

RESEARCH ARTICLE

How might ants use panoramic views for route navigation?

Andrew Philippides¹, Bart Baddeley¹, Ken Cheng² and Paul Graham^{3,*}

¹School of Informatics, University of Sussex, Falmer, Brighton BN1 9QJ, UK, ²

INTRODUCTION

Many ants use fixed routes to travel between their nest and a profitable foraging ground, e.g. *Cataglyphis bicolor* (Santschi, 1913), *Formica rufa* (Rosengren and Fortelius, 1986), *Cataglyphis fortis* (Wehner et al., 1996), *Melophorus bagoti* (Köhler and Wehner, 2005). Experienced ants use visual landmarks to guide their routes and experiments show that portions of the route can be performed out of sequence and independently of path integration (Collett et al., 1992; Collett et al., 1998; Andel and Wehner, 2004; Köhler and Wehner, 2005). Interest in how this robust behaviour is produced by small-brained animals with their low-resolution visual system has intrigued biologists and roboticists. The subsequent cross-fertilisation of ideas has influenced the approach we have taken in this paper; a computational analysis of the visual information available to ants as they navigate routes through complex natural environments.

Laboratory and field studies using artificial landmarks at the nest (Wehner and Rüber, 1979; Wehner et al., 1996; Åkesson and Wehner, 2002; Narendra et al., 2007) or a feeder (Wolf and Wehner, 2000; Durier et al., 2003; Graham et al., 2004) have shown how visual landmark information can be used to guide the search for an important location. The basic mechanism is for a single view of the world to be stored at the goal location. The difference between the current view of the world and the view from the goal location can subsequently be used to drive the search for that goal, so-called ‘view-based homing’ or ‘snapshot matching’ (Cartwright and Collett, 1983). It is natural to ask whether view-based homing can also be used to guide long natural habitual routes.

Our knowledge of the mechanisms underpinning visually guided routes is less extensive than our understanding of view-based homing. A variety of experiments have highlighted how simple procedural rules can be used to associate directional information with visual landmarks and so guide small portions of a route (Collett

et al., 1992; Collett et al., 1998; Collett et al., 2001; Pratt et al., 2001; Graham and Collett, 2002; Graham et al., 2003; Harris et al., 2007; Collett, 2010). However, we do not have a general understanding of the mechanisms by which ants navigate habitual routes using information from natural visual panoramas. Similarly, we know little about how ants extract information from natural visual environments. It has long been suggested that in natural environments the skyline profile could provide a characteristic signature for a location (Wehner and Rüber, 1979; Wehner et al., 1996) or provide easily identifiable discrete landmarks that can be associated with appropriate directions (Fourcassié, 1991; Fukushi, 2001). Recently, Graham and Cheng demonstrated that a skyline profile generated by an artificial arena can functionally mimic a natural panoramic scene even when colour cues, the distance distribution of objects and orientation relative to a celestial compass radically differ from the ants’ familiar foraging locations (Graham and Cheng, 2009b). However, we do not know what information is extracted from a skyline and how this might be used for route guidance.

Our attempts to address these questions are fundamentally hampered by our lack of understanding of the ants’ perspective of their environment. Here we begin the process of quantifying the visual information available to ants as they move through the world. Specifically, we have asked over what range can the comparison of the current view of the world with a remembered view of the world provide useful navigational information. This is a fundamental and powerful question because it relates to the issue of the level of world knowledge necessary for route performance.

Our approach follows that of Jochen Zeil and colleagues (Zeil et al., 2003; Stürzl and Zeil, 2007), who captured sets of images using a panoramic imaging device within a natural environment of significance to behaving animals; in their case, ground-nesting solitary wasps. By measuring the difference between a reference

image and images from surrounding points, they defined an image difference function (IDF) and showed that, over a few metres, image differences increase monotonically with increasing distance. Insects can return to a goal by monitoring the difference between what they currently see and the stored reference image, then moving so that the difference decreases. Therefore, the presence of a smoothly increasing IDF showed that the information needed for view-based homing is available in unprocessed natural scenes and can be utilised over behaviourally significant distances. This result relies on the camera being aligned to an external frame of reference for all images. However, Zeil and colleagues also showed that the alignment of the camera when the reference image was taken (a proxy for heading information) can be robustly recovered at locations near the goal by comparing the reference image to rotated versions of the current image (Zeil et al., 2003). The orientation at which the current image best matches the reference image will be close to the orientation of the reference image. The implications for route following are stark: an insect with a visual system that is fixed relative to its body axis can recover a heading by rotating until it finds the best match between the current scene and a stored snapshot. Therefore the correct heading for a portion of a learnt route can be specified by a snapshot stored when the insect was previously moving in the correct direction along the route.

The study by Zeil and colleagues (Zeil et al., 2003) laid the foundations for a quantitative analysis of real visual environments

was removed manually. Subsequently, the first row of pixels was set to ON and any 'holes' in the binary map (OFF pixels surrounded by ON pixels) were filled in. Finally, any 'floating' objects (sets of ON pixels surrounded by OFF pixels) were removed, resulting in a binary 'template' image showing ground and foliage as ON and sky as OFF. The skyline was then defined as the height of the highest ON pixel in each azimuth. Skyline extraction is intentionally simple as we did not want the results to be dependent on sophisticated processing or optimised parameter choices. Possible variations in skyline extraction can be safely ignored as we tested variants of the skyline extraction process (e.g. taking the lowest ON pixel at each azimuth, not removing floating objects before selecting the highest ON pixel) and found a negligible effect on the overall results. It is likely that skyline extraction would be straightforward for ants. They have dichromatic vision with peak sensitivities in the UV and green range. A simple UV–green opponent channel would be perfectly suited to extracting the skyline (Möller, 2002).

Whole images

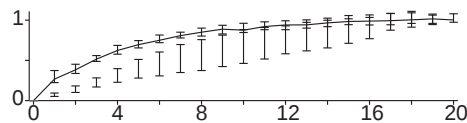
To evaluate properly the information available in an unprocessed scene we had to mitigate any influence of varying light levels or persistent light gradients from sun position, as they may have biased our recorded catchment areas. Firstly, contrast was normalised using histogram equalisation (`histeq` function in Matlab, MathWorks, Natick, MA, USA) of grayscale images, resulting in integer-valued pixels in the range 0–255. Subsequently, using the binary image template delineating sky from not-sky, we homogenised the sky to an intensity of 250 (several other sky intensities were tested – 50, 100, 150, 200 – with little effect on the overall results).

Catchment areas

For each transect we calculated two catchment areas for every image along the transect. The first was based on the IDF, the pixel-based r.m.s. difference between an image and a reference image when the two images are aligned to a common heading. The catchment area of an image is defined as the region within which an agent could return to the location where the reference image was taken by descending the gradient of the IDF (Zeil et al., 2003). The catchment area is defined as the number of consecutive locations spreading out from the reference image where the IDF gradient is positive, relative to the direction of movement, on either side of that location. In Figs 2, 3 and 4, we report the median radius of the catchment areas. This process was undertaken for whole images and scenes encoded as 1D skylines (e.g. Fig. 3A).

The second, the rotational catchment area, is determined from the rotational IDF (RIDF). The RIDF is calculated by evaluating the r.m.s. difference between a reference image and the current image rotated (*in silico*) in steps of 1 deg of azimuth, resulting in a 1×360 RIDF (e.g. Fig. 4A). The minimum value in the RIDF defines an orientation of the current image that gives the closest match with the reference image. In the vicinity of the reference image these orientations will be similar (Zeil et al., 2003). We defined a rotational catchment area as the region spreading out from the location of the reference image where the minimum in the RIDF is less than 45 deg from the true orientation of the reference image.

In a secondary analysis of the information available in the RIDF we applied a simple behaviourally plausible heuristic when calculating the rotational catchment areas. Starting from the goal locations we moved out along the transect. For each image we calculated the RIDF and from this extracted the three most prominent



minima (Fig. 4A). Each minimum represents an orientation where there is a locally optimal match with the reference image. Rather than simply taking the best match, we took the minimum that is closest in orientation to the orientation at the previous location. This heuristic reduces aliasing by favouring RIDF minima close to the previous heading.

transect. The size of catchment areas is strongly dependent on the environment type and, as expected, catchment areas increase as the environments become more open. Nonetheless, even for the cluttered environment, where the appearance of the world can change very quickly (Fig. 1), there is still a gradient in the function relating median IDF to distance up to 6 m from the reference image.

Gradient descent of IDF

We began our analysis of the information available in natural panoramic scenes by considering the IDF for all possible goal positions along our four transects. The IDF is generated by evaluating the pixel-wise r.m.s. difference between the reference image and all other images on the route. The catchment area of each goal image is estimated by looking for changes in the sign of the gradient of the IDF (e.g. Fig. 2A). Estimating the catchment area in this way gave us a direct measure of how useful a single stored snapshot would be for navigation in that environment. We performed this analysis with whole images that had been minimally processed to avoid systematic bias from varying light levels or gradients in sky intensities. We also measured catchment areas for panoramic images sparsely encoded as a 1D skyline profile representing the height of foliage against the sky for each azimuthal direction. For these image sets the IDF was generated by calculating the r.m.s. difference of skyline heights in the reference and route images.

Fig. 2 shows examples of IDF functions for whole images and skylines from a single goal at the midpoint of each transect (Fig. 2A–D), plots of how image difference relates to distance from the reference image (Fig. 2E–H) and also the distribution of catchment areas (Fig. 2I–L) for all possible goals along the

Using a goal image to retrieve orientation

The analysis of IDFs in natural environments shows that panoramic scenes contain information useful for view-based homing over a behaviourally relevant scale. However, if we consider in more detail what is actually required for route guidance, then perhaps stored views can be used more simply. Route guidance requires, in the main, an ant to decide in which direction to go rather than to home accurately to a specific goal or sub-goal. Therefore we investigated the range over which a stored view can be used to recover the orientation at which the reference image was taken. As ants are constrained by their morphology to travel in the direction of their

used for experiments by Kohler and Wehner (Kohler and Wehner, 2005). They were able to show the utility of a skyline code for uniquely defining a place. Their simulation, however, only included grass tussocks local to the region of interest and the reported catchment areas were around 2m. A more realistic simulation of this particular semi-arid environment would have included medium (bushes) and large (trees) objects at a variety of distances (Muser, 2005). Zeil and Stürzl, in accordance with our Fig. 5, have shown that, in natural environments, a rich depth structure underpins robust view-based homing and directly influences catchment areas (Zeil and Stürzl, 2007). Therefore, it is likely that a more realistic world model would have given different results.

By analysing the information available in natural panoramic scenes we have shown that ants might not require a large set of views corresponding to a dense series of locations in order to control a habitual route. In the next section we discuss possible route guidance mechanisms that utilise remembered views.

Implications for route guidance mechanisms

We have shown that stored panoramic scenes from natural environments contain information that can be utilised over reasonable distances for descent in image difference (a proxy for view-based homing methods) and as a visual compass to recover an orientation. These two uses of stored images represent two very different mechanisms by which an ant could control a route. With the first, one would use a view-based homing algorithm to navigate to a sequence of views that are stored from points along the route. However, this intuitive strategy may not be as straightforward as it seems. Images need to be aligned to an external frame of reference requiring neural or behavioural mechanisms. Moreover, attempts to model route guidance as the sequential matching of a series of views (Smith et al., 2007; Smith et al., 2008; Vardy, 2006) have revealed non-trivial issues such as knowing when a sub-goal has been reached and reliably crossing boundaries into the catchment area of the next stored view. It may be that chaining snapshots is overly complicated for route guidance. Routes do not require the accuracy inherent in view-based homing (Collett et al., 1992). Rather, routes require simply the recall of headings appropriate to the current world location. Relevant to this, we have shown that stored natural views contain enough information to be used to recall route-appropriate headings (Zeil et al., 2003; Labrosse, 2006) without a global compass reference. Another attractive property of a mechanism that uses stored views to recall an orientation is that information from comparisons with multiple views can be sensibly polled. For

instance, stored views that represent broadly similar directions could be simultaneously compared with the current view. Heading could then be set by some average, perhaps weighted by similarity, of the outputs across the multiple comparisons. Indeed, a mechanism like this, but with images stored at different headings, can even be used to home to a discrete location (Graham et al., 2010).

Using a stored image to retrieve an orientation is a neat way of controlling a route and, as described above, using this method it would be possible to combine results from simultaneous comparisons with multiple stored views. However, the problem of how to select the most appropriate views to define a route still remains. For instance, an arbitrarily chosen view along a route may contain objects that only appear transiently in the view sequence, which will reduce its effective catchment area. One possibility is that the views along a route segment could be averaged into a single view to remove noisy transient components to focus on the low frequency signal that should robustly define the correct orientation for that portion of the route. Given appropriate processing during route learning, it might be possible to learn a function that maps properties of a changing scene to appropriate headings along the route. Recent abstract models of route guidance based on experimental findings in ants (Harris et al., 2007; Collett, 2010) have shown how route control can be simplified when the route can be described by a smooth function that maps changes in (some aspect of) the visual scene onto a navigational instruction. Theoretical investigation into whether a moving average could extract a useful description of the views experienced along complex natural routes promises to be fascinating.

How to encode visual scenes

Inspired by recent behavioural findings with bees (Towne and Moscrip, 2008) and ants (Graham and Cheng, 2009a; Graham and Cheng, 2009b), we asked whether encoding panoramic scenes as 1D skylines influenced their catchment areas. Our results show that there is not a significant loss in performance between whole image and skyline when performing gradient descent on the IDF. For recovering orientation using the skyline, there is a small drop in performance because aliasing of prominent features is more likely when they are defined solely by their height, disregarding intensity and shape information. Despite this, there are compelling reasons why encoding panoramic scenes as skylines may be functionally successful. The skyline is an economical encoding that is probably

independent of lighting conditions and time of day. We have shown how simple behavioural heuristics can alleviate some aliasing problems, therefore rendering skyline encoding a plausible explanation of how ants parameterise natural scenes.

Insect navigation is a beautiful behaviour that continues to provide insight into insect perception and cognition. We believe a necessary complement to future behavioural studies of navigation is an understanding of the sensory ecology within which an animal behaves. To this end we have analysed ants'-perspective views of a habitat within which desert ant navigation is well studied. Our analysis points to specific route guidance mechanisms that we can look for in future behavioural experiments.

ACKNOWLEDGEMENTS

We thank Matthew Collett for helpful discussions about this work. The work was supported by the Australian Research Council (DP0770300 to K.C.), The Royal Society, and the BBSRC and EPSRC through the Cognitive Foresight scheme (BBF0100521).

REFERENCES

- Åkesson, S. and Wehner, R. (2002). Visual navigation in desert ants *Cataglyphis fortis*: are snapshots coupled to a celestial system of reference? *J. Exp. Biol.* **205**, 1971-1978.
- Andel, D. and Wehner, R. (2004). Path integration in desert ants, *Cataglyphis*: how to make a homing ant run away from home. *Proc. Biol. Sci.* **271**, 1485-1489.
- Baddeley, B., Graham, P., Philippides, A. and Husbands, P. (in press). Holistic visual encoding of ant-like routes: navigation without waypoints. *Adapt. Behav.* Doi: 10.1177/1059712310395410.
- Basten, K. and Mallot, H. A. (2010). Simulated visual homing in desert ant natural environments: efficiency of skyline cues. *Biol. Cybern.* **102**, 413-425.
- Cartwright, B. A. and Collett, T. S. (1983). Landmark learning in bees. *J. Comp. Physiol. A*, **151**, 521-543.
- Cartwright, B. A. and Collett, T. S. (1987). Landmark maps for honeybees. *Biol. Cybern.* **57**, 85-93.
- Cheng, K., Narendra, A., Sommer, S. and Wehner, R. (2009). Traveling in clutter: Navigation in the Central Australian desert ant *Melophorus bagoti*. *Behav. Processes* **80**, 261-268.
- Collett, M. (2010). How desert ants use a visual landmark for guidance along a route. *Proc. Natl. Acad. Sci. USA* **107**, 11638-11643.
- Collett, M., Collett, T. S., Bisch, S. and Wehner, R. (1998). Local and global vectors in desert ant navigation. *Nature*, **394**, 269-271.
- Collett, T. S., Dillmann, E., Giger, A. and Wehner, R. (1992). Visual landmarks and route following in desert ants. *J. Comp. Physiol. A*, **170**, 435-442.
- Collett, T. S., Collett, M. and Wehner, R. (2001). The guidance of desert ants by extended landmarks. *J. Exp. Biol.* **204**, 1635-1639.
- Durier, V., Graham, P. and Collett, T. S. (2003). Snapshot memories and landmark guidance in wood ants. *Curr. Biol.* **13**, 1614-1618.
- Fourcassié, V. (1991). Landmark orientation in natural situations in the red wood ant *Formica lugubris* Zett (Hymenoptera; Formicidae). *Ethol. Ecol. Evol.* **3**, 89-99.
- Fukushi, T. (2001). Homing in wood ants, *Formica japonica*: use of the skyline panorama. *J. Exp. Biol.* **204**, 2063-2072.
- Graham, P. and Cheng, K. (2009a). Which portion of the natural panorama is used for view-based navigation in the Australian desert ant? *J. Comp. Physiol. A*, **195**, 681-689.
- Graham, P. and Cheng, K. (2009b). Ants use the panoramic skyline as a visual cue during navigation. *Curr. Biol.* **19**, R935-R937.
- Graham, P. and Collett, T. S. (2002). View-based navigation in insects: how wood ants (*Formica rufa* L.) look at and are guided by extended landmarks. *J. Exp. Biol.* **205**, 2499-2509.
- Graham, P., Fauria, K. and Collett, T. S.

Fo1009las1009las1,ueesert ab8w 18.989795 1m.