, Australia

*Corresponding author: HaDi MaBouDi, maboudi@gmail.com

Short title: *Insect-inspired model for active vision*

Abstract: Bees' remarkable visual learning abilities make them ideal for studying active information acquisition and representation. Here, we develop a biologically inspired model to examine how flight behaviours during visual scanning shape neural representation in the insect brain, exploring the interplay between scanning behaviour, neural connectivity, and visual encoding efficiency. Incorporating non-associative learning—adaptive changes without reinforcement—and exposing the model to sequential natural images during scanning, we obtain results that closely match neurobiological observations. Active scanning and non-associative learning dynamically shape neural activity, optimising information flow and representation. Lobula neurons, crucial for visual integration, self-organise into orientation-selective cells with sparse, decorrelated responses to orthogonal bar movements. They encode a range of orientations, biased by input speed and contrast, suggesting co-evolution with scanning behaviour to enhance visual representation and support efficient coding. To assess the significance of this spatiotemporal coding, we extend the model with circuitry analogous to the mushroom body, a region linked to associative learning. The model demonstrates robust performance in pattern recognition, implying a similar encoding mechanism in insects. Integrating behavioural, neurobiological, and computational insights, this study highlights how spatiotemporal coding in the lobula efficiently compresses visual features, offering broader insights into active vision strategies and bio-inspired automation.

Impact statements: Active vision dynamically refines spatiotemporal neural representations, optimising visual processing through scanning behaviour and non-associative learning, providing insights into efficient sensory encoding in dynamic environments.

Keywords: active vision, image statistics, lobula, mushroom bodies, non-associative learning, orientation selective neurons, scanning behaviour, visual recognition.

Introduction

Bees are capable of remarkable cognitive feats, particularly in visual learning (Chittka, 2022; Srinivasan, 2010; Turner, 1911; Von Frisch, 1914; Wehner, 1967). They can not only learn to associate a colour or orientation of a bar with reward (Dyer et al., 2011; Guiraud et al., 2025b; MaBouDi et al., 2020b; Srinivasan, 1994; Stach et al., 2004) but are also able to identify specific features to categorise visual patterns, by finding the relevant stimuli properties (Benard et al., 2006; Guiraud et al., 2025a; Stach et al., 2004). Furthermore, bees have demonstrated the capacity to grasp abstract concepts (Avarguès-Weber et al., 2011; Giurfa et al., 2001; Guiraud et al., 2018; MaBouDi et al., 2020c; Menzel, 2012) and solve numerosity tasks by sequentially scanning the elements within a stimulus (MaBouDi et al., 2020a). These exceptional capabilities position bees as a valuable animal model for investigating the principles of visual learning through the analysis of their behavioural responses (Menzel and Giurfa, 2006; Srinivasan, 2010). Nevertheless, it remains unclear how bees, despite their supposedly low visual acuity (Gribakin, 1975; Spaethe and Chittka, 2003a; Srinivasan and Lehrer, 1988) and limited neural resources, recognise complex patterns and perceive the intricacies of the natural world encountered during foraging (Chittka and Niven, 2009; Giurfa, 2013).

The natural scene that animals encounter is structured differently from random or artificial ones (Clark et al., 2014a; Ruderman and Bialek, 1994; Simoncelli and Olshausen, 2001; Zimmermann et al., 2018). It has been hypothesised that visual sensory neurons evolve to exploit statistical regularities in natural scenes, efficiently encoding information through their spatiotemporal structures (Barlow, 1961). Over evolutionary time, insect visual neurons have developed mechanisms that provide robust and efficient responses to naturalistic inputs (Dyakova et al., 2019, 2015; Dyakova and Nordström, 2017; Juusola et al., 2025; Song and Juusola, 2014; Zheng et al., 2006). For instance, Song and Juusola (2014) showed that fly photoreceptors extract more information from naturalistic time series than from artificial stimuli or white noise, yielding stronger responses with a higher signal-to-noise ratio (Song and Juusola, 2014). Additionally, numerous studies have demonstrated that insect sensory pathways and their associated behaviours dynamically adapt to varying environmental conditions, adjusting their responses based on input parameters such as contrast, spatial frequency, and spatiotemporal correlations (Arenz et al., 2017; Brinkworth and O'Carroll, 2009; Clark et al., 2014b; Dyakova et al., 2019; Dyakova and Nordström, 2017; Juusola et al., 2025; Juusola and Song, 2017; Schwegmann et al., 2014; Serbe et al., 2016; Song and Juusola, 2014, 2014; van Hateren, 1997; van Hateren, 1992). Experience-dependent adaptation has been observed in fly photoreceptors and motion-sensitive neurons in the lobula plate, enabling efficient visual processing under varying conditions. For instance, photoreceptors adapt their response dynamics to different light intensities, optimizing sensitivity to natural stimuli (Juusola and Hardie, 2001a, 2001b). Similarly, motion-sensitive neurons such as the H1 neuron adjust their response properties based on prior motion exposure,

enhancing motion detection in dynamic environments (Maddess et al., 1997). This dynamic plasticity allows insects to process ecologically relevant information in real time. However, the precise neural mechanisms underlying natural scene processing remain elusive and require further investigation. Here we examine how insect visual circuitry has adapted to regularities in natural scenes, focusing on the efficient coding strategies and robust response mechanisms that enhance visual pattern recognition.

In animal vision, active sampling strategies—wherein animals actively scan their environment to extract visual information over time—are widely observed across species (Land, 1999; Land and Nilsson, 2012; Severance and Washburn, 1907; Varella et al., 2024; Washburn, 1926, 1916; Yarbus, 2013). Primates employ eye movements, including saccades and microsaccades, to enhance fine spatial resolution and improve the encoding of natural stimuli (Anderson et al., 2020; Land, 1999; Näher et al., 2023; Rucci et al., 2007; Rucci and Victor, 2015). Similarly, insects utilise active vision strategies, incorporating characteristic head and body movements or specific approach trajectories to optimise visual processing during behavioural tasks (Bertrand et al., 2021; Chittka and Skorupski, 2017; Dawkins and Woodington, 2000; Egelhaaf et al., 2009; Land, 1973; Land and Nilsson, 2012; Langridge et al., 2021; MaBouDi et al., 2021b). Recent studies have shown that *Drosophila* generate photomechanical photoreceptor microsaccades and can move their retinas to stabilise their retinal images, achieving hyperacute vision and enhancing depth perception (Fenk et al., 2022; Hardie and Franze, 2012; Juusola et al., 2017; Kemppainen et al., 2022b, 2022a). Likewise, honeybee vision may require sequential sampling and integration of colour information due to their limited ability to discriminate between similar hues in brief flashes (<50 ms) (Nityananda et al., 2014). To overcome this constraint, bees engage in systematic scanning movements, continuously sampling their surroundings to construct an internal neural representation of their environment (Boeddeker et al., 2015; Collett et al., 1993; Doussot et al., 2021; Guiraud et al., 2018; Kemppainen et al., 2022a; Langridge et al., 2018; Lehrer and Collett, 1994; MaBouDi et al., 2020a; Werner et al., 2016). For instance, bumblebees enumerate visual elements sequentially rather than processing them in parallel, suggesting a reliance on scanning behaviour for feature extraction parallel (MaBouDi et al., 2020a), and their flight trajectories further indicate that they prioritise specific pattern regions before making a decision, rather than processing the entire pattern globally (Langridge et al., 2021; MaBouDi et al., 2021b). Given the low-resolution nature of compound eyes and the potentially reduced parallel processing capacity in insects compared to vertebrates, it is likely that bees rely on active vision and sequential sampling to construct a more robust neural representation of their (Chittka and Skorupski, 2017; Nityananda et al., 2014). These active strategies, akin to primate eye movements, play a crucial role in early visual processing, redundancy reduction, and efficient encoding of visual stimuli (Doussot et al., 2021; Kuang et al., 2012; Odenthal et al., 2021). However, it remains poorly understood how such mechanisms allow bees to overcome representational constraints, detect visual regularities, and

113 solve complex discrimination tasks. Understanding these strategies is key to uncovering the
114 fundamental principles of insect vision and their broader implications for visual processing across
115 biological and artificial systems.

116 Building on our previous work analysing bee flight paths during a simple visual task (MaBouDi et al.,
117 2021b), we further investigated the main circuit elements that contribute to active vision in achromatic
118 pattern recognition, focusing on a simplified yet biologically plausible model. Our primary objective
119 was to determine how bees' scanning behaviour contributes to the functional organisation and
120 connectivity of neurons in the visual lobe. We hypothesised that the bees' scanning behaviours have
121 co-adapted to sample complex visual features in a way that efficiently encodes them into
122 spatiotemporal patterns of activity in the lobula neurons, facilitating distinct and specific
123 representations that support learning in the bees' compact brain. To test this, we developed a
124 neuromorphic model of the bee optic lobes incorporating efficient coding principles via a novel model
125 of non-associative plasticity. This model demonstrates how spatial scanning behaviour in response
126 to naturalistic visual inputs has shaped the connectivity within the medulla (the 2nd optic ganglion) to
127 facilitate an efficient representation of these inputs in the lobula (the 3rd optic ganglion). This efficiency
128 is achieved through the self-organisation of a specific set of orientation-selective neuron in the lobula,
129 highlighting the combined impact of scanning behaviour and non-associative learning on shaping the
130 neural circuitry within the bees' optic lobes.

131 To evaluate the proposed visual network, we enhance our visual processing framework by
132 incorporating a secondary decision-making module inspired by insect associative learning
133 mechanisms, grounded in previous neurobiological evidence (Cassenaer and Laurent, 2012; Fiala
134 and Kaun, 2024; Fisher et al., 2015; Li et al., 2017; Okada et al., 2007; Paulk et al., 2009; Paulk and
135 Gronenberg, 2008). Visual input and flight dynamics for the model were derived from our observations
136 of bee behaviour during a visual discrimination task (MaBouDi et al., 2021b). This allowed us to
137 evaluate and test the hypothesis of active sampling from our model against real-world behaviour
138 results (MaBouDi et al., 2021b), as well as other published visual discrimination tasks performed by
139 bees (Benard et al., 2006; Dyer et al., 2005; Guiraud et al., 2022, 2025a; Srinivasan, 2010, 1994;
140 Zhang and Horridge, 1992). Furthermore, we conducted a detailed analysis comparing the neural
141 response features emerging from our model with existing neurobiological findings (James and Osorio,
142 1996; Paulk et al., 2008; Seelig and Jayaraman, 2013; Yang and Maddess, 1997). This alignment
143 enhances the credibility of our model in capturing essential aspects of neural processing underlying
144 active vision.

145

146

Results

A bio-inspired neural network for active vision

To investigate how bee scanning behaviour optimises neural activity in the visual lobes and enhances visual information processing for efficient pattern recognition, we developed a neural network model inspired by the key morphological and functional characteristics of the bee brain (Figure 1A, C). The network abstracts the circuitry responsible for the initial processing of visual input in the bee's lamina and medulla (the first and second optic ganglia). To mimic temporal encoding during scans (Figure 1B), we introduced a progressive time delay of 1–5 'temporal instances' between the outputs of medulla neurons and wide-field neurons in the lobula (the third optic ganglion; Figure 1D, see Methods). This temporal structuring facilitates sequential sampling of specific locations along the scan trajectory, gradually integrating visual information into a coherent internal representation that emerges as the final output of the lobula neurons.

Building on findings from bee scanning behaviour (MaBouDi et al., 2021b), the model extracts image input in five sequential patches of 75 × 75 pixels, sampled at a speed of 0.1 m/s, corresponding to a lateral displacement of 15 pixels between consecutive patches (Figure 1B; see Methods for details). The green pixel intensities of each patch modulate the membrane potentials of 5,625 (75 × 75) grid photoreceptors within the simulated bee's single eye. These photoreceptor responses converge onto 625 lamina neurons via recurrent neural connectivity, providing a feedforward mechanism for transferring visual information. The lamina neurons then project to 250 small-field medulla neurons through a simple feedforward pathway (Figure 1C, see Methods).

For each of the five sequential patches that compose a full scan, medulla neuron responses are computed using a spiking neural model. These responses are progressively integrated into the synapses of their corresponding lobula neuron with a structured time delay. As depicted in Figure 1D, the synaptic weights dynamically encode the visual information at different temporal instances (T, 2T, 3T, 4T, 5T), effectively aligning sequentially sampled spatial information into a temporally coherent representation. This ensures that the lobula neuron accumulates and processes the underlying medulla input signals at a synchronised time point, mirroring mechanisms that may occur in biological systems (see Discussion). Additionally, lateral inhibitory connections (red connections in Figure 1C) are proposed between lobula neurons to reduce correlation between them, enhancing redundancy reduction in the process.

It is important to note that this proposed spatiotemporal coding is a simplification. In the bee brain, similar processes are likely mediated through dendritic and synaptic latencies, as well as intermediate neuron transmission within the medulla, influenced by non-associative learning in the visual lobe (Figure 1C, D). We hypothesise that connectivity in the medulla and lobula can be refined through exposure to sequences of time-varying images, incorporating non-associative learning rules and efficient coding principles. These mechanisms are optimised and shaped through a generative

183   learning process to align with the statistical properties of natural scenes, enhancing the system's

184   capacity for processing complex visual inputs (see next section).

185
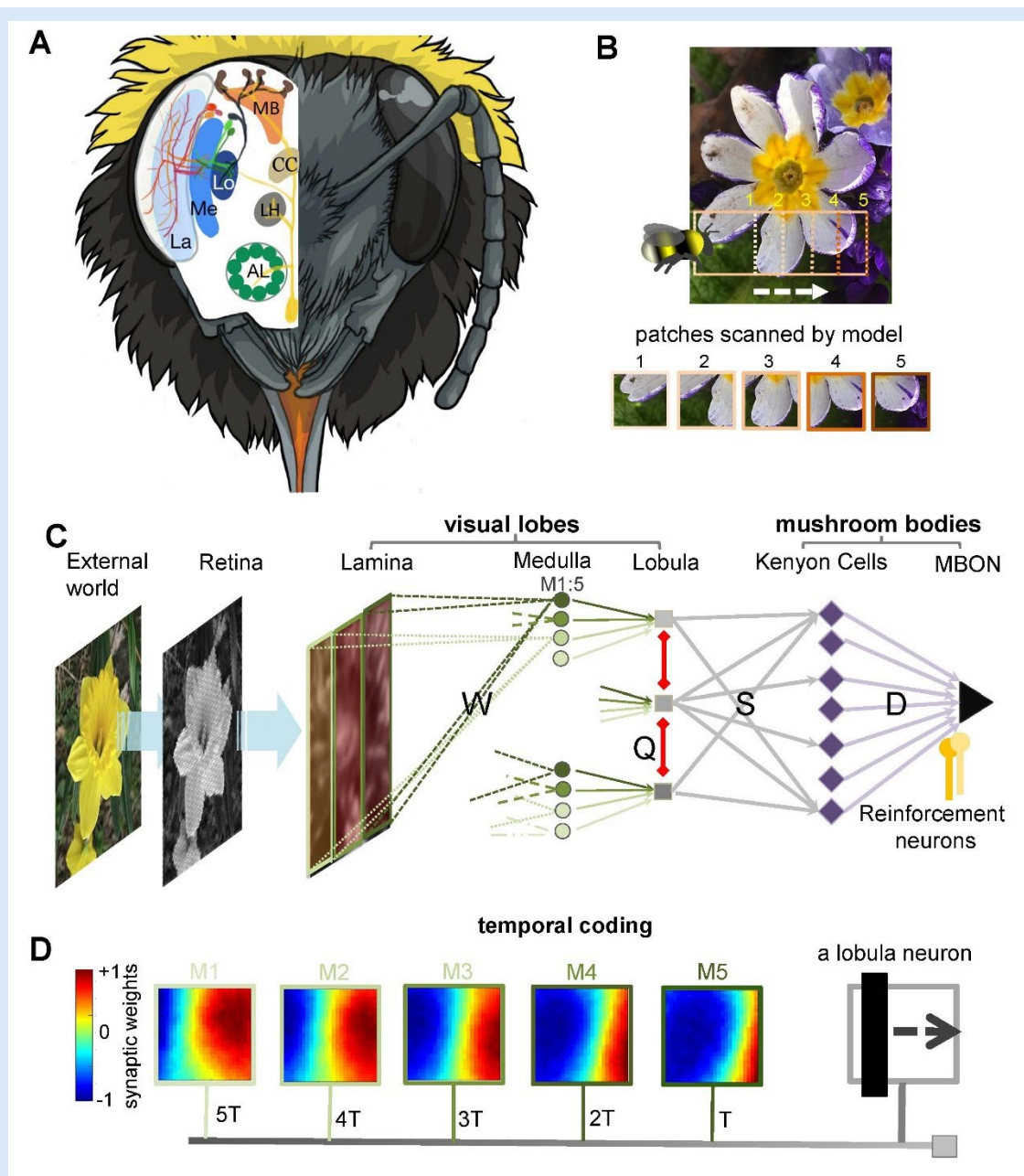
Figure 1. Neural network of active vision inspired by neurobiology and flight dynamics of bees. (A) The right side displays the front view of the bumblebee head showing the component eye and antenna. Left hand side presents a schematic view of the bee's brain regions. Part of neural pathways from the retina to the mushroom bodies are also represented. Labels: AL – Antennal lobe; LH – Lateral horn; CC – Central complex; La – Lamina; Me – Medulla; Lo – Lobula; MB – Mushroom body. Figure was designed by Alice Bridges **(B)** A representation of the modelled bee's scanning behaviour of a flower demonstrating how a sequence of patches project to the simulated bee's eye with lateral movement from left to right. Below are five image patches sampled by the simulated bee. **(C)** Representation of the neural network model of active vision inspired by micromorphology of the bee brain that underlie learning, memory and experience-

dependent control of behaviour. The photoreceptors located in the eye are excited by the input pattern. The activities of photoreceptors change the membrane potential of a neuron in the next layer, Lamina. The lamina neurons send signals (through W connectivity matrix) to the medulla neurons to generate spikes in this layer. Each wide-field lobula neuron integrates the synaptic output of five small-field medulla neurons. The lobula neurons are laterally inhibited by local lobula interconnections (via Q connectivity matrix). Lobula neurons project their axons into the mushroom body, forming connections with Kenyon Cells (KCs) through a randomly weighted connectivity matrix, S. The KCs all connect to a single mushroom body output neuron (MBON) through random synaptic connections D. A single reinforcement neuron (yellow neuron) modulates the synaptic weights between KCs and MBON by simulating the release of octopamine or dopamine when presented with specific visual stimuli (see Method section). **(D)** A temporal coding model that is proposed as the connectivity between medulla and lobula neurons. Each matrix shows the inhibitory (blue) and excitatory (red) connectivity between lamina neurons to a medulla neuron at a given time delay. In this model, the five small-field medulla neurons that are activated by the locally visual input, at different times of scanning, send their activities to a wide-field lobula neuron with a synaptic delay such that the lobula neuron receives all medulla input signals at the same instance (*i.e.* in the presented simulation the lobula neuron is maximally activated by the black vertical bar passing across the visual field from the left to right. Each underlying medulla neuron encodes the vertical bar in a different location of the visual field).

186

187 The neural representation of the visual inputs was subsequently transmitted and processed in the
188 mushroom body - the learning centre of the bee brain (Ehmer and Gronenberg, 2002; Li et al., 2017;
189 Paulk et al., 2008; Paulk and Gronenberg, 2008; Schmalz et al., 2022) (Figure 1C). To simplify the
190 model, we incorporated a single mushroom body output neuron (MBON), whose firing rate reflects
191 the simulated bee's preference for a given visual input. By adjusting synaptic weights within the
192 mushroom body, the network was trained to classify visual patterns as either positive (low MBON
193 firing rates) or negative (high MBON firing rates; see Discussion). Following non-associative learning
194 and extensive exposure to natural images, the entire network was trained and tested on various
195 pattern recognition tasks commonly used in experimental studies (Benard et al., 2006; Dyer et al.,
196 2005; Srinivasan, 2010, 1994; Zhang and Horridge, 1992), including the discrimination of "plus" and
197 "multiplication sign" patterns, as previously examined in real bumblebees (MaBouDi et al., 2021b)
198 (Figure 4A).

199 To evaluate the performance of the active vision model, we analysed MBON activity, which functions
200 as a decision-making unit. A lower MBON response from its baseline activity to a particular pattern
201 indicates preference, whereas a higher response suggests rejection. After multiple training trials
202 through a novel associative learning (see Methods and Discussion), the MBON exhibited a distinct
203 response pattern, with reduced activity towards the chosen visual stimulus and increased activity
204 towards the rejected one. This suggests that the MBON plays a role akin to a decision neuron in
205 pattern recognition tasks. Importantly, no reinforcement learning, or synaptic updates were applied
206 during the testing phase, ensuring that the observed responses reflected the network's learned
207 capacity for visual discrimination rather than online adaptation.

208
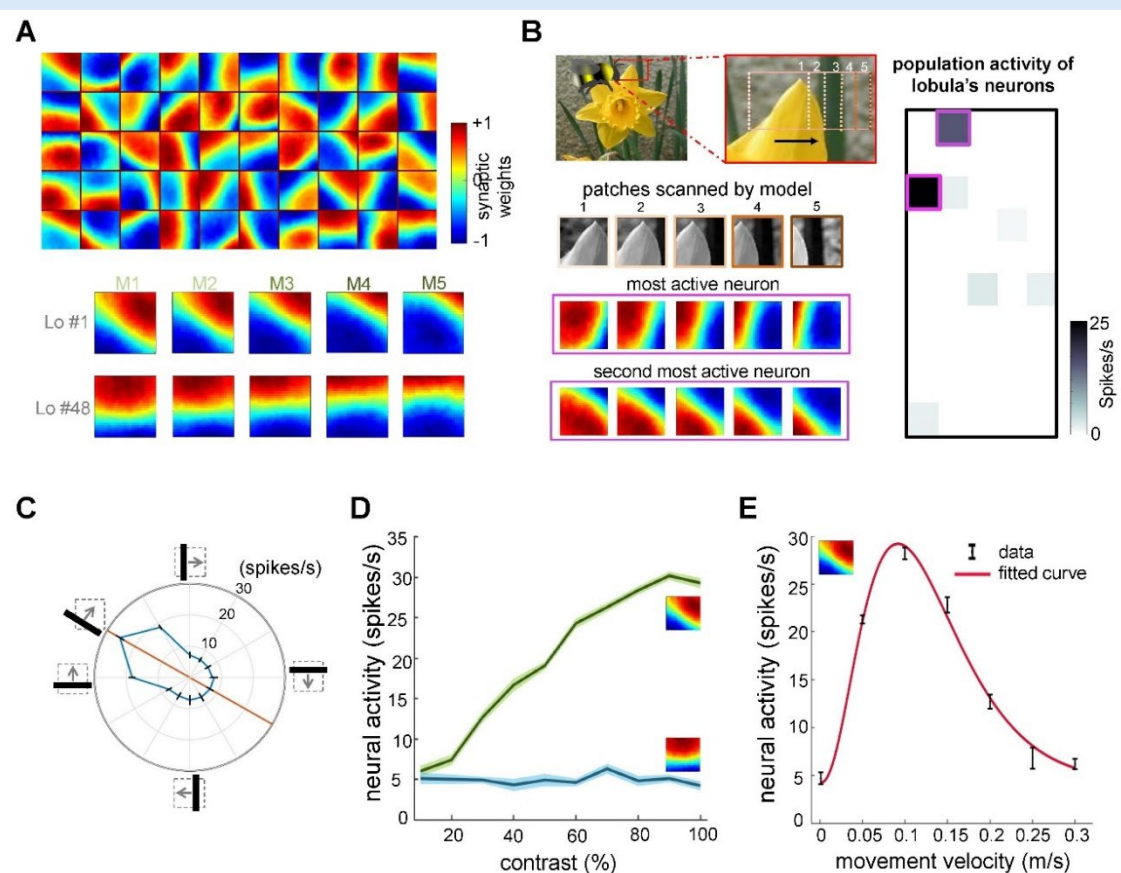
Figure 2. Neural responses of the simulated bee model to visual patterns. (A) Top: each square in the matrix corresponds to a single time slice of the obtained spatiotemporal receptive field of a lobula neuron (5x10 lobula neurons) that emerged from non-associative learning in the visual lobes after exposing the model to images of flowers and nature scenes (see Video 3). Bottom: spatiotemporal receptive field of two example lobula neurons are visualised in the five-time delay slices of the matrices of synaptic connectivity between lamina and five medulla neurons (See Figure 1D). The lobula neuron integrates signals from these medulla neurons at each of five time periods as the simulated bees scan a pattern (time goes from left to right). Blue and red cells show inhibitory and excitatory synaptic connectivity, respectively. The first example lobula neuron (#1) encodes the 150° angled bar moving from lower left to the upper right of the visual field. The second example lobula neuron (#48) encodes the movement of the horizontal bar moving up in the visual field. **(B)** An example of an image sequence projected to the simulated bee' eye with lateral movement from left to right. Below shows the five images patched sampled by the simulated bee. The right side presents the firing rate of all lobula neurons responding to the image sequence. The spatiotemporal receptive field of two highest active neurons to the image sequence are highlighted in purple. **(C)** The polar plot shows the average orientation selectivity of one example lobula neuron (#1) to differently angled bars moving across the visual field in a direction orthogonal to their axis (average of 50 simulations). This neuron is most sensitive to movement when the bar orientation is at 150°. **(D)** The spiking response of the lobula neuron to the preferred orientation raised as the contrast was increased, whereas the response of the lobula neuron to a non-preferred orientation is maintained irrespective of contrast. **(E)** The average velocity-sensitivity curve (±SEM) of the orientation-sensitive lobula neuron (#1) is obtained from the responses of the lobula neuron to optimal (angle of maximum sensitivity) moving stimuli presented to the model at different velocities. The red line shows the Gamma function fitted to the data.

209

Video 1: Spatiotemporal dynamics of receptive fields in 50 lobula neurons emerging from non-associative learning and active scanning. Each square in the matrix represents a single time slice of the spatiotemporal receptive field for a lobula neuron (5 × 10 array of neurons). These receptive fields illustrate the connectivity matrix between five medulla neurons and their corresponding lobula neuron, operating under a temporal coding structure. In this framework, each of the five medulla neurons sequentially transfers a portion of the visual input to the lobula neuron through excitatory (red) and inhibitory (blue) synaptic connections. These receptive fields develop within the visual lobes after the model is exposed to natural images, including flowers and scenery. As the simulated bee scans a visual pattern, lobula neurons dynamically integrate inputs from medulla neurons over time, forming a temporally structured neural representation of the visual scene.

210

211 ***Non-associative learning shapes spatiotemporal coding in the lobula to align with the***
212 ***statistical features of natural scenes.***

213 The synaptic weights in the optic lobe were updated through exposure to natural images during the
214 model's lateral scanning process (see Methods, Figure 1B). While lamina-to-medulla connections are
215 structured based on temporal coding (Figure 1D), lobula neurons are configured to laterally inhibit
216 each other, facilitating competitive interactions. Synaptic connections were updated using Oja's
217 implementation of Hebb's rule (Oja, 1982). Simultaneously, a symmetric inhibitory spike-timing-
218 dependent plasticity (iSTDP) rule was applied to lateral inhibitory connections among lobula neurons
219 (Vogels et al., 2011). These local synaptic plasticity rules, which govern interactions between lamina,
220 medulla, and lobula neurons, support non-associative learning—i.e., synaptic modifications occurring
221 in the absence of reward (see Methods). Together, these plasticity mechanisms drive the network
222 toward an efficient representation of visual input, reducing redundancy while preserving essential
223 visual information.

224 Figure 2A illustrates the receptive fields of lobula neurons, which exhibit spatiotemporal orientation
225 selectivity after training on 100 flower and natural images (comprising 50,000 time-varying image
226 patches). Each square in the figure represents one of the 50 lobula neurons, with the heat map
227 indicating the synaptic weights of the corresponding lamina neurons (connected via the medulla
228 neurons). To aid interpretation, the lower portion of Figure 2A provides examples of two individual
229 lobula neurons, detailing their lamina synaptic weights for each of the five medulla neurons. For a
230 more dynamic representation of these receptive fields over time, see Video 1. The receptive fields of
231 lobula neurons are characterised by an elongated "on" area (regions representing positive synaptic
232 weights) adjacent to an antagonistic "off" area (regions with negative synaptic weights). These regions
233 are generally aligned along a specific orientation, and their balance changes dynamically across time-
234 delayed instances of medulla responses. For instance, one lobula neuron responds most strongly to
235 a 135° bar moving orthogonally to its "on" or "off" areas while exhibiting little or no response to other
236 orientations (Figure 2C). The population of 50 lobula neurons demonstrates specificity to both
237 orientation and direction, closely resembling neuronal responses observed in bees and other insects

238 (James and Osorio, 1996; Paulk et al., 2008; Seelig and Jayaraman, 2013; Yang and Maddess,
239 1997).

Figure 2- figure supplement 1: (A) Spatiotemporal receptive fields of lobula neurons emerging from non-associative learning when the number of lobula neurons is set to 100 (see Video 2). The details follow those described in Figure 2A. **(B)** Spatiotemporal receptive fields of 50 lobula neurons trained with shuffled natural images. The neurons fail to develop meaningful connections, resulting in random synaptic weight distributions, indicating that spatial coherence in training images is essential for efficient feature extraction.

240

Video 2: Spatiotemporal dynamics of receptive fields in 100 lobula neurons emerging from non-associative learning and active scanning. This follows the same structure as Video 1, depicting the receptive field evolution in a larger population of 100 lobula neurons under the temporal coding framework and non-associative learning.

241

242 To illustrate how lobula neurons process natural visual inputs, Figure 2B depicts the sequence of
243 image patches scanned during a simulated horizontal movement. The results show that only a small
244 subset of lobula neurons respond at any given moment, indicating that their activity is decorrelated
245 and relatively selective—an outcome of non-associative learning mechanisms in the visual lobe (see
246 Discussion). Notably, the two most active lobula neurons captured distinct structural features of the
247 flower petal: one neuron's receptive field aligned with the left 45° edge of the petal, while another
248 matched the right-angled edge. This suggests that the model effectively extracts distinct visual
249 features with a minimal number of filters (lobula neurons).
250 To examine the selectivity of lobula neurons further, we analysed the spiking activity of a
251 representative neuron tuned to a 150° orientation. As expected, the neuron showed maximal firing

252 (26 spikes/sec) when presented with a 150° moving bar. It also exhibited moderate responses to a

253 horizontal bar and a 120° moving bar (18 spikes/sec) but remained largely unresponsive to other

254 orientations (Figure 2C). Consistent with experimental findings (Yang and Maddess, 1997), the firing

255 rate of lobula neurons increased with contrast at their preferred orientations, whereas responses to

256 non-preferred orientations remained unchanged across contrast levels (Figure 2D).

257 Finally, Figure 2E highlights the velocity sensitivity of lobula neurons. Each neuron responds

258 maximally at a specific velocity, demonstrating a tuning curve that aligns with known insect neural

259 responses (Paulk et al., 2008; Yang and Maddess, 1997). This reinforces that our model successfully

260 captures key quantitative properties of lobula edge detector neurons, including their joint selectivity

261 for orientation, contrast, and motion velocity.

262 The model demonstrates robustness in generating spatiotemporal receptive fields, even as the

263 number of lobula neurons increases. Training the non-associative learning model with a larger lobula

264 neuronal population while maintaining the same underlying structure from photoreceptors to the

265 medulla, enhances the diversity of orientation-selective responses (see Figure 2 -Supplementary

266 figure 1A and Video 2). As the number of lobula neurons increases, their tuning properties become

267 more distributed, enabling a finer and more precise encoding of different orientations and motion

268 patterns. This scalability highlights the model's ability to generalise its representation of natural scene

269 statistics while achieving varying levels of resolution in visual encoding.

270

271 ***Lobula neuron responses become sparse and decorrelated through non-associative learning***

272 ***with natural images***

273 To assess the impact of training on the population activity of lobula neurons, we quantified their

274 response sparsity and decorrelation before and after learning. Figure 3A presents the correlation

275 matrix of lobula neurons in response to 10,000 sequential scans of natural images after training. The

276 results reveal a highly decorrelated response pattern, with a strong diagonal structure indicating that

277 each lobula neuron maintains a distinct response profile. This demonstrates that non-associative

278 learning enhances both the selectivity and independence of neural representations, allowing the

279 network to develop more efficient and diverse feature encoding.

280 Further supporting this, Figure 3B displays the sparseness index of lobula neurons under different

281 training conditions. Before training, lobula neuron activity was broadly distributed, as reflected in the

282 high sparseness index (see Methods). Training on natural images significantly reduced the

283 sparseness index, indicating that lobula neurons developed a more selective and efficient coding

284 scheme, where only a small subset of neurons responded to any given input. In contrast, training on

285 shuffled natural images led to only a moderate reduction in sparseness, suggesting that the

286 implemented local synaptic plasticity rules alone are insufficient for optimal feature encoding without

287 exposure to structured natural inputs. Notably, while shuffling natural images preserves pixel intensity

288 and overall distribution, it disrupts spatial correlations and higher-order structures present in natural

11

289 scenes. Consequently, training on these shuffled datasets results in non-structured receptive fields,
290 leading to broadly distributed and less selective coding in the lobula neurons (Figure 2-
291 Supplementary figure 1B).

292

Figure 3. Effect of non-associative learning on lobula neuron activity and response sparseness. (A) Correlation matrix of lobula neuron responses after training with natural images. The near-diagonal structure indicates that neurons develop distinct and strongly uncorrelated responses, suggesting an efficient, decorrelated representation of visual input. (B) Sparseness index of lobula neurons before and after training with different image sets. Before training, neural responses are broadly distributed. Training with shuffled natural images does not change the sparseness of lobula population, whereas training with natural images significantly increases response sparseness, indicating that exposure to structured visual inputs enhances efficient coding. Error bars represent SEM. Asterisks (*) indicate p-values < 0.05, while 'n.s.' denotes non-significant results.

293

294 These findings underscore the role of non-associative learning in shaping neural representations,
295 fostering both sparsity and decorrelation in lobula neurons. Such sparse coding is crucial for efficient
296 sensory processing, as it minimises redundancy while preserving essential visual information.
297 Moreover, it optimises metabolic efficiency in neural networks, aligning with coding strategies
298 observed in biological visual systems (see Discussions).

299

300 ***Active vision enhances visual discrimination through sequential scanning***
301 To replicate bee behavioural findings reported in the literature (Benard et al., 2006; Dyer et al., 2005;
302 MaBouDi et al., 2021b; Srinivasan, 1994, 2010; Zhang and Horridge, 1992), we implemented
303 computational plasticity within the mushroom body circuitry, leveraging sparse lobula neurons that
304 emerge through non-associative learning—critical for encoding both appetitive and aversive values.
305 Specifically, we incorporated classical spike-timing-dependent plasticity (STDP), modulated by
306 dopamine, to regulate synaptic modifications between mushroom body Kenyon cells (KCs) and
307 extrinsic MBONs in response to negative (unrewarded) patterns. Additionally, we introduced a novel

308 STDP-based plasticity rule, modulated by octopamine (See Methods, Figure 9), which we hypothesise
309 induces synaptic depression among KC-MBON connections in response to positive (rewarded)
310 patterns. These plasticity mechanisms allowed us to explore synaptic dynamics underlying the
311 discrimination of rewarded and non-rewarded stimuli (see Discussion).
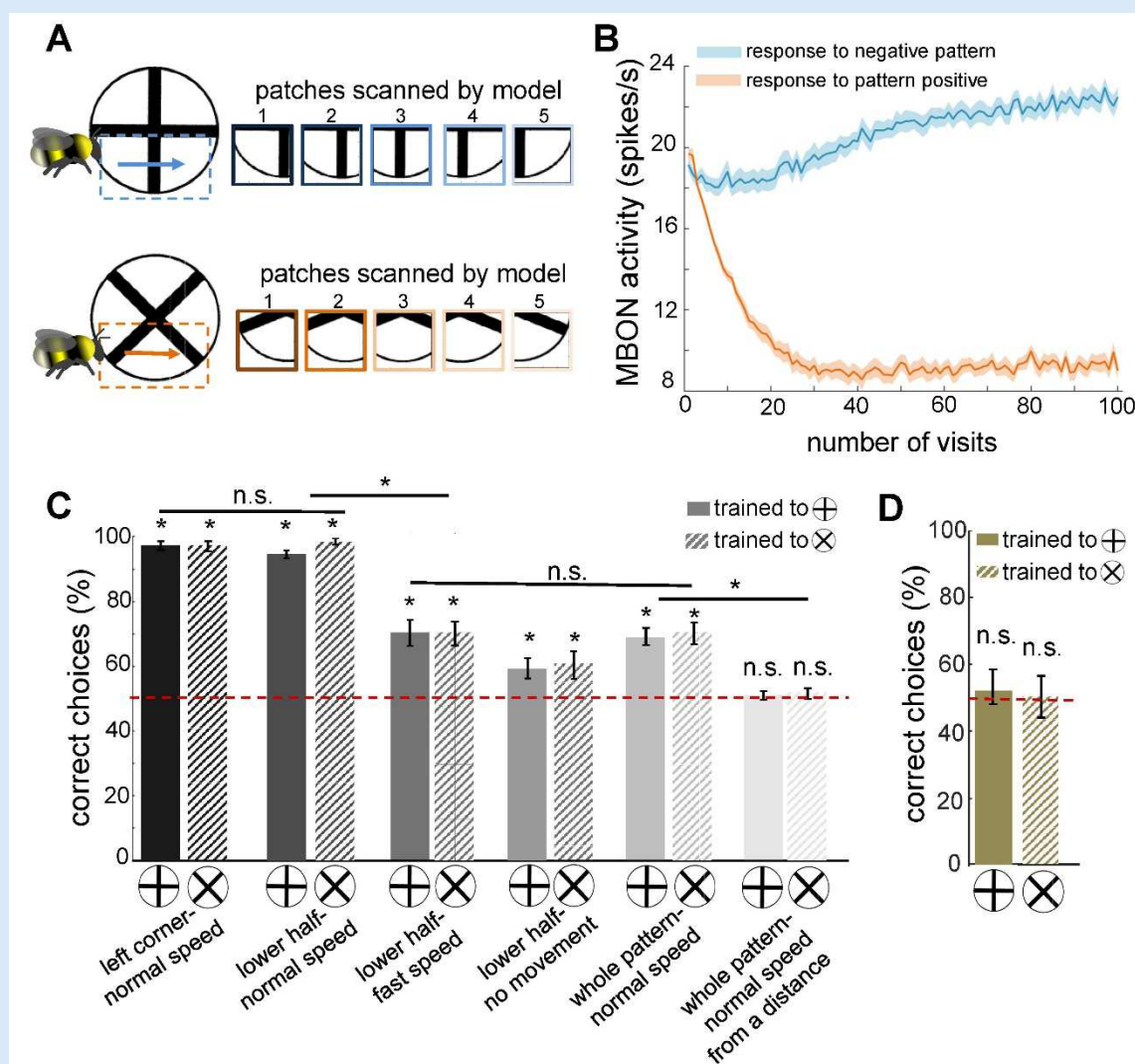
Figure 4. Simulated bees' performance in a pattern recognition task using different scanning strategies. Twenty simulated bees, with random initial neuronal connectivity in mushroom bodies (see Methods) and a fixed connectivity in the visual lobe that were shaped from the non-associative learning, were trained to discriminate plus from a multiplication symbol (100 random training exposures per pattern). The simulated bees scanned different regions of the patterns at different speeds. **(A)** Top and below panels show the five image patches sampled from the plus and multiplication symbols by simulated bees, respectively. It is assumed that the simulated bees scanned the lower half of the patterns with lateral movement from left to right with normal speed (0.1 m/s). **(B)** The plot shows the average responses of the MBON to rewarding multiplication and punishing plus patterns during training procedure (multiplication symbol rewarding, producing an Octopamine release by the reinforcement neuron, and the plus symbol inducing a Dopamine release). This shows how the response of the MBON to the rewarding plus was decreased while its response to the punishing multiplication pattern was increased during the training. The MBON equally responded to both multiplication and plus before the training (at number of visits =0). **(C)** The performance of the simulated bees in discriminating the right-angled plus and a 45° rotated version of the same cross (i.e. multiplication

symbol) (MaBouDi et al., 2021b; Srinivasan, 1994), when the stimulated bees scanned different regions of the pattern (left corner, lower half, whole pattern) at different speeds: no speed 0.0m/s (i.e. all medulla to lobula temporal slices observed the same visual input), normal speed at 0.1 m/s and fast speed at 0.3m/s), and from a simulated distance of 2cm from stimuli (default) and 10cm (distal view). The optimal model parameters were for the stimulated bees at the default distance when only a local region of the pattern (bottom half or lower left quadrant) was scanned at a normal speed. **(D)** Mean performance (±SEM) of two groups of simulated bees in discriminating the plus from multiplication patterns when their inhibitory connectivity between lobula neurons were not modified by non-associative learning rules. Asterisks (*) indicate p-values < 0.05, while 'n.s.' denotes non-significant results.

312

313   The model was trained using a differential conditioning paradigm, where the correct stimulus (S⁺) was

314   paired with a reward and the incorrect stimulus (S⁻) was associated with punishment. For simplicity,

315   we denote positive patterns as S⁺ and negative patterns as S⁻. During associative learning

316   simulations, only the synaptic weights between KCs and MBONs corresponding to the presented

317   patterns were updated, while randomly weighted connections between the lobula and KCs were

318   incorporated to ensure sparse activity in KCs (Figure 1C; see Methods). To capture individual

319   variability observed in bees, we repeated the simulations with different initial conditions, including

320   random neural connectivity between lobula neurons and Kenyon cells, as well as between KCs and

321   MBONs. The model's performance was then assessed across multiple visual discrimination

322   paradigms (Figures 4, 6 and 7).

323   In the initial implementation, simulated bees were trained to distinguish a plus sign from a

324   multiplication sign, with training focused on the lower half of the plus pattern (Figure 4A). Following

325   training, MBON activity decreased in response to the plus (S⁺) while increasing in response to the

326   multiplication sign (S⁻), whereas prior to training, MBON responses to both patterns were similar

327   (Figure 4B). This demonstrates that the model successfully discriminated between S⁺ and S⁻ through

328   temporal coding and sequential scanning of the visual pattern (Figure 4C). In contrast, a model with

329   fixed, random connectivity in the visual lobe failed to differentiate between the plus and multiplication

330   patterns (Figure 4D). This underscores the importance of structured connectivity that emerges in the

331   bee visual lobes through non-associative learning—specifically, the development of spatiotemporal

332   receptive fields in lobula neurons—for successful visual learning.

333   Our model further revealed that rewarding patterns elicited a reduction in extrinsic neuron responses,

334   while punished patterns led to increased responses (Figure 4B), a phenomenon consistent with neural

335   activity recorded from alpha lobe PE1 neurons in the mushroom body (Okada et al., 2007). Initially,

336   the simulated bees performed worse than real bees in the plus versus multiplication discrimination

337   task (Figure 4C, last pair of bars). Using experimentally derived parameters from bumblebee

338   studies—including an average scanning speed of 0.1 m/s (referred to as normal speed) for whole-

339   pattern scanning and a viewing distance of 20 mm—simulated bees achieved a correct choice rate

340   of only 63% for the plus stimulus and 60% for the reciprocal cross-protocol (averaged over 20

341   simulations; see Figure 4C, fifth pair of bars).

342   However, when the experimental conditions were adjusted so that simulated bees scanned only the
343   lower half of the patterns or focused on the lower left corner—consistent with real bee behaviour
344   (MaBouDi et al., 2021)—correct choice performance improved significantly, reaching ≥96% and
345   ≥98%, respectively (Figure 4C, first two pairs of bars). Conversely, increasing scanning speed
346   (resulting in larger separations between sampled image patches) reduced accuracy to 70%, while
347   stationary simulated bees—those that did not actively scan—achieved only 60% correct choices.
348   In additional experiments where the model bees were trained from a greater distance (>100 mm),
349   they failed to discriminate between patterns. This aligns with behavioural findings in bees, where they
350   initially select a pattern at random and move closer to scan it before making a decision (MaBouDi et
351   al., 2021). Studies show that bumblebees approach stimuli to scan patterns at close range, and their
352   initial approach choices are random (Guiraud et al., 2018; MaBouDi et al., 2021b). These findings
353   underscore the critical role of active vision in visual learning and discrimination, demonstrating that
354   the model effectively captures key aspects of biological visual processing, including sequential
355   sampling, spatial integration, and plasticity-driven adaptation.
356
357   ***Scanning strategy modulates lobula neural representations: Enhanced selectivity with***
358   ***localised sampling***
359   To further investigate how scanning behaviour influences the distinctiveness of neural representations
360   in the lobula and impacts performance in visual learning tasks, we analysed activity patterns and
361   response magnitudes under two different scanning conditions from the first experiment (Figure 4C).
362   Heatmaps of lobula neuron activity (Figure 5A, B) illustrate neural responses to two stimulus
363   conditions: scanning only the lower half of the pattern at normal speed (Panel A) and scanning the
364   entire pattern at normal speed (Panel B).
365   The activity maps indicate that different subsets of lobula neurons responded preferentially to the
366   lower half of the stimulus, exhibiting stronger activation in a smaller neuronal subset. In contrast,
367   scanning the entire pattern resulted in a more widespread and overlapping activation across the lobula
368   neuron population. Corresponding spike rate plots (Figure 5, right panels) further highlight these
369   differences, revealing that neurons displayed stronger selectivity when scanning was restricted to the
370   lower half at normal speed compared to whole-pattern scanning. These findings suggest that lobula
371   neurons exhibit stimulus-specific selectivity that is modulated by the spatial dynamics of scanning
372   behaviour.
373   To quantify the separability of these population responses, we computed the angular distance (θ)
374   between their activity vectors over 20 stimulations using the cosine similarity metric (see Methods).
375   The results (Figure 5C) indicate that neural responses exhibited a significantly larger angular
376   distance when the model scanned only the lower half of the stimuli at normal speed compared to
377   scanning the entire of the same pattern. This suggests that neural activity population were more

378   distinct when scanning was confined to a localised region of these visual patterns, reinforcing the

379   idea that restricted sampling enhances neural discriminability.
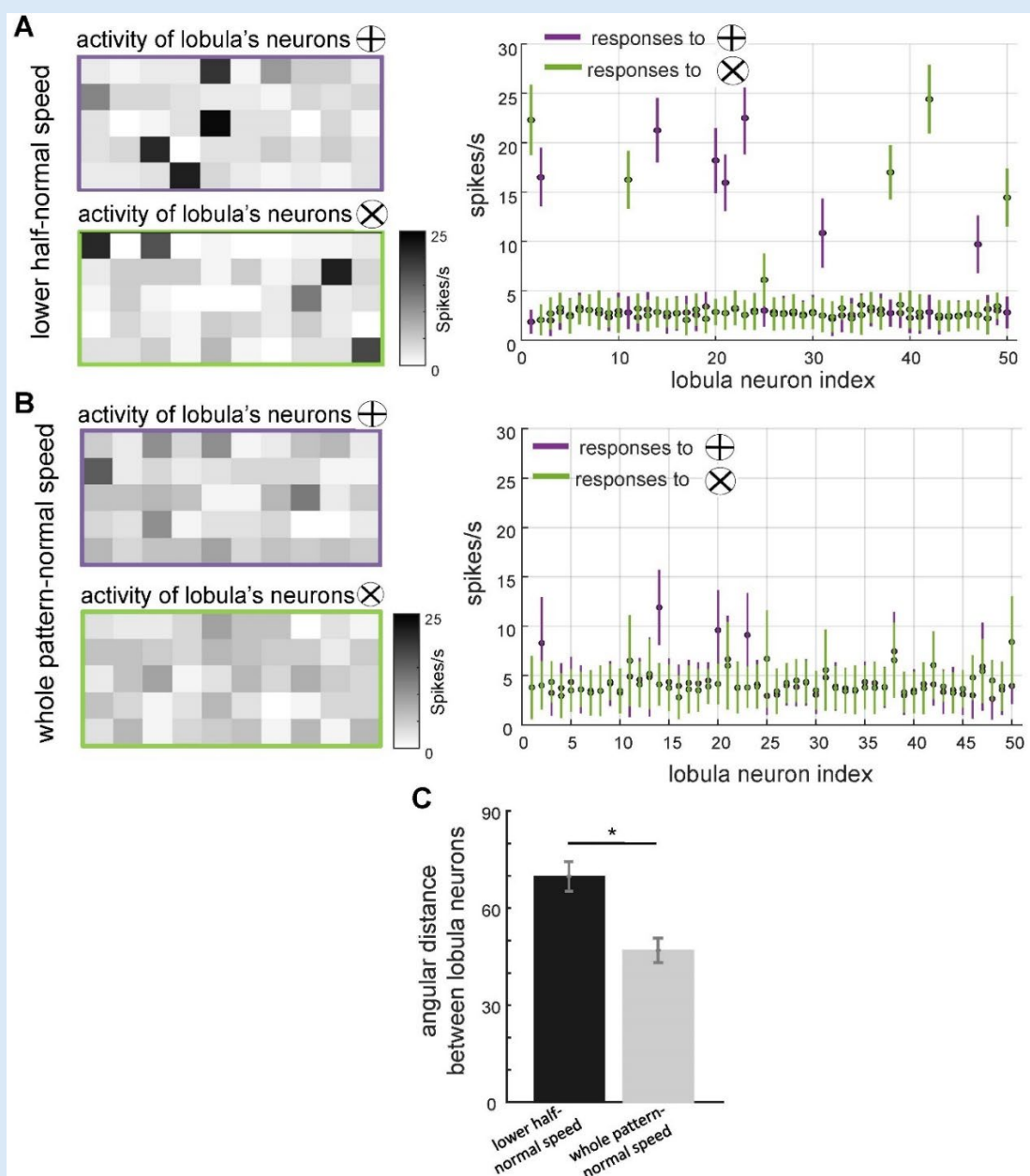
Figure 5. The effect of scanning behaviour on the spatiotemporal encoding of visual patterns in lobula neurons responses. (A, B) Neural responses of simulated lobula neurons to different scanning conditions. Left panels: heatmaps showing the spiking activity of 50 lobula neurons in response to visual patterns when scanning either the lower half of the pattern at normal speed **(A)** or the whole pattern at normal speed **(B)**. Right panels: the mean and standard deviation of the spike rate responses of individual lobula neurons to two distinct visual patterns (plus and multiplication), with colour-coded responses (purple for ⊕, green for ⊗). Scanning behaviour significantly alters the neural responses, with distinct sets of neurons preferentially responding to each stimulus. **(C)** Mean angular distance between lobula neuron responses for different scanning conditions. Lower half-normal speed scanning results in greater separation between neural representations, suggesting that scanning of local region enhances feature selectivity. Error bars represent SEM. Asterisks (*) indicate p-values < 0.05, while 'n.s.' denotes non-significant results.

380

381 These findings demonstrate that lobula neurons encode visual stimuli in a structured manner, with

382 response contrast influenced by both spatial and temporal properties of scanning behaviour

383 (Figures 4, 5). By selectively sampling specific regions of a stimulus, the system enhances the

384 differentiation of these visual patterns, supporting the hypothesis that active vision plays a crucial

385 role in effective neural coding and discrimination.

386

387 ***Neural network model of active vision bee behaviours across various visual experiments***

388 In this study, we evaluated our model - using scans from the lower half of the visual field - by

389 comparing its performance with results from bee experiments reported in the literature. (Note: bees

390 may exhibit variations in scanning behaviour under different patterns and training conditions; see

391 Discussion). Our simulated bees successfully discriminated between angled bars (Hateren et al.,

392 1990), a 22.5° angled cross from a 90° rotated version (Srinivasan, 1994), and spiral patterns (Zhang

393 and Horridge, 1992) (Figure 6A). When trained on grating patterns with -45° versus +45° orientations,

394 the simulated bees successfully identified the correct pattern. Moreover, it demonstrated the ability to

395 transfer the learned rule to novel patterns the model had never encountered during training, including

396 single-bar patterns (Figure 6B). This suggests that the model captures key aspects of visual

397 generalisation observed in real bees. Figure 6C shows that the proposed model not only learned to

398 identify the correctly oriented bar pattern but also to distinguish the rewarding pattern from a novel

399 one (two circles). Notably, the model exhibited a 22% lower preference for the negatively trained

400 pattern compared to the novel pattern, validating the implementation of the rejection behaviour and

401 demonstrating that the model can simultaneously learn rewarding and aversive stimuli.

402 This was further explored by training the network with patterns containing two oriented bars in each

403 lower quadrant (Figure 6D) (Benard et al., 2006; Stach et al., 2004; Zhang and Horridge, 1992). The

404 simulated bees discriminated these training patterns with over 99% accuracy, but performance

405 dropped to an average of 61% when presented with a simplified variant. When tested with the original

406 positive pattern and novel patterns containing only one correct orientation, the bees showed a high

407 preference for the correct stimulus. Similarly, the simulated bees exhibited a clear preference for a

408 pattern with a single correct feature over the trained negative pattern, indicating that the model can

409 extract multiple features during scanning.

410

Figure 6. Proposed neural network performance to published bee pattern experiments. Twenty simulated bees, with random initial neuronal connectivity in mushroom bodies (see Methods), were trained to discriminate a positive target pattern from a negative distractor pattern (50 training exposures per pattern). The simulated bees' performances were examined via unrewarded tests, where synaptic weights were not updated (average of 20 simulated pattern pair tests per bee). All simulations were conducted under the assumption that model bees viewed the targets from a distance of 2 cm while flying at a normal speed of 0.1 m/s. During this process, the bees scanned the lower half of the pattern. **(A)** Mean percentage of correct choices (±SEM) in discriminating bars oriented at 90° to each other, 25.5° angled cross with a 45° rotated version of the same cross, and a pair of mirrored spiral patterns (MaBouDi et al., 2021b; Srinivasan, 1994). The simulated bees achieved greater than chance performances. **(B)** Performance of

simulated bees trained with a generalisation protocol (Benard et al., 2006). Trained to 6 pairs of perpendicular oriented gratings (10 exposures per grating). Simulated bees then tested with a novel gating pair, and a single oriented bar pair. The simulated bees performed well in distinguishing between the novel pair of gratings; less well, but still significantly above chance, to the single bars. This indicates that the model can generalise the orientation of the training patterns to distinguish the novel patterns. **(C)** Mean performance (±SEM) of the simulated bees in discriminating the positive orientation from negative orientation. Additionally, the performance in recognising the positive orientation from the novel pattern, and preference for the negative pattern from a novel pattern. Simulated bees learnt to prefer positive patterns, but also reject negative patterns, in this case preferring novel stimuli. **(D)** Performance of simulated bees trained to a horizontal and -45° bar in the lower pattern half versus a vertical and +45° bar (Stach et al., 2004). The simulated bees could easily discriminate between the trained bars, and a colour inverted version of the patterns. They performed less well when the bars were replaced with similarly oriented gratings, but still significantly above chance. When tested on the positive pattern vs. a novel pattern with one correctly and one incorrectly oriented bar, the simulated bees chose the positive patterns (fourth and fifth bars), whereas with the negative pattern versus this same novel pattern the simulated bees rejected the negative pattern in preference for the novel pattern with single positive oriented bar (two last bars). **(E)** The graph shows the mean percentage of correct choices for the 20 simulated bees during a facial recognition task (Dyer et al., 2005). Simulated bees were trained to the positive (rewarded) face image versus a negative (non-rewarded) distractor face. The model bee is able to recognise the target face from distractors after training, and also to recognise the positive face from novel faces even if the novel face is similar to the target face (fourth bar). However, it failed to discriminate between the positive and negative faces rotated by 180°. **(F)** The model was trained on spatially structured patterns from Stach et al. (2004), requiring recognition of orientation arrangements across four quadrants. Unlike bees, the model failed to discriminate these patterns, highlighting its limitations in integrating local features into a coherent global representation. Asterisks (*) indicate p-values < 0.05, while 'n.s.' denotes non-significant results.

411

412 To present a more complex pattern recognition task, we replicated a facial recognition experiment

413 performed on honeybees (Dyer et al., 2005) by training the neural network with images of two human

414 faces (Figure 6E). As with honeybees, our simulated bees were able to identify the positive trained

415 face from the negative one, as well as distinguish two novel faces and a caricature. Both the real and

416 simulated bees failed to discriminate the faces when rotated through 180°. These results demonstrate

417 that complex visual features can be condensed through spatiotemporal encoding in the lobula

418 neurons into specific and distinct neuronal representations that are critical for learning in the miniature

419 bee brain.

420 To evaluate the model's ability to discriminate more complex visual stimuli, we trained it on a set of

421 patterns from (Stach et al., 2004), which require recognising the spatial arrangement of orientations

422 across four quadrants (Figure 6F). Unlike previous experiments where scanning behaviour was

423 confined to specific regions, the complexity of these stimuli necessitated training the model using

424 whole-pattern exposure. The results indicate that the model failed to replicate the bees' ability to

425 discriminate these patterns (Figure 6F), despite successfully distinguishing between the plus and

426 multiplication signs when scanning the entire pattern (Figure 4C), even if its performance remained

427 lower than that of real bees. Moreover, the model was unable to generalise this capability to more

428 configurations testing patterns, suggesting that while it can process simple spatial features through

429   sequential scanning, it lacks the longer and more dynamic scanning required for assembling and

430   integrating local features into a coherent global representation, as observed in bees (see Discussion).

431

432   ***What is the minimally sufficient number of lobula neurons and the necessary connectivity for***

433   ***active vision in bees?***

434   As reported above, our model successfully accomplishes various pattern recognition tasks (Figures

435   4, 6). We then asked whether our neural networks could perform with a very limited number of lobula

436   neurons that transfer visual information to the mushroom body. To investigate this, we ran the non-

437   associative learning process with different numbers of lobula neurons – specifically, 4, 16, or 36

438   neurons (the original model had 50). The visual network was subsequently trained using the same

439   set of natural images and protocol as the original model (Figure 7A).
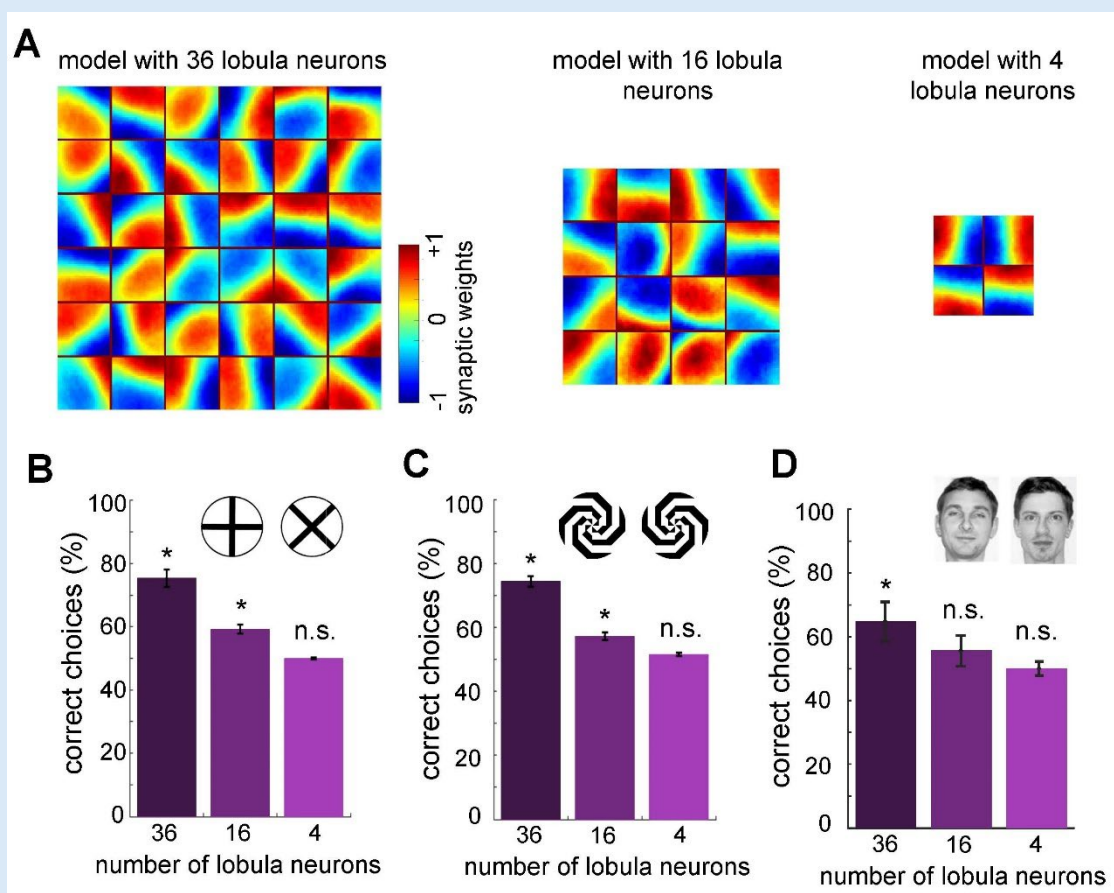
Figure 7. Minimum number of lobula neurons that are necessary for pattern recognition.
(A) Obtained spatiotemporal receptive field of lobula neurons when the number of lobula neurons were set at 36, 16 or 4 during the non-associative learning in the visual lobe (See Figure 5A). This shows the models with lower number of lobula neurons encode less variability of orientations and temporal coding of the visual inputs (see Videos 5,6) **(B, C & D)** The average correct choices of the three models with 36, 16 or 4 lobula neurons after training to a pair of plus and multiplication patterns (B) mirrored spiral patterns (C) and human face discrimination (D). The model with 36 lobula neurons still can solve pattern recognition tasks at a level above chance. It indicates that only 36 lobula neurons that provide all inputs to mushroom bodies are sufficient for the simulated bees to be able to discriminate between patterns. Asterisks (*) indicate p-values < 0.05, while 'n.s.' denotes non-significant results.

20

440

Video 3: Spatiotemporal dynamics of receptive fields in 36 lobula neurons emerging from non-associative learning and active scanning. This follows the same structure as Video 1, depicting the receptive field evolution in a larger population of 36 lobula neurons under the temporal coding framework and non-associative learning.

441

Video 4: Spatiotemporal dynamics of receptive fields in 16 lobula neurons emerging from non-associative learning and active scanning. This follows the same structure as Video 1, depicting the receptive field evolution in a larger population of 16 lobula neurons under the temporal coding framework and non-associative learning (compare to Videos 1, 3).

442

Video 5: Spatiotemporal dynamics of receptive fields in only 4 lobula neurons emerging from non-associative learning and active scanning. This follows the same structure as Video 1, depicting the receptive field evolution in a larger population of 4 lobula neurons under the temporal coding framework and non-associative learning (compare to Videos 1, 3 & 4).

443

444 Interestingly, the non-associative learning process led to the emergence of distinct spatiotemporal
445 structures in the lobula neurons. We found that reducing the number of lobula neurons decreased the
446 variability in their spatiotemporal receptive (Figure 7A and Videos 3, 4 ,5). In particular, when the
447 network was limited to four neurons, it could not encode the full spatiotemporal structure of the training
448 patterns obtained with the model with 50 lobula neurons - only vertical and horizontal receptive fields
449 were produced (Video 1, 5). As expected, the overall performance of the model decreased as the
450 number of lobula neurons was reduced. Although the model with 16 lobula neurons demonstrated the
451 ability to discriminate more complex patterns beyond the plus and multiplication signs (Figure 7B, C),
452 it remained insufficient for recognising highly complex stimuli such as human faces (Figure 7D, 7E).

453

454 Moreover, to investigate the effect of inhibitory neurons within the visual lobe on lobula neuron output,
455 we trained the model using the same protocol but fixed the synaptic weights of the inhibitory
456 connections (i.e., these weights were not updated during exposure to the training images). These
457 fixed inhibitory connections limited the ability of the lobula neuron population to encode moving
458 orientations (Figure 8 and Video 6), indicating that the plasticity of inhibitory interneurons in the visual
459 lobe plays a crucial role in facilitating an efficient representation of the visual environment. While this
460 suggests that increasing the network size can enhance the model's capacity for discriminating
461 complex visual patterns, the results (Figure 4) indicate that scanning behaviour plays a crucial role in
462 overcoming this limitation. Specifically, adopting a more targeted scanning strategy can improve
463 discriminability by directing visual sampling to the most informative regions of the stimulus (see
464 Discussion).

465

466 Taken together, these findings demonstrate that our assumption regarding non-associative plasticity
467 in the visual lobe successfully replicates the neural responses of lobula neurons across various

468 patterns and conditions. This plasticity yields a sparse, uncorrelated representation of the visual input

469 that benefits subsequent learning processes in the mushroom body. Importantly, these results closely

470 align with theoretical studies (see Discussion), further supporting the effectiveness of the active vision

471 in capturing the underlying principles of information encoding in the insect visual system.

472

473
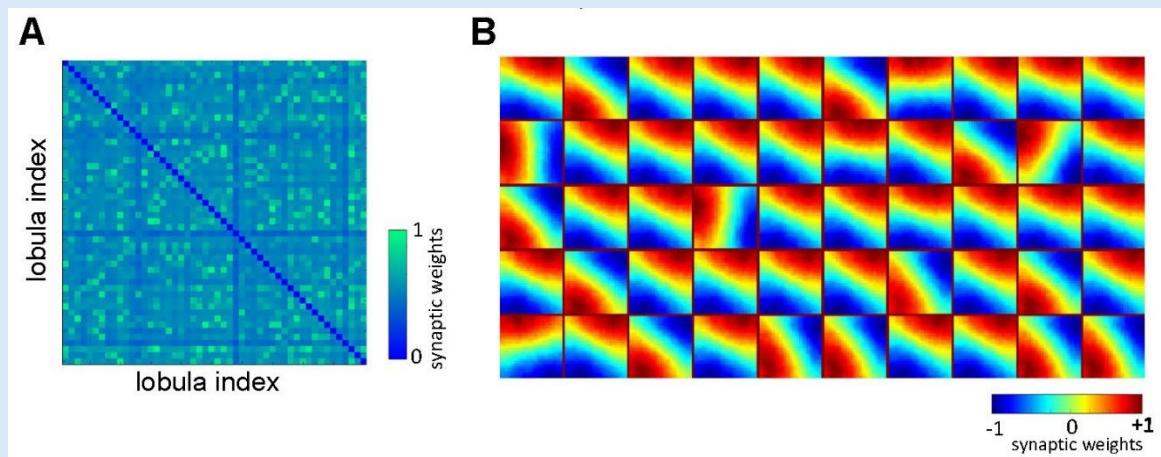
Figure 8. The role of lateral inhibitory connections between lobula neurons. Obtained spatiotemporal receptive field of lobula neurons (B) when the lateral inhibitory connectivity between lobula neurons is fixed (A) during the non-associative learning (see Video 6 and Figure 2A & Video 1).

474

Video 6: Spatiotemporal dynamics of receptive fields in 50 lobula neurons emerging from non-associative learning and active scanning. This follows the same structure as Video 1, but with fixed lateral inhibitory connectivity between lobula neurons during non-associative learning. The video illustrates how receptive fields evolve under the temporal coding framework, providing a comparison to Video 1, where lateral inhibition was plastic.

475

476 **Discussion**

477 In this study, we investigated the core computational requirements for visual pattern recognition by

478 examined a minimal neural network Inspired by the active scanning flight behaviours of bees

479 (MaBouDi et al., 2021b). We developed a novel model based on the insect visual system, simulating

480 how a small population of lobula neurons encodes the visual environment through spatiotemporal

481 responses. By incorporating non-associative learning, the model self-organises its connectivity within

482 the visual lobe, generating efficient environmental representations (Figures 2, 3, 5). The process leads

483 to the emergence of orientation-selective cells in the lobula, which are essential for encoding complex

484 visual scenes (Figure 2).

485 Our simulations reveal that a small subset of lobula neurons, sensitive to specific orientations and

486 velocities, can condense complex visual environments into spatiotemporal representations expressed

487   as firing rates (Figure 2). These sparse representations effectively discriminate between the plus and
488   multiplication patterns used in behavioural experiments (MaBouDi et al., 2021b) while also
489   generalising to novel stimuli—including successful recognition of human faces—highlighting the
490   model's broader applicability (Figures 4, 6).

491   Furthermore, our findings highlight the crucial role of bee movement, or active vision, in optimising
492   the analysis and encoding of environmental information (Figure 4). Spatiotemporal encoding in the
493   visual lobe emerges as a key mechanism driving the efficiency of minimal intelligent systems. Our
494   study underscores the fundamental computational principles of visual pattern recognition, particularly
495   the interplay between active vision and spatiotemporal encoding in insect information processing.
496   These insights not only advance our understanding of biological vision but also inspire the
497   development of novel computational models for visual recognition tasks (Figures 3, 4, 5).

498   The question of how animals cope with a noisy and complex natural world has long been a central
499   topic in neuroscience and behavioural ecology (Barlow, 1961; Gibson, 1979; Menzel and Giurfa,
500   2006; Srinivasan, 2010). One key theoretical framework addressing this challenge is the efficient
501   coding hypothesis, which posits that early sensory systems compress incoming information into a
502   more efficient format, optimising the transmission of relevant signals to higher brain regions (Barlow,
503   1961). According to this hypothesis, individual visual neurons should maximise their output capacity
504   (e.g., reaching their maximum firing rate) when responding to natural stimuli while population
505   responses should exhibit statistical independence (Simoncelli and Olshausen, 2001). Despite the
506   relative simplicity of our model compared to recently available full *Drosophila* connectome data (Lin
507   et al., 2024; Schlegel et al., 2024), our findings suggest that insects optimise visual coding through
508   non-associative learning while actively exploring their environment (see below). Specifically, we
509   demonstrate that neural features of the insect brain, combined with active vision, facilitate this
510   optimisation by developing uncorrelated and sparse coding in the lobula (Figures 2, 3, 5). This
511   supports the idea that efficient coding is not merely a passive process but one actively shaped by an
512   animal's interactions with its surroundings. However, bees exhibit a remarkably diverse behavioural
513   repertoire despite their small brains—ranging from fine-scale object inspections to long-distance
514   navigation (Chittka, 2022; Juusola et al., 2025; Menzel, 2012; Srinivasan, 2010). This behavioural
515   diversity makes them an excellent model for investigating how ecological constraints shape neural
516   computation and, ultimately, efficient coding. Understanding how insects dynamically refine sensory
517   representations in response to environmental demands offers broader insights into the fundamental
518   principles of neural information processing in biological systems.

519   The non-associative model presented in this study operates as a linear generative model that
520   effectively captures the receptive fields of lobula neurons, linking the spatiotemporal statistics of
521   natural environments to principles of efficient neural coding (Barlow, 1961; Olshausen, 2003).

522 Following training, lobula neuron activity in response to naturalistic spatiotemporal signals becomes
523 highly decorrelated, with only a limited subset of neurons selectively responding to specific visual
524 stimuli (Figure 5). This sparse coding strategy enhances energy efficiency by minimising overall
525 neural activity while maintaining distinct stimulus encoding. By ensuring that only a small fraction of
526 neurons is active at any given moment, this mechanism optimises information processing, reduces
527 redundancy, and enhances metabolic efficiency, reinforcing the adaptive advantages of efficient
528 coding in visual systems. Our model introduces a novel generative framework that can be extended
529 to other species, including primates, to investigate how movement contributes to visual-spatial
530 encoding in larger brains. Given the ubiquity of active vision across the animal kingdom (Land, 1973;
531 Land and Nilsson, 2012; Washburn, 1926, 1916; Yarbus, 2013), the principles identified in this study
532 may be broadly conserved across different taxa, underscoring the fundamental role of sensorimotor
533 interactions in shaping neural representations. These insights provide a valuable foundation for future
534 comparative studies on the interplay between movement, efficient coding, and sensory processing
535 across diverse neural architectures.

536 Our findings align with previous studies on bumblebees' discrimination of plus and multiplication sign
537 patterns (MaBouDi et al., 2021b), demonstrating improved model performance when scanning the
538 lower half of patterns at specific velocities (Figure 3). However, bees exhibit variations in scanning
539 behaviour depending on pattern complexity and training (Giurfa et al., 1999; Guiraud et al., 2018,
540 2025b). Research has shown that both honeybees and bumblebees solve visual tasks by extracting
541 localised or elemental features within patterns, adapting their discrimination strategies accordingly
542 (Giurfa et al., 1999; MaBouDi et al., 2021b; Stach et al., 2004; Stach and Giurfa, 2005). This suggests
543 that bees develop tailored flight manoeuvres during training, optimising their scanning behaviour to
544 maximise visual information extraction. Although our model simplifies visual flight dynamics by
545 employing a five-step constant-speed horizontal scan (Figure 1B), this abstraction was useful for
546 isolating key computational principles. However, its failure to solve the complex pattern recognition
547 task in Figure 6F underscores its limitations, suggesting that incorporating longer and more dynamic
548 scanning strategies could enhance visual processing capacity. Extending the scanning duration while
549 integrating multiple visual features across quadrants—alongside mechanisms such as working
550 memory and sequential learning—could improve performance by enabling the model to retain and
551 integrate previously acquired visual information. Real-world insect vision, however, relies on more
552 flexible and adaptive scanning behaviours shaped by flight speed, head movements, and
553 environmental feedback. Future work should leverage recent advances in insect connectomics, which
554 reveal a diverse range of neuron types—including small object-detecting neurons, motion-sensitive
555 neurons, and colour-processing cells—alongside machine learning techniques for analysing animal
556 movement to develop a more comprehensive flight dynamics model. Incorporating variable scan

557 trajectories and real-time sensorimotor feedback will offer deeper insights into how active vision
558 optimises information acquisition and enhances learning in dynamic environments.

559

560 A key advantage of our model lies in its ability to leverage sparsity and selectivity to efficiently process
561 sequential visual data (Figures 2B, 3, 6), in contrast to models that rely on pixel-wise image
562 representation for training. Traditional models that directly process raw pixel values often require
563 substantial computational resources and struggle with scalability, particularly when handling large-
564 scale visual inputs (Amin et al., 2024, 2024; Ardin et al., 2016; Baddeley et al., 2012). In contrast, our
565 model extracts and encodes sequential visual information within a small population of lobula neurons
566 (Figures 2B, 6), significantly compressing the visual input while preserving essential features. This
567 sparse representation reduces redundancy, enhances computational efficiency, and ensures that only
568 the most informative aspects of the scene are processed for pattern recognition. Furthermore, our
569 model demonstrates adaptive selectivity, dynamically adjusting to different visual inputs by learning
570 lobula responses optimised to the statistical features of the scene. Unlike pixel-based models, which
571 require processing every individual pixel in an image, our approach extracts compact, high-contrast
572 signals that are more robust to noise and enhance generalisability. This is particularly relevant for bio-
573 inspired visual processing, as it aligns with known sparse and decorrelated representations in
574 biological vision systems. By integrating biologically inspired sparse coding, adaptive selectivity, and
575 motion-driven encoding, our model provides a robust alternative to conventional pixel-based
576 architectures. This not only improves computational efficiency but also enhances discriminability and
577 generalisation, making it well-suited for real-world applications, including robotic vision and
578 autonomous navigation, where rapid adaptation to dynamic environments is crucial.

579

580 The results of our model suggest that passive visual exposure to natural images alters the connectivity
581 in the visual lobes, leading to enhanced pattern recognition abilities (Figures 4, 6). Notably, these
582 synaptic connections develop independently of the initial connectivity profiles of the simulated bees.
583 We propose that, beyond the general gross neuroanatomy of the insect optic lobes—which has been
584 preserved since the Cambrian period for efficient neural representation (Ma et al., 2012)—the specific
585 visual experiences encountered by real bees during early life play a crucial role in shaping their
586 individual visual representations. This, in turn, may influence their subsequent performance in
587 behavioural tasks (Hertel, 1983, 1982; MaBouDi et al., 2017a; Vetter and Visscher, n.d.). There is
588 direct empirical evidence for such neural developmental processes in olfactory systems of bees,
589 where early passive exposure improves subsequent odour discrimination (Arenas and Farina, 2008;
590 Locatelli et al., 2013). Our previous research in olfactory coding demonstrated that the iSTDP learning
591 rule can establish specific connectivity in the sensory system and enhance the separability of odour
592 representations in antennal lobe outputs (MaBouDi et al., 2017a). A similar mechanism was observed
593 here among lobula neurons, where only a limited subset is activated by specific visual inputs, resulting

25

594   in sparse and distinct outputs to the mushroom body learning centres (Figures 2B, 5). The receptive
595   fields of lobula cells, maintained with the fixed lateral connectivity, shows that inhibition is required for
596   orientation selectivity and temporal coding in the visual lobe (Fisher et al., 2015). These findings
597   highlight the critical role of inhibitory connections within the visual lobes. Accordingly, our model
598   predicts that bees with limited early-life visual experiences will perform worse in visual learning and
599   memory compared to bees with rich visual experiences. Further behavioural and neurobiological
600   studies are needed to test this prediction.

601

602   Mushroom bodies are critical centres for associative learning and memory in insects (Fiala and Kaun,
603   2024; Heisenberg, 2003; Menzel, 2022, 2012). Synapses between Kenyon Cells and extrinsic
604   mushroom body neurons follow a Hebbian STDP rule (Aso et al., 2014; Cassenaer and Laurent, 2007;
605   Markram et al., 1997); however, the STDP rule alone cannot maintain associative learning (Abbott
606   and Nelson, 2000; Meeks and Holy, 2008). In insects, associative learning appears to rely on the
607   neurotransmitters octopamine and dopamine to signal unconditioned appetitive and aversive values
608   (Cognigni et al., 2018; Davidson et al., 2023; Fisher, 2024; Hammer, 1993; Hammer and Menzel,
609   1995; Matsumoto et al., 2015; Mohammad et al., 2024; Perry and Barron, 2013; Schwaerzel et al.,
610   2003; Selcho, 2024). These neurotransmitters are released into the mushroom body lobes, where
611   Kenyon cells connect to mushroom body output neurons (MBON) (Burke et al., 2012; Menzel, 2022;
612   Okada et al., 2007; Strube-Bloss et al., 2011). Using in vivo electrophysiology in locusts, Cassenaer
613   and Laurent (2012) reported that octopamine depresses synapses underlying STDP rule, leading to
614   a lower response in MBONs when octopamine is present.

615

616   Following this observation, we modelled associative learning by pairing the positive pattern with the
617   reward via octopamine-modulated STDP (Equation 4; Figure 1). In this formulation, the temporal
618   ordering of pre- or postsynaptic spikes depresses the synaptic connection between Kenyon cells and
619   the MBONs. Conversely, synapses are updated according to classical STDP when negative patterns
620   are paired with the punishment (Equation 3; Figure 9). This combination produces a complex interplay
621   between synaptic changes and reinforcer signals, enabling the model to not only learn to select the
622   positive patterns but to reject incorrect ones (Figure 4B, 6C). The resulting changes in MBON
623   response to positive patterns during associative learning are consistent with the PE1 extrinsic neuron
624   in the honeybee brain, which exhibits a lower response to the positive patterns (Ce et al., 2022; Fiala
625   and Kaun, 2024; Okada et al., 2007) (Figure 4B).

626

627   However, further studies are required to investigate the novel combination of octopamine and
628   dopamine modulation of STDP that is introduced in this study. Combining non-associative learning in
629   the optic lobes with supervised learning in the mushroom bodies produced a model capable of not
630   only discriminating simple patterns but also generalisation (Figure 6B), and correct judgments in

26

631   conflicting stimulus experiments (Figure 6D). The real power of this approach is exemplified in the

632   facial recognition task (Figure 6E), where the complexity of the human face is reduced to a set of

633   sparse lobula neuron activations that can be learnt by the mushroom bodies. Moreover, the

634   spatiotemporal receptive fields formed during non-associative learning respond differently for different

635   faces, allowing fine differences to be encoded. Although real bees rarely have to discriminate between

636   human faces, these processes likely enable bees to select rewarding flowers without requiring a

637   complex visual memory within their miniature brains.

638

639   We used natural scenes with a statistical structure similar to those that visual systems have adapted

640   to over evolutionary time (Geisler, 2008; Hyvärinen et al., 2009; MaBouDi et al., 2016; Simoncelli and

641   Olshausen, 2001). Because bee navigation and foraging primarily involve locating food among a

642   variety of flowers, our non-associative network was trained with a set of different flower images. As

643   with all theoretical models, this is a simplification, since real bees navigate a 3D environment with a

644   large field of view. Here, we assume that receptive field formation in real bees is comparable to our

645   2D simulations. Nevertheless, further studies are necessary to refine and expand our model based

646   on a more comprehensive understanding of the function and structure of the bee eye components

647   (Juston et al., 2013; Juusola et al., 2017; Kemppainen et al., 2022b; Viollet and Franceschini, 2010).

648   Moreover, investigating the neural mechanisms underlying visual learning in the bee brain will allow

649   us to fine-tune our model's architecture and parameters, leading to a more faithful representation of

650   the bee visual system.

651

652   In this study, we restricted the model's input to green photoreceptors to align with the known visual

653   processing mechanisms of honeybees. This decision was based on the hypothesis that bee pattern

654   recognition primarily relies on the green component of visual input, as green-sensitive photoreceptors

655   are the most abundant, comprising approximately two-thirds of the ommatidia in the compound eye

656   (Briscoe and Chittka, 2001; Giger and Srinivasan, 1996; Spaethe et al., 2001; Spaethe and Briscoe,

657   2004). There is also empirical evidence that the green channel provides the predominant input to

658   movement and edge detection, as well as detailed spatial information for visual discrimination.

659   Moreover, natural images exhibit a strong correlation among colour channels, meaning that excluding

660   certain channels does not substantially alter the structure of the visual scene. By focusing on the

661   green photoreceptor input, our model remains biologically plausible while ensuring computational

662   efficiency. Future work could explore the contributions of other photoreceptor types to assess their

663   impact on visual pattern recognition and potential interactions between colour and spatial information

664   processing in the insect visual system.

665

666   Our model provides a functional abstraction of the insect visual system, focusing on core

667   computational principles rather than replicating the detailed structural connectivity available from

668   recent connectome studies (Lin et al., 2024; Schlegel et al., 2024). By prioritising the identification of

669   fundamental mechanisms underlying active vision and visual learning, our approach avoids the

670   challenges associated with highly parameterised models that can be difficult to interpret

671   mechanistically. Integrating known physiological properties and behavioural findings from bees, our

672   model generates testable hypotheses on how motion-driven visual processing enhances pattern

673   recognition. This functional simplification allows us to isolate key mechanisms that might otherwise

674   be obscured in large-scale anatomical reconstructions. Additionally, our study emphasises the critical

675   role of motion in visual recognition—an aspect often overlooked in static-image studies—and

676   demonstrates how sequential visual input dynamically shapes neural encoding, reinforcing the

677   importance of active vision in efficient sensory processing.

678

679   Recent studies have shown that bees often employ efficient, low-cost strategies to solve cognitive

680   tasks (Cope et al., 2018; Guiraud et al., 2018; Langridge et al., 2021; MaBouDi et al., 2023a, 2023b,

681   2021a, 2020b; Roper et al., 2017; Vasas et al., 2019).Understanding these cognitive strategies not

682   only advances our knowledge of neural computation in miniature brains but also provides a framework

683   for improving artificial intelligence and autonomous systems. Our study highlights the minimal neural

684   architectures required for visual learning and lays the foundation for bio-inspired, unsupervised

685   machine learning algorithms. By emphasising active vision through movement-driven pattern

686   recognition, our model offers insights into solutions for key AI challenges, such as visual invariance

687   and robust 3D environmental understanding. Moreover, engineering implementations of eye

688   micromovements have been shown to enhance edge and bar discrimination, improving the visual

689   processing efficiency of flying robots (Juston et al., 2013; Viollet and Franceschini, 2010). Additionally,

690   the non-associative learning model and local plasticity rules explored in this work closely align with

691   unsupervised learning techniques, particularly sparse coding models, where sparsity constraints

692   enhance efficiency by reducing redundancy and promoting selective coding. This bio-inspired

693   framework enables the extraction of latent structures in high-dimensional temporal data, with

694   applications ranging from sensory signal processing to more adaptive and robust autonomous

695   perception. Bridging biological and machine intelligence through evolutionarily optimised

696   computational strategies paves the way for the next generation of AI, driving advancements in

697   robotics, autonomous navigation, and real-world learning systems (de Croon et al., 2022;

698   Manoonpong et al., 2021; Serres and Viollet, 2018; Webb, 2020).

699

700   **Materials and Methods**

701   ***Network topology of active vision model***

702   The model architecture of the bee visual pathway is illustrated in Figure 1A. The bumblebee has a

703   pair of compound eyes that are composed of ~5,500 ommatidia (Spaethe and Chittka, 2003b;

704   Streinzer et al., 2013). Each eye contains three different types of photoreceptors, short, medium and

705   long wavelength sensitive peaking in the UV, blue and the green respectively (Menzel and Blakers,

706   1976; Skorupski et al., 2007). Since the green photoreceptors are those that predominantly mediate

707   visual pattern recognition (Giger and Srinivasan, 1996; Spaethe et al., 2001), we modelled that 75x75

708   green photoreceptors in one eye component are activated by the pixel values of the input pattern.

709   Photoreceptors then project to 625 (25x25) neurons in the lamina, which is the first centre of visual

710   processing. In this model, each lamina neuron, r_l^{La}, receives input from a non-overlapping 3x3 grid of

711   neighbouring photoreceptors, corresponding to adjust ommatidia. The response of a lamina neuron

712   is computed as $r_l^{La} = f(\sum_{p=1}^{P=9} r_p \; ; \; A_0, m, b)$. Here, the activation function f is defined as:

713    $$f(r; A_0, m, b) = A_0 / (1 + exp(mr + b),$$

714   where r is the input to the lamina neuron La, and A_0 represent the maximum possible activity of

715   lamina neurons. The parameters m and b define the shape of the activation function, controlling its

716   steepness and midpoint, respectively. For simplicity, the activation function in our model is fixed with

717    $A_0 = 1$, $m = -1$, and $b = 0.5$. This function imposes a constraint where weak inputs result in low

718   activity, while stronger inputs drive the response toward its maximum value in a sigmoidal manner.

719   Each photoreceptor's output, $r_{p,}$ is derived directly from the pixel intensity at the corresponding location

720   in the input image, representing the green channel's brightness. The values are normalised between

721   0 and 1, ensuring a continuous response that reflects the natural variation in luminance.

722   In this study, each spiking neuron operates according to the integrate-and-fire model. The dynamics

723   of the subthreshold membrane potential of a neuron, $u(t)$ is described by the following standard

724   conductance- based leaky integrate-and-fire model: $\tau \frac{du(t)}{dt} = -u(t) + R.I(t)$, where $R = 10$ and $\tau =$

725    $10ms$ are the resistance and membrane time constant of the neuron respectively. Here, the input $I(t)$

726   exhibit the total synaptic input to the cell from presynaptic neurons. The membrane potential is reset

727   to the base activity, $v_0 = -80 \; mV$, if it exceeds the threshold, $V_T = 0 \; mV$.

728   Each medulla neuron is activated by the summed activity of lamina neurons through the synaptic

729   connectivity matrix W. The input of the $m - $ the medulla neuron, I_m^{Me}, is calculated $I_m^{Me} =$

730    $\sum_{l=1}^{L} W_{m,l} \; r_l^{La}$. The value $W_{l,m}$ specifies the strength of a synaptic input from the $l -$th lamina neuron

731   to the $m - $the medulla neuron. To account for the inherent variability in neural responses, we

732   incorporated stochasticity by adding signal noise generated from a Poisson distribution to the output

733   of each neuron. The Poisson distribution was chosen because it closely models the statistical

734   fluctuations observed in biological neural firing, where the variance of the response scales with the

735   mean activity. This noise was applied independently to each neuron, ensuring that the variability in

736   responses remains biologically plausible while preserving the overall signal structure. By introducing

737 this element, the model better reflects the natural dynamics of neural processing, capturing the
738 probabilistic nature of spike generation and sensory encoding in biological systems.

739 We propose a temporal coding model that captures the interaction between medulla and lobula
740 neurons in the visual pathway, incorporating sequential scanning of visual stimuli to optimise
741 information processing. In this model, each wide-field lobula neuron receives synaptic input from M
742 small-field medulla neurons, with a structured progressive delay T (Figure 1D). Here, $M = 5$
743 represents the number of temporal instances within the model's input sequence. Each medulla neuron
744 is activated by visual information sampled from one of M overlapping segments of an image patch,
745 determined by the scanning speed, and follows the hierarchical processing pathway from
746 photoreceptors to lamina neurons (Figure 1C). While individual medulla neurons encode only a
747 fraction of the visual stimulus, the lobula neuron integrates input from all medulla neurons to generate
748 spiking activity, thereby forming a holistic representation of the entire visual scene. The synaptic
749 transmission between medulla and lobula neurons incorporates structured temporal delays at distinct
750 instances $(T, 2T, 3T, 4T, 5T)$, ensuring that sequentially acquired visual information is temporally
751 aligned. This results in the synchronized activation of the lobula neuron at a single unified time point,
752 effectively integrating spatially and temporally structured input into a cohesive internal representation.
753 By simulating the dynamic interplay between spatial sampling and temporal integration, this model
754 mirrors the way bees may optimise visual processing through active vision. The resulting alignment
755 of visual signals enhances feature extraction and pattern recognition, providing a biologically plausible
756 mechanism for encoding complex scenes efficiently (Figures 2, 3, 6).

757 The model incorporates lateral inhibitory connections between lobula neurons (Figure 1C, depicted
758 in red), $Q = [q_{i,j}]$, where $q_{i,j}$ represents the lateral connectivity between $i-$th and $j-$th lobula
759 neurons. This connectivity along with $W_{m,l}$ are updated during a non-associative learning process, to
760 reduce redundancy and decorrelate overlapping inputs (see next subsection). This inhibitory
761 mechanism enhances contrast and improves pattern recognition by selectively amplifying novel
762 spatial features (Figures 3, 6).

763 The processed visual information is then transformed to the Kenyon cells (KCs) in the mushroom
764 body. The synaptic connectivity matrix $S^{Lo \rightarrow KC} = [s_{o,k}]$ determines the excitatory connections
765 between lobula neurons and KCs in the mushroom body, following previous findings on sparse,
766 random connectivity (Caron et al., 2013; Szyszka et al., 2005). Kenyon cells exhibit sparse activity,
767 with fewer than 5% of KCs activated per stimulus, ensuring high selectivity for particular image
768 features (Honegger et al., 2011). This sparsity emerges naturally from the random connectivity and
769 thresholder activation dynamics, reinforcing the sparseness in the model. Each connection weight
770 $s_{o,k}$ is randomly initialized from a uniform distribution in the range $[0, w_{max}]$, where $, w_{max}$ is a scaling
771 factor ensuring the limitation of input to the KCs. For each simulated bee, the connectivity matrix

772   $S^{Lo \to KC}$ is randomly reinitialised, ensuring that each instance of the model has a unique but statistically

773   comparable connectivity structure. This reinitialization reflects individual variations in synaptic wiring

774   and allows us to assess the robustness of the model's pattern recognition ability across different

775   randomly generated network configurations.

776   All Kenyon cells project to a single mushroom body output neuron (MBON), which is the final output

777   of the of the MBON, I_{MBON}, is computed by the KC-MBON connections D such

778   that $I_{MBON} = \sum_{k=1}^{K} D_k r_k^{KC}$, where r_k^{KC} is the spiking activity of the k-th KCs. Finally, a reinforcement

779   neuron makes reinforcement-modulated connections with the KCs and MBON in the presence of the

780   positive and negative patterns (see the next section).

781   Each neuron type is denoted with a superscript corresponding to its respective processing stage. For

782   example, $r_{p,}$ represents the response of an individual photoreceptor, while r^{La} denotes the activity of

783   a lamina neuron, where 'La' refers to the lamina layer. Similarly, other superscripts—Me, Lo, KC, and

784   MBON—correspond to specific network components: the medulla, lobula, Kenyon cells, and

785   mushroom body output neurons, respectively. This ensures consistency in notation throughout the

786   model description.

787   ***Training the network via a non-associative learning***

788   We trained the model using 50,000 time-varying image patches randomly sampled from a dataset of

789   100 natural flowers and scene images. During each training step, the model received an input

790   sequence of five sequential 75 × 75-pixel patches, extracted by shifting 15 pixels across the image

791   from the left or right or the reverse orientation (Figures 1B, 2B). This patchwise input simulates the

792   sequential scanning behaviour of bees as they explore visual stimuli.

793   Using the described network architecture, each time-varying patch dynamically drives spiking activity

794   in lobula neurons as the simulated movement progresses. At the start of training, all inhibitory

795   connection strengths Q were initialized with values randomly drawn from a uniform distribution

796   between 0 and 1. The feedforward synaptic weights W were initialized using Gaussian white noise

797   $\mathcal{N}(0,1)$. As training progressed, the evoked neural responses of lobula neurons to time-varying

798   patches were used to iteratively update both the inhibitory weights (Q) and feedforward connections

799   (W) simultaneously (see Discussion section).

800   After the image presentation, the feed-forward weight W is updated according to Oja's

801   implementation of the Hebbian learning rule (MaBouDi et al., 2017a; Oja, 1982) via

802   $$\Delta W_{i,j} = \gamma r_j^{Me}(r_i^{La} - r_j^{Me} W_{i,j}) \qquad \text{(Equation 1)}$$

803   Here, the r_j^{Me} and r_i^{La} represent the activities of the j-th medulla and i-th lamina neurons, respectively.

804   The positive constant γ defines the learning rate.

805   At the same time of processing, the lateral inhibitory connectivity in the lobula is modified by inhibitory

806   spike-time-dependent plasticity (iSTDP) (Vogels et al., 2011). Here, we model *non-associative*

807   *learning* in lobula by a symmetric iSTDP between presynaptic of the inhibitory neurons and

808   postsynaptic lobula neurons. In this learning rule, both temporal ordering of pre- or post- synaptic

809   spikes potentiates the connectivity and the synaptic strength of j-th inhibitory neuron onto i-th lobula

810   neuron ($Q_{i,j}$) is updated as follows:

811
$$\Delta Q_{i,j} = \eta(r_i^{Lo} * r_j^{In} - \alpha) \qquad \text{(Equation 2)}$$

812   where r_i^{Lo} and r_j^{In} exhibit the mean firing rate of the lobula and inhibitory neurons, respectively. The

813   depression factor α controls the target activity rate of the lobula neurons. Here, η is the learning rate.

814   To simplify, a one-to-one connection between the inhibitory and lobula neuron is assumed in the

815   model such that the activity of the j-th inhibitory neuron is equal to the activity of the j-th lobula neuron.

816   The training is terminated when the synaptic weights over time are changed less than a small

817   threshold (0.001). In the training process, synaptic weights were constrained within the range $[-1, +1]$

818   for the W and $[0, +1]$ for Q to ensure stable convergence and to reflect biological limitations in

819   synaptic transmission strength.

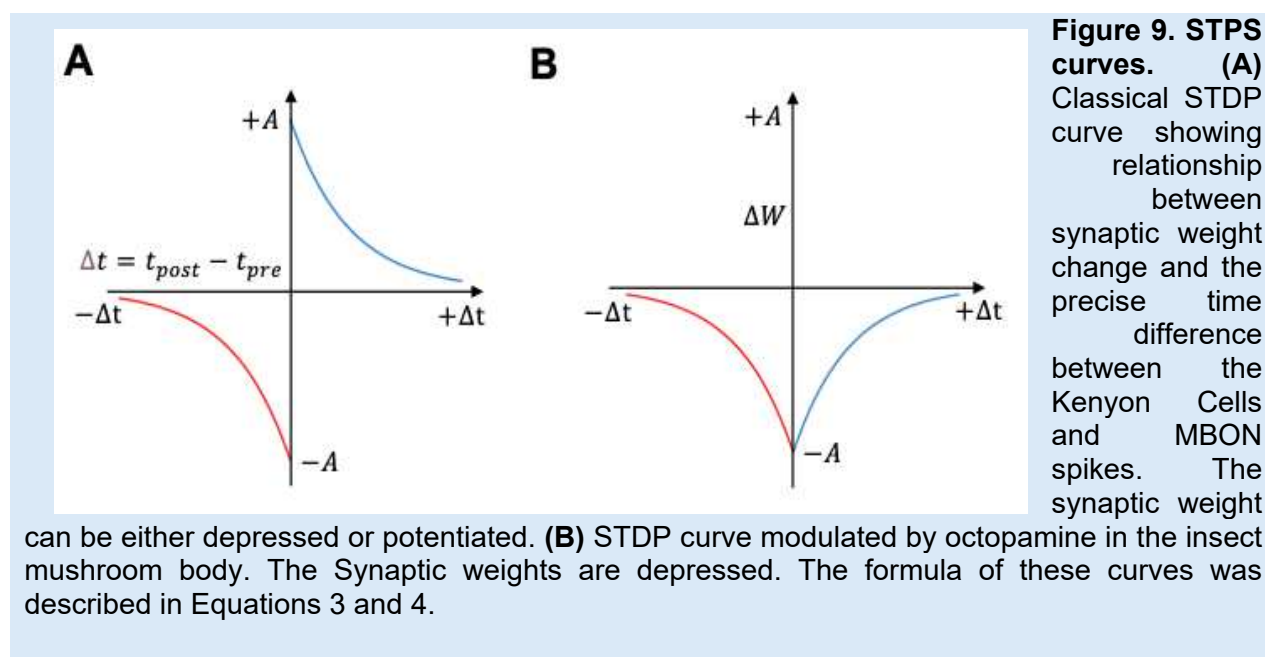820   ***Associative learning in Mushroom Bodies***

821   To verify if the lobula neurons can reproduce empirical behavioural results in different visual tasks,

822   the model is enriched with associative learning process in the mushroom bodies (a bio-inspired

823   supervised learning). When the training process of the non-associative learning is terminated, we use

824   a reward-based synaptic wright modification rule in KCs-MBON connection (D), such that, if a stimulus

825   is rewarding (*i.e.* positive), the corresponding synapses between activated neurons will be weakened

826   while for a stimulus paired with punishment (*i.e.* negative), activated synapses are strengthened

827   (Cassenaer and Laurent, 2012) (see Discussion section). The model behaves as the activity of

828   mushroom body neurons in decreasing their firing rate in responding to the positive stimuli during

829   training (Okada et al., 2007). In this model, two reinforcement neurons modulated strengths of

830   synaptic connectivity at the output of the KCs in response to both reward and punishment. In the

831   presence of the negative patterns, the synaptic strengths from the KCs to the MBON are modified,

832   and modulated by dopamine, based on the classical STDP (Song et al., 2000; Zhang et al., 1998)

833   (Figure S3A):

834
$$STDP_{Dop}(\Delta t) = \begin{cases} A\,e^{-\Delta t/\tau}, & \Delta t > 0 \\ -A\,e^{\Delta t/\tau}, & \Delta t < 0 \end{cases}, \qquad \text{(Equation 3)}$$

835 where $\Delta t = t_{post} - t_{pre}$ implies the difference between the spike time of pre- and post- synaptic

836 neurons. Further, applying the synaptic plasticity rule modulated by octopamine (octopamine

837 modulated STDP) observed in the presence of rewarding stimuli to the synapses between KCs and

838 MBON (Cassenaer and Laurent, 2012), the change in synaptic weight can be summarised as (Figure

839 4B):

840
$$STDP_{OCT}(\Delta t) = \begin{cases} -A\, e^{-\Delta t/\tau}, & \Delta t > 0 \\ -A\, e^{\Delta t/\tau}, & \Delta t < 0 \end{cases}, \quad \text{(Equation 4)}.$$

841 Here, $A = 0.01$ and $\tau = 20\,ms$ exhibit the maximum magnitude and time constant of the STDP

842 function for the synaptic potentiation or depression.

Figure 9. STPS curves. (A) Classical STDP curve showing relationship between synaptic weight change and the precise time difference between the Kenyon Cells and MBON spikes. The synaptic weight can be either depressed or potentiated. **(B)** STDP curve modulated by octopamine in the insect mushroom body. The Synaptic weights are depressed. The formula of these curves was described in Equations 3 and 4.

843

844 To train the model in different conditions of scanning, the flight-scan forms of the positive and negative

845 patterns were presented to the model. Each set of flight-scan input contained a set of five patches

846 with size 75x75 pixels were selected from the test patterns by shifting 15 pixels over each pattern

847 from the left to right (Figure 4A). The numbers of shifted pixels control the speed of scanning. The

848 activity of the MBON was used to assess the performance of the model. Following the training, the

849 performance of the model was calculated from a decrease in firing rate of the MBON to a pattern that

850 had been rewarding and/or an increase in firing rate of MBON to a pattern that had been punishing in

851 training. The bee's final behavioural decision is proposed to come from a simple integration of these

852 different valence-encoding neurons.

853

854 ***Quantifying neural population sparseness and response separability***

855 To quantify population sparseness in the lobula following training on natural images, we employed

856 the Treves-Rolls sparseness index (MaBouDi et al., 2017b; Willmore and Tolhurst, 2001), defined as:

857
$$SI = \frac{\left(\sum_{j=1}^{N} r_j / N\right)^2}{\left(\sum_{j=1}^{N} r_j^2\right)/N}$$

858 where r_j represents the firing rate of the j-th lobula neuron, and N is the total number of lobula

859 neurons.

860 This metric provides insight into the population coding strategy of lobula neurons, distinguishing

861 between broad distributed representations and sparse selective responses:

- *Maximum SI = 1 (low sparseness):* Achieved when all neurons respond equally, indicating a
  fully distributed code where the entire population is uniformly active across all stimuli.
- *Minimum SI = 1/N (high sparseness):* Occurs when only a single neuron is active while all
  others remain silent, reflecting a highly selective encoding scheme.

866

867 To further assess the separability between response population of lobula neurons in response to

868 different stimuli, we computed the angular distance (θ) using the cosine similarity formula: $\theta =$

869 $\cos^{-1}(R_1.R_2/(|R_1||R_1|))$, where R_1 and R_2 represent the activity vectors of two neural populations,

870 and $|R_1| \ and \ |R_2|$ denote their responses Euclidean norms. This measure captures the geometric

871 distinction between response patterns, with larger angles indicating greater separability between

872 neural representations of different stimuli. A higher angular distance suggests that the population

873 responses are more distinct, reflecting improved stimulus discrimination within the lobula.

874

875 ### *Simulation and Statistical Analysis*

876 To assess the model's performance across experiments, we conducted 20 independent simulations

877 for each condition, ensuring statistical reliability and robustness. In each simulation, the synaptic

878 connectivity matrix between lobula neurons and Kenyon cells was randomly initialised using a uniform

879 distribution within a biologically plausible range, mimicking the individual variability observed in real

880 bees. This stochastic initialisation prevented bias in learning outcomes and allowed us to examine

881 how the model generalises across different neural configurations.

882 Each simulated bee underwent multiple training exposures to visual stimuli. Following training, once

883 the model parameters were fixed at the final stage of training, the model was tested across 50

884 repetitions per condition with testing patterns to account for variability in responses. The model's

885 performance was evaluated by averaging across simulations, with standard error of the mean (SEM)

886 reported in the figures to provide a statistically robust representation of discrimination accuracy.

887 Statistical analyses were conducted to compare pattern discrimination across conditions. To maintain

888 clarity and focus on the modelling findings, detailed significance values are not reported in the main

34

889 text. Instead, figures indicate p-values < 0.05 with '*' and non-significant results with 'n.s.'. Since data

890 distributions did not always meet normality assumptions, we used the Wilcoxon signed-rank test for

891 matched data and the Wilcoxon rank-sum test (Mann-Whitney U test) for independent samples. For

892 comparisons across multiple groups, we applied the Kruskal-Wallis test, followed by Dunn's post-hoc

893 test when necessary.

894 ***Computing Environment***: All modelling and visualisation were performed using MATLAB

895 (RRID:SCR_001622) and Python (RRID:SCR_008394). MATLAB was also used for statistical

896 analysis.

897

Acknowledgements

906

Conflict of interest statement

908 All authors declare no conflict of interest.

909

Data and code accessibility

911 The code developed for this research project has been made openly accessible on GitHub:

912 https://github.com/hadiimaboudi/neuromorphic_model_active-vision-

913

Reference

915 Abbott LF, Nelson SB. 2000. Synaptic plasticity: taming the beast. *Nature neuroscience* **3**:1178–
916 1183.
917 Amin AA, Kagioulis E, Domcsek N, Nowotny T, Graham P, Philippides A. 2024. Investigating the
918 limits of familiarity-based navigation. *Artificial Life* 1–17.
919 Anderson AG, Ratnam K, Roorda A, Olshausen BA. 2020. High-acuity vision from retinal image
920 motion. *Journal of vision* **20**:34–34.
921 Ardin P, Peng F, Mangan M, Lagogiannis K, Webb B. 2016. Using an insect mushroom body circuit
922 to encode route memory in complex natural environments. *PLoS computational biology*
923 **12**:e1004683.
924 Arenas A, Farina WM. 2008. Age and rearing environment interact in the retention of early olfactory
925 memories in honeybees. *J Comp Physiol A* **194**:629–640. doi:10.1007/s00359-008-0337-z

926 Arenz A, Drews MS, Richter FG, Ammer G, Borst A. 2017. The temporal tuning of the *Drosophila*
927 motion detectors is determined by the dynamics of their input elements. *Current Biology* **27**:929–
928 944.
929 Aso Y, Hattori D, Yu Y, Johnston RM, Iyer NA, Ngo T-T, Dionne H, Abbott LF, Axel R, Tanimoto H.
930 2014. The neuronal architecture of the mushroom body provides a logic for associative learning.
931 *elife* **3**:e04577.
932 Avarguès-Weber A, Dyer AG, Giurfa M. 2011. Conceptualization of above and below relationships
933 by an insect. *Proceedings of the Royal Society of London B: Biological Sciences* **278**:898–905.
934 doi:10.1098/rspb.2010.1891
935 Baddeley B, Graham P, Husbands P, Philippides A. 2012. A model of ant route navigation driven by
936 scene familiarity. *PLoS computational biology* **8**:e1002336.
937 Barlow HB. 1961. Possible principles underlying the transformation of sensory messages. *Sensory
938 communication* **1**.
939 Benard J, Stach S, Giurfa M. 2006. Categorization of visual stimuli in the honeybee *Apis mellifera*.
940 *Anim Cogn* **9**:257–270. doi:10.1007/s10071-006-0032-9
941 Bertrand OJ, Doussot C, Siesenop T, Ravi S, Egelhaaf M. 2021. Visual and movement memories
942 steer foraging bumblebees along habitual routes. *Journal of Experimental Biology* **224**:jeb237867.
943 Boeddeker N, Mertes M, Dittmar L, Egelhaaf M. 2015. Bumblebee Homing: The Fine Structure of
944 Head Turning Movements. *PLOS ONE* **10**:e0135020. doi:10.1371/journal.pone.0135020
945 Brinkworth RS, O'Carroll DC. 2009. Robust models for optic flow coding in natural scenes inspired
946 by insect biology. *PLoS computational biology* **5**:e1000555.
947 Briscoe AD, Chittka L. 2001. THE EVOLUTION OF COLOR VISION IN INSECTS. *Annual Review of
948 Entomology* **46**:471–510. doi:10.1146/annurev.ento.46.1.471
949 Burke CJ, Huetteroth W, Owald D, Perisse E, Krashes MJ, Das G, Gohl D, Silies M, Certel S,
950 Waddell S. 2012. Layered reward signalling through octopamine and dopamine in *Drosophila*.
951 *Nature* **492**:433–437.
952 Caron SJC, Ruta V, Abbott LF, Axel R. 2013. Random convergence of olfactory inputs in the
953 *Drosophila* mushroom body. *Nature* **497**:113–117. doi:10.1038/nature12063
954 Cassenaer S, Laurent G. 2012. Conditional modulation of spike-timing-dependent plasticity for
955 olfactory learning. *Nature* **482**:47–52. doi:10.1038/nature10776
956 Cassenaer S, Laurent G. 2007. Hebbian STDP in mushroom bodies facilitates the synchronous flow
957 of olfactory information in locusts. *Nature* **448**:709–713. doi:10.1038/nature05973
958 Ce H, V R, Ey R, Sh D, Mp N, A F. 2022. Visualization of learning-induced synaptic plasticity in
959 output neurons of the *Drosophila* mushroom body γ-lobe. *Scientific reports* **12**. doi:10.1038/s41598-
960 022-14413-5
961 Chittka L. 2022. The Mind of a Bee. Princeton University Press. doi:10.1515/9780691236247
962 Chittka L, Niven J. 2009. Are bigger brains better? *Curr Biol* **19**:R995–R1008.
963 doi:10.1016/j.cub.2009.08.023
964 Chittka L, Skorupski P. 2017. Active vision: a broader comparative perspective is needed.
965 Clark DA, Fitzgerald JE, Ales JM, Gohl DM, Silies MA, Norcia AM, Clandinin TR. 2014a. Flies and
966 humans share a motion estimation strategy that exploits natural scene statistics. *Nature
967 neuroscience* **17**:296–303.
968 Clark DA, Fitzgerald JE, Ales JM, Gohl DM, Silies MA, Norcia AM, Clandinin TR. 2014b. Flies and
969 humans share a motion estimation strategy that exploits natural scene statistics. *Nat Neurosci*
970 **17**:296–303. doi:10.1038/nn.3600
971 Cognigni P, Felsenberg J, Waddell S. 2018. Do the right thing: neural network mechanisms of
972 memory formation, expression and update in *Drosophila*. *Current Opinion in Neurobiology* **49**:51–
973 58. doi:10.1016/j.conb.2017.12.002
974 Collett TS, Fry SN, Wehner R. 1993. Sequence learning by honeybees. *J Comp Physiol A* **172**:693–
975 706. doi:10.1007/BF00195395
976 Cope AJ, Vasilaki E, Minors D, Sabo C, Marshall JAR, Barron AB. 2018. Abstract concept learning
977 in a simple neural network inspired by the insect brain. *PLOS Computational Biology* **14**:e1006435.
978 doi:10.1371/journal.pcbi.1006435

979 Davidson AM, Kaushik S, Hige T. 2023. Dopamine-Dependent Plasticity Is Heterogeneously
980 Expressed by Presynaptic Calcium Activity across Individual Boutons of the *Drosophila* Mushroom
981 Body. *eNeuro* **10**. doi:10.1523/ENEURO.0275-23.2023

982 Dawkins MS, Woodington A. 2000. Pattern recognition and active vision in chickens. *Nature*
983 **403**:652–655. doi:10.1038/35001064

984 de Croon GCHE, Dupeyroux JJG, Fuller SB, Marshall JAR. 2022. Insect-inspired AI for autonomous
985 robots. *Science Robotics* **7**:eabl6334. doi:10.1126/scirobotics.abl6334

986 Doussot C, Bertrand OJN, Egelhaaf M. 2021. The Critical Role of Head Movements for Spatial
987 Representation During Bumblebees Learning Flight. *Frontiers in Behavioral Neuroscience* **14**.

988 Dyakova O, Lee Y-J, Longden KD, Kiselev VG, Nordström K. 2015. A higher order visual neuron
989 tuned to the spatial amplitude spectra of natural scenes. *Nature Communications* **6**:8522.
990 doi:10.1038/ncomms9522

991 Dyakova O, Müller MM, Egelhaaf M, Nordström K. 2019. Image statistics of the environment
992 surrounding freely behaving hoverflies. *Journal of Comparative Physiology A* **205**:373–385.

993 Dyakova O, Nordström K. 2017. Image statistics and their processing in insect vision. *Current
994 Opinion in Insect Science*, Neuroscience * Pheromones **24**:7–14. doi:10.1016/j.cois.2017.08.002

995 Dyer AG, Neumeyer C, Chittka L. 2005. Honeybee (*Apis mellifera*) vision can discriminate between
996 and recognise images of human faces. *Journal of Experimental Biology* **208**:4709–4714.

997 Dyer AG, Paulk AC, Reser DH. 2011. Colour processing in complex environments: insights from the
998 visual system of bees. *Proceedings of the Royal Society of London B: Biological Sciences* **278**:952–
999 959. doi:10.1098/rspb.2010.2412

1000 Egelhaaf M, Kern R, Lindemann JP, Braun E, Geurten B. 2009. Active Vision in Blowflies:
1001 Strategies and Mechanisms of Spatial Orientation In: Floreano D, Zufferey J-C, Srinivasan MV,
1002 Ellington C, editors. Flying Insects and Robots. Berlin, Heidelberg: Springer Berlin Heidelberg. pp.
1003 51–61. doi:10.1007/978-3-540-89393-6_4

1004 Ehmer B, Gronenberg W. 2002. Segregation of visual input to the mushroom bodies in the
1005 honeybee (*Apis mellifera*). *J of Comparative Neurology* **451**:362–373. doi:10.1002/cne.10355

1006 Fenk LM, Avritzer SC, Weisman JL, Nair A, Randt LD, Mohren TL, Siwanowicz I, Maimon G. 2022.
1007 Muscles that move the retina augment compound eye vision in *Drosophila*. *Nature* **612**:116–122.
1008 doi:10.1038/s41586-022-05317-5

1009 Fiala A, Kaun KR. 2024. What do the mushroom bodies do for the insect brain? Twenty-five years
1010 of progress. *Learn Mem* **31**:a053827. doi:10.1101/lm.053827.123

1011 Fisher YE. 2024. Octopamine enhances learning. *National Science Review* **11**:nwae185.
1012 doi:10.1093/nsr/nwae185

1013 Fisher YE, Silies M, Clandinin TR. 2015. Orientation Selectivity Sharpens Motion Detection in
1014 *Drosophila*. *Neuron* **88**:390–402. doi:10.1016/j.neuron.2015.09.033

1015 Geisler WS. 2008. Visual Perception and the Statistical Properties of Natural Scenes. *Annual
1016 Review of Psychology* **59**:167–192. doi:10.1146/annurev.psych.58.110405.085632

1017 Gibson JJ. 1979. The ecological approach to visual perception.

1018 Giger AD, Srinivasan MV. 1996. Pattern recognition in honeybees: chromatic properties of
1019 orientation analysis. *J Comp Physiol A* **178**:763–769. doi:10.1007/BF00225824

1020 Giurfa M. 2013. Cognition with few neurons: higher-order learning in insects. *Trends in
1021 Neurosciences* **36**:285–294. doi:10.1016/j.tins.2012.12.011

1022 Giurfa M, Hammer M, Stach S, Stollhoff N, MüLler-Deisig N, Mizyrycki C. 1999. Pattern learning by
1023 honeybees: conditioning procedure and recognition strategy. *Animal Behaviour* **57**:315–324.
1024 doi:10.1006/anbe.1998.0957

1025 Giurfa M, Zhang S, Jenett A, Menzel R, Srinivasan MV. 2001. The concepts of 'sameness' and
1026 'difference' in an insect. *Nature* **410**:930–933. doi:10.1038/35073582

1027 Gribakin FG. 1975. Functional morphology of the compound eye of the bee. *The compound eye and
1028 vision of insects* **154**:176.

1029 Guiraud M, Roper M, Chittka L. 2018. High-Speed Videography Reveals How Honeybees Can Turn
1030 a Spatial Concept Learning Task Into a Simple Discrimination Task by Stereotyped Flight
1031 Movements and Sequential Inspection of Pattern Elements. *Front Psychol* **9**.
1032 doi:10.3389/fpsyg.2018.01347

1033   Guiraud M, Roper M, Wolf S, Woodgate JL, Chittka L. 2022. Discrimination of edge orientation by
1034   bumblebees. *Plos one* **17**:e0263198.
1035   Guiraud M-G, Gallo V, Quinsal-Keel E, MaBouDi H. 2025a. Bumble bee visual learning: simple
1036   solutions for complex stimuli. *Animal Behaviour* **221**:123070.
1037   Guiraud M-G, MaBouDi H, Woodgate J, Bates OK, Rodriguez OR, Gallo V, Barron AB. 2025b. How
1038   bumblebees manage conflicting information seen on arrival and departure from flowers. *Anim Cogn*
1039   **28**:11. doi:10.1007/s10071-024-01926-x
1040   Hammer M. 1993. An identified neuron mediates the unconditioned stimulus in associative olfactory
1041   learning in honeybees. *Nature* **366**:59–63. doi:10.1038/366059a0
1042   Hammer M, Menzel R. 1995. Learning and memory in the honeybee. *J Neurosci* **15**:1617–1630.
1043   Hardie RC, Franze K. 2012. Photomechanical Responses in *Drosophila* Photoreceptors. *Science*
1044   **338**:260–263. doi:10.1126/science.1222376
1045   Hateren J van, Srinivasan MV, Wait PB. 1990. Pattern recognition in bees: orientation
1046   discrimination. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and*
1047   *Behavioral Physiology* **167**:649–654.
1048   Heisenberg M. 2003. Mushroom body memoir: from maps to models. *Nat Rev Neurosci* **4**:266–275.
1049   doi:10.1038/nrn1074
1050   Hertel H. 1983. Change of synapse frequency in certain photoreceptors of the honeybee after
1051   chromatic deprivation. *Journal of comparative physiology* **151**:477–482.
1052   Hertel H. 1982. The effect of spectral light deprivation on the spectral sensitivity of the honey bee.
1053   *Journal of comparative physiology* **147**:365–369.
1054   Hyvärinen A, Hurri J, Hoyer PO. 2009. Natural image statistics: A probabilistic approach to early
1055   computational vision. Springer Science & Business Media.
1056   James AC, Osorio D. 1996. Characterisation of columnar neurons and visual signal processing in
1057   the medulla of the locust optic lobe by system identification techniques. *Journal of Comparative*
1058   *Physiology A* **178**:183–199. doi:10.1007/BF00188161
1059   Juston R, Kerhuel L, Franceschini N, Viollet S. 2013. Hyperacute edge and bar detection in a
1060   bioinspired optical position sensing device. *IEEE/ASME Transactions on Mechatronics* **19**:1025–
1061   1034.
1062   Juusola M, Dau A, Song Z, Solanki N, Rien D, Jaciuch D, Dongre SA, Blanchard F, Polavieja GG
1063   de, Hardie RC, Takalo J. 2017. Microsaccadic sampling of moving image information provides
1064   *Drosophila* hyperacute vision. *eLife Sciences* **6**:e26117. doi:10.7554/eLife.26117
1065   Juusola M, Hardie RC. 2001a. Light adaptation in *Drosophila* photoreceptors: I. Response dynamics
1066   and signaling efficiency at 25 C. *The Journal of general physiology* **117**:3–25.
1067   Juusola M, Hardie RC. 2001b. Light adaptation in *Drosophila* photoreceptors: II. Rising temperature
1068   increases the bandwidth of reliable signaling. *The Journal of general physiology* **117**:27–42.
1069   Juusola M, Song Z. 2017. How a fly photoreceptor samples light information in time. *The Journal of*
1070   *Physiology* **595**:5427–5437.
1071   Juusola M, Takalo J, Kemppainen J, Haghighi KR, Scales B, McManus J, Bridges A, MaBouDi H,
1072   Chittka L. 2025. Theory of morphodynamic information processing: linking sensing to behaviour.
1073   *Vision Research* **227**:108537.
1074   Kemppainen J, Mansour N, Takalo J, Juusola M. 2022a. High-speed imaging of light-induced
1075   photoreceptor microsaccades in compound eyes. *Communications Biology* **5**:203.
1076   Kemppainen J, Scales B, Razban Haghighi K, Takalo J, Mansour N, McManus J, Leko G, Saari P,
1077   Hurcomb J, Antohi A, Suuronen J-P, Blanchard F, Hardie RC, Song Z, Hampton M, Eckermann M,
1078   Westermeier F, Frohn J, Hoekstra H, Lee C-H, Huttula M, Mokso R, Juusola M. 2022b. Binocular
1079   mirror–symmetric microsaccadic sampling enables *Drosophila* hyperacute 3D vision. *Proc Natl Acad*
1080   *Sci USA* **119**:e2109717119. doi:10.1073/pnas.2109717119
1081   Kuang X, Poletti M, Victor JD, Rucci M. 2012. Temporal encoding of spatial information during
1082   active visual fixation. *Curr Biol* **22**:510–514. doi:10.1016/j.cub.2012.01.050
1083   Land MF. 1999. Motion and vision: why animals move their eyes. *J Comp Physiol A* **185**:341–352.
1084   Land MF. 1973. Head Movement of Flies during Visually Guided Flight. *Nature* **243**:299–300.
1085   doi:10.1038/243299a0
1086   Land MF, Nilsson D-E. 2012. Animal eyes. Oxford University Press.

1087 Langridge KV, Wilke C, Riabinina O, Vorobyev M, de Ibarra NH. 2018. Approach direction prior to
1088 landing explains patterns of colour learning. *bioRxiv* 381210.
1089 Langridge KV, Wilke C, Riabinina O, Vorobyev M, Hempel de Ibarra N. 2021. Approach Direction
1090 Prior to Landing Explains Patterns of Colour Learning in Bees. *Frontiers in Physiology* **12**.
1091 Lehrer M, Collett TS. 1994. Approaching and departing bees learn different cues to the distance of a
1092 landmark. *J Comp Physiol A* **175**:171–177. doi:10.1007/BF00215113
1093 Li L, MaBouDi H, Egertová M, Elphick MR, Chittka L, Perry CJ. 2017. A possible structural correlate
1094 of learning performance on a colour discrimination task in the brain of the bumblebee. *Proc R Soc B*
1095 **284**:20171323. doi:10.1098/rspb.2017.1323
1096 Lin A, Yang R, Dorkenwald S, Matsliah A, Sterling AR, Schlegel P, Yu S, McKellar CE, Costa M,
1097 Eichler K. 2024. Network statistics of the whole-brain connectome of *Drosophila*. *Nature* **634**:153–
1098 165.
1099 Locatelli FF, Fernandez PC, Villareal F, Muezzinoglu K, Huerta R, Galizia CG, Smith BH. 2013.
1100 Nonassociative plasticity alters competitive interactions among mixture components in early
1101 olfactory processing. *Eur J Neurosci* **37**:63–79. doi:10.1111/ejn.12021
1102 Ma X, Hou X, Edgecombe GD, Strausfeld NJ. 2012. Complex brain and optic lobes in an early
1103 Cambrian arthropod. *Nature* **490**:258–261. doi:10.1038/nature11495
1104 MaBouDi H, Barron AB, Li S, Honkanen M, Loukola OJ, Peng F, Li W, Marshall JAR, Cope A,
1105 Vasilaki E, Solvi C. 2021a. Non-numerical strategies used by bees to solve numerical cognition
1106 tasks. *Proceedings of the Royal Society B: Biological Sciences* **288**:20202711.
1107 doi:10.1098/rspb.2020.2711
1108 MaBouDi H, Dona H, Gatto E, Loukola OJ, Buckley E, Onoufriou PD, Skorupski P, Chittka L. 2020a.
1109 Bumblebees use sequential scanning of countable items in visual patterns to solve numerosity
1110 tasks. *Integrative and Comparative Biology*.
1111 MaBouDi H, Marshall JA, Dearden N, Barron AB. 2023b. How honey bees make fast and accurate
1112 decisions. *eLife* **12**:e86176.
1113 MaBouDi H, Marshall JAR, Barron AB. 2020b. Honeybees solve a multi-comparison ranking task by
1114 probability matching. *Proceedings of the Royal Society B: Biological Sciences* **287**:20201525.
1115 doi:10.1098/rspb.2020.1525
1116 MaBouDi H, Roper M, Guiraud M, Marshall JAR, Chittka L. 2021b. Automated video tracking and
1117 flight analysis show how bumblebees solve a pattern discrimination task using active vision. *bioRxiv*
1118 2021.03.09.434580. doi:10.1101/2021.03.09.434580
1119 MaBouDi H, Shimazaki H, Amari S, Soltanian-Zadeh H. 2016. Representation of higher-order
1120 statistical structures in natural scenes via spatial phase distributions. *Vision research* **120**:61–73.
1121 MaBouDi H, Shimazaki H, Giurfa M, Chittka L. 2017a. Olfactory learning without the mushroom
1122 bodies: Spiking neural network models of the honeybee lateral antennal lobe tract reveal its
1123 capacities in odour memory tasks of varied complexities. *PLoS computational biology* **13**:e1005551.
1124 MaBouDi H, Shimazaki H, Giurfa M, Chittka L. 2017b. Olfactory learning without the mushroom
1125 bodies: spiking neural network models of the honeybee lateral antennal lobe tract reveal its
1126 capacities in odour memory tasks of varied complexities. *PLoS computational biology* **13**:e1005551.
1127 MaBouDi H, Solvi C, Chittka L. 2020c. Bumblebees learn a relational rule but switch to a win-
1128 stay/lose-switch heuristic after extensive training. *bioRxiv* 2020.05.08.085142.
1129 doi:10.1101/2020.05.08.085142
1130 Maddess T, Laughlin SB, Horridge GA, Levick WR. 1997. Adaptation of the motion-sensitive neuron
1131 H1 is generated locally and governed by contrast frequency. *Proceedings of the Royal Society of*
1132 *London Series B Biological Sciences* **225**:251–275. doi:10.1098/rspb.1985.0061
1133 Manoonpong P, Patanè L, Xiong X, Brodoline I, Dupeyroux J, Viollet S, Arena P, Serres JR. 2021.
1134 Insect-inspired robots: bridging biological and artificial systems. *Sensors* **21**:7609.
1135 Markram H, Lübke J, Frotscher M, Sakmann B. 1997. Regulation of synaptic efficacy by
1136 coincidence of postsynaptic APs and EPSPs. *Science* **275**:213–215.
1137 Matsumoto Y, Matsumoto C-S, Wakuda R, Ichihara S, Mizunami M. 2015. Roles of octopamine and
1138 dopamine in appetitive and aversive memory acquisition studied in olfactory conditioning of
1139 maxillary palpi extension response in crickets. *Front Behav Neurosci* **9**.
1140 doi:10.3389/fnbeh.2015.00230

1141 Meeks JP, Holy TE. 2008. Pavlov's moth: olfactory learning and spike timing–dependent plasticity.
1142 *Nature neuroscience* **11**:1126–1127.
1143 Menzel R. 2022. In Search for the Retrievable Memory Trace in an Insect Brain. *Front Syst*
1144 *Neurosci* **16**. doi:10.3389/fnsys.2022.876376
1145 Menzel R. 2012. The honeybee as a model for understanding the basis of cognition. *Nat Rev*
1146 *Neurosci* **13**:758–768. doi:10.1038/nrn3357
1147 Menzel R, Blakers M. 1976. Colour receptors in the bee eye — Morphology and spectral sensitivity.
1148 *J Comp Physiol* **108**:11–13. doi:10.1007/BF00625437
1149 Menzel R, Giurfa M. 2006. Dimensions of cognition in an insect, the honeybee. *Behav Cogn*
1150 *Neurosci Rev* **5**:24–40. doi:10.1177/1534582306289522
1151 Mohammad F, Mai Y, Ho J, Zhang X, Ott S, Stewart JC, Claridge-Chang A. 2024. Dopamine
1152 neurons that inform *Drosophila* olfactory memory have distinct, acute functions driving attraction and
1153 aversion. *PLOS Biology* **22**:e3002843. doi:10.1371/journal.pbio.3002843
1154 Näher T, Zhang Y, Pandinelli M, Fries P. 2023. Primate Saccade Rhythmicity. *bioRxiv* 2023–09.
1155 Nityananda V, Skorupski P, Chittka L. 2014. Can bees see at a glance? *J Exp Biol* **217**:1933–1939.
1156 doi:10.1242/jeb.101394
1157 Odenthal L, Doussot C, Meyer S, Bertrand OJN. 2021. Analysing Head-Thorax Choreography
1158 During Free-Flights in Bumblebees. *Front Behav Neurosci* **14**. doi:10.3389/fnbeh.2020.610029
1159 Oja E. 1982. Simplified neuron model as a principal component analyzer. *Journal of mathematical*
1160 *biology* **15**:267–273.
1161 Okada R, Rybak J, Manz G, Menzel R. 2007. Learning-Related Plasticity in PE1 and Other
1162 Mushroom Body-Extrinsic Neurons in the Honeybee Brain. *J Neurosci* **27**:11736–11747.
1163 doi:10.1523/JNEUROSCI.2216-07.2007
1164 Olshausen BA. 2003. Principles of image representation in visual cortex. *The visual neurosciences*
1165 **2**:1603–1615.
1166 Paulk AC, Dacks AM, Phillips-Portillo J, Fellous J-M, Gronenberg W. 2009. Visual Processing in the
1167 Central Bee Brain. *J Neurosci* **29**:9987–9999. doi:10.1523/JNEUROSCI.1325-09.2009
1168 Paulk AC, Gronenberg W. 2008. Higher order visual input to the mushroom bodies in the bee,
1169 Bombus impatiens. *Arthropod Structure & Development* **37**:443–458. doi:10.1016/j.asd.2008.03.002
1170 Paulk AC, Phillips-Portillo J, Dacks AM, Fellous J-M, Gronenberg W. 2008. The Processing of
1171 Color, Motion, and Stimulus Timing Are Anatomically Segregated in the Bumblebee Brain. *J*
1172 *Neurosci* **28**:6319–6332. doi:10.1523/JNEUROSCI.1196-08.2008
1173 Perry CJ, Barron AB. 2013. Neural mechanisms of reward in insects. *Annual review of entomology*
1174 **58**:543–562.
1175 Roper M, Fernando C, Chittka L. 2017. Insect Bio-inspired Neural Network Provides New Evidence
1176 on How Simple Feature Detectors Can Enable Complex Visual Generalization and Stimulus
1177 Location Invariance in the Miniature Brain of Honeybees. *PLOS Computational Biology*
1178 **13**:e1005333. doi:10.1371/journal.pcbi.1005333
1179 Rucci M, Iovin R, Poletti M, Santini F. 2007. Miniature eye movements enhance fine spatial detail.
1180 *Nature* **447**:852–855.
1181 Rucci M, Victor JD. 2015. The unsteady eye: an information-processing stage, not a bug. *Trends*
1182 *Neurosci* **38**:195–206. doi:10.1016/j.tins.2015.01.005
1183 Ruderman DL, Bialek W. 1994. Statistics of natural images: Scaling in the woods. *Phys Rev Lett*
1184 **73**:814–817. doi:10.1103/PhysRevLett.73.814
1185 Schlegel P, Yin Y, Bates AS, Dorkenwald S, Eichler K, Brooks P, Han DS, Gkantia M, Dos Santos
1186 M, Munnelly EJ. 2024. Whole-brain annotation and multi-connectome cell typing of *Drosophila*.
1187 *Nature* **634**:139–152.
1188 Schmalz F, El Jundi B, Rössler W, Strube-Bloss M. 2022. Categorizing visual information in
1189 subpopulations of honeybee mushroom body output neurons. *Frontiers in Physiology* **13**:866807.
1190 Schwaerzel M, Monastirioti M, Scholz H, Friggi-Grelin F, Birman S, Heisenberg M. 2003. Dopamine
1191 and octopamine differentiate between aversive and appetitive olfactory memories in *Drosophila*. *J*
1192 *Neurosci* **23**:10495–10502.
1193 Schwegmann A, Lindemann JP, Egelhaaf M. 2014. Temporal statistics of natural image sequences
1194 generated by movements with insect flight characteristics. *PLoS ONE* **9**:e110386.
1195 doi:10.1371/journal.pone.0110386

1196 Seelig JD, Jayaraman V. 2013. Feature detection and orientation tuning in the *Drosophila* central
1197 complex. *Nature* **503**:262–266. doi:10.1038/nature12601
1198 Selcho M. 2024. Octopamine in the mushroom body circuitry for learning and memory. *Learn Mem*
1199 **31**:a053839. doi:10.1101/lm.053839.123
1200 Serbe E, Meier M, Leonhardt A, Borst A. 2016. Comprehensive characterization of the major
1201 presynaptic elements to the *Drosophila* OFF motion detector. *Neuron* **89**:829–841.
1202 Serres JR, Viollet S. 2018. Insect-inspired vision for autonomous vehicles. *Current opinion in insect*
1203 *science* **30**:46–51.
1204 Severance E, Washburn MF. 1907. The loss of associative power in words after long fixation. *The*
1205 *American Journal of Psychology* 182–186.
1206 Simoncelli EP, Olshausen BA. 2001. Natural image statistics and neural representation. *Annual*
1207 *review of neuroscience* **24**:1193–1216.
1208 Skorupski P, Döring TF, Chittka L. 2007. Photoreceptor spectral sensitivity in island and mainland
1209 populations of the bumblebee, Bombus terrestris. *Journal of Comparative Physiology A* **193**:485–
1210 494.
1211 Song S, Miller KD, Abbott LF. 2000. Competitive Hebbian learning through spike-timing-dependent
1212 synaptic plasticity. *Nature neuroscience* **3**:919–926.
1213 Song Z, Juusola M. 2014. Refractory Sampling Links Efficiency and Costs of Sensory Encoding to
1214 Stimulus Statistics. *Journal of Neuroscience* **34**:7216–7237. doi:10.1523/JNEUROSCI.4463-
1215 13.2014
1216 Spaethe J, Briscoe AD. 2004. Early duplication and functional diversification of the opsin gene
1217 family in insects. *Molecular Biology and Evolution* **21**:1583–1594.
1218 Spaethe J, Chittka L. 2003a. Interindividual variation of eye optics and single object resolution in
1219 bumblebees. *Journal of Experimental Biology* **206**:3447–3453.
1220 Spaethe J, Chittka L. 2003b. Interindividual variation of eye optics and single object resolution in
1221 bumblebees. *Journal of Experimental Biology* **206**:3447–3453.
1222 Spaethe J, Tautz J, Chittka L. 2001. Visual constraints in foraging bumblebees: flower size and
1223 color affect search time and flight behavior. *Proceedings of the National Academy of Sciences*
1224 **98**:3898–3903.
1225 Srinivasan MV. 2010. Honey bees as a model for vision, perception, and cognition. *Annual review of*
1226 *entomology* **55**:267–284.
1227 Srinivasan MV. 1994. Pattern recognition in the honeybee: recent progress. *Journal of Insect*
1228 *Physiology* **40**:183–194.
1229 Srinivasan MV, Lehrer M. 1988. Spatial acuity of honeybee vision and its spectral properties. *J*
1230 *Comp Physiol* **162**:159–172. doi:10.1007/BF00606081
1231 Stach S, Benard J, Giurfa M. 2004. Local-feature assembling in visual pattern recognition and
1232 generalization in honeybees. *Nature* **429**:758–761. doi:10.1038/nature02594
1233 Stach S, Giurfa M. 2005. The influence of training length on generalization of visual feature
1234 assemblies in honeybees. *Behavioural Brain Research* **161**:8–17. doi:10.1016/j.bbr.2005.02.008
1235 Streinzer M, Brockmann A, Nagaraja N, Spaethe J. 2013. Sex and Caste-Specific Variation in
1236 Compound Eye Morphology of Five Honeybee Species. *PLOS ONE* **8**:e57702.
1237 doi:10.1371/journal.pone.0057702
1238 Strube-Bloss MF, Nawrot MP, Menzel R. 2011. Mushroom Body Output Neurons Encode Odor–
1239 Reward Associations. *J Neurosci* **31**:3129–3140. doi:10.1523/JNEUROSCI.2583-10.2011
1240 Szyszka P, Ditzen M, Galkin A, Galizia CG, Menzel R. 2005. Sparsening and temporal sharpening
1241 of olfactory representations in the honeybee mushroom bodies. *J Neurophysiol* **94**:3303–3313.
1242 doi:10.1152/jn.00397.2005
1243 Turner CH. 1911. Experiments on Pattern-Vision of the Honey Bee. *Biological Bulletin* **21**:249–264.
1244 doi:10.2307/1536017
1245 Van Hateren JH. 1997. Processing of natural time series of intensities by the visual system of the
1246 blowfly. *Vision research* **37**:3407–3416.
1247 van Hateren JH. 1992. Theoretical predictions of spatiotemporal receptive fields of fly LMCs, and
1248 experimental validation. *Journal of Comparative Physiology A* **171**:157–170.
1249 Varella TT, Takahashi DY, Ghazanfar AA. 2024. Active sampling as an information seeking strategy
1250 in primate vocal interactions. *Communications Biology* **7**:1098.

1251 Vasas V, Peng F, MaBouDi H, Chittka L. 2019. Randomly weighted receptor inputs can explain the
1252 large diversity of colour-coding neurons in the bee visual system. *Scientific reports* **9**:1–13.
1253 Vetter RS, Visscher PK. n.d. Influence of Age on Antennal Response of Male Honey Bees, *Apis*
1254 *mellifera*, to Queen Mandibular Pheromone and Alarm Pheromone Component. *J Chem Ecol*
1255 **23**:1867–1880. doi:10.1023/B:JOEC.0000006456.90528.94
1256 Viollet S, Franceschini N. 2010. A hyperacute optical position sensor based on biomimetic retinal
1257 micro-scanning. *Sensors and Actuators A: Physical* **160**:60–68.
1258 Vogels TP, Sprekeler H, Zenke F, Clopath C, Gerstner W. 2011. Inhibitory Plasticity Balances
1259 Excitation and Inhibition in Sensory Pathways and Memory Networks. *Science* **334**:1569–1573.
1260 doi:10.1126/science.1211095
1261 Von Frisch K. 1914. Der farbensinn und formensinn der biene. Рипол Классик.
1262 Washburn MF. 1926. The animal mind: A text-book of comparative psychology. Macmillan.
1263 Washburn MF. 1916. Movement and mental imagery: Outlines of a motor theory of the complexer
1264 mental processes. Houghton Mifflin.
1265 Webb B. 2020. Robots with insect brains. *Science* **368**:244–245.
1266 Wehner R. 1967. Pattern recognition in bees. *Nature* **215**:1244–1248.
1267 Werner A, Stürzl W, Zanker J. 2016. Object Recognition in Flight: How Do Bees Distinguish
1268 between 3D Shapes? *PLoS ONE* **11**:e0147106. doi:10.1371/journal.pone.0147106
1269 Willmore B, Tolhurst DJ. 2001. Characterizing the sparsenessof neural codes. *Network:*
1270 *Computation in Neural Systems* **12**:255.
1271 Yang E-C, Maddess T. 1997. Orientation-sensitive Neurons in the Brain of the Honey Bee (*Apis*
1272 *mellifera*). *Journal of Insect Physiology* **43**:329–336. doi:10.1016/S0022-1910(96)00111-4
1273 Yarbus AL. 2013. Eye movements and vision. Springer.
1274 Zhang LI, Tao HW, Holt CE, Harris WA, Poo M. 1998. A critical window for cooperation and
1275 competition among developing retinotectal synapses. *Nature* **395**:37–44.
1276 Zhang SW, Horridge GA. 1992. Pattern recognition in bees: size of regions in spatial layout.
1277 *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* **337**:65–
1278 71.
1279 Zheng L, de Polavieja GG, Wolfram V, Asyali MH, Hardie RC, Juusola M. 2006. Feedback Network
1280 Controls Photoreceptor Output at the Layer of First Visual Synapses in *Drosophila*. *Journal of*
1281 *General Physiology* **127**:495–510. doi:10.1085/jgp.200509470
1282 Zimmermann MJ, Nevala NE, Yoshimatsu T, Osorio D, Nilsson D-E, Berens P, Baden T. 2018.
1283 Zebrafish differentially process color across visual space to match natural scenes. *Current Biology*
1284 **28**.
1285