Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

VPCE: A visual clustering model of place cells with multiscales, multifields, and spatial elongation

Chance J. Hamilton

chamilton4@usf.edu

University of South Florida

Alfredo Weitzenfeld

University of South Florida

Research Article

Keywords: hippocampus, place cells, multiscale, multifields, visual clustering, visual encoding, spatial representation, bio-inspired navigation

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11640306/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

VPCE: A visual clustering model of place cells with multiscales, multifields, and spatial elongation

Chance J. Hamilton[1*] and Alfredo Weitzenfeld[1]

[1*]Bellini College of Artificial Intelligence, Cybersecurity and Computing, University of South Florida, Tampa, FL, USA.

*Corresponding author(s). E-mail(s): chamilton4@usf.edu;
Contributing authors: aweitzenfeld@usf.edu;

Abstract

Until recently, most studies of hippocampal place cells have represented spatial distributions using single place fields at a single scale, while representing the fields as circular. More recent recordings in rats and bats, particularly in large environments, have shown multi-field and multi-scale spatial distributions, with some cells showing elongated fields. Based on most of these studies, traditional computational models of place cells represent place-cell spatial distributions as circular single-field, single-scale maps for individual place cells. While there are some models that do consider multiple scales, most of them are based on circular single fields, with the field locations and the number of scales specified in advance. In this work, we introduce VPCE (Visual Place Cell Encoding), an adaptive visual clustering model of place-cell spatial distributions, where multiple field scales result from the clustering of observations rather than being set in advance, and where an individual place cell may produce spatially elongated fields that are active in multiple separate regions in the environment, at varying scales. Such a multiscale and multifield population of place cells is derived exclusively from visual and depth observations, without coordinates, odometry, path integration, or reward. Agglomerative clustering of the observations is used to yield a clustering hierarchy at varying scales. We evaluated the model in multiple simulated environments and feature-space configurations against properties reported in experimental studies in both rats and bats, where smaller fields greatly outnumbered larger ones, with many place cells being active in separate regions, at multiple scales, and producing multiple fields spatially elongated near walls.

Keywords: hippocampus, place cells, multiscale, multifields, visual clustering, visual encoding, spatial representation, bio-inspired navigation

1 Introduction

Hippocampal place cells have been studied for several decades. In the classical description, a place cell fires when an animal occupies a specific location in an environment, the cell's place field (Muller et al. 1987; O'Keefe and Burgess 1996; Burgess and O'Keefe 1996). While most studies have shown that place field sizes increase along the dorsoventral axis of the hippocampus (Jung et al. 1994; Kjelstrup et al. 2008), recent recordings in large environments have shown that a single dorsal CA1 cell may be active at several locations and at more than one scale (Eliav et al. 2021; Harland et al. 2021).

Most existing experimental studies of place cells have been done in small arenas, and they established that the enclosure shapes the field. In a cylinder roughly a meter across, a CA1 cell fires over one compact region, and that region is stable across sessions, rotates with the cues on the enclosure wall, and changes when the enclosure changes shape (Muller et al. 1987; Muller and Kubie 1987). O'Keefe and Burgess (1996) showed that field shape is predicted by distance and direction to the walls rather than by absolute position, while also conforming to the wall shape. Muller et al. (1987) described crescent-shaped fields pressed against the wall of a cylinder, and linear fields running nearly the full length of a wall in rectangular enclosures, while other fields were circular or elliptical. When a rectangular box is lengthened along one axis, some fields stretch along that axis and others develop a second peak (O'Keefe and Burgess 1996). Across a population, CA1 fields vary substantially in shape as well as in size (Cohen and Drugowitsch 2025).

Burgess and O'Keefe (1996) and Hartley et al. (2000) proposed a mechanism by which the walls of an enclosure could determine where a place field lies and what shape the field takes. Both models assume a population of input cells, known as boundary vector cells, each firing when a wall lies at one particular distance in one particular direction. A place field is then the region where the summed activity of the boundary vector cells rises above a threshold. Lever et al. (2009) later recorded cells with exactly the predicted tuning in the subiculum: a subicular cell fires wherever a wall lies at the cell's preferred distance and in the cell's preferred direction, measured relative to the room rather than to the animal's head, and inserting a new wall into the enclosure creates a second region of firing beside the new wall. In entorhinal cortex, border cells fire along a wall of the enclosure and continue to do so when the enclosure is stretched (Solstad et al. 2008). On this account, the shape of a field reflects the geometry the animal senses in the environment.

Knierim and Hamilton (2011) distinguished two frames of reference for animal perception: distal cues, which lie beyond the area the animal moves through, and proximal cues, which lie on the surface it moves over. (We use the term landmark for a distal cue throughout this paper.) Whether a place field is controlled by the distal or the proximal cues is not settled. Early studies reported stronger control by distal cues, but in cue-conflict experiments, where the two sets of cues are rotated against each other, salient proximal cues can control the fields as well, and some cells follow the distal cues while others follow the proximal cues (Knierim and Hamilton 2011). Scaplen et al. (2014) found that the effect of a visual cue on place cells depends on the size of the cue, the perspective from which the animal views it, and the animal's

prior experience of it. The same object can therefore anchor a field under one set of conditions and not another. Being visible is not sufficient for a cue to control a place field, which raises a question on what kinds of visual information, such as size and shape, place fields are built from.

Wang et al. (2020) recorded CA1 cells in an environment whose floor was divided into regions of different texture, and found that the edges of place fields concentrated at the boundaries between the textures, finding that the arrangement of proximal cues is reflected in the structure of the place code. Skaggs and McNaughton (1998) recorded CA1 while rats moved through an environment containing two visually identical compartments, finding that many cells had similar fields in both compartments. Similarity in what the animal sees at separate locations is one plausible contributor, and whether such similarity is on its own sufficient is an open question. Field shape and the number of fields a cell has are considered to be both tied to the environment. Studies have shown that field size is set by position along the dorsoventral axis of the hippocampus, ranging from tens of centimeters dorsally to several meters ventrally (Jung et al. 1994; Kjelstrup et al. 2008).

Recordings in dorsal CA1 alone, in environments far larger than the classical smaller environments, found a wide range of field sizes within a single region of the hippocampus. Eliav et al. (2021) recorded dorsal CA1 in bats flying a 200 m tunnel. Field sizes ranged from 0.6 m to 32 m and were distributed log-normally, so small fields were far more common than large ones, that is, finding multiscale fields that span a wide range of sizes. Individual cells in the tunnel had several fields each, that is, multifields. Furthermore, the fields of a single cell were themselves multiscale: a cell's largest field was on average 4.4 times its smallest. To test whether this depended on the size of the environment, the authors recorded further bats flying in a 6 m segment of the same tunnel. There, most cells had only one or two fields, and the fields of a cell covered a narrower range of sizes. The same species, in a smaller space, produced a narrower range of scales. The wide range appeared from the first day of recording and in laboratory-born bats that had never been in a large space, so it did not depend on prolonged experience of such a space.

Harland et al. (2021) recorded dorsal CA1 in rats exploring an 18.6 m^2 open arena, which they call a megaspace. Most cells were multifield: 82% had between two and five subfields, where in a classic enclosure under 1 m^2 94% of cells had a single field. The fields of a cell were also multiscale. Subfield areas ranged from 0.023 m^2 to 3.46 m^2, and 79% of multifield cells had subfields whose areas ranged over more than 0.6 m^2. Across the population, small subfields greatly outnumbered large ones: the distribution of subfield sizes in the megaspace was fitted by a negative exponential, with 78% of subfields at or below 1 m^2, whereas in the smaller environments of the same study it was fitted by a Gaussian. Comparing four environments from 0.54 to 18.6 m^2, the number of subfields per cell grew linearly with the area of the environment.

Two further findings concern how the subfields were arranged, in relation to each other and in relation to the walls. Subfields overlapped far more in the megaspace than in the small environment, and Harland et al. (2021) argue that it is this overlap across many cells, rather than any difference in firing rate between a single cell's subfields, that allows the population to specify position. There was also a small accumulation

of subfields near the walls, and subfield area correlated positively, though only moderately, with distance to the nearest wall ($r = 0.36$ in the megaspace, $r = 0.40$ in the small environment). Harland et al. (2021) suggest that fields gathered at the walls because the walls were cue-rich while the center of the room was cue-poor.

Taken together, the two studies describe a population whose range of field sizes and number of fields per cell change with the size of the environment, and whose fields are arranged with respect to the walls in both large and small enclosures. All of these recordings were made in dorsal CA1, so none of these properties follows from where a cell sits on the dorsoventral axis, and in the bats the wide range of scales was present from the first recording. Three properties of the place code therefore appear to track the environment rather than the anatomy: the range of field sizes, the number of fields per cell, and the shape of a field near a boundary.

The goal of the current paper is to describe a new Visual Place Cell Encoding (VPCE) computational model able to produce multiscales and multifields as described by the experimental studies in rats (Harland et al. 2021) and bats (Eliav et al. 2021). Most existing computational models of place cells have not accounted for these properties. Furthermore, most of these models consider each cell to have a single circular field at one scale, and the models that use several scales fix the scales, and the number and position of the fields, in advance. The structure of the representation is therefore chosen by the modeler rather than produced by the environment, which leaves open the question of what shapes the recorded fields.

The VPCE computational model considers that an agent (animal or robot) records a panoramic set of camera views and a range scan at evenly spaced positions covering the environment. The observations are then grouped by similarity using agglomerative clustering. Each group defines a place cell that is active wherever the agent's current observation resembles the group, and the region of the environment over which the place cell is active produces the cell's place field. The model does not require knowing the position of the agent, and does not have knowledge of odometry, path integration, or learning rewards. Similarly, field sizes, field numbers, and field locations are not set in advance. While it is not fully known how distal and proximal cues shape place cell fields, including their size and number, in this work we propose a computational model where multiscale place fields, multifield activity, and elongated fields can arise from visual appearance and depth alone.

This work makes three contributions:

1. A new Visual Place Cell Encoding (VPCE) computational model that derives multiscale place fields, multifield activity, and elongated fields directly from visual and depth observations.
2. An evaluation of the VPCE model results across environments that differ in shape and in scale, with and without landmarks, and considering different kinds of visual and range information used alone and in combination.
3. A comparison of the resulting field-size distribution, field coverage, and field elongation with the recordings of Eliav et al. (2021) and Harland et al. (2021).

The rest of the paper is organized as follows. Section 2 reviews previously developed models. Section 3 describes the VPCE model. Section 4 describes the simulated

environments and reports the effect of field selection, distribution of field sizes, coverage of each environment, field elongation, and multifield activity. Section 5 provides the conclusion and discusses the results, limitations of the model, and future work. The paper includes a short appendix (Appendix A) defining all symbols and terms used in the paper.

2 Related work

The introduction to the paper described three properties of the recorded place code that change with the environment: the range of field sizes, the number of fields per cell, and the shape of a field near a boundary. This section reviews various existing place cell computational models and their support of spatial navigation, and their underlying modeling of either single or multiple scales and fields. Existing computational models have variations in place cell spatial representations and many have been built in support of robot navigation.

A number of models build a place field out of boundary vector cells, tuned to the distance and direction of a wall (Hartley et al. 2000; Barry et al. 2006), and representations of this kind have been used to steer a robot around obstacles (Edvardsen et al. 2020). The tuning of these cells broadens in proportion to the distance they prefer, so a cell responding to a nearby wall is more sharply tuned than one responding to a distant wall, and when an enclosure is stretched the simulated fields stay at a fixed distance from a wall while some elongate along the stretched axis or develop a second peak (Burgess et al. 2000). The boundary vector cells are assumed as input rather than derived from what the agent senses, and the size of the resulting fields follows from the widths of the tuning curves.

Arleo and Gerstner (2000) and Barrera and Weitzenfeld (2008) described single scale models of spatial cognition pairing sensory input with path integration. Tejera et al. (2018) combined models of single-scale place cells, multi-scale grid cells and head-direction cells to estimate robot positions more accurately. A cell in these models becomes active because the agent believes it is in a particular place, so the representation is only as good as its position estimates.

Milford et al. (2004) developed RatSLAM based on single-scale place cells to build a map of experiences from camera images using a model of the hippocampus. Milford and Wyeth (2012) extended the model to match sequence-based place recognition with routes across large changes in lighting and season. Both models still depend on odometry, and both are evaluated on how accurately the robot locates itself rather than on the structure of the representation they produce.

Erdem and Hasselmo (2012, 2014) describe a planning model where the agent stands still and simulates straight trajectories in candidate directions, using grid cells to track how far each simulated trajectory has reached, and coarser levels were added so that these trajectories could extend toward more distant goals, with the reach of each level fixed in advance.

Cueva and Wei (2018) and Banino et al. (2018) developed networks trained to keep track of their own position with grid-like units. Stachenfeld et al. (2017) described place fields as predictive maps of where an animal is likely to go next.

Mathis et al. (2012) analyzed how accurately an animal's position can be recovered from a population of spatially tuned place cells, finding that high resolution and large range cannot both be obtained at a single spatial period. The analysis was developed for grid codes that outperform codes in which each cell has a single field, predicting that small spacings should outnumber large ones. That prediction parallels the distributions of field size later reported in large environments (Eliav et al. 2021; Harland et al. 2021).

Llofriu et al. (2015) built a reinforcement learning model where multiple-scale place cells supply the state representation and the fields are arranged in layers of different sizes. They found that larger scales, and combinations of multiple scales, sped up learning relative to a single scale. Scleidorovich et al. (2020) separated the contribution of each layer, where small fields produced shorter paths, large fields reduced both the learning time and the number of cells required, and combining scales traded one against the other. Scleidorovich et al. (2022) found that by varying their location and size, small fields near goals and obstacle corners and large fields in open spaces improved path length, learning time and cell count together, over uniform populations of single size fields. Eliav et al. (2021) produced recordings that point the same way, where a population with fields at many scales reached a mean decoding error below 2 m with roughly twenty times fewer neurons than any single-scale population.

In each of these models the scales are settled before the agent is placed in the environment, where layers are assigned a field size in advance (Llofriu et al. 2015; Scleidorovich et al. 2020), and where the field size distribution is not uniform, the size of a field is set in advance from its distance to the nearest subgoal (Scleidorovich et al. 2022).

Hausler et al. (2020) developed work on localization driving layers of grid-like units from camera input, with the spacing of each layer chosen beforehand. Fan et al. (2017) developed a system that groups images by appearance and treats the size of a cluster as a spatial scale, with scales resulting from image grouping rather than fixed by the designer. The clusters are formed by fuzzy c-means with the number of clusters estimated in advance, and the scales chosen to maximize place-recognition performance along the route being learned. The system decides whether a new image matches a place already visited, with a scale describing how many images are grouped together rather than the size of a region of the environment.

It is an open question whether the scales an animal uses are fixed or determined by the environment (Fan et al. 2017). In many of the models previously described, scales are preset by the designer, by an estimate made before the images are grouped, or by an objective the scales need to satisfy. The same open question applies to the number of places at which a single cell is active.

Cells with more than one field have been shown to appear in recordings made in a space larger than the arenas the classical account was built from. Fenton et al. (2008) compared a 68 cm cylinder with a chamber roughly six times its area and found 4.1 fields per cell in the chamber against 1.3 in the cylinder, with 89% of place cells having more than one field in the chamber against 28% in the cylinder. Park et al. (2011) reported the same in a 1.8×1.4 m box, where roughly four in five CA1 cells had several fields, as did cells in CA3 and the dentate gyrus. Multiple fields have also been reported for rats running a 48 m track (Rich et al. 2014).

In none of these studies did the fields of one cell form a regular pattern. Only 7.6% of cells in the larger chamber showed the spatial autocorrelation characteristic of grid cells (Fenton et al. 2008), and the spacing between fields was likewise irregular in the box (Park et al. 2011). Because a cell's fields are scattered rather than tiled, the cell does not by itself say where the animal is. Harland et al. (2021) found that the peak firing rates of a cell's own subfields differ too little to distinguish them, and argued that position is specified by the overlap of subfields across many cells rather than within any one cell.

It is not clear what produces the several fields of a single cell, with recordings suggesting that there may be more than one source. One possibility is environment geometry, where lengthening a rectangular box can cause some fields to develop a second peak (O'Keefe and Burgess 1996). Boundary vector cell models have also been shown to produce second peaks, with simulated fields becoming bimodal along the stretched axis (Burgess et al. 2000). Another possibility is appearance, where an environment built with two visually identical compartments may cause more cells to produce similar fields in both sections (Skaggs and McNaughton 1998). Environment size is another possibility, where the number of fields per cell grows with the available area (Fenton et al. 2008; Harland et al. 2021).

Models have addressed multifields far less than they have addressed multiscale. The clearest exception is the megamap by Hedrick and Zhang (2016), an attractor network in which each cell is given many irregularly placed fields so that a fixed population can represent a space far larger than a conventional attractor map allows. How many fields a cell receives is drawn from a distribution chosen by the modeler, and within a simulation every field of a cell has the same width. The number of fields a cell has, and the size of each, are therefore settled before the agent is placed in the environment.

Most of the models reviewed in this section are given in advance the structure of the representation: the scales a population spans, the number of places a cell responds to, and the shape of each field. The present work develops the spatial structure based instead on available perceptions.

Section 3 describes the VPCE model, where perceptions from the environment generate how many fields a cell has, how large they are, and what shape they take.

3 Research methods

The visual clustering model groups the positions an agent samples in an environment by the similarity of the observations recorded there, so that positions whose observations are similar enough to be treated as the same place become part of the same place field.

The complete processing pipeline of the model is shown in Figure 1 and consists of six stages; stages 1 to 5 are each described in the subsection of the same title. (1) Environments and sampling: the agent records one observation per sampled position, consisting of eight camera views and one 360° range scan. (2) Feature spaces: the observations are described in four feature spaces, used alone and in combination. (3) Agglomerative clustering: positions are compared by the distance between the observation vectors recorded at the positions and are merged into a binary hierarchy

that holds clusters of multiple sizes, from a single position up to the whole environment. (4) Place fields: the members of a cluster give a centroid and a width in feature space, which define a place cell whose activity depends on the observation received. Evaluating that activity on the observations recorded throughout the environment estimates where in the environment the cell is active. The estimated place field is the region where the activity rises above a threshold set at half the cell's peak. (5) Field selection: the candidate fields within a range of sizes are selected, with a scale corresponding to the field's size. (6) Selected fields: the hierarchy holds clusters of multiple sizes and scales.

At no stage does the model receive coordinates, odometry, path integration, or reward. A selected field can occupy several separate regions, the model's counterpart of a multifield place cell. We call the fields selected in one environment, for one feature-space configuration, a set of selected fields.

3.1 Environments and sampling

Table 2 lists eight environments that differ in size, geometry, and landmarks: two circular enclosures of identical shape differing fourfold in area and two rectangular environments, 10×10 m and 10×2 m, each built twice, once with eight landmarks and once with none. Each landmark is a 0.75 m square panel with a distinct texture on the inner face of the wall. We refer to the panels as landmarks and use the term landmark interchangeably with distal cue. Panel 1 of Figure 1 shows `corr_lm8_l10w2` from above, with the environment's landmarks on the walls, and the eight camera views recorded at one sampled position. In the circular enclosures, the landmarks are evenly spaced around the wall, starting at a bearing of 90°. In the rectangular environments, the landmarks are placed at equal intervals along the perimeter, which puts two on each wall of the 10×10 m environment. An environment without landmarks shares the walls and the sampled positions of the counterpart environment with landmarks exactly, so the landmarks are the only difference within a pair.

Landmarks have the same physical size in every environment, comparable to the room cues in Harland et al. (2021).

The agent visited positions spaced at regular intervals throughout each environment, recording one observation at each. The positions cover the whole environment apart from a 0.2 m buffer along the walls, which leaves the agent room to maneuver without colliding with one.

The spacing was chosen separately for each environment, so that in every case the agent visited approximately 30,100 evenly spaced positions. The environments differ in area, so a spacing that yields this many positions in the small circular enclosure would yield many more in the large one. The number of positions is held constant because agglomerative clustering of N positions produces a hierarchy of exactly $2N-1$ clusters (Section 3.3), so that every environment offers the same number of clusters from which candidate fields are drawn and a difference between environments is not simply a difference in how densely each was sampled.

The simulator also records the distance from each sampled position to the nearest wall, which is used in the analysis of field shape (Section 4.4).

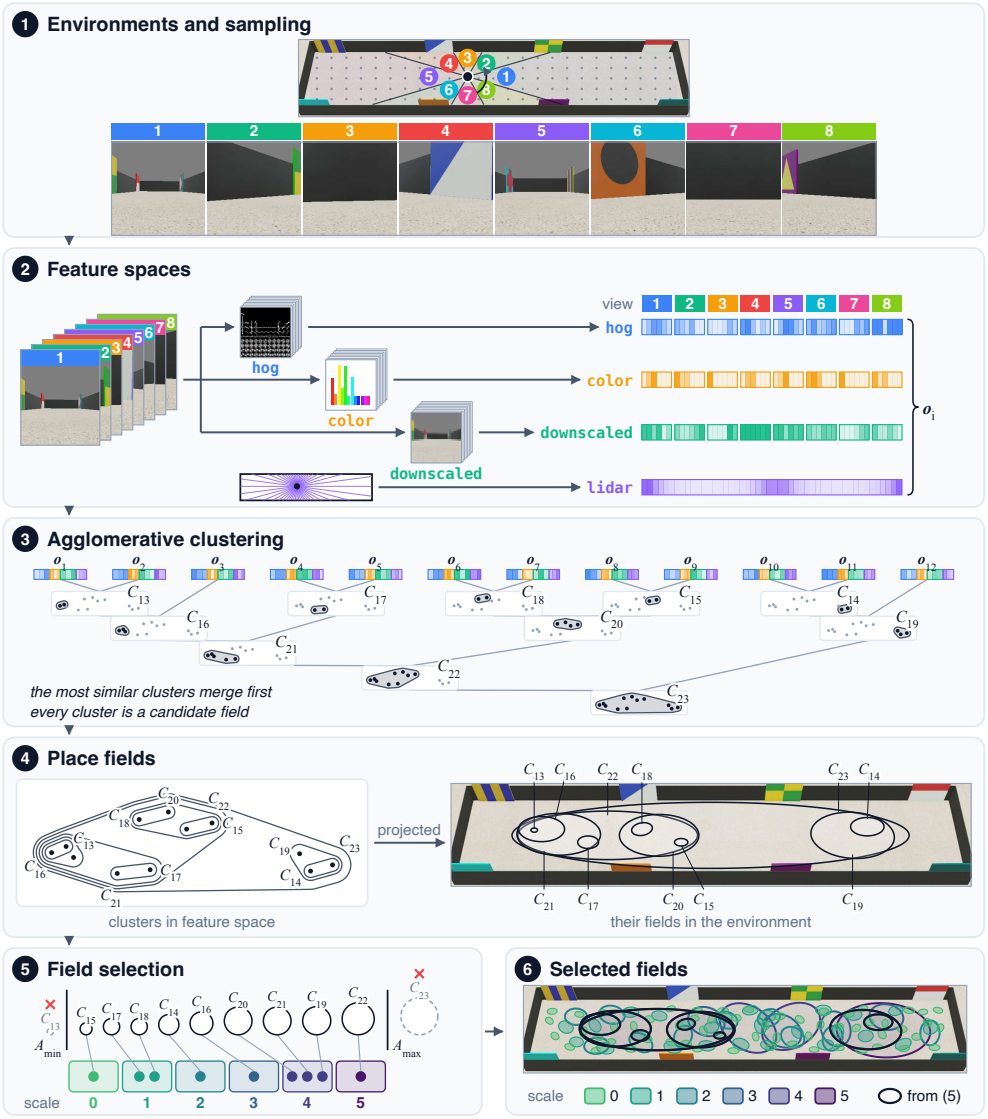


Fig. 1 The model, from the agent's camera views and range scans to the selected fields. **(1)** `corr_lm8_l10w2` (Table 2) seen from above, and the eight camera views recorded at one sampled position, in the order of the directions the views face; a range scan (`lidar`) is also recorded at each position. **(2)** Each view yields a vector in the `hog`, `color`, and `downscaled` feature spaces, and the range scan yields a vector in the `lidar` feature space; the vectors in use form $\mathbf{o}_i$. **(3)** Agglomerative clustering of 12 observation vectors, which begin as the single-member clusters $\mathcal{C}_1$ to $\mathcal{C}_{12}$. Each box shows the members (black) of one cluster formed by a merge, below the vectors or clusters merged to form the cluster, and is labeled $\mathcal{C}_{13}$ to $\mathcal{C}_{23}$ in the order of the merges. **(4)** The same clusters in feature space (left) and the fields of the clusters in the environment (right), labeled as in (3). A field is the region where a cluster's place-cell activity is above a threshold. **(5)** The fields in order of size: fields below A_{min} or above A_{max} are rejected (crosses), and the remaining fields are sorted by size into scales 0 (finest) to 5 (coarsest). **(6)** The selected fields of the whole environment, colored by scale; the selected fields shown in (5) are outlined.

3.2 Feature spaces

At each position, the agent records eight camera images and one planar range scan. The camera images are taken at headings 0°, 45°, …, 315°; each image is 224 × 224 pixels with a 45° horizontal field of view, so the eight views tile the full circle without overlap. The range scan has 360 beams spanning the full circle, with a maximum range of 40 m, and is taken at a fixed world heading. Each position is described in four feature spaces, three computed from each camera image and one from the range scan:

Edge orientation (`hog`). A histogram of oriented gradients (Dalal and Triggs 2005) computed on the grayscale image resized to 64 × 128 pixels, using 8 × 8-pixel cells, 16 × 16-pixel blocks at an 8-pixel stride, and nine orientation bins, so that each bin collects the gradients whose direction falls within one 20° interval: 3,780 values per image.
Color (`color`). A joint RGB histogram with eight bins per color channel, normalized to unit length: 512 values per image.
Coarse layout (`downscaled`). The image downsampled to 32×32 pixels and flattened: 3,072 values per image.
Depth (`lidar`). The distance measured along each of the 360 beams of the range scan.

For the three image feature spaces, the eight per-heading vectors are concatenated in a fixed heading order. Every feature space therefore describes the full circle around a position, rather than the view in whichever direction the agent happens to face. This presupposes that the agent knows which way it is facing, a signal the head-direction system supplies in the animal: head-direction cells fire as a function of the direction in which the head points, independently of the animal's location (Taube et al. 1990).

The model is run separately on six different feature-space configurations, `hog`, `color`, `downscaled`, `lidar`, `visual` (combining `hog`, `color`, and `downscaled`), and `all` (combining `visual` and `lidar`), as described in Table 3. At each position, every feature space in a configuration produces a vector. Those vectors differ widely in magnitude between feature spaces, so each is rescaled before they are combined, and the observation vector at position i is the concatenation of the rescaled vectors,

$$\mathbf{o}_i \;=\; \big[\,\mathbf{f}_i(k)/m_k\,\big]_{k\in\mathcal{K}}, \qquad m_k \;=\; \sqrt{\frac{1}{N}\sum_{j=1}^{N}\big\|\mathbf{f}_j(k)\big\|^2}, \tag{1}$$

where $\mathbf{f}_i(k)$ is the vector of feature space k at position i, $\mathcal{K}$ is the set of feature spaces the configuration uses (Table 3), and N is the number of sampled positions (Section 3.1). Dividing by m_k gives every feature space a mean squared magnitude of 1, so none dominates the distance between two positions.

3.3 Agglomerative clustering

Let $\mathbf{o}_i \in \mathbb{R}^D$ denote the observation vector recorded at position $\mathbf{p}_i \in \mathbb{R}^2$, for $i = 1, \ldots, N$, where D is the number of values in the observation vector (Table 3). The dissimilarity between two positions is the Euclidean distance between the observation vectors of the two positions, $\|\mathbf{o}_i - \mathbf{o}_j\|$: the greater the distance, the more the two

observations differ, that is, the larger the dissimilarity. In the `color` feature space, for example, the observation at a position describes the colors the agent sees there. If the same landmark is observed from two positions, the color histograms of the two positions have much of the histogram mass in the same bins, and the distance between the two histograms is small. If one of the positions observes a landmark of different colors instead, or bare wall, the mass falls in other bins and the distance is large. The dissimilarity depends only on the observations: two positions are similar if the observations at the two positions are alike, however far apart the positions may be, and dissimilar if the observations differ, however close together the positions may be.

Positions are grouped by agglomerative clustering using Ward's minimum-variance criterion (Ward 1963). A cluster $\mathcal{C}$ is a set of observation vectors that the algorithm assigns to one group, and the positions at which those vectors were recorded are the members of the cluster; $|\mathcal{C}|$ is the number of members, and the centroid $\boldsymbol{\mu}_{\mathcal{C}}$ is the mean observation vector of the members. Each of the N observation vectors begins as a single-member cluster, $\mathcal{C}_i = \{\mathbf{o}_i\}$ for $i = 1, \ldots, N$. At each step, the two clusters $\mathcal{C}_a$ and $\mathcal{C}_b$ whose merger least increases the total within-cluster sum of squares are merged into a new cluster, that is, the pair minimizing

$$\Delta_{ab} \;=\; \frac{|\mathcal{C}_a|\,|\mathcal{C}_b|}{|\mathcal{C}_a| + |\mathcal{C}_b|}\,\|\boldsymbol{\mu}_a - \boldsymbol{\mu}_b\|^2, \tag{2}$$

where $\boldsymbol{\mu}_a$ and $\boldsymbol{\mu}_b$ are the centroids of $\mathcal{C}_a$ and $\mathcal{C}_b$. The new clusters are numbered $\mathcal{C}_{N+1}, \ldots, \mathcal{C}_{2N-1}$ in the order in which the merges form them (Figure 2). Because Δ_{ab} grows with the sizes of the two clusters as well as with the distance between their centroids, the closest pair of clusters is not always the next pair to merge (Figure 2a). To compute the hierarchy, we use the Lance–Williams formulation of Ward's criterion, which takes as input the Euclidean distances between observation vectors rather than the squared distances (Murtagh and Legendre 2014).

Merging continues until a single cluster remains. The result is a binary tree of $2N - 1$ clusters (Figure 2b), which contains groups of every size, from a single position up to the whole environment.

The hierarchy holds every grouping from a single observation up to all of them, and clusters near either extreme are of no use. One with very few members is too specific: it would activate almost nowhere. The cluster holding every observation is too general: it resembles the environment as a whole and is active almost everywhere. We therefore discard clusters with too few or too many members, and treat each remaining cluster as a candidate place field.

Field selection chooses the subset of candidate fields that forms the environment's set of selected fields. Thus, clustering and projection produce the place fields, while place-field selection determines which place fields are used to map the environment.

3.4 Place fields

A cluster is a group of similar observations. Each cluster defines one place cell of the model, whose activity is set by two quantities: a centroid and a width. The cell's place field is the region of the environment where that activity is high (Section 3.5).

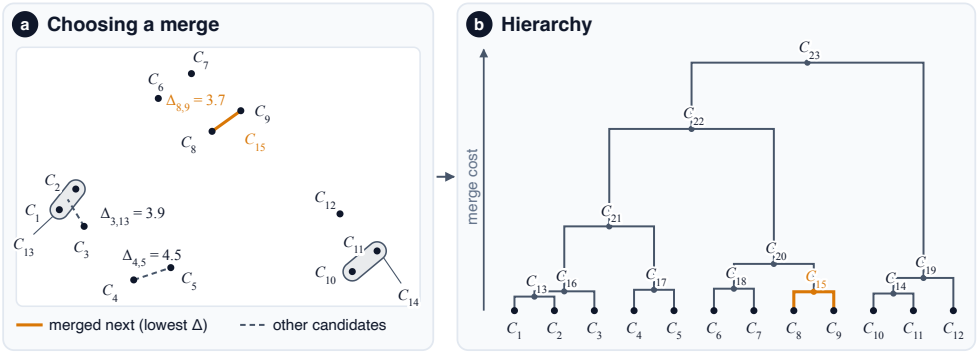


Fig. 2 Agglomerative clustering with Ward's criterion, shown on 12 observation vectors in a two-dimensional sketch of feature space. Clusters are labeled $\mathcal{C}_1$ to $\mathcal{C}_{12}$, the single-member clusters, one for each observation vector, and $\mathcal{C}_{13}$ to $\mathcal{C}_{23}$, the clusters formed by the merges, in the order the merges occur. **(a)** Feature space after two merges: the single-member clusters $\mathcal{C}_1$ to $\mathcal{C}_{12}$ (black dots) and the merged clusters $\mathcal{C}_{13}$ and $\mathcal{C}_{14}$ (outlined). Each line joins the centroids of a candidate pair $\mathcal{C}_a$ and $\mathcal{C}_b$ and is labeled with the cost Δ_{ab} of merging the pair (Equation (2)), in arbitrary units. The pair with the lowest cost, $\mathcal{C}_8$ and $\mathcal{C}_9$ (orange), is merged next and forms $\mathcal{C}_{15}$; the other candidates are dashed. **(b)** The hierarchy of all 11 merges, with the leaves $\mathcal{C}_1$ to $\mathcal{C}_{12}$ and each merge drawn at a height that increases with the merge cost; the merge that forms $\mathcal{C}_{15}$ is orange.

Place-cell activity.

Let $\boldsymbol{\mu}_{\mathcal{C}}$ be the mean observation vector of the members of cluster $\mathcal{C}$, and $\mathbf{o}(\mathbf{p})$ the observation recorded at position $\mathbf{p}$. The place-cell activity of $\mathcal{C}$ is a radial basis function of the distance between the observation and the centroid,

$$a_{\mathcal{C}}(\mathbf{p}) \;=\; \exp\!\left(-\frac{\|\mathbf{o}(\mathbf{p})-\boldsymbol{\mu}_{\mathcal{C}}\|^2}{2\sigma_{\mathcal{C}}^2}\right), \tag{3}$$

where $\sigma_{\mathcal{C}}$ is the width of the activity. The activity reaches a maximum of 1 where the observation matches the centroid exactly, and decreases as the observation becomes less similar to it (Figure 3b). We calculate $\sigma_{\mathcal{C}}$ separately for each cluster, so that the cluster's field includes 80% of the cluster's members (Figure 3a,b); for clusters with more than 512 members, $\sigma_{\mathcal{C}}$ is calculated from 512 randomly selected members.

In the `color` feature space, for example, a cluster whose members all face the same landmark has a centroid carrying that landmark's colors. The cell is then active wherever the agent sees those colors and inactive elsewhere, so two positions far apart that both face the landmark receive the same activity. The Gaussian form follows O'Keefe and Burgess (1996), who account for recorded place fields with Gaussian tuning curves over distance to walls; here the tuning curve is defined over observations instead.

We evaluate the activity of every place cell on all of the observations collected in an environment (Figure 3c), and use those activities to estimate the cell's place field, as described in Section 3.5.

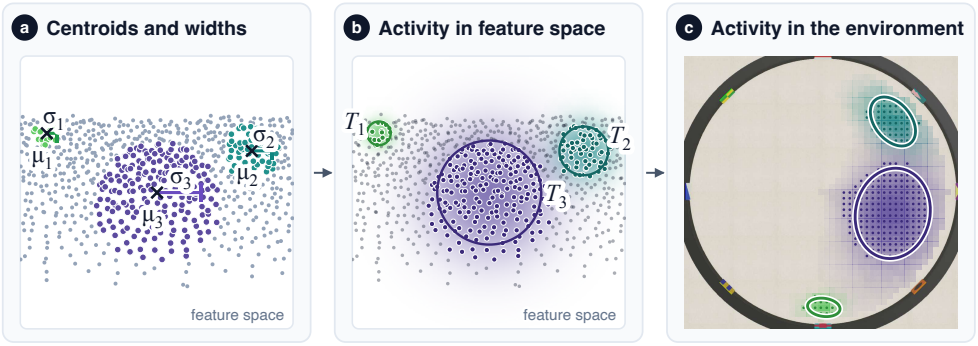


Fig. 3 How place fields are defined from clusters (Equation (3)). Three clusters of different sizes are shown, indexed 1 to 3 and each cluster drawn in a different color. **(a)** The members of each cluster and the other observations (gray) in feature space, with each cluster's centroid (×; $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$, $\boldsymbol{\mu}_3$) and width (bars; σ_1, σ_2, σ_3). **(b)** The place-cell activity of each cluster over feature space. Each circle marks where the activity falls to the cluster's field threshold (T_1, T_2, T_3) and encloses 80% of the cluster's members. **(c)** The activity projected into the environment. Each field is outlined by the field's fitted ellipse (Section 3.5), and dots mark the cluster's members.

3.5 Field geometry

A cluster's field is bounded by a threshold T on the place-cell activity, which we set at half of the cluster's peak activity,

$$T \;=\; 0.5 \max_{\mathbf{p}} a_{\mathcal{C}}(\mathbf{p}), \tag{4}$$

the maximum being taken over all sampled observations. A bin belongs to the field when its activity is at least T. The factor 0.5 is a free parameter of the model.

Activity map.

The environment is discretized into a grid of square bins, and the cluster's activity is evaluated on the observations in each bin. The field mask $M_{\mathcal{C}}$ is the set of bins whose activity is at least T, and this mask is the estimate of the field that the measurements below are taken from. Bin size is set from the sampling interval, so that the resolution of those measurements is the same fraction of the environment in every enclosure (Table 2).

Area, center, and shape.

Each field is summarized by its area, its center, and an ellipse with two semi-axes and an orientation (Figure 4).

All of these quantities are measured from the field mask. Every bin covers the same area, which in an environment of area A_{env} sampled at N positions is A_{env}/N, so the area of a field is the number of bins in its mask times the area of one bin,

$$A \;=\; |M_{\mathcal{C}}| \, \frac{A_{\text{env}}}{N}, \tag{5}$$

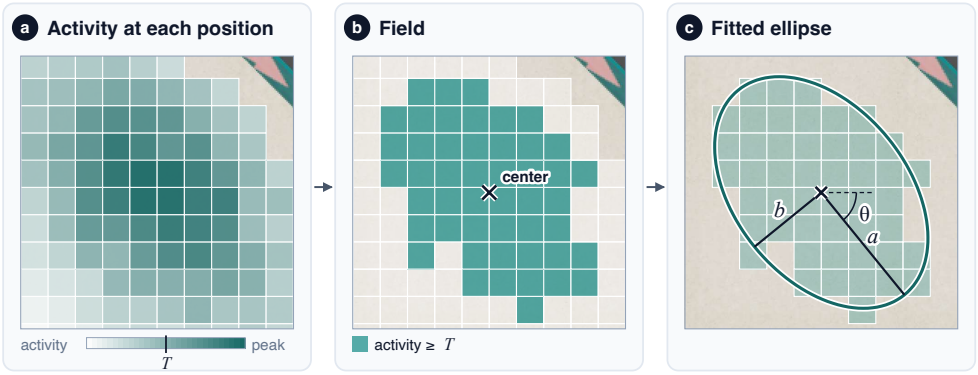


Fig. 4 How the geometry of a field is measured, shown on a close-up of the environment around one field. Each square is one sampled position. **(a)** The place-cell activity of one cluster at each position, shaded from 0 to the peak activity; T marks the field threshold on the scale. **(b)** The field, the positions whose activity is at least the field threshold T (teal), and the center of the field, the mean of these positions (cross). **(c)** The best-fit ellipse from the second-order central moments of the field, scaled to the area of the field, with the major and minor semi-axes a and b, whose ratio a/b is the elongation, and the orientation θ of the major axis.

where $|M_C|$ is the number of bins in which the cluster is active. The center (c_x, c_y) of the field is the mean position of those bins (Figure 4b). Field size is also expressed as an equivalent radius $r_{\text{eq}} = \sqrt{A/\pi}$, the radius of a circle of area A, which is the quantity the scale assignment of Section 3.6 is based on.

The shape of the region is described by an ellipse rather than a circle. Recorded place fields are not circular: their shape depends on the geometry of the enclosure, and they stretch along walls (Muller et al. 1987; O'Keefe and Burgess 1996). A circle reduces a field to a single radius and discards that structure. Cohen and Drugowitsch (2025) give a further reason to measure shape: for a population encoding position in two dimensions, a spread of field sizes adds little unless field shape varies as well. We calculate the best-fit ellipse from the field's second-order central moments, a classic region-based approach outlined by Teague (1980). The ellipse is scaled so that its area equals A (Figure 4c), and it gives a major and a minor semi-axis $a \geq b$ and the orientation θ of the major axis; we define the elongation as a/b. Elongation is the measure of field shape analyzed in Section 4.4, where it is related to field scale and to the distance of the field's center from the wall.

Subfields.

A cluster's place-cell activity can occupy several separate regions of the environment. We call each 8-connected region of the mask M_C that is at least as large as the lower size bound A_{min} (Equation (6)) a subfield and treat each subfield as an individual field: the subfield's area, center, semi-axes, and orientation are measured in the same way from the subfield's own positions. Regions smaller than A_{min} are not considered fields. A place cell with two or more selected subfields shows multifield activity, in analogy to the recordings of a single place cell with several fields (Harland et al. 2021; Eliav et al. 2021).

3.6 Field selection

From the candidate place fields produced by agglomerative clustering (one for every subfield of every cluster carried through projection), a subset is reported as the environment's set of selected fields. Candidate fields are tested against a size criterion, which corresponds to a property by which place fields are identified in the recorded literature, and the candidate fields that pass the criterion become selected fields. Field selection keeps multifield activity (Section 3.5). Each subfield of a cluster is tested against the criterion separately, corresponding to individual cells with several fields having been recorded in large environments (Harland et al. 2021; Eliav et al. 2021).

Size.

A field must satisfy $A_{\min} \le A \le A_{\max}$, with

$$A_{\min} = \frac{0.023}{18.6}\, A_{\text{env}} \approx 1.24 \times 10^{-3}\, A_{\text{env}}, \qquad A_{\max} = 0.20\, A_{\text{env}}. \tag{6}$$

where A is the area of the field (Section 3.5) and A_{env} is the area of the environment (Table 2). The lower bound corresponds to the smallest subfield reported by Harland et al. (2021), $0.023\,\text{m}^2$ in the $18.6\,\text{m}^2$ environment of that study (Fig. 3A of Harland et al. (2021)), expressed as a fraction of the $18.6\,\text{m}^2$ environment. The upper bound is set from Harland et al. (2021), the only one of the two large-environment studies to report field area: the largest subfield there was $3.46\,\text{m}^2$, 19% of that environment, and the bound lies just above it. Fields in the 200 m tunnel of Eliav et al. (2021) are one-dimensional, measured as extents along the tunnel rather than as areas, so that study does not calibrate an area bound; its largest field, 32 m, does show that fields in large environments can occupy a substantial fraction of the space. A field at the upper bound already spans a fifth of the environment, and activity spread over so large a region specifies position only coarsely. Both bounds are expressed as fractions of environment area rather than as absolute areas because field size itself scales with the environment (Muller and Kubie 1987; Burgess and O'Keefe 1996).

Scales.

Each selected field is assigned to one of a geometric series of scales,

$$\beta \;=\; \left\lfloor \frac{\ln(r_{\text{eq}}/r_{\min})}{\ln \rho} \right\rfloor, \qquad \rho = 1.6, \tag{7}$$

where β is the scale of the field, numbered from 0 for the finest, r_{eq} is the equivalent radius of the field (Section 3.5), and $r_{\min} = \sqrt{A_{\min}/\pi}$ is the equivalent radius of the smallest allowed field, so that neighboring scales differ in radius by a factor of ρ. The ratio ρ is based on the argument of Wei et al. (2015), who derive from an economy of neurons that neighboring grid scales should differ by a constant factor of about $\sqrt{e} \approx 1.65$. Scales group the selected fields for analysis only; scales do not determine which fields are selected.

3.7 Parameters

Table 1 lists the parameters of the model, the value used for all reported results, and the basis for each value. The size bounds and the scale ratio are taken from published measurements and derivations; the remaining parameters are either physical, fixed by the simulated sensors and the sampling design, or conventions with no measured counterpart. Table A1 (Appendix A) lists the symbols and terms used in the paper.

Table 1 Model parameters. *Physical*: fixed by the simulated sensors or sampling design. *Literature*: taken from a published measurement or derivation. *Convention*: a chosen value with no measured counterpart.

Symbol	Parameter	Value	Basis
Sampling and sensing (Sections 3.1 and 3.2)			
N	Positions per environment	$\approx$ 30,100	Physical
–	Camera headings	8, at 45°	Physical
–	Camera resolution and field of view	224×224 px, 45°	Physical
–	Range beams and maximum range	360, 40 m	Physical
–	Wall buffer	0.2 m	Physical
Place fields (Sections 3.4 and 3.5)			
–	Fraction of a cluster's members inside its field	80%	Convention; sets $\sigma_{\mathcal{C}}$ (Equation (3))
–	Members sampled to set $\sigma_{\mathcal{C}}$	512	Convention; applies to larger clusters only
T	Field threshold (Equation (4))	0.50× peak activity	Convention
Field selection (Section 3.6)			
$A_{\min}$	Lower size bound	$1.24 \times 10^{-3}\, A_{\text{env}}$	Literature (Harland et al. 2021)
$A_{\max}$	Upper size bound	$0.20\, A_{\text{env}}$	Literature (Harland et al. 2021)
ρ	Ratio between scales	1.6	Literature (Wei et al. 2015)

4 Experiments and results

Experiments were run in Webots R2025a (Michel 2004) using the REALM simulation framework (USF BioRobotics Lab 2026). Running the model once produces the fields selected in one environment under one feature-space configuration, which we call a *set of selected fields*. Each of the eight environments in Table 2 was run under each of the six feature-space configurations in Table 3, giving $8 \times 6 = 48$ sets of selected fields,

Table 2 Environments. Each identifier gives the geometry (`circ`, circular; `corr`, rectangular), the number of landmarks (`lm`), and the radius (`r`) or the length and width (`l`, `w`) in meters. Landmarks is the number of textured panels on the wall, area is the area of the enclosure, and positions is the number of sampled positions at which the agent recorded observations.

Environment	Geometry	Landmarks	Area (m^2)	Positions
`circ_lm8_r3`	Circular, radius 3 m	8	28.27	30,129
`circ_lm8_r6`	Circular, radius 6 m	8	113.10	30,129
`corr_lm8_l10w10`	Rectangular, 10×10 m	8	100.00	29,929
`corr_lm8_l10w2`	Rectangular, 10×2 m	8	20.00	30,175
`circ_lm0_r3`	Circular, radius 3 m	0	28.27	30,129
`circ_lm0_r6`	Circular, radius 6 m	0	113.10	30,129
`corr_lm0_l10w10`	Rectangular, 10×10 m	0	100.00	29,929
`corr_lm0_l10w2`	Rectangular, 10×2 m	0	20.00	30,175

Table 3 Feature spaces and combinations of feature spaces (Section 3.2). Dimension is the number of values in the observation vector at one position, derived from the eight camera views captured there or, for `lidar`, from the range scan.

Feature space	Content	Dimension
`hog`	Edge orientation histograms	30,240
`color`	RGB color histogram	4,096
`downscaled`	Downsampled image	24,576
`lidar`	Range scan	360
`visual`	`hog` + `color` + `downscaled`	58,912
`all`	`visual` + `lidar`	59,272

and every result in this section is computed from those 48 sets. All sets were built with the parameter values in Table 1. We refer to the environments as the small and large circular enclosures (radius 3 m and 6 m) and the wide and narrow rectangular environments (10×10 m and 10×2 m), each with or without landmarks, and to the feature-space configurations by the content of each configuration: edge orientation (`hog`), color (`color`), coarse layout (`downscaled`), depth (`lidar`), the three visual descriptors combined (`visual`), and all feature spaces combined (`all`). We report five analyses: (1) the effect of field selection on the candidate fields, (2) the distribution of field sizes across scales, (3) how the fields cover each environment, (4) field elongation and the relation of elongation to distance from the boundary, and (5) the multifield activity the model produces.

4.1 Effect of field selection

Agglomerative clustering produces a large pool of candidate fields. Every cluster in the hierarchy is projected into the environment, and each subfield of the cluster becomes

a candidate field (Section 3.5). Across the 48 sets of selected fields, 122,868 clusters produced 123,722 candidate fields, between 2,492 and 2,773 per set (median 2,562).

Not all candidate fields are reasonable components of a representation of the environment. Many are smaller than the smallest place fields recorded in large environments. The model therefore selects a subset of the candidate fields with a biologically motivated size criterion (Section 3.6), which keeps fields within the range of sizes recorded in large environments. The selected fields are the ones that map the environment. We recorded for every candidate field whether the candidate field was rejected, and at which size bound.

Of all candidate fields, 63.1% were rejected by the size criterion and 36.9% (45,678) were selected. The selected fields include 1.2% of all candidate fields that were selected as subfields of multifield place cells (Section 4.5).

Size.

Almost all rejections by the size criterion occurred at the lower bound: 62.9% of candidate fields fell below the lower bound, and 0.2% (242) exceeded the upper bound. A large rejection at the lower bound is expected from how the hierarchy is built. Agglomeration begins with every sampled position as a single-position cluster, so most clusters in the hierarchy contain few positions and project to very small fields. In the small circular enclosures, for example, the median candidate field had an equivalent radius of 0.088 m to 0.094 m across feature spaces, below the lower bound of 0.105 m, and the same held in every environment. The lower bound, which corresponds to the smallest subfield recorded by Harland et al. (2021) scaled to environment area, removed the fine-grained candidate fields. The upper bound rarely acted, but this follows from the structure of the hierarchy rather than from the observations: a binary tree holds few clusters large enough to project to a field above a fifth of the environment (Section 3.3). The bound nonetheless removes the few oversized fields the hierarchy does produce, so every selected field lies within the range of sizes recorded in large environments.

Scales.

Every candidate field within the size bounds was selected, so the number of selected fields at each scale is the number of candidate fields within the bounds that the hierarchy produced at that scale. The number of selected fields fell by a factor of about 170 from the finest scale to the coarsest, from 27,376 at scale 0 to 159 at scale 5. A coarse field and the finer fields nested within the coarse field were retained together, as in recordings in which small and large fields overlap (Harland et al. 2021) and fields at different dorsoventral levels coexist (Kjelstrup et al. 2008).

Differences between environments and feature spaces.

The proportion of candidate fields selected depended more on the environment than on the feature space. The proportion selected was 37–40% in the circular enclosures and the wide rectangular environment, and 34.3% and 29.5% in the narrow rectangular environment with and without landmarks. In the narrow rectangular environment, the size criterion removed the largest share of candidate fields (65.7–70.5%). Across feature

spaces, the share selected ranged from 35.2% for coarse layout to 39.7% for color, and combining all feature spaces (35.7%) did not select more fields than color alone. Across individual sets, the share selected ranged from 23.3% (coarse layout, narrow rectangular environment without landmarks) to 43.4% (depth, large circular enclosure without landmarks). Sets of selected fields in the narrow rectangular environment were therefore built from a smaller share of the environment's candidate fields than the sets in the circular enclosures, and the difference should be kept in mind when the two geometries are compared below.

4.2 The distribution of field sizes

The 48 sets of selected fields contained 45,678 fields in all, between 645 and 1,159 per set (median 954). All sets but one contained fields at all six scales. The edge-orientation set of the narrow rectangular environment with landmarks had no fields at the coarsest scale (Figures 5 to 8).

Shape of the distribution.

In every set, small fields greatly outnumbered large ones. Pooled across sets, 60.0% of fields lay at the finest scale and 24.1%, 9.9%, 4.1%, 1.5%, and 0.3% at the five coarser ones, so each scale held roughly 0.4 times as many fields as the next finer scale. In every environment and for every feature space, the share of fields was highest just above the smallest allowed size and fell steeply toward larger sizes (Figure 9). The median field area was 2.0 times the lower size bound (range 1.95–2.51 across sets), and 48% of fields were less than twice the bound. Both recording studies found the same pattern of many small fields and few large ones. Eliav et al. (2021) described field sizes in the tunnel with a log-normal distribution, and in the megaspace 78% of subfields were at or below $1\,\mathrm{m}^2$, with sizes fitted by a negative exponential (Harland et al. 2021). The model agrees with both studies in the predominance of small fields.

We do not fit a parametric form to the field-size distribution. What the model shares with the recordings is the shape of the distribution rather than a particular family: small fields predominate, but the coarser scales are occupied as well, in all but one of the 48 sets. The model therefore produces a range of scales rather than small fields alone.

Spread of field sizes.

Field sizes spanned a wide range. The largest field in a set was typically 147 times the smallest in area (12.1 times in radius). Fields were small relative to the environment: the median field covered 0.25% of the environment's area, and 98% of fields lay in the four finest scales, all of which are below 5.3% of the environment's area. In the megaspace, 78% of subfields were at or below $1\,\mathrm{m}^2$, or 5.4% of the megaspace (Harland et al. 2021). Relative to environment size, the model's sets of selected fields therefore contained proportionally more small fields than the recorded population.

Outside the narrow rectangular environment, the spread differed little between feature spaces. In the circular enclosures and the wide rectangular environment, the median ratio of the largest to the smallest field ranged from 144 (color) to 154 (edge orientation), and every feature space occupied all six scales. Across all environments,

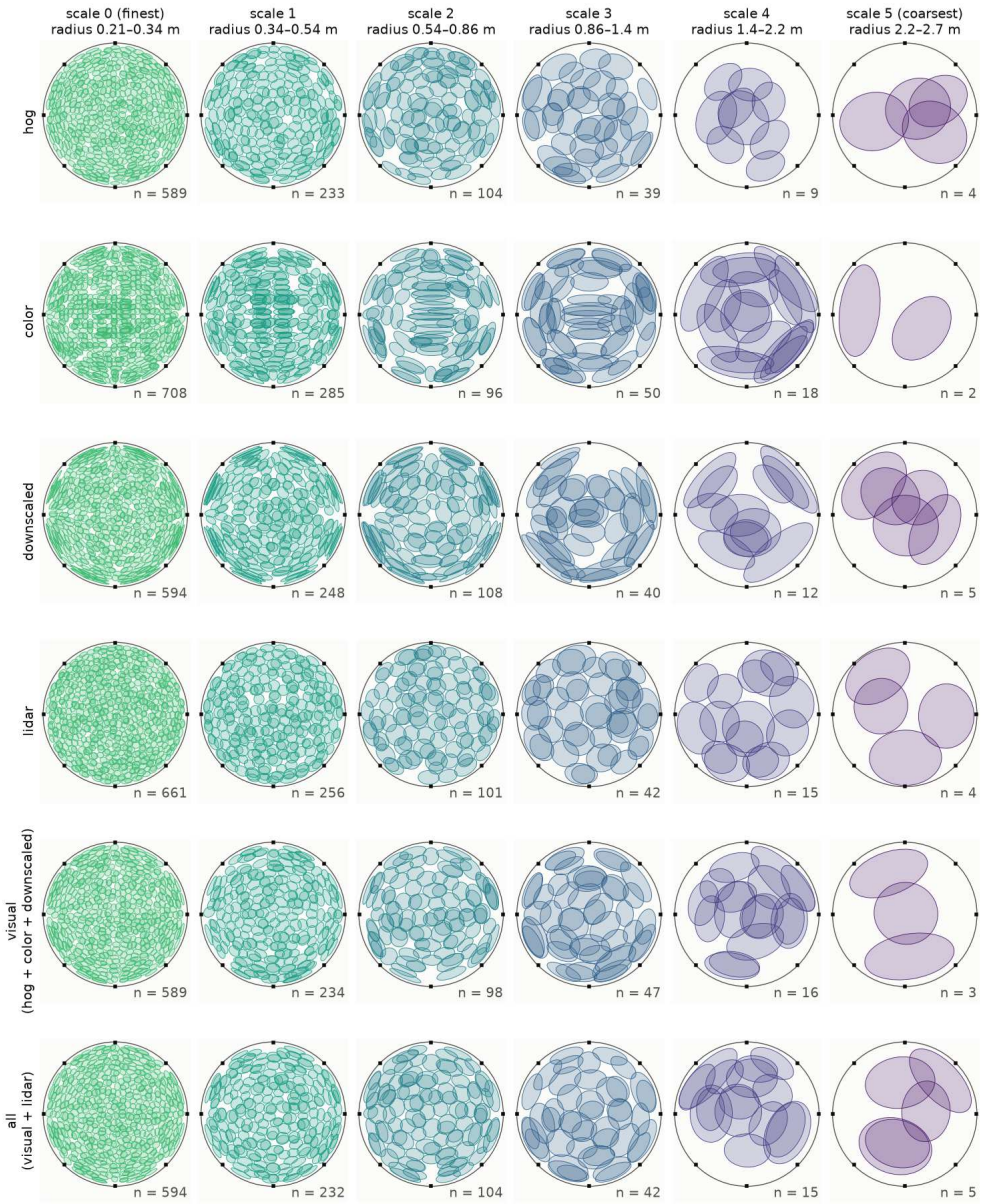


Fig. 5 Selected fields in `circ_lm8_r6` (Table 2), arranged by feature-space configuration (rows) and scale (columns). Each panel shows the enclosure from above with the enclosure's eight landmarks (black squares), and each selected field is drawn as the field's fitted ellipse, colored by scale from the finest (green) to the coarsest (purple). The heading of each column gives the range of equivalent radius that defines the scale, and n is the number of selected fields in the panel.

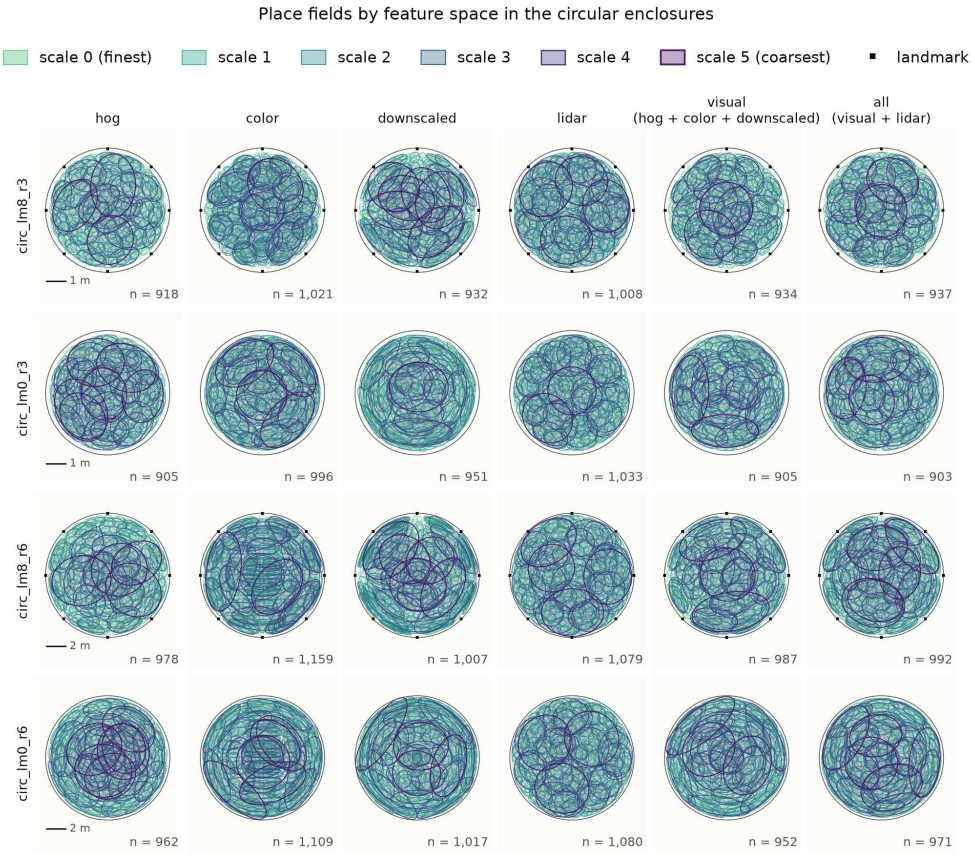


Fig. 6 Selected fields in the four circular enclosures (Table 2), with one row for each environment and one column for each feature-space configuration. Each panel shows the enclosure from above, with the enclosure's landmarks (black squares) where the enclosure has landmarks, and draws every selected field of the set as the field's fitted ellipse, with the six scales overlaid and colored from scale 0, the finest (green), to scale 5, the coarsest (purple). The scale bar in the first column applies to all panels in the same row, and n is the number of selected fields in the set.

color yielded the most fields (median 1044 per set) and all feature spaces combined the fewest (924). The narrow rectangular environment with landmarks was the exception. There, the ratio varied more between feature spaces, from 110 for the combined visual descriptors to 153 for coarse layout, with 138 for depth.

Geometry and landmarks.

The eight environments differ in shape, in area, and in whether landmarks are present (Table 2). We report the sets of selected fields across all eight rather than treating environment area as a controlled variable. Between the small and large circular enclosures, the median field area rose from 0.071 to 0.287 m^2 both with and without landmarks, in proportion to area. Because the size bounds are fixed fractions of environment area (Equation (6)), the increase in median field area reflects the scaling of the bounds.

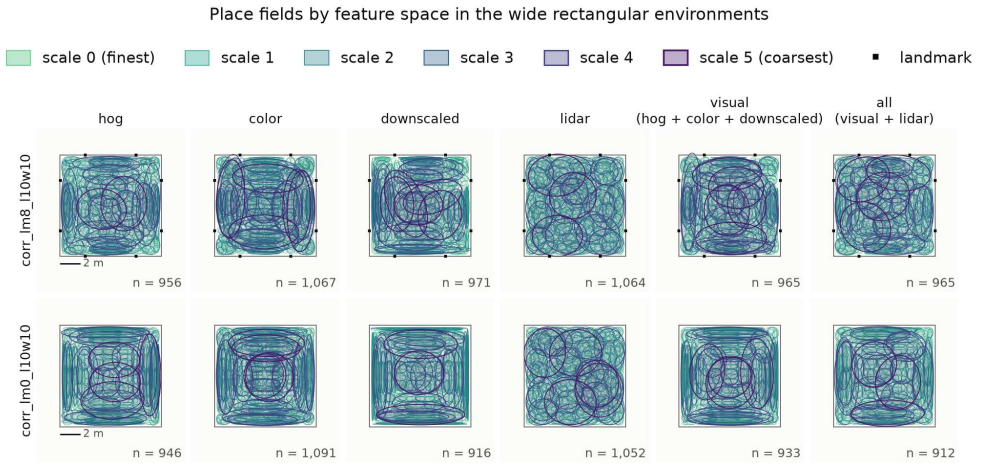


Fig. 7 Selected fields in the two wide rectangular environments (Table 2), with one row for each environment and one column for each feature-space configuration, drawn as in Figure 6. The scale bar in the first column applies to all panels in the same row, and n is the number of selected fields in the set.

Relative to the lower bound, the median field was nearly the same size in both enclosures (2.03 and 2.07 times the bound with landmarks; 2.02 and 2.04 without), and the median field covered 0.25–0.26% of each enclosure. The spread also changed little: the ratio of the largest to the smallest field rose from 147 to 156 with landmarks and from 144 to 148 without. The model therefore produced nearly the same relative distribution of field sizes in both enclosures.

Landmarks can affect the sets of selected fields only through the visual feature spaces, because the range scan does not detect the landmark panels. Accordingly, the depth sets were nearly the same with and without landmarks in every geometry, differing by at most 25 fields (2.4%). For the feature spaces that include visual descriptors, the effect of landmarks depended on geometry. In the narrow rectangular environment, landmarks increased the number of selected fields by factors of 1.0 (color) to 1.5 (coarse layout), whereas in the wide rectangular environment and the circular enclosures landmarks changed the number of selected fields little (factors of 1.0 to 1.1 and 0.98 to 1.05). The median field size relative to the lower bound was essentially unchanged by landmarks in every geometry (2.0–2.1 times the bound).

4.3 Coverage of the environment

Harland et al. (2021) describe the megaspace representation as an ensemble of overlapping small and large fields from many simultaneously active cells. We quantified the overlap among the selected fields with the redundancy

$$R = \frac{\sum_{k=1}^{K} A_k}{A_{\mathrm{cov}}}, \qquad A_{\mathrm{cov}} = \Big| \bigcup_{k=1}^{K} F_k \Big|, \tag{8}$$

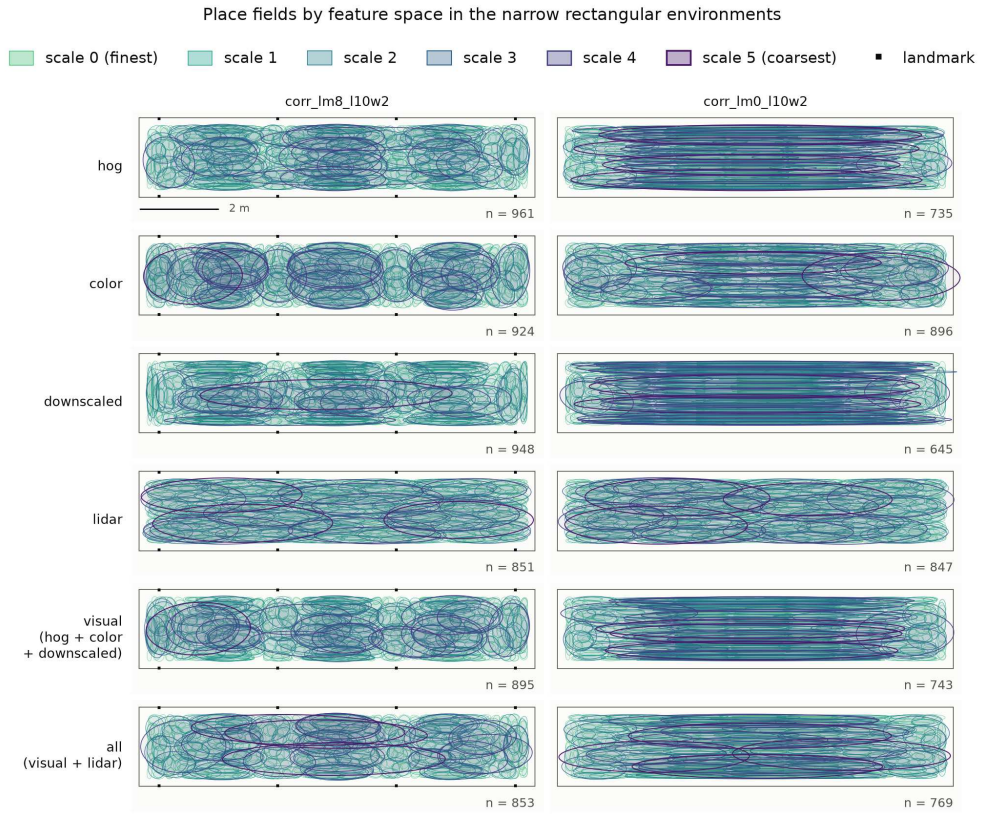


Fig. 8 Selected fields in the two narrow rectangular environments (Table 2), with one column for each environment and one row for each feature-space configuration, drawn as in Figure 6. The scale bar in the first panel applies to all panels, and n is the number of selected fields in the set.

where K is the number of fields, F_k is the region occupied by field k, A_k is the area of field k, and A_{cov} is the area covered by at least one field, computed by drawing each field as the field's fitted ellipse. R is at least 1, because a sum of areas cannot be smaller than the area of the union of the regions. $R = 1$ means that the fields tile the covered area without overlap, and $R = 3$ means that the covered area is, on average, three fields deep. We computed R in two ways. The across-scale redundancy (ASR) uses all fields of a set of selected fields together. The within-scale redundancy (WSR) applies the same ratio to the fields of each scale separately and takes the median over scales. A high across-scale redundancy with a within-scale redundancy near 1 indicates overlap that arises from nesting across scales, whereas high values of both redundancies would indicate redundant fields of the same size.

Across the 48 sets of selected fields, the across-scale redundancy had a median of 7.4 (interquartile range 7.1–7.6), whereas the within-scale redundancy had a median of 1.53 and never exceeded 1.80 (Table 4). The covered area was therefore, on average, about seven fields deep when fields of all scales were counted, but about one and a half fields deep within a single scale: fields of the same size overlapped far less than fields of

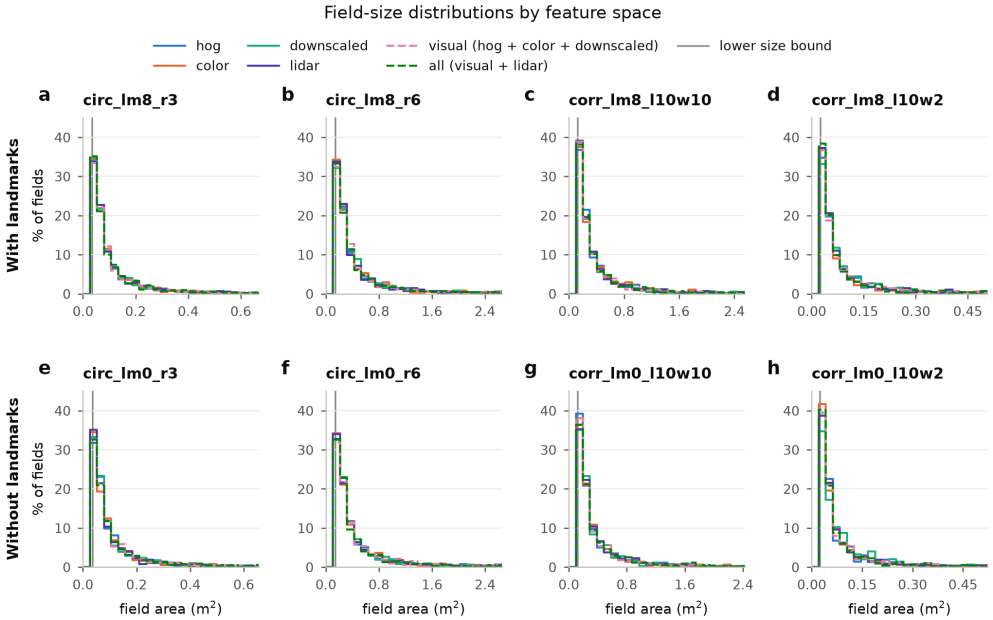


Fig. 9 Distribution of the area of the selected fields in each environment (Table 2); panels (a–d) show the environments with landmarks and panels (e–h) the environments without landmarks. Each line is the histogram for one feature-space configuration and gives the percentage of the configuration's selected fields in each bin of field area: solid lines for `hog`, `color`, `downscaled`, and `lidar`, and dashed lines for the combinations, `visual` and `all`. The vertical gray line marks the lower size bound.

Table 4 Across-scale redundancy (ASR) and within-scale redundancy (WSR) of the selected fields (Equation (8)) in each environment (Table 2) and feature-space configuration (Table 3). ASR uses all selected fields of a set; WSR is computed for each scale separately, and the median over scales is given.

	`hog`		`color`		`downscaled`		`lidar`		`visual`		`all`	
Environment	ASR	WSR	ASR	WSR	ASR	WSR	ASR	WSR	ASR	WSR	ASR	WSR
`circ_lm8_r3`	6.6	1.48	8.0	1.70	7.1	1.48	7.5	1.54	7.1	1.51	7.2	1.51
`circ_lm8_r6`	6.7	1.44	8.2	1.80	7.6	1.62	7.5	1.53	7.3	1.48	7.6	1.52
`corr_lm8_l10w10`	7.2	1.55	8.4	1.79	7.4	1.67	7.5	1.58	7.9	1.55	7.6	1.47
`corr_lm8_l10w2`	7.8	1.79	8.6	1.78	7.4	1.73	7.2	1.50	7.8	1.76	7.5	1.52
`circ_lm0_r3`	7.5	1.49	7.5	1.57	6.7	1.55	7.4	1.60	7.2	1.45	7.1	1.50
`circ_lm0_r6`	7.5	1.53	7.6	1.63	6.9	1.55	7.6	1.56	7.2	1.50	7.6	1.46
`corr_lm0_l10w10`	6.7	1.43	7.8	1.74	6.7	1.51	7.6	1.52	6.8	1.47	6.8	1.41
`corr_lm0_l10w2`	7.0	1.45	7.4	1.63	6.9	1.57	7.4	1.52	6.6	1.40	6.9	1.43

Table 5 Spearman correlation of elongation with scale and with wall distance in each environment (Table 2), combining the selected fields of all feature-space configurations. Fields is the number of selected fields. Scale and wall distance give the correlations over all selected fields; the last five columns give the correlation with wall distance within each of scales 0–4, which holds field size approximately constant. An asterisk marks a correlation that fails the Benjamini–Hochberg test at $q = 0.05$, that is, a correlation that is not significant after adjustment for multiple comparisons. Every entry is a Spearman rank correlation coefficient and is therefore dimensionless.

		Spearman correlation of elongation with						
				wall distance at scale				
Environment	Fields	scale	wall distance	0	1	2	3	4
`circ_lm8_r3`	5,750	−0.04	−0.58	−0.60	−0.58	−0.58	−0.40	−0.39
`circ_lm8_r6`	6,202	0.07	−0.62	−0.64	−0.60	−0.60	−0.55	−0.61
`corr_lm8_l10w10`	5,988	0.07	−0.56	−0.58	−0.54	−0.52	−0.64	−0.77
`corr_lm8_l10w2`	5,432	0.09	−0.41	−0.40	−0.45	−0.48	−0.41	−0.50
`circ_lm0_r3`	5,693	−0.03	−0.58	−0.55	−0.62	−0.61	−0.66	−0.62
`circ_lm0_r6`	6,091	0.02*	−0.65	−0.65	−0.66	−0.64	−0.58	−0.73
`corr_lm0_l10w10`	5,850	−0.01*	−0.58	−0.58	−0.54	−0.62	−0.66	−0.85
`corr_lm0_l10w2`	4,635	0.10	−0.14	−0.11	−0.18	−0.22	−0.27	−0.47
All environments	45,678	0.03	−0.36	–	–	–	–	–

different sizes. Harland et al. (2021) describe the same structure: overlapping fields of different sizes rather than many fields of the same size. The within-scale redundancy varied little between environments and feature spaces (1.40–1.80). The across-scale redundancy was most uniform for depth (7.2–7.6 in every environment) and was the same in the narrow rectangular environment as elsewhere (median 7.4 in both).

4.4 Field shape near a boundary

For each selected field we recorded the field's elongation a/b and the distance from the field's center to the nearest point on the boundary. Wall distance is normalized within each environment, from 0 at the closest a field center can lie to the boundary to 1 at the center of a circular enclosure or the midline of a rectangular environment. We computed two Spearman correlations, scale against elongation and wall distance against elongation, for each environment, for each environment and feature space, and for each environment and scale, the last holding field size approximately constant. We adjusted the p values for multiple comparisons with the Benjamini–Hochberg procedure and report a correlation as significant when the correlation's Benjamini–Hochberg q value is below 0.05. The analysis is descriptive, with no null model. Table 5 summarizes the correlations for each environment, and Table 6 gives the correlation with wall distance for each environment and feature space.

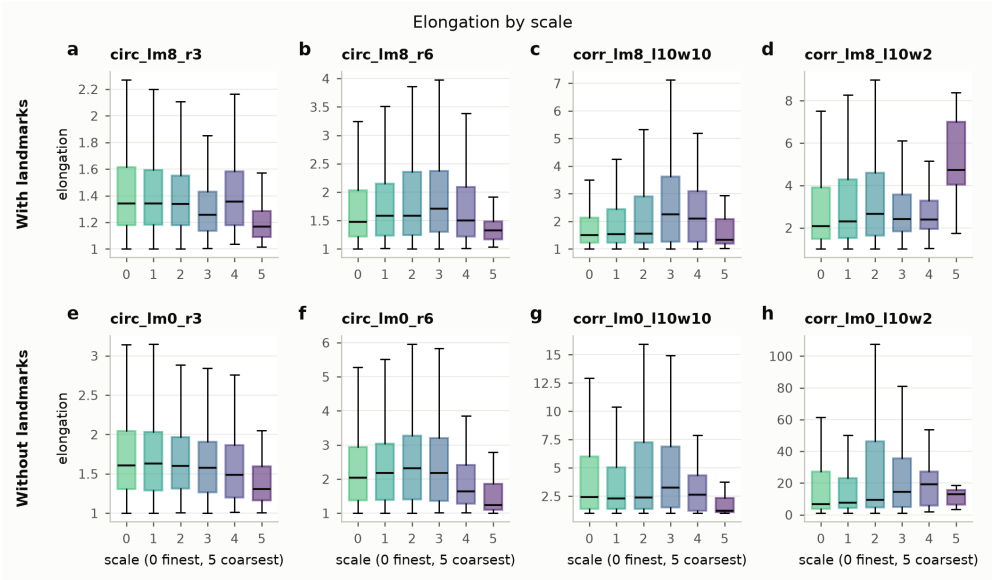


Fig. 10 Elongation of the selected fields at each scale, from 0 (finest) to 5 (coarsest), in each environment (Table 2); panels (a–d) show the environments with landmarks and panels (e–h) the environments without landmarks. Elongation is the ratio of a field's major to minor semi-axis, and elongation equals 1 for a circular field. Each box combines the selected fields of all feature-space configurations at one scale, spans the interquartile range of the fields' elongation, and marks the median elongation with a black line. Boxes are colored by scale from the finest (green) to the coarsest (purple).

Elongation and scale.

Most fields were elongated: at scales 0–4, the median elongation within a scale lay between about 1.3 and 3.3 in the circular enclosures and the wide rectangular environment and between 2.1 and 19.2 in the narrow rectangular environment (Figure 10). Field shape therefore varied across the population, as Cohen and Drugowitsch (2025) report for CA1 place fields. Elongation did not, however, vary systematically with scale. The correlation between scale and elongation was close to zero in every environment (Table 5) and for most feature spaces. Of the 48 correlations computed for each environment and feature space, 20 were significant, and all were small: 6 for color (-0.21 to 0.27), 5 for coarse layout (-0.08 to 0.27), 4 for all feature spaces combined (-0.11 to 0.19), 2 each for edge orientation (0.12 to 0.17) and the three visual descriptors combined (0.08 to 0.14), and 1 for depth (-0.08). In every geometry, fields were less elongated with landmarks than without (Figure 10).

Elongation and distance from the wall.

Recorded place fields near a boundary can conform to the boundary, taking crescent-shaped or linear forms along the wall (Muller et al. 1987). The model's fields were likewise more elongated near a wall (Figure 11); the analysis measured elongation only, not alignment with the wall. Elongation decreased with wall distance in every environment (Table 5), most strongly in the circular enclosures and the wide rectangular

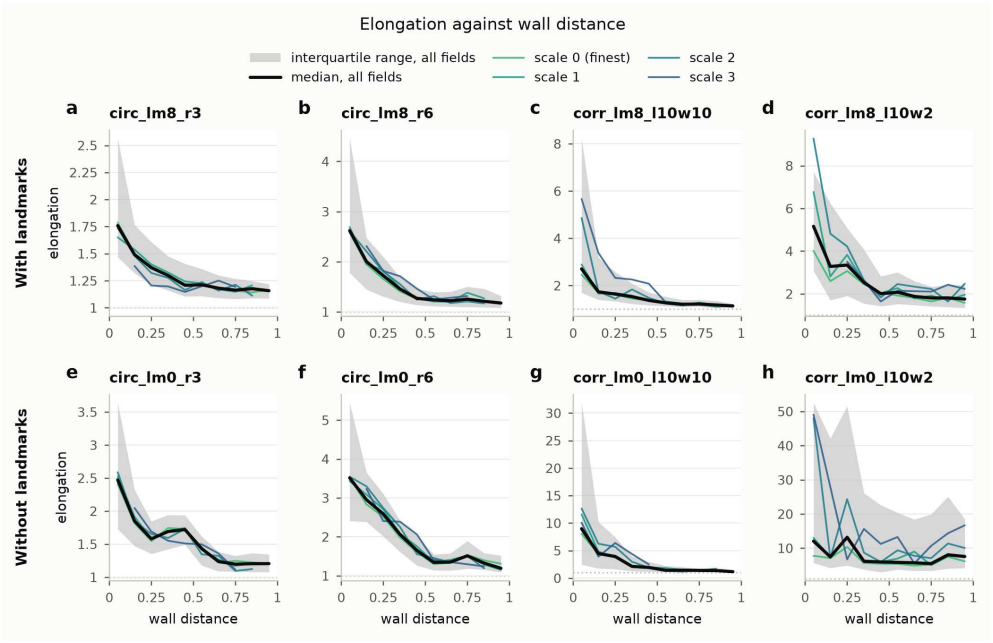


Fig. 11 Elongation of the selected fields against wall distance in each environment (Table 2); panels (a–d) show the environments with landmarks and panels (e–h) the environments without landmarks. Wall distance runs from 0, the closest a field center can lie to the wall, to 1 at the center of a circular enclosure or the midline of a rectangular environment. Each panel combines the selected fields of all feature-space configurations in bins of wall distance. The black line is the median elongation in each bin and the gray band the interquartile range of elongation in each bin; each colored line is the median for the fields of one scale, for scales 0 to 3, the two coarsest scales having too few fields per bin to plot. The dotted line marks an elongation of 1, a circular field.

environment and least in the narrow rectangular environment. The relation was not an effect of field size, because the relation held within each of scales 0–4 in every environment (Table 5). In the circular enclosures and the wide rectangular environment, median elongation fell from about 1.8–9.0 among the fields nearest the wall to about 1.1–1.2 at the center or midline.

Differences between feature spaces.

The strength of the relation with wall distance differed between feature spaces (Table 6). The relation was strongest for depth and, among the visual descriptors, strongest for coarse layout and weakest for color. Landmarks weakened the relation for color and coarse layout, most clearly in the circular enclosures, but not for depth. The absence of a landmark effect for depth is expected, because the range scan does not detect the landmark panels; the range scan does, however, measure distance to the walls directly (Section 3.2). Combining feature spaces gave correlations between the correlations of depth and the correlations of the visual descriptors. In the narrow rectangular environment, the relation was weaker for every feature space (-0.14 to -0.58).

Table 6 Spearman correlation of elongation with wall distance in each environment (Table 2) and feature-space configuration (Table 3). An asterisk marks a correlation that fails the Benjamini–Hochberg test at $q = 0.05$, that is, a correlation that is not significant after adjustment for multiple comparisons.

	Spearman correlation of elongation with wall distance					
Environment	`hog`	`color`	`downscaled`	`lidar`	`visual`	`all`
`circ_lm8_r3`	−0.62	−0.37	−0.75	−0.96	−0.60	−0.60
`circ_lm8_r6`	−0.82	−0.28	−0.83	−0.97	−0.76	−0.78
`corr_lm8_l10w10`	−0.67	−0.22	−0.71	−0.85	−0.58	−0.60
`corr_lm8_l10w2`	−0.50	−0.45	−0.58	−0.35	−0.43	−0.35
`circ_lm0_r3`	−0.36	−0.74	−0.95	−0.96	−0.71	−0.74
`circ_lm0_r6`	−0.77	−0.64	−0.92	−0.97	−0.88	−0.90
`corr_lm0_l10w10`	−0.72	−0.50	−0.78	−0.85	−0.70	−0.68
`corr_lm0_l10w2`	−0.20	−0.14	−0.17	−0.36	−0.22	−0.27

4.5 Multifield activity

A single cluster can be active in several separate places, as a multifield place cell is (Section 3.5). Such cells are common in large environments: most dorsal CA1 cells in the megaspace had several subfields (Harland et al. 2021), and cells in the 200 m tunnel had a mean of 4.9 fields per flight direction (Eliav et al. 2021). Across the 48 sets of selected fields, 637 place cells had two or more selected subfields. The 1,483 subfields of the multifield place cells made up 3.2% of all selected fields and 1.2% of all candidate fields, the orange and violet bars of Figure 12, which Table 7 lists set by set.

Multifield activity arose only from the visual feature spaces (269 place cells from color, 151 from edge orientation, 144 from coarse layout, 56 from the three visual descriptors combined, and 17 from all feature spaces combined) and never from depth. The dependence of multifield activity on appearance is consistent with the similar fields that a higher-than-chance fraction of CA1 cells have in two visually identical compartments (Skaggs and McNaughton 1998), although the multifield analysis identifies where multifield activity occurs, not the cause of multifield activity. Multifield activity was most frequent in the narrow rectangular environment, where 8.6% of selected fields belonged to multifield place cells, less frequent in the wide rectangular environment (2.8%), and rare in the circular enclosures (1.2%), and 28 of the 48 sets contained none. By environment, the narrow rectangular environment with landmarks had 324 multifield place cells, the wide rectangular environment without landmarks 88, the wide rectangular environment with landmarks 71, the large circular enclosures 66 with landmarks and 43 without, the narrow rectangular environment without landmarks 38, the small circular enclosure with landmarks 7, and the small circular enclosure without landmarks none. Figure 13 shows every multifield place cell of the narrow rectangular environment with landmarks, the environment with the most multifield place cells. The share of selected fields in multifield place cells also increased with scale, from 1.5% at the finest scale to 9.6% at scale 3 and 9.1% at scale 4, and fell to 2.5% at the coarsest.

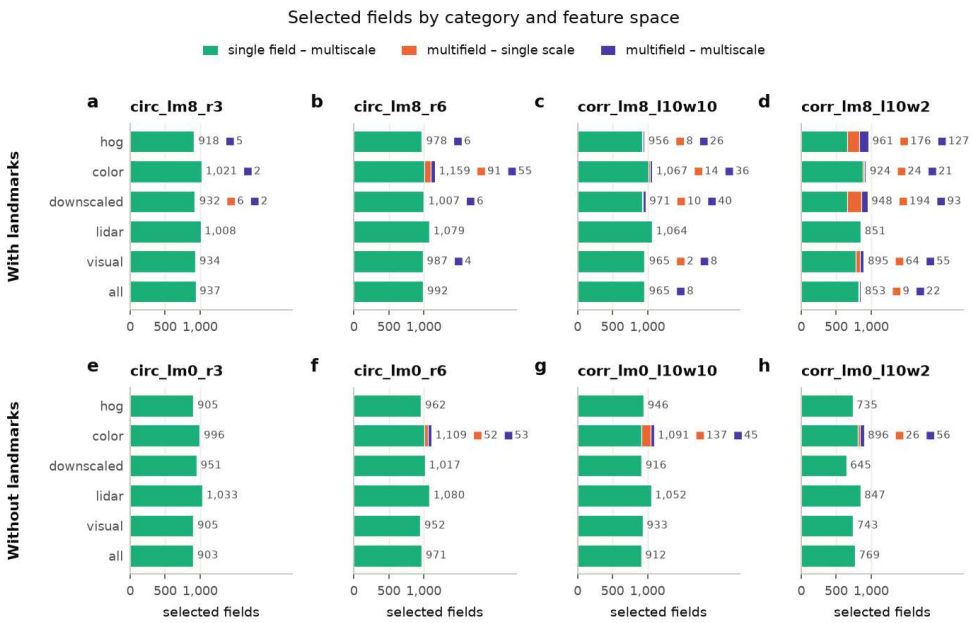


Fig. 12 Selected fields of each kind in each set, with one panel for each environment (Table 2); panels (a–d) show the environments with landmarks and panels (e–h) the environments without landmarks. Each bar is one feature-space configuration and stacks the fields of single-field place cells (green), the subfields of multifield place cells whose subfields all lie at one scale (orange), and the subfields of multifield place cells whose subfields lie at two or more scales (violet). The number at the end of each bar is the number of selected fields in the set, followed, where the set has any, by the numbers of orange and violet fields.

The subfields of a place cell could lie at different scales: 276 of the 637 multifield place cells had subfields at two or more scales, including 125 of the 324 in the narrow rectangular environment with landmarks (Figure 13b). In that environment under edge orientation, 131 place cells had multifield activity, 93 with two subfields, 35 with three, and 3 with four, and the subfields of a place cell recurred along the length of the environment at intervals close to the spacing of the landmarks (median 2.95 m, interquartile range 2.91 m to 2.99 m, against a landmark spacing of 3.00 m). In that set, the largest subfield of a place cell was on average 2.5 times the smallest subfield. Over all multifield place cells the mean of the largest-to-smallest ratio was 3.6, but the median of the ratio was 1.4: most place cells had subfields of similar size, and a few had subfields that differed greatly.

The properties of multifield activity follow from how place-cell activity is defined. A place cell's activity depends only on how closely the observation at a position resembles the centroid of the place cell's cluster (Equation (3)), so a place cell has several subfields wherever separate places look alike. Places that look alike are common where appearance repeats, as along the narrow rectangular environment with its regularly spaced landmarks, whereas depth, the feature space that reflects the geometry around

Table 7 Selected fields of each kind in each set, the numbers behind the bars of Figure 12. Single field counts the fields of single-field place cells; multifield, single scale the subfields of multifield place cells whose subfields all lie at one scale; and multifield, multiscale the subfields of multifield place cells whose subfields lie at two or more scales. The three field-type columns sum to the total selected fields of the set. A dash marks none. The 28 sets with no multifield place cells are omitted: every `lidar` set, every set of `circ_lm0_r3`, every set but `color` of the other three environments without landmarks, `visual` and `all` of `circ_lm8_r3`, and `all` of `circ_lm8_r6`.

		Field type			Total
Environment	Feature space	Single field	Multifield, single scale	Multifield, multiscale	selected fields
`circ_lm8_r3`	`hog`	913	–	5	918
	`color`	1,019	–	2	1,021
	`downscaled`	924	6	2	932
`circ_lm8_r6`	`hog`	972	–	6	978
	`color`	1,013	91	55	1,159
	`downscaled`	1,001	–	6	1,007
	`visual`	983	–	4	987
`corr_lm8_l10w10`	`hog`	922	8	26	956
	`color`	1,017	14	36	1,067
	`downscaled`	921	10	40	971
	`visual`	955	2	8	965
	`all`	957	–	8	965
`corr_lm8_l10w2`	`hog`	658	176	127	961
	`color`	879	24	21	924
	`downscaled`	661	194	93	948
	`visual`	776	64	55	895
	`all`	822	9	22	853
`circ_lm0_r6`	`color`	1,004	52	53	1,109
`corr_lm0_l10w10`	`color`	909	137	45	1,091
`corr_lm0_l10w2`	`color`	814	26	56	896

each position, produced no multifield activity: in the model, multifield activity arose only from appearance.

The model's multifield activity differs from the recordings in two respects. First, most of the model's multifield place cells had subfields of similar size, whereas the fields of one neuron in the 200 m tunnel differed in size by a mean factor of 4.4 (Eliav et al. 2021). The mean of the model's largest-to-smallest ratio, 3.6, approached the 4.4 of the tunnel only because a few place cells had subfields that differed greatly, although 276 of the 637 multifield place cells had subfields at different scales. In most place cells, then, the range of scales lies across place cells more than within a place cell. Second, multifield activity was most frequent in the narrow rectangular environment, the smallest environment, so across the environments studied here the frequency of

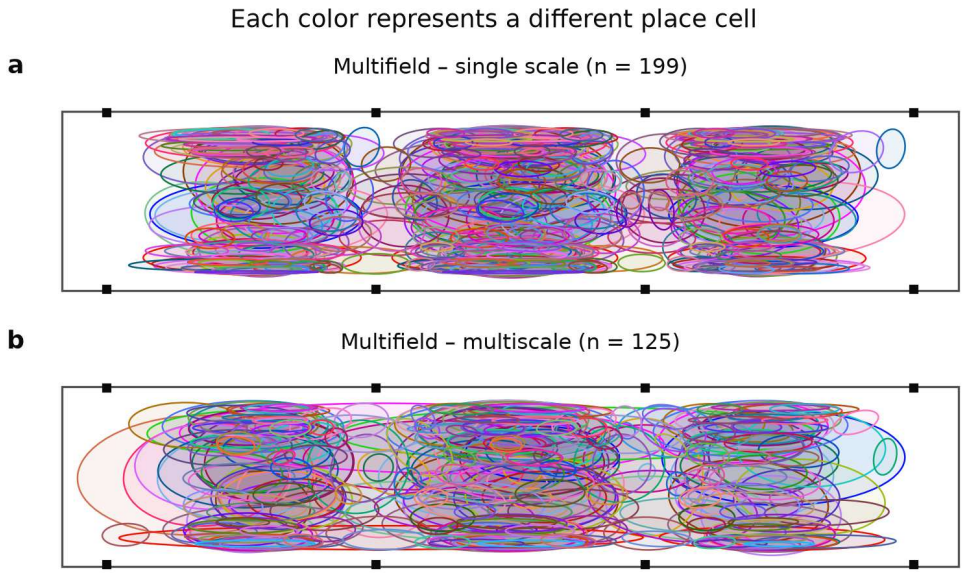


Fig. 13 Multifield place cells in `corr_lm8_l10w2` (Table 2), the environment with the most multifield place cells, from all feature-space configurations. Each color represents a different place cell, and each ellipse of that color is one of the cell's place fields, drawn as the field's fitted ellipse: an individual place cell may generate multiple place fields, one in each place whose views resemble the cell's centroid. **(a)** Multifield place cells whose subfields all lie at one scale. **(b)** Multifield place cells whose subfields lie at two or more scales. Each panel shows the environment from above with the environment's landmarks (black squares), and n is the number of place cells in the panel.

multifield activity did not increase with area as the number of subfields per cell does in the megaspace recordings (Harland et al. 2021).

The multifield activity of the present model is limited. Only 3.2% of selected fields belonged to multifield place cells, and all multifield place cells came from the visual feature spaces. The limited multifield activity follows from the clustering step, which groups positions by the similarity of the observations alone, so that a place cell has several subfields only where separate places look alike.

5 Conclusion and discussion

This paper asked whether the structure seen in the place code is available in what an agent perceives. An agent records a panoramic set of camera views and a range scan at evenly spaced positions, agglomerative clustering groups those observations by similarity, and each cluster is projected back into the environment as a place field. No position, odometry, path integration or reward enters at any stage, and no field size, count or location is set beforehand.

Across eight environments and six feature-space configurations, the model produced fields with three attributes that Eliav et al. (2021) report for bats and Harland et al. (2021) for rats: a range of field sizes, multifield activity in which a single cell is active in several separate places, and a change in field shape near a boundary.

Small fields greatly outnumbered large ones, as they do in the 200 m tunnel of Eliav et al. (2021) and the megaspace of Harland et al. (2021). At a typical location in the covered area, about seven place fields were active when all scales were counted together, and only about one and a half within any single scale (medians of 7.4 and 1.53; Section 4.3). The overlap therefore comes from fields of different sizes covering the same place rather than from fields of one size piling up, which is the overlapping multiscale ensemble that Harland et al. (2021) describe in the megaspace. Across the 48 sets, 637 place cells were active in more than one region, the model's counterpart of the multifield cells both studies record. Fields were also more elongated close to a wall, where Muller et al. (1987) found that recorded fields can conform to the boundary, taking crescent or linear forms along it.

Coverage of this kind is what multiscale navigation models exploit, with coarse fields carrying an agent across open space and fine fields resolving position where precision is needed (Scleidorovich et al. 2020, 2022). A population that already has it is a candidate state representation for a learning agent. We discuss this further in Section 5.3.

The feature-space configurations show what each kind of sensory information contributes. Depth alone gave the most uniform coverage and the strongest relation between elongation and wall distance, as expected of a signal that measures distance to the walls directly, and depth never produced multifield activity. The visual descriptors yielded similar size distributions but depended more on the environment. In the narrow rectangular environment, landmarks increased the number of fields obtained from the visual feature spaces by up to 1.5-fold while leaving the depth sets nearly unchanged. Color related elongation to wall distance most weakly and was the feature space most affected by landmarks. The visual feature spaces were thus the ones affected by landmarks, and the only ones that produced multifield activity.

The model's fields differed from the recorded ones in two respects. Relative to the environment they were smaller. The median selected field covered 0.25% of its environment, whereas 78% of the subfields Harland et al. (2021) measured in the megaspace were at or below $1\,\mathrm{m}^2$, which is 5.4% of that environment. Because the size bounds scale with environment area, the relative distribution of field sizes was nevertheless nearly the same in the small and large circular enclosures.

The second difference concerns how often multifield activity occurred. In the model, 3.2% of selected fields belonged to a cell active in more than one region. In the megaspace, Harland et al. (2021) found that 82% of cells had between two and five subfields. The dependence on environment size also ran the other way. Multifield activity in the model was most frequent in the narrow rectangular environment, the smallest of the eight, and did not increase with area, whereas Harland et al. (2021) found the number of subfields per cell growing with the area of the environment.

The goal of the current paper was to describe the VPCE model showing its capability to produce place cell multiscale, multifield and spatial elongations. While many of the results described in the paper were contrasted and compared to the experimental studies of Harland et al. (2021) and Eliav et al. (2021), there are important differences between the different studies, including whether the visual and range perceptions described in the paper are actually available and exploited by either rats or

bats in similar or different form, such as echolocation in bats as opposed to visual perception. These considerations require further investigation.

The rest of this section discusses some of the model limitations and future work planned including: (1) the model does not incorporate self-motion information, (2) the multifield activity is based exclusively on clustering from visual and range similarity without consideration for other possibilities, such as place cell remapping, and (3) the model has not yet been used in a robot navigation task.

5.1 Self-motion and temporal structure

The model deliberately receives no self-motion signal. In the animal, locomotion generates self-motion signals continuously, and they reach the hippocampal formation alongside vision. Place fields can persist in darkness (Quirk et al. 1990), so vision is not required for localized firing. When rats drove a car around a track instead of walking, which removes the self-motion signals that walking generates, Terrazas et al. (2005) found that fewer CA1 cells expressed place fields and that the fields that were expressed were substantially larger, with the spatial scale of the code appearing to follow the gain of a self-motion velocity signal. McNaughton et al. (2006) likewise attributed the increase in field size along the dorsoventral axis to variation in the gain of a movement-speed signal. Locomotion therefore contributes not only to whether a cell has a field, but to how large that field is. In virtual reality, Chen et al. (2013) found that visual information alone was sufficient for localized firing in 25% of place cells, whereas the remainder also required movement-related information, and Ravassard et al. (2013) found a smaller share of cells active on a virtual track than on a real one. The present model is based exclusively on perceptual input corresponding to vision and range only. Robotic models of spatial cognition, by contrast, have typically combined place cells with path integration (Arleo and Gerstner 2000; Barrera and Weitzenfeld 2008; Tejera et al. 2018).

Additionally, we plan to extend the model with the temporal structure that locomotion provides, but without knowledge of any coordinates. Sampling at evenly spaced positions will be replaced by an exploration phase in which the agent moves throughout the environment randomly, so that observations arrive as a trajectory. A temporal term will then be added to the Ward criterion (Equation (2)), so that observations experienced close together in time are less costly to merge. With the temporal term, positions that the agent visits in succession are more likely to be grouped into the same cluster. Positions that look alike but lie far apart, which the present model can place in the same cluster (Section 4.5), may then be separated. Robotic place recognition uses temporal order in a related way, matching sequences of views rather than single views (Milford and Wyeth 2012). The extension also links the model to accounts in which place fields reflect the transition structure of experience (Stachenfeld et al. 2017). By varying the weight of the temporal term relative to sensory similarity, we can ask whether field scale depends on that term as place-field size depends on the gain of self-motion signals in the animal.

What makes a place cell active in several places rather than one has not been clearly established. Eliav et al. (2021) and Harland et al. (2021) describe where a cell is active and how its subfields are arranged, but neither identifies the signal that

binds those subfields to one cell. If the order in which an agent encounters places carries information about which fields belong together, a model given that order should produce multifield cells closer to those the two studies record than the present model does. It is an open question whether a self-motion signal can account for multifield activity.

5.2 Multifield place cells

Eliav et al. (2021) and Harland et al. (2021) both found cells with several fields in large environments. The VPCE model produced such cells, but at a far lower rate and without the dependence on environment size that these authors reported.

The experimental literature points to more than one source of multiple fields. Geometry is one. When a rectangular box is lengthened along one axis, O'Keefe and Burgess (1996) found that some fields develop a second peak. Repeated appearance is another. Skaggs and McNaughton (1998) found that a higher-than-chance fraction of CA1 cells have similar fields in two visually identical compartments, although other cells distinguish the two, a difference those authors attribute to learned expectations and a memory of movement. The VPCE model's multifield activity exploits the similarity in perception. It arises wherever separate places produce similar observations, which is why it appeared in the visual feature spaces but not under depth alone (Section 4.5). The temporal term of Section 5.1 would supply a counterpart to the movement-related signals that Skaggs and McNaughton (1998) invoke. By varying its weight it would show how much of the model's multifield activity survives when places that look alike are visited at different times.

A further mechanism is needed to assign several fields to one place cell. We plan to supply one with a mixture step that combines individual fields into a single place cell, as a linear combination of several radial basis functions. Two measurements would constrain the mixture: the number of subfields a cell has, which Harland et al. (2021) found to be between two and five for 82% of cells in the megaspace, and the spread of subfield sizes within a cell, for which Eliav et al. (2021) report a mean ratio of 4.4 between a cell's largest and smallest field in the tunnel. We also plan to evaluate the decoding from field activation to agent location. Furthermore, we would like to explore whether the mixture step could reduce the number of place cells required while keeping the decode robust.

One model of this kind already exists. In the megamap of Hedrick and Zhang (2016), each cell is given many irregularly placed fields so that a fixed population can represent a large space, but those fields are of similar size and their number is drawn from a distribution chosen by the modeler. The mixture step proposed here would instead draw on fields the model has already produced, which differ in size, and would take the number of subfields per cell and the spread of their sizes from the two recording studies.

Remapping is a further possible source. Moving between environments, the hippocampus assigns a largely independent firing pattern to each, and Sanders et al. (2020) argue that the reassignment follows the animal's belief about which hidden state it occupies rather than any objective property of the surroundings. The division can fall within a single space. In a four-compartment environment, Spiers et al. (2015)

found place fields repeating across compartments, a cell firing at the same relative position in several of them, while altering one compartment left firing in the others unchanged. A cell with one field per compartment is a multifield cell, produced by the space being parsed into repeated local contexts rather than by a single map holding several fields. The VPCE model has no notion of context: every observation from an environment enters one hierarchy, and a cluster spanning two distant regions becomes one cell active in both. We plan to ask whether multifield activity also arises when an environment is divided into contexts.

In the robotic models that motivated this work, fields were predefined and activity was computed from the robot's position (Scleidorovich et al. 2020, 2022), so each cell was tied to one region. A multifield cell is active at several places, so a single cell cannot say which of the several places the agent occupies. In robotics this is perceptual aliasing, and it is addressed by adding information beyond a single view, whether by matching sequences of views (Milford and Wyeth 2012) or by combining appearance with odometry (Milford et al. 2004).

A population resolves the same ambiguity without adding information. No single multifield cell specifies a location, but the set of cells active at a location does, because different cells are ambiguous in different ways and only one location is consistent with all of them. This is the ensemble code Harland et al. (2021) describe, and it is why a multiscale population can be decoded with far fewer neurons than in a single-scale population (Section 2).

For a reinforcement-learning agent, multifield cells could therefore cut either way: fewer cells for the same spatial resolution, or confusion between places that require different actions. The navigation experiments of Section 5.3 include and remove multifield cells in turn, so that the two effects can be distinguished.

5.3 Robot spatial navigation and learning

Place fields support spatial navigation in the animal, which makes a navigation task a natural setting in which to measure what a population of fields is good for. The functional value of a multiscale population has been measured that way in robotic models (Section 2), where the fields were predefined and place-cell activity was computed from the robot's position (Scleidorovich et al. 2020, 2022). In the present model the activity of a place cell is a function of the current observation (Equation (3)), so it can be computed from what the agent senses without an estimate of position. The sets of selected fields could therefore serve as a state representation for an agent equipped with cameras and a range sensor but no global localization system.

We plan to use a navigation task as the setting in which to evaluate the model representation. The sets of selected fields would form the state representation of a reinforcement-learning agent, and we would measure how the structure of that representation affects learning: the time taken to learn a route, the number of place cells required, the optimality of the paths found, and the robustness of learning to noise in the observations. The contribution of each part of the structure can be isolated by including or removing it in turn: the range of scales, the multifield cells of Section 5.2, and the temporal term of Section 5.1.

In the successor representation, a place field reflects where the animal is likely to go next, so the code depends on the animal's higher-level behavior (Stachenfeld et al. 2017). Danieli and Lepperød (2026) give behavior a more direct role, with signals tied to reward and to boundaries recruiting cells and shrinking their fields at important locations. The fields described here are produced from observations alone. Placing them in a learning agent would show whether behavior such as planning affects learning.

In robotics, appearance-based mapping systems pair visual input with odometry (Milford et al. 2004), and systems that select spatial scales from images do so in order to recognize places already visited (Fan et al. 2017). The representation described here differs in what it is asked to produce: a population of fields covering a two-dimensional environment, whose number and sizes follow from the observations without a target count or an objective to optimize, and whose geometry is compared against the place fields Eliav et al. (2021) and Harland et al. (2021) measured. How much the structure of such a population matters to an agent that has to learn is what the planned experiments would establish.

Declarations

Funding. The authors did not receive support from any organization for the submitted work.

Conflict of interest. The authors have no competing interests to declare that are relevant to the content of this article.

Ethics approval and consent to participate. Not applicable. This study reports simulation results and involved no human participants and no animals.

Consent for publication. Not applicable.

Data availability. The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

Materials availability. Not applicable.

Code availability. The implementation of the model is available at https://github.com/CJHRobotics/REALM-VPCE.

Author contribution. Both authors contributed to the study conception and design. The model, the simulations and the analysis were carried out by Chance J. Hamilton. The first draft of the manuscript was written by Chance J. Hamilton, and both authors commented on later versions. Both authors read and approved the final manuscript.

Appendix A Symbols and terms

Table A1 defines the symbols and terms used in the paper, with the section or equation in which each symbol or term is introduced.

Table A1: Symbols and terms.

Symbol or term	Meaning	Defined in
Sampling and observations		

Symbol or term	Meaning	Defined in
Landmark	A distal cue: one lying beyond the area the animal moves through. In the model, a textured panel on the wall of an environment	Section 1
Proximal cue	A cue lying on the surface the animal moves over	Section 1
Wall buffer	The 0.2 m band along the walls in which no position is sampled	Section 3.1
N	Number of sampled positions in an environment	Section 3.1
$\mathbf{p}$, $\mathbf{p}_i$	A sampled position on the floor; the ith sampled position	Section 3.3
Feature space	One description of an observation: `hog`, `color`, `downscaled`, or `lidar` (Table 3)	Section 3.2
Feature-space configuration	A feature space, or a combination of feature spaces, on which the model is run (Table 3)	Section 3.2
$\mathbf{f}_i(k)$	Vector of feature space k at position i	Equation (1)
m_k	Root-mean-square magnitude of feature space k over all positions, by which the vectors of feature space k are divided before feature spaces are combined	Equation (1)
$\mathcal{K}$	The set of feature spaces a configuration uses	Equation (1)
$\mathbf{o}_i$, $\mathbf{o}(\mathbf{p})$	Observation vector at position i; at position $\mathbf{p}$	Equation (1)
D	Number of values in an observation vector	Section 3.3
Clustering		
$\mathcal{C}$, $\lvert\mathcal{C}\rvert$	A cluster: a set of observation vectors assigned to one group by agglomerative clustering, whose positions are its members; the number of members	Section 3.3
$\mathcal{C}_i$	The ith cluster: $\mathcal{C}_i = \{\mathbf{o}_i\}$ for $i \leq N$, and for $i > N$ the cluster formed by the $(i - N)$th merge	Section 3.3
$\boldsymbol{\mu}_\mathcal{C}$	Centroid of a cluster: the mean observation vector of the cluster's members	Section 3.3
Δ_{ab}	Increase in the total within-cluster sum of squares when clusters $\mathcal{C}_a$ and $\mathcal{C}_b$ are merged (Ward's criterion)	Equation (2)
Agglomerative hierarchy	The binary tree of $2N - 1$ clusters produced by merging, from single positions to the whole environment	Section 3.3
Candidate field	A field projected from a retained cluster of the hierarchy, one for each of the cluster's subfields	Sections 3.3 and 3.6
Place fields		
$a_\mathcal{C}(\mathbf{p})$	Place-cell activity of cluster $\mathcal{C}$ at position $\mathbf{p}$, a radial basis function of the distance between $\mathbf{o}(\mathbf{p})$ and $\boldsymbol{\mu}_\mathcal{C}$	Equation (3)
$\sigma_\mathcal{C}$	Width of the place-cell activity, set so that the field includes 80% of the cluster's members	Section 3.4
T	Field threshold: half of a cluster's peak activity over all sampled observations; a bin belongs to the field when its activity is at least T	Equation (4)
Activity map	Place-cell activity on a grid of square spatial bins covering the environment	Section 3.5

Symbol or term	Meaning	Defined in
$M_{\mathcal{C}}$	Field mask: the bins whose activity is at least T	Section 3.5
Field geometry		
A	Area of a field: the number of bins in its mask times the area of one bin	Equation (5)
r_{eq}	Equivalent radius: the radius of a circle of area A	Section 3.5
(c_x, c_y)	Center of a field	Section 3.5
θ	Orientation of a field	Section 3.5
Fitted ellipse	The best-fit ellipse from a field's second-order central moments, scaled to the field's area	Section 3.5
a, b	Major and minor semi-axes of a field's fitted ellipse	Section 3.5
Elongation	The ratio a/b	Section 3.5
Subfield	An 8-connected region of a field mask at least A_{min} in area, treated as an individual field	Section 3.5
Place cell (of the model)	The place cell defined by a cluster, with place-cell activity $a_{\mathcal{C}}$	Section 3.4
Multiscale	A set of fields spanning a range of sizes; used both of a population and of the fields of one cell	Section 1
Multifield cell	A cell with more than one field, as reported in the recordings	Section 1
Multifield activity	In the model, activity of a place cell with two or more selected subfields	Section 3.5
Field selection		
Field selection	Selection of candidate fields by the size criterion	Section 3.6
Selected field	A candidate field that passes the size criterion	Section 3.6
Set of selected fields	The fields selected in one environment under one feature-space configuration	Section 4
A_{env}	Area of the environment	Equation (6)
A_{min}, A_{max}	Lower and upper size bounds	Equation (6)
r_{min}	Equivalent radius of the lower size bound	Equation (7)
β	Scale of a field, from 0 (finest) to 5 (coarsest)	Equation (7)
ρ	Ratio between the radii of neighboring scales (1.6)	Equation (7)
Analysis		
R	Redundancy: the summed area of a group of fields divided by the area the fields cover	Equation (8)
K; F_k, A_k; A_{cov}	Number of fields; region and area of field k; area covered by at least one field	Equation (8)
ASR, WSR	Across-scale redundancy, over all fields of a set; within-scale redundancy, the median over scales of the redundancy of each scale	Section 4.3
Wall distance	Distance from a field center to the nearest boundary, normalized within each environment from 0 (closest to the boundary) to 1 (center or midline)	Section 4.4
q	Benjamini–Hochberg q value; a correlation is significant when $q < 0.05$	Section 4.4

References

Arleo A, Gerstner W (2000) Spatial cognition and neuro-mimetic navigation: a model of hippocampal place cell activity. Biological Cybernetics 83(3):287–299. https://doi.org/10.1007/s004220000171

Banino A, Barry C, Uria B, et al (2018) Vector-based navigation using grid-like representations in artificial agents. Nature 557(7705):429–433. https://doi.org/10.1038/s41586-018-0102-6

Barrera A, Weitzenfeld A (2008) Biologically-inspired robot spatial cognition based on rat neurophysiological studies. Autonomous Robots 25(1–2):147–169. https://doi.org/10.1007/s10514-007-9074-3

Barry C, Lever C, Hayman R, et al (2006) The boundary vector cell model of place cell firing and spatial memory. Reviews in the Neurosciences 17(1-2):71–97. https://doi.org/10.1515/REVNEURO.2006.17.1-2.71

Burgess N, O'Keefe J (1996) Neuronal computations underlying the firing of place cells and their role in navigation. Hippocampus 6(6):749–762. https://doi.org/10.1002/(SICI)1098-1063(1996)6:6⟨749::AID-HIPO16⟩3.0.CO;2-0

Burgess N, Jackson A, Hartley T, et al (2000) Predictions derived from modelling the hippocampal role in navigation. Biological Cybernetics 83(3):301–312. https://doi.org/10.1007/s004220000172

Chen G, King JA, Burgess N, et al (2013) How vision and movement combine in the hippocampal place code. Proceedings of the National Academy of Sciences 110(1):378–383. https://doi.org/10.1073/pnas.1215834110

Cohen Z, Drugowitsch J (2025) A unifying theory of receptive field heterogeneity predicts hippocampal spatial tuning. bioRxiv https://doi.org/10.1101/2025.07.26.666958, preprint

Cueva CJ, Wei XX (2018) Emergence of grid-like representations by training recurrent neural networks to perform spatial localization. In: International Conference on Learning Representations (ICLR)

Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), IEEE, pp 886–893, https://doi.org/10.1109/CVPR.2005.177

Danieli K, Lepperød ME (2026) Flexible navigation with neuromodulated cognitive maps. PLOS Computational Biology 22(7):e1013487. https://doi.org/10.1371/journal.pcbi.1013487

Edvardsen V, Bicanski A, Burgess N (2020) Navigating with grid and place cells in cluttered environments. Hippocampus 30(3):220–232. https://doi.org/10.1002/hipo.23147

Eliav T, Maimon SR, Aljadeff J, et al (2021) Multiscale representation of very large environments in the hippocampus of flying bats. Science 372(6545):eabg4020. https://doi.org/10.1126/science.abg4020

Erdem UM, Hasselmo M (2012) A goal-directed spatial navigation model using forward trajectory planning based on grid cells. European Journal of Neuroscience 35(6):916–931. https://doi.org/10.1111/j.1460-9568.2012.08015.x

Erdem UM, Hasselmo ME (2014) A biologically inspired hierarchical goal directed navigation model. Journal of Physiology-Paris 108(1):28–37

Fan C, Chen Z, Jacobson A, et al (2017) Biologically-inspired visual place recognition with adaptive multiple scales. Robotics and Autonomous Systems 96:224–237. https://doi.org/10.1016/j.robot.2017.07.015

Fenton AA, Kao HY, Neymotin SA, et al (2008) Unmasking the CA1 ensemble place code by exposures to small and large environments: more place cells and multiple, irregularly arranged, and expanded place fields in the larger space. Journal of Neuroscience 28(44):11250–11262. https://doi.org/10.1523/JNEUROSCI.2862-08.2008

Harland B, Contreras M, Souder M, et al (2021) Dorsal CA1 hippocampal place cells form a multi-scale representation of megaspace. Current Biology 31(10):2178–2190.e6. https://doi.org/10.1016/j.cub.2021.03.003

Hartley T, Burgess N, Lever C, et al (2000) Modeling place fields in terms of the cortical inputs to the hippocampus. Hippocampus 10(4):369–379. https://doi.org/10.1002/1098-1063(2000)10:4⟨369::AID-HIPO3⟩3.0.CO;2-0

Hausler S, Chen Z, Hasselmo ME, et al (2020) Bio-inspired multi-scale fusion. Biological Cybernetics 114(2):209–229. https://doi.org/10.1007/s00422-020-00831-z

Hedrick KR, Zhang K (2016) Megamap: flexible representation of a large space embedded with nonspatial information by a hippocampal attractor network. Journal of Neurophysiology 116(2):868–891. https://doi.org/10.1152/jn.00856.2015

Jung MW, Wiener SI, McNaughton BL (1994) Comparison of spatial firing characteristics of units in dorsal and ventral hippocampus of the rat. The Journal of Neuroscience 14(12):7347–7356. https://doi.org/10.1523/JNEUROSCI.14-12-07347.1994

Kjelstrup KB, Solstad T, Brun VH, et al (2008) Finite scale of spatial representation in the hippocampus. Science 321(5885):140–143. https://doi.org/10.1126/science.1157086

Knierim JJ, Hamilton DA (2011) Framing spatial cognition: neural representations of proximal and distal frames of reference and their roles in navigation. Physiological Reviews 91(4):1245–1279. https://doi.org/10.1152/physrev.00021.2010

Lever C, Burton S, Jeewajee A, et al (2009) Boundary vector cells in the subiculum of the hippocampal formation. The Journal of Neuroscience 29(31):9771–9777. https://doi.org/10.1523/JNEUROSCI.1319-09.2009

Llofriu M, Tejera G, Contreras M, et al (2015) Goal-oriented robot navigation learning using a multi-scale space representation. Neural Networks 72:62–74. https://doi.org/10.1016/j.neunet.2015.09.006

Mathis A, Herz AVM, Stemmler M (2012) Optimal population codes for space: grid cells outperform place cells. Neural Computation 24(9):2280–2317. https://doi.org/10.1162/NECO_a_00319

McNaughton BL, Battaglia FP, Jensen O, et al (2006) Path integration and the neural basis of the 'cognitive map'. Nature Reviews Neuroscience 7(8):663–678. https://doi.org/10.1038/nrn1932

Michel O (2004) Cyberbotics Ltd. Webots™: Professional mobile robot simulation. International Journal of Advanced Robotic Systems 1(1):39–42. https://doi.org/10.5772/5618

Milford MJ, Wyeth GF (2012) SeqSLAM: visual route-based navigation for sunny summer days and stormy winter nights. In: IEEE International Conference on Robotics and Automation (ICRA), IEEE, pp 1643–1649, https://doi.org/10.1109/ICRA.2012.6224623

Milford MJ, Wyeth GF, Prasser D (2004) RatSLAM: a hippocampal model for simultaneous localization and mapping. In: IEEE International Conference on Robotics and Automation (ICRA), IEEE, pp 403–408, https://doi.org/10.1109/ROBOT.2004.1307183

Muller RU, Kubie JL (1987) The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. The Journal of Neuroscience 7(7):1951–1968. https://doi.org/10.1523/JNEUROSCI.07-07-01951.1987

Muller RU, Kubie JL, Ranck JBJr. (1987) Spatial firing patterns of hippocampal complex-spike cells in a fixed environment. The Journal of Neuroscience 7(7):1935–1950. https://doi.org/10.1523/JNEUROSCI.07-07-01935.1987

Murtagh F, Legendre P (2014) Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? Journal of Classification 31(3):274–295. https://doi.org/10.1007/s00357-014-9161-z

O'Keefe J, Burgess N (1996) Geometric determinants of the place fields of hippocampal neurons. Nature 381(6581):425–428. https://doi.org/10.1038/381425a0

Park E, Dvorak D, Fenton AA (2011) Ensemble place codes in hippocampus: CA1, CA3, and dentate gyrus place cells have multiple place fields in large environments. PLoS ONE 6(7):e22349. https://doi.org/10.1371/journal.pone.0022349

Quirk GJ, Muller RU, Kubie JL (1990) The firing of hippocampal place cells in the dark depends on the rat's recent experience. The Journal of Neuroscience 10(6):2008–2017. https://doi.org/10.1523/JNEUROSCI.10-06-02008.1990

Ravassard P, Kees A, Willers B, et al (2013) Multisensory control of hippocampal spatiotemporal selectivity. Science 340(6138):1342–1346. https://doi.org/10.1126/science.1232655

Rich PD, Liaw HP, Lee AK (2014) Large environments reveal the statistical structure governing hippocampal representations. Science 345(6198):814–817. https://doi.org/10.1126/science.1255635

Sanders H, Wilson MA, Gershman SJ (2020) Hippocampal remapping as hidden state inference. eLife 9:e51140. https://doi.org/10.7554/eLife.51140

Scaplen KM, Gulati AA, Heimer-McGinn VL, et al (2014) Objects and landmarks: hippocampal place cells respond differently to manipulations of visual cues depending on size, perspective, and experience. Hippocampus 24(11):1287–1299

Scleidorovich P, Llofriu M, Fellous JM, et al (2020) A computational model for spatial cognition combining dorsal and ventral hippocampal place field maps: multiscale navigation. Biological Cybernetics 114(2):187–207. https://doi.org/10.1007/s00422-019-00812-x

Scleidorovich P, Fellous JM, Weitzenfeld A (2022) Adapting hippocampus multi-scale place field distributions in cluttered environments optimizes spatial navigation and learning. Frontiers in Computational Neuroscience 16:1039822. https://doi.org/10.3389/fncom.2022.1039822

Skaggs WE, McNaughton BL (1998) Spatial firing properties of hippocampal CA1 populations in an environment containing two visually identical regions. The Journal of Neuroscience 18(20):8455–8466. https://doi.org/10.1523/JNEUROSCI.18-20-08455.1998

Solstad T, Boccara CN, Kropff E, et al (2008) Representation of geometric borders in the entorhinal cortex. Science 322(5909):1865–1868. https://doi.org/10.1126/science.1166466

Spiers HJ, Hayman RMA, Jovalekic A, et al (2015) Place field repetition and purely local remapping in a multicompartment environment. Cerebral Cortex 25(1):10–25.

https://doi.org/10.1093/cercor/bht198

Stachenfeld KL, Botvinick MM, Gershman SJ (2017) The hippocampus as a predictive map. Nature Neuroscience 20(11):1643–1653. https://doi.org/10.1038/nn.4650

Taube JS, Muller RU, Ranck JBJr. (1990) Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. The Journal of Neuroscience 10(2):420–435. https://doi.org/10.1523/JNEUROSCI.10-02-00420.1990

Teague MR (1980) Image analysis via the general theory of moments. Journal of the Optical Society of America 70(8):920–930. https://doi.org/10.1364/JOSA.70.000920

Tejera G, Llofriu M, Barrera A, et al (2018) Bio-inspired robotics: a spatial cognition model integrating place cells, grid cells and head direction cells. Journal of Intelligent & Robotic Systems 91(1):85–99. https://doi.org/10.1007/s10846-018-0852-2

Terrazas A, Krause M, Lipa P, et al (2005) Self-motion and the hippocampal spatial metric. The Journal of Neuroscience 25(35):8085–8096. https://doi.org/10.1523/JNEUROSCI.0693-05.2005

USF BioRobotics Lab (2026) REALM: Robotic Environment for Autonomous Learning and Mapping. https://github.com/CJHRobotics/REALM, software

Wang CH, Monaco JD, Knierim JJ (2020) Hippocampal place cells encode local surface-texture boundaries. Current Biology 30(8):1397–1409.e7. https://doi.org/10.1016/j.cub.2020.01.083

Ward JHJr. (1963) Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 58(301):236–244. https://doi.org/10.1080/01621459.1963.10500845

Wei XX, Prentice J, Balasubramanian V (2015) A principle of economy predicts the functional architecture of grid cells. eLife 4:e08362. https://doi.org/10.7554/eLife.08362