The Information Content of Panoramic Images I: The Rotational Errors and the Similarity of Views in Rectangular Experimental Arenas

Wolfgang Stürzl and Allen Cheung
The Australian National University

Ken Cheng
Macquarie University

Jochen Zeil
The Australian National University

Animals relocating a target corner in a rectangular space often make rotational errors searching not only at the target corner but also at the diagonally opposite corner. The authors tested whether view-based navigation can explain rotational errors by recording panoramic snapshots at regularly spaced locations in a rectangular box. The authors calculated the global image difference between the image at each location and the image recorded at a target location in 1 of the corners, thus creating a 2-dimensional map of image differences. The authors found the most pronounced minima of image differences at the target corner and the diagonally opposite corner—conditions favoring rotational errors. The authors confirmed these results in virtual reality simulations and showed that the relative salience of different visual cues determines whether image differences are dominated by geometry or by features. The geometry of space is thus implicitly contained in panoramic images and does not require explicit computation by a dedicated module. A testable prediction is that animals making rotational errors in rectangular spaces are guided by remembered views.

Keywords: view-based homing, spatial orientation, navigation, geometry

Rats, rhesus monkeys, and human children have been trained to find a reward hidden close to one of the corners of a rectangular box. In learning the task, they made systematic errors at the corner diagonally opposite the correct corner (reviews in Cheng, 2005; Cheng & Newcombe, 2005). Cheng (1986), for instance, trained rats to find food located at one constant corner of a rectangular arena. Aside from the shape of the box providing geometric cues, many visual features were arrayed around the walls. One wall was white, whereas the others were black, and panels with distinct patterns stood in the corners. In learning the task, the rats made systematic rotational errors, with the majority of the errors at the diagonally opposite corner. This corner stood in the same geometric relation to the shape of the arena as the target corner, but it contained different visual features. The rotational error has also been found in rhesus monkeys (Gouteux, Thinus-Blanc, & Vauclair, 2001) and in children (Hermer & Spelke, 1994, 1996; Learmonth, Nadel, & Newcombe, 2002; Learmonth, Newcombe, & Huttenlocher, 2001; reviewed by Cheng & Newcombe, 2005). Animals make rotational errors even when the inside of the box contains distinct visual features, such as differently colored walls or visually distinct objects in the four corners. One influential idea arising from the observation of such rotational errors is the concept of a geometric module: Animals encode the geometry of such experimental spaces in a dedicated module separately from the features they contain (Cheng, 1986; Wang & Spelke, 2002, 2003). Others, however, disagree (Newcombe, 2002).

Similar experimental designs, in which the arrangement of objects or features has to be memorized in relation to the environment in which they are placed, also have been used to study the organization of human spatial memory (e.g., Burgess, 2006). Humans also make errors of judgment, including rotational errors of the kind animals make, after disorientation or when environmental geometry and feature arrangements are put into conflict. These alignment, spatial updating, and reorientation effects, together with pointing errors, indicate that both egocentric (view-based) and allocentric representations are at work in parallel, at least in human spatial memory (Burgess, 2006).

Very little attention has been paid to the question of what visual cues animals, including humans, actually have available in these tasks and how salient they are. Here, we approach the question of why rotational errors occur from a purely sensory perspective by investigating the principal visual cues available for navigating in rectangular arenas, similar to those used by Cheng (1986). We used both tools from view-based robotic navigation and from virtual reality computer modeling to ask how well a location in a rectangular box is defined by the view taken from it (see Stürzl & Zeil, 2007; Zeil, Hofmann, & Chahl, 2003). In what follows, we were guided by two hypotheses: (a) Rotational errors arise from inherent similarities of views in rectangular spaces, and (b) the visual salience of upper and lower edges of the experimental environment determine whether view similarities are dominated by the rectangular space itself or by its internal visual features. In short, we tested whether geometric information can be acquired without explicit computations of surface layout and a fortiori without a dedicated module. In a companion article, we employed virtual reality simulations and simple performance models to test the validity of these hypotheses in nonrectangular test spaces (Cheung, Stürzl, Zeil, & Cheng, 2007).

Experiment 1: Image Differences in a Rectangular Arena

In this experiment, we used a robotic arm to move a camera for acquiring panoramic images to defined positions in a rectangular arena like that used in Cheng's (1986) study. The image at a target corner served as a reference. Images obtained at any particular location in the arena were compared with the reference image to map image differences across the experimental space.

Method

Data Acquisition

Panoramic images were recorded inside a rectangular box that had similar dimensions to that used in the original experiment by Cheng (1986). The box was 120 cm long × 60 cm wide and 30 cm high and was made of white, slightly shiny laminated boards, which were covered inside with black cardboard if needed (see Figure 1A). The floor of the box was covered with light brown wrapping paper. To minimize visual cues from the outside of the box, a rectangular white cloth screen (180 cm × 90 cm) was suspended from the ceiling in such a way that it floated about 5 cm above the upper edge of the box. The box was positioned on a table symmetrically below a rectangular ceiling light.
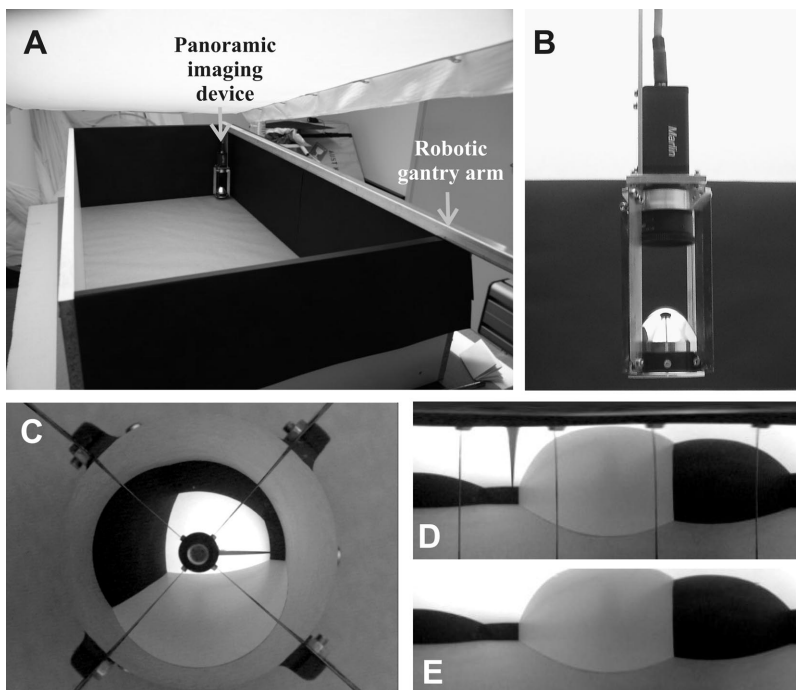
Figure 1. Experimental arena and recording technique. (A) The experimental box was constructed from laminated wooden boards, centered under a rectangular ceiling light and lined with black and white cardboard. In the case shown, the box had one white and three black walls. Panoramic images were recorded at regularly spaced locations inside the box by moving a panoramic imaging device with the aid of a robotic gantry. During recording, a white screen was lowered from above to a height of about 5 cm above the upper edge of the box to exclude any external features. (B) A close-up view of the panoramic imaging device, which consisted of a digital video camera looking down onto a parabolic mirror surface. (C) The raw panoramic image as seen by the camera. (D) The same view after the image has been transformed to a cylindrical array of pixels. (E) The panoramic image after pixels viewing the gantry and recording system have been "cleaned up." For further details, see the *Method* section.

We recorded panoramic images with a color firewire CCD-camera (Marlin MF-046C, Allied Vision Technologies, image size 640 × 480 pixels) viewing a convex mirror (see Figure 1B) with constant gain in elevation (Chahl & Srinivasan, 1997) at regular positions in a 23 × 11 grid with 5-cm spacing. The viewpoint of this panoramic imaging system was about 5 cm off the floor of the box. The imaging device was moved inside the arena (reposition accuracy < 0.1 mm) with the aid of a computer-controlled robotic gantry (for details, see Zeil et al., 2003). The vertical field of view of the imaging system extended from −52.5° (below the horizontal) to +62.5° (above the horizontal). Color images were converted to 8-bit gray values (Figure 1C) and unwarped to 1° resolution, resulting in rectangular panoramic images of 360 pixel azimuth and 115 pixel elevation (see Figure 1D). The images thus represent the highest spatial frequencies that can be resolved by a rat's eye (e.g., Keller, Strasburger, Cerutti, & Sabel, 2000). Unwarping parameters were chosen in such a way that for an image recorded at the center of the box the unwarped image and its 180° rotated version had maximum similarity. The image regions viewing the four thin vertical blades connecting the mirror to the camera and those viewing the gantry arm were filled and padded with gray values obtained by row-wise linear interpolation of neighboring pixel values. In other words, views of the recording system itself were removed from the images (Figure 1E).

We recorded three sets of images (23 × 11 = 253 images each) with different camera settings for three internal box conditions: (a) box with four black walls, (b) box with three black walls and one long white wall, and (c) box with three black walls and one long wall with a pattern of vertical black and white stripes (stripe width 2.5 cm). The camera settings and their consequences were as follows:

1. Low camera gain (gain 1).[1] This results in comparatively dark images with only slight differences between black and white walls and between black walls and the ground. The ceiling light outside the box remains visible in the images (except for images recorded close to the center of the box).

2. High camera gain (gain 2.25). This setting achieves higher sensitivity to low light intensities, but also leads to saturation at high intensities. The difference between black and white walls is increased, slight differences in brightness of the walls and the floor become visible, together with weak reflections on the white wall. The part of the image viewing the ceiling saturates, and the ceiling light is no longer visible. The brightness difference between the ceiling and the upper edge of the box is increased. Compared with gain 1, this camera setting basically tests for effects of quantization and the visibility of the ceiling light.

3. Gamma correction switched on while camera gain is set to 1. This setting maps light intensity to pixel values in a nonlinear way. The 10-bit intensity values of the camera are mapped to 8-bit pixel values of the image according to $\mathbf{I}[8\text{bit}] = 255\ (\mathbf{I}[10\text{bit}]/1023)^{0.45}$ (Allied Vision Technologies, 2004). This compresses the higher dynamic range and results in higher resolution of low light intensities and lower resolution for high light intensities without causing saturation. Gamma-corrected images thus have strong differences between white and black walls, and the ceiling light is still clearly visible through the cloth screen. We included this camera setting because biological photoreceptors often have nonlinear intensity scaling.

Image Preprocessing

In addition to using three different camera settings described above, we also tested the effect of manipulating the contrast between the black and white walls by applying a lower threshold to the 8-bit gray values of images recorded with camera gain 1. All pixel values less than 32 (out of 256 values) were set to 32. This operation, which we refer to as *threshold 32*, simulates the opposite to the high intensity saturation provided by gain 2.25.

We also applied a simple edge detector to the unwarped panoramic images, which can be thought of as mimicking edge enhancement by lateral inhibition in the visual system:

1. Images were filtered with a linear Difference of Gaussians (DoG) filter (size 9 × 9 pixels, standard deviations of the two Gaussians were $\sigma_1 = 1$ pixel and $\sigma_2 = 2$ pixels). This operation removes low spatial frequencies, removes all homogeneous surfaces from the image, and causes image difference functions (IDFs) to become flat in the center of the box.

2. We computed a binarized image by applying a threshold: All pixel values below threshold were set to 0; values greater than threshold were set to 1. The threshold was chosen so that the upper edge of the box would be emphasized in all images. This operation ensures that only strong responses by the linear DoG filter are regarded as edges. Without thresholding, rotational errors were less pronounced, but IDFs still exhibited clear minima.

3. We blurred the images with a Gaussian filter with standard deviation $\sigma = \sqrt{3}$ pixels. This additional operation enhances the "smoothness" of the IDF and thereby reduces the number of local minima that can interfere with successful homing. It also compensates for small errors that are introduced by small remaining misalignments of the camera system with respect to the box reference frame.

We investigated what effect other filter parameters than the ones we used would have on the shape of IDFs for edge images; we found that results did not appear to be very sensitive to variations in filter parameters. Edge detection using a "standard" edge detector (Canny, 1986) usually gave similar results.

Determining Difference Functions

The image difference between two images $\mathbf{I}^A$ and $\mathbf{I}^B$, each consisting of N pixels, was calculated as the mean square pixel difference (MSPD) over all corresponding pairs of pixels, that is,

$$\text{MSDP}(\mathbf{I}^A, \mathbf{I}^B) = \frac{1}{N}\sum_{i=1}^{N}(I_i^A - I_i^B)^2$$

Because we did not make any assumptions about an animal's orientation in the box, we computed the MSPD for all possible relative image orientations (in steps of 1°) and then used the minimum MSPD as the image difference at that particular location. The IDF is the two-dimensional distribution of image differences between a reference image and the other images of the recording grid inside the box; that is, the IDF is a function of location (for

---

[1] A video camera with a CCD sensor chip maps light intensity (irradiance) linearly to sensor voltage. Before analog to digital conversion, the voltage is amplified according to the selected gain.

further details, see Stürzl & Zeil, 2007; Zeil et al., 2003). In view-based matching, a reference image is taken at the target location, and the animal or robot uses differences between the reference image and the images at any particular location (i.e., the IDF) to determine the next step to take (see Vardy & Möller, 2005).

Homing Algorithm and the Catchment Areas of Panoramic Snapshots

We wanted to analyze whether an animal, in principle, could use the distribution of image differences inside a rectangular box to relocate a target position (the location at which a reference image was stored). To do this, we estimated the local slope of the IDF by comparing the difference of all images in a certain neighborhood with the reference image and by computing the direction of movement for each position on the sampling grid. We chose the neighborhood to be either the eight next neighbors or the 24 nearest grid positions, including next-but-one neighbors. The local direction of movement (depicted as vectors in the figures) was then made toward the position within the local neighborhood with the greatest drop in image difference. A larger neighborhood reduces the chance of "getting stuck" in a local minimum of the IDF. We denote those regions inside the box as *catchment areas* that, on the basis of the locally estimated movement directions, lead to a single minimum of the IDF.

Results and Discussion

The Shape of Image Difference Functions in a Rectangular Box

Box with four black walls with and without corner features. We explain our basic results and our conventions in Figure 2. Panoramic views of the box (camera gain 1) from the training or target corner are shown in Figure 2A. The pictures are surrounded by a rectangle to represent the internal condition of the box; in this case, the box carried four black walls and had either no corner features (left column) or included corner features (right column). The IDF (Figure 2B) for the reference image taken in the top right corner of the box (d in Figure 2C) has a clear maximum in the center of the box and four minima in the corners for both box conditions. The IDF is elongated in the *x*-direction (primary axis) of the box, as can be seen best in its two-dimensional representation shown in Figure 2C, together with our labels for the four corners (a–d), the location of minima (crossed lines), and the transects through the minima (continuous and dashed lines). Figure 2D displays the IDF across the transects at the locations of minima with five IDF values shown at 5-cm intervals, except for values that exceed the *y*-axis scale. The results of two local homing algorithms are shown in Figures 2E and F. They differ in the extent of the local area over which image differences are compared to establish the direction of the slope of the IDF (next neighbor homing in Figure 2E and next-but-one neighbor homing in Figure 2F). In Figures 2E and F, the local direction of decreasing image differences are indicated by arrows. The differently shaded areas denote the catchment area of a given minimum, that is, the range over which an agent following the local directions would reach this particular minimum.

In this basic configuration in which the box either does not contain any internal features, except for the upper and lower edges, or does not contain apparently very salient visual features marking each corner, a number of observations are noteworthy: The image differences relative to the reference image vary smoothly, with maximal values forming an elongated "hill" in the center of the box and decreasing toward the four corners. Minimal values are reached both in the target corner (d) and in the corner diagonally opposite (b), corresponding to the corners visited by animals when they exhibit a rotational error. For the box without internal features (Figure 2, left column), this is not surprising as the two diagonally opposite corners are visually indistinguishable. However, even apparently distinct corner features do not change this basic property of panoramic image differences (Figure 2, right column). In both box conditions, an agent placed in the center of the box and moving to minimize image differences would clearly end up in one of the two corners. If released repeatedly in slightly varying positions, or if allowed to search around the box, such an agent would show a rotational error: It would end up most often in the target corner (d) and the one diagonally opposite (b), as these have larger catchment areas compared with the other two geometrically incorrect corners (Figures 2D, E, and F), irrespective of whether all corners are marked by distinct visual features.

Boxes with three black walls. One of the most surprising and consistent findings in animal studies is that animals continue to make rotational errors, even if the box carries more dominant features than individual corner marks (e.g., by having distinctly colored walls). We set out to test to what extent the IDF in a rectangular box changes when these features are present. We therefore collected two further sets of images in a box with three black walls and one white wall and a box with three black walls and one wall carrying a pattern of 2.5-cm wide black and white vertical stripes. The IDFs, their minima, and the catchment areas for these two situations are shown in Figure 3. The first result to note is that the conditions for a rotational error are still met in both cases, but that in the presence of a white wall, the difference function has two pronounced and one shallow minima. However, it is still the case that the two deepest minima are at the target location (d) and the corner diagonally opposite (b). Therefore, even in a situation when the inside of the box contains large and conspicuous features (i.e., cues to its orientation and to the identity of corners), navigation by global image differences still leads to rotational errors. Most surprising at first sight, a highly structured and seemingly more salient pattern, such as the vertical stripes in the second case, actually generates more distinct rotational error conditions than a homogeneous white wall. Although one would have thought that more internal visual structure would reduce rotational errors, this counterintuitive result is due to the fact that global pixel-by-pixel image differences actually become smaller with the addition of the black pixels in the striped pattern.

In a box with three black walls, the difference function may depend on which corner is chosen as the target. As can be seen by comparing the difference functions in Figure 3 with those shown in Figure 4, where the target is in corner b, this is indeed true for the box with a white wall but not for the box with a wall with a stripe pattern (see Figure 4).

The dependence on target location in the box with one white and three black walls prompted us to look again at individual data in Cheng's (1986) reference memory experiment (Experiment 2).
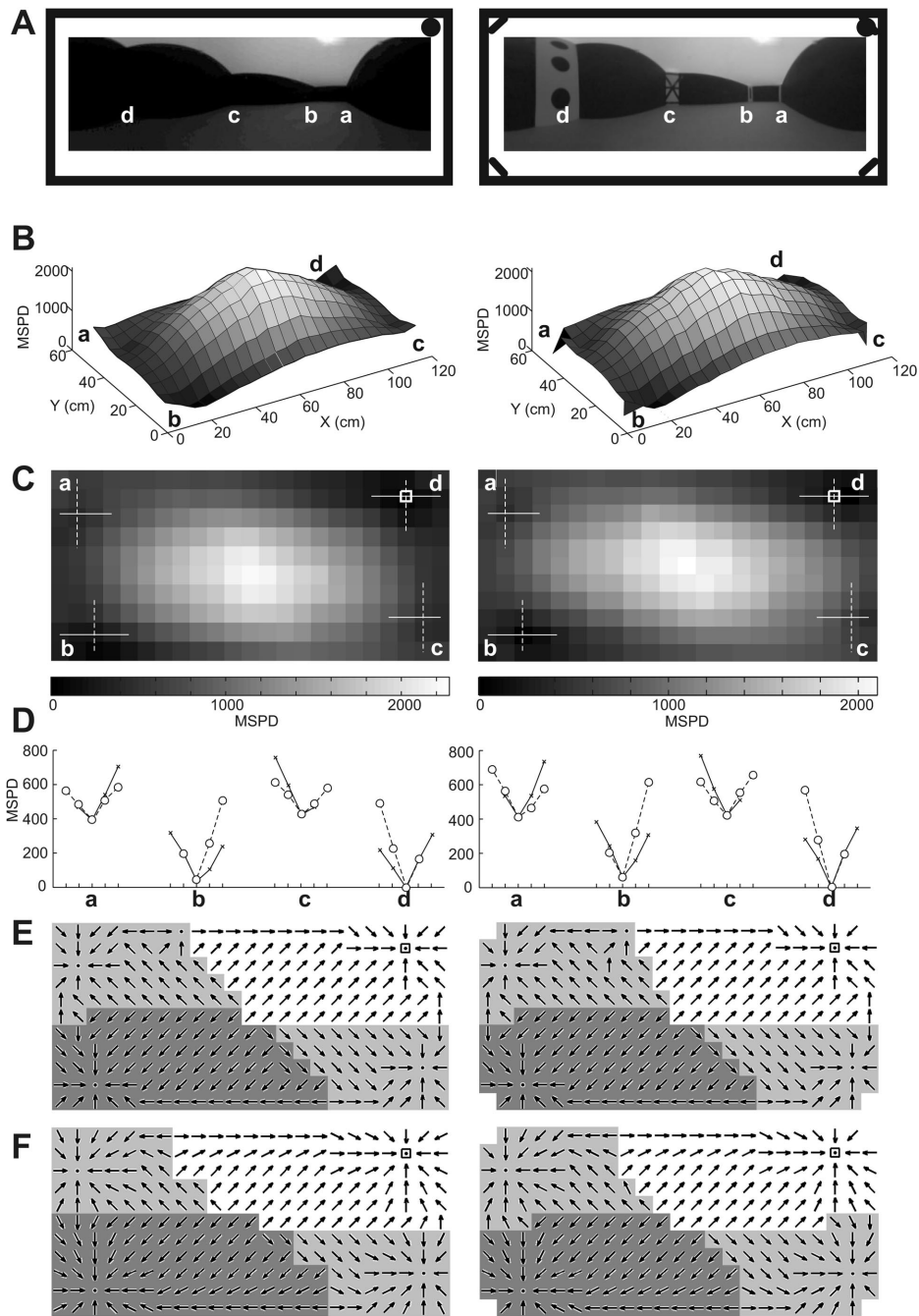
Figure 2. The image difference function (IDF) in a box with four black walls without (on the left) and with distinct corner features (on the right). (A) Panoramic views of the box as seen from the target location. The target location is marked with a black circle and the condition of the box walls is shown in the color of the frame. (B and C) *The difference function:* Global image differences (mean squared pixel differences [MSPDs]) between the reference image and images recorded at regular grid locations inside the box are shown as a three-dimensional surface (B) and as a two-dimensional gray-level map (C). In C, the locations of minima are labeled by crossed lines and letters a–d. The target location d is marked by a small square. (D) Difference functions over the horizontal (crosses and solid lines) and vertical (circles and dashed lines) transects through the neighborhood of minima in the four locations a–d, as indicated by continuous and dashed lines in C. For each transect, five IDF values are shown at 5-cm intervals, except for values that exceed the *y*-axis scale. Note that the difference function minimum is most pronounced in the target location d and in the diagonally opposite location b in both box conditions. (E and F) *The catchment areas of the reference image:* Local arrows point in the direction of the largest local decrease of image differences, between the reference image at d and the images taken at different grid positions. Arrows thus point along the slope of the IDF. The differently shaded areas mark those domains in the box from which the gradient leads to one minimum. (E) Local image differences are determined only for the 8 neighboring locations. (F) Local image differences are determined for 24 neighboring locations, including next-but-one neighbors. Camera gain setting was 1. For further details, see the *Method* section.
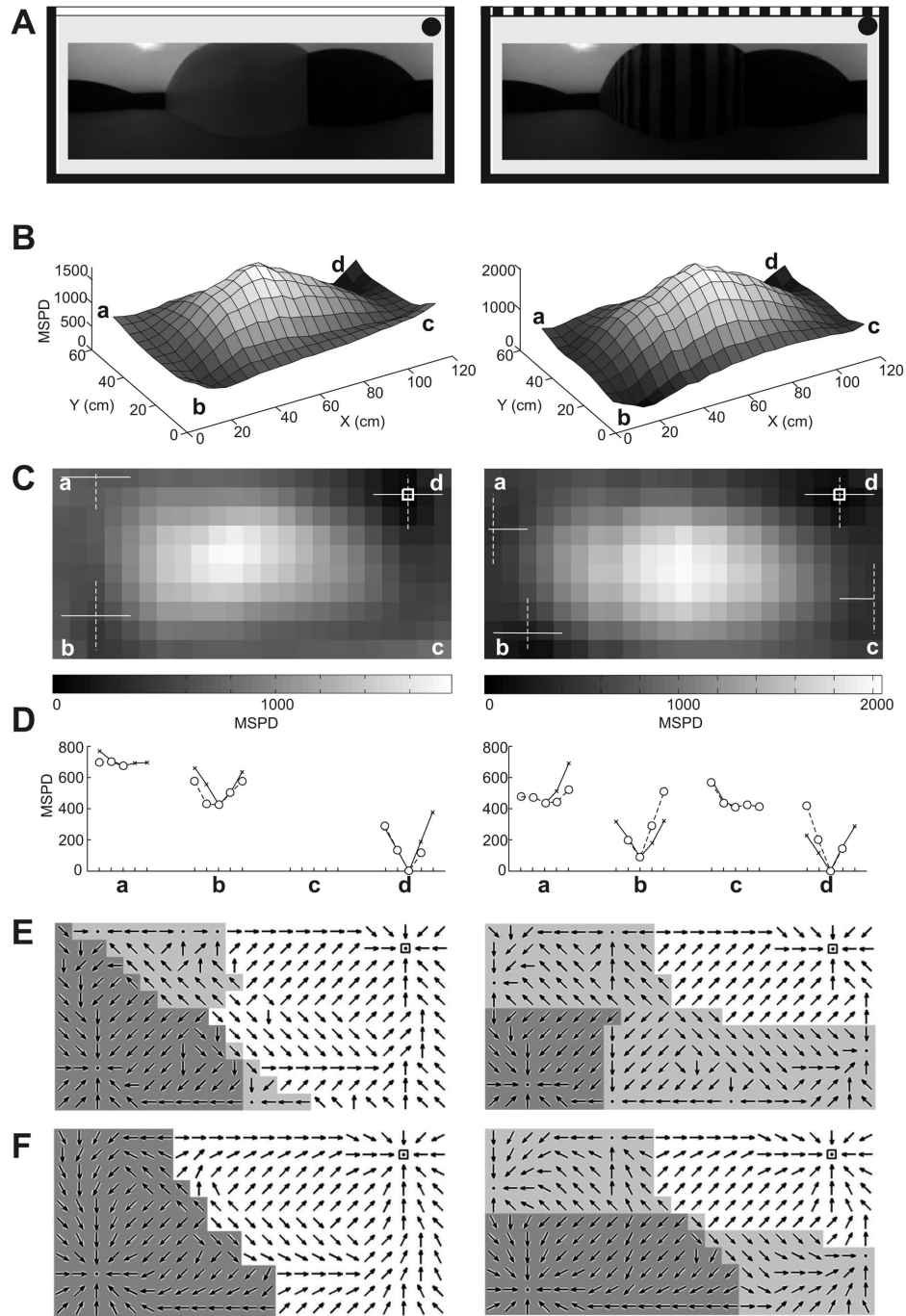
Figure 3. Difference functions for a box with different internal cues. (Left) Box with one white and three black walls. (Right) Box with one wall carrying black and white vertical stripes and three black walls. Conventions and camera settings as in Figure 2. Note that in both cases, the difference function has its most pronounced minima at the target location d and the location diagonally opposite b.

Although individual data were not published, they are available. When rats were tested in the box with three black walls and one white wall, target locations were counterbalanced across rats, so that two rats had targets near a white wall and two had targets at a corner surrounded by two black walls. We tallied rotational errors as the proportion of all errors. From Figures 3 and 4 (left),

the errors should be more evenly distributed for rats with all black corners, as the nontarget minima are more similar. For the two rats trained to a corner with a white wall, proportions of rotational errors were 1.00 and 0.94. For the two rats trained to an all-black corner (excepting the panel at the corner), the proportions of rotational errors were 0.39 and 0.66. These differences, even with
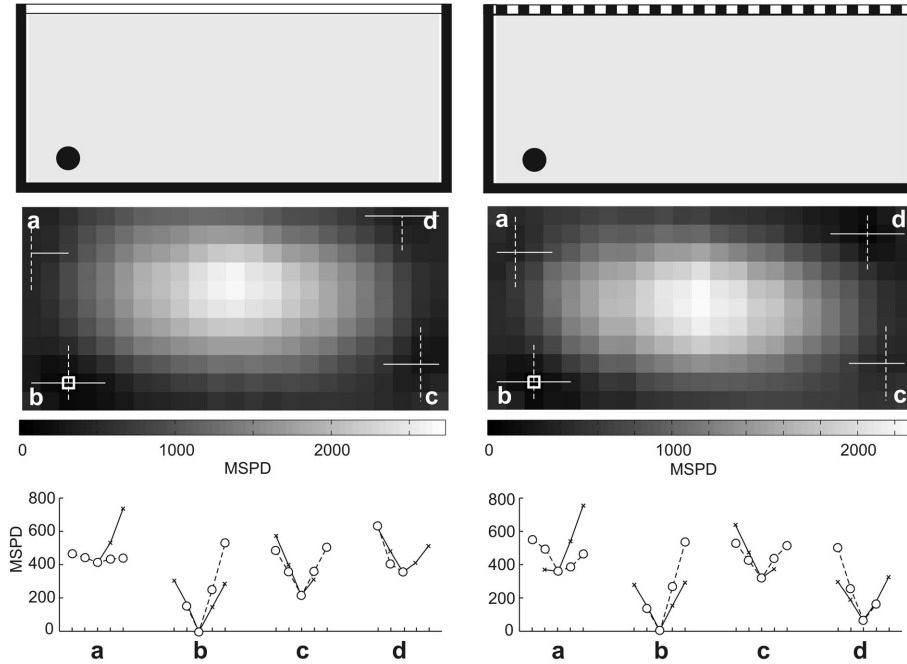
Figure 4. How difference functions depend on the target location. Compared with the situation in Figure 3, the target location is now at location b. Conventions and camera settings as in Figure 2. Note that for the box with one white wall (left), the two most pronounced minima are now located at the target location and location c, but not in the diagonally opposite corner. This is not so for the box with a striped wall.

the tiny sample, are actually statistically significant ($p < .05$, one-tailed). Thus, view-based navigation might well play a role in explaining this pattern. However, this dependence on the location of the target corner also is influenced heavily by the intensity scaling and preprocessing of the image, as we will show next.

The Effects of Nonlinear Intensity Scaling

Our results so far indicate that the geometry of a rectangular box is not necessarily a distinct set of cues that is separable from other features and requires special computations to extract. Rather, geometry is contained in the views animals have of the box as demonstrated by the fact that view-based image matching can result in rotational errors. Our second hypothesis was that the relative salience of the visual cues determines whether the features or the shape of the box dominate the image transformations experienced by an animal searching within such an experimental space. We investigated this conjecture by testing the influence of simple image preprocessing strategies on the shape of the IDFs. This also allowed us to check how sensitive our results are to the distribution of pixel values in panoramic images.

We investigated the effects of different ways of coding intensity on the IDFs in two of the cases considered earlier: for the box with three black walls and one white wall, with the target location in two different corners (top and bottom panel, Figure 5). For each of these conditions we compared four different settings of gain, threshold, and gamma transformation (as indicated in the first column of Figure 5; see *Method* section for details). The results are expressed as before as a gray-level coded map of image differences (Figure 5, second column), as transects through IDFs across

the corner minima (Figure 5, third column), and as a map of local slopes (Figure 5, fourth column).

The results show that image contrast matters in the box with one white wall for both target locations: For location d (top right box corner, flanked by a white and a black wall), the rotational error condition persists across different ways of coding intensity, whereas for location b (bottom left corner in the box, flanked by two black walls), only the threshold 32 operation produces rotational error conditions both in terms of depth of minimum and size of catchment area. Note that this operation minimizes any shading variations on the black walls and effectively decreases the difference in pixel values between the black and white walls. As a consequence, the visual salience of the internal cues is reduced.

We conclude from these results that, in certain conditions, the distribution of image contrast does not significantly change the shape of IDFs, but that in other box conditions, slight variations in nonlinear intensity scaling, especially of the gray levels of black walls, do lead to qualitative differences (e.g., in the case of target location b in a box with one white and three black walls).

The Effects of Image Preprocessing

Next, we asked how well corner locations are defined when edges provide the only salient cues in the box. We reasoned that this would also allow us to disentangle the effects of geometric and featural cues. We find indeed that the more internal contours are emphasized, the weaker rotational error conditions become because the views in different corners become unique. In turn, rotational error conditions are more pronounced when the edges of the box are more salient.
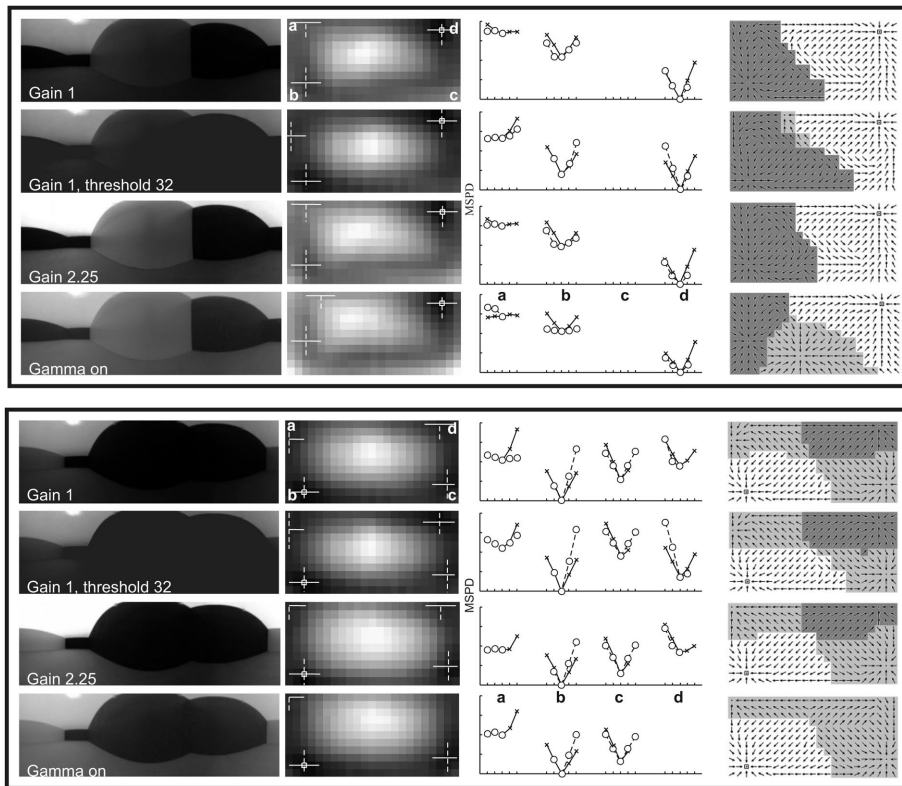
Figure 5. How the relative depth of difference function minima depends on the nonlinear intensity scaling in a box with one white wall. The target location is either at d (top four rows) or b (bottom four rows). Conventions as in Figure 2, with the catchment areas (right column) determined over 24 nearest neighbor locations. Settings vary along rows in both cases (for details, see the *Method* section), as indicated in the first column. Note that gain settings have little influence on difference functions in the case of target location d, but do affect the results qualitatively in the case of target location b.

The images in the second row of Figure 6 show the reference images with gain 1 after edge detection for the four standard box conditions (shown schematically in the first row, Figure 6). Before we applied our edge-detection algorithm, we preprocessed the images by applying gain 1 (top row of images), which emphasizes box contours; gain 2.25, which increases the contribution of small intensity values; or a gamma transformation, which does so in a nonlinear way (images in the bottom row). The transects through the IDF across the corner minima in Figures 6A to C show that gain 1 images produce the most pronounced conditions for rotational errors (clear minima at diagonally opposite locations, Figure 6A), whereas these conditions become weaker or disappear with gamma-processed images in which internal features contribute to the image (Figure 6C).

Experiment 2: Virtual Reality Simulations

There are limits to how far we can test our hypotheses in "real-world" experiments. This is especially true when we ask how different visual cues in experimental boxes contribute to image differences and how these cues may be used to drive behavior. Manipulations addressing these questions are difficult to implement in real arenas (e.g., swapping the brightness levels of the white wall and the ceiling). In virtual reality, such manipulations are straightforward. We therefore reconstructed a rectangular box in a virtual reality environment to test whether it would allow us to answer these questions. We designed these virtual reality experiments in two steps: We first asked whether IDFs have the same properties in the real world and in virtual reality, and then went on to test our second hypothesis, namely that the visual appearance of the edges of the box is primarily responsible for rotational errors.

Method

All virtual reality simulations and subsequent data processing were carried out using a Moebius PC with an Intel Pentium 4 CPU (3.2GHz, 1GB RAM) and an nVidia GeForce 6600 graphics card. Virtual environments were generated using software written in OpenGL (open source graphics library), developed in a Visual C++ programming environment (Microsoft Visual Studio .NET 2003). Image transformations and subsequent data processing were carried out using software written in MATLAB R13 (Mathworks, Natick, MA).

Construction of Virtual Reality Environments

The virtual reality images of experimental boxes were constructed as follows: Texture bitmaps were created in MATLAB for all surfaces (including ground and sky), using desired gray pixel
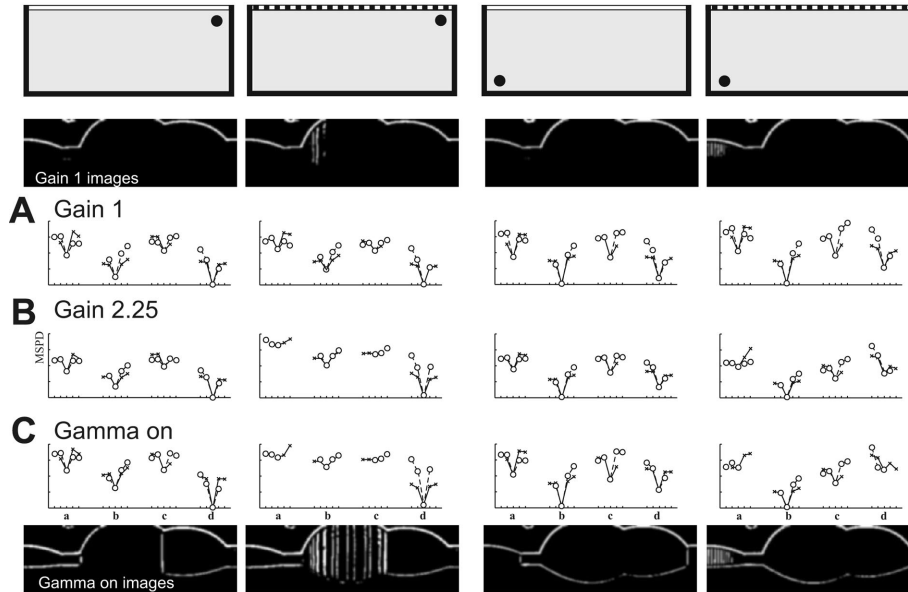
Figure 6. Difference functions after edge detection and their dependence on gain. Box conditions are shown in top row pictograms. Images have been processed (for details, see the *Method* section) to contain mainly the edges of the scene (top row are gain 1 images, bottom row are gamma-transformed images). A–C show difference functions over transects through the minima for three different image preprocessing regimes. Note that with the top edge of the box dominating the scene (A), the most pronounced minima of the difference function lie at the target location and at the diagonally opposite corner. The top edge thus produces conditions favoring rotational errors. Gamma transformation (C) highlights in addition the internal edges of the box and the grating pattern to the extent that conditions leading to rotational errors become weaker (columns 1 and 3) or even disappear (columns 2 and 4). For further details, see text.

values (0–255). The textures were single homogeneous gray values for each entire rectangular surface. In-house software (Allen Cheung, The Australian National University) was developed in OpenGL (in a Visual C++ environment), which allowed relevant shapes (e.g., triangles, rectangles) to be drawn at specified three-dimensional locations and orientations and the predefined textures to be pasted onto these shapes. No light source was added so that there was no possibility of distortion of the pixel values.

To capture a fully panoramic view of this environment, we used six perpendicular viewports, each spanning 100° with 200 × 200 pixel resolution in a manner similar to Neumann (2002). The result may be described conceptually as a "viewing cube." The OpenGL software was compiled into a MATLAB executable file (MEX file) and took position and orientation values as input, returning the image as its output.

Further software was developed in MATLAB that allowed the viewing cube image to be unwarped into spherical coordinates to mimic as closely as possible the experiments in the real box. Unwarping involved a single mapping between viewing cube pixel values and the final image (in spherical coordinates). Pixel values at noninteger positions were bilinearly interpolated from the nearest four pixels (forming a square). The final panoramic image extended 180° in elevation and 360° in azimuth, with 1° resolution. The viewpoint of the imaging apparatus described previously was estimated to be approximately 5 cm from the ground surface. This was the viewing height adopted for the virtual reality simulations. The reference and sampling positions were identical to that of

experiments in the real box. All further processing of virtual images were identical to that of real images unless otherwise specified.

Procedures for Simulations

We first simulated the standard rectangular box with one white wall and three black walls. The procedures for generating IDFs and catchment areas followed those of Experiment 1, except that pixel values were derived from virtual reality rather than a gantry-controlled camera system. We then used virtual reality to radically alter the salience of external and internal cues to test our hypothesis that this may have a crucial effect on the IDF in rectangular spaces. A standard virtual reality environment was compared with a manipulated environment that increased the salience of internal walls and box shape-related cues. This was done by interchanging the brightness levels of the ceiling and the white wall. This operation enhances the internal (featural) cue of the white wall and reduces the salience of the external (geometric) cue of the upper edge of the box. We thus predicted that it would reduce the proportion of rotational errors.

Results and Discussion

Comparison of Virtual and Real Boxes

Figure 7 shows panoramic views of the standard box with one white and three black walls from the center of the box (Figures 7A
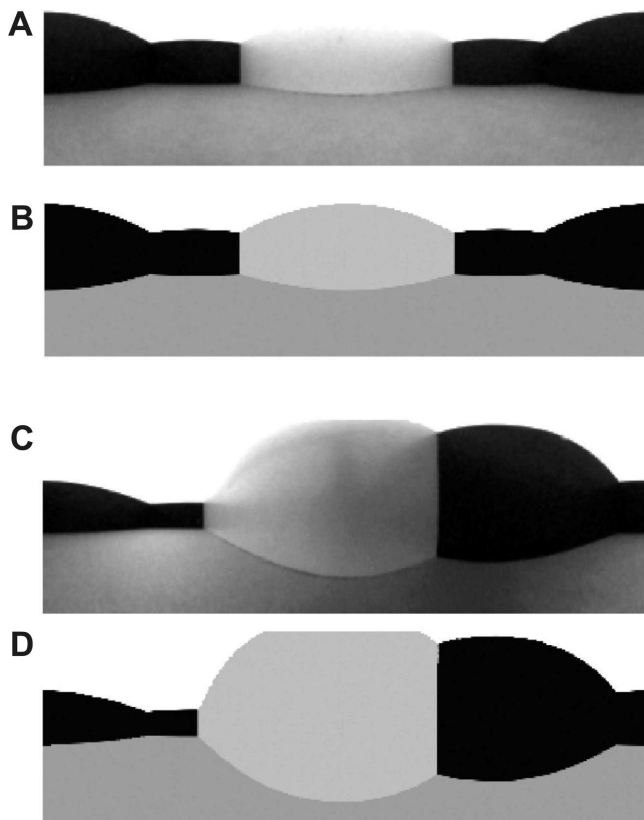
Figure 7. Comparison of "real-world" (gain 2.25) and virtual (computer-generated) views of a rectangular box. (A) View from the center of the real box. (B) The corresponding view in the virtual box. (C) View from the target location in the real box. (D) The corresponding view in the virtual box. Note the effects of illumination and wall reflections in the real images.

and B) and from the target corner (Figures 7C and D). Figures 7A and C are images taken inside the real box, and Figures 7B and D are the equivalent views in the virtual box. Comparing these images demonstrates clearly that the virtual environment allowed us to remove spurious cues inside the box because we could make surfaces completely homogeneous.

To explore how the properties of our virtual environment compare with those of the real box, we determined the IDF in the virtual box by using the reference image, taken with gain 1 setting, from the real box (Figure 8, left column) and comparing it with the IDF using the reference image from the virtual box (Figure 8, right column). Two results are noteworthy first: Both the overall shape of the IDFs in these two cases (compare left and right panels of Figures 8B and C) and the relative differences of the minimum MSPD values of corner minima (compare left and right panels of Figures 8C and D) are similar. However, there are also differences: The gradients of image differences in the vicinity of minima are steeper in the IDF determined with the virtual compared with the IDF determined with the real reference image (Figure 8D), which is a consequence of the fact that brightness transitions at the edges are harder in virtual compared with real images. Furthermore—and surprisingly—the IDF for the virtual reference image has two additional shallow local minima at locations marked by circles in

Figure 8E. These properties also are reflected in the catchment areas for the two situations (Figure 8E). These local minima do not occur with the real reference image (and are also absent in the real box; see Figure 3, left column), possibly because slight variations of brightness on the walls (see Figure 7) make real box views more distinct compared with the homogeneous surface brightness in the virtual box, which can lead to spurious local image similarities. We concluded that the conditions in real and virtual experimental spaces are sufficiently similar for virtual reality experiments to be a useful tool for the investigation of view-based homing in such simple visual environments.

Altering the Salience of Internal and External Cues

Figure 9 compares the standard box with one white and three black walls (left column) and the same virtual box in which the brightness of the "sky" and the brightness of the white wall have been interchanged (right column). As predicted, the data in Figure 9 show that increasing the salience of internal cues reduced rotational errors. In the box with the brightness values of the sky and white wall interchanged, the IDF has only one minimum, which is located in the target corner. The catchment area of the reference image practically spans the whole box (Figures 9E and F), with the two remaining shallow minima located along the same white wall. Most significant, there is no detectable minimum at the diagonally opposite corner. In short, salience of external, box-related cues produces conditions favoring rotational errors, whereas salience of internal cues eliminates rotational errors.

General Discussion

We have shown in robotic experiments and in virtual reality simulations that panoramic snapshots contain information about the geometry of rectangular experimental arenas as they are used in many studies of spatial memory in animals. Our results suggest that an animal that is sensitive to image transformations in such rectangular spaces, and that moves to minimize image differences when relocating a goal, would exhibit rotational errors on the basis of purely pixel-by-pixel global image matching. The assumption is that animals memorize the panoramic view at the target location and use it as a reference when moving around the rectangular space. Relocating the target location would then be achieved by moving in such a way as to minimize image differences (Cartwright & Collett, 1983; Franz, Schölkopf, Mallot, & Bülthoff, 1998; Vardy & Möller, 2005; Zeil et al., 2003). We have shown that no feature extraction is needed for this process to work. The processing of the images is simple and does not go beyond edge detection, a characteristic of early visual processing in many animals. In particular, there is no need to explicitly compute or extract geometric properties such as shape parameters or featural cues. Results varied, but under many conditions, especially those that de-emphasized internal features, the conditions for rotational errors were prevalent. The depth of the local minimum at the location of rotational errors, and hence the extent to which the rotational ambiguity may affect searching animals, depends on the visual salience of the top and bottom edges of the box relative to internal visual cues arrayed on the walls. Our analysis thus suggests that panoramic image differences may explain why some animals make rotational errors in
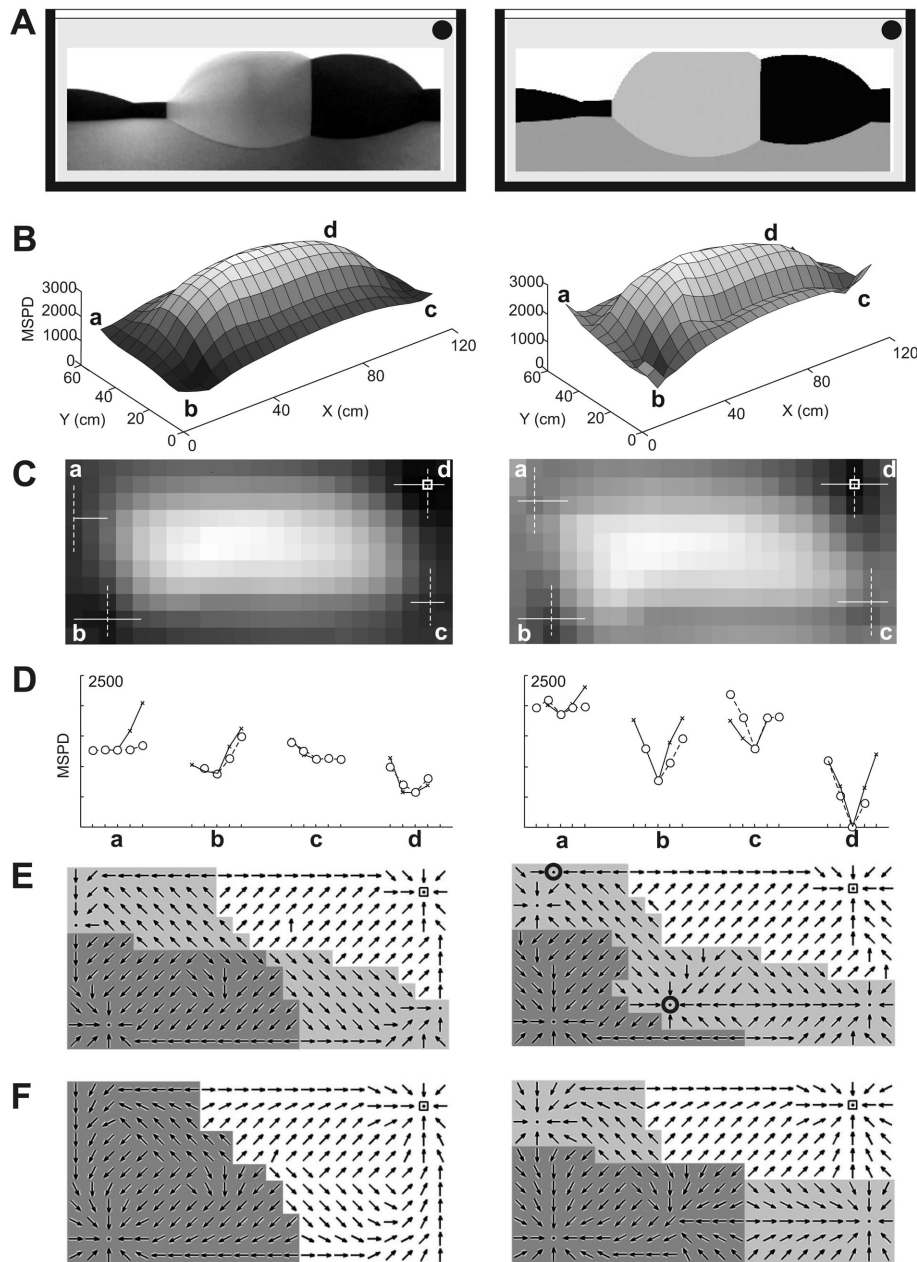
Figure 8. Difference functions in a virtual reality box, with the reference image (A) taken in the real box (left column, gain 1) or in the virtual box (right column). The left image in Figure 8A is the same image as the one in Figure 3A, but it appears brighter because of different gray value mapping to make comparison easier. Conventions as in Figure 2. Note the presence of additional local minima in the fully virtual case. Virtual images were clipped to 115° elevation range to match real image size. For details, see text.

these rectangular boxes when attempting to relocate a hidden reward close to one of the corners. Therefore, it is not necessary to invoke a distinct geometric module for computing and representing geometric properties to account for the rotational errors of animals in these simple environments.

Although the empirical literature on the use of geometry in navigation has burgeoned, the question of how information on the geometry of a space is extracted and encoded has been little

addressed. Gallistel (1990) hypothesized that animals in experimental spaces use the principal axes of space for orientation by extracting a global shape parameter, whereas others (Pearce, Good, Jones, & McGregor, 2004; Tommasi & Polli, 2004) recently have suggested that local geometric cues such as corner angles deliver the geometric cues. So far, experimental results are ambiguous on this issue, and Cheng (2005) therefore proposed that geometric and featural cues are encoded together, but that some computational
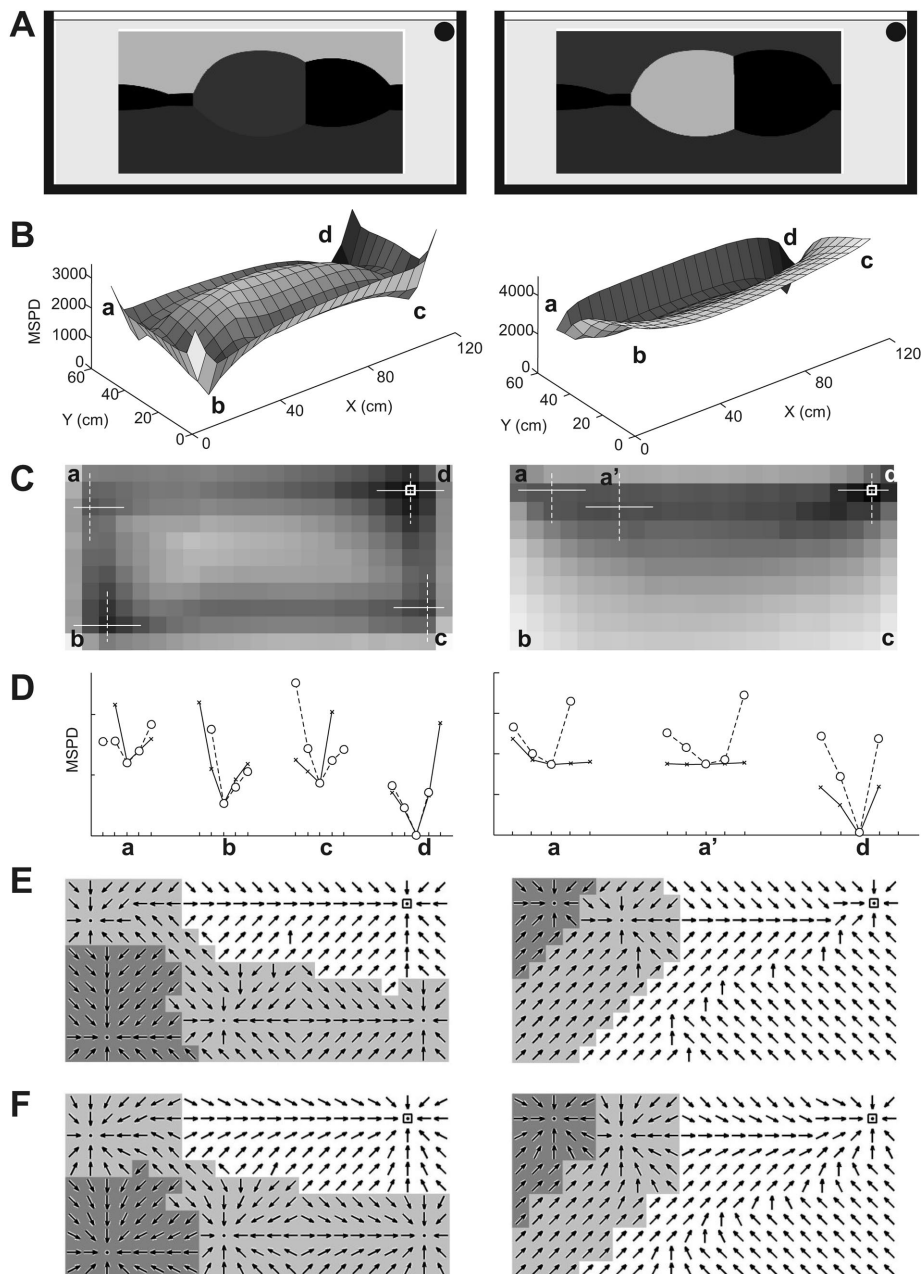
Figure 9. A test for the role of salience, by emphasizing or de-emphasizing the top edge of the box. Conventions as in Figure 2. (Left) Difference function with the top edge of the box seen against a bright "sky" (see A). (Right) Difference function after pixel values between "sky" and the white inner wall of the box have been swapped. Note that conditions favoring rotational errors prevail when the top edge of the box is most salient (left), whereas a unique minimum develops at the target location when internal box features are most salient (right). The gray levels in both images were scaled linearly, and they maintained the full elevation range of 180° possible in the virtual environment.

process based on shape parameters produces rotational errors. Our present, entirely view-based interpretation is simpler and more parsimonious than current theories that explicitly extract local features or global geometric properties. It is more parsimonious because no image segmentation is needed, and the early visual processing used in our simulations are known for most visual systems, including those of all the animals studied in the geometry literature.

Our analysis brings the novel insight that geometry and features are contained in the same, early visual representation, namely panoramic snapshot views, and that how animals search in rectangular boxes may depend on the relative salience of those cues.

The explicit extraction of any kind of information on environmental shape is not necessary. Instead, the IDFs in rectangular or elongated boxes are oriented reflecting the geometry of the space in which they are determined (see also Collett & Zeil, 1997; Zeil et al., 2003). Thus, complicated schemes of geometric representation are not needed when the geometry of space is implicitly represented in the orderly transformations of images as an animal moves through that space (e.g., Gibson, 1950).

In our modeling, we have used panoramic views. Some animals may not have panoramic views, and this may have implications for the relative salience of geometric and featural cues. Sovrano and Vallortigara (2006), working with chicks, argued that restricted visual fields lead to relatively more salience for geometry in smaller spaces and relatively more salience for features in larger spaces. At a given distance from a corner, a narrow visual field takes in less of the geometry of larger spaces because it may view only one or two corners. In a smaller space, the same visual field may view more corners and, hence, more geometric characteristics (see Sovrano & Vallortigara, Figure 3). Indeed, chicks learn geometry better in a smaller than in a larger space, and they learn features better in a larger than in a smaller space (Chiandetti, Regolin, Sovrano, & Vallortigara, 2007). Modeling the information content of restricted visual fields, however, requires information on the viewing directions of searching animals, which is not available at the moment.

Outlook

The information content of panoramic images will not account for all experiments done in enclosed spaces. However, the possibility that search in these spaces is dominated by view similarities does generate eminently testable predictions. As we have demonstrated, the detailed distribution of light and contrast in these experimental environments plays a crucial role in modifying the relative salience of cues specifying geometry and features. This observation immediately suggests systematic ways of manipulating the salience of internal and box-related cues. For instance, the salience of roof and floor may be reduced by making them the same intensity as the majority of the walls. Our prediction would be that reducing the salience of internal cues would increase the proportion of rotational errors. Given that view-based homing strategies in visually cluttered and complicated environments have been very successful in modeling insect search behavior and in robotic implementations (e.g., Cartwright & Collett, 1983; Franz et al., 1998; Stürzl & Zeil, 2007; Vardy & Möller, 2005; Zeil et al., 2003), it would be important in future studies to specify carefully the visual input given to animals, including considerations of type, position, and intensity of illumination, all of which influence the visual salience of cues in these experimental spaces. In addition, it will be important to accurately monitor how animals actually move in these environments on a moment-to-moment basis and to learn more about early visual processing in animals involved in these experiments. As we have shown, the extent to which view-based navigation can explain how animals orient in rectangular arenas depends heavily on how they process, store, and recall remembered views.

We also have shown that virtual reality experiments are a powerful tool for testing view-based explanations of animal search behavior in simple experimental spaces. We have provided a proof of concept and show in the companion article (Cheung et al., 2007) how this tool can now be used to probe many variants of experiments designed to unravel the contributions of space shape and feature cues to the task of localizing a goal. It will be especially useful to understand and predict the search behavior of animals in nonrectangular spaces, the visual appearance of which is difficult to analyze in any other way (e.g., Graham, Good, McGregor, & Pearce, 2006; Pearce et al., 2004; Tommasi & Polli, 2004). This topic is the subject of the companion article (Cheung et al., 2007).

Conclusions

In artificially constructed spaces, animals sometimes confuse locations that stand in the same geometric relation to the shape of the space, even when the two locations differ in many local features. We have shown that explicit computations of the geometry of space are not needed to explain such rotational errors. Instead, under a range of conditions, the information contained in panoramic views can lead to rotational errors. By comparing the information content of real images with virtual reality simulations, we have demonstrated that the latter provide a powerful technique for exploring the potentials and the limits of view-based search strategies in confined experimental spaces.

References

Allied Vision Technologies. (2004). *AVT Marlin technical manual.* Stadtroda, Germany: http://www.alliedvisiontec.com/produktinfos.html.

Burgess, N. (2006). Spatial memory: How egocentric and allocentric combine. *Trends in Cognitive Sciences, 10,* 551–557.

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 8,* 679–714.

Cartwright, B. A., & Collett, T. S. (1983). Landmark learning in bees. *Journal of Comparative Physiology A, 151,* 521–543.

Chahl, J. S., & Srinivasan, M. V. (1997). Reflective surfaces for panoramic imaging. *Applied Optics, 36,* 8275–8285.

Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition, 23,* 149–178.

Cheng, K. (2005). Reflections on geometry and navigation. *Connection Science, 17,* 5–21.

Cheng, K., & Newcombe, N. S. (2005). Is there a geometric module for spatial orientation? Squaring theory and evidence. *Psychonomic Bulletin & Review, 12,* 1–23.

Cheung, A., Stürzl, W., Zeil, J., & Cheng, K. (2008). The information content of panoramic images: II. View-based navigation in nonrectangular experimental arenas. *Journal of Experimental Psychology: Animal Behavior Processes.*

Chiandetti, C., Regolin, L., Sovrano, V. A., & Vallortigara, G. (2007). Spatial reorientation: The effects of space size on the encoding of landmark and geometry information. *Animal Cognition, 10,* 159–168.

Collett, T. S., & Zeil, J. (1997). Selection and use of landmarks by insects. In M. Lehrer (Ed.), *Orientation and communication in arthropods* (pp. 41–65). Basel: Birkhäuser Verlag.

Franz, M. O., Schölkopf, B., Mallot, H. A., & Bülthoff, H. H. (1998). Where did I take that snapshot? Scene-based homing by image matching. *Biological Cybernetics, 79,* 191–202.

Gallistel, C. R. (1990). *The organization of learning.* Cambridge, MA: MIT Press.

Gibson, J. J. (1950). *Perception of the visual world.* Boston: Houghton Mifflin.

Gouteux, S., Thinus-Blanc, C., & Vauclair, J. (2001). Rhesus monkeys use geometric and nongeometric cues during a reorientation task. *Journal of Experimental Psychology: General, 130,* 505–519.

Graham, M., Good, M. A., McGregor, A., & Pearce, J. M. (2006). Spatial learning based on the shape of the environment is influenced by properties of the objects forming the shape. *Journal of Experimental Psychology: Animal Behavior Processes, 32,* 44–59.

Hermer, L., & Spelke, E. S. (1994). A geometric process for spatial reorientation in young children. *Nature, 370,* 57–59.

Hermer, L., & Spelke, E. (1996). Modularity and development: The case of spatial reorientation. *Cognition, 61,* 195–232.

Keller, J., Strasburger, H., Cerutti, D. T., & Sabel, B. A. (2000). Assessing spatial vision—Automated measurement of the contrast-sensitivity function in the hooded rat. *Journal of Neuroscience Methods, 97,* 103–110.

Learmonth, A. E., Nadel, L., & Newcombe, N. S. (2002). Children's use of landmarks: Implications for modularity theory. *Psychological Science, 13,* 337–341.

Learmonth, A. E., Newcombe, N. S., & Huttenlocher, J. (2001). Toddlers' use of metric information and landmarks to reorient. *Journal of Experimental Child Psychology, 80,* 225–244.

Neumann, T. R. (2002). Modeling insect compound eyes: Space-variant spherical vision. In H. H. Bülthoff, S.-W. Lee, T. Poggio, & C. Wallraven (Eds.), *Proceedings of the 2nd International Workshop on Biologically Motivated Computer Vision* (BMCV 2002), LNCS 2525 (pp. 360–367). Berlin: Springer-Verlag.

Newcombe, N. S. (2002). The nativist–empiricist controversy in the context of recent research on spatial and quantitative development. *Psychological Science, 13,* 395–401.

Pearce, J. M., Good, M. A., Jones, P. M., & McGregor, A. (2004). Transfer of spatial behavior between different environments: Implications for theories of spatial learning and for the role of the hippocampus in spatial learning. *Journal of Experimental Psychology: Animal Behavior Processes, 30,* 135–147.

Sovrano, V. A., & Vallortigara, G. (2006). Dissecting the geometric module: A sense linkage for metric information and landmark information in animals' spatial orientation. *Psychological Science, 17,* 616–621.

Stürzl, W., & Zeil, J. (2007). Depth, contrast and view-based homing in outdoor scenes. *Biological Cybernetics, 96,* 519–531.

Tommasi, L., & Polli, C. (2004). Representation of two geometric features of the environment in the domestic chick (*Gallus gallus*). *Animal Cognition, 7,* 53–59.

Vardy, A., & Möller, R. (2005). Biologically plausible visual homing methods based on optical flow techniques. *Connection Science, 17,* 47–89.

Wang, R. F., & Spelke, E. S. (2002). Human spatial representation: Insights from animals. *Trends in Cognitive Sciences, 6,* 376–382.

Wang, R. F., & Spelke, E. S. (2003). Comparative approaches to human navigation. In K. J. Jeffery (Ed.), *The neurobiology of spatial behaviour* (pp. 119–143). Oxford, England: Oxford University Press.

Zeil, J., Hofmann, M. I., & Chahl, J. S. (2003). Catchment areas of panoramic snapshots in outdoor scenes. *Journal of the Optical Society of America, A20,* 450–469.