Biological Cybernetics manuscript No.
(will be inserted by the editor)

Michael Mangan · Barbara Webb

Modelling Place Memory in Crickets

Received: date / Revised: date

Abstract Insects can remember and return to a place of interest using the surrounding visual cues. In previous experiments we showed that crickets could home to an invisible cool spot in a hot environment, and that they did so most effectively with a natural scene surround, though they were also able to home with distinct landmarks or blank walls. Here we compare six different models of visual homing with the same visual environments used for testing the crickets. Only models deemed biologically plausible for use by insects were implemented. The Average Landmark Vector model and First Order Differential Optic Flow are unable to home better than chance in at least one of the visual environments. Second Order Differential Optic Flow and Gradient Descent on image differences can home better than chance in all visual environments, and best in the natural scene environment, but do not quantitatively match the distributions of the cricket data. Two models - Centre of Mass Average Landmark Vector, and Run Down on image differences - could produce the same pattern of results as observed for crickets. Both performed best using simple binary images, and were robust to changes in resolution and image smoothing.

1 Introduction

The ability of insects to return to a location of interest such as a feeder or nest using visual cues has been well documented in natural settings (Wehner 2003; Collett and Collett 2002). Additionally, it has been shown that cockroaches (Mizunami et al 1998), and crickets (Wessnitzer et al 2008) can learn to return to a target location in a novel laboratory experiment analogous to the Morris water-maze used to assess place memory in mammals (Morris et al 1982). This experimental paradigm placed insects in a hostile environment (a hotplate maintained above 40°C) from which they would seek an escape. An invisible cool spot, maintained at a moderate temperature represented the only refuge. Over successive trials the time taken by the insects to re-locate the cool spot decreased significantly. Removal of all visual cues through trials performed in the dark prevented learning, and the search is affected by rotation of the visual surroundings.

Many visual homing models have been proposed as possible strategies employed by insects. These can be broadly split into two classes: feature-based models and image-based models. Feature-based models, for example the snapshot model (Cartwright and Collett 1983), extract features from the visual surround such as angular size and bearing of prominent landmarks. Comparison of the features extracted at the goal location with the corresponding features extracted from an image taken at a displaced location allows a homing vector to be calculated. Image-based models, by contrast, use comparison of the raw images as seen from the goal and the displaced location. For example, Zeil et al (2003) show that the pixel-wise root mean square (RMS) error between a panoramic reference image and image viewed from a displaced location increases monotonically with distance. Homing can then be achieved through some form of gradient descent where either agent movements allow the home direction to be inferred (Zeil et al 2003) or where simulated agent movements allow a home vector to be estimated (Franz et al 1998).

If insects are indeed using a feature-based technique to return to locations of importance then it might be expected that performance would be best when distinct landmarks, easily segmented from the background, are presented. In contrast, an image-based algorithm will

M. Mangan
Informatics Forum,
10 Crichton Street,
Edinburgh,
EH8 9AB,
Scotland,
UK
Tel.: +44 131 650 2690
Fax: +44 131 651 1426
E-mail: m.mangan@sms.ed.ac.uk

Table 1: P-values calculated using Wilcoxon rank-sum test when comparing homing times across trials 7 - 10 within the four experimental paradigms.

	NS	**BW**	**DL**	**DC**
NS	X	<0.01	<0.01	<0.01
BW	<0.01	X	0.54	<0.01
DL	<0.01	0.54	X	<0.01
DC	<0.01	<0.01	<0.01	X

perform more successfully when a complex scene is presented, as it uses information from each pixel directly. In the hotplate experiments described above (Wessnitzer et al 2008), crickets were tested with both distinct landmarks and a natural scene stimulus. Learning was observed in both cases but the natural scene elicited greater improvement in homing times and more consistent learning. Figure 1 shows box plots of time taken to locate the cool spot by crickets during the final four out of ten learning trials, when their homing times had stabilised. It is clear that crickets locate the cool spot fastest in the Natural Scene surround followed by Blank Walls, Distinct Landmarks and then the Dark Control. Statistical comparisons (Table 1) show that the Natural Scene surround elicits significantly better results than all other paradigms, and performance in the dark is significantly worse than any visual condition.

Surprisingly, results in the Blank Walls surround and Distinct Landmarks are not significantly different, but it should be mentioned, as noted in Wessnitzer et al (2008) and below, that the Blank Walls environment did not eliminate all possible visual cues, as a combination of shadows in the canopy and light gradients across the arena wall remained.

In the current study we implement six biologically plausible models of visual homing, including both image-based and feature-based model types, and test them using the visual environments that were presented to the crickets in Wessnitzer et al (2008). A direct comparison can then be made between the homing path lengths recorded using the competing visual homing models and the homing times of crickets under varying visual conditions. The models will thus be assessed on their ability to reproduce the performance shown in figure 1 and table 1.

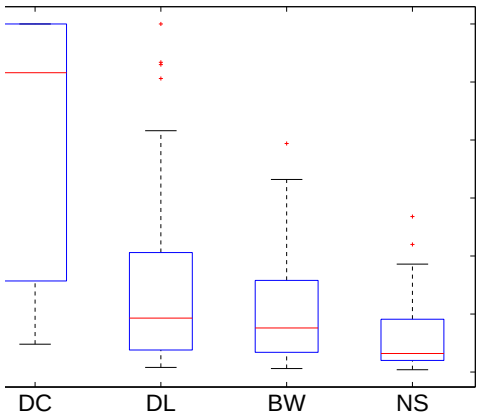


Fig. 1: Box plots showing times taken to locate the cool spot for crickets across trials 7-10 (i.e. after learning) in the different visual surroundings (*n=12*). Boxes show lines at the lower quartile, median, and upper quartile values. Whiskers indicate the most extreme values within 1.5 times the interquartile range from the ends of the box. Outliers are shown as red + signs. Testing within the natural scene (NS) produces fastest homing times, followed by blank walls (BW), distinct landmarks (DL) and the dark control (DC) respectively.

2 General methods

2.1 Image Databases

Three image databases were collected from within the cricket experimental arena on a 2cm*2cm grid using a Khepera II mobile robot and a custom built panoramic camera (see figure 2). Images were collected with the arena configured as for the cricket trials giving three sets of 208 images: natural scene (NS), distinct landmarks (DL), and blank walls (BW). All images were captured with the camera in the same orientation (see discussion). It is worth noting that the BW database images are not uniform, as might be expected, but have a clear intensity gradient. This seems to have been sufficient for homing in the cricket, and as we shall demonstrate, also suffices for homing in some of the tested models.

2.2 Homing process

Visual homing models, as described in the literature, frequently differ not only in how the home vector is determined but also in how it is used to generate motion. Here we tried to maintain consistency across the implementations so that only the relative efficacy of each model's method of determining the home direction will contribute to the results. A block diagram of the homing process for one time step is shown in figure 3.

Visual Input is received from the image database in the form of an unprocessed image (Figure 2 (c), (d), and (e)) as would be supplied by a robot positioned at the corresponding grid position. The Cricket Eye Model then unwraps and unwarps a ring corresponding to 20° above and below the image horizon. Note that as the camera turret is mounted above the robot base unit and with the mirror located above the camera to prevent image

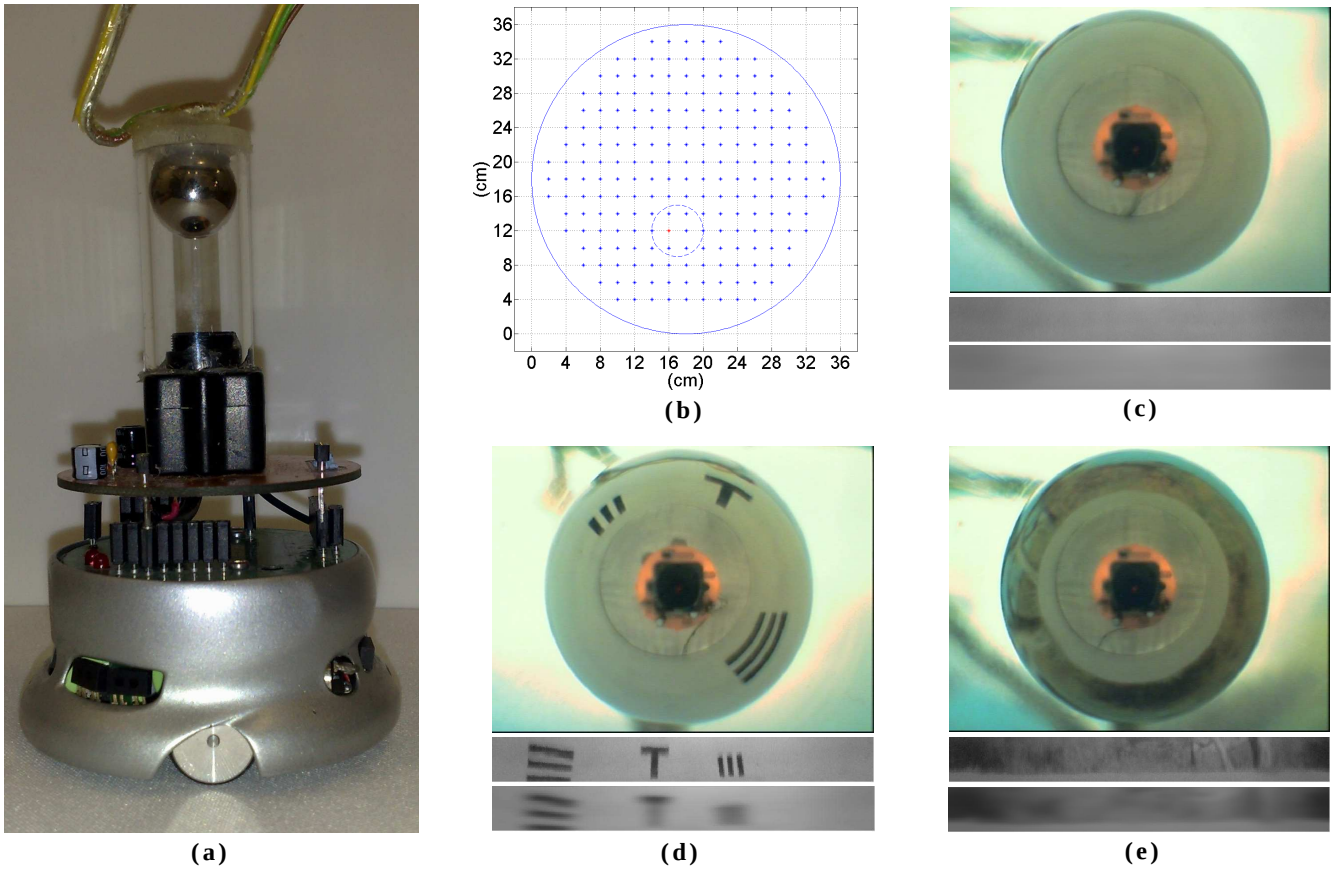


Fig. 2: (a) Khepera II robot base with custom panoramic camera turret used to record image database. (b) Cricket arena diagram showing 2cm*2cm grid (blue stars) where images where taken. The cool-spot perimeter is shown by the inner blue circle and the home position used in the modelling study by the red star. (c), (d) and (e) Sample images from the BW, DL, and NS image-sets respectively at the home location. Below is shown the images' unwrapped at 1° resolution to the maximum image size of +/- 20° around the horizon prior to smoothing, and also the same sample images post smoothing using the Butterworth filter.

interference from cables, the image horizon is approximately X cm above the arena floor. Images are initially unwrapped at 1° resolution in both azimuth and elevation. We were not able to find a precise estimate for the visual acuity of the ventral areas of the cricket species *Gryllus bimaculatus* eye, but interommitidial angles of 1° have been observed in the dorsal rim area (Labhart et al 2001). Note that the models are also tested with lower resolutions (see Section 2.4). Images are then blurred using a first order Butterworth filter where the cut-off frequency is defined using the acceptance angle of 6° as observed in *Gryllus campestris* (Labhart et al 1984). The Cricket Eye Model images sampled at the home positions within each of the test environments are shown in figure 2 under the corresponding original images.

The specific Homing Model under test is then used to calculate the home vector at the current location. The Motor Output then selects the cardinal direction most closely matching the home vector and updates the agent position to the nearest grid location in the defined direction. As described below, different levels of noise can be added to the home vector direction after it has been calculated and before the movement is determined. Note however that the Run Down model (section 3) is an exception as it does not calculate an explicit home vector but instead moves first and then evaluates, on the basis of image difference, whether to continue in the same direction or randomly try a new direction. If the agent attempts to move to a location outwith the image-database (equivalent to the cricket encountering the arena wall) then the agent is forced to move to the closest available location to the right. This procedure keeps the agent within the image-database and simulates a simple wall following response when the wall is encountered.

The process described above iterates until either the cool spot location (grid position 16,12) is found or the path-length exceeds 300 steps. This stop condition was selected as crickets trials were ended after 300 seconds

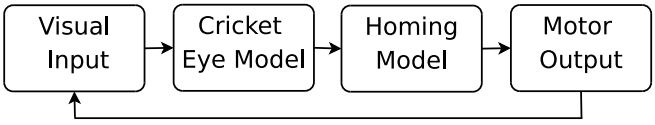
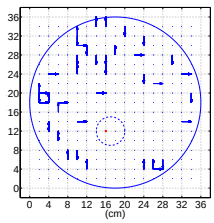
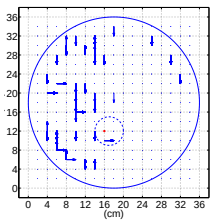


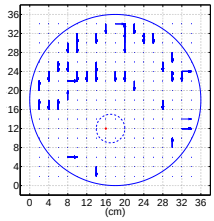
Fig. 3: System Overview: all processing steps are kept consistent except the method used to determine the homing direction.



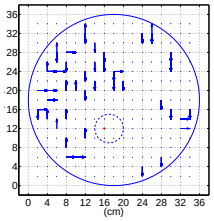
(a) DC Start locations



(b) BW Start locations



(c) DL Start locations



(d) NS Start locations

Fig. 4: Cricket start locations and orientations in the final four trials are approximated giving the 48 image database start positions and orientations shown for each test environment. These start-positions, and where appropriate orientations, are used in all subsequent model testing.

if the cool-spot had not been located. The ability of the homing models to replicate the cricket behaviour is assessed by recording the homing path lengths produced within each of the three environments. Homing trials are initiated from the same 48 positions (and where appropriate, orientations) from which cricket trials were initiated. It should be noted that as cricket start locations were chosen at random in the behavioural study this leads to somewhat different start positions and orientations within each environment. The start locations and orientations used within each environment are shown in figure 4. We also calculate the home vector from every grid position so as to visualise the overall effectiveness of the homing method. Home vector plots are shown with the average angular error (AAE) which is calculated by taking the mean error between home vector calculated at each location and the known ideal home vector and provides a simple measure of model accuracy.

2.3 Control trials

In the original experiments a completely dark arena acted as the control. This would produce a uniformly black image set, making any visual homing process ineffective. To generate comparable control data for the simulation, we generated random search paths from same 48 start positions as the cricket, as outlined previously. At each time step the agent moves randomly in one of the four possible directions with equal probability, until either the cool spot is encountered or the path-length exceeds 300 steps. This process produces path lengths with comparable median (289) and upper and lower quartiles (300 and 125) as the observed path durations of crickets in the dark (median 258, quartiles 300 and 79).

2.4 Parameter Tuning

In their original forms, the different visual homing models implemented in this study utilise various further pre-processing steps such as image smoothing, or using only a certain area of the image, to improve performance. Rather than make any assumptions about such image processing in crickets, which might bias the results towards one or other homing model, we instead use an optimisation procedure to tune the pre-processing parameters individually for each model. That is, optimisation is done by exhaustively searching through all possible parameter combinations and determining which parameters (if any) produce the same performance trend across the different visual environments as we observed for the crickets. Parameters are optimised according to two successive criteria:

1. Median path-lengths in all visual environments must statistically outperform the control.
2. Median path-lengths in the NS environment must statistically outperform homing within both the DL and BW environments.

Statistical comparison of the homing performance is performed using the Wilcoxon rank-sum test. For each model, the outcome from the parameter setting that produces the closest match of medians and interquartile differences to the cricket data will be presented in the results, and the pattern of parameter settings that pass or fail the criteria discussed.

This search through parameter space also allows us to compare the relative robustness of the the different models. The free parameters optimised by the models are:

Image Smoothing: In many studies, visual homing is performed on images that have been highly low-pass filtered; a processing step easily performed in neural hardware. Vardy (2005) outlines a Gaussian low-pass filtering scheme where images are convolved with the kernel:

$$G = [0.005\ 0.061\ 0.242\ 0.383\ 0.242\ 0.061\ 0.005] \quad (1)$$

in the x and then the y direction. It is shown that successive applications of this filter is comparable to convolving with a single larger Gaussian. Thus smoothing the cricket eye images using Gaussians of various sizes is achieved by optimising for 0, 1, 3 or 5 applications of this Gaussian Filter.

Down-sampling rate: As the cricket eye model unwraps the raw image at a resolution of 1°, down-sampling can be thought of as resetting the visual acuity to 1°, 2°, or 4°. This is achieved by sampling every 1, 2 or 4 pixels both horizontally and vertically from those supplied by the cricket eye model. It is worth noting that these values are closely matched to the visual acuity of *Gryllus campestris* (1°), the honeybee eye *Apis meliphora* (1.7°), and desert ant *Cataglyphis bicolor* (4°) (Land 1997). Although it should also be noted that the acceptance angle of the honeybee and ant eye are 2.6° and 3° degrees respectively compared with the cricket eye acceptance angle of 6° which would result in less initial smoothing.

Image area used: Each model also optimises for the portion of the input image used to calculate the home vector. The input image from the Cricket Eye Model consists of +/- 20° of elevation around the horizon sampled at 1° azimuth giving a maximum image size of 360*41 pixels. Each model then selects whether to use:

1. 10° above the horizon.
2. 10° below the horizon.
3. Horizon pixels only
4. +/- 5° around the horizon.
5. +/- 10° around the horizon.
6. +/- 20° around the horizon.

Note the optic flow models cannot use the horizon pixels only image area as it does not allow vertical image gradients to be calculated.

Addition of noise to home vectors: The addition of noise to the homing signal aids certain models that otherwise become trapped in deterministic loops in the grid leading to high failure rates despite good general approximation of the home direction. At each homing iteration, noise is generated through the addition of an error term to the derived home vector. The error term is randomly selected from a circular normal (Von Mises) distribution with a mean of zero and variable concentration parameter κ (analogous to standard deviation in a non-circular normal distribution). Models optimise for the concentration parameter κ which ranges from 0 (no noise) to 90° in increments of 10°. The maximum concentration parameter setting of 90° results in a 62% chance that additional noise corrupts the home vector by more than 90°, resulting in movement in a random orientation with respect to the generated home vector.

Image Type: Some models can operate with black and white, rather than greyscale, images. For the COMALV and gradient descent methods we optimise for either image type, where black and white images are generated by thresholding the output of the cricket eye model at the median greyscale value. The ALV model inherently converts images to black and white to define landmarks and thus only optimises for greyscale images. The differential optic flow models can only operate on greyscale images.

3 Models of Visual Homing

As the explicit aim of this study is to investigate the strategies employed by crickets when returning to the cool spot, only visual homing models that can be considered "biologically plausible" are implemented. A criterion for the selection of such models was outlined by Vardy (2005):

1. As the insect brain has limited neural capacity, models must not be so computationally complex that no convincing argument can be made for their implementation in the neural hardware of an insect.
2. As the retinotopic mapping is maintained throughout sensory pathways from the insect eye through the optic lobes then all calculations required by the model must be theoretically possible using local retinotopic calculations rather than global searches in the image space.

Applying the above criterion, the following six models were selected for use in this study:

1. Average Landmark Vector Model.
2. Centre-of-Mass Average Landmark Vector Model.
3. Differential Optic Flow Models:
   (a) First Order
   (b) Second Order
4. Gradient Descent Models using:
   (a) GradDescent
   (b) RunDown

For each model we present a brief outline of how it works, the results for the optimised model after parameter tuning, and discussion of why the homing behaviour succeeds or fails in the different environments.

3.1 Average Landmark Vector Model

The Average Landmark Vector (ALV) model is a derivative of the classic snapshot model (Cartwright and Collett 1983) offering an extremely parsimonious system which also bypasses the correspondence problem (Lambrinos et al 2000). At the home location (H) the processed home image is further reduced to a 1D vector often referred to as the horizon ring. Vertical edges in the

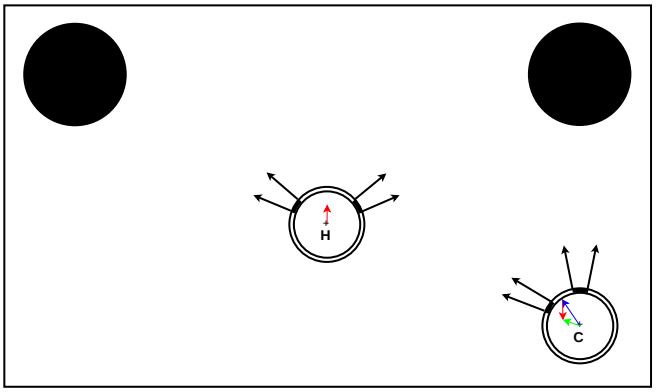


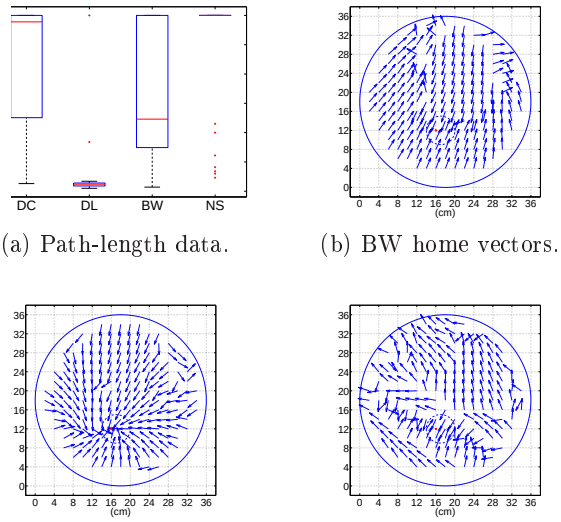
Fig. 5: The ALV is calculated at the home location H (red arrow) and also at the current location C (blue arrow). Through a simple vector subtraction the home vector (green arrow) is calculated.

horizon ring are then identified, and a unit vector drawn towards each edge (Figure 5). Taking the mean of the vectors across the entire image provides the home location average landmark vector (ALV_H). When the agent is moved to a distant location (C) the current average landmark vector (ALV_C) is calculated in the same manner. The home vector can then be calculated through a simple vector subtraction ($h = ALV_C - ALV_H$).

The ALV model has been shown to successfully home in simulated environments consisting of distinct landmarks within an infinite horizon background (Lambrinos et al 2000). Homing has also been successfully achieved on a mobile robot in a desert habitat with black cylinders provided as landmarks, and in a university lobby (Möller et al 2001).

Simple edge detection techniques such as those employed by Lambrinos et al (2000) are found inadequate for homing within the cricket surrounds, specifically within NS. Therefore, the edge detection procedure implemented by Möller et al (2001) to achieve successful homing in the university lobby is adopted. Firstly, the snapshot image image is column averaged producing a 1-D greyscale image. Regions of this 1-D snapshot image showing consistent intensity levels are identified as stable points in the image, and their intensity values recorded. The midpoint between the largest difference in stable points is then defined as the threshold value. Selection of the threshold value in this manner is intended to offer robustness to light and contrast variance. Any points where the image intensity crosses this threshold are considered edges and are associated with a unit vector. The threshold calculated at the home location is then used throughout when defining all subsequent edges at displaced locations.

Figure 6 shows the homing path lengths produced by the ALV model in the various image databases, and the home vectors generated in each environment. Table 2 shows the p-values of the statistical comparisons of homing paths.



(a) Path-length data. (b) BW home vectors.

(c) DL home vectors. (d) NS home vectors.

Fig. 6: (a) ALV Model path lengths within the three image databases and control condition. Optimal parameter settings: downsampling rate = 1, No. of Gaussians=1, image region used = horizon pixels only, and concentration parameter κ=10° using greyscale images. (b), (c), (d) Home vectors generated using the optimal parameter settings in the BW, DL and NS surrounds respectively. Note that the single bar at 300 for the NS results indicates that all homing trials save outliers reached the stop condition. AAE: BW=58°, DL=12°, NS=93°

Table 2: P-values calculated for comparisons between path-lengths the various image databases using the ALV model. Note the significant difference of NS and control is in the wrong direction (NS worse than control).

	DC	**DL**	**BW**	**NS**
DC	X	<0.01	0.12	<0.01
DL	<0.01	X	<0.01	<0.01
BW	0.12	<0.01	X	<0.01
NS	<0.01	<0.01	<0.01	X

It is clear that the ALV Model does not reproduce the same performance trend as observed in crickets. Indeed no parameter setting produced shorter path lengths in all visual environments when compared to the control; the first performance criterion. This failure is caused by inaccurate home vector generation within NS due to the difficulty in consistently defining edges within such an environment. The slightly improved path lengths produced within BW are attributable to the background intensity gradient outlined previously which the ALV Model defines as a single landmark that is generally detected across arena positions allowing somewhat accurate home vector generation. In contrast, within the distinct landmarks surround, where edges are easily identified, home

Table 3: Comparison of the number of edges identified in the home image compared to the mean and standard deviation found at all subsequent image positions in each environment.

	Home Edges	μ edges	σ
DL	6	5.9320	0.8897
BW	2	1.9612	0.3922
NS	4	4.4951	1.8401

vectors are accurate from most regions of the arena, resulting in short path lengths.

Table 3 shows the mean number of edges detected across image positions within each of the image databases compared with the number of edges found in the home image of those image-sets. Within NS the high degree of variance in detected edges catastrophically affects correct home-vector calculation. Despite a higher degree of edge detection accuracy within BW, the small number of detectable edges also results in inaccurate home vector calculation when edges are incorrectly defined. Furthermore, the reduced intensity range within BW images renders the model susceptible to noise and makes accurate ALV computation difficult resulting in a high AAE. In contrast, within DL six edges are detected in the snapshot image. The larger number of distinguished edges in this environment increases the robustness of the ALV when edges are incorrectly identified, or missed, resulting in the improved performance displayed. It may be possible to improve the performance of the ALV Model in NS using a more sophisticated feature extraction algorithm. However as homing is close to optimal in DL it seems unlikely that any such enhancement would produce statistically superior homing in the NS surround. We thus dismiss the ALV Model as a strategy used by the crickets to relocate the cool spot.

3.2 Centre-of-Mass Average Landmark Vector Model

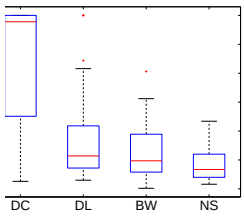
The Centre-of-Mass Average Landmark Vector (COMALV) Model (Hafner 2001) as its name suggests is conceptually similar to the ALV model. Vectors are again derived at both the home and current locations and the home vector through vector subtraction. However rather than using edges or landmarks to calculate these vectors, the COMALV model stores the vector projecting to the 'centre of mass' in each image:

$$COMALV = \sum \left(I(\theta) \begin{pmatrix} cos(\theta) \\ sin(\theta) \end{pmatrix} \right) \quad (2)$$

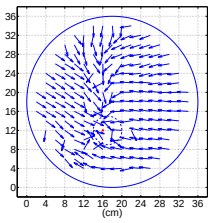
where $I(\theta)$ is the image intensity value at the bearing indicated by θ in the one-dimensional input image.

The COMALV Model was originally derived through the use of a learning procedure on an artificial neural network, but in this study we directly implemented the above equations. The COMALV model has performed successful homing trials in simulation and also on a mobile robot within an office environment. Furthermore it is computationally cheap and bypasses both correspondence and feature extraction issues. The model however failed to home successfully in various image databases without re-tuning of parameters (Vardy 2005). Nevertheless as Vardy's image database consisted of an office environment where the homing capability of insects is unknown the COMALV is included in this study.

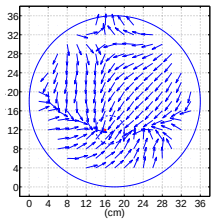
Figure 7 shows the homing path lengths produced by the COMALV model in the various image databases, and the home vectors generated in each environment. Table 4 shows the p-values of the statistical comparisons of homing paths.



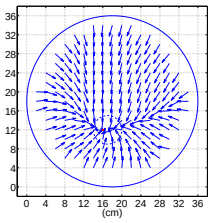
(a) Path-length data.



(b) BW home vectors.



(c) DL home vectors.



(d) NS home vectors.

Fig. 7: (a) COMALV Model path lengths within the three image databases and control condition. Optimal parameter settings: downsampling rate = 2, No. of Gaussians=0, image region used = +/- 20° around the horizon, concentration parameter κ=80°, using black and white images. (b), (c), (d) Home vectors generated using the optimal parameter settings in the BW, DL and NS surrounds respectively. AAE: BW=32°, DL=38°, NS=14°

Table 4: P-values calculated for comparisons between path-lengths the various image databases using the COMALV model.

	DC	DL	BW	NS
DC	X	<0.01	<0.01	<0.01
DL	<0.01	X	0.28	<0.01
BW	<0.01	0.28	X	0.01
NS	<0.01	<0.01	0.01	X

Parameter optimisation of the COMALV model found 80 parameter combinations that pass both performance criteria. These settings include every possible smoothing and acuity variant at least once indicating a robustness to these pre-processing procedures. In contrast, the only successful image region is +/-20° around the horizon, and the image type is restricted to black and white images. All parameter settings require some level of noise for successful homing.

The selection of only the maximum image region setting is due the interaction of the landmarks and the background intensity gradient within the DL surround. COMALVs obtained within both BW and NS generally orient towards an attractor in the environment when sampled across image locations. For example within BW, COMALVs are oriented towards the peak of the background intensity gradient outlined previously. An attractor is necessary such that when the snapshot COMALV is subtracted from the current COMALV an appropriate angular offset is present resulting in correct home vector computation. However within DL, for smaller image region settings, no such attractor exists as the T shaped landmark and background intensity peak coincide. This flattens the intensity gradient and removes the presence of a prominent COMALV within DL. However, when +/-20° images are presented, the proportion of blank arena wall to landmarks is increased, introducing a prominent attractor to which COMALVs orient and resulting in the improved homing observed.

It was found that the use of greyscale images had a catastrophic effect in the BW environment, caused by a subset of home vectors located near the home position the snapshot which are inverted with respect to the true home direction across many parameter settings. Homing agents would be deflected from the goal. The incorrect home vector direction was a result of the magnitude of the current COMALVs at these locations exceeding that of home COMALV; often only by a small amount but this is sufficient to produce a small but incorrectly oriented home vector. This problem could be partially circumvented by adding sufficient noise that agents near the goal would sometimes reach it instead of being deflected, but such high noise levels degraded performance in the DL surround to chance levels. No parameter combination using greyscale images could be found where path lengths within both DL and BW are statistically superior to the control.

The impact of these anomalous home vectors within the BW surround is probably magnified by the use of an image-database rather than a fully autonomous robot study where images would be generated repeatedly across trials. There also exists a number of simple modifications to the COMALV Model that may help overcome such deficiencies such as defining a minimum threshold between magnitudes that should be reached before home vectors are computed, weighing trust in home vector relative to magnitude, image normalisation prior to COMALV calculation, or the use of a momentum component that would push agents past erroneous home vectors. However none of these model extensions were implemented in this study, given that the use of blank and white images was sufficient to produce cricket-like results.

3.3 Differential Optic Flow Models

Building upon the finding that successful block-matching models of visual homing are dependent upon low rather than high frequency components of images, Vardy derived two equivalent models based on classic differential optic flow techniques (Vardy 2005). The differential models perform only local searches for image correspondences and therefore fulfill biological plausibility constraints failed by block-matching methods.

The First Order (FO) model rests on the assumption that pixel intensities are maintained across images such that:

$$H(x, y) = C(x + u, y + v) \tag{3}$$

where H is the intensity of the pixel at image position (x, y) in the home image and C is the intensity of the same pixel at its new location in the current image given by summing the previous pixel location with the translation vector (u, v) caused by agent movement.

The Second Order (SO) model assumes that intensity gradients rather than pixel intensities are maintained across images such that:

$$H_x(x, y) = C_x(x + u, y + v) \tag{4}$$

$$H_y(x, y) = C_y(x + u, y + v) \tag{5}$$

where H_x and H_y are the partial derivatives of the pixel intensity at image position (x, y) in the home image and C_x and C_y are the partial derivatives of the intensity of the same pixel at its new location in the current image given by summing the previous pixel location with the translation vector (u, v) caused by agent movement.

Differential models seek to calculate the translation vector (u, v) of each pixel by calculating the intensity gradients (FO model), or the second derivative of the intensity gradient (SO model), surrounding the pixel in question. This allows the translation vector orientation to be calculated locally, which is then converted into a home-vector through an approximate vector mapping technique. That is, knowledge of the robot hardware allows the translation vector existing in image space to be transformed into a home vector in robot space. As differential models derive home vectors at all pixels in the image, the overall home-vector is computed by taking the mean of all home vectors across pixel locations.

Differential methods have classically been applied to optic flow problems where pixel translation is small between successive images and therefore intensity gradients are robust ensuring good translation vector calculation. The success of these models in homing tasks therefore is somewhat counter-intuitive as the scale of agent translation between image captures causes large pixel translations between home and current image. Such image shifts are shown to have catastrophic effects on correct translation vector accuracy. However Vardy demonstrated that incorrect home vectors are uncorrelated and therefore when averaged they generally cancel each other out. Moreover in the focus of expansion and contraction in the image, pixel movement remains small such that the small image translation assumption is valid. This allows correct home vectors to be calculated at pixel locations within these regions. These correct and correlated home vectors dominate when home-vectors are averaged across pixel locations (known as the democracy effect) producing accurate overall home-vector and the positive homing results in the environments outlined below.

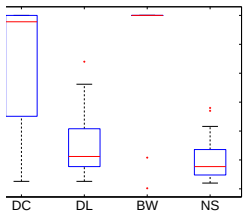
Although the procedure outlined above may sound computationally complex, differential optic flow models remain biologically plausible. The majority of the model calculations are local and are ideally suited to parallel computation as could be performed retinotopically by insects. The differential visual homing models have been shown to home successfully within a number of indoor image databases such as an office environment, and a university hall-way (Vardy 2005).

First Order Differential Model

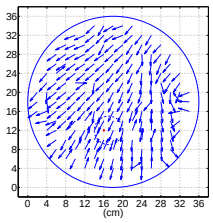
Figure 8 shows the homing path lengths produced by the FO model in the various image databases, and the home vectors generated in each environment. Table 5 shows the p-values of the statistical comparisons of homing paths.

Parameter optimisation of the FO Model found no parameter settings that passed the first performance criterion where improved path-lengths are sought in all visual environments when compared with the control. Despite performing excellently within both NS and DL, the FO Model fails to generate accurate home vectors within the BW surround, where it never outperforms the control. This is because the lack of significant intensity variations within the BW environment does not allow a sufficient number of correct pixel-wise home-vectors to be calculated such that when pixel-wise home vectors are averaged, the correct home vector prevails.

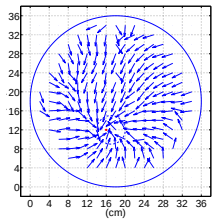
The failure of the FO Model to successfully outperform the control when homing within the BW surround contrasts with the observed behaviour of the cricket and we therefore dismiss the FO Model as a candidate visual homing technique used by crickets.



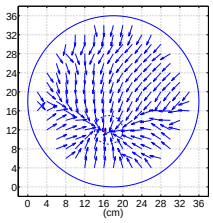
(a) Path-length data.



(b) BW home vectors.



(c) DL home vectors.



(d) NS home vectors.

Fig. 8: (a) FO Model path lengths within the three image databases and control condition. Optimal parameter settings: downsampling rate = 2, No. of Gaussians=1, image region used = +/-10° around the horizon, and concentration parameter κ=90° using greyscale images. (b), (c), (d) Home vectors generated using the optimal parameter settings in the BW, DL and NS surrounds respectively. AAE: BW=66°, DL=20°, NS=15°

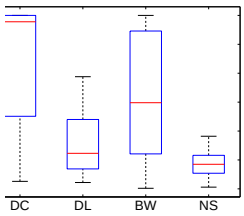
Table 5: P-values calculated for comparisons between path-lengths the various image databases using the FO model. Note the significant difference of BW and control is in the wrong direction (BW worse than control).

	DC	**DL**	**BW**	**NS**
DC	X	<0.01	<0.01	<0.01
DL	<0.01	X	<0.01	<0.01
BW	<0.01	<0.01	X	<0.01
NS	<0.01	<0.01	<0.01	X

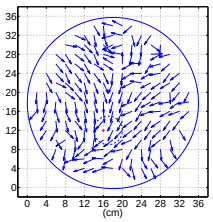
Second Order Differential Model

Figure 9 shows the homing path lengths produced by the SO model in the various image databases, and the home vectors generated in each environment. Table 6 shows the p-values of the statistical comparisons of homing paths. Note the significant difference between BW and control is in the wrong direction (BW worse than control)

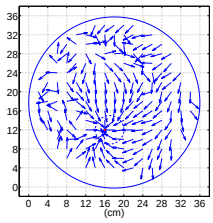
Parameter tuning of the SO Model found five parameter settings that passed both performance criteria. These are comprised of two distinct settings buoyed by the addition of large noise terms. The first setting applies one Gaussian filter, to maximally sampled +/- 20° with concentration parameters κ=70°, 80°, and 90°. The second setting applies no smoothing +/- 20° images downsampled at 2° with concentration parameters of κ=70°



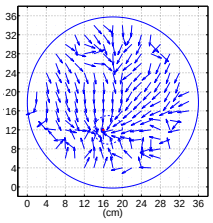
(a) Path-length data.



(b) BW home vectors.



(c) DL home vectors.



(d) NS home vectors.

Fig. 9: (a) SO Model path lengths within the three image databases and control condition. Optimal parameter settings: downsampling rate = 1, No. of Gaussians=1 and image region used = +/- 20° around the horizon, and concentration parameter κ=80° using greyscale images. (b), (c), (d) Home vectors generated using the optimal parameter settings in the BW, DL and NS surrounds respectively. AAE: BW=53°, DL=46°, NS=30°

and 80°. The SO Model was tuned for only image type=greyscale as explained in the methods.

As with the FO Model, homing within BW rarely outperforms the control due to the same absence of robust intensity gradients required by differential optic flow models to accurately compute home vectors. The increased accuracy of the home vectors generated by the SO Model within BW may have been expected as some image gradients are likely to be maintained even where individual pixel intensities are not. The use of minimal levels of downsampling and image smoothing increases the robustness of such gradients where they exist. Moreover, the use of the largest image region increases the influence of correct and correlated home vectors where they can be generated. However, despite this improvement the resultant path-lengths still fail to out-perform the control without the addition of a substantial noise term.

3.4 Gradient Descent Models

Gradient descent models of visual homing have their roots in the finding that the pixel-wise intensity difference between aligned images taken from different locations tends to increase smoothly and monotonically with distance (Zeil et al 2003). Plotting the difference between

Table 6: P-values calculated for comparisons between path-lengths the various image databases using the SO model.

	DC	**DL**	**BW**	**NS**
DC	X	<0.01	0.01	<0.01
DL	<0.01	X	<0.01	<0.01
BW	0.01	<0.01	X	<0.01
NS	<0.01	<0.01	<0.01	X

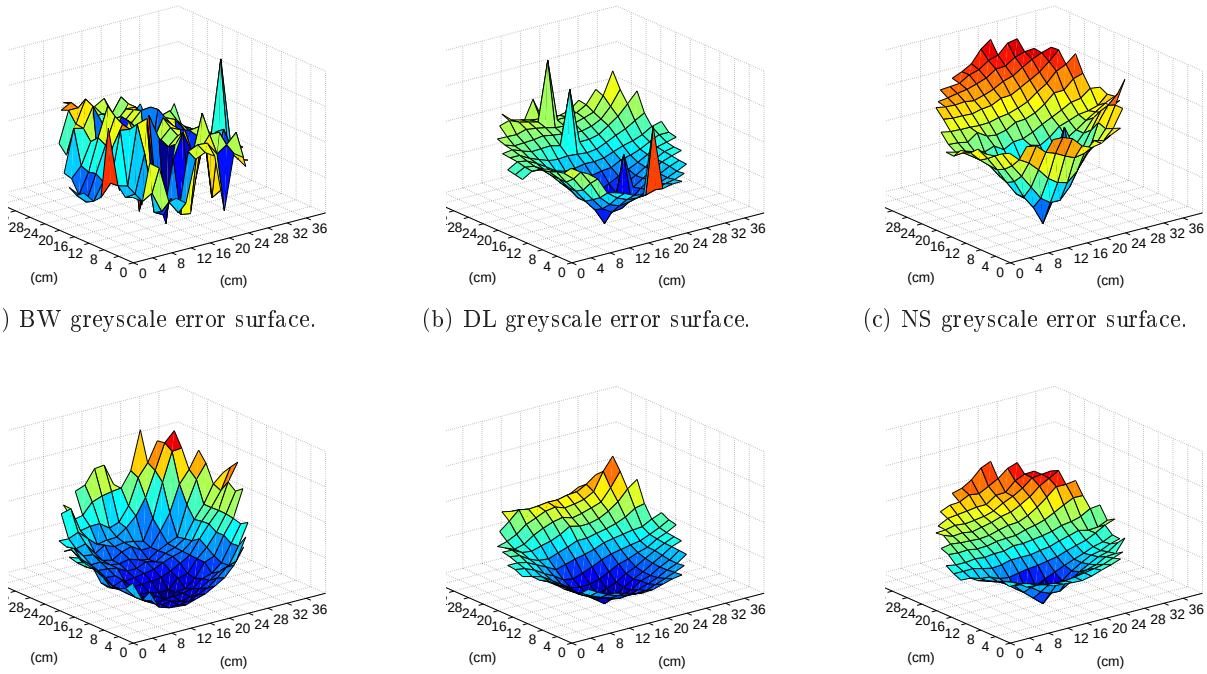
images across locations therefore reveals a sloping surface where the minimum corresponds to the target location. Figure 10 shows the difference surfaces calculated within the various image databases using the error metrics further detailed below.

By sampling the image difference at a number of locations in the environment (by either simulated or actual agent movement), simple gradient descent algorithms can utilise the error slope to return to the goal. Gradient descent models have been shown capable of homing in natural scenes (Zeil et al 2003), in indoor environments using image databases (Zampoglou et al 2006), (Vardy 2005), and also on a mobile robot in a laboratory environment (Zampoglou et al 2006). Furthermore Gradient Descent models are computationally cheap, and ideally suited to parallel computation as is expected in a retinotopic structure.

Previous Gradient Descent models generally make use of only greyscale images and both the pixelwise root-mean-square (RMS) (?)Zeil2003) error and sum-squared difference (SSD) (?)Vardy2005 have been implemented as error metrics. In this study the error metric shall be the pixelwise RMS for use when both greyscale and black and white images are presented. It should be noted that use of the SSD metric was investigated but changing the error metric had no effect on the resultant path lengths.

GradDescent Homing Method

Vardy (2005) outlined the GradDescent method of homing which samples image differences at multiple locations such that an explicit home vector can be derived. Thus homing can proceed as for the other models implemented in this work. GradDescent samples image differences in all four cardinal directions surrounding the current location. Comparison of the values in the cardinal directions with that at the central point allows a difference vector to be drawn in each direction. Taking the mean of the difference vectors indicates the orientation producing the largest increase in image difference and thus the home-direction is correspondingly calculated by inverting the mean difference vector. In the implementation of GradDescent in this study when locations outside of the current image database are to be sampled the difference values are set to the maximum possible value such that the agent moves away from the wall.



(a) BW greyscale error surface. (b) DL greyscale error surface. (c) NS greyscale error surface.

(d) BW black and white error surface. (e) DL black and white error surface. (f) NS black and white error surface.

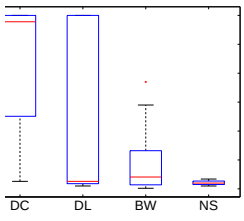
Fig. 10: RMS error surfaces within the three image data-bases. Note that the parameter settings are constant for both image types used: Number of Gaussians=0, downsampling rate=1, image regions=+/- 20° around the horizon. (a) - (c) Error surfaces generated using greyscale images in the BW, DL and NS surrounds respectively. (d) - (f) Error surfaces generated using black and white images in the BW, DL and NS surrounds respectively.

Figure 11 shows the homing path lengths produced by the GradDescent algorithm in the various image-sets, and the home vectors generated in each environment. Table 7 shows the p-values of the statistical comparisons of homing paths.

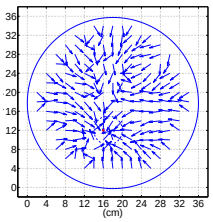
Table 7: P-values calculated for comparisons between path-lengths the various image databases using the GradDescent model.

	DC	**DL**	**BW**	**NS**
DC	X	<0.01	<0.01	<0.01
DL	<0.01	X	0.3	<0.01
BW	<0.01	0.3	X	0.03
NS	<0.01	<0.01	0.03	X

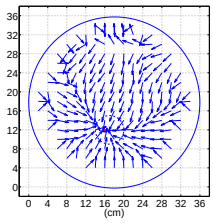
Parameter tuning of the GradDescent method found six parameter settings that passed both performance criteria. The image region used is limited to +/-20° around the horizon, and the downsampling rate to 2 or 4. However the model seems robust to smoothing with all settings except no smoothing being used. The image type used is fixed as black and white. All six parameter settings require a small amount of additive noise (max κ=30°, μ=16.67°, σ=8.16° across the parameter set). When greyscale images are used, homing paths within BW fail to statistically outperform the control across all parameter settings. This failure is caused by the error surface generated within BW (Figure 10(a)) which appears noisy and littered with local minima making successful and repeatable homing almost impossible. Conversion of the input images to black and white removes a dimensionality of the data producing smoothed error surfaces across image-sets, but with particularly profound effect on the BW error surface (Figure 10(d)). The use of black and white images in preference to greyscale, improves homing across image-sets to the extent that statistically separating the performance between paths generated within NS and those generated within the other surrounds becomes the main cause of model failure. The use of the largest image type in conjunction with downsampling increases the proportion of blank arena wall in comparison to landmarks and thus degrades the error surface in DL, without overly effecting the error surface in NS. With the further addition of noise statistical separation can be achieved, at the expense of a much higher variance in the DL paths than was found in the cricket data.



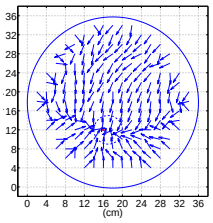
(a) Path-length data.



(b) BW home vectors.



(c) DL home vectors.



(d) NS home vectors.

Fig. 11: (a) GradDescent Model path lengths within the three image databases and control condition. Optimal parameter settings: acuity = 4°, No. of Gaussians=1, image region=+/- 20° around the horizon, concentration parameter κ=10° and using black and white image type. (b), (c), (d) Home vectors generated using the optimal parameter settings in the BW, DL and NS surrounds respectively. AAE: BW=23°, DL=26°, NS=27°

RunDown Homing Method

In contrast to the other methods described so far, the RunDown homing method (Zeil et al 2003) does not compute a home vector at all but instead uses agent movement to assess the current error gradient. The agent's first movement in our simulation is determined by the initial orientation of the cricket in the corresponding trial (figure 4). At each iteration the current image difference is compared to the that computed when at the previous location. If the movement resulted in a decrease in error then the agent continues in its current direction. However, if the error value is equal or larger the agent performs a 90° turn, with the orientation selected at random. The entire procedure is then repeated until the home location is found. The random turning inherently incorporates noise into this approach.

The resulting path-lengths, and statistical comparisons are shown in figure 12 and table 8 respectively. Optimisation of parameters for the RunDown algorithm found six parameter settings successfully passing the performance criteria. The RunDown method seems robust to most parameter settings with all possible smoothing, downsampling and image regions (excluding horizon pixels only) being used. However, as with the GradDescent method, if greyscale images are used, path-lengths in BW are never significantly better than the control. With

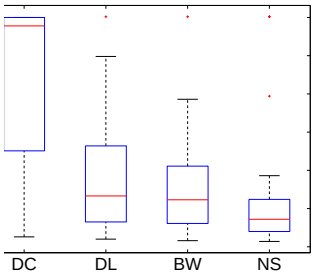


Fig. 12: Box plots showing path-lengths obtained using the RunDown model within the three image databases and control condition. Optimal parameter settings: downsampling rate = 1, No. of Gaussians=3 and image region used = +/-20° around the horizon using black and white images.

Table 8: P-values calculated for comparisons between path-lengths the various image databases using RunDown Model.

	DC	**DL**	**BW**	**NS**
DC	X	<0.01	0.02	<0.01
DL	<0.01	X	0.01	<0.01
BW	0.02	0.01	X	<0.01
NS	<0.01	<0.01	<0.01	X

black and white images, many parameter settings pass the first criteria of significantly shorter paths in all visual surrounds than the control, but though most combinations also show the trend of shorter paths for NS, this is not statistically significant except for the six settings mentioned.

4 Conclusions

Six biologically plausible models of visual homing have been implemented and tested using image data-sets taken directly from the visual environments used in homing experiments on crickets. Each model was assessed for its ability to replicate the performance trends observed for crickets homing with different visual surrounds (Wessnitzer et al 2008). That is: that homing was better with any visual surround than in the dark and that homing was better with a natural scene than with distinct landmarks and blank walls. Homing paths were initiated from the same start-points, and where appropriate start directions, as the crickets. Additionally, across the model implementations, aspects of visual pre-processing and movement control were either held constant or optimised for each model by an exhaustive parameter search.

Table 9 summarises the results of the parameter search, indicating the number of possible parameter permutations (which, as explained in the methods, could vary

Homing model	Number of parameter permutations tested	Control outperformed in all visual surrounds	DL and BW outperformed by NS
ALV	**720**	**0**	**0**
COMALV	**1440**	**102**	**80**
FO	**600**	**0**	**0**
SO	**600**	**5**	**5**
GradDescent	**1440**	**426**	**6**
RunDown	**144**	**41**	**6**

Table 9: Summary of the model parameter tuning results. The maximum parameter permutations are shown for each model. The the numbers in the subsequent columns indicate the number of parameter permutations that failed to pass the performance indicated by the heading.

for different models) and the number of permutations that met the performance criteria for each model.

The Average Landmark Vector and First Order Differential models were found to be incapable of reproducing the performance trends of crickets. The ALV model was unable to home in the natural scene, and always produced the best homing with distinct landmarks. Similar results might be expected for any feature-based homing algorithm that relies on reliable extraction of landmark features. The FO model in contrast homes well within both DL and NS but performance within the BW surround is statistically worse than the control condition. This can be explained by the fact that the useful homing information for this model is contained largely in the focus of expansion and contraction. In the BW environment, the foci always coincide with regions of bare arena wall where intensity values are broadly similar and thus correspondences are difficult to accurately assess. We therefore dismiss the Average Landmark Vector and First Order Differential models as homing strategies employed by crickets.

The increased robustness of image gradients compared with absolute pixel intensities within the blank walls surround enabled the Second Order Differential Model to produce some parameter combinations that passed the performance criteria. However this is dependent on large levels of additive noise which are not only required to improve homing within BW but also to increase the performance gap between NS and DL. As a result, even with the optimal parameter setting, which minimised for both the difference in medians and interquartile range between cricket and model data, the median and variance generated within BW are substantially greater than the cricket homing times when tested in the same environment. We note that though not included as a specific criterion, crickets show no significant difference for DL and BW whereas this model always produces significantly worse results in BW than DL.

Gradient descent based models of visual homing were found capable of reproducing the performance trend of crickets. However success could not be found when grey-scale images were used due to the noisy error surface within the BW surround. Through the use of black and white images rather than grey-scale error surfaces are smoothed such that homing is successful in all environments. Indeed this smoothing makes performance within NS and the other tests environments difficult to statistically separate. The GradDescent homing method successfully achieved statistical significance between NS and both DL and BW by downsampling the image at the maximum rate, which reduced the robustness of the error surface within DL, in combination with a small noise term. However, even with the optimal parameter choice, this produces much higher variance in path lengths in DL than observed for crickets. Given that this method also requires the sampling of all cardinal directions before moving in the home direction, it is seems less plausible than the simple RunDown method.

The RunDown method could successfully reproduce the right pattern of statistical differences across the visual environments and also provided a reasonable match to the observed median and spread in the cricket data. The COMALV Model also accurately reproduced the performance of crickets. The performance criterion was reached for 6 out of 144 possible parameter combinations for RunDown and for 80 out of 1440 possibilities for COMALV. For both methods, the performance criteria were only met when black and white images were used as input. This reduction in information dimensionality reduces the effects of noise in the blank walls environment. Both models were successful for all possible levels of image smoothing and image resolution suggesting they may be robust for different insect eye models, such as bee or ant. COMALV required some noise to prevent it from being trapped in loops; RunDown is an inherently noisy procedure. The COMALV algorithm worked only with the largest image type, whereas RunDown was also successful with more restricted fields of view, excepting when only horizon pixels were used.

From these results we conclude that the 'place memory' observed in our original cricket experiment (Wessnitzer et al 2008) can be explained by visual homing (rather than requiring more explicit spatial representations); and that simple calculation of either the image 'centre of mass' or the image difference is not only sufficient, but produces results closer to those observed for the cricket than more complex algorithms requiring feature extraction or optic flow calculation. In passing it is

worth noting that the initially unexpected cricket homing in the Blank Walls environment is accounted for by these models without including the canopy area in the images; the very slight light gradient that existed across the arena was sufficient for homing. This may be important for eliminating unintended cues in any visual orientation experiments on insects.

Both successful methods are computationally cheap. The COMALV method is also cheap in requiring only one vector, rather than a home image, to be stored: although parallel retinotopic processing in the insect brain may mean image storage is also relatively cheap. The reliance of the COMALV Model on the entire image region suggests a possible experimental design to separate the COMALV and RunDown models. If the field of the view of the insect could be limited, either through eye capping or physical barriers, and its impact on the homing ability of insects observed, then the likelihood that the COMALV Model is the homing strategy in use could be inferred. Such screening experiments have sought to infer the portion of the visual scene used by homing wood ants (Fukushi 2001) and desert ants (P. Graham, personal communication), and may offer a modelling environment able to distinguish the homing models.

One limitation of the all models described is that the orientation of the images had to be kept consistent throughout. Such image alignment is generally required for both image-based and feature-based homing models. While it is theoretically possible to infer the rotation of a current image in relation to a reference image such that images may be aligned prior to comparison, and some algorithms carry out this step, it is a non-trivial problem. Usually it is addressed by assuming the insect or agent has a compass to provide rotation information. However within the cricket arena no compass cues (magnetic or polarised) are available and there is no evidence of crickets rotating to align images from the behavioural data.

A possible solution to the alignment problem, where compass information is not available, is that insects may store multiple snapshots at the goal location, while oriented in different directions. Thus when the insect is performing a subsequent homing run, the snapshot memory most closely matching the current world view would be used as a reference for calculating the current error or home vector. In this way locations of importance, which are visited repeatedly would have multiple, strongly reinforced memories and could be approached from many directions without the need for a compass at all. This type of gradual learning of a location may account for the learning curve observed in the cricket behavioural experiments in Wessnitzer et al (2008) and will be subject of further study.

References

Cartwright B, Collett TS (1983) Landmark learning in bees. Journal of Comparative Physiology A 151:521–543, DOI 10.1007/BF00605469

Collett T, Collett M (2002) Memory use in insect visual navigation. Nat Rev Neurosci 3(7):542–552, DOI 10.1038/nrn872, URL http://dx.doi.org/10.1038/nrn872

Franz M, Schölkopf B, HMallot, Bülthoff H (1998) Where did i take that snapshot? scene-based homing by image matching. Biological Cybernetics 79:191–202

Fukushi T (2001) Homing in wood ants, formica japonica: use of the skyline panorama. The Journal of Experimental Biology 204:2063–2072

Hafner V (2001) Adaptive homing - robotic exploration tours. Adaptive Behaviour 9:131–141

Labhart T, Hodel B, Valenzuela I (1984) The physiology of the cricket's compound eye with particular reference to the anatomically specialized dorsal rim area. Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology 155:289–296

Labhart T, Petzold J, Helbling H (2001) Spatial integration in polarization-sensitive interneurones of crickets: a survey of evidence, mechanisms and benefits. The Journal of Experimental Biology 204:2423–2430

Lambrinos D, Moller R, Labhart T, Pfeifer R, Wehner R (2000) A mobile robot employing insect strategies for navigation. Robotics and Autonomous Systems 30:39–64, DOI 10.1016/S0921-8890(99)00064-0

Land M (1997) Visual acuity in insects. Annu Rev Entomol 42:147–177

Mizunami M, Weibrecht JM, Strausfeld NJ (1998) Mushroom bodies of the cockroach: their participation in place memory. The Journal of Comparative Neurology 402(4):520–537, DOI 10.1002/(SICI)1096-9861(19981228)402:4<520::AID-CNE6>3.0.CO;2-K

Möller R, Lambrinos D, Roggendorf T, Pfeifer R, Wehner R (2001) Insect Strategies of Visual Homing in Mobile Robots, AAAI Press / The MIT Press, chap 3, pp 37–66

Morris RG, Garrud P, Rawlins JN, O'Keefe J (1982) Place navigation impaired in rats with hippocampal lesions. Nature 297(5868):681–683

Vardy A (2005) Biologically plausible methods for robot visual homing. PhD thesis, Ottawa-Carleton Institute for Computer Science

Wehner R (2003) Desert ant navigation: how miniature brains solve complex tasks. Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology Volume 189(Number 8):579–588, DOI 10.1007/s00359-003-0431-1

Wessnitzer J, Mangan M, Webb B (2008) Place memory in crickets. Proceedings of the Royal Society B: Biological Sciences 275:915–921

Zampoglou M, Szenher M, Webb B (2006) Adaptation of controllers for image-based homing. Adaptive Behaviour 14(4):381–399, DOI 10.1177/1059712306072338

Zeil J, Hofmann M, Chahl J (2003) Catchment areas of panoramic snapshots in outdoor scenes. Optical Society of America Journal 20:450–469