Looking down

Looking down: a model for visual route following in flying insects

Bioinspiration & Biomimetics

OPEN ACCESS

PAPER

Looking down: a model for visual route following in flying insects

J Stankiewicz and B Webb*

School of Informatics, University of Edinburgh, 10 Crichton Street, Edinburgh EH8 9AB, United Kingdom
* Author to whom any correspondence should be addressed.

E-mail: B.Webb@ed.ac.uk

RECEIVED
1 April 2021

REVISED
25 May 2021

ACCEPTED FOR PUBLICATION
9 July 2021

PUBLISHED
4 August 2021

Keywords: insect navigation, biorobotics, visual route following, UAV, honey bee

Supplementary material for this article is available online

Abstract

Insect visual navigation is often assumed to depend on panoramic views of the horizon, and how these change as the animal moves. However, it is known that honey bees can visually navigate in flat, open meadows where visual information at the horizon is minimal, or would remain relatively constant across a wide range of positions. In this paper we hypothesise that these animals can navigate using view memories of the ground. We find that in natural scenes, low resolution views from an aerial perspective of ostensibly self-similar terrain (e.g. within a field of grass) provide surprisingly robust descriptors of precise spatial locations. We propose a new visual route following approach that makes use of transverse oscillations to centre a flight path along a sequence of learned views of the ground. We deploy this model on an autonomous quadcopter and demonstrate that it provides robust performance in the real world on journeys of up to 30 m. The success of our method is contingent on a robust view matching process which can evaluate the familiarity of a view with a degree of translational invariance. We show that a previously developed wavelet based bandpass orientated filter approach fits these requirements well, exhibiting double the catchment area of standard approaches. Using a realistic simulation package, we evaluate the robustness of our approach to variations in heading direction and aircraft height between inbound and outbound journeys. We also demonstrate that our approach can operate using a vision system with a biologically relevant visual acuity and viewing direction.

1. Introduction

Bees and wasps use a combination of path integration (PI) and visual memory to navigate in their local environments [76]. A common experimental paradigm to test the visual memory navigation system in isolation is to displace an animal from its nest [11, 37, 74]. Animals that have not previously departed their nests become lost after such a manipulation while those with foraging experience can often find their way home after displacements of several hundred meters [11, 35]. In [11], bees that are displaced to previously visited locations return to their nest quickly and directly. By contrast, bees that are transported to new locations embark on long convoluted routes which tend to straighten out when their flight paths coincide with familiar ground. This appears to be a feature of flat landscapes, whereas in environments with more clutter, both bees [44] and ants [42, 86] show an ability to travel towards their nests from new locations, presumably guided by prominent landmarks on the horizon.

In recent years a popular hypothesis for visual memory navigation has emerged, based principally on models of desert ant behaviour: that routes are learnt by periodically storing view snapshots [29, 34]. Modelling approaches have bolstered this theory by demonstrating that an agent using a simple strategy of moving in the direction of greatest visual familiarity can traverse previously learned routes [3]. This process is known as *visual route following*. Rotational scanning allows the agent to compare view familiarity at different heading angles, forming a *rotational image difference function* [85]. While desert ants are known to stop and rotate their body or head [81], this *scanning* action is sporadic and infrequent

in nature. By contrast, contemporary visual route following models require agents to scan periodically and often. Despite these extra scans, visual homing models are less robust than desert ants, sometimes getting lost several times along a route [2]. Furthermore, the bulk of modelling work has been conducted in basic simulation environments, which have limited textural information, static lighting conditions, simplistic agent dynamics and perfect orientation control. In a handful of exceptions to this [20, 27, 28, 47, 72], visual route following models have been deployed on robotic platforms. However, they have yet to be tested over distances of greater than 10 m and in the range of environments that Hymenoptera inhabit.

Visual route following approaches have also been applied to an aerial context [18, 41, 47, 62, 72]. In [41], the mushroom body model from [2] was deployed on a simulated agent that moves at a fixed offset from the ground. The authors demonstrate that visual route following can operate in digitally reconstructed habitat of a honey bee experimentation site. In [18], a simulated agent is used to consider the challenge of extending visual homing to a three-dimensional space. Here it is demonstrated that helical flight paths could be used to scan in the vertical and horizontal dimensions simultaneously. The authors in [47] use a robotic gantry to demonstrate that visual homing can be tolerant to height offsets between the inbound and outbound routes without the use of special flight paths. In the above studies a forwards facing camera system has been adopted. Previous work has shown that a downward view can be a useful perspective for navigation [12, 13, 21], and it has been speculated that the holonomic nature of insect flight (i.e. a honey bee's heading direction is often offset from its velocity direction) provides a good basis for translational visual scanning [9]. However, these visual route following elements have yet to be combined in a robotic agent with the purpose of addressing biologically relevant questions.

In this paper, we explore the effectiveness of a navigational strategy based on translational visual scanning of the downward view. This imposes the constraint that the view matching procedure should have a degree of translational invariance to increase the useful volumetric range of a given view memory (often referred to as the '*catchment area*'), a key evaluation metric for visual homing algorithms [45, 85]. We therefore introduce an alternative to whole-image pixel-wise mean squared error (MSE), which has been the default method of choice for previous navigation models based on visual matching. Despite its convenience and simplicity, the MSE method makes little reference to known processing steps in biological vision. Orientated edge detectors have been previously proposed as a likely basis for visual processing in natural scenes in insects [4, 8]. Isotropic bandpass filters which provide this functionality are known to feature in low level mammalian vision processing [5, 25, 26] and have also been detected in the brains of honey bees [82] and other flying insects [56].

Modelling approaches have shown that a population of 42 orientated bandpass *ring neurons* in fruit flies [56] can discriminate between locations in natural scenes [15] and perform general pattern recognition tasks [16]. Coarse receptive fields seem to be an advantageous property for comparing natural scenes with this network [78]. While less is known about properties of ring neurons in the bee brain, the bandpass neurons from [82] have been modelled and shown to be a useful processing step for the location invariant identification of patterns of orientated bars [53].

In recent years, studies have shown that *Haar* like wavelet features can improve view discrimination in the task of visual homing while reducing memory requirements [31, 39]. Here, we adopt a bandpass filtering approach, using a computational method that was originally inspired by the orientated bandpass filter properties of V1 neurons in vertebrates [48]. The complex wavelet structural similarity (CWSSIM) image comparison method was chosen because it has been shown to have a greater level of translational invariance than whole image comparison metrics such as the MSE [54, 75, 87]. In the following sections we first provide an overview of how this image matching approach is integrated into a behavioural control schema for route following, and then describe in detail its implementation and evaluation on a flying robotic platform.

2. Concept overview

The method for navigation that we propose in this paper is illustrated in figure 1. In common with a number of previous models [2, 3, 28] we assume that during a learning phase, under guidance of other navigational mechanisms, the insect captures a sequence of visual snapshots as it moves along a route towards a goal (see figure 1(a)). Then, on subsequent journeys, it uses image comparison to evaluate the familiarity of its current view, and thus determine its next action so as to arrive at the goal. In a departure from most previous work, we assume the insect utilises a narrow field of view directed towards the ground to capture snapshots, rather than panoramic views of the horizon (see figure 1(b)). We also propose the use of isotropic orientated bandpass filtering to obtain greater translation invariance in image matching. This allows robust familiarity detection of ground-directed views when the insect enters a catchment area near the same location. The set of catchment areas thus forms a breadcrumb trail towards the goal, or with sufficient overlap, a continuous ridge, as shown in the familiarity manifold in figure 1(c). We propose that if a honey bee returning to its nest intercepts this familiarity ridge, it can follow the ridge home by using transverse oscillations modulated by the view

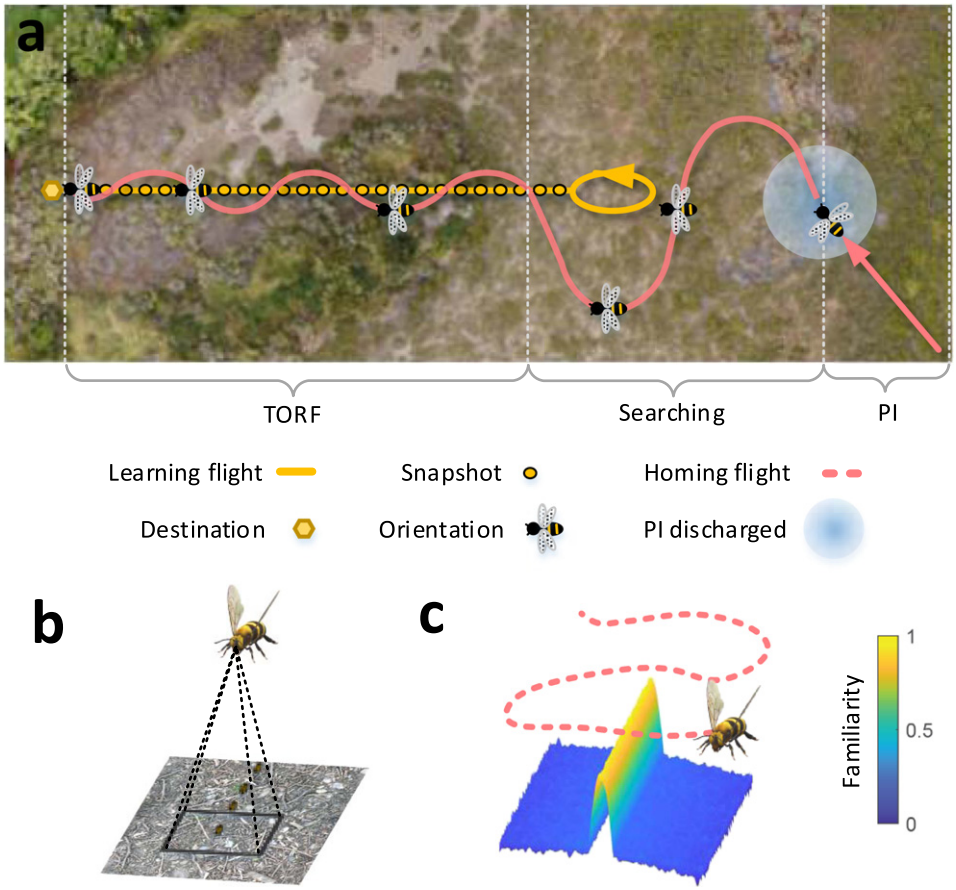

Figure 1. Concept overview. (a) Schematic of our approach. View snapshots are acquired on the inbound portion of the learning flight. Following a long foraging bout, the PI circuit is discharged before the animal reaches its destination. A searching phase is then conducted until a familiar ground view is obtained, and from this point the animal uses TORF to reach its goal. The background aerial image was modified from an OpenAerialMap image. Reproduced from [1]. CC BY 4.0. (b) Ventral views are acquired at regular intervals on the inbound portion of the learning flight. (c) The search space of an effective view matching scheme features a prominent ridge of familiarity as represented by the familiarity manifold. The agent centres its oscillatory flight path on the familiarity ridge in order to stay in contact with this guide.

familiarity to remain centred on the ridge. We refer to this process as transverse oscillating route following (TORF).

We suggest this navigational mechanism is particularly appropriate for a honey bee trying to relocate a previously visited destination in a relatively open environment, complementing PI which facilitates long excursions but accumulates error. Learning flights in bees are relatively straight [10, 43] which increases the likelihood that a returning insect would intercept the familiarity ridge, either on the inbound route itself or during a systematic search pattern [51]. Our approach also exploits the holonomic flight patterns observed in many insects [19, 46, 67], as it assumes transverse sweeps can be made while maintaining a relatively consistent heading direction. Finally, as discussed above, there is some biological precedence for the visual preprocessing method we propose [82], but we also evaluate it against some alternatives, under the constraints imposed by the TORF method and by deployment on a quadcopter.

3. Methods

3.1. View matching pipeline

The view matching pipeline is required to evaluate the familiarity of a current view based on previously memorised views. The TORF approach places two main requirements on this pipeline: (1) it must be accurate at discriminating between familiar and unfamiliar views (2) it must provide a degree of translational invariance, that is, views acquired near to a snapshot location should be classified as familiar.

We used the processes outlined in [6, 48, 49] to implement the *complex wavelet* transform with a filter mask arrangement (figure S1(b) https://stacks.iop.org/BB/16/055007/mmedia). This process requires the algorithm designer to select a predefined number of levels and orientations. The number of levels defines how many times the image is downsampled and the number of orientations specifies the preferred directions of the complex wavelet isotropic analysing function (see figure

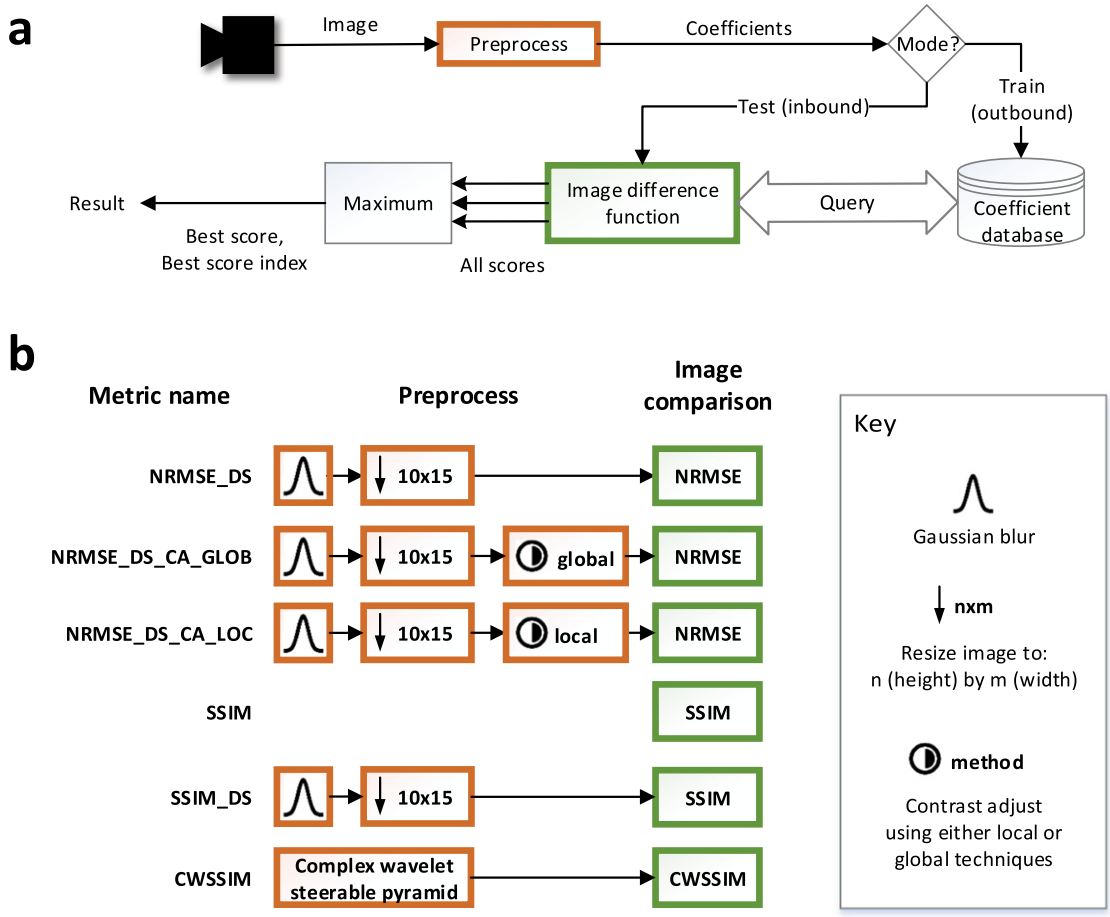

Figure 2. (a) The image processing pipeline. After preprocessing, images are either stored (on 'train' routes) or compared to the database (on 'test' routes). (b) Overview of six image comparison pipelines that are tested in this work.

S1(d) for a visual impression). We used five levels to provide a coarse insect eye resolution (equivalent to approximately 2° acuity). We used two orientations which equates to subbands that respond strongly to edges in the horizontal and vertical directions. In order to quantify the similarity of two views, the normalised correlation was applied to coefficients on the lowest level of the *complex wavelet* transform in both the horizontal and vertical subbands. This results in a familiarity score for each subband. The CWSSIM score is the mean of these two values. Full implementation details are included in section S 1.2.1.1.

We created a flexible view matching pipeline framework (see figure 2(a)) that enables the comparison of CWSSIM with state-of-the-art MSE and structural similarity (SSIM) approaches. We added Gaussian blurring, downsampling and contrast adjustment operations to the MSE pipelines as these methods have previously been shown to be beneficial for the task of view matching. An overview of the configurations we trialled are summarised in figure 2(b) and further implementation details are provided in section S 1.2. The framework features two modes of operation, test and train. In both modes, incoming images are optionally preprocessed. In train mode, images are sequentially indexed and stored in a database in the processed state. In test mode, new images are assessed for familiarity against each of the training images. The pipeline outputs the maximum familiarity score and its corresponding database index. Although not necessary for following the route, this index transpires to be a useful piece of information for informing the agent to slow down towards the end of the route as described later in section 3.2.1.

3.2. Transverse oscillating route following

The TORF route following procedure subscribes to the familiarity output of the view matching pipeline and uses this signal to locate and traverse learned routes. A top down view of a trajectory generated by TORF is included in figure 3(a).

3.2.1. Implementation details

3.2.1.1. Control The agent is velocity controlled in the horizontal (XY) plane and maintains a fixed height above the ground via position control in the Z axis. The horizontal control law independently commands a longitudinal (V_x) and a transverse (V_y) component of velocity (relative to the agent's body frame) which creates an angle of attack (α) with respect to the agent's heading (see figure 3). Here 70° was used

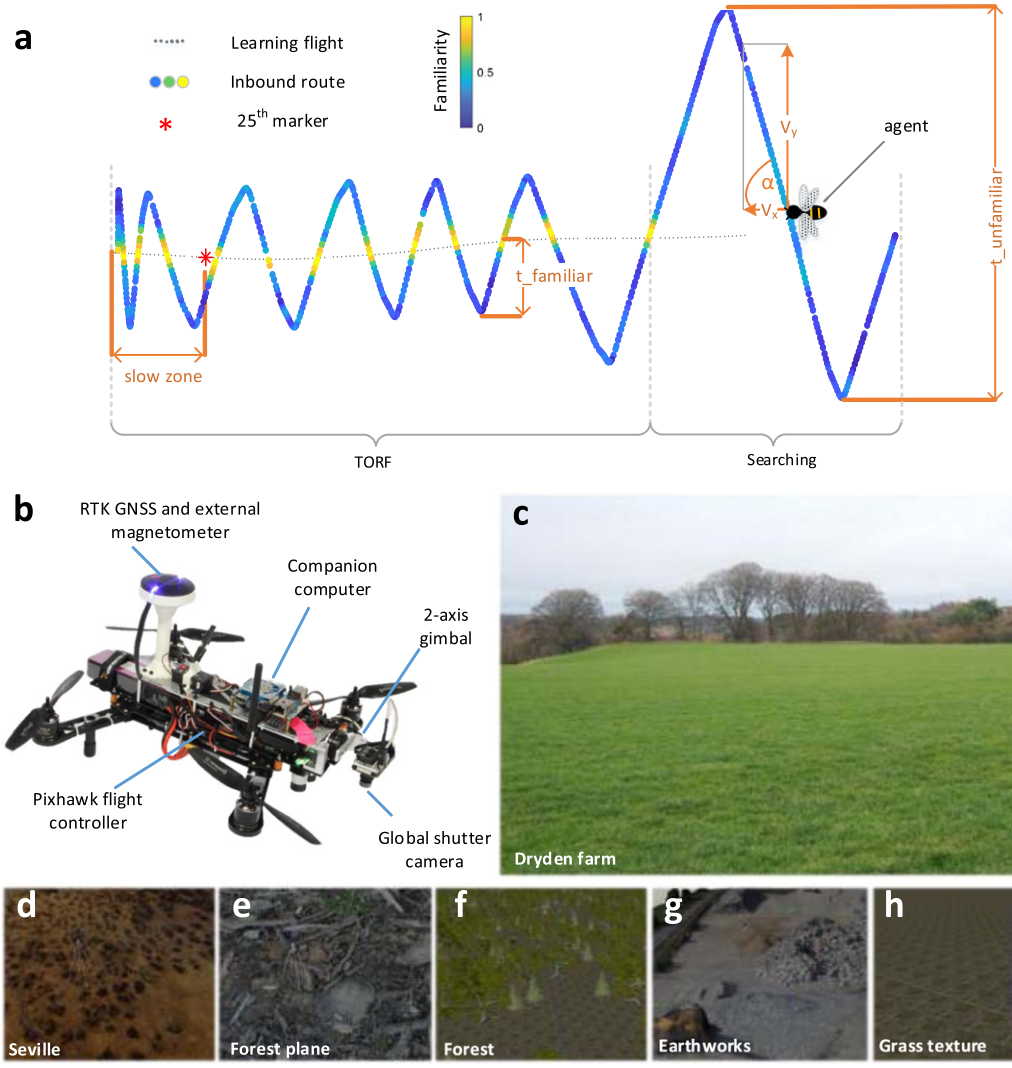

Figure 3. TORF schematic, biorobotic hardware and test locations. (a) Top down view from a visual homing run that has been annotated with the key parameters for tuning the visual route following concept. (b) Photograph of the biorobot, annotated with key components. (c)–(h) Photograph (c) and renderings (d)–(h) of the real and simulated test locations respectively. Images are labelled with their associated environment name as referenced throughout this paper.

for α unless otherwise stated. The default magnitude of commanded horizontal velocity of the aircraft is 1 ms^{-1}.

The agent's heading setpoint is constant throughout the inbound portion of the learning flight and the inbound route (equal in both cases). To date, the heading (yaw angle) of insects returning to their nests has only been recorded over short ranges [67]. This fixed heading strategy is therefore entirely speculative. However, insects are capable of perceiving their heading in the global coordinate system with a projected accuracy of $\pm 2^{\circ}$ [22]. This strategy is therefore possible within the confines of honey bee physiology.

The oscillatory motion of the trajectory is produced by inverting the sign of V_y forming a transverse motion *switchback*. The segment of motion between consecutive switchback events is referred to as a *sidesweep*.

3.2.1.2. Searching pattern A simple searching pattern enables the biorobot to locate the familiarity ridge. If the familiarity signal has not exceeded the familiarity threshold *t_unfamiliar* (10 s) after the previous switchback, a new switchback event is triggered. Note that this strategy was implemented to minimise time spent searching during TORF trials, rather than as a biologically informed mechanism.

3.2.1.3. Homing via TORF During TORF the switchback timer, *t_familiar* (1 s), is triggered when a predefined familiarity threshold (0.82 by default) is exceeded. A switchback event is initiated once *t_familiar* elapses. Note that once *t_familiar* is activated, if the highest familiarity score on that sidesweep is exceeded, *t_familiar* is reset such that the oscillation is centered on the most familiar location.

3.2.1.4. Terminating TORF In addition to the best score of each sidesweep, the index of the most familiar training image is returned. This value can be used as a quasimetric that describes the distance to the nest. When this index falls below a threshold

value (25 by default) V_x is proportionally reduced causing α to increase. If the matched image index falls below six, the final sidesweep is initiated. On this final sidesweep, the trial is terminated once the familiarity threshold is exceeded or a secondary timer elapses. It is assumed that the insect would be in visual contact with its target at this stage causing a separate short range visual homing behaviour to initialise (e.g. [67]). Note that while there is some evidence that disrupting the sequence of views experienced by an insect can impact its homing behaviour [55], the ability to index image sequences (or more loosely, to distinguish those images that were stored nearest to the nest) is purely speculative.

3.2.1.5. Learning flight Because of the directed nature of honey bee vision we assume that view memories are acquired on the inbound portion of their learning flights. However, for all trials in all sections other than section 4.2.3 we rotated outbound images by 180° so that they could be used as view memories on the inbound route. This was purely to reduce the duration of each test.

3.3. The biorobot

A quadcopter was used to deploy the behavioural models presented in this work. A photograph of the biorobot is included in figure 3(b), its construction and software development is described in detail in [58]. Of note, we made use of the PX4 flight controller ecosystem [14] which includes state estimation, flight control, and datalogging. It also includes realistic simulation functionality. We used a mix of real hardware and simulation trials when testing our concept. The hardware solution was used to verify that the system operates in the physical world, whereas the simulation environment was invaluable for rapid development, parameter tuning and performing large batches of tests [58].

The vision system comprised of a global shutter camera (MatrixVision Bluefox2, 200w) which was mounted on a custom built gimbal. We used a lens that provided a 42° horizontal field of view. Image processing took place on the onboard companion computer (Odroid XU4). The image sampling rate is set at 10 Hz, a somewhat arbitrary value but one that matches the sampling rate of the model in [2] and therefore memory capacity assumptions can be retained as well as being well within the refresh rate of a honey bee's vision system. Images were subscribed at a software adjusted resolution of 235 × 150 (downsampled from the native resolution of 752 × 480).

3.3.1. Biorobot test procedures

Evaluation flights conducted in this work were either *active* (where the agent was steered home via TORF) or *passive* (where the agent oscillated over the outbound route using an open-loop control strategy (i.e. without the aid of visual feedback)). Both flight types involved a predefined outbound route with a constant heading. At the end of this route the biorobot's commanded heading is offset by 180°. The biorobot is typically allowed to drift for a configurable amount of time before the homing procedure starts. This drift time provides some variance in the starting conditions. In an active flight, the biorobot is driven according to the procedure outlined in section 3.2. In a passive flight the aircraft sweeps over the outbound route but this is achieved with a predefined open loop flight pattern, by default switchbacks are triggered every ten seconds following a first switchback after five seconds.

3.3.2. Test location

Dryden farm was the test site used for the outdoor experiments. A photograph of this location is shown in figure 3(c) and satellite images of the site is included in figure S4. This site comprises a relatively flat field which is mown annually. We tested at various grass lengths and found that the vegetation height had no discernible effect on the performance of our model. Vertical land marks in this scene are restricted to the horizon at an effectively infinite distance from the agent. A series of simulated worlds were also used to evaluate our model. Renderings of these worlds are included in figures 3(d)–(h).

3.4. Evaluation procedures

3.4.1. View matching pipeline metrics

There are three principle objectives of the view matching procedure:

(a) All inbound/outbound route crossover points should be detected.
(b) There should be no false positive (FP) detections.
(c) Views should have wide catchment areas.

The metrics described in the following section were established to evaluate these criteria.

3.4.1.1. Detection of crossover points The quantity and location of instances where the inbound route bisects the outbound route are manually identified. Ten datapoints either side of the crossover points are labelled as the *valid detection region*.

3.4.1.2. Median outlier detection The MATLAB® *isoutlier* function was used to make median outlier detections (MODs) of spikes in the familiarity value, as a criteria for recognition. We set the outlier threshold at three scaled median absolute deviations above the median inbound familiarity value. Any datapoint that exceeds this threshold was counted as a true positive (TP) if it was in a valid detection region, otherwise it was counted as a FP.

3.4.1.3. Custom catchment area metric A shortcoming of the MOD metric is the arbitrary selection of the number of scaled median absolute deviations for the familiarity threshold. In order to ensure that this didn't unduly penalise a pipeline's score, we created a custom metric in which the familiarity threshold is

calculated by taking the highest familiarity value of any datapoint outside of the *valid detection* region. A second feature of our custom metric is that it only includes datapoints that monotonically decrease from the crossover point. This is a desirable property because it allows navigation algorithms to use gradient descent [85].

Our custom metric counts out from the crossover point centres (in both directions) until either: the familiarity score is no longer monotonically decreasing; or the familiarity score falls below a threshold set by the highest familiarity score occurring outside any valid detection region (see figure S11). The total number of valid detection regions with one or more valid datapoints in is referred to as the recognition rate (RR). The total count of valid datapoints inside all of the *valid detection* regions that obey the monotonic slope constraint is known as the familiarity count (FC). This value provides an indication of the catchment area of a view matching pipeline.

3.4.2. Route following metrics

3.4.2.1. Route success This metric is a boolean success/fail measure that evaluates the overall efficacy of TORF. A route trial is considered successful if the agent detects all crossover points along the route and if the trial is successfully terminated within 2 m of the destination. If these conditions are not met we consider the homing process on that trial to be a failure. In a set of trials the reliability of TORF can be measured by identifying the percentage of routes that were classified as successful.

3.4.2.2. Homing error The homing error is the Euclidean distance in the *XY* plane between the start and end locations of a trial in metres. Assuming that the route has been successful, this metric primarily quantifies how well the stopping procedure is working but also acts as an automatic way of detecting route failure (i.e. without the need to manually check if each crossover point has been detected).

3.4.2.3. Mean familiarity ridge width This metric quantifies the average width of the familiarity ridge for each of the crossover points correctly identified along a route. It allows a comparison of TORF in different environments or with different parameter configurations. The width of a given familiarity ridge is quantified using the sum of the transverse components of velocity at each point in the familiarity ridge, multiplied by the time interval between image samples (nominally 0.1 s). We then find the mean of all the familiarity ridge widths for each route under test. Note that we don't include the data from the terminating sidesweep as this only covers half of the route. Note also that since detection of crossover points requires a manual process, this metric is not suitable for tests involving high quantities of trials.

4. Results

4.1. Offline comparison of view matching pipelines

We began with an offline comparison of the six view matching pipelines specified in figure 2 with a dataset collected from the biorobot, generated according to the passive procedure outlined in section 3.3.1. For each pipeline, the score for every possible combination of outbound and inbound views was computed, resulting in a view correspondence matrix (figure S9). The maximum value in each correspondence matrix column will be the best match amongst all outbound views for that inbound view, and is logged as the familiarity score for the corresponding inbound route location. A spatial representation of the familiarity scores along the inbound route is included in figure 4 and the familiarity time series for each pipeline is shown in figure 5. The results of the view matching pipeline metrics (section 3.4.1), which summarise the overall effectiveness of each view matching pipeline for the complete inbound route, are shown in table 1. The same analysis was performed on a second route and the results are included in table ST1.

CWSSIM, NRMSE_DS_CA_GLOB, NRMSE_DS_CA_GLOB and SSIM_DS are each able to detect 100% of the crossover points. The failure of NRMSE_DS to locate any crossover points confirms that contrast adjustment is essential when using MSE with a real-world agent [66]. CWSSIM was the only method that didn't produce any FPs according to the MOD metric. This indicates that CWSSIM provides a higher degree of discriminability than the other methods.

Both the custom metric and the MOD method indicate that CWSSIM has around twice the catchment area of NRMSE_DS_CA_GLOB, the next best method. The increased catchment areas of CWSSIM are also perceptible in time series plots of figures 5 and S9. Previous work has shown that reducing the image resolution can help to increase catchment areas [79] with MSE approaches. We use a grid search approach to find the optimal combination of blur and image resolution for NRMSE_DS_CA_GLOB (see figure S10). This exercise confirmed that the resultant catchment area was still well below that of the CWSSIM. Note that the level of downsampling we have selected to apply to the MSE pipelines in this section matches the coefficient space of the sparsest level of the complex wavelet pyramid (10×15 pixels). Given the above analysis we conclude that CWSSIM outperforms MSE and standard SSIM approaches, at least in our chosen test environment. Altogether our results show that CWSSIM is an effective image comparison method for use in appearance based navigation pipelines. We therefore adopted it for the remainder of this paper.

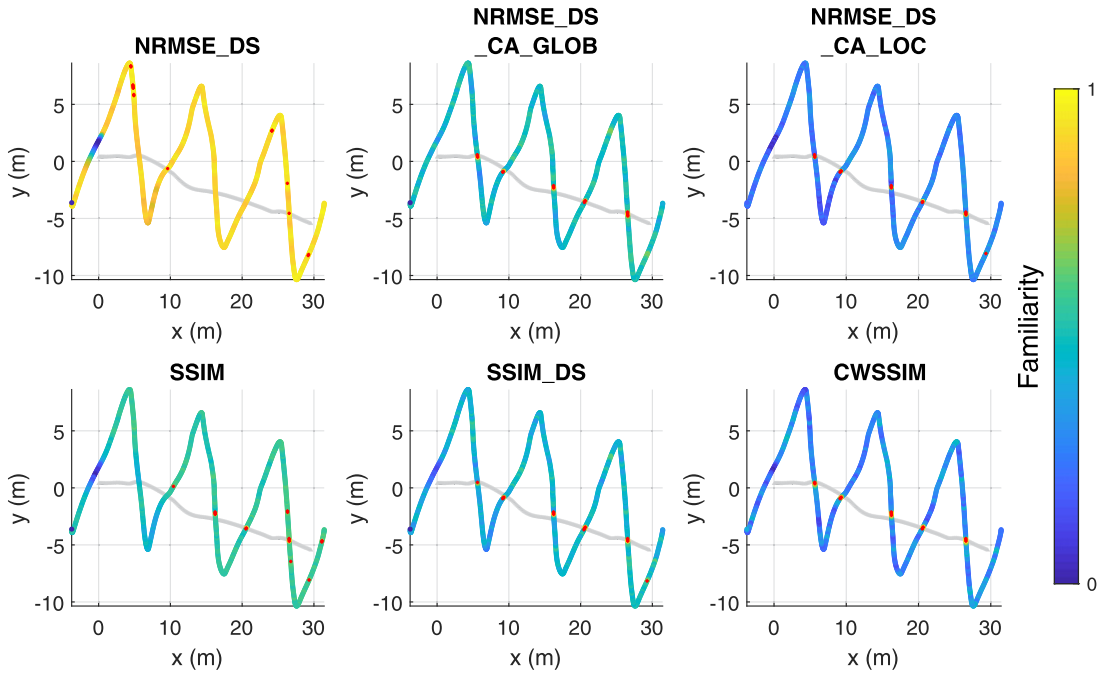

Figure 4. Plots of the familiarity scores that were generated offline for each pipeline on the same passive sidesweep flight performed at Dryden farm. The outbound route is the straight path with small grey markers. The inbound route is the zig-zagging formation with large filled markers. Each inbound datum is colour coded according to the pipeline under test. Red markers indicate the top 15 scores for each pipeline on each journey segment.

4.2. TORF experiments

4.2.1. Hardware implementation tests

The TORF homing algorithm outlined in section 2 was evaluated on the biorobot at Dryden farm using the active flight control procedure detailed in section 3.3.1. Following an outbound (learning) route of 20 m, the biorobot returned to within 1.5 m of the starting location under the guidance of TORF on five out of five test runs. A representative trajectory is included in figure 6(a), and the displacement of the landing site from the target location for each test run is shown in figure 6(b); the full set of results are plotted in S5.

On the day we collected the main set of results (07/12/2019), there were 16 kph wind gusts at the test location. The effect of these gusts can be observed in the outbound routes by uncommanded lateral translations (see (a) and (d) from figure S5). While this is not an in depth examination of environmental robustness, the results go some way towards demonstrating the tolerance of the system to windy conditions. This is particularly encouraging for the view matching subsystem which could be vulnerable to in-scene motion. Additional trajectories detailing the operation of TORF over some longer outbound routes are included in figure S6.

In some instances the biorobot overshot the route end (see figures S5(a) and (e) but in each case the procedure still terminates within 1.5 m of the goal because of the translational invariance of visual memories in the longitudinal direction (i.e. the overshoot is within the catchment area of the route). This indicates that the parameters generated in the simulation environment (section 4.2.2) did not translate into the real-world setting, possibly because of the previously mentioned wind gusts or the tendency of the simulated gimbal to loose alignment with gravity (investigated in S 1.7).

4.2.2. Simulated implementation tests

4.2.2.1. Batch testing We simulated TORF 64 times in the Seville world. A plot of the resultant trajectories is included in figure 6(c). Following a trial and error tuning process (adjusting t_familiar and α to get the default values outlined in sections 3.2.1.3 and 3.2.1.1 respectively), 100% of trials successfully tracked the ridge of familiarity. A histogram of the homing errors for each trial is included in figure 6(d). The worst case homing error is 80 cm which represents a significant reduction compared with PI circuits which are likely to produce errors in the order of tens of meters [59].

4.2.2.2. Effect of topography The TORF procedure was also simulated in five virtual worlds with a range of visual texture and height profiles (see renderings in figures 3(d)–(h) for a visual impression). As before, the biorobot was commanded to fly at 5 m above the ground. In one exception to this, the aircraft flew at an elevation of 10 m in the forest world in order to avoid collisions with the tree tops.

We first note that TORF is not suitable for use in simple virtual worlds featuring tiled texture maps (see the bottom row in figure S8). While it might seem obvious that using repeating patterns will cause problems in visual place recognition tasks, we include this

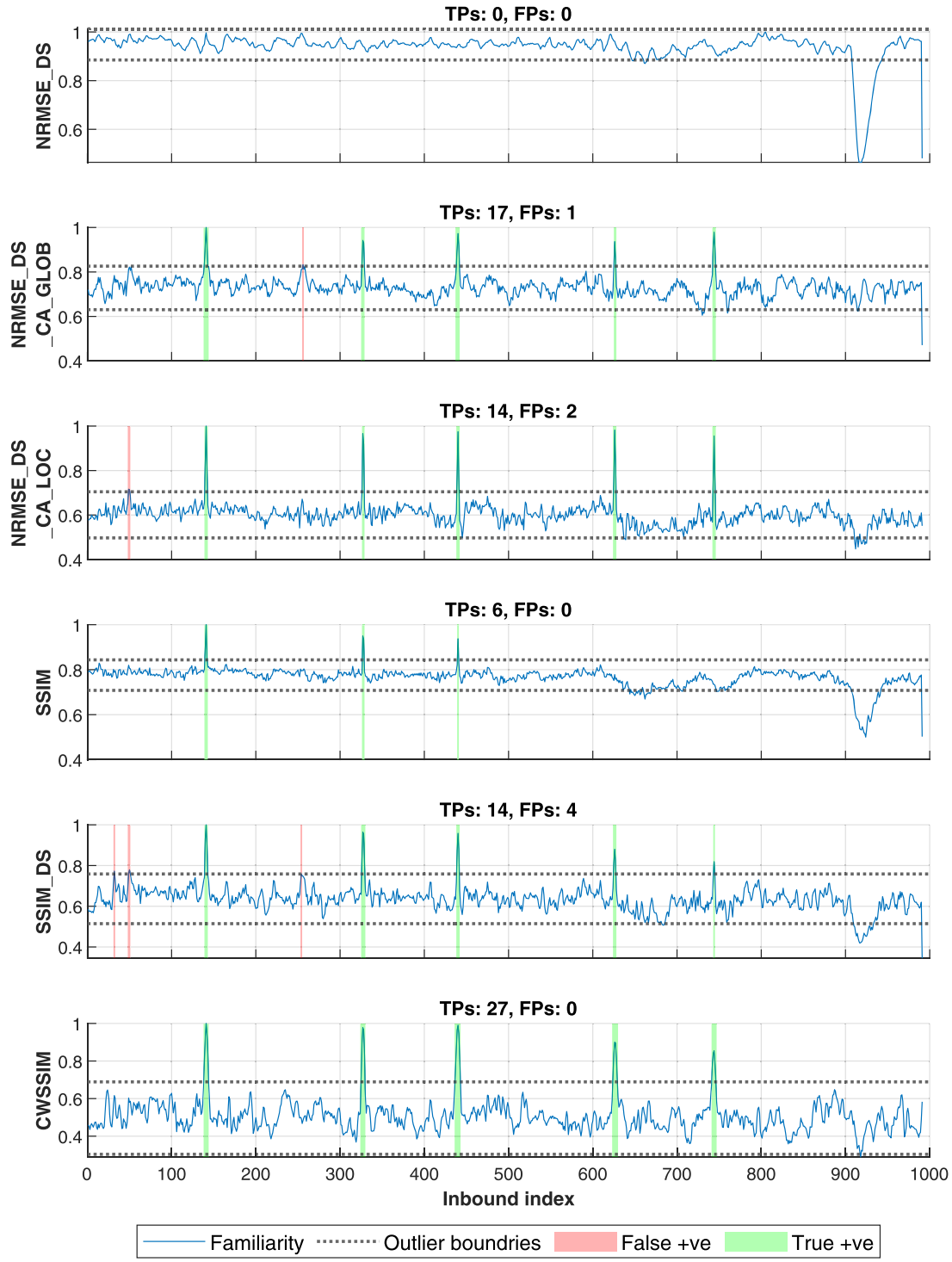

Figure 5. MOD applied to each pipeline's inbound familiarity score time series. Horizontal dashed lines indicate three median absolute deviations from the median value of that familiarity time series. Datapoints that have been labelled as a TP or FP are denoted by colour patches as indicated in the key.

result to highlight the benefits of using realistic simulation environments for vision behaviour modelling exercises. In all other worlds we found that TORF successfully detected 100% of encountered crossover points.

In figure 6(e), we found that the homing error was greater in the forest world than in other simulated worlds. The homing error provides a measure of how close the agent gets to its destination and therefore, how accurate the *route termination* procedure is. As outlined in section 4.2.2, our hypothesis is that the discrepancy between the forest and other simulated worlds can be explained by the fact that the agent has to fly at a greater height in the forest world. This is because misalignments between the camera's viewing direction and gravity can lead to localisation offsets with TORF. This effect is scaled by height

Table 1. Pipeline performances according to the offline evaluation process. The RR indicates how many crossover points were detected and the FC shows how many locations were correctly identified as familiar. High values in these metrics are therefore desirable. Low FP MOD scores indicate that a pipeline is reliable and a high number of TPs indicate that a pipeline has a wide catchment area.

	Custom metric			MOD	
Strategy	Crossover detection rate (%)	RR	FC	TP	FP
NRMSE_DS	0	0	0	0	0
NRMSE_DS_CA_GLOB	100	5	15	17	1
NRMSE_DS_CA_LOC	100	5	14	14	2
SSIM	60	3	7	6	0
SSIM_DS	100	5	14	14	4
CWSSIM	100	5	32	27	0

because for a given gimbal gravity offset angle, the horizontal error is related to the aircraft's height. In order to confirm and quantify this effect we systematically raised the commanded aircraft ground height in the *forest plane* world and monitored the effect this had on the homing error of a TORF mission. The results are summarised in figure S12.

We also compared the mean familiarity ridge width (section 3.4.2) across all simulated routes collected for this section. A boxplot of this metric across the different worlds is included in figure 6(f). In general there are only minor variations in the mean familiarity ridge widths between the different worlds, suggesting that TORF copes well with a range of terrain textures and depth profiles. There are a couple of variations to note. Firstly, the familiarity ridge appears to be wider in the real-world setting. This is somewhat surprising because the amplitude of the CWSSIM output is more consistent at crossover points in simulated trials. We attribute the better performance in the real world to the gimbal which was less jittery than the simulation environment. Amongst the simulated worlds the forest environment had the widest mean familiarity ridge width. While this can be explained by the aircraft flying 5 m higher in this environment, we note that the robustness of the model to the height variations caused by the tree canopy is an encouraging result.

4.2.2.3. Following curved routes In environments with physical obstacles the ability to follow curved routes could be required. We conducted tests with a sinusoidal learning flight path and simulated TORF's ability to recapitulate this route. As with other tests, the aircraft's heading remained constant, rather than tracking the profile of the route. We found that our homing procedure could follow routes with small deviations. However, to retain robust performance on routes with tight radii, the angle of attack of the transverse oscillations must be increased as illustrated with the sample routes in figures 6(g) and (h). Note that this process only works for turns with acute angles (relative the agent's heading). Furthermore, increasing the angle of attack has the adverse effect of increasing the overall distance travelled.

4.2.3. Robustness of transverse oscillating route following

So far we have demonstrated that the TORF procedure can be used to guide an aircraft along previously traversed routes. In this section we examine the impact of adjusting biologically significant parameters of the robotic system in order to gain an understanding of whether TORF could be realised with the anatomy and physiology of honey bees. Due to the requirement to accurately adjust the robotic system and perform several trials at each configuration we have performed this work in simulation. The general approach was to systematically adjust the parameter under test across a meaningful range of values. Ten trials were performed at each parameter setting and we used the Seville mesh throughout as it has the greatest degree of realism in terms of ground structure texture and depth profile.

4.2.3.1. Effect of view acuity In order for CWSSIM to be considered a biologically plausible snapshot mechanism for scene recognition in foraging insects it must be able to operate on images captured with the angular acuities commensurate with the compound eyes of honey bees. Since visual acuity can be modified by adjusting either the field of view or the image resolution we tried adjusting both of these independently while keeping the other variable constant (figures 7(a) and (b)).

We first fixed the image size to 15×10 pixels and adjusted the simulated camera's field of view so that we could try out different visual acuities. As shown in figure 7(a), the scheme is tolerant to acuities as low as 11°. This is the lowest angular acuity that can be generated with a field of view lower than 180° (i.e. without resorting to special optics). This indicates that the regions in the compound eye's of honey bees with very low acuity, such as the ventral section, can provide useful information for navigation purposes.

To test the effect of changing the resolution we first fixed the field of view of the camera at $42^\circ \times 42^\circ$. We started with an image resolution of 11×11 pixels and decremented this by a row and column of pixels between trial sets. The results are plotted in figure 7(b). We found that FPs were produced by

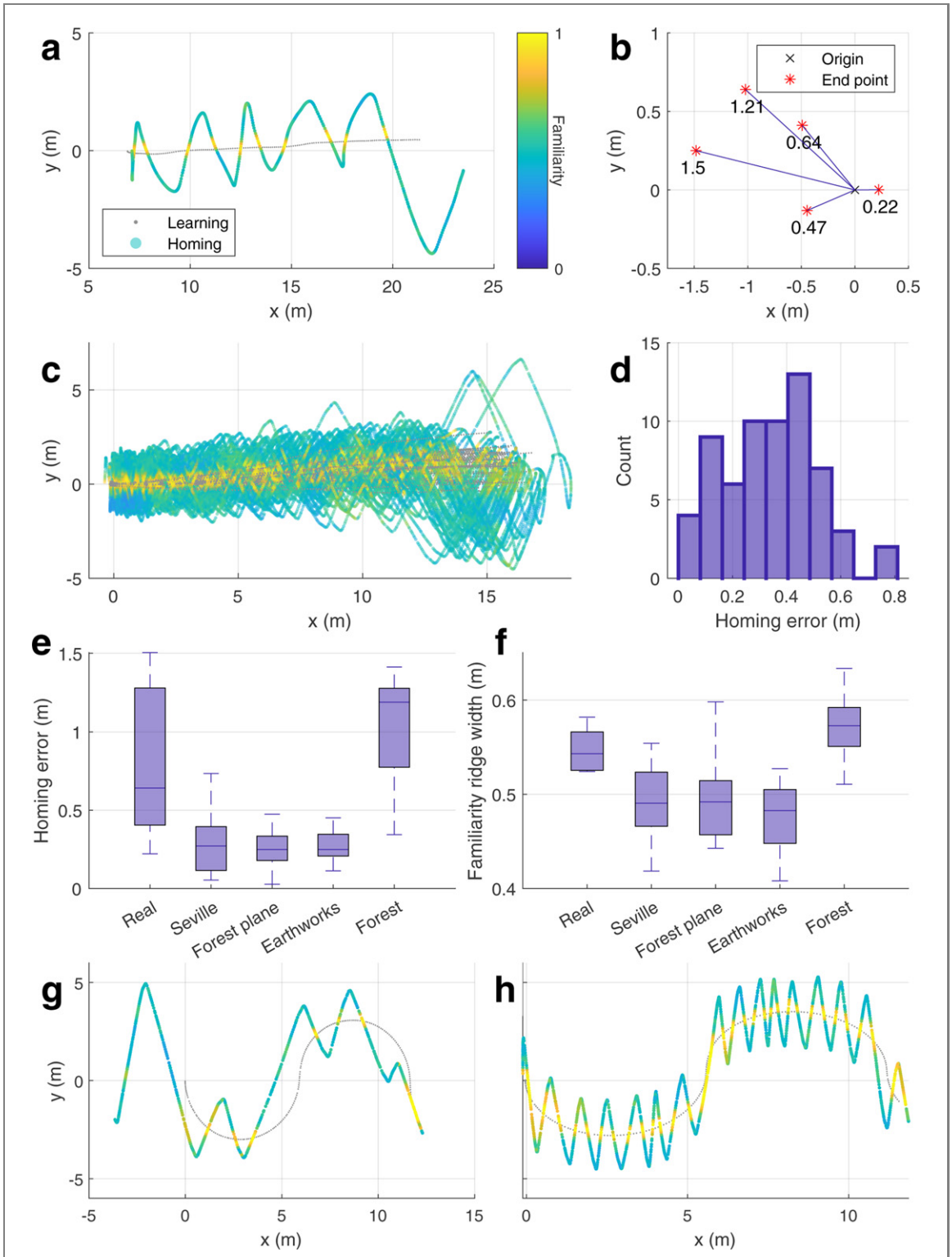

Figure 6. Efficacy of TORF. Note that the axes in panels (a), (c), (g) and (h) have an unequal aspect ratio but the axis grid tick interval is consistent for each plot. (a) Top down plot of a representative TORF trial with the biorobot. The homing route data are coloured according to their associated familiarity (see colour bar). (b) Summary of the landing sites of the five 20 m routes performed on the biorobot. There is a tendency to overshoot in the *x* dimension. (c) Top down plot of 64 trials in the Seville simulation environment (same colour scheme as (a)). (d) Histogram of homing errors of the trials plotted in (c). (e)–(f) Boxplot showing *homing error* and *familiarity ridge width* distributions respectively for different environments ($n = 5$ and $n = 20$ for real and simulated environments respectively). (g)–(h) Plots of the TORF homing procedure operating on curved outbound routes featuring a tight radii using the default transverse angle of attack ((g): $\alpha = 70$) and an increased transverse angle of attack ((h): $\alpha = 85$) (same colour scheme as (a).

the CWSSIM pipeline when image resolutions below 9×9 pixels were used, this resulted in the agent failing to home. At 60° below the horizon, the visual acuity is approximately 3° [71]. A 9×9 patch in this region would therefore occupy a $27° \times 27°$ field of view. We conclude that CWSSIM could be achieved with the resolution and acuity of the fronto-ventral portion of a honey bee's vision system.

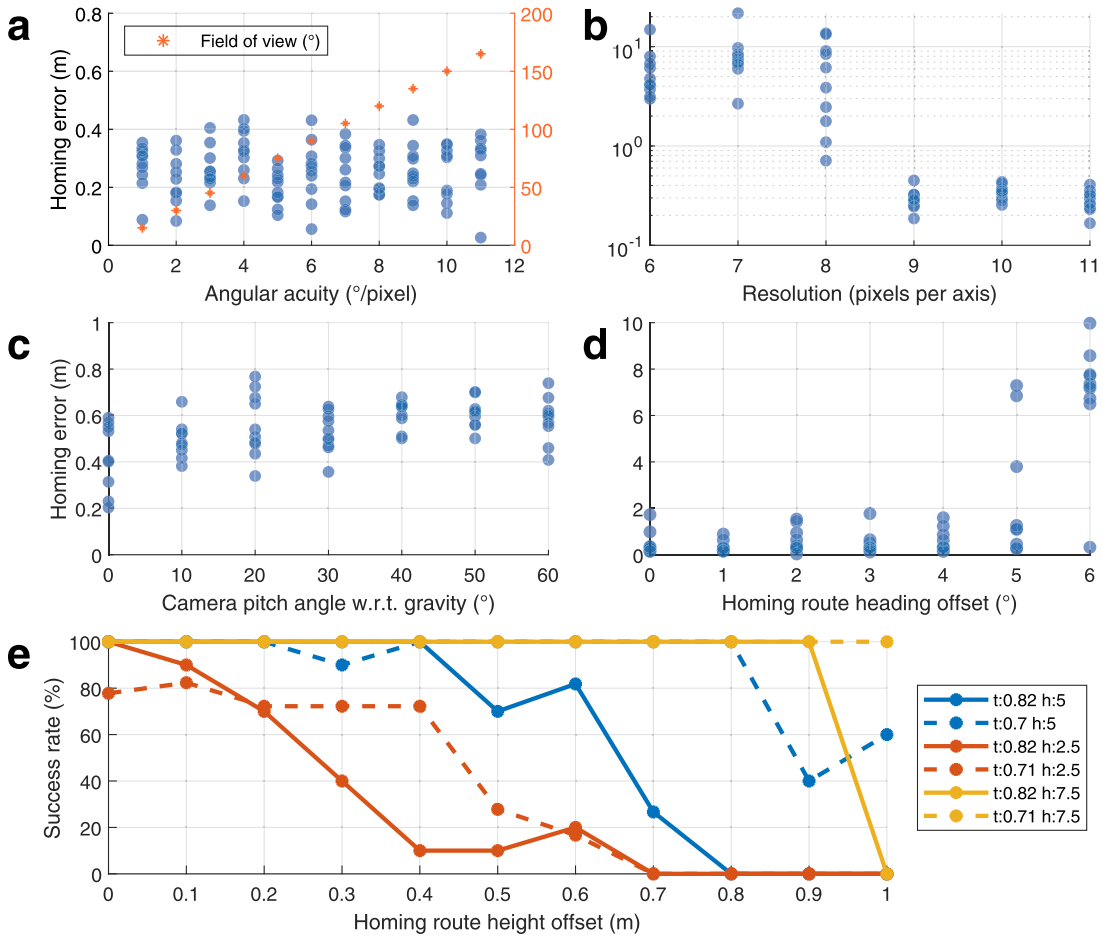

Figure 7. Robustness of TORF, $n = 10$ for each trial configuration. (a) Scatter plot showing the homing error for trials performed at different angular acuities. Reducing the acuity was achieved by increasing the field of view (field of view value specified by the right hand axis) while keeping the number of image pixels constant. (b) Scatter plot showing the homing error for trials performed at different camera resolutions with a constant camera field-of-view of 42° × 42°. (c) Scatter plot of the homing error for a series of simulated trials in which the biorobot's camera pitch angle is systematically raised. (d) Scatter plot showing the homing error for trials where the homing route heading deviates from the learning route. (e) Effect of height variation between homing and learning routes. For each configuration of familiarity threshold (t) and inbound height (h), a fixed offset was made to the height setpoint on the outbound route in increments of +10 cm (the x-axis). The simulation was considered a pass if the homing error was less than 1 m.

4.2.3.2. Effect of camera pitch angle We have shown that TORF is robust when the camera's view direction is aligned with gravity (facing the ground). Ommatidial maps of the eyes of bees however indicate that they have little or no visual acuity directly beneath them [71]. In order to establish whether TORF can operate successfully with biologically plausible view directions, the pitch angle of the camera was raised from alignment with gravity in intervals of 10°. Due to the limited size of the Seville mesh, the highest pitch elevation trialled was 60°. Note that at 30° below the horizon, the inter-facet angle of bumble bees and honey bees is between 1 and 2° [71]. According to section 4.2.3, this is well within the required acuity for TORF to operate.

The results are summarised in figure 7(c). It is apparent that in the Seville simulation environment, TORF with CWSSIM is equally effective across each of the trialled pitch angles. We therefore conclude that the distribution of ommatidia in the eyes of honey bees provides enough directed visual acuity towards the ground plane to support a TORF-like visual homing strategy.

4.2.3.3. Effect of learning/homing route heading variation Honey bees are thought to be able to estimate their global heading with a tolerance of around ±2° [22]. In order to verify that our homing system could operate in this range we performed trials with a fixed heading offset between the learning and homing routes. The results are included in figure 7(d). This data indicates that our approach is robustly tolerant to image alignment offsets of around 4°. Beyond this, the number of flights with at least 1 failure to identify the familiarity ridge steadily increases. While this shows that our system operates within the required tolerance, it does not leave a lot of scope for resilience to environmental perturbations. We conclude that a TORF-like visual homing strategy could exist in honey bees provided that they are capable of controlling their heading with an accuracy of around ±2°

in the natural world, or that they have a mitigation strategy for large rotational disturbance events.

4.2.3.4. Effect of learning/homing route height variation Our biorobot uses a height control system which includes a time-of-flight range sensor for state estimation. This is clearly not a biologically plausible mechanism and is therefore currently a deficiency of our model. It is not trivial to design a biologically plausible height control mechanism and the variability in height of the return flights of honey bees remains unmeasured. Nevertheless, we can quantify the sensitivity of TORF to variation between the height of inbound and outbound routes as a means of evaluating the biological relevance of TORF.

A series of simulations was conducted with a fixed constant height offset between learning and the homing route pairs. Preliminary trials indicated that the higher the ground height of the learning route, the less sensitive CWSSIM is to absolute height variations on the inbound route. We therefore conducted the process at three different heights. Finally, we have been relatively conservative with the familiarity threshold throughout this paper to minimise the risks of FPs. In this experiment we investigated the impact of reducing this threshold.

The results are included in figure 7(e). It is apparent that with the default familiarity threshold, the homing success rate is 100% when the inbound route is within 10% of the height of the outbound route. Reducing the familiarity threshold can increase the homing reliability at greater height offsets but this is at the expense of reliability at smaller height offsets. Given that our model fails after a single FP, it is plausible that a biological system could make use of a more flexible familiarity classification process. Once travelling along a route for example, the threshold could be reduced. Alternatively, a Bayesian approach could be used to incorporate familiarity information from the previous sidesweep. Altogether, TORF and by extension CWSSIM display some robustness to height variation but it is likely that honey bees would have to rely on additional mechanisms to stay on a route if a TORF-like visual homing strategy does form part of their navigational repertoire.

5. Discussion

Honey bees that are displaced from their nests in wide open meadows can home using visual navigation mechanisms [11]. This paper presents a new model of visual route following which complements existing paradigms. A transverse oscillatory motor pattern is combined with an image matching process that acts on a set of previously experienced downward views to guide the animal efficiently along a familiar inward path. Our approach yields robust performance when tested on a flying robotic platform over distances of up to 30 m. Our results indicate that the ostensibly self-similar textures of the meadow topography provide surprisingly unique and stable descriptors, even in windy conditions and with ultra-low resolution vision sensors. This suggests that downward views could provide an important source of information to the route following system of honey bees and other flying insects.

Our model builds on previous work showing that a memory of a set of familiar views along a route can guide subsequent travel along that route for an insect. However, previous models of this behaviour tend to rely on the agent regularly scanning by rotating on the spot [3], a motion that is not readily observed during insect flight. Alternatively it has been suggested that mental rotation of views [40] or rotation invariant view matching procedures [60, 65, 68] could be used. However, these approaches have to date assumed uniform spatial sampling, and neglect the fact that animals do scan when they are visually disorientated [81]. Moreover, the frequency-based encodings proposed in these models seem at odds with the general view that biological vision systems rely on locally supported Gabor-like basis functions [23, 26, 69]. We have suggested an alternative encoding, based on wavelet decomposition, which provides a degree of translational invariance. This allows an agent maintaining a fixed heading to reliably recognise when it crosses a familiar path, and to follow it home.

5.1. Using downward views

Most previous attempts to implement visual route following in flying agents ([18, 41, 47] for example) have maintained the panoramic view input assumption from terrestrial navigation models. As previously suggested in [12, 21], we have found that a narrow downward viewing frustum is sufficient for aerial agents to navigate in a familiar area, for example, from within the general range of their nest that they could obtain after using PI to home. We suggest that such views could be a better source of information when the task is to locate a precise destination in the open spaces that honey bees often inhabit, where panoramic views of the horizon may be similar over large areas. Furthermore, it is known that these animals attend ground level visual features for navigation [38]. Nevertheless, views of the horizon could have a role in both locating the familiarity ridge and in maintaining a consistent heading in the relevant direction via the visual compass. Indeed, adding information from a wider range of view directions could increase robustness of the TORF procedure. While previous work has shown that panoramic visual information can work in simulated route following, this tends to require a critical density of vertical objects [18, 41] which is not the case with TORF. Although, it should be noted that while our approach is insensitive to the level of horizontal clutter, it does depend on the agent being able to plot a course that follows a reliable ground height. It would be interesting in future

to systematically compare our approach with those outlined in [18, 41] in environments of different levels of vertical clutter. It would also be interesting to examine how a displaced agent could use different visual navigation approaches with different fields of visual attention to locate a destination most effectively. This is something that has been previously examined in terrestrial animals [77, 79].

5.2. Wavelet-based image matching

The complex wavelet steerable pyramid provides an efficient means of decomposing images at multiple scales. Combined with the SSIM component of the CWSSIM algorithm, it provides an image matching method with a degree of translational invariance, as set out in section 2.2 of [87]. Our results indicate that CWSSIM provides approximately double the catchment area of MSE based methods for a given reference image. Crucially, CWSSIM's output monotonically declines with horizontal distance from a reference image. By increasing the width of the monotonic region of the search space generated by an image/image difference pipeline combination, the number of view memories that is required can be reduced. Our work also demonstrates for the first time that CWSSIM can operate on low resolution images commensurate with the insect vision system.

The CWSSIM approach is biologically relevant in that orientated bandpass filters are known to receive input from the invertebrate vision system [56, 82] and mammalian vision systems [5, 25, 26]. However, as implemented here, the CWSSIM pipeline (figure S1(a) involves a fast Fourier transform operation which we do not consider to be a biologically plausible process. This operation however, is purely for performance reasons and unlike view matching schemes which operate in the frequency domain [60, 65, 68], our view matching is conducted in the spatial domain (i.e. following an inverse fast Fourier transform). Nevertheless, it would be useful to verify biological plausibility by implementing a bandpass filter array that operates on a biologically relevant eye model model such as those detailed in [17, 63].

We have not as yet exploited the multi-resolution aspect of our wavelet function since adding finer levels added computational burden. As discussed in [39], weighting coefficients at different scales could provide more scene recognition discriminability. In initial trials, we found that adding the higher resolution frequency bands tended to reduce the catchment area of reference images. However, it is possible that multi-resolution approaches would be beneficial in images with wider fields of view. The low frequency information in ultra-low resolution images could for example be used to prune image search spaces or network parameters, an approach that has been recently investigated in deep neural network architectures [32, 70].

An additional feature of CWSSIM is that it is inherently amenable to compression. High value CWSSIM coefficients represent strong edges at a particular scale and low value coefficients can be discarded with limited impact on the view comparison process. Previous work has shown that surprisingly few coefficients can describe a view memory [16, 53] and that in mammalian vision, surprisingly few V1 cells are active in when a mouse is presented with a visual scene [83]. The level of compression that is achievable with a biologically relevant circuit will be examined in future work. Using a mushroom body network to sparsely encode views could further reduce the memory overheads for a given number of views and also speed up processing time by making the comparison of the current view to all stored views a simple parallel operation.

5.3. Transverse oscillating route following

TORF differs from the existing visual route following approaches in that it uses a translational rather than a rotational image difference function to evaluate view familiarity over time. We have achieved translational scanning via a component of velocity that is transverse to the agent's principle direction of travel. While this additional motion increases the overall distance that an agent travels along a particular route, faster journey times are obtained by eliminating the need to stop and scan. While flying insects are more obviously suited to holonomic translational motion, it is interesting to note that oscillations have also been reported in desert ants on the micro-scale [33], although recent works suggest that these oscillations could be emerging from an opponency process that acts between the left and right eye visual streams [80] or steering circuits [24, 50]. It would be interesting to evaluate the visual opponent process from [80] onboard a quadcopter to see if TORF levels of reliability are attainable with this approach.

A criticism of previous implementations of visual route following algorithms is that they have often been deployed in low polygon count simulation environments [2, 3]. In recent years high fidelity reconstructions of animal habitats have been used to improve simulation environment realism [52, 64]. However, these tools still amount to a simplification of the real world as they provide light intensity scales with a low dynamic range, built in colour constancy properties and a lack of in-scene physics. Deploying visual route following algorithms on mobile robots forces the algorithm designer to consider how real world signals can be reliably used [27, 28], but this approach requires considerably more design effort and the validation process is similarly laborious.

In this paper we have made use of the open source Pixhawk flight controller and PX4 software stack [14] to deploy our algorithm in both simulation and on real hardware. This approach has enabled us to verify that our algorithm can operate in real world

conditions while maintaining the flexibility to also systematically examine the impact of adjusting key parameters. We were able to incorporate the 3D meshes from [52] in order to maximise the biological relevance of the simulated worlds. We have made our codebase open source and we encourage other researchers to make use of this platform and the underlying robotics framework. In future we aim to further explore the impact of real world lighting schemes by adding different time intervals between the learning and homing flights.

In our current implementation of TORF, we opted to use transverse oscillations with relatively large amplitudes in order to minimise the impact of rolling oscillations that were experienced at the simulated gimbal whenever the sign of transverse velocity was reversed. As a consequence, a large angle of attack was set to ensure smooth gimbal action at crossover locations. We predict that the amplitude and angle-of-attack of the transverse oscillations would be considerably smaller in nature, such that flying insects remain on the familiarity ridge most of the time, and their velocity vector is more biased towards their direction of travel. Preliminary routes with an improved gimbal system are included in section S 1.4.2. In future work we will try to optimise the oscillation profile of our approach by minimise the energy expended on lateral motion while preserving the homing reliability. A more direct flight path could also make TORF a more attractive solution for robotic applications such as point-to-point navigation under a tree canopies.

Given the lack of supporting biological data, we have used a simple route termination profile which simply used open loop control to center on the route familiarity ridge when the agent detected it was nearing the end of the route. An obvious way to improve this method would be to use a closed loop system which continues to oscillate until the nest itself is detected. This method would require the agent to move backwards in the longitudinal direction so that nest overshoot scenarios can be recovered from. In recent work, the use of negative training views have been explored, whereby the agent learns views that inform the animal it is going in the wrong direction [30]. It would be interesting to develop this approach in the TORF route termination process. I.e. in the learning flight the agent would over run the nest slightly in order to acquire some downward views that can be used to inform the agent that it has overshot its goal. If this strategy is implemented in the biological system, it should be detectable in honey bee learning flights.

A requirement of the TORF method is that the homing agent must assume the same heading as when the route was learned. In this work we have minimised this challenge by using straight learning flights and by assuming that the animal maintains a fixed, route-aligned heading on both the learning and homing flight. This has the advantage of reducing the search effort to find the route, and then once on the route it allows the agent to use purely translational motion to stay on the route. The straightness of the route was justified by noting that the inbound leg of honey bee [10] and bumble bee [43] learning flights appear to be straight. However, in more cluttered environments it is also likely that honey bees would have to change course in the process of avoiding obstacles resulting in curved routes. In section 4.2.2, we demonstrated TORF's ability to handle curved routes, it is also possible that using the indexing approach outlined in section 3.2.1, changes in heading could be built-in to the route memory.

While the straightness constraint of a TORF learning flight can be justified with biological data, there is currently no evidence to suggest that honey bees maintain a constant alignment with their route direction during homing flights. There are several datasets which track the heading of various flying insects on the final approach ($<$50 cm) to their nest [46, 67, 84]. These studies indicate that while the insects demonstrate TORF-like goal-directed zig-zags and high degrees of sideslip on straight sections of flight, the range of headings observed is far greater than the 4° tolerance we specified in section 4.2.3. It therefore seems likely that a honeybee's route following strategy can cope with a wider range of headings than is predicted by our model. It is worth noting that the average width of the *familiarity corridor* found in figure 6(f) is around 50 cm which could provide scope for a joint exploration of heading and translation exploration. For example, heading exploration is performed at the switchback points and translational exploration is achieved in the intervening straight sections.

We have shown in this work that TORF is resilient to height variations of approximately 10% which would cover small disturbances and modest errors in height estimation. However, it is reasonable to assume that obstacles, height discontinuities and the landing approach would cause considerably more than 10% height variation. The degree to which a honey bee can gauge its height is unclear since optic flow signals are confounded by speed and height [57]. In [18], helical flight paths are suggested for extending the search space to the third dimension. We propose the indexing approach of section 3.2.1 could be used to manage large height variations along a route in conjunction with the height estimation mechanisms proposed in [61].

5.4. Ethological implications

We have shown that low resolution images of the ground can be used as robust descriptors in the natural world provided that the right visual processing is used. Navigation can be achieved using a chain of these memories over long distances without any

other navigation cues. This result casts an interesting lens on previous studies which have assumed that flat grassland does not afford this opportunity. For example in [37], the authors state:

> *'We were concerned about the possibility that ground structures around the hive might provide guiding posts over larger distances than the estimated 60 m radius around the hive. There were no structures on the ground 2–3 m in diameter close to the hive, and larger patches of grassland with slightly different vegetation were distributed rather evenly in the whole study area. Bees approached the hive from all directions, thus excluding the possibility that they might have seen a particular spatial arrangement of grassland patches. Therefore, we conclude that bees were not guided to the hive by any beacons found close to the hive, in its surroundings, or in the profile of the horizon.'*

The authors use this to justify their claim that honey bees must be using a cognitive map to find their way. Our work demonstrates that even with low resolution vision, arbitrary patches within continuous grassland could be reliably distinguished and used to navigate across this terrain without a cognitive map. Other paradigms have used tents as beacons in grasslands [7, 36] in order to show that the visual odometer is the dominant cue for judging range on foraging journeys. We suggest that these results should be interpreted with caution because the meadow surface could be providing many other effective *beacons*. In general we suggest that ethologists should consider any terrain as a potential source of location information for honey bees and other flying insects, even if this is not obvious to the human eye.

If the TORF mechanism that we have proposed is manifested in the nervous system of honey bees it should be possible to observe: (1) the presence of transverse oscillations in the flight path and (2) that the return paths of a particular animal have a high degree of spatial overlap on successive foraging bouts. If return flights were not found to overlap this would indicate a more horizon-centric viewing strategy such as that proposed in [64]. It is worth re-iterating that we believe that a blend of strategies may be observed in different colonies depending on the hive surroundings. Our hypothesis is that honey bees process information from all viewing directions simultaneously and the incoming information is weighted downstream of the retina.

Unfortunately the trajectories of honey bees have not yet been captured in sufficient detail over a suitable range for models like TORF to be evaluated. We hope that high-resolution multi-camera rigs or other technological advances (e.g. [73]) will be able to capture these details in the medium-term future. Conventional experimental approaches may also help to interrogate the plausibility of TORF, for example, by disrupting the ground level information that a tracked insect experiences on its route home or by examining the impact that masking a honey bee's ventral vision has on its ability to navigate close to its nest. We predict that the navigation abilities of displaced honey bees would be temporarily but significantly disrupted if a substantially large area around their nest was modified. For example by mowing or ploughing the field such that the ground structures around the nest are fundamentally altered.

Acknowledgments

This work was supported by EPSRC Centre for Doctoral Training in Robotics and Autonomous Systems (EPSRC Grant EP/L016834/1) and BBSRC Grant BB/R005036/1. Material in the results section was presented in the Insect Navigation Conference (online) and is an edited version of the third chapter of the first author's PhD Thesis [58]. We would like to thank Professor Andy Philippides for the useful suggestions he provided during the defence of this thesis. We also thank Evripidis Gkanias for proofreading the manuscript.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://datashare.ed.ac.uk/handle/10283/3815.

ORCID iDs

J Stankiewicz https://orcid.org/0000-0003-2419-9168
B Webb https://orcid.org/0000-0002-8336-6926

References

[1] Alley R 2021 Eric armishaw reserve aerial image from openaerialmaps https://map.openaerialmap.org/#/179.62646484375,-40.06125658140473,6/square/311333000/5fd32175f00c200007292ba3?_k&tnqx3d;7zzzdg (accessed 20 May 2021)
[2] Ardin P, Peng F, Mangan M, Lagogiannis K and Webb B 2016 Using an insect mushroom body circuit to encode route memory in complex natural environments *PLoS Comput. Biol.* **12** e1004683
[3] Baddeley B, Graham P, Husbands P and Philippides A 2012 A model of ant route navigation driven by scene familiarity *PLOS Computational Biology* **202** 87–95
[4] Bell A J and Sejnowski T J 1997 The 'independent components' of natural scenes are edge filters *Vis. Res.* **37** 3327–38
[5] Bredfeldt C E and Ringach D L 2002 Dynamics of spatial frequency tuning in macaque V1 *J. Neurosci.* **22** 1976–84
[6] Briand T, Vacher J, Galerne B and Rabin J 2014 The Heeger & Bergen pyramid based texture synthesis algorithm *Image Process. Line* **4** 276–99

[7] Chittka L, Geiger K and Kunze J 1995 The influences of landmarks on distance estimation of honey bees *Anim. Behav.* **50** 23–31

[8] Collett M, Chittka L and Collett T S 2013 Spatial memory in insect navigation *Curr. Biol.* **23** R789–800

[9] Collett T S and Collett M 2002 Memory use in insect visual navigation *Nat. Rev. Neurosci.* **3** 542–52

[10] Degen J *et al* 2015 Exploratory behaviour of honeybees during orientation flights *Anim. Behav.* **102** 45–57 cited by 38

[11] Degen J *et al* 2016 Honeybees learn landscape features during exploratory orientation flights *Curr. Biol.* **26** 2800–4

[12] Denuelle A and Srinivasan M V 2016 A sparse snapshot-based navigation strategy for UAS guidance in natural environments *2016 IEEE Int. Conf. Robotics and Automation (ICRA)* pp 3455–62

[13] Denuelle A, Thurrowgood S, Kendoul F and Srinivasan M 2015 A view-based method for local homing of unmanned rotorcraft *ICARA 2015-Proc. 2015 6th Int. Conf. Automation, Robotics and Applications* pp 443–9

[14] PX4 Development Team 2020 *PX4* https://px4.io/ (accessed 20 June 2021)

[15] Dewar A D M, Wystrach A, Graham P and Philippides A 2015 Navigation-specific neural coding in the visual system of drosophila *Biosystems* **136** 120–7 Selected papers presented at the Eleventh International Workshop on Neural Coding, Versailles, France, 2014

[16] Dewar A D M, Wystrach A, Philippides A and Graham P 2017 Neural coding in the visual system of *Drosophila melanogaster*: how do small neural populations support visually guided behaviours? *PLoS Comput. Biol.* **13** 1–21

[17] Dickson W B, Straw A D and Dickinson M H 2008 Integrative model of drosophila flight *AIAA J.* **46** 2150–64

[18] Differt D and Stürzl W 2020 A generalized multi-snapshot model for 3D homing and route following *Adapt. Behav.* https://doi.org/10.1177/1059712320911217

[19] Dittmar L, Stürzl W, Baird E, Boeddeker N and Egelhaaf M 2010 Goal seeking in honeybees: matching of optic flow snapshots? *J. Exp. Biol.* **213** 2913–23

[20] Gaffin D D and Brayfield B P 2016 Autonomous visual navigation of an indoor environment using a parsimonious, insect inspired familiarity algorithm *PLoS One* **11** 1–25

[21] Gaffin D D, Dewar A, Graham P and Philippides A 2015 Insect-inspired navigation algorithm for an aerial agent using satellite imagery *PLoS One* **10** e0122077

[22] Gkanias E, Risse B, Mangan M and Webb B 2019 From skylight input to behavioural output: a computational model of the insect polarised light compass *PLoS Comput. Biol.* **15** 1–30

[23] Gladilin E and Eils R 2015 On the role of spatial phase and phase correlation in vision, illusion, and cognition *Front. Comput. Neurosci.* **9** 45

[24] Goulard R, Buehlmann C, Niven J E, Graham P and Webb B 2021 A unified mechanism for innate and learned visual landmark guidance in the insect central complex *bioRxiv Preprint* https://doi.org/10.1101/2021.01.28.428620

[25] Hubel D H and Wiesel T N 1959 Receptive fields of single neurones in the cat's striate cortex *J. Physiol.* **148** 574–91

[26] Jones J P and Palmer L A 1987 An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex *J. Neurophysiol.* **58** 1233–58

[27] Knight J C, Sakhapov D, Domcsek N, Dewar A D M, Graham P, Nowotny T and Philippides A 2019 Insect-inspired visual navigation on-board an autonomous robot: real-world routes encoded in a single layer network *The 2019 Conf. Artificial Life* vol 31 (MIT Press) pp 60–7

[28] Kodzhabashev A and Mangan M 2015 Route following without scanning *Biomimetic and Biohybrid Systems* ed S P Wilson, P F Verschure, A Mura and T J Prescott (Berlin: Springer) pp 199–210

[29] Kohler M and Wehner R 2005 Idiosyncratic route-based memories in desert ants, *Melophorus bagoti*: how do they interact with path-integration vectors? *Neurobiol. Learn. Mem.* **83** 1–12

[30] Le Möel F and Wystrach A 2020 Opponent processes in visual memories: a model of attraction and repulsion in navigating insects' mushroom bodies *PLoS Comput. Biol.* **16** 1–24

[31] Lee C and Kim D 2018 Visual homing navigation with Haar-like features in the snapshot *IEEE Access* **6** 33666–81

[32] Liu Z, Xu J, Peng X and Xiong R 2018 Frequency-domain dynamic pruning for convolutional neural networks *Advances in Neural Information Processing Systems* ed S Bengio, H Wallach, H Larochelle, K Grauman, N Cesa-Bianchi and R Garnett vol 31 (Red Hook: Curran Associates, Inc)

[33] Léo C, Schwarz S and Wystrach A 2020 How oscillations are controlled in navigating ants (poster) *Insect Navigation Workshop (Online)*

[34] Mangan M and Webb B 2012 Spontaneous formation of multiple routes in individual desert ants (*Cataglyphis velox*) *Behav. Ecol.* **23** 944–54

[35] Menzel R 2019 The waggle dance as an intended flight: a cognitive perspective *Insects* **10** 424

[36] Menzel R, Fuchs J, Nadler L, Weiss B, Kumbischinski N, Adebiyi D, Hartfil S and Greggers U 2010 Dominance of the odometer over serial landmark learning in honeybee navigation *Naturwissenschaften* **97** 763–7

[37] Menzel R *et al* 2005 Honey bees navigate according to a map-like spatial memory *Proc. Natl Acad. Sci.* **102** 3040–5

[38] Menzel R *et al* 2019 Guidance of navigating honeybees by learned elongated ground structures *Front. Behav. Neurosci.* **12** 322

[39] Meyer S, Nowotny T, Graham P, Dewar A and Philippides A 2020 Snapshot navigation in the wavelet domain *Biomimetic and Biohybrid Systems* ed V Vouloutsi, A Mura, F Tauber, T Speck, T J Prescott and P F M J Verschure (Berlin: Springer) pp 245–56

[40] Möller R 2012 A model of ant navigation based on visual prediction *J. Theor. Biol.* **305** 118–30

[41] Müller J, Nawrot M, Menzel R and Landgraf T 2017 A neural network model for familiarity and context learning during honeybee foraging flights *Biol. Cybern.* **112** 113–26

[42] Narendra A, Gourmaud S and Zeil J 2013 Mapping the navigational knowledge of individually foraging ants, *Myrmecia croslandi Proc. R. Soc. B.* **280** 20130683

[43] Osborne J L, Smith A, Clark S J, Reynolds D R, Barron M C, Lim K S and Reynolds A M 2013 The ontogeny of bumblebee flight trajectories: from naïve explorers to experienced foragers *PLoS One* **8** 1–11

[44] Pahl M, Zhu H, Tautz J and Zhang S 2011 Large scale homing in honeybees *PLoS One* **6** 1–7

[45] Philippides A, Baddeley B, Cheng K and Graham P 2011 How might ants use panoramic views for route navigation? *J. Exp. Biol.* **214** 445–51

[46] Philippides A, de Ibarra N H, Riabinina O and Collett T S 2013 Bumblebee calligraphy: the design and control of flight motifs in the learning and return flights of bombus terrestris *J. Exp. Biol.* **216** 1093–104

[47] Philippides A, Steadman N, Dewar A, Walker C and Graham P 2016 Insect-inspired visual navigation for flying robots *Biomimetic and Biohybrid Systems* ed N F Lepora, A Mura, M Mangan, P F Verschure, M Desmulliez and T J Prescott (Berlin: Springer) pp 263–74

[48] Portilla J and Simoncelli E P 2000 Parametric texture model based on joint statistics of complex wavelet coefficients *Int. J. Comput. Vis.* **40** 49–70

[49] Portilla J, Strela V, Wainwright M J and Simoncelli E P 2003 Image denoising using scale mixtures of Gaussians in the wavelet domain *IEEE Trans. Image Process.* **12** 1338–51

[50] Rayshubskiy A, Holtz S L, D'Alessandro I, Li A A, Vanderbeck Q X, Haber I S, Gibb P W and Wilson R I 2020 Neural circuit mechanisms for steering control in walking

drosophila *bioRxiv Preprint* https://doi.org/10.1101/2020.04.04.024703 (accessed 26 July 2021)
[51] Reynolds A M, Smith A D, Menzel R, Greggers U, Reynolds D R and Riley J R 2007 Displaced honey bees perform optimal scale-free search flights *Ecology* **88** 1955–61
[52] Risse B, Mangan M, Stürzl W and Webb B 2018 Software to convert terrestrial LiDAR scans of natural environments into photorealistic meshes *Environ. Model. Software* **99** 88–100
[53] Roper M, Fernando C and Chittka L 2017 Insect bio-inspired neural network provides new evidence on how simple feature detectors can enable complex visual generalization and stimulus location invariance in the miniature brain of honeybees *PLoS Comput. Biol.* **13** 1–23
[54] Sampat M P, Wang Z, Gupta S, Bovik A C and Markey M K 2009 Complex wavelet structural similarity: a new image similarity index *IEEE Trans. Image Process.* **18** 2385–401
[55] Schwarz S, Mangan M, Webb B and Wystrach A 2020 Route-following ants respond to alterations of the view sequence *J. Exp. Biol.* **223** jeb218701
[56] Seelig J and Jayaraman V 2013 Feature detection and orientation tuning in the drosophila central brain *Nature* **106** 262–6
[57] Serres J R and Ruffier F 2017 Optic flow-based collision-free strategies: from insects to robots *Arthropod Struct. Dev.* **46** 703–17 From Insects to Robots
[58] Stankiewicz J 2020 Using a quadcopter to model the visual navigation behaviours of flying insects *PhD Thesis* University of Edinburgh
[59] Stankiewicz J and Webb B 2020 Using the neural circuit of the insect central complex for path integration on a micro aerial vehicle *Biomimetic and Biohybrid Systems* ed V Vouloutsi, A Mura, F Tauber, T Speck, T J Prescott and P F M J Verschure (Berlin: Springer) pp 325–37
[60] Stone T, Mangan M, Wystrach A and Webb B 2018 Rotation invariant visual processing for spatial memory in insects *Interface Focus* **8** 20180010
[61] Straw A D, Lee S and Dickinson M H 2010 Visual control of altitude in flying drosophila *Curr. Biol.* **20** 1550–6
[62] Strydom R, Denuelle A and Srinivasan M 2016 Bio-inspired principles applied to the guidance, navigation and control of UAS *Aerospace* **3** 21
[63] Stürzl W, Boeddeker N, Dittmar L and Egelhaaf M 2010 Mimicking honeybee eyes with a 280° field of view catadioptric imaging system *Bioinspir. Biomim.* **5** 036002
[64] Stürzl W, Grixa I, Mair E, Narendra A and Zeil J 2015 Three-dimensional models of natural environments and the mapping of navigational information *J. Comp. Physiol.* A **201** 563–84
[65] Sturzl W and Mallot H A 2006 Efficient visual homing based on Fourier transformed panoramic images *Robot. Auton. Syst.* **54** 300–13
[66] Stürzl W and Zeil J 2007 Depth, contrast and view-based homing in outdoor scenes *Biol. Cybern.* **96** 519–31
[67] Stürzl W, Zeil J, Boeddeker N and Hemmi J M 2016 How wasps acquire and use views for homing *Curr. Biol.* **26** 470–82
[68] Sun X, Yue S and Mangan M 2019 A decentralised neural model explaining optimal integration of navigational strategies in insects *eLife* **9** e54026
[69] Swindale N V 1997 Visual cortex: a cat's-eye view of the visual system *Curr. Biol.* **7** R387–389
[70] Tan S, Wu W, Shao Z, Li Q, Li B and Huang J 2020 CALPA-NET: channel-pruning-assisted deep residual network for steganalysis of digital images (arXiv:1911.04657)
[71] Taylor G J, Tichit P, Schmidt M D, Bodey A J, Rau C and Baird E 2019 Bumblebee visual allometry results in locally improved resolution and globally improved sensitivity *eLife* **8** 1–32
[72] van Dalen G J J, McGuire K N and de Croon G C H E 2018 Visual homing for micro aerial vehicles using scene familiarity *Unmanned Syst.* **06** 119–30
[73] Vo-Doan T T and Straw A D 2020 Millisecond insect tracking system (arXiv:2002.12100)
[74] von Frish K 1967 *The Dance Language and Orientation of Bees* (Oxford: Oxford University Press)
[75] Wang Z and Bovik A C 2009 Mean squared error: love it or leave it ? *IEEE Signal Process. Mag.* **26** 98–117
[76] Webb B 2019 The internal maps of insects *J. Exp. Biol.* **222** jeb188094
[77] Wystrach A, Beugnon G and Cheng K 2012 Ants might use different view-matching strategies on and off the route *J. Exp. Biol.* **215** 44–55
[78] Wystrach A, Dewar A D M and Graham P 2014 Insect vision: emergence of pattern recognition from coarse encoding *Curr. Biol.* **24** R78–80
[79] Wystrach A, Dewar A, Philippides A and Graham P 2016 How do field of view and resolution affect the information content of panoramic scenes for visual navigation? A computational investigation *J. Comp. Physiol.* A **202** 87–95
[80] Wystrach A, Le Moël F, Clement L and Schwarz S 2020 A lateralised design for the interaction of visual memories and heading representations in navigating ants *bioRxiv Preprint* https://doi.org/10.1101/2020.08.13.249193 (accessed 26 July 2021)
[81] Wystrach A, Philippides A, Aurejac A, Cheng K and Graham P 2014 Visual scanning behaviours and their role in the navigation of the Australian desert ant *Melophorus bagoti J. Comp. Physiol.* A **200** 615–26
[82] Yang E-C and Maddess T 1997 Orientation-sensitive neurons in the brain of the honey bee (*Apis mellifera*) *J. Insect Physiol.* **43** 329–36
[83] Yoshida T and Ohki K 2020 Natural images are reliably represented by sparse and variable populations of neurons in visual cortex *Nat. Commun.* **11** 872
[84] Zeil J 1993 Orientation flights of solitary wasps (Cerceris; Sphecidae; Hymenoptera) *J. Comp. Physiol.* A **172** 207–22
[85] Zeil J, Hofmann M I and Chahl J S 2003 Catchment areas of panoramic snapshots in outdoor scenes *J. Opt. Soc. Am.* A **20** 450
[86] Zeil J, Narendra A and Stürzl W 2014 Looking and homing: how displaced ants decide where to go *Phil. Trans. R. Soc.* B **369** 20130034
[87] Wang Z and Simoncelli E P 2005 Translation insensitive image similarity in complex wavelet domain *Proc. (ICASSP '05). IEEE Int. Conf. Acoustics, Speech, and Signal Processing* vol 2 pp 573–6