Extracting Number-Selective Responses from Coherent Oscillations in a Computer Model

Jeremy A. Miller
jeremymiller@ucla.edu
Garrett T. Kenyon
gkenyon@lanl.gov
Los Alamos National Laboratory, Los Alamos, New Mexico 87544, U.S.A.

Cortical neurons selective for numerosity may underlie an innate number sense in both animals and humans. We hypothesize that the number-selective responses of cortical neurons may in part be extracted from coherent, object-specific oscillations. Here, indirect evidence for this hypothesis is obtained by analyzing the numerosity information encoded by coherent oscillations in artificially generated spikes trains. Several experiments report that gamma-band oscillations evoked by the same object remain coherent, whereas oscillations evoked by separate objects are uncorrelated. Because the oscillations arising from separate objects would add in random phase to the total power summed across all stimulated neurons, we postulated that the total gamma activity, normalized by the number of spikes, should fall roughly as the square root of the number of objects in the scene, thereby implicitly encoding numerosity. To test the hypothesis, we examined the normalized gamma activity in multiunit spike trains, 50 to 1000 msec in duration, produced by a model feedback circuit previously shown to generate realistic coherent oscillations. In response to images containing different numbers of objects, regardless of their shape, size, or shading, the normalized gamma activity followed a square-root-of-\(n\) rule as long as the separation between objects was sufficiently large and their relative size and contrast differences were not too great. Arrays of winner-take-all numerosity detectors, each responding to normalized gamma activity within a particular band, exhibited tuning curves consistent with behavioral data. We conclude that coherent oscillations in principle could contribute to the number-selective responses of cortical neurons, although many critical issues await experimental resolution.

1 Introduction

Both animals and humans possess an innate number sense that has been characterized in a variety of behavioral experiments (Dehaene, 1997, 2001; Nieder, 2005). Monkeys, for example, can distinguish between images
containing different numbers of objects regardless of their shape, size, or distribution, demonstrating an ability to utilize an abstract sense of number independent of nonnumerical factors, such as the total area of illumination or total integrated contour length (Brannon & Terrace, 1998; Nieder & Miller, 2004a). Likewise, preverbal human infants respond specifically to changes in numerosity (Feigenson, Dehaene, & Spelke, 2004), and the ability to perform numerosity discriminations under controlled conditions has been documented across a wide range of animal taxa, including birds, rats, dolphins, cats, and primates (Davis & Perusse, 1988; Nieder, 2005). Adult humans also possess an innate number sense, distinct from arithmetical computation, that can be revealed by measuring performance on numerosity discrimination tasks in which explicit counting is prevented, either by appropriate experimental protocols (Cordes, Gelman, Gallistel, & Whalen, 2001; Whalen, Gallistel, & Gelman, 1999) or by studying cultures that lack numerically precise words for all but the first few integers (Pica, Lemer, Izard, & Dehaene, 2004).

The evolution of an innate number sense is clearly adaptive, as it provides, for example, a basis for making optimal foraging, fight-or-flight, and mating decisions (e.g., “Which branch has more fruit?” or “Which troupe has more individuals?”). In both animals and humans, performance on numerosity discrimination tasks typically depends on the ratio of the quantities being compared, so that 2 versus 4, 4 versus 8, 8 versus 16, and so on are all distinguished with approximately equal levels of accuracy. Similarly, behavioral experiments exhibit both a prominent size effect (e.g., 3 versus 4 can be distinguished more reliably than 4 versus 5) as well as a prominent distance effect (e.g., 3 versus 5 can be distinguished more reliably than 3 versus 4). Combined with data showing that the latency of such judgments depends only weakly on numerosity, the above results indicate that animals and humans, especially preverbal human infants, utilize an innate, qualitative sense of number in which individual items are processed in parallel without resort to explicit counting or reliance on the formal mathematical concept of an integer.

Recordings from single neurons in the monkey prefrontal and posterior parietal cortex have identified cells that respond selectively to a preferred number of visual items, usually between one and five, independent of nonnumerical factors such as shape, total area, or boundary length, or the distribution of the individual objects in the scene (Nieder, Freedman, & Miller, 2002; Nieder & Miller, 2003, 2004b). Similar to behavioral measures, number-selective neurons exhibit both a size effect—tuning curves become broader as their preferred numerosity increases (e.g., the cells respond to a wider range of numerosities)—and a distance effect—the overlap between tuning curves falls off with increasing numerical separation. The response properties of number-selective cortical neurons make them attractive candidates for mediating the innate number sense documented in related behavioral studies.
Here, we explore the hypothesis that coherent, object-specific oscillations implicitly encode numerosity in a manner that might contribute to the number-selective responses of cortical neurons. Neurons in the early visual system often exhibit coherent oscillations at frequencies within the gamma- or upper gamma band, here defined broadly as between 40 and 120 Hz (Castelo-Branco, Neuenschwander, & Singer, 1998; Gray, König, Engel, & Singer, 1989; Gray & Singer, 1989; Kreiter & Singer, 1996; Neuenschwander, Castelo-Branco, & Singer, 1999; Neuenschwander & Singer, 1996; Singer & Gray, 1995). Typically, gamma-band oscillations recorded from visual neurons responding to the same object are strongly correlated, even though their common phase drifts randomly over time, as revealed by the absence of periodic structure in stimulus-locked measures, such as the peristimulus time histogram (PSTH), as well as by the declining amplitude of successive periodic side bands in autocorrelograms that otherwise exhibit strong oscillations, an indication that phase coherence is lost after several cycles. Moreover, the oscillatory responses evoked by different objects do not remain phase-locked, so that the cross-correlogram between two recordings sites may be flat even though the autocorrelograms recorded at each site show strong periodic modulations at roughly equal frequencies. It follows from the object-specific nature of coherent oscillations that if only a single item is present in a scene, all of the stimulated cells will be modulated in phase, and the average spectral amplitude in the gamma band, computed from the summed activity across all stimulated neurons, will be maximal when normalized by the total number of spikes. If the scene instead contains two objects of similar size and luminance, there will be two sources of oscillatory activity whose relative phase will vary randomly over time. Thus, in the case of two objects, the total normalized power in the gamma band, assuming the sum includes all stimulated neurons, will be reduced by approximately the square root of 2. Likewise, for \( n \) distinct objects of similar size, the normalized gamma power will be reduced by approximately the square root of \( n \) (i.e., the average of \( n \) imaginary numbers possessing equal magnitudes and random phases). If it is assumed that all distinct visual items, regardless of their shape, orientation, shading, or distribution, evoke coherent, object-specific oscillations, the resulting estimate of numerosity will be automatically independent of such nonnumerical factors. Although the status of gamma activity in normal visual processing remains hotly debated (Sejnowski & Paulsen, 2006; Shadlen & Movshon, 1999; Singer & Gray, 1995), we may nonetheless ask to what extent numerosity can be encoded by coherent, object-specific oscillations.

To investigate the numerosity encoded by coherent oscillations, we analyzed the normalized gamma power from multiunit spike trains obtained from a circuit model previously shown to generate realistic spatiotemporal correlations (Kenyon, Harvey, Stephens, & Theiler, 2004; Kenyon et al., 2003; Kenyon, Theiler, George, Travis, & Marshak, 2004; Kenyon, Travis et al., 2004). To obtain similar measurements from an experimental preparation
would have required recording simultaneously from large numbers of oscillatory neurons all of the same type and distributed across the visual field. Our theoretical study used computer-generated spike trains to investigate the feasibility of the hypothesis that gamma-band oscillations can, in principle, encode numerosity as a prelude to a more rigorous, yet technically difficult experimental study. Our results suggest that the normalized gamma activity, summed over all stimulated neurons and divided by the total spike count, permits single-trial discriminations to be made between images containing different numbers of objects. Because the circuit model was uniform and isotropic and produced object-specific, coherent oscillations in response to all sufficiently large, contiguous features, the resulting numerosity information was encoded in a manner that was inherently independent of the shape, orientation, shading, and distribution of the individual items in the scene.

Previous attempts to compute numerosity using artificial neural networks implicitly utilized separate nodes for every enumerable object at each image location, either by starting with an explicit location map or by limiting the implementation to a 1D retina in which all possible lengths could be accounted for by appropriate winner-take-all filters (Dehaene & Changeux, 1993; Verguts & Fias, 2004). Our approach is fundamentally different. By utilizing the numerosity information encoded by stimulus-specific, coherent oscillations, it was possible to achieve invariance with respect to nonnumerical factors, such as size, shape, shading, and distribution, in a fully 2D model without introducing a combinatorially large number of additional elements. Finally, to compare our results to experimental measurements, we constructed an array of winner-take-all numerosity detectors that used only the normalized gamma activity present on single trials. The resulting theoretical tuning curves were consistent with behavioral data, supporting the hypothesis that gamma-band oscillations can, in principle, provide an implicit representation of numerosity that might contribute to the innate number sense shared by animals and humans.

2 Methods

2.1 Retinal Model. A negative feedback circuit, originally developed to account for high-frequency oscillations between retinal ganglion cells (Kenyon et al., 2003), was used to generate artificial spike trains with realistic spatiotemporal correlations in a manner consistent with known anatomy and physiology. Coherent oscillations generated by the model were of similar amplitude, frequency, and duration to those measured experimentally (Kenyon et al., 2003). Moreover, the gamma-band oscillations between model neurons were strongly size dependent and remained coherent only within contiguously stimulated regions (Kenyon, Harvey et al., 2004; Kenyon, Travis et al., 2004), thereby capturing important aspects of the object-specific, spatiotemporal correlations between neurons in the cat.
retina (Neuenschwander et al., 1999). The circuit model thus provided a reasonable basis for estimating the numerosity information encoded by coherent oscillations in the central nervous system. Because details of the circuit model have been presented previously, only an abbreviated description is provided below (see also Figure 1).

Input to the model was conveyed by an array of external currents proportional to the pixel-by-pixel gray-scale value of a two-dimensional image. These external currents directly stimulated the model bipolar cells and approximated their light-modulated synaptic input from cone photoreceptors. The bipolar cells produced excitatory postsynaptic potentials in both ganglion cells and amacrine cells according to a random process (Freed, 2000). The axon-bearing amacrine cells were electrically coupled to neighboring ganglion cells and to each other, and made strong inhibitory connections onto the surrounding ganglion cells and axon-bearing amacrine cells. This feedback circuit produced robust, physiologically realistic oscillations in response to large stimuli. When several ganglion cells were activated by a stimulus, they in turn activated neighboring axon-bearing amacrine cells via gap junctions (Dacey & Brace, 1992; Jacoby, Stafford, Kouyama, & Marshall, 1996; Vaney, 1994). The stimulated cells were then hyperpolarized by the ensuing wave of axon-mediated inhibition, thus setting up the next cycle of the oscillation. Spike generation in ganglion cells was modeled as a leaky integrate-and-fire process with a membrane time constant of 5 msec, consistent with published physiological data (O’Brien, Isayama, Richardson, & Berson, 2002). The model also contained local nonspiking amacrine cells that generated randomly distributed inhibitory postsynaptic potentials that helped make spontaneous firing asynchronous and acted to increase

Figure 1: Schematic. (a,b) Circuit diagram. The model contained five distinct cell types—bipolar (BP), small amacrine (SA), large amacrine (LA), polyaxonal amacrine (PA), and ganglion cells (GC)—whose implementation was based on the anatomy and physiology of the motion processing pathway in the cat inner retina. (a) Excitation, feedforward, and feedback inhibition. BPs (relay neurons) were activated directly by the stimulus and in turn excited the other cell types (open triangles). Both the SAs and LAs provided feedforward inhibition to the GCs (closed circles), while all three amacrine cell types (retinal interneurons) made inhibitory feedback synapses onto the BPs. (b) Serial inhibition. Interactions between retinal interneurons were arranged as a negative feedback loop. (c) Oscillatory circuit. Local excitation, via gap junctions, elicited long-range, axon-mediated inhibition, producing realistic coherent oscillations. (d) Winner-take-all numerosity detectors. Normalized gamma activity, a measure of the coherent oscillations across all stimulated cells, is postulated to decline as the inverse square root of the number of objects. Bands representing two and four objects are illustrated. The normalized gamma activity present on any given trial was guaranteed to fall within the range of one and only one numerosity detector.
a) Excitation, Feedforward & Feedback Inhibition

b) Serial Inhibition

c) Oscillatory Circuit

d) winner-take-all numerosity detectors
the overall dynamic range, but were not otherwise critical for generating oscillatory spike trains.

The behavior of the model in response to a variety of stimuli, the robustness of the model with respect to both numerical and physiological parameters, and the connection of the model to anatomical and electrophysiological data, are described in detail elsewhere (Kenyon, Harvey et al., 2004; Kenyon & Marshak, 1998; Kenyon et al., 2003; Kenyon, Theiler et al., 2004; Kenyon, Travis et al., 2004). Some model parameters were modified from previously published values. Specifically, all axon-mediated feedback interactions were assigned a fixed synaptic delay of 2 msec and a finite transmission delay, corresponding to an action potential conduction velocity of 4 ganglion cells per time step (~1 mm/msec), was incorporated. The time constant of the axon-bearing amacrine cells was also increased slightly, from 5 to 7.5 msec. By increasing the overall delay of the axon-mediated negative feedback loop, the above changes tended to strengthen the coherent oscillations in the retinal model, but this effect was compensated for by reducing the strength of the feedback inhibition onto the model ganglion cells by a factor of 0.44. Overall, the above changes made no qualitative difference in the behavior of the coherent oscillations in the dynamic feedback circuit but were implemented to make the model more realistic and provided a better fit to the high frequency resonance in the responses of cat Y ganglion cells to rapidly modulated stimuli (Frishman, Freeman, Troy, Schweitzer-Tong, & Enroth-Cugell, 1987; Miller, Denning, George, Marshak, & Kenyon, 2006).

2.2 Common Input Model. To control for the possibility that the dynamics of our retinal circuit model, which depended on long-range inhibitory feedback from spiking amacrine cells, could have introduced spurious higher-order correlations that might have unintentionally augmented the numerosity information encoded by coherent oscillations, we generated a second set of artificial spike trains in which the model ganglion cells were simply passive recipients of a common oscillatory input that uniformly modulated the firing rates of all neurons responding to the same contiguous stimulus. Starting with a frequency spectrum approximated by a single Lorentzian function, a best fit was obtained to the normalized Fourier amplitudes produced by the circuit model in response to an 8 x 8 stimulus. The specific form of the Lorentzian function used was \( (c/\pi) \frac{(0.5a)}{(0.25a^2 + b^2 + 0.25a^2)} \), where \( f \) is the frequency of the three parameters \( a \), \( b \), and \( c \), roughly corresponding to the width, location, and amplitude, which were determined empirically to be 12.9844 Hz, 73.4968 Hz, and 6.4584, respectively. To obtain a time-dependent oscillatory signal with realistic temporal correlations, we performed an inverse Fourier transform by assigning random phases, chosen uniformly between 0 and \( 2\pi \), to the individual frequency components. Parallel sets of artificial spike trains with realistic spatiotemporal correlations were then constructed by applying the same time-dependent oscillatory signal to an 8 x 8 array of Poisson spike generators that simultaneously modulated their instantaneous firing rates.
A constant offset was added to the oscillatory signal to fix the baseline firing rate of the Poisson spike generators at the same value exhibited by ganglion cells in the retinal model, here equal to 31 Hz, well within the range of observed resting firing rates for Y ganglion cells in the cat retina (Passaglia, Enroth-Cugell, & Troy, 2001). Each distinct object was represented by a separate $8 \times 8$ array of Poisson spike generators, obtained by repeating the above process using an independent distribution of random phase assignments, ensuring that the gamma-band oscillations arising from separate stimuli were uncorrelated. Additional mathematical details of a related procedure for generating artificial spike train via common input modulation are described elsewhere (Kenyon, Theiler et al., 2004).

2.3 Stimulation and Data Analysis. Both natural and computer-generated images containing various numbers of distinct objects, encompassing a range of sizes, shapes, shadings, and separations, were presented to the circuit model. Multiunit spike trains were constructed by combining the output of all stimulated output neurons into a single time series. For natural images, a binary mask was used to determine which neurons to include in the multiunit spike train. Nearly identical results to those presented here were obtained when an activity-dependent mask, which selected only those cells that fired a suprathreshold number of spikes over a brief time interval, was used instead of a fixed mask, but the latter method allowed more controlled comparisons across stimulus conditions. In most experiments, stimuli were presented for 200 msec prior to collecting data, so that only steady-state activity during the plateau portion of the response was considered. Multiunit spike trains obtained during the plateau portion of the response were typically 1000 msec, but shorter durations, corresponding to 125, 250, and 500 msec, were examined as well. In one set of experiments, we also considered the numerosity encoded by gamma activity during the transient portion of the response, using multiunit spike trains that included only the first 50 msec following stimulus onset. In all cases, single-trial frequency spectra were obtained as follows: the autocorrelation function was first computed directly from the multiunit spike train and then Fourier transformed, yielding a conventional power spectrum. The result was then divided by the total number of spikes (power in the DC band) to control for variations in the combined area of stimulation. Finally, the square root was taken of each frequency component, so that the final values reflected normalized spectral amplitudes. For each stimulus condition, corresponding to different numbers of objects in the input image, means and standard deviations of the normalized frequency spectra were computed by analyzing the results from 25 independent trials. The peak amplitude in the upper gamma band typically fell between the discrete steps along the frequency axis. To compensate, the maximum lag value of the autocorrelation function was adjusted for each stimulus condition so that the spectral peak was maximal. This latter procedure had no qualitative impact on our results or conclusions but was conducted to ameliorate aliasing effects due to the
use of discrete frequency components. Finally, to obtain a scalar estimate of the normalized gamma activity, the normalized spectral amplitudes were averaged from 65 to 85 Hz, a range that approximately enclosed the peak oscillatory response. The predicted gamma activity for multiple objects was determined by dividing the normalized gamma activity produced in response to a single object by the square root of the number of objects in the corresponding input image.

Comparisons to behavioral data were facilitated by constructing artificial winner-take-all numerosity detectors that responded to single-trial estimates of normalized gamma activity falling in a given span (see Figure 1d). Input ranges were nonoverlapping, such that one and only one detector responded to each stimulus presentation, corresponding to the node whose target normalized gamma activity was closest to the actual normalized gamma activity present on that trial. The probability of discharge for each detector in response to images containing various numbers of objects was computed from the mean and standard deviation of the normalized gamma activity under each stimulus condition, under the assumption that all measured quantities were gaussian distributed. The target normalized gamma activity for each numerosity detector was determined separately in each experiment by the mean value produced by images containing the preferred number of items.

3 Results

To determine whether coherent oscillations could, in principle, encode numerosity, we stimulated the model circuit with computer-generated images containing different numbers of identical square spots, with each spot covering an array of $8 \times 8$ output cells (see Figure 2a). Normalized frequency
spectra, computed from multiunit spike trains 1000 msec in duration obtained during the plateau portion of the response and consisting of all stimulated cells, exhibited different amounts of gamma activity depending on the number of objects in the image, being maximal when only one object was present and declining as the number of objects increased (see Figure 2b). To obtain a scalar estimate of gamma activity, the spectral amplitude was averaged over the response peak, 65 to 85 Hz. The mean normalized gamma activity declined as the number of objects increased in a manner well predicted by the square-root-of-\( n \) rule (see Figure 2c). Error bars denote one standard deviation of the corresponding distribution, not the standard deviation of the mean. There was good discrimination between images containing either one or two objects, as the corresponding means were well outside the associated error bars, but the overlap between neighboring distributions grew steadily as more objects were added, consistent with the size effect observed in a variety of animal and human studies (Nieder, 2005).

To quantify the numerosity information encoded by coherent oscillations, we constructed an array of winner-take-all numerosity detectors that responded to the normalized gamma activity present on single stimulus trials. One and only one detector was activated by each stimulus presentation, corresponding to the element whose target activity, defined as the mean response to images containing the preferred number of objects, was closest to the normalized gamma activity present on that trial. Using the means and standard deviations of the normalized gamma activity for each stimulus condition, percentage activation rates were determined for each detector in response to images containing from one to four objects (see Figure 2d, solid lines). When only one object was present in the input image, the 1-object detector was always activated, corresponding to 0% false negatives. In response to images containing two, three, or four objects, the activation rates for the 2-, 3-, and 4-object detectors were 84%, 67%, and 45%, respectively, indicating an increasing percentage of false negatives for detectors preferring larger numerosities. Likewise, the percentage of false positives, given by the integrated response over all detectors other than the correct node, increased with preferred numerosity as well.

The theoretical tuning curves computed from the single-trial normalized gamma activity (see Figure 2d, dashed lines) agreed quite well with previously published behavioral data (Nieder & Miller, 2004a). Monkeys were trained to report whether the number of items in a test stimulus, given by the value along the abscissa, equaled the number of items in the target image, with each curve representing a different preferred numerosity. The behavioral and model data both exhibited strong size and distance effects. Tuning curves became broader as the preferred numerosity increased, and the overlap between pairs of tuning curves tended to zero as the absolute difference in their preferred numerosities became larger. Because our theoretical analysis involved no fine-tuning to fit the experimental results—the
parameters of the circuit model having been established previously on the basis of electrophysiological recordings—the close correspondence between the model and behavioral data suggests that the encoding of numerosity may be a robust property of coherent, stimulus-specific oscillations. As an additional test of robustness, we obtain qualitatively similar results after repeating the above experiment using the same basic circuit but with slightly different, previously published parameter values (see section 2), confirming that the encoding of numerosity by coherent oscillations did not require precise tuning of the model.

The innate number sense, as characterized both behaviorally and electrophysiologically, depends only on numerosity and is independent of the specific properties of the individual objects in the scene, such as their shape or location. We asked whether normalized gamma activity obtained from the model circuit encoded numerosity in a manner that was likewise robust against nonnumerical factors (see Figure 3). The first set of input images contained various numbers of objects encompassing different shapes, sizes, and orientations, such that the total stimulated area remained invariant (see Figure 3a). This control was necessary to ensure that the normalized gamma activity was in fact specific for numerosity, and not simply a decreasing function of the total stimulated area. As a second, complementary control, we used images in which object size increased in proportion to the number of items present, so that the total stimulated area grew supralinearly (see Figure 3b). Third, to ensure that high-contrast edges were not required for normalized gamma activity to encode numerosity, we repeated the basic experimental protocol using objects that stimulated the same number of cells but possessed continuously shaded boundaries, falling off over one-quarter cycle of a cosine function so that stimulus intensity was described by a series of square, isoluminant contours with the same total integrated strength as the standard $8 \times 8$ spot (see Figure 3c). For both the binary and gray-scale images, the normalized gamma activity encoded numerosity in a manner independent of object-specific properties (see Figure 3d), consistent with the innate number sense characterized experimentally. In addition, the decline in normalized gamma activity as a function of the number of objects was well predicted by a square-root-of-$n$ rule.

In the above experiments, only stimulated cells contributed to the normalized gamma activity. The restriction to stimulated cells was necessary because when all cells were included in the calculation, coherent oscillations no longer provided a useful measure of numerosity, probably because the oscillatory signals arising from the subset of stimulated neurons were overwhelmed by the baseline of ongoing activity. Background firing rates in the model were rather high, approximately 31 Hz, consistent with experimentally measured values (Passaglia et al., 2001). However, oscillatory activity in the retina can propagate to the visual cortex (Castelo-Branco et al., 1998), where the background firing rates of projection neurons may be much lower. Thus, the extraction of numerosity from coherent oscillations.
Figure 3: Invariance of numerosity encoding. (a) Constant area. Individual object sizes were decreased so as to keep the total stimulated area invariant. (b) Supralinear growth. Object sizes increased in proportion to the number of objects in the image. (c) Shaded boundaries. Sharp, high-contrast edges were replaced by smooth gradations. (d) Normalized gamma activity encoded numerosity in a manner consistent with a square-root-of-$n$ rule regardless of object size, shape, or shading.

at points further along the visual processing pathway may not require the elimination of unstimulated cells from the analysis. Alternatively, it may be possible to identify stimulated cells from the same data used to estimate numerosity. As a preliminary test of this possibility, we reanalyzed the responses to the gray-scale objects (see Figure 3c) using an activity-dependent mask, whose threshold was chosen to best discriminate between stimulated and unstimulated cells based on 100 msec spike trains taken from the middle of each trial, to determine which individual units to include in the analysis. In this simple test, we obtained virtually identical results to those shown above, since the activity-dependent mask rarely differed by more than one or two cells from the fixed binary mask specified a priori. Although
a detailed account of how numerosity might be extracted physiologically is beyond the scope of this study, an activity-dependent mask might be implemented physiologically by facilitating synapses (Beierlein and Connors, 2002) or dendritic nonlinearities (Polsky, Mel, & Schiller, 2004).

Behavioral experiments indicate that numerical discriminations do not depend on the absolute difference in the number of items but rather on their ratio (Nieder, 2005; Nieder & Miller, 2003). To determine whether coherent oscillations encode numerosity in a manner consistent with the ratio law observed experimentally, we examined the discrimination mediated by normalized gamma activity in response to computer-generated images containing 1, 2, 4, 8, and 16 identical square spots, each covering an array of $8 \times 8$ model ganglion cells (see Figure 4). As in the above experiments, frequency spectra were computed from mass spike trains that included all stimulated cells. Spectral amplitudes in the upper gamma band, normalized by the total number of spikes, continued to decrease as the number of objects was raised, up to and including the maximum numerosity tested (see Figure 4a). As a function of the number of objects, the mean normalized spectral amplitude between 65 Hz to 85 Hz (gamma activity) adhered quite closely to a square-root-of-$n$ rule, even for images containing up to 16 items (see Figure 4b). Any discrimination procedure based on the percentage change in normalized gamma activity will therefore depend only on the ratio of the number of items being compared. Plotted on a logarithmic scale, model numerosity detectors exhibited symmetrical tuning curves whose half-width at half-maximum (HWHM) was approximately constant (see Figure 4c). The total area under each tuning curve was always unity; the apparent differences were linear distortions due to the logarithmic scale. To the extent that the HWHM remains linearly proportional to the number of items, numerosity discrimination would be expected to obey a ratio law.

We also asked whether the retinal model could encode the numerosity of more realistic gray-scale images, using a photograph that contained five airplanes on a tarmac (see Figure 5a). A somewhat artificial image was employed for these experiments to achieve good object specificity. To simulate both single- and multiple-plane images, we summed the spikes from either one or five distinct groups of selected ganglion cells, indicated by the corresponding binary masks (see Figures 5b and 5c). As with computer-generated images, the normalized frequency spectra computed from the multiunit spike trains showed a clear effect of numerosity, decreasing substantially when the data consisted of responses to five as opposed to a single airplane (see Figure 5d). In additional, the normalized gamma activity computed from the multiunit spike train containing responses to all five planes fell well within the error bars of the value predicted from a simple square-root-of-$n$ rule (see Figure 5e). To ensure that the numerosity encoded in the multiunit spike train corresponding to a single plane was not confounded by the presence of additional items, we repeated the experiment using a new image in which the other four planes were digitally removed,
Figure 4: Ratio law. The circuit model was stimulated with images containing 1, 2, 4, 8, or 16 identical square spots each covering an 8 × 8 array of output neurons. (a) Normalized frequency spectra. Gamma-band activity declined monotonically for numerosities up to 16. (b) Normalized gamma activity (black bars) closely adhered to a square-root-of-\(n\) rule (gray bars). (c) Percentage activation of winner-take-all numerosity detectors. On a logarithmic scale, tuning curves separated by powers of 2 were all approximately of the same width and had similar overlap with their neighbors, consistent with the ratio law observed experimentally.
obtaining equivalent results (not shown). These findings suggest that coherent oscillations may encode numerosity even in response to nonidealized gray-scale images.

The above experiments all used an analysis window of 1000 msec during the plateau portion of the response. To be of maximum behavioral significance, however, the numerosity information encoded by coherent oscillations should be available on shorter timescales as well. We therefore used the circuit model to investigate how numerosity information encoded by gamma activity during the plateau portion of the response might degrade
as the size of the analysis window or, equivalently, the length of the multiunit spike train record was reduced (see Figure 6). The mean normalized gamma activity was roughly constant for analysis window sizes down to approximately 200 msecs and then began rising steadily as the spike train duration was decreased still further, reflecting an overall increase in broadband noise (see Figure 6a). The standard deviation in the distribution of single-trial normalized gamma activities also increased steadily as the spike train duration was reduced (see Figure 6a, error bars). The degradation of numerosity encoding with smaller window sizes was reflected in the corresponding tuning curves, which became broader as the analysis time was decreased (see Figure 6b). However, considerable numerosity information was still present even at the shortest analysis times tested. In particular, reasonable numerical discrimination during the plateau portion of the response was still possible even for analysis times of only 125 msec.

So far we have considered only the sustained portion of the multiunit spike train recorded after the initial response transients have settled to a stable plateau level. One advantage of focusing on the oscillations present during the asymptotic portion of the response is that correlations due to stimulus coordination become negligible after approximately 100 msec, as revealed by a nearly flat shift predictor once the plateau period has been reached (Kenyon et al., 2003). When instead we analyzed multiunit spike trains recorded during the transient portion of the response, the predicted performance on numerical discrimination tasks was greatly impaired (data not shown). Although the model predicted large-amplitude oscillations during the response transient, there was very little specificity between objects due to strong stimulus coordination. However, the model consisted of an unrealistically homogeneous population whose onset latencies were all identical, potentially exacerbating the problem of stimulus coordination during the response transient. When a gaussian spread of onset latencies, with width equal to ±5 msec, was imposed on the model input cells, reasonable performance on the numerical discrimination task was obtained (see Figure 7). Normalized spectra showed a clear trend toward lower amplitudes as the number of objects was increased (see Figure 7a). The peaks became broader in part due to the relatively low resolution (20 Hz) at which individual frequency components could be resolved using 50 msec analysis windows. Normalized gamma activity also showed a clear decline with increasing numerosity (see Figure 7b), although the widths of the corresponding distributions, indicated by the attached errors bars, were rather large due to the reduced signal-to-noise in numerical estimates based on 50 msec multiunit spike records. Similarly, the theoretical tuning curves predicted by a bank of winner-take-all numerosity detectors were much broader than in the standard case (see Figure 7c), although some level of numerosity discrimination was still clearly evident. Our results suggest that as long as there is a sufficient spread of onset latencies to support adequate object specificity during the transient portion of the response, coherent
Figure 6: Effect of analysis window size. (a) Normalized gamma activity versus spike train duration as a function of numerosity. Regardless of the analysis window size, normalized gamma activity declined with increasing numerosity. Distribution widths (error bars) increased for shorter spike trains. Base-line gamma activity also increased as the analysis window was reduced below approximately 200 msec. (b) Percentage activation of artificial numerosity detectors. Tuning curves became broader as the analysis time was decreased, although considerable numerosity information was still present even for analysis windows of 250 msec.
Figure 7: Rapid encoding of numerosity during the response transient. Multi-unit spike trains in response to images containing one to four identical square spots but limited in duration to the first 50 msec following stimulus onset. In order to reduce stimulus coordination, onsets were jittered by an average of ±5 msec independently for each input cell. (a) Normalized frequency spectra. Amplitudes declined with increasing numerosity. (b) Normalized gamma activity declined with numerosity in a manner roughly consistent with a square-root-of-\(n\) rule. Distribution widths (error bars) were very large, reflecting reduced signal-to-noise. (c) Percentage activation of winner-take-all numerosity detectors. Tuning curves, though broad, nonetheless indicated numerosity discriminations could be supported by coherent oscillations during the response transient.
oscillations could begin to encode numerosity very shortly after stimulus onset, consistent with the measured onset latencies of number-selective cortical neurons, which are on the order of 100 msec (Nieder & Miller, 2004b).

Previous analysis of the circuit model suggests that coherent oscillations between separate objects will become phase locked as their spacing is reduced (Kenyon, Travis et al., 2004). The resulting breakdown in object specificity will necessarily confound the numerosity information encoded by coherent oscillations. To quantify this effect, we examined how the normalized gamma activity was affected as a group of four objects, each covering an 8 x 8 array of model ganglion cells, was moved closer together (see Figure 8a). As expected, at high packing densities, the normalized gamma activity became strongly dependent on the interobject spacing, with the transition between object-specific and nonspecific activity occurring at separations approximately three to five times the distance between neighboring output neurons (see Figure 8b). However, even at the highest densities examined, the four objects still could be distinguished from a single object on the basis of normalized gamma activity alone, although the exact numerosity was no longer represented unambiguously. Thus, even though strict specificity broke down as the objects were packed closer than approximately five interneuron spacings, significant numerosity information was still retained even for very small separations. Moreover, as long as the objects in the image were sufficiently well separated, the numerosity encoded by the normalized gamma activity in the circuit model was independent of their spatial distribution.

The above experiments all employed objects possessing the same maximum stimulus intensity. Would coherent oscillations continue to encode numerosity at stimulus intensities corresponding to different levels of luminance contrast? Computer-generated images containing from one to four identical objects, with each again covering an array of 8 x 8 output neurons, were presented to the circuit model with the maximum stimulus intensity set to 1/2 or 3/2 times the standard value (see Figure 9, columns 1 and 2, respectively). For a single object, the averaged plateau firing rates were 50, 62, and 87 Hz, corresponding to stimulus intensities of 1/2, 1, and 3/2, respectively, relative to a baseline firing rate of 31 Hz. Despite dividing by the total number of events in the multiunit spike train, the normalized spectral amplitudes were not contrast invariant (see Figure 9a; compare columns 1 and 2). Coherent oscillations in the circuit model are more sensitive to contrast than are the firing rates (Kenyon, Theiler et al., 2004), possibly contributing to the lack of contrast invariance. Nonetheless, for any given contrast, the normalized spectral amplitudes still declined steadily with increasing object number. Likewise, the normalized gamma activity followed a square-root-of-n rule over a threefold range of stimulus intensities (see Figure 9b). In addition, the numerosity discrimination exhibited by the winner-take-all detectors was also roughly invariant with respect to stimulus intensity, as revealed by the similarity of the corresponding tuning
Figure 8: Effects of object spacing. (a) The spacing between four identical objects, each covering an $8 \times 8$ array of output neurons, was reduced from 8 to 0, with 1 corresponding to the distance between two neighboring output neurons. (b) Normalized gamma activity increased as the spacing between objects was reduced. The oscillations evoked by separate objects became partially phase-locked as they were moved closer together, obscuring numerosity information, although discrimination between one and four objects was still possible even at the smallest nonzero separations tested.
Number-Selective Responses 1787

curves (see Figure 9c). These results imply that coherent, object-specific oscillations can encode the numerosity of distinct visual items at various luminance contrasts, as long as the differences in the relative luminance contrast between objects in the same image are not too large. Otherwise, some mechanism for contrast normalization, or for making the image representation more binary, would be required to apply the model as currently formulated. We are unaware of any behavioral experiments that explicitly address the issue of how relative luminance contrast affects innate numerical comparisons.

It was also useful to consider how numerosity estimates based on normalized gamma activity became degraded as the disparity in the relative sizes of the different objects in the image was made progressively larger. The model circuit was presented with sets of images containing various numbers of objects with equal intensity but whose relative areas differed by a factor of up to 16 (see Figure 10). We expected that the strong size dependence of the coherent oscillations, coupled with surround inhibition, would act to partially preserve numerosity information in the normalized gamma activity even when the image contained objects of different sizes. For input images consisting of one relatively large spot, covering a 16 × 16 array of model ganglion cells, in combination with either one, two, or three much smaller spots, each covering only a 4 × 4 array of model ganglion cells, the numerosity information conveyed by the normalized gamma activity was seriously confounded (see Figure 10a). Some numerosity information was still preserved, as it was possible to partially distinguish between one versus four objects (see Figure 10a, right column), but overall such large discrepancies in relative size had a major impact on the accuracy of numerical comparisons. The predicted performance on numerosity discrimination tasks improved dramatically as the difference in the relative sizes of the individual objects was reduced. For example, the numerosity encoded by gamma activity in response to images containing objects whose relative size differed by a factor of 4 supported discriminations comparable to that obtained when relative size differed by only a factor of 2 (see Figure 10b versus Figure 10c). Although performance on numerosity discrimination tasks, when based solely on the normalized gamma activity, clearly degraded as the size disparity between the individual objects became very large, coherent oscillations continued to encode potentially useful numerosity information over an approximately fourfold range of relative object sizes. We are unaware of any behavioral experiments that explicitly examine the issue of how numerical discriminations are affected by large differences in relative object size, but our results clearly predict a significant deterioration in task performance as such disparities become exceedingly large.

Finally, to increase the generality of our conclusions, we considered the possibility that the numerosity encoded by normalized gamma activity in the retinal model was influenced by simulation artifacts. Although the
pairwise correlations between model neurons have been shown to be consistent with those measured experimentally, our simulations, based as they were on complex, nonlinear feedback dynamics, might have introduced spurious higher-order correlations that unintentionally favored the extraction of numerosity from multunit spike trains over single stimulus trials. As a control, we constructed a new set of multunit spike trains based on a much simpler dynamical process: a common oscillatory modulation applied to an array of identical Poisson spike generators. To model stimuli containing \( n \) objects, we constructed \( n \) separate \( 8 \times 8 \) arrays of Poisson spike generators, with each \( 8 \times 8 \) array driven by an independent source of common modulatory input. This procedure yielded a set of \( n \) multiunit spike trains with realistic gamma-band oscillations that were mutually uncorrelated with each other. As with the data generated by the more complex circuit model, the multiunit spike trains produced by common modulatory input supported reasonable numerosity discrimination (see Figure 11). Indeed, the diminished normalized spectral amplitudes with increasing numerosity (see Figure 11a), the corresponding decline in normalized gamma activity (see Figure 11b, black bars), the general adherence to a square-root-of-\( n \) rule (see Figure 11b, shaded bars), and the average responses of the numerosity detectors to single-stimulus presentations (see Figure 11c), were all nearly identical to corresponding measures obtained from the dynamic feedback model. However, because the correlations between the separate Poisson spike generators were due entirely to a common modulatory input, the possibility of higher-order artifacts being present in these experiments was greatly reduced.

4 Discussion

To test the hypothesis that coherent oscillations can implicitly encode numerosity, images containing different numbers of objects, encompassing various shapes, sizes, and shadings, as well as different packing densities, intensities, and areas of illumination, were used to stimulate a model feedback circuit that generated gamma-band activity consistent with experimental data. As a function of the number of objects in the input image,
Figure 10: Numerosity discriminations involving stimuli of different relative size. All images contained one large object, covering a $16 \times 16$ array of output neurons, plus zero to three smaller objects. Relative sizes illustrated by insets at the end of each row. (a) Top row: Ratio of large to small object size: 16:1. (b) Middle row: Ratio reduced to 4:1. (c) Bottom row: Ratio reduced to 2:1. Left column: Normalized gamma activity. The square-root-of-$n$ rule was strongly violated at very large ratios of relative object size but was in reasonable accord with the data for relative size ratios less than approximately 4:1. Right column: Percentage activation of winner-take-all (Polsky et al., 2004) numerosity detectors. Tuning curves became reasonably distinct for relative size ratios less than approximately 4:1.
the normalized gamma activity measured during the plateau portion of the response followed a square-root-of-\( n \) rule rather closely, consistent with the expectation that the gamma-band oscillations arising from within a single object add in phase, whereas the oscillations arising from separate objects add with random phase. The square-root-of-\( n \) rule was found to be generally robust to the shape, size, shading, contrast, and distribution of the individual objects with several notable exceptions. Decreasing the separation between objects confounded numerosity estimates at high packing densities, as the oscillations evoked by separate objects became partially coherent (Kenyon, Travis et al., 2004). The inability of the feedback circuit to completely segment cluttered natural scenes, especially those containing large, overlapping objects, makes it unlikely that the model could have been successfully applied to more realistic images. Likewise, numerosity comparisons involving sets of objects encompassing very different relative sizes or very different relative contrasts (here modeled as stimulus intensity) were also confounded, as the brightest and largest objects could, if the relative disparity was sufficiently great, dominate the normalized gamma activity. Finally, using shorter analysis windows reduced the signal-to-noise ratio and thus led to degraded numerosity discrimination. Nonetheless, except in the most limiting cases, it was usually possible to discriminate between one and several objects based on the normalized gamma activity alone, virtually regardless of nonnumerical factors.

To quantitatively assess the numerosity information encoded by coherent, object-specific oscillations, we constructed winner-take-all numerosity detectors based on the normalized gamma activity integrated across all stimulated neurons. An individual numerosity detector was activated whenever the normalized gamma activity on that trial fell nearest to its target value, given by the mean gamma activity evoked by images containing its preferred number of objects. The accuracy of these detectors followed the same general trends for numerosity discrimination observed behaviorally (Nieder & Miller, 2003). As the number of objects increased, the tuning curves broadened, so that it was progressively more difficult to distinguish sets of objects differing by a fixed number of items. Likewise, sets of objects became more distinguishable as their numerical separation increased. These two effects, known as the size and distance effects, respectively, were both clearly evident in our results and suggest that coherent, object-specific oscillations can encode numerosity in a manner consistent with behavioral data.

Previous attempts to account for an innate number sense have used one-dimensional neural networks based on arrays of position and size filters (Dehaene & Changeux, 1993). However, to implement an analogous two-dimensional model, large numbers of additional shape and orientation filters, representing all possible 2D objects, would presumably be required. Coherent, object-specific oscillations, on the other hand, can in principle represent numerosity regardless of the shape, size, or locations of the
individual visual items. Our proposed mechanism for encoding numerosity, based on normalized gamma activity, thus eliminates the need to allocate large numbers of additional neurons in order to represent every possible object to be enumerated. Other modeling efforts have shown how number-selective responses might be learned de novo by employing an appropriate...
plasticity rule (Verguts & Fias, 2004), but these experiments assumed as input a “location map” in which objects were represented a priori as single points, leaving open the question of how numerosity detectors might be constructed from low-level, retinotopic inputs. A completely different approach to an innate number sense is represented by accumulator models, which posit that objects are enumerated sequentially, with the tally being summed into a noisy short-term buffer (Gallistel & Gelman, 2000). Both behavioral and electrophysiological experiments, however, have cast doubt on accumulator models (Nieder, 2005). The response latencies of number-selective cortical neurons, as well as those measured behaviorally, suggest that numerosity is processed in parallel rather than sequentially. Moreover, the response curves recorded from number-selective cortical neurons suggest that numerosity is represented by a population code distributed across multiple broadly tuned elements.

Our findings suggest that coherent oscillations encode numerosity in a form that can be extracted from multiunit spike trains over physiologically meaningful timescales. One advantage of using normalized gamma activity for encoding numerosity is that the complexity of the model does not increase as a function of either the number of objects to be detected or the number of possible combinations of object size, shape, or distribution. By employing numerosity detectors based purely on the normalized gamma activity, we could naturally take advantage of the translational and rotational invariance of low-level visual circuitry, thus automatically generalizing over a wide range of nonnumerical properties. The issue of how numerosity information encoded by coherent oscillations might be extracted by downstream elements was not addressed in this study, although it has been suggested that increased gamma activity or synchrony could differentially activate target structures (Alonso, Usrey, & Reid, 1996; Kenyon, Fetz, & Puff, 1990; Usrey, Alonso, & Reid, 2000). In addition, known physiological mechanisms such as facilitating synapses (Beierlein &
Connors, 2002), nonlinear dendritic processing (Polsky et al., 2004), and dynamical resonances (Miller et al., 2006; Frishman et al., 1987; Rager & Singer, 1998; Vigh, Solessio, Morgans, & Lasater, 2003) might confer a differential sensitivity to oscillatory input.

Innate numerical competency extends well beyond the few examples considered here. Rather than being limited to simultaneously presented visual items, the number sense characterized extensively in animals and humans encompasses multiple sensory and motor modalities (Dehaene, 1997), such as the ability to respond to a particular number of tones (Meck & Church, 1983) or to generate a specified number of bar presses (Mechner & Guevrekian, 1962). Coherent oscillations arising early in the visual processing pathway can at best account for only a narrow aspect of this much wider phenomenon. Nonetheless, we may ask to what extent the mechanisms illustrated by our restricted model might be expanded to include a broader range of behaviors.

Although our results were obtained using a model of the inner retina, gamma-band oscillations of cortical origin (Castelo-Branco et al., 1998; Gray et al., 1989; Gray & Singer, 1989; Kreiter & Singer, 1996; Singer & Gray, 1995) could likely encode similar, and indeed much higher-quality, numerosity information. The superior processing resources available to the cerebral cortex could presumably generate coherent oscillations that were much more object specific. In particular, by utilizing additional segmentation cues, such as shape, texture, color, depth, and motion, cortically generated coherent oscillations might encode numerosity much more reliably than was true of our comparatively simple feedback circuit. From this point of view, the much higher frequencies of the coherent oscillations arising in the retina versus those arising in the cerebral cortex (Castelo-Branco et al., 1998) suggest a possible strategy for distinguishing between numerosity information encoded at these two very distinct processing stages. Although their status remains speculative, theoretical models have also been proposed to explain how coherent oscillations could be used to represent temporal sequences (Hopfield & Brody, 2001), suggesting a possible mechanism for extending these findings to numerosity discrimination tasks involving nonsimultaneous stimuli. Oscillatory responses have been observed in a variety of brain regions, including auditory (Brosch, Budinger, & Scheich, 2002), motor (Murthy & Fetz, 1996a, 1996b), olfactory (Eeckman & Freeman, 1990), and somatosensory (Jones & Barth, 1997) cortical areas. The principles hypothesized to underlie the responses of number-selective neurons to visual inputs may therefore apply to the encoding of numerosity in other sensory modalities as well.

Acknowledgments

We acknowledge technical assistance from Mark Flynn, Chris Wood, James Theiler, David Marshak, John George, Reid Porter, and Bryan Travis. This
work was supported in part by the Department of Energy Office of Nonproliferation Research and Engineering.

References


