Supplementary Information
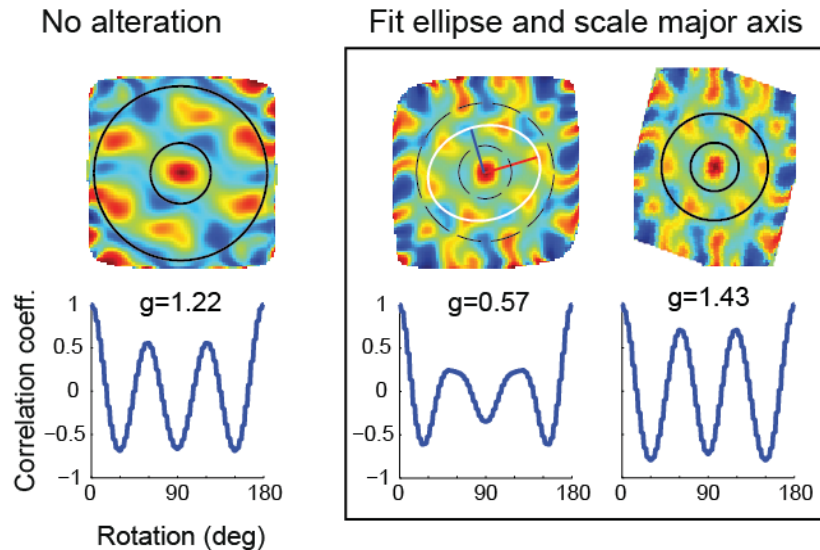

Supplementary Figure 1 | Gridness score calculation. Grid cells were classified using a gridness score[1,2]. An annulus was fit around the six peaks nearest to the center peak of the spatial autocorrelation and the rotational correlation was calculated at 30, 60, 90, 120, and 150 degrees. The minimum correlation coefficient from rotations of 60 and 120 degrees and the maximum correlation coefficient from rotations of 30, 90, and 150 degrees were identified, and the gridness score was equal to the minimum correlation coefficient of the set of values at 60 and 120 degrees minus the maximum correlation coefficient of the set of values at 30, 90, and 150 degrees. The inner radius of the annulus was determined as half of the mean of the distance to the six nearest peaks and the outer radius was varied from 2 bins (see Supplementary Table 1) past the inner radius to the outer edge of the autocorrelation, taking the best resulting score. The score was calculated without scaling and by an ellipse fit-and-scale method like that used in Brandon et al., 2011, to compute gridness for grids that have some ellipticity. As shown in the example at the right, ellipticity causes firing fields of outer rings to enter into the annulus, increasing the correlation at 90 degrees and thus decreasing the gridness score. The minor axis of an ellipse was solved for after specifying the far peak as the major axis. After determining the ellipse axes, the autocorrelation image was scaled along the major axis so that the ellipse became a circle. The higher of the two calculated gridness scores was used for each actual or surrogate autocorrelation image.
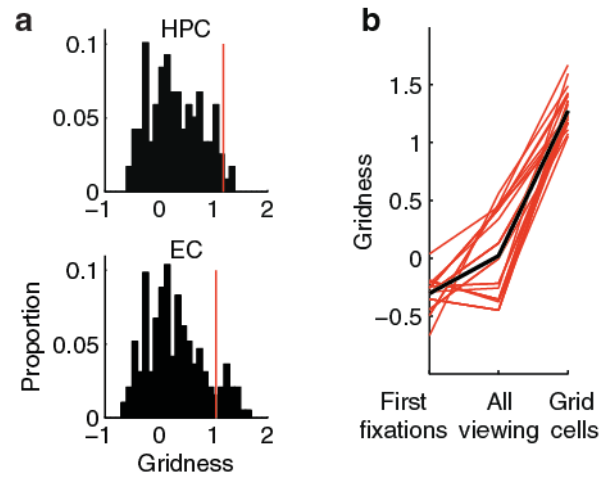
Supplementary Figure 2 | Gridness scores. a. Distribution of gridness scores for EC ($n = 193$) and HPC ($n = 119$) neurons, cells with significant gridness scores lay to the right of the red lines (the smallest of the significant grid scores for each group). The variances of EC and HPC gridness scores were not significantly different (Levene's test, $F(1,310) = 0.51$, $P = 0.48$). **b.** Shown are the gridness scores for the autocorrelations of initial fixation and aggregate viewing distributions for all grid cells ($n = 23$). For all grid cells, viewing patterns had lower gridness scores. The black line represents the average gridness scores.
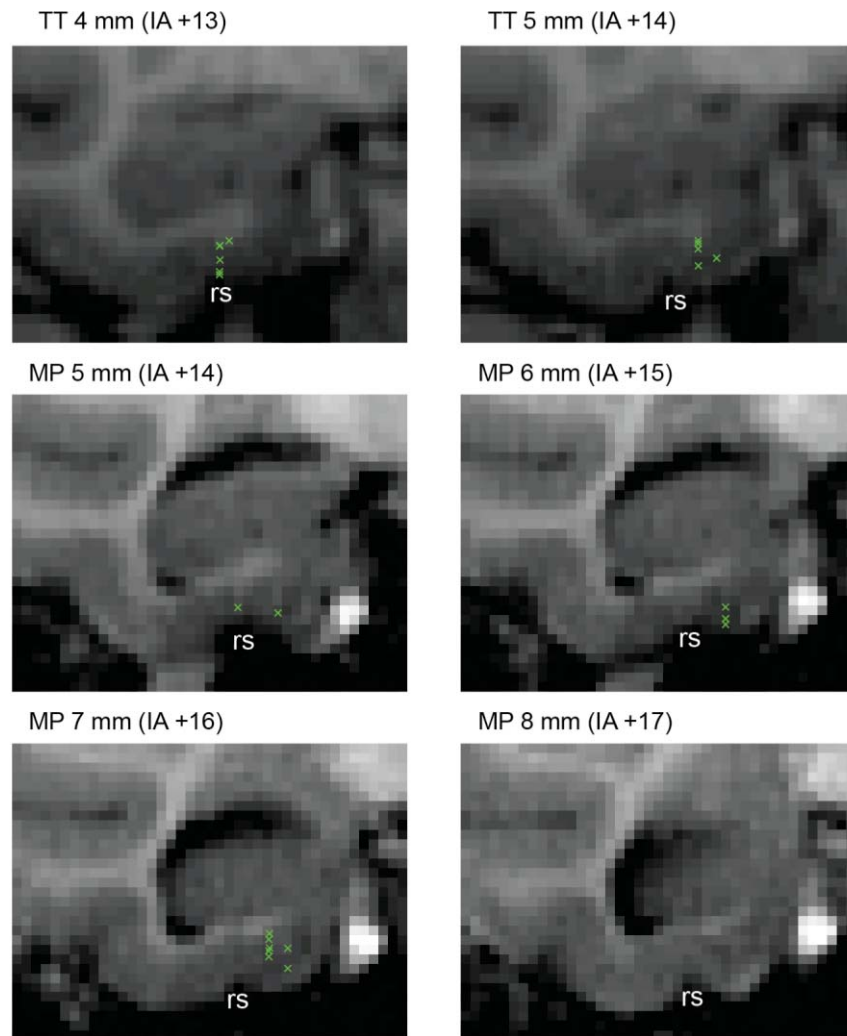
Supplementary Figure 3 | Locations of entorhinal grid cells. Green markers label the locations of all 23 grid cells found in two monkeys (MP and TT). The monkey name, location anterior to the posterior border of the entorhinal cortex, and interaural location[3] are given above each panel. rs: rhinal sulcus. 7 mm corresponded roughly to the midway point between the anterior and posterior borders of the EC. No grid cells were found anterior to this location.
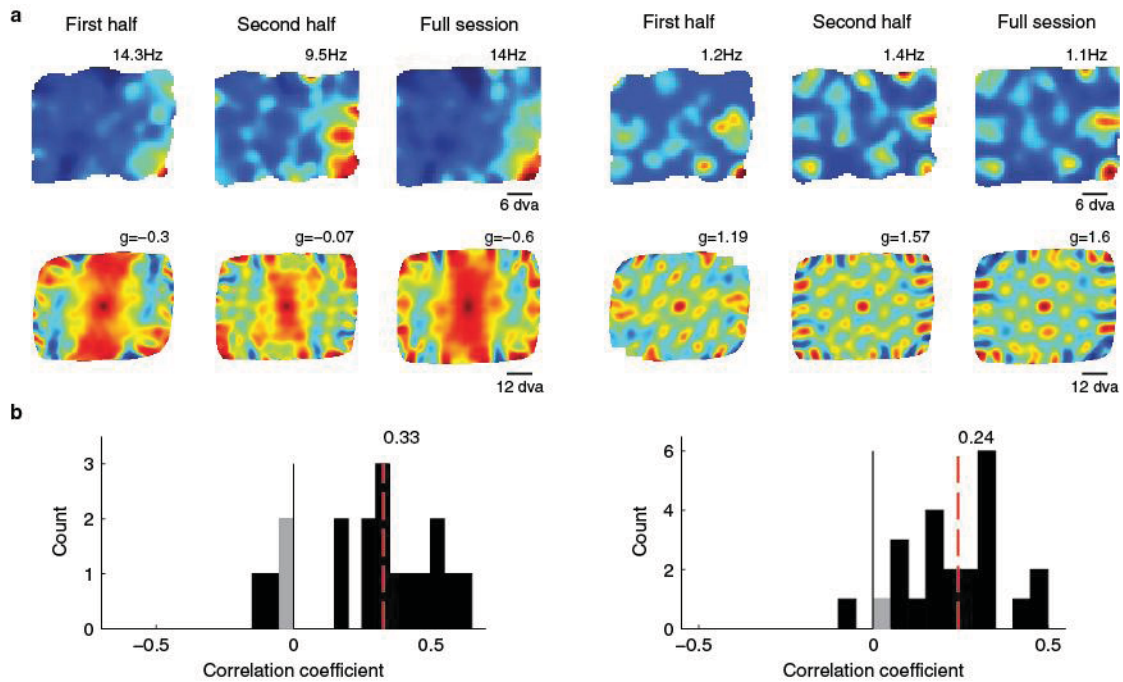
Supplementary Figure 4 | Stability of firing rate maps. a. Stability of spatial representation can be seen in the firing rate and autocorrelation plots using separate partitions of the data for an example cell with a significant border score, left (within-session correlation $\rho = 0.54$), and grid cell, right (within-session correlation $\rho = 0.19$). **b.** For the populations of units, firing rate maps were significantly correlated across the session. Correlation coefficients are shown in the histograms for cells in the EC with significant border scores (left) and significant gridness scores (right). The median correlation coefficients are represented by dashed red lines. Bars corresponding to cells with significant correlation coefficients ($P < 0.05$) are black. The population of cells with significant border scores had significant positive correlations as a population ($P < 0.01$, $n = 18$, signed rank test) and for 14 of 18 cells individually ($P < 0.05$). The population of grid cells also had significant positive correlations ($P < 1x10^{-5}$, $n = 23$, signed rank test) and for 21 of 23 grid cells individually ($P < 0.05$). See Supplementary Methods for more details on methods used.
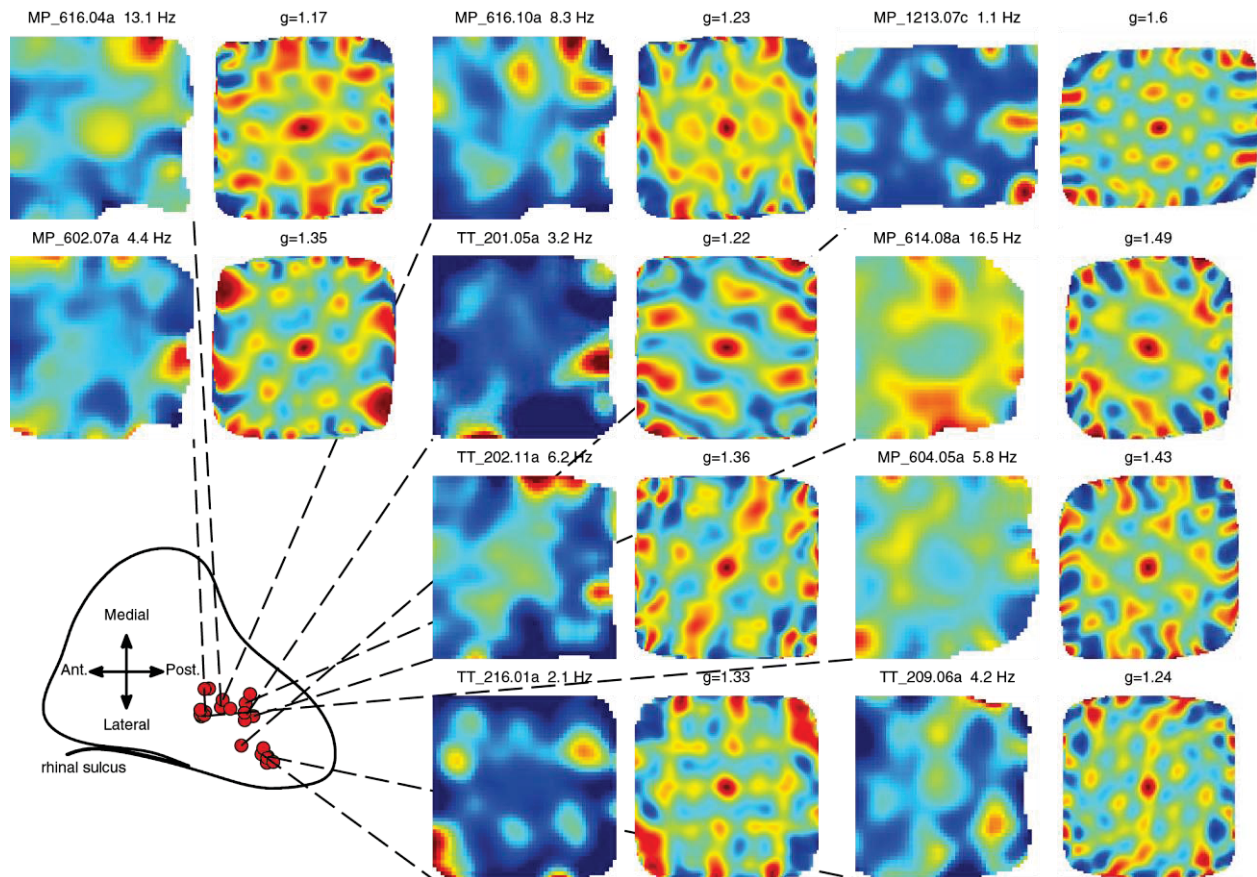
Supplementary Figure 5 | Grid cell spatial firing rate maps and autocorrelograms. Shown are the rate map, autocorrelogram, and approximate recording location placed on a prototype map of the EC[4] (see also Supplementary Fig. 3) for a selection of cells with significant gridness scores. Red circles indicate the recording locations of all grid cells.
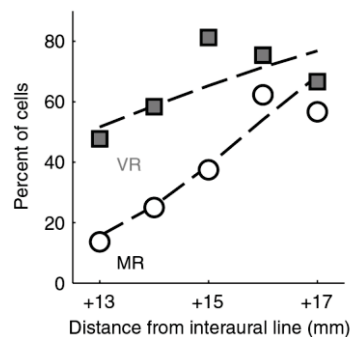
Supplementary Figure 6 | Anterior-posterior gradient of visual and memory responses. In recordings at more anterior locations within the EC, cells were more likely to be visually responsive (VR) and among those, more likely to show a recognition memory response (MR). Dashed lines are logistic regression curves.
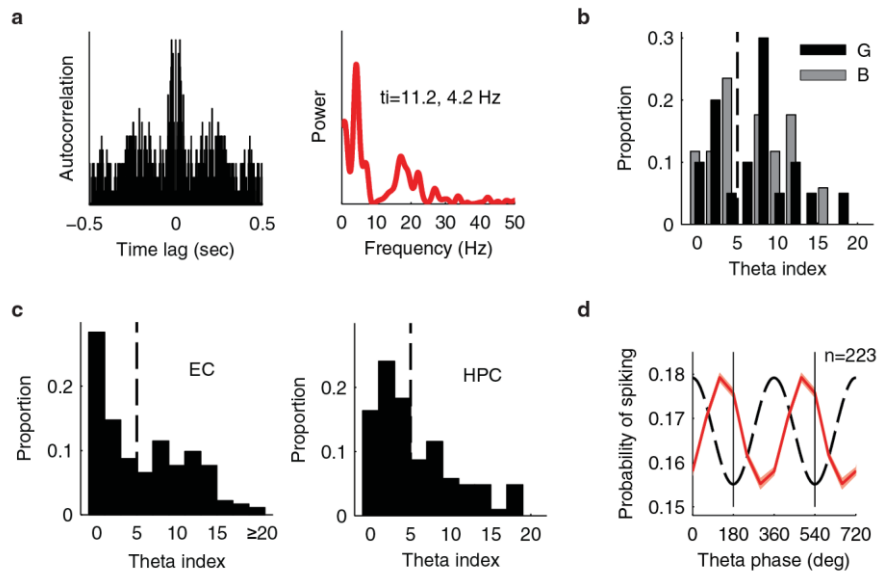
Supplementary Figure 7 | Theta-band modulation of spiking and theta phase-locking to the LFP. a. Autocorrelogram and power spectrum of the autocorrelation for the example grid cell shown in Figure 3c. The theta index computed from the power spectrum is 11.2 at a peak frequency of 4.2 Hz. **b.** Histogram of theta indices for EC grid cells (G, $n = 20/23$ cells with a calculable theta index) and cells with significant border scores (B, $n = 17/18$ cells with a calculable theta index). Both groups showed theta-modulation. **c.** Histograms of theta indices for all entorhinal cortex ($n = 183/223$ cells with a calculable theta index) and hippocampus ($n = 104/119$ cells with a calculable theta index) neurons. Neurons were considered theta-modulated if the theta index was above 5 (to the right of the dashed line). **d.** The population of EC neurons as a whole shows phase-locking near the trough of theta. Shown is the probability of spiking (mean ± s.e.m.) with respect to the theta-band LFP recorded at an electrode 150 or 300 microns away for all EC single units (cf. Figure 3b).
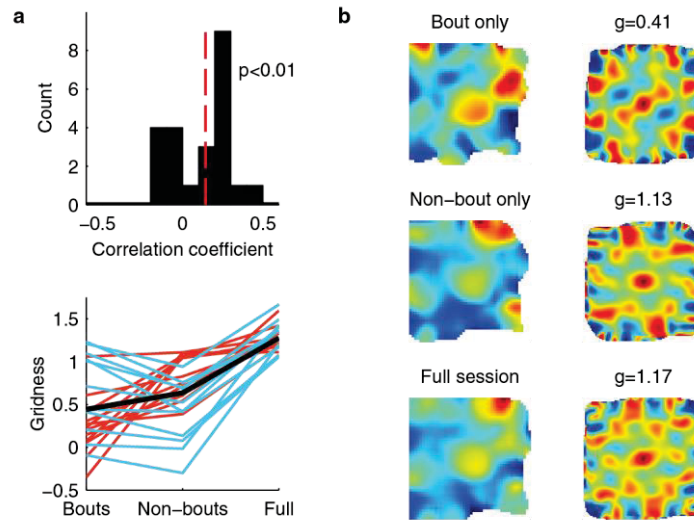
Supplementary Figure 8 | Evidence for grid-like representations in both the presence and absence of theta bouts. a. Top: Histogram of bout compared to non-bout rate map correlations (Pearson's r). Rate maps were significantly correlated between bout and non-bout periods ($P < 0.01$, $n = 23$, signed rank test; the red dashed line indicates the median, 0.14). Bottom: Gridness was not significantly different between bouts and non-bouts ($P = 0.11$, Wilcoxon rank-sum test). Gridness was higher during bouts compared to non-bouts for 11 of 23 grid cells (blue lines), and lower for 12 of 23 (red lines); the black line represents the average gridness scores. **b.** The rate map (left) and autocorrelogram (right) for one grid cell are shown during theta bouts (top), during periods without theta bouts (middle), and during the full session (bottom).
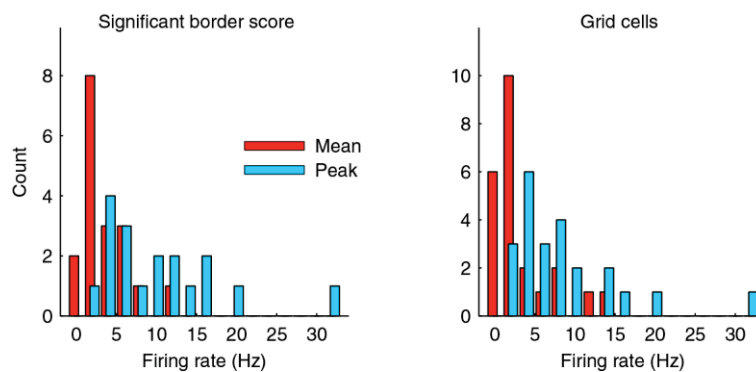
Supplementary Figure 9 | Firing rates of cells with significant gridness scores or border scores. Left: histograms of the mean (3.41 ± 0.71) and peak rate (10.4 ± 1.74) are shown for the 18 cells with significant border scores. The average peak to mean ratio was 3.97 ± 0.55. Right: histograms of the mean (3.31 ± 0.73) and peak rate (8.55 ± 1.45) are shown for the 23 grid cells. The average peak to mean ratio was 4.72 ± 1.39 (mean $\pm$ s.e.m.).

Supplementary Methods

Behavioral task and parameters

We examined neuronal activity in the entorhinal cortex (EC) and hippocampus of three male monkeys performing the Visual Preferential Looking Task (VPLT), using experimental control software (CORTEX; http://dally.nimh.nih.gov/)[5,6]. Performance on this task represents recognition memory and is sensitive to the integrity of the hippocampus and adjacent cortical regions[7]. For each recording session the monkeys (*Macaca mulatta*) performed the VPLT, in which 200 novel, complex visual images obtained from a database of over 25,000 images constituted the stimuli of one session. In this task, each image was viewed twice with up to 10 intervening stimuli between presentations. On each trial, after a 1-second fixation period, an 11 x 11 degree image appeared on the screen and the monkey was allowed to explore the image until the tracked eye position moved outside of a 12x12 degree region, or a maximum of 5 seconds (7 seconds in TT, EC recordings). For the experiments with the $33.\overline{3}$ x 25 degree images covering the entire viewable region of the monitor, 36 novel images were shown twice with 10 seconds of viewing required for each presentation. The gaze location was recorded with an infrared eye-tracking system (ISCAN, Inc.). Stimuli were presented on a 19 inch diagonal CRT monitor at a 120 Hz refresh rate. The monkeys were head-fixed and the distance of the monitor to the eyes was maintained at 60 centimeters.

Electrophysiological recordings

We recorded spikes (250-8000 Hz) and local field potentials (LFPs, 0.7-170 Hz) from the EC with a laminar electrode array mounted on a tungsten microelectrode (AXIAL array: 12-site, 30 μm diameter, 150 μm spacing, FHC, Inc.). Recordings were performed using hardware and software from Plexon, Inc., and all subsequent analyses were performed using MATLAB (The MathWorks, Natick, MA). Spikes were recorded from the hippocampus with 4 independently moveable tungsten microelectrodes. All recordings were performed in the left hemisphere. Recording sites were planned with the aid of MRI. Electrodes inside a stainless steel

guide tube were slowly lowered through a craniotomy at the beginning of each recording session. Electrodes were then advanced out of the guide tube to the recording site while the animals performed unrelated tasks. Electrodes in the hippocampus were advanced until single units could be isolated. The laminar array used to record from the EC was advanced very slowly until noise on the tungsten tip was observed due to reaching the ventral surface (the array was then retracted until no such signal was present). Recordings were initiated after stable spiking could be verified (usually about 30 minutes). Spikes were sorted offline into distinct clusters using principle components analysis (Offline Sorter, Plexon, Inc.). Sorted clusters were classified as single units if the root-mean-square error (RMSE) of the waveforms from the mean waveform was below 0.3 after normalizing to the mean waveform of the cluster. This secondary criterion helped to restrict waveform groups to signals from single neurons, but may have included clusters with waveforms recorded from multiple neurons in some cases. In addition, units with fewer than 150 spikes and less than 500 seconds of total sampling were excluded from the analyses to reduce sampling error. We recorded from three monkeys: 223 single units from the EC of 2 monkeys (MP and TT) and 119 single units from the hippocampus of 2 monkeys (MP and IW). For the EC recordings, 193 single units were recorded from the posterior region of the EC (up to IA +16) and the other 30 units were recorded from more anterior regions.

Theta phase-locking was examined by comparing spike times to the instantaneous phase of the theta-band LFP (the phase of the analytic signal taken after filtering from 3-12 Hz with a 4th-order Butterworth filter). For this analysis, the LFPs were recorded on an adjacent electrode 150 microns away (or 300 microns away if the nearest neighbor electrode was not in use). Sixty-degree bins were used to compute the probability of spiking[8,9].

On average, the EC units exhibited a match suppression effect as reported previously[10]. As a population, the EC units exhibited a decreased firing rate for the repeated stimulus from about 250 to 500 msec after stimulus onset. To analyze the firing rates for individual units, spike counts from 250 to 500 msec after stimulus onset were compared with a Student's t-test to spike counts 500 to 250 msec before stimulus onset. Variable

trial lengths were allowed and only trials that lasted at least 310 msec were included in the analysis. Units were classified as visually responsive if either the encoding (novel stimulus presentation) or recognition (repeated stimulus presentation) trials had a significant difference in rate compared to baseline. Of the visually responsive units, units with memory responses were identified using a Student's t-test to compare the same time windows after stimulus onset with paired novel and repeat trials. The regions of significant differences between novel and repeat viewings for individual units in Figure 2 were computed using a cluster-based non-parametric permutation test ($P_{\text{two-sided}} < 0.05$)[11]. Logistic regression was used to examine the relationship between anterior-posterior location and proportion of visually responsive or differentially responsive cells. Significance of the anatomical location as a predictor was calculated using a chi-squared test with 1 degree of freedom on the increase in deviance of the fit without the predictor.

Layer classification

Location within the EC was classified with the aid of current source density (CSD) analysis, using simultaneously recorded local field potentials (LFPs)[12]. During the fixation periods of visual exploration, two regions of excitation (intracellular current sinks) were observed. Spatial regions excited first were classified as superficial layers and regions excited later in time were classified as deep layers.

Rate maps and autocorrelations

Rate maps were computed with a Gaussian smoothing procedure (Supplementary Table 1)[13]. Data from the image presentation periods were concatenated across trials. Spikes and eye position sampled at 1 kHz were binned into square bins over the viewing region. The resulting values were smoothed by convolving with a 2D Gaussian kernel approximation out to ±2 standard deviations around the center bin and divided to estimate the firing rate over space. To reduce sampling error, a minimum looking time of 100 msec per 1 deg^2 area was required. Pixels below the minimum looking time requirement were removed from rate maps if they were part

of a contiguous 8-connected neighborhood (islands of a single bin were not removed). Distributions of mean and peak rates of the rate maps for cells with significant gridness scores and border scores are shown in Supplementary Figure 9.

Spatial autocorrelations were computed by shifting a copy of the discretized rate map with respect to itself as in the autocorrelation function, but with the values at each x and y translation computed using the Pearson product-moment correlation coefficient formula[13]. A set number of valid (above the minimum looking time) pixels was required to compute an autocorrelation value for a given translation..

Stability of rate maps (Supplementary Fig. 4b) was analyzed by removing the initial 120 seconds of the session (25% of the session on average), dividing the remaining data in half, computing the rate map for each half with smoothing performed after dividing binned spikes by binned time, and taking the Spearman rank correlation coefficient between the corresponding pixels in the two rate maps.

Grid cell classification

The two gridness calculation approaches (with or without ellipse correction) were applied to the original data and 100 time-shifted versions the data. The spike train was shifted in 100 linearly-spaced increments from 20 seconds into the session to the end of the session minus 20 seconds and the resulting rate maps, spatial autocorrelations, and gridness scores were calculated. Units with a gridness score above the 95th percentile of the surrogate data for each unit were considered grid cells.

To validate the time-shifting method, we designed an alternative test that uses a model of spiking to take into account other covariates. We modeled firing rate for each neuron using the average firing rate during times after saccade onset and stimulus onset. Simulated spikes were generated based on an inhomogeneous Poisson process with 1 msec time bins and having average firing rate as a function of time, $\lambda(t)$:

$$\lambda(t) = \sum_{i}^{N_i} \lambda_p(t - t_i)\left(u(t - t_i) - u\left(t - (t_i + l_i)\right)\right) \sum_{k}^{N_k} \lambda_s(t - t_k)\left(u(t - t_k) - u\left(t - (t_k + l_k)\right)\right)$$

For the ith image presentation of N_i presentations of length l_i beginning at time t_i and N_k saccades of length l_k starting at time t_k with normalized mean rate $\lambda_s(t)$ where $u(t)$ is the Heaviside step function and $\lambda_p(t)$ is the mean rate with respect to stimulus onset. The saccade rate function was normalized to its maximum to modulate the rate based on stimulus onset time. One hundred simulations were run for each grid cell and the gridness score was calculated for each resulting rate map. Neurons with gridness scores above the 95th percentile were classified as grid cells. With this test, 19 of 23 grid cells classified with the time-shifting test were still classified as grid cells and 1 new grid cell was identified. The number of grid cells with this test, 20 of 193 (10.4%), was significantly above what would be expected by chance with a P-value of 0.0006 using a chi-squared test of independence.

Border scores

Cells with a border score above the 95th percentile of spatial information rate for 100 time-shifted permutations (the same permutations used for gridness scores) were classified as significantly preferring to fire along borders. The border score was the same as that used in Solstad et al. 2008, however where Solstad et al. used a rate map threshold of 30% of the maximum rate and a border score cutoff of 0.5 we used a rate map threshold equal to the mean firing rate over all bins and a permutation distribution to define significant border scores[14]. Significant border scores ranged from 0.53 to 0.69. The border score, b, was calculated by comparing the maximum coverage of a given stimulus edge by a single firing field, c_M, to the firing rate-weighted average

distance to the nearest edge, d_m : $\text{b} = \dfrac{c_M - d_m}{c_M + d_m}$

Firing field size and spacing

The distance between firing field centers was set to the best outer radius determined through the gridness score calculation procedure[13]. This helped reduce error from peak detection, though some peak detection error

remained through determining the inner radius. Correlations between anatomical location and field spacing were calculated using the Spearman's rho rank correlation coefficient. Only grid cells recorded with the smaller images were used in the correlation (22 of 23 grid cells, 11 per monkey).

Grid cell locations

The posterior-anterior anatomical location of the EC was estimated by taking fourteen 1 mm slices from an MRI scan to constitute the range of the EC. The extent of the EC was identified by comparing anatomical features on the MRI scans to an atlas[3]. The fundus of the rhinal sulcus was identified in MRI scans to estimate the medial distance of the recording locations. Locations of all grid cells were estimated using MRI scans and known electrode positions (Supplementary Fig. 3). Grid cells were identified in all recorded layers of the EC.

Theta-band analyses

For each LFP recording, Morlet wavelets were used to calculate power over time for 30 $\log_2$-spaced frequencies from 2 to 16 Hz. For each frequency of interest, a chi-squared distribution with 2 degrees of freedom was fit to have a mean equal to the mean power and contiguous time points with power above the 95th percentile for at least 3 cycles were defined as bouts. Theta bouts were defined as any contiguous time points of bouts from 3-12 Hz (3.08-12.01 after $\log_2$-spacing)[15]. The percentage of time in a theta bout was higher during pre-stimulus fixation periods ($0.57 \pm 0.07\%$) compared to image viewing ($0.43 \pm 0.09\%$) ($P < 1\times10^{-5}$, Wilcoxon rank-sum test), suggesting that bouts may have been modulated by saccades.

Theta modulation of individual neurons was quantified by the theta index of the power spectrum of the spike train autocorrelation[8,13]. The autocorrelation was computed using a bin size of 2 ms and truncated to ± 0.5 seconds. The signal was zero-padded to 2^{16} samples and tapered with a Hann window before computing the FFT and the resulting power spectrum. The theta index was calculated as the ratio between the average power

from ± 1 Hz of the tallest local maximum (if any local maxima existed) between 3 and 12 Hz and the mean power from 0 to 125 Hz. If no local maxima existed, the theta index could not be calculated.

Theta phase-locking was examined by comparing spike times to the instantaneous phase of the theta-band LFP (the phase of the analytic signal taken after filtering from 3-12 Hz with a 4th-order Butterworth filter). For this analysis, the LFPs were recorded on an adjacent electrode 150 microns away (or 300 microns away if the nearest neighbor electrode was not in use). Sixty-degree bins were used to compute the probability of spiking and histograms[8,9].

Saccade analyses

Saccades were detected by thresholding the eye movement speed at 25 degrees per second. To reduce detection of saccades due to noise in the recorded eye position, thresholding was performed after filtering the gaze location with a 30th-order low-pass (30 Hz cutoff) linear phase digital finite impulse response filter. For the analysis of saccade rate (Fig. 3d), only image viewing trials at least 500 msec long with at least 2 detected saccades were included.

Supplementary Table 1 | Rate map and autocorrelation parameters.

Parameter	Image size	
	small	large
bin width (deg)	0.25	0.5
kernel dimensions (pixels)	13x13	13x13
Gaussian standard deviation (deg)	0.75	1.5
min. looking time (msec/deg^2)	100	100
min. overlap for autocorr. (pixels)	10	50

References

1. Sargolini, F. *et al.* Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science* **312**, 758–62 (2006).
2. Brandon, M. P. *et al.* Reduction of theta rhythm dissociates grid cell spatial periodicity from directional tuning. *Science* **332**, 595–9 (2011).
3. Saleem, K. S. & Logothetis, N. K. *A Combined MRI and Histology Atlas of the Rhesus Monkey Brain in Stereotaxic Coordinates*. **2006**, (Academic Press: London, 2007).
4. Insausti, R. & Amaral, D. G. Entorhinal cortex of the monkey: IV. Topographical and laminar organization of cortical afferents. *J. Comp. Neurol.* **509**, 608–41 (2008).
5. Wilson, F. A. & Goldman-Rakic, P. S. Viewing preferences of rhesus monkeys related to memory for complex pictures, colours and faces. *Behav. Brain Res.* **60**, 79–89 (1994).
6. Jutras, M. J. & Buffalo, E. A. Recognition memory signals in the macaque hippocampus. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 401–6 (2010).
7. Zola, S. M. *et al.* Impaired recognition memory in monkeys after damage limited to the hippocampal region. *J. Neurosci.* **20**, 451–63 (2000).
8. Yartsev, M. M., Witter, M. P. & Ulanovsky, N. Grid cells without theta oscillations in the entorhinal cortex of bats. *Nature* **479**, 103–107 (2011).
9. Mizuseki, K., Sirota, A., Pastalkova, E. & Buzsáki, G. Theta oscillations provide temporal windows for local circuit computation in the entorhinal-hippocampal loop. *Neuron* **64**, 267–80 (2009).
10. Suzuki, W. A., Miller, E. K. & Desimone, R. Object and place memory in the macaque entorhinal cortex. *J. Neurophysiol.* **78**, 1062–81 (1997).
11. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–90 (2007).
12. Pettersen, K. H., Devor, A., Ulbert, I., Dale, A. M. & Einevoll, G. T. Current-source density estimation based on inversion of electrostatic forward solution: effects of finite extent of neuronal activity and conductivity discontinuities. *J. Neurosci. Methods* **154**, 116–33 (2006).
13. Langston, R. F. *et al.* Development of the spatial representation system in the rat. *Science* **328**, 1576–80 (2010).
14. Solstad, T., Boccara, C. N., Kropff, E., Moser, M. & Moser, E. I. Representation of Geometric Borders in the Entorhinal Cortex. *Science* **1109**, 1865–1868 (2008).
15. Ekstrom, A. D. *et al.* Human hippocampal theta activity during virtual navigation. *Hippocampus* **15**, 881–9 (2005).