

Natural versus Synthetic Stimuli for Estimating Receptive Field Models: A Comparison of Predictive Robustness

Vargha Talebi and Curtis L. Baker Jr

McGill University, McGill Vision Research Unit, Department of Ophthalmology, Montreal, Quebec H3A 1A1, Canada

An ultimate goal of visual neuroscience is to understand the neural encoding of complex, everyday scenes. Yet most of our knowledge of neuronal receptive fields has come from studies using simple artificial stimuli (e.g., bars, gratings) that may fail to reveal the full nature of a neuron's actual response properties. Our goal was to compare the utility of artificial and natural stimuli for estimating receptive field (RF) models. Using extracellular recordings from simple type cells in cat A18, we acquired responses to three types of broadband stimulus ensembles: two widely used artificial patterns (white noise and short bars), and natural images. We used a primary dataset to estimate the spatiotemporal receptive field (STRF) with two hold-back datasets for regularization and validation. STRFs were estimated using an iterative regression algorithm with regularization and subsequently fit with a zero-memory nonlinearity. Each RF model (STRF and zero-memory nonlinearity) was then used in simulations to predict responses to the same stimulus type used to estimate it, as well as to other broadband stimuli and sinewave gratings. White noise stimuli often elicited poor responses leading to noisy RF estimates, while short bars and natural image stimuli were more successful in driving A18 neurons and producing clear RF estimates with strong predictive ability. Natural image-derived RF models were the most robust at predicting responses to other broadband stimulus ensembles that were not used in their estimation and also provided good predictions of tuning curves for sinewave gratings.

Introduction

A primary objective in visual neurophysiology is to create neuronal models with sufficient generality to predict responses to arbitrary stimuli (Rust and Movshon, 2005). However, a fundamental outstanding question is what kind of stimulus is most appropriate for the creation of such models (Carandini et al., 2005). Much of what we know about visual processing has been obtained from responses to simple artificial stimuli such as bars and sinusoidal grating patterns that rarely occur in the natural environment in which visual systems evolved (Felsen and Dan, 2005). Natural images are statistically much richer, functionally more relevant, and in some cases more effective for driving visual cortex neurons, thereby potentially revealing more complex underlying visual mechanisms.

System identification methods provide quantitative functional models describing how sensory neurons integrate signals from different receptive field (RF) locations and times to generate a response. Most common is reverse correlation (Ringach and Shapley, 2004), which has been used to map the spatiotemporal

receptive field (STRF) of sensory neurons. However reverse correlation has major drawbacks stemming from its requirement of spectrally white stimuli; white noise can sometimes be ineffective for driving visual neurons, particularly in later processing stages (Alonso and Martinez, 1998; Felsen and Dan, 2005), and it lacks the rich features found in natural images, potentially leaving complex RF properties uncharacterized. Recent studies have begun to use photographs of the natural environment as stimuli for system identification (Ringach et al., 2002; Touryan et al., 2005).

Here we assess models derived from system identification using synthetic and natural stimuli in three ways. First, how good is the models' predictive ability for the same stimuli used to estimate them? Second, how robust are these models when used to predict neuronal responses to other stimuli? Third, how well do these models characterize a cell's optimal tuning properties to sinewave gratings? To address these questions, we have employed a regression algorithm with regularization (Theunissen et al., 2001; Wu et al., 2006; see also <http://strflab.berkeley.edu>) to estimate the full three-dimensional (3D) STRF for linear–nonlinear (LN) models of simple type cells in A18 of the cat. White noise has been widely used and is theoretically optimal (Marmarelis and Marmarelis, 1978). Short bars are quasi-white stimuli that better drive neuronal responses (DeAngelis et al., 1993a). Natural image stimuli allow us to estimate models under more realistic conditions (Willmore et al., 2010). LN models of STRFs are conceptually intuitive and provide a compact description of simple cells and a common ground for comparison of different approaches. Secondary visual cortex is a more complex, intermediate-level processing stage with neurons selective for complex patterns (Hegde and Van Essen, 2000; Baker and Mareschal, 2001), thus providing a more demanding test of these different approaches.

Received Sept. 12, 2011; revised Oct. 25, 2011; accepted Nov. 21, 2011.

Author contributions: V.T. and C.L.B. designed research; V.T. performed research; V.T. analyzed data; V.T. and C.L.B. wrote the paper.

This work was supported by Canadian Institutes of Health Research Grant MA 9685 to Curtis L. Baker Jr. and Natural Sciences and Engineering Research Council of Canada Scholarship D2-348949-2007 to V.T. Special thanks to Michael Oliver of the J. Gallant Laboratory at the University of California at Berkeley, Guangxing Li of the McGill Vision Research Unit, and Linda Domazet.

The authors declare no competing financial interests.

Correspondence should be addressed to Vargha Talebi, McGill University, McGill Vision Research Unit, Department of Ophthalmology, 687 Pine Avenue West, Room H4-14, Montreal, QC H3A 1A1, Canada. E-mail: vargha.talebi@mail.mcgill.ca

DOI:10.1523/JNEUROSCI.4661-12.2012

Copyright © 2012 the authors 0270-6474/12/321560-17\$15.00/0

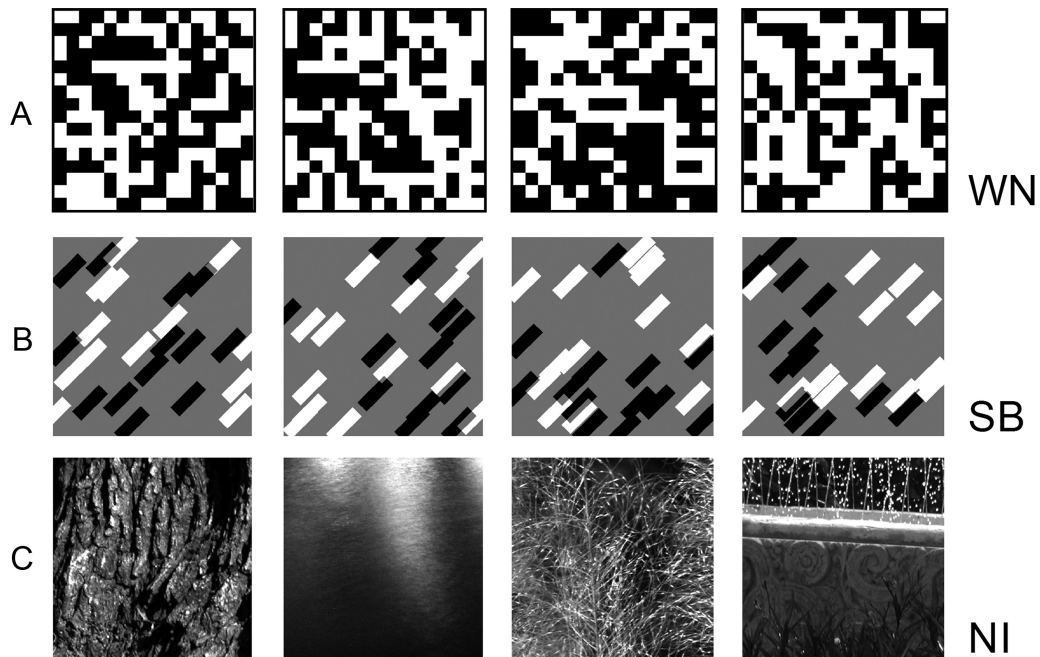


Figure 1. Example images from broadband stimulus ensembles. Shown are four frames of each type of broadband stimulus used in this study. **A**, Dense binary white noise (WN). **B**, Sparsely structured short bars (SB). **C**, Natural images (NI).

We demonstrate that short bars and natural images are more successful than white noise in producing clear RF estimates with good predictive power. Models derived from natural images, however, generalize better to other stimulus types as well as provide good predictions of tuning curves for sine-wave gratings.

Materials and Methods

Animal preparation. Anesthesia was induced by isoflurane/oxygen 3–5% inhalation, followed by intravenous (i.v.) cannulation and bolus i.v. injection of thiopentone sodium (8 mg/kg) or propofol (5 mg/kg). Surgical anesthesia was maintained with supplemental doses of thiopentone sodium or propofol as required. Atropine sulfate (0.05 mg/kg i.v.) or glycopyrrolate [30 μ g intramuscular (i.m.)] and dexamethasone (0.2 mg/kg i.v. or 1.8 mg i.m.) were administered, and a tracheal cannula or intubation tube was inserted. A craniotomy (A3/L4) over cortical area 18 (Tusa et al., 1979) was performed, followed by a small durotomy. The cortical surface was protected with 2% agarose (Sigma, type 1-A) capped with petroleum jelly. Local injections of bupivacaine (0.50%), a long lasting anesthetic, were administered at all surgical sites. Throughout the surgical procedure, body temperature was thermostatically maintained at 37°C, and heart rate was monitored (Vet/Ox Plus 4700; Heska).

After completion of surgery, animals were connected to a respirator (Ugo Basile 6025) and paralyzed with a bolus i.v. injection of gallamine triethiodide (to effect), followed by infusion (10 mg \cdot kg⁻¹ \cdot h⁻¹). Sodium pentobarbital (1.0 mg \cdot kg⁻¹ \cdot h⁻¹), or in later experiments propofol (5.3 mg \cdot kg⁻¹ \cdot h⁻¹), was supplemented with fentanyl citrate (7.4 μ g \cdot kg⁻¹ \cdot h⁻¹) following a bolus injection (2.5 μ g/kg). Both anesthesia regimes were further supplemented with oxygen/nitrous oxide (70:30) and a continuous infusion of lactated dextrose-saline (2 ml/h i.v.) was supplied. Expired CO₂, EEG, EKG, body temperature, blood oxygen, heart rate, and airway pressure were monitored and maintained at appropriate levels.

Corneas were initially protected with topical carboxymethylcellulose (1%) and subsequently with neutral contact lenses. Spectacle lenses, selected with slit retinoscopy, were used to bring objects at a distance of 57 cm into focus. Artificial pupils (2.5 mm) were placed in front of the eyes. The area centralis was determined by back projection of the optic disk onto a tangent screen (Nikara et al., 1968; Fernald and Chase, 1971).

Daily maintenance included topical atropine sulfate (1%) and phenylephrine hydrochloride (2.5%) to dilate the pupils and retract the nictitating membrane respectively, as well as glycopyrrolate (16 μ g) and dexamethasone (1.8 mg) administered i.m. All animal procedures were approved by the McGill University Animal Care Committee and are in accordance with the guidelines of the Canadian Council on Animal Care.

Visual stimuli. Visual stimuli were generated on a Macintosh computer (MacPro, 2.66 GHz Quad Core Intel Xeon, 6 GB, NVIDIA GeForce GT 120) using custom software written in Matlab (MathWorks) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). A CRT monitor (NEC FP1350, 20", 640 \times 480 pixels, 75 Hz, 36 cd/m²) placed at a viewing distance of 57 cm was used to display the stimuli. The monitor's gamma nonlinearity was measured with a photometer (United Detector Technology) and corrected with inverse lookup tables.

Drifting sinewave gratings (3 Hz, 30% contrast) were presented within a cosine-tapered circular window against a uniform background at the mean luminance of the pattern. The same mean luminance was also maintained during intervals between stimuli and presented as blank conditions for measurement of spontaneous activity.

Three types of broadband stimulus patterns were employed for system identification: white noise, short bars, and natural images (Fig. 1). "White noise" stimuli (Fig. 1A) were dense noise patterns with random equiprobable black and white checks. "Short bar" stimuli (Fig. 1B) consisted of sparse equiprobable white and black bars (3:1 aspect ratio) placed randomly without constraint on a background of mean luminance. The bar density was chosen so that \sim 20% of the image area was filled with bars. "Natural image" stimuli (Fig. 1C) were constructed from high-quality digital photographs (McGill Calibrated Color Image Database; Olmos and Kingdom, 2004), each of which was converted to monochrome and divided into 480 \times 480 pixel images. Nearly blank images (e.g., sky, water) were rejected by setting a root mean square (RMS) energy threshold (0.03 standard deviation of pixel values). Remaining images were RMS-normalized with mean luminance removed. In each case, a stimulus ensemble consisted of fresh independent patterns in each frame, comprising sets of 375 images that were presented as 5 s movies. Each stimulus presentation was preceded by a 1 s mean luminance blank screen for measurement of spontaneous activity. Stimulus images were displayed in a 480 \times 480 pixel window and randomly changed on each

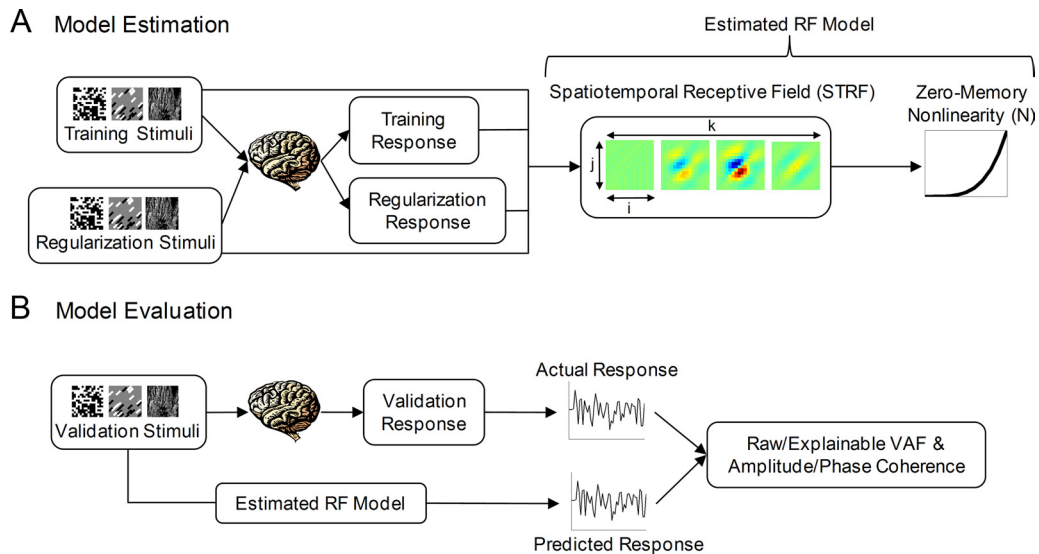


Figure 2. System identification procedure. Neural responses of a simple type cell to three independent ensembles of a given broadband stimulus type were used to estimate receptive field models and evaluate their predictive abilities. **A**, Model estimation. Training and regularization datasets were used to estimate a 3D (space – i , space – j , time – k) spatiotemporal receptive field, STRF. A subsequent zero-memory nonlinearity (N) was fit from comparison with measured responses. The STRF and N together make up the estimated RF model. **B**, Model evaluation. The estimated RF model's response to the validation stimuli generate a predicted response, which was compared with the actual (measured) validation response. The quality of prediction was quantified with raw/explainable VAF and amplitude/phase coherence analyses.

frame refresh, with mean luminance maintained between presentations. Other types of stimuli were also pilot tested, including space–time correlated noise and “cat cam” movies (Einhauser et al., 2002), but both these stimuli were relatively poor at driving A18 neurons and therefore abandoned.

Extracellular single unit recording. In early experiments, extracellular recordings were obtained using single-channel, glass-coated, platinum-iridium or parylene-coated tungsten microelectrodes (Frederick Haer), and in later experiments with silicon axial multi-electrodes (NeuroNexus A1×16) or multi-shank tetrodes (NeuroNexus A4×1-tet). Electrodes were advanced using a stepping motor microdrive (M. Walsh Electronics, uD-800A). A primary, single-channel recording pathway incorporated an audio monitor, a window discriminator (Frederick Haer) to isolate single units, and a delay-triggered oscilloscope to monitor isolation. Spike times were recorded at a resolution of 100 μ s (Instrutech, ITC-18) and time referenced to the stimulus using an optical photo sensor (TAOS T2L 12S) placed on the corner of the CRT monitor within which images contained stimulus timing information. Later experiments also incorporated a secondary, parallel, multi-channel recording pathway (Plexon Recorder, version 2.3), in which complete raw signals were acquired for all 16 channels at 40 kHz, and stored to hard disk for subsequent spike sorting and detailed analysis. One of the 16 channels was also routed to the primary recording pathway for online analysis to guide the recording protocol.

Manually controlled bar-shaped stimuli were used to assess the approximate location, orientation preference, and ocular dominance of isolated neurons. The CRT monitor was centered on the receptive field, and all subsequent stimuli were presented monocularly to the dominant eye. Neurons were first characterized with conventional tuning curve measurements using sinuswave grating patterns to determine optimal spatial frequency, orientation, and temporal frequency. The cell's receptive field was further localized by displaying small grating patches at a grid of spatial locations, and the monitor repositioned as necessary. Each of the three broadband stimulus types (white noise, short bars, or natural images) was then presented. The check size (white noise) and the bar width (short bars) were set to $1/4$ to $1/6$ of the spatial period of the neuron's optimal grating, and bar orientation was set to the neuron's optimal grating orientation (DeAngelis et al., 1993a). For each broadband stimulus type, three independent datasets were collected for training, regularization, and validation (Fig. 2A,B), each requiring about 20–30 min to acquire. The training stimuli consisted of 20 image ensem-

bles (total of 7500 unique images) repeated five times. The regularization and validation stimuli each consisted of five image ensembles (1875 unique images each) repeated 20 times. This trade-off of stimulus diversity versus repetitions was aimed at maximizing the informativeness provided by unique images (training), but minimizing response variance (regularization, validation). This procedure was repeated for all three stimulus types whenever it was possible to maintain isolation for sufficient time.

A total of 281 datasets were collected from 73 neurons in 13 animals of either sex. Other data were also collected from these animals as part of several on-going projects in the same laboratory. Cells having average spike frequencies of <1 spike/s in response to the stimulus image ensemble were considered to be unresponsive.

Data analysis. In experiments with records of raw broadband responses, signals were reanalyzed *post hoc* to extract spike waveforms, which were carefully classified using Plexon Offline Sorter (Plexon, version 2.8.8) software. Conservative thresholds were set for clear separation of distinct signals to ensure that analyses were performed only on spikes from single neurons. All spike time data, whether from single-channel or multi-channel recordings, were then analyzed in a common manner with custom software written in Matlab (MathWorks). Gradient descent and regularization (see below in this section) were implemented using functions from the STRFlab ToolBox (<http://strflab.berkeley.edu>).

Responses to sinuswave gratings were analyzed conventionally to produce tuning curves of average spike frequency as a function of spatial frequency and orientation. Spatial frequency tuning curves were fit with a Gaussian function to yield estimates of optimal frequency, bandwidth, and response amplitude (DeAngelis et al., 1994):

$$R(sf) = ke^{-\left(\frac{sf-SF_{opt}}{\alpha}\right)^2} + R_0, \quad (1)$$

where k = maximum response amplitude; sf = measured spatial frequency; SF_{opt} = optimal spatial frequency; 1.65α = tuning bandwidth in octaves; R_0 = spontaneous response; and $R(sf)$ = fitted response as a function of spatial frequency.

Orientation tuning curves were characterized by a vector-based summation method to indicate optimal orientation, optimal direction, orientation bias, and direction bias (Worgotter and Eysel, 1987; Leventhal et al., 2003):

$$OB = \frac{\sum_k R_k e^{i2\theta_k}}{\sum_k R_k} \quad (2)$$

$$DB = \frac{\sum_k R_k e^{i\theta_k}}{\sum_k R_k}, \quad (3)$$

where R_k is the response at stimulus orientation θ_k and OB and DB are complex values whose magnitudes ($|OB|$, $|DB|$) are the orientation bias and direction bias respectively. $|OB|$ and $|DB|$ have a bounded range between 0 and 1 (dimensionless), where 0 indicates no orientation/direction selectivity and 1 represents absolute selectivity. Cells with orientation/direction bias values greater than 0.1 are considered to be orientation/direction sensitive (Leventhal et al., 1995). The angles of OB and DB provide the optimal orientation and optimal direction, respectively. Neurons were classified as simple or complex type cells on the basis of poststimulus time histogram (PSTH) modulation by an optimal grating (Skottun et al., 1991), and only simple type cells were used for subsequent analysis.

For responses to image ensembles, spike times were collected into PSTHs binned at the stimulus refresh rate (bin width of 13.3 ms) to create analog responses that were then averaged across repetitions. To reduce the number of estimated parameters in the STRF, images within each ensemble were spatially downsampled in proportion to the stimulus check size or bar width to yield image sizes typically in the range of 12^2 to 24^2 .

Each neuron's response to an image ensemble was estimated within the framework of a generalized linear model (GLM) (Fig. 2A), consisting of a linear STRF followed by a zero-memory nonlinearity:

$$w(t) = \sum_{i=1}^M \sum_{j=1}^M \sum_{k=1}^8 h(i, j, k) s(i, j, k - t) \quad (4)$$

$$r(t) = |w^a(t)|^+ + n(t), \quad (5)$$

where $h(i, j, k)$ = linear filter (STRF) "weight" of i, j th pixel and k th time index; $s(i, j, k)$ = stimulus image at i, j th pixel and k th time index; i, j = indices of (downsampled) pixels, typically ranging from 1 to M ; k = time (lag) index, ranging from 0 to 7; M = downsampled image size, typically circa 12–24; $w(t)$ = response of linear filter as a function of time (t); $[\dots]^+$ denotes half-wave rectification; a = exponent of power law nonlinearity; $n(t)$ = noise; and $r(t)$ = model response as function of time (t).

The spatiotemporal filter weights h_{ijk} were optimized with iterative gradient descent to minimize the mean square error between the responses $r(t)$ of the model filter and those measured from the neuron, using the entire training dataset on each iteration. GLMs are guaranteed to have a unique global minimum (i.e., convex problem) (McCullagh and Nelder, 1989), making them amenable to gradient descent optimization methods.

Since the number of parameters being fit (e.g., $16 \times 16 \times 8 = 2048$) is on the order of the number of data points ($375 \times 20 = 7500$) and the noise $n(t)$ is not negligible, this optimization can lead to the fitted filter weights in part reflecting the particular noise in the training dataset rather than the system function. This "overfitting" can be circumvented by employing regularized methods that incorporate a constraint or a priori assumption typically through the use of a penalty function that discourages high-valued coefficients or through enforcing priors such as smoothness or sparseness (Willmore and Smyth, 2003; Wu et al., 2006). Recent studies have successfully used a number of different forms of regularized methods including ridge regression (Machens et al., 2004), Tikhonov-Miller regularization (Smyth et al., 2003), and early stopping (Willmore et al., 2010) to reconstruct RFs of sensory neurons. We implemented regularization here through early stopping (Hagiwara, 2002), which halts the gradient descent algorithm prematurely before it begins fitting to the noise. To achieve this early stopping, the model estimate from each iteration of the gradient descent was tested for its predictive ability on the regularization dataset—when the prediction error ceased to decline and started to increase, the gradient descent algorithm was halted. Such early

stopping regularization acts to avoid fitting unnecessarily large values to the filter weight parameters h_{ijk} (Hagiwara, 2002), resulting in estimated receptive fields that look less "noisy" while having better predictive ability on novel datasets. Early stopping has proven highly effective in machine learning and neural network modeling (Bishop, 2006).

The power law parameter (a) of the zero-memory nonlinearity was then fit using a simplex algorithm (Nelder-Mead method—Matlab's *fminsearch*), between the measured neuronal responses (training dataset) and their predicted values based on convolution of the STRF with the training image ensembles. For some datasets in preliminary analyses, other nonlinear functions (e.g., Naka-Rushton, third-order polynomial) were also fit but did not provide significant improvements.

For comparison with traditional approaches, we also analyzed response to white noise stimulus ensembles using conventional reverse correlation (Marmarelis and Marmarelis, 1978). All aspects of the reverse correlation analysis (i.e., binning, averaging across repetitions, down-sampling, zero-memory nonlinearity estimation) were the same as those used for the GLM approach. Note that the results from reverse correlation with white noise stimuli should be identical to those from the gradient descent, iterative regression analysis (without regularization), if the system noise, $n(t)$, is negligible and the dataset is sufficiently large.

To measure the estimated RF model's predictive ability for novel stimulus ensembles, the validation dataset was used to compare with a predicted response based on simulated responses of the final estimated LN model, i.e., convolution with the STRF, followed by the zero-memory nonlinearity (Fig. 2B). The predictive accuracy was quantified as "variance accounted for" (VAF), calculated as the square of the correlation coefficient (R), expressed as a percentage:

$$R = \frac{\sum [(x - \bar{x})(y - \bar{y})]}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} \quad (6)$$

$$VAF = [100(R)^2], \quad (7)$$

where x, y = actual and predicted responses, respectively; \bar{x}, \bar{y} = mean of the actual and predicted responses, respectively; and R = correlation coefficient. The VAF is the percentage of variance in the actual measured response that is accounted for in the predicted response.

We often observed that the estimated models more accurately predicted the timing of a response (i.e., phase) than its amplitude, as can be seen in the example of Figure 3A. To explore this phenomenon, we measured amplitude and phase coherence (Drongelen, 2007), which essentially splits the VAF into amplitude and phase components, measured as a function of temporal frequency (ω). These quantities were calculated by taking the average cross-spectrum between actual and predicted responses (S_{xy}) normalized by the square root of the average power spectra for actual (S_{xx}) and predicted (S_{yy}) responses:

$$C(\omega) = \frac{\langle S_{xy}(\omega) \rangle}{\sqrt{\langle S_{xx}(\omega) \rangle \langle S_{yy}(\omega) \rangle}}, \quad (8)$$

where $C(\omega)$ is a complex-valued function of temporal frequency ω , its magnitude is the amplitude coherence, and its angle is the phase coherence ($\langle \dots \rangle$ denotes average across trials).

The amplitude coherence ranges from zero to an ideal value of unity. The phase coherence normally ranges from 0° (ideal) to 90° (random) but can have values up to 180° (anti-phase). The example in Figure 3B shows amplitude/phase coherence values represented in a polar plot in which the plotted points are amplitude/phase coherence values at different temporal frequencies. These plots are further summarized by two indices: the average amplitude coherence ("amp COH") and the average phase coherence ("phase COH"). Amp COH and phase COH are computed by averaging the amplitude and phase (collapsed to 0 – 180°) coherence functions across temporal frequencies.

In practice, the above "raw" VAF is never 100% for two reasons: first, the neural responses are very noisy; second, the LN model is undoubtedly inadequate due to other neural nonlinearities not instantiated in the model architecture. We attempt to disambiguate these two sources of reduced VAF by the method of David and Gallant (2005), illustrated in Figure 3, C and D.

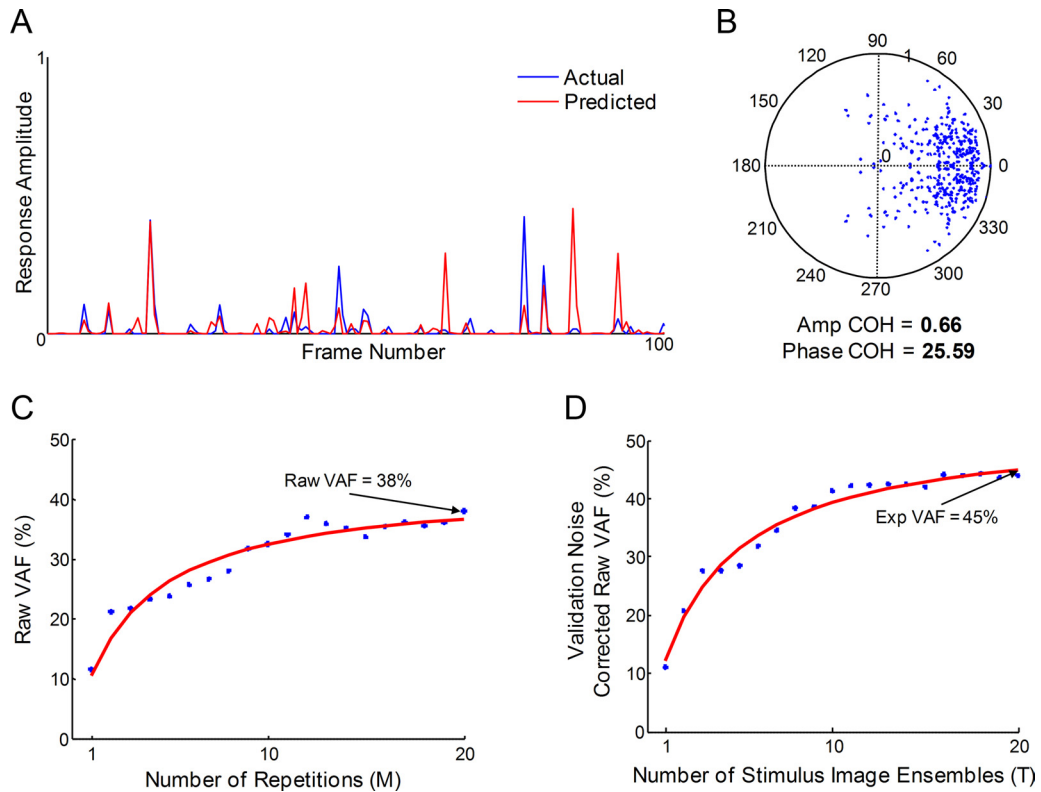


Figure 3. Quantifying indices for predictive power. Accuracy of estimated RF models was assessed using measures of raw/explainable VAF and amplitude/phase coherence. **A**, Example of actual (blue) and predicted (red) neuronal response amplitudes graphed against stimulus frame number. It is apparent that the estimated RF model does a better job of predicting when responses occur than predicting their amplitudes. **B**, Amplitude and phase coherence values at different temporal frequencies displayed as a polar plot. Better RF model performance yields more points clustered around ideal values of 1 for amplitude and 0° for phase. Average amplitude and phase coherence values across temporal frequencies, amp COH and phase COH, were used as summary statistics. Amp COH is a dimensionless quantity between 0 and 1, while phase COH ranges between 0° and 180°. **C**, The noise ceiling in the validation dataset was determined by calculating the VAF between predicted and actual responses, as the number of repetitions used in the actual response was increased. The results were fit with a smooth curve (solid red line) and the final plotted point taken as the raw VAF (38%), i.e., when the actual response was averaged across all repetitions. **D**, The noise ceiling in the training dataset was determined by calculating the VAF between actual and predicted responses, as the number of stimulus image ensembles used in the STRF estimation was increased. The explainable VAF (45%) is defined as the plateau value of the fitted curve (solid red line).

A “noise ceiling” is determined by incrementally increasing the amount of data used to validate and train the STRF. The noise ceiling in the validation dataset is determined by calculating the raw VAF between the predicted and actual responses averaged across an increasing number of repetitions. The resulting points (Fig. 3C) are fit with a smooth asymptotic curve (solid red line):

$$\frac{1}{R^2} = \frac{1}{R_{\max}^2} + \frac{A}{M}, \tag{9}$$

where R_{\max}^2 = the fraction of variance that would be explained if there were no noise in the validation dataset; A = constant, reflecting trial-to-trial variability; M = number of repetitions; and R^2 = the squared correlation coefficient (i.e., raw VAF) between actual and predicted responses. The final plotted point corresponds to the raw VAF (38%) when the actual response is averaged across all repetitions.

The noise ceiling in the training set is determined in the opposite fashion by calculating the raw VAF between the actual and predicted responses from STRFs estimated using increasing numbers of image ensembles. These training VAFs are then corrected for the amount of noise in the validation dataset by solving Equation 9 for R_{\max}^2 and using the fitted value of A . The resulting validation noise-corrected VAFs (i.e., R_{valcorr}^2) are plotted (Fig. 3D) and fit with the smooth asymptotic curve (solid red line):

$$\frac{1}{R_{\text{valcorr}}^2} = \frac{1}{R_{\text{ideal}}^2} + \frac{B}{T}, \tag{10}$$

where R_{ideal}^2 = ideal squared correlation coefficient value if there were no corrupting noise (i.e., “explainable” VAF); B = constant, reflecting trial-to-

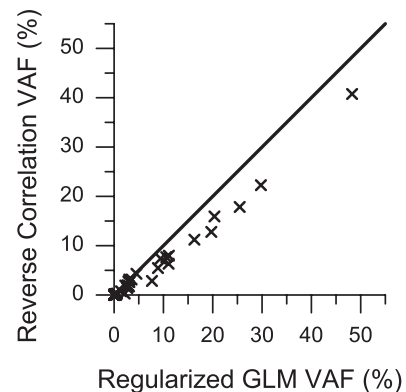


Figure 4. VAFs for reverse correlation and regularized GLM. Comparing reverse correlation and regularized GLM system identification methods on RF models derived from white noise. Points lie below the 1:1 equality line (diagonal), indicating that regularized GLM provides RF models with better predictive performance.

trial variability; T = number of stimulus image ensembles; and R_{valcorr}^2 = squared correlation coefficient value from the training dataset that has been corrected for validation dataset noise. The explainable VAF (exp VAF) (45%) is taken as the fitted plateau value (R_{ideal}^2) of the curve. Thus, the explainable VAF provides an estimate of the fraction of total response variance that could theoretically be predicted in the absence of neuronal noise (David and Gallant, 2005).

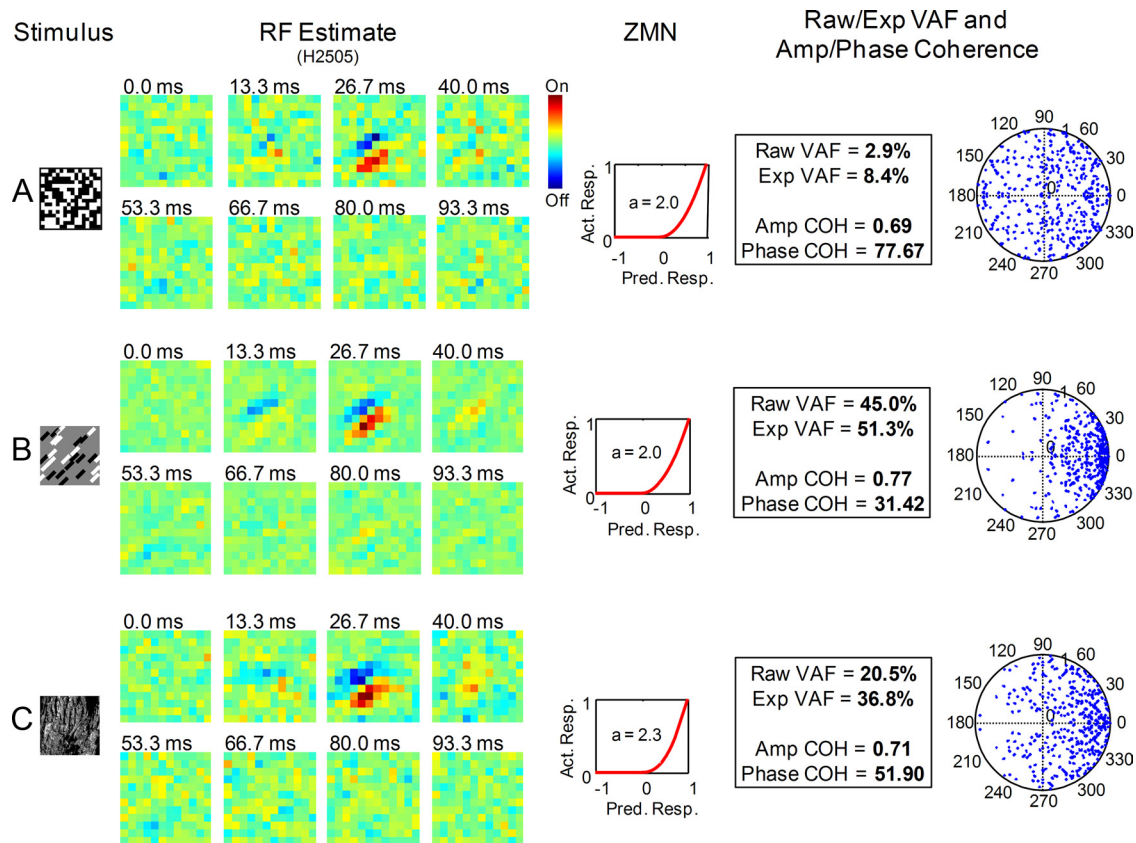


Figure 5. Example RF estimates. RF estimates for a simple cell derived using three stimulus types. Each row depicts results for one stimulus type showing, from left to right, the following: an example stimulus image; a spatiotemporal RF estimate across 8 time lags (0.0 ms to 93.3 ms); fitted zero-memory nonlinearity, ZMN, of the actual response (Act. Resp.) against the predicted response (Pred. Resp.) with the exponent of the power law nonlinearity (a); raw/explainable VAFs and amplitude/phase COHs to an independent validation dataset; and polar plots of amplitude/phase coherence values at different temporal frequencies. Ideal values of amplitude and phase coherence are 1 and 0° , respectively. **A**, White noise stimuli result in a RF estimate with flanking elongated ON and OFF regions (red and blue areas, respectively), with a right oblique orientation and phase progression across successive time lags. However, the RF estimate is noisy-looking with low VAFs (raw VAF = 2.9%, exp VAF = 8.4%) and amplitude/phase coherence values that are highly scattered (amp COH = 0.69, phase COH = 77.67). **B**, Short bar stimuli result in a very clear, similarly structured RF estimate with high VAFs (raw VAF = 45.0%, exp VAF = 51.3%) and amplitude/phase coherence values clustering near optimal (amp COH = 0.77, phase COH = 31.42). **C**, Natural image stimuli result in a RF estimate that is noisier than for short bars, but the spatial structure of the ON and OFF regions is clear, the VAFs are reasonable (raw VAF = 20.5%, exp VAF = 36.8%), and the amplitude/phase coherence values are somewhat clustered (amp COH = 0.71, phase COH = 51.90). For all three stimulus types, the power law nonlinearity is expansive, with an exponent ranging from 2.0 to 2.3.

Each RF model was used in simulations to predict responses to the same stimulus image ensemble used to estimate it, as well as to the other stimulus types. For example, a RF model derived from white noise stimuli was used to predict responses to another ensemble of white noise, as well as to ensembles of short bars and natural images. VAFs were computed between the simulated predictions from the three stimulus types and the actual measured responses in the respective validation datasets. Comparison of these VAFs indicated how well the RF models derived from the different stimulus types could generalize to other stimuli that were not used to estimate them.

Further simulations were conducted to assess how well the estimated RF models could predict grating tuning curves for spatial frequency and orientation. Raw VAFs were computed between the simulated tuning curve predictions and the actual measured tuning curves. In addition, characteristic parameters were calculated for each tuning curve (as described above) and systematically compared, i.e., optimal spatial frequency, bandwidth, response amplitude, optimal orientation, orientation bias, optimal direction, and direction bias.

Extensive validation tests, using both hardware and software models, were performed to ensure the accuracy of the RF model estimates and the quantifying metrics. A hardware FPGA model of a simple type visual cortex neuron (Li et al., 2010) was used to test the data acquisition system, stimulus presentation software, and data analysis programs. Responses collected from the FPGA model verified that our system identification software could correctly yield the spatial RF and the temporal latency. Using software-

simulated models of a noiseless simple cell with specified delays and the above GLM system identification, we were able to accurately extract the model's spatiotemporal filter at the model's simulated delay. Quantifying measures (i.e., raw/explainable VAF and amplitude/phase coherence) were validated by adding varying amounts of noise to the model's response. As expected, amplitude and phase coherence values were nominal when there was no noise and declined progressively with increasing noise. In a model with no noise, the raw and explainable VAFs were nearly the same, with values of $\sim 100\%$. When noise was added, the raw VAF dropped but the explainable VAF was maintained at nearly 100%.

Results

For comparison with traditional methods, neuronal responses to white noise were analyzed using reverse correlation as well as regularized GLM. Figure 4 compares VAFs for these two system identification methods in a scatter plot—each point represents results from a different neuron. Note that all the points lie below the 1:1 equality line, indicating that RF models from a regularized GLM approach yield better predictive ability than those from reverse correlation. Indeed, regularized methods have been shown to produce accurate estimates of RFs using fewer stimuli than other methods (Willmore and Smyth, 2003). All subsequent system identification analysis used regularized GLM, both for

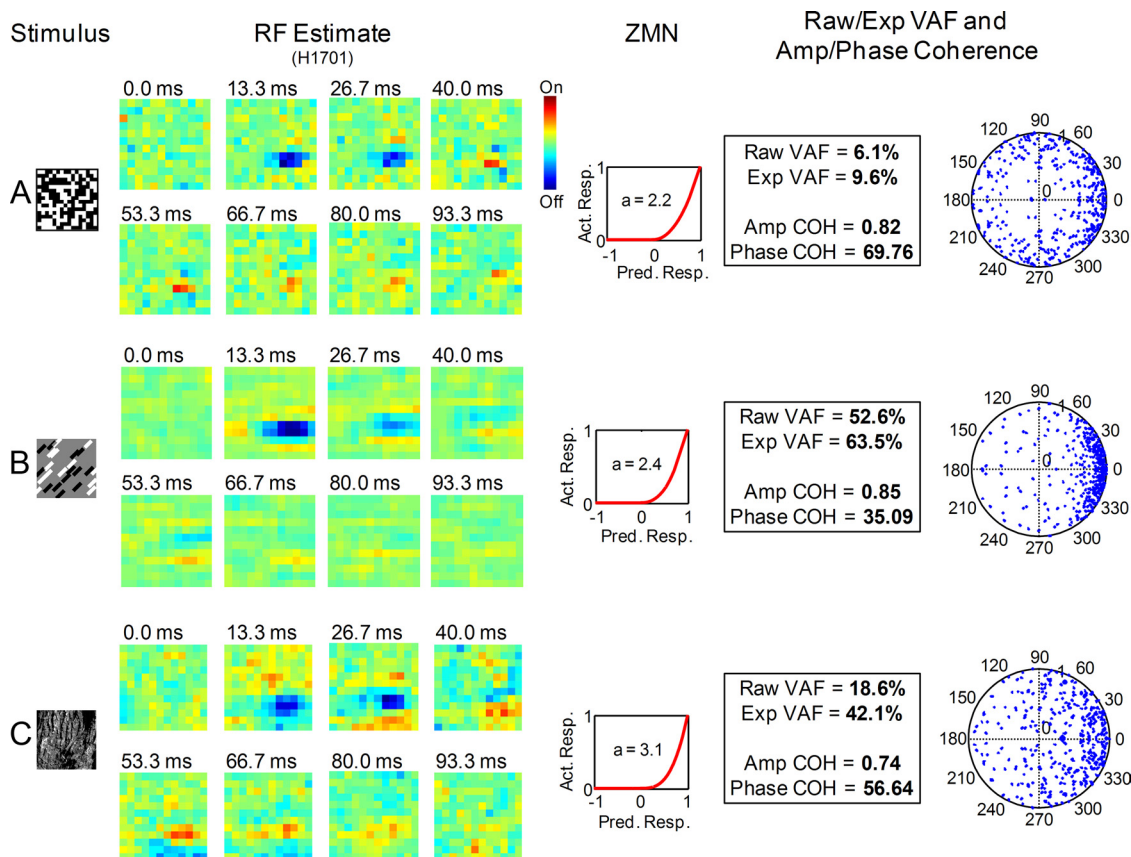


Figure 6. Example RF estimate as in Figure 5 for a neuron giving somewhat different RF estimates for different stimulus types. All RF estimates show a horizontally oriented OFF zone accompanied by weaker flanking ON regions and an upward phase progression across time lags. **A**, RF estimate derived from white noise has minimal flanking ON regions in comparison to the other stimuli and a noisy-looking structure with low VAFs (raw VAF = 6.1%, exp VAF = 9.6%). Amplitude and phase coherence values are scattered (amp COH = 0.82, phase COH = 69.76); however, a larger percentage of points lie at a greater radial distance from the origin, indicating that the amplitude could be reasonably predicted. **B**, RF estimate derived from short bars has slightly increased flanking ON regions and an elongated OFF response. It is also less noisy-looking with very high VAFs (raw VAF = 52.6%, exp VAF = 63.5%) and amplitude/phase coherence values clustering near optimal (amp COH = 0.85, phase COH = 35.09). Again notice the particularly high amplitude coherence values. **C**, RF estimate from natural images has the most apparent flanking ON regions. It is also noisy-looking with reduced raw VAF (18.6%) but a quite improved explainable VAF (42.1%). Amplitude/phase coherence values are somewhat clustered about optimal (amp COH = 0.74, phase COH = 56.64). For all three stimulus types, the power law nonlinearity is expansive with an exponent ranging from 2.2 to 3.1. Act. Resp., Actual response; Pred. Resp., predicted response.

this reason and because it provides valid results for stimuli that are spectrally not white.

RF estimates were derived from the three broadband stimulus types whenever possible and evaluated based on how well they could act as models to predict responses to an independently collected validation dataset, using raw/explainable VAFs and amplitude/phase coherence values as criteria. To assess the robustness of these RF models for other types of stimuli, they were subsequently used to predict responses to the other types of stimulus ensembles, as well as spatial frequency and orientation tuning for sinewave gratings.

Evaluation of receptive field estimates derived from different types of broadband stimuli

Examples of RF estimates and indices of their predictive performance are illustrated for four neurons in Figures 5–8, showing in each case results for white noise, short bars, and natural images in three rows. Each row shows, from left to right, an example stimulus image, spatial RF estimates at each of eight temporal lags, the fitted zero-memory nonlinearity (ZMN), and predictive performance indices (VAFs, COHs).

For the neuron in Figure 5, all three types of stimulus ensembles produce RF estimates with similar spatiotemporal structures: flanking elongated ON and OFF regions with a right-

oblique orientation and a phase progression across successive time lags. However short bars (Fig. 5B) give the cleanest-looking RF estimate, which also yields high VAFs (raw VAF = 45.0%, exp VAF = 51.3%) and amplitude/phase coherence values that cluster nearer the optimal values of unity and 0° (amp COH = 0.77, phase COH = 31.42). Natural images (Fig. 5C) produces RF estimates with moderate VAFs (raw VAF = 20.5%, exp VAF = 36.8%) and amplitude/phase coherence values that cluster about optimal (amp COH = 0.71, phase COH = 51.90). White noise (Fig. 5A) produces a noisy-looking RF estimate with very low VAFs (raw VAF = 2.9%, exp VAF = 8.4%) and highly scattered amplitude/phase coherence values (amp COH = 0.69, phase COH = 77.67).

Figure 6 shows results from a cell with a clear, horizontally oriented OFF zone accompanied by weaker flanking ON regions and an upward phase progression across temporal lags. However in this case the three kinds of stimuli produce RF estimates with varying types of structure. For example, flanking ON response regions are minimal for white noise (Fig. 6A), increase slightly for short bars (Fig. 6B), and are very apparent with natural images (Fig. 6C). Also, the OFF response estimated from short bars (Fig. 6B) seems to be more elongated than the others. Short bars (Fig. 6B) produces a clear RF estimate with high VAFs (raw VAF = 52.6%, exp VAF = 63.5%) and amplitude/phase coherence values that cluster

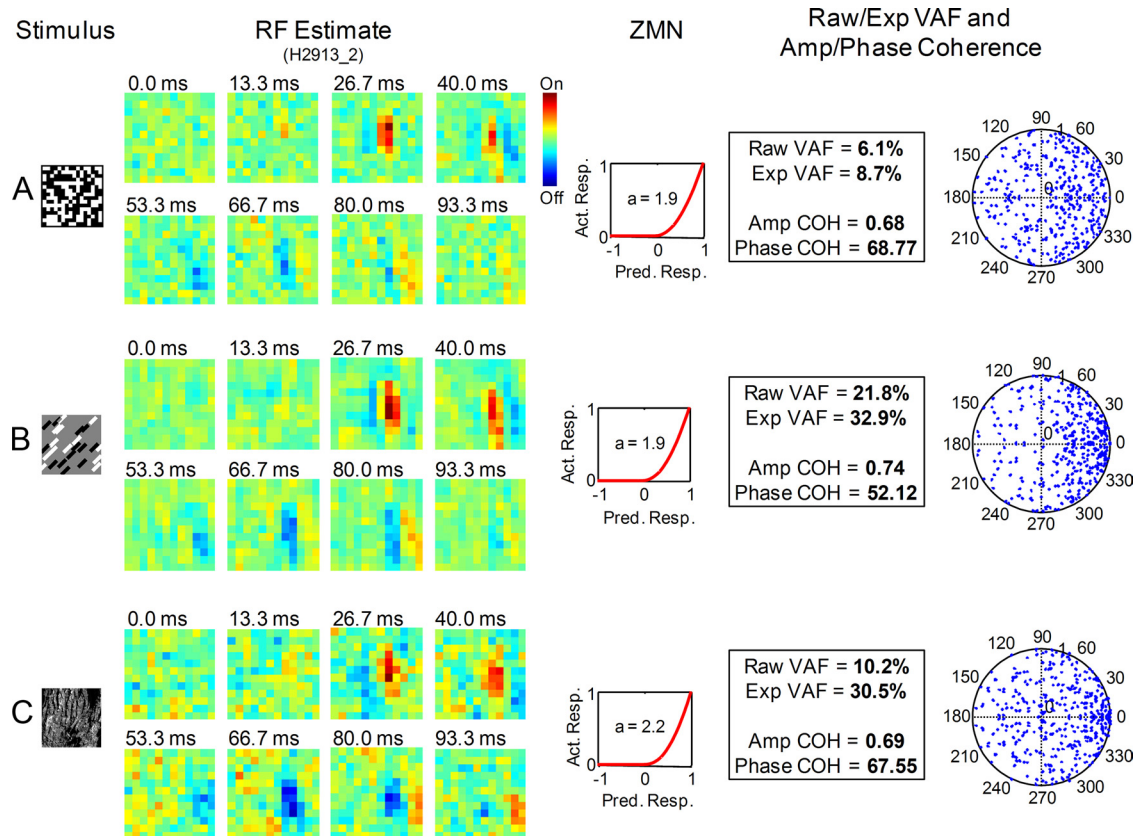


Figure 7. Example RF estimate as in Figure 5 for another neuron. All RF estimates have vertically oriented, flanking ON and OFF zones that show a phase reversal at later time lags. **A**, RF estimate derived from white noise performs poorly with very low VAFs (raw VAF = 6.1%, exp VAF = 8.7%) and scattered amplitude/phase values (amp COH = 0.68, phase COH = 68.77). **B**, RF estimate derived from short bars performs much better with higher VAFs (raw VAF = 21.8%, exp VAF = 32.9%) and amplitude/phase values that cluster loosely near optimal (amp COH = 0.74, phase COH = 52.12). **C**, RF estimate from natural images at first glance performs poorly (raw VAF = 10.2%), but when a noise ceiling is calculated the explainable VAF (30.5%) is nearly the same as that of short bars. Amplitude/phase coherence values show less clustering near optimal (amp COH = 0.69, phase COH = 67.55). For all three stimulus types, the power law nonlinearity is close to half-square, with an exponent ranging from 1.9 to 2.2. Act. Resp., Actual response; Pred. Resp., predicted response.

near optimal (amp COH = 0.85, phase COH = 35.09). Note the particularly high amplitude coherence values, indicating that in this example the prediction was good not only for the timing of this response but also its amplitude. Natural images (Fig. 6C) produce a noisy-looking RF estimate with reduced raw VAF (18.6%), but nevertheless the explainable VAF (42.1%) is quite good. Amplitude/phase coherence values are somewhat clustered about optimal (amp COH = 0.74, phase COH = 56.64), although not as well as for short bars. White noise (Fig. 6A) again performs poorly, with a noisy-looking RF estimate and very low VAFs (raw VAF = 6.1%, exp VAF = 9.6%). The amplitude/phase coherence values, however, are clustered at higher amplitude but very scattered phase (amp COH = 0.82, phase COH = 69.76), indicating that in this case the amplitude of the response could be more reliably predicted than the timing.

Figure 7 shows an example of a cell for which all three stimulus types produce RF estimates with vertically oriented, flanking ON and OFF zones that show a phase reversal at later time lags. Short bars (Fig. 7B) again produce the cleanest-looking RF estimate with reasonable VAFs (raw VAF = 21.8%, exp VAF = 32.9%) and amplitude/phase coherence values that cluster loosely about optimal (amp COH = 0.74, phase COH = 52.12). Natural images (Fig. 7C) produce a somewhat noisy RF estimate with a low raw VAF (10.2%) but a much improved explainable VAF (30.5%) that is nearly equal to that for short bars and amplitude/phase coherence values showing less clustering near optimal (amp COH = 0.69, phase COH = 67.55). White noise (Fig. 7A)

again performs poorly with a noisy RF estimate, very low VAFs (raw VAF = 6.1%, exp VAF = 8.7%), and scattered amplitude/phase coherence values (amp COH = 0.68, phase COH = 68.77).

For the somewhat unusual cell in Figure 8, all three stimulus types produce RF estimates with a similar-looking punctate structure that is lacking clearly oriented domains, although short bars reveal a slight left-oblique orientation. In each case, the RF is predominantly OFF at shorter time lags followed by reversal to ON at later time lags. Short bars (Fig. 8B) again produce the cleanest-looking RF estimate with very high VAFs (raw VAF = 60.5%, exp VAF = 62.4%) and amplitude/phase coherence values that cluster near optimal (amp COH = 0.76, phase COH = 14.13). Interestingly, in this case white noise (Fig. 8A) outperforms natural images (Fig. 8C) with higher VAFs (white noise: raw VAF = 53.1%, exp VAF = 55.2%; natural images: raw VAF = 43.4%, exp VAF = 50.1%) and more tightly clustered phase coherence values (white noise: phase COH = 17.81, natural images: phase COH = 36.84). Amp COH, however, is better predicted by natural images than by white noise, with values of 0.80 and 0.68, respectively.

Examining explainable VAFs over our sample population of simple cells (Fig. 9A), it is apparent that on average short bars yield the highest VAFs (exp VAF = 53%), followed by natural images (exp VAF = 41%) and finally white noise (exp VAF = 19%). However the average amp COH (Fig. 9B) values are about the same (~ 0.5) for all three stimulus types, indicating that approximately half of the response amplitude is predicted. Average

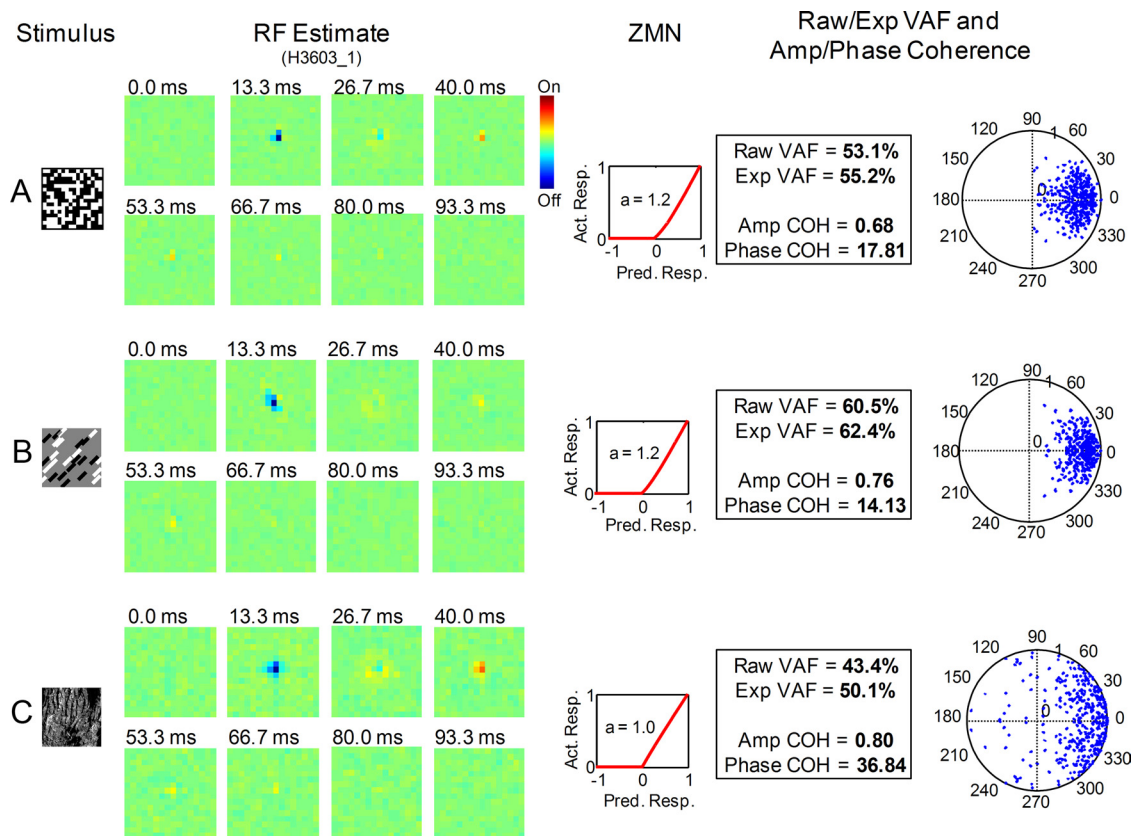


Figure 8. Example RF estimate as in Figure 5, for a neuron with a nonoriented RF. All RF estimates have a similar-looking, orientationally isotropic punctate structure with a clear phase reversal at later time lags. **A**, White noise produces a very clear RF estimate with high predictive power (raw VAF = 53.1%, exp VAF = 55.2%) and amplitude/phase coherence values that tightly cluster near optimal (amp COH = 0.68, phase COH = 17.81). **B**, RF estimate derived from short bars performs slightly better than white noise, with higher VAFs (raw VAF = 60.5%, exp VAF = 62.4%) and amplitude/phase coherence values that also tightly cluster near optimal (amp COH = 0.76, phase COH = 14.13). **C**, RF estimate derived from natural images has the lower predictive power than the other stimulus types; however, when compared to other simple cells the VAFs are high (raw VAF = 43.4%, exp VAF = 50.1%) with amplitude/phase coherence values that are clustered (amp COH = 0.80, phase COH = 36.84). For all three stimulus types, the power law nonlinearity is close to a linear half-wave rectification, with an exponent ranging from 1.0 to 1.2. Act. Resp., Actual response; Pred. Resp., predicted response.

phase COH (Fig. 9C), however, follows similar trends as VAFs, with short bars achieving the best average phase COH (phase COH = 60.57), followed by natural images (phase COH = 66.38) and finally white noise (phase COH = 77.33). These results would seem to suggest that short bars are the best of these stimuli for system identification, as they lead to models that better predict responses to independently collected validation datasets. However, a more important question is how well the RF models' predictive performance can generalize to other types of stimuli.

Predicting responses to other broadband stimulus image ensembles

To assess the robustness of the estimated RFs, each was used in model simulations to predict responses to the other two stimulus types as well as to another ensemble of the same stimulus used for its estimation. For example, a RF model derived from white noise was used to predict responses to short bars and natural images as well as to another ensemble of white noise. Results are summarized in Figure 10, where each row shows results for RF estimates derived from one of the three

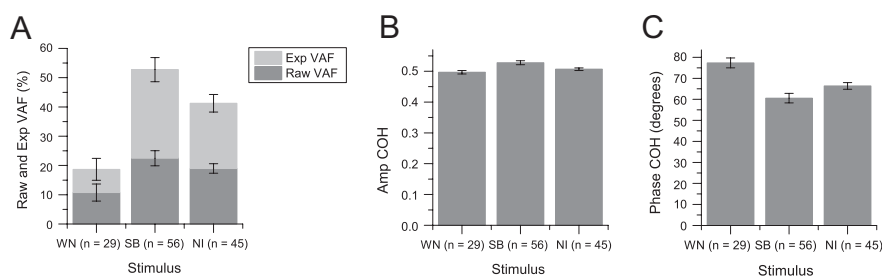


Figure 9. Average VAFs and coherence values. Raw/explainable VAFs and amplitude/phase coherence values averaged across 73 simple cells. Error bars represent standard error. **A**, On average, short bars (SB) produce the highest VAFs, followed by natural images (NI) and finally white noise (WN). This indicates that in general, short bars can better predict responses to an independently collected validation dataset of the same stimulus type, while natural images perform reasonably well and white noise does quite poorly. **B**, Average amplitude coherences are roughly equal (~0.5) across all stimulus types, indicating that predictions of neuronal response amplitude are independent of the type of stimulus used for system identification. **C**, On average, short bars achieve phase coherences closest to zero, followed by natural images and finally white noise. Unlike response amplitude, response timing is dependent on the type of stimulus used for system identification.

stimulus types, and each column compares different stimulus ensembles used in the predictions. Columns **A** and **B** compare the like stimulus (i.e., the stimulus used to create the model) with an unlike stimulus (i.e., one of the other types), and column **C** compares the two unlike types of stimuli. For RF models derived from white noise (top row, columns **A** and **B**), all points lie along the horizontal axis, indicating that a white noise-derived RF can pre-

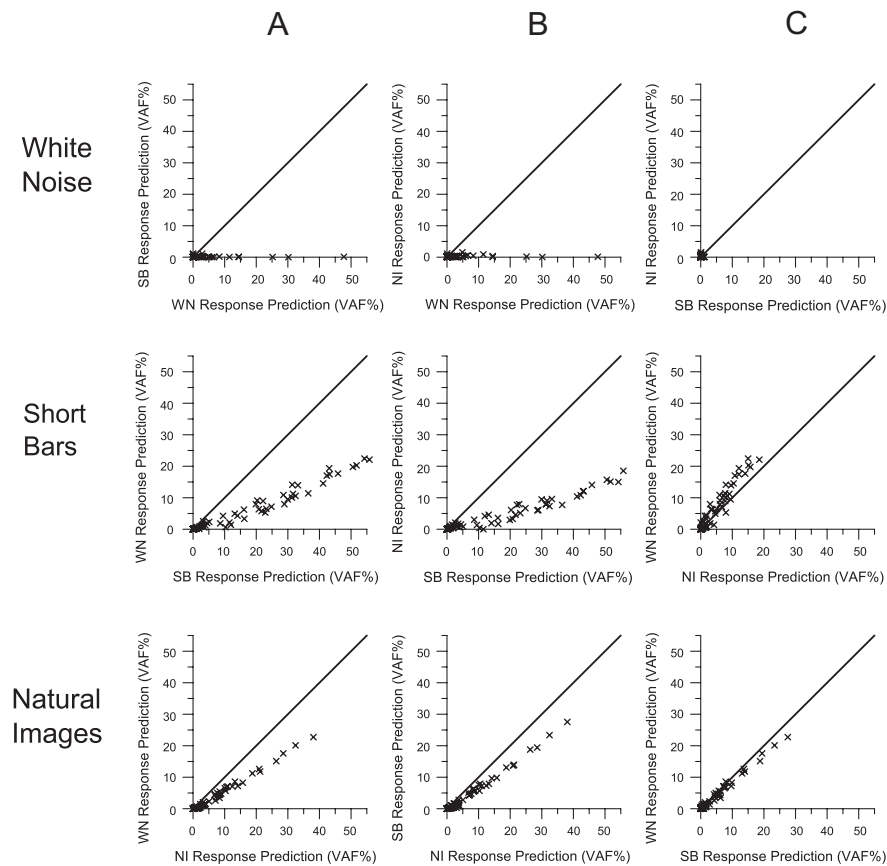


Figure 10. Generalization of predictive power to other stimulus ensembles. RF models derived from one stimulus type are used to predict responses to other types of stimuli as well as the same type of stimulus used to estimate it. Each plotted point is the VAF from predicted responses of model simulations using RF models derived from the different stimulus types. Diagonal line indicates 1:1 equality. For white noise (WN) (top row, columns *A* and *B*), all points lie along the horizontal axis, indicating that RF models derived from white noise are only capable of predicting responses to another ensemble of white noise and fail to predict short bar or natural image responses. When the responses to the two unlike types of stimuli (short bars and natural images) are plotted against one another (column *C*), all points lie at the origin, indicating that a white noise-derived RF performs equally as poorly for both of the unlike stimuli. For short bars (SB) (middle row, columns *A* and *B*), the points are closer to the 1:1 equality line, indicating that RF models derived from short bars can, to some extent, predict responses to white noise and natural images. When the two unlike types of stimuli (white noise and natural images) are plotted against one another (column *C*), points lie slightly above the 1:1 equality line, indicating that a short bar-derived RF does a marginally better job at predicting responses to white noise than to natural images. For natural images (NI) (bottom row, columns *A* and *B*), the points are very close to the 1:1 equality line, indicating that RF models derived from natural images do the best job of generalizing to white noise and short bars. When the two unlike types of stimuli (white noise and short bars) are plotted against one another (column *C*), points lie nearly on the 1:1 equality line, indicating that a natural image-derived RF does an almost equally good job at predicting responses to short bars as to white noise.

dict responses to another ensemble of white noise (although usually quite poorly) but fails entirely to predict responses to short bars (column *A*) or natural images (column *B*). When the VAFs for predictions of the two unlike types of stimuli (short bars and natural images) are plotted against one another (column *C*), all the points lie at the origin, indicating that a white noise-derived RF performs equally poorly for both of the unlike stimuli.

For RF models derived from short bars (Fig. 10, middle row, columns *A* and *B*), the points fall along a locus below the 1:1 equality line, indicating that a RF estimate derived from short bars can, to a limited extent, predict responses to white noise (column *A*) and natural images (column *B*) in a manner proportionate to its predictive power for short bar responses. When VAFs for predictions of the two unlike types of stimuli (white noise and natural images) are plotted against one another (column *C*), points lie slightly above the 1:1 equality line, indicating that a short bar-derived RF does a marginally better job at predicting responses to white noise than to natural images.

Natural image-derived RF models (Fig. 10, bottom row, columns *A* and *B*) produced scatter plots falling on a locus close to the 1:1 equality line, indicating that RF models derived from natural images do the best job of generalizing to white noise (column *A*) and short bars (column *B*). When VAFs for predictions of the two unlike types of stimuli (white noise and short bars) are plotted against one another (column *C*), the points lie almost on the 1:1 equality line, indicating that a natural image-derived RF does an almost equally good job at predicting responses to short bars as to white noise.

The preceding results suggest that natural images yield RF models that are more robust, i.e., their predictive power generalizes better to other types of stimulus image ensembles. But these results do not necessarily indicate how well these models generalize to more commonly used narrowband visual stimuli, such as sinewave gratings.

Predicting grating tuning curves for spatial frequency and orientation

RF estimates derived from each of the three stimulus types were used in model simulations to predict spatial frequency and orientation tuning curves for sinewave gratings. Predicted tuning curves were compared with actual tuning curves, as measured with sinewave gratings on the same neurons. The goodness of prediction of the measured tuning curves was quantified as VAF (see Materials and Methods). Examples of actual spatial frequency and orientation tuning curves, overlaid with predicted tuning curves, are shown in Figure 11 for the same four cells as those in Figures 5–8.

Figure 11*A* shows spatial frequency and orientation response curves for the same neuron as that in Figure 5. The spatial frequency tuning curve is best predicted by a short bar-derived RF (VAF = 83.0%), closely followed by natural images (VAF = 79.8%) and finally white noise (VAF = 67.5%). The short bar- and natural image-derived RF estimates both accurately predict the optimal spatial frequency and response amplitude but underestimate the tuning bandwidth. The white noise-derived RF accurately predicts the response amplitude but underestimates the optimal spatial frequency and overestimates the tuning bandwidth. For this cell's orientation tuning, the natural image-derived RF replicates the actual tuning almost perfectly (VAF = 96.5%), followed by short bars (VAF = 88.9%) and finally white noise (VAF = 66.0%). This cell's direction selectivity was also predicted by all three, consistent with the systematic phase progression across temporal lags that can be seen in the RF estimates in Figure 5. Note the relatively good tuning curve predictive performance of the RF model from natural images, notwithstanding its rather noisy-looking appearance (Fig. 5*C*). In this example, all three stimulus types performed reasonably well

at predicting spatial frequency and orientation tuning curves, but this was not always the case.

For the neuron in Figure 11*B* (same cell as that in Fig. 6), the spatial frequency tuning curve is best estimated by a natural image-derived RF (VAF = 62.8%) and somewhat well by short bars (VAF = 43.9%), but quite poorly by white noise (VAF = 8.3%). The natural image-derived RF estimate better predicts the neuron's tuning bandwidth compared to short bars, resulting in a better VAF. The white noise-derived RF estimate once again mispredicts the optimal spatial frequency, leading to a low VAF. For this neuron's orientation tuning, a short bar-derived RF best predicts the actual tuning (VAF = 75.0%) followed by white noise and natural images, which perform almost equally well (VAF = 59.7% and 57.3%, respectively). Note that the natural image-derived RF is the only one that accurately predicts the optimal orientation, although unlike white noise and short bars it fails to capture the narrow orientation bandwidth. All three predictions capture the strong direction selectivity in accordance with the phase progression across successive time lags, which is readily evident in the RF estimates shown in Figure 6.

The spatial frequency tuning curve for the cell in Figure 11*C* (same neuron as that in Fig. 7) is very well estimated from a natural image-derived RF (VAF = 88.3%) and reasonably well from short bars (VAF = 62.3%), while white noise fails completely to predict the actual tuning (VAF = 1.0%). The short bar-derived RF fails to capture the tuning bandwidth, while the white noise-derived RF overestimates the optimal spatial frequency. For this cell's orientation tuning, a natural image-derived RF replicates the actual tuning almost perfectly (VAF = 92.3%), followed by short bars (VAF = 58.7%) and white noise (VAF = 44.9%). The short bar-derived RF predicts the upper lobe of the actual orientation tuning well, but incorrectly predicts the direction selectivity despite the lack of phase progression across time in the RF estimate (Fig. 7*B*). The white noise-derived RF, on the other hand, somewhat predicts the nondirectionality but fails to predict the optimal orientation and orientation bandwidth. Again, note the excellent tuning curve predictive performance of the RF model from natural images, regardless of its rather noisy-looking appearance (Fig. 7*C*).

The neuron in Figure 11*D* is somewhat unusual for A18 in that its selectivity for grating spatial frequency and orientation are relatively low; however, this poor tuning is in accord with its

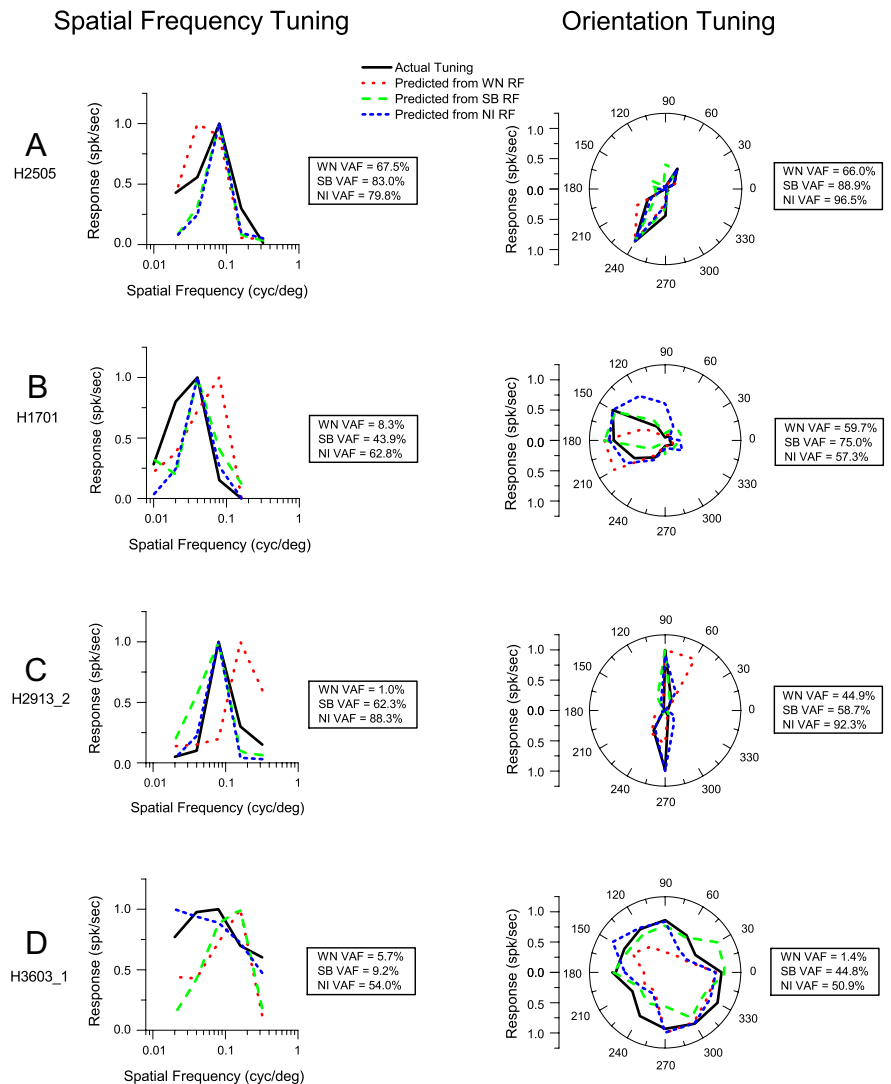


Figure 11. Example tuning curves. Actual and predicted spatial frequency and orientation tuning curves for four simple cells (same neurons as those in Figs. 5–8) using RF estimates derived from all three stimulus types. **A**, A short bar (SB)-derived RF estimate most accurately predicts the actual spatial frequency tuning (VAF = 83.0%), while natural images (NI) perform nearly as well (VAF = 79.8%), followed by white noise (WN) (VAF = 67.5%). Orientation tuning was replicated nearly perfectly using a natural image-derived RF estimate (VAF = 96.5%), followed by short bars (VAF = 88.9%) and finally white noise (VAF = 66.0%). This cell's direction selectivity was also predicted by all three stimulus types. **B**, A natural image-derived RF estimate best predicts the actual spatial frequency tuning (VAF = 62.8%), followed by short bars (VAF = 43.9%) and then white noise (VAF = 8.3%). Orientation tuning however, is best predicted by a short bar-derived RF estimate (VAF = 75.0%), followed by white noise (VAF = 59.7%) and natural images (VAF = 57.3%), which perform nearly equally. All three stimulus types capture the strong direction selectivity of this neuron. **C**, The spatial frequency tuning is very well predicted using a natural image-derived RF estimate (VAF = 88.3%) and reasonably well from short bars (VAF = 62.3%), while white noise fails entirely (VAF = 1.0%). Orientation tuning is nearly perfectly predicted with a natural image-derived RF estimate (VAF = 92.3%), followed by short bars (VAF = 58.7%) and then white noise (VAF = 44.9%). The nondirectionality of this neuron is predicted by natural images and somewhat by white noise, but fails for short bars. **D**, A natural image-derived RF estimate is the only one that can reasonably predict the actual spatial frequency tuning with a VAF = 54.0%, compared to short bars (VAF = 9.2%) and white noise (VAF = 5.7%). Orientation tuning, on the other hand, can be reasonably predicted by a natural image (VAF = 50.9%) or short bar-derived (VAF = 44.8%) RF estimate, while white noise fails entirely (VAF = 1.4%). All stimulus types capture this neuron's lack of selectivity for orientation and direction.

RF estimates (Fig. 8) that show little orientation and lack well-formed, spatially antagonistic regions. The spatial frequency response is very much better estimated by a natural image-derived RF (VAF = 54.0%) as opposed to short bars (VAF = 9.2%) or white noise (VAF = 5.7%), which both overestimate the optimal spatial frequency and tuning bandwidth. Its orientation tuning is also better predicted by a natural image-derived RF (VAF =

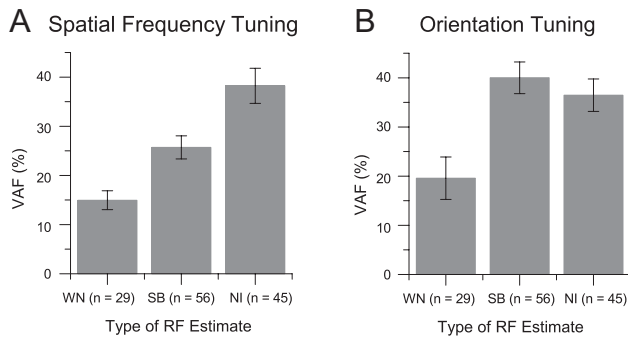


Figure 12. Average VAFs for tuning curves. Average VAFs for predictions of grating spatial frequency and orientation tuning curves. Error bars represent standard error. **A**, RF models derived from natural images (NI) better predict spatial frequency tuning curves, followed by short bars (SB) and white noise (WN). **B**, RF models derived from short bars better predict orientation tuning curves, although natural images perform nearly equally (within the error margin) while white noise again performs poorly.

50.9%), followed by short bars (VAF = 44.8%). Interestingly, white noise failed entirely to predict the actual response curve (VAF = 1.4%) despite the clean-looking RF estimate (Fig. 8A) and good VAFs for validation data. This neuron is a particularly good example of how clean-looking RF estimates and good VAFs for validation data do not guarantee good predictive power for other kinds of stimuli.

Figure 12 summarizes average VAFs between actual and predicted tuning curves for each of the three stimulus types. Across the sample population of simple cells, estimated RFs derived from natural images were best at predicting spatial frequency responses, and white noise was the poorest (Fig. 12A). Orientation tuning responses were predicted about equally well by short bars and natural images, with white noise again performing poorly (Fig. 12B).

Measurements of grating tuning curves are often made primarily for the purpose of extracting characteristic parameters, e.g., optimal values or bandwidths. To address this issue, predicted spatial frequency tuning curves were assessed on how accurately they could be used to estimate a cell's optimal spatial frequency, tuning bandwidth, and response amplitude. For orientation, predicted tuning curves were assessed on how well they could provide a cell's optimal orientation, orientation bias, optimal direction, and direction bias.

Parameters of the spatial frequency tuning curves (i.e., optimal spatial frequency, bandwidth, and response amplitude) estimated from Gaussian curve fits (see Materials and Methods) are summarized in Figure 13. Each row is for RF estimates derived from one of the three stimulus types, and each column examines a specific parameter of the tuning curve, i.e., optimal spatial frequency (column A), bandwidth (column B), and response amplitude (column C). The scatter plots show the predicted parameter values plotted against the actual values for all neurons in our sample. The histograms in column A are measured as the perpendicular distance to the 1:1 equality line. In columns B and C, the histograms indicate the distribution of prediction errors (residuals). For all parameters of the spatial frequency tuning curves examined, natural image-derived RFs were the best predictors as they produced the least amount of deviation from the 1:1 equality line, and histograms with more tightly tuned distributions (bottom row, columns A–C), while white noise yielded the worst predictors (top row, columns A–C). This is further confirmed by examining the mean and standard deviations of the histograms (shown right of histograms), which are measures of

accuracy and precision, respectively. Natural image-derived RF estimates lead to spatial frequency tuning parameters with mean and standard deviation values closest to zero (i.e., highest degree of accuracy and precision), while short bars result in intermediate values and white noise to values furthest from zero. Note one seemingly interesting result, that white noise-derived RF estimates lead to optimal spatial frequency predictions that are highly biased toward higher values (top row, column A).

Parameters of orientation tuning curves (i.e., optimal orientation, orientation bias, optimal direction, and direction bias) estimated from vector-based summation (see Materials and Methods) are summarized in Figures 14 and 15. Figure 14 examines optimal orientation (column A) and orientation bias (column B) where each row is for RF estimates derived from one of the three stimulus types. The scatter plots show predicted parameter values plotted against the actual values for all neurons in our sample, with the histograms below indicating the distribution of prediction errors (residuals). Optimal orientation values range between 0° and 180°, and orientation bias values range between 0 and 1 (dimensionless). Optimal orientation (column A) is best predicted using estimated RFs derived from short bars, as there is less deviation from the 1:1 equality line and the residual histogram is tightly tuned with mean and standard deviation values closest to zero. Natural image-derived RFs perform very similarly to short bars, although the residual histogram is somewhat more broadly spread. White noise performs poorly with a broad residual histogram. Orientation bias (column B) is best predicted using estimated RFs derived from natural images, as there is less deviation from the 1:1 equality line and the residual histogram is tightly tuned with a standard deviation value closest to zero. Note that since many points in the natural image scatter plot (bottom row, column B) fall below the 1:1 equality line, the mean of the residual histogram is slightly skewed to the right, indicating a tendency to underestimate orientation selectivity. Short bar-derived RFs also perform reasonably well, although again they somewhat underestimate orientation selectivity. Once again, white noise-derived RFs perform poorly, with a residual histogram that is spread across many bins.

Figure 15 examines optimal direction (column A) and direction bias (column B), where each row is for RF estimates derived from one of the three stimulus types. The scatter plots show the predicted parameter values plotted against the actual values for all neurons in our sample, with the histogram below indicating the distribution of prediction errors (residuals). Optimal direction values range between 0° and 360°, and direction bias values range between 0 and 1 (dimensionless). Optimal direction (column A) is predicted equally well using estimated RFs derived from short bars or natural images, as both have scatter plots with points that lie close to the 1:1 equality line and residual histograms with similar mean and standard deviation values. White noise-derived RFs once again do a poor job, with predictions of optimal direction that are highly scattered and a residual histogram with high standard deviation. Direction bias (column B) is best predicted by using estimated RFs derived from natural images, followed by short bars that perform reasonably well and finally by white noise that performs poorly. Natural images have the least amount of deviation from the 1:1 equality line and a residual histogram that is tightly tuned with a standard deviation value closest to zero.

Discussion

We compared the utility of three types of broadband stimuli (white noise, short bars, and natural images) to provide spatio-

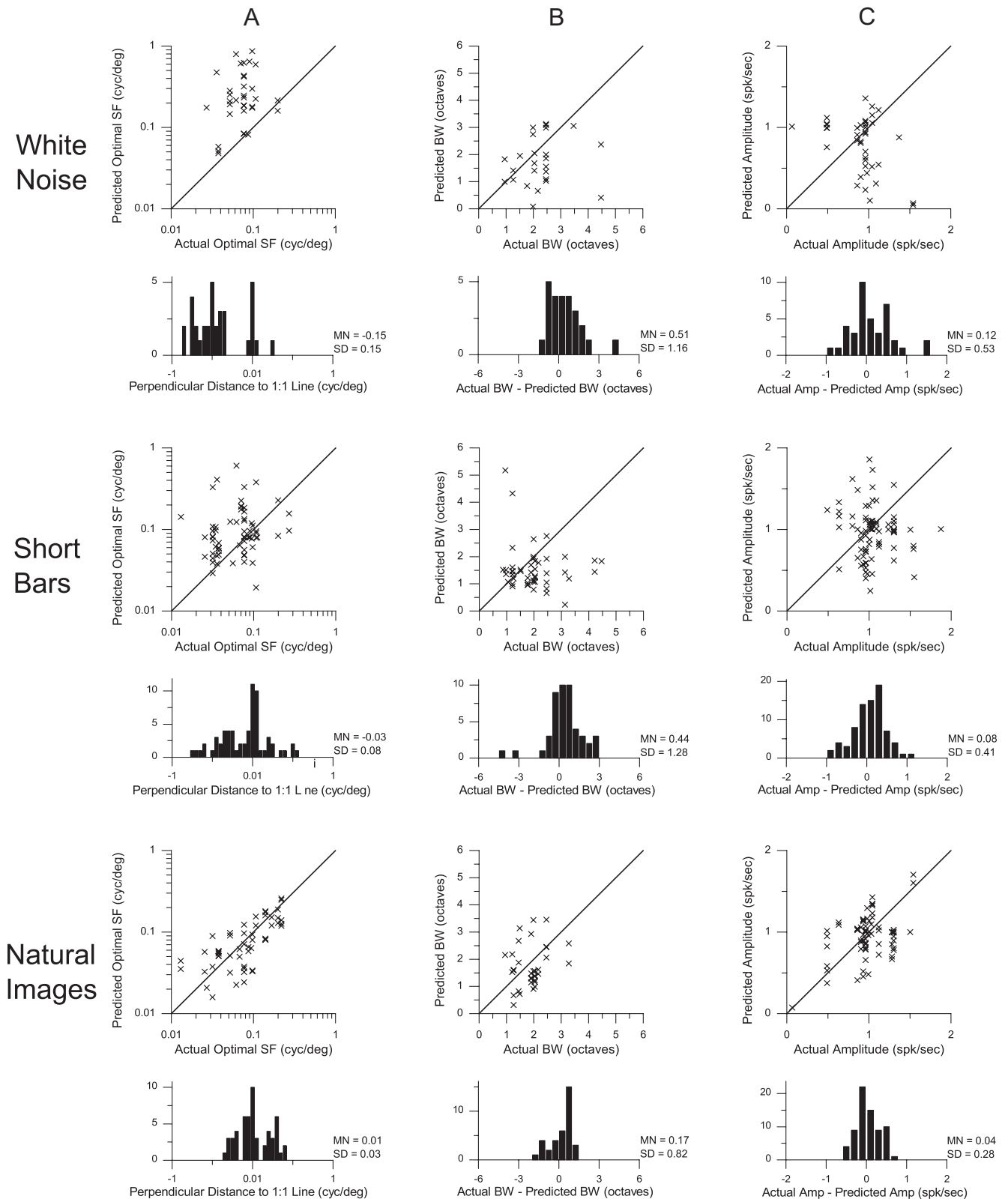


Figure 13. Prediction of spatial frequency tuning parameters. Actual and predicted values of optimal spatial frequency (SF) (column **A**), bandwidth (BW) (column **B**), and normalized response amplitude (Amp) (column **C**) from Gaussian curve fits to spatial frequency tuning curves. Each row is for RF models derived from the three stimulus types (white noise, short bars, and natural images). Scatter plots show predicted values fitted against actual values, and the histograms below indicate the distribution of prediction errors. All spatial frequency parameters (columns **A–C**) are best predicted by natural image-derived RF estimates, followed by short bars and finally white noise. In all cases, natural images have the least amount of deviation from the 1:1 equality line (diagonal) and least spread across histogram bins. The means (MN) and standard deviations (SD) of the histograms confirm that natural image-derived RF estimates lead to spatial frequency tuning parameters with the highest degree of accuracy and precision (i.e., mean and standard deviations closest to zero), followed by short bars and finally white noise.

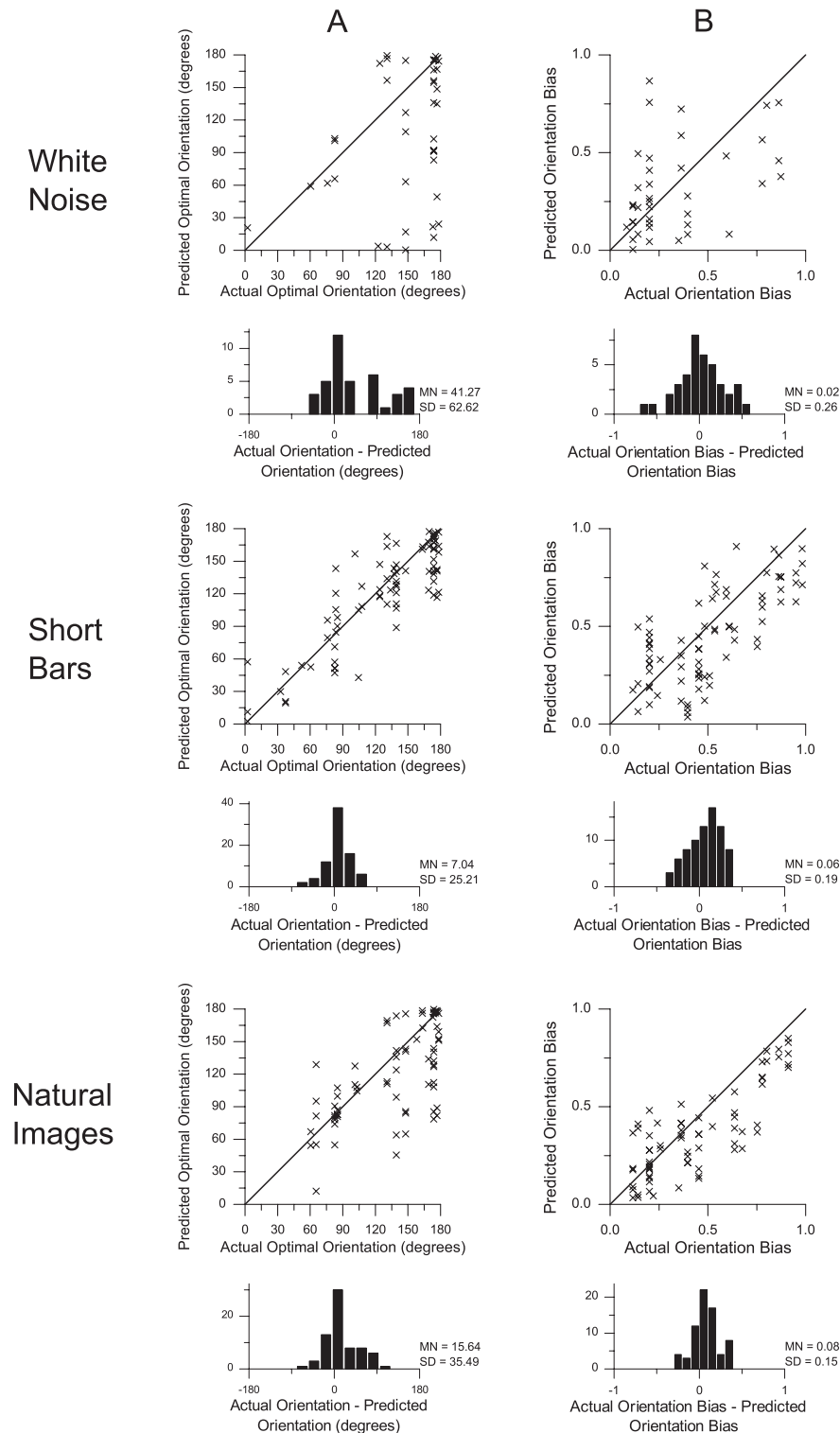


Figure 14. Prediction of orientation tuning parameters. Actual and predicted values of optimal orientation (column **A**) and orientation bias (column **B**) from vector-based summations of orientation tuning curves. Each row is for RF models derived from one of the three stimulus types (white noise, short bars, and natural images). Scatter plots show predicted values plotted against actual values, and the histograms below indicate the distribution of prediction errors. Optimal orientation values range between 0° and 180° and orientation bias values range between 0 and 1 (dimensionless), where values greater than 0.1 indicate selectivity. Optimal orientation (column **A**) is best predicted by short bar-derived RF estimates, followed by natural images and finally white noise. Short bars have the least deviation of points from the 1:1 equality line (diagonal) and a residual histogram that is tightly tuned with mean (MN) and standard deviation (SD) values closest to zero when compared to the other two stimuli. Orientation bias (column **B**) is best predicted by natural image-derived RF estimates, followed by short bars and finally white noise. Natural images have points that are closer to the 1:1 equality line and a residual histogram that is tightly tuned with a standard deviation value that is closest to zero when compared to the other two stimuli.

temporal RF models with high predictive ability for responses to other kinds of stimuli. Dense white noise stimuli often elicited poor responses and RF estimates, while sparsely structured short bars and complex natural image stimuli were more successful in driving A18 neurons and producing clear RF estimates. The RF models derived from short bar and natural image stimuli were able to predict responses relatively well, yielding explainable VAFs of ~40–50%. Models from all three types of stimuli could more reliably predict the timing of a response than its amplitude. Natural image-derived RF models, however, were the most robust at predicting responses to other types of broadband stimuli than those used for their estimation, and also performed well in predicting tuning curves for sinuswave gratings.

Comparison of stimuli

Wu et al. (2006) delineated three major classes of stimuli typically used in visual system identification: white noise, parametric noise (i.e., stimuli with random series of structured patterns such as bars, sinusoidal gratings, sum of sinusoids), and natural images. Although many studies have used one of these classes of stimuli (Jones and Palmer, 1987; Ringach et al., 1997; Smyth et al., 2003) or compared two of them (David et al., 2004; Felsen et al., 2005; Yeh et al., 2009), to our knowledge this work is novel in its comparison of examples from all three classes.

The markedly poor performance of white noise would seem contrary to the numerous successful visual system identification studies performed using this stimulus (Ringach and Shapley, 2004). However, many of these studies were in retina (Yasui et al., 1979; Hida and Naka, 1982) or LGN (Reid and Shapley, 1992, 2002), where receptive field properties are relatively simple. In A17, the best white noise responses have been reported to be in layer 4 (Alonso et al., 2001), which is a principal thalamorecipient target. The poor performance of white noise in A18 may be related to the greater complexity of receptive fields, perhaps analogous to a similar comparison between V1 and V2 in the primate (Willmore et al., 2010).

A seemingly more appropriate stimulus for system identification would be short bars, since our results showed them to produce the clearest-looking RF estimates with the best predictive ability (Fig. 9). Short bar-derived RFs were also more successful at predicting orientation tuning; however, this result might not be surprising due to the necessity of selecting

short bar ensembles whose orientations and bar widths were matched to each neuron's measured parameters. Furthermore, our results indicate that RF models derived from short bars have limited capacity in predicting responses to other types of broadband stimuli.

Natural images produced RF models with reasonable VAFs, but more importantly they were the most robust at predicting responses to other types of broadband stimuli (Fig. 10). They also performed very well in prediction of sinewave grating responses. It is unclear from these analyses whether the apparently greater "noise" in the natural image RF estimates is due to variance in the estimates themselves (arising, for example, from somewhat weaker responses) or if these estimates depict genuine fine-grain RF structure that plays a constructive role in robust predictive power—this will be a matter of future investigation.

From these results we cannot definitively say which aspects of the different broadband stimuli led to their varying utility in system identification; they were selected because of their common use in visual neurophysiology and differed from one another in various ways (e.g., spatial spectrum, RMS energy, and higher-order image statistics). We conjecture that natural images performed best due to their sparseness and richer spatial features. Visual neurons presumably evolved to efficiently encode the rich spatial structure of natural scenes whose image statistics are a highly constrained manifold of the space of possible images (Simoncelli and Olshausen, 2001). One key feature of natural images is their sparseness, which may provide an important constraint for efficient coding (Field, 1994). This sparseness arises from a number of higher-order statistics of natural images, including their pronounced contrast modulations (Johnson and Baker, 2004), abrupt edges with local phase alignments across spatial frequencies (Field, 1993; Olshausen and Field, 1996), and local correlations of luminance and contrast (Johnson and Baker, 2004; Mante et al., 2005; Frazor and Geisler, 2006).

System identification

Most early efforts at neural system identification only attempted to estimate RFs but did not assess them by measuring how well they could account for responses to arbitrary stimuli. An ultimate measure of the accuracy of a RF model is how well it can predict responses, particularly to stimuli that were not used for estimation (Wu et al., 2006). Only a few studies have

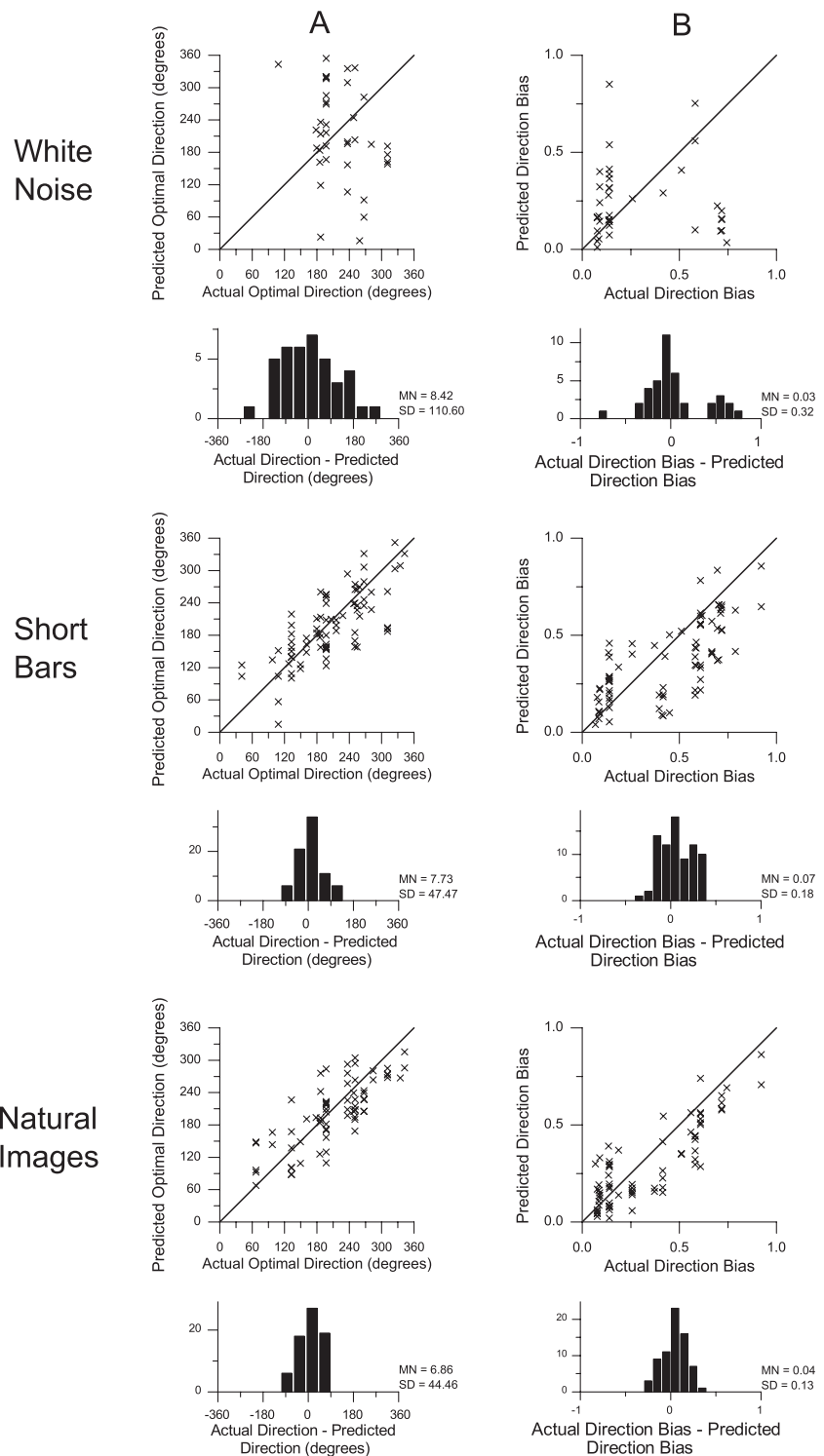


Figure 15. Prediction of direction tuning parameters. Actual and predicted values of optimal direction (column **A**) and direction bias (column **B**) from vector-based summations of orientation tuning curves. Each row is for RF models derived from one of the three stimulus types (white noise, short bars, and natural images). Scatter plots show predicted values plotted against actual values, and the histograms below indicate the distribution of prediction errors. Optimal direction values range between 0° and 360° and direction bias values range between 0 and 1 (dimensionless), where values greater than 0.1 indicate selectivity. Optimal direction (column **A**) is equally well predicted by short bar and natural image-derived RF estimates. Both scatter plots have points that lie close to the 1:1 equality line (diagonal) and residual histograms that are tightly tuned, with similar mean (MN) and standard deviation (SD) values. White noise-derived RF estimates perform poorly, with predictions of optimal direction that are very scattered and a residual histogram with high standard deviation. Direction bias (column **B**) is best predicted by natural image-derived RF estimates, closely followed by short bars and finally white noise. Natural image results have the least amount of deviation from the 1:1 equality line and a residual histogram that is tightly tuned with a standard deviation value closest to zero when compared to the other two stimuli.

employed a rigorous procedure with independent datasets to fit model parameters, minimize overfitting, and evaluate predictions (Theunissen et al., 2001; David et al., 2004; David and Gallant, 2005; Willmore et al., 2010).

The predictive power of a model is best assessed using measures that take into account sampling limitations and experimental noise (David and Gallant, 2005). Although several types of measures exist (Hsu et al., 2004; Mante et al., 2008), we have chosen to use explainable VAF that we computed through a noise ceiling analysis (David and Gallant, 2005). Explainable VAF helps separate how much of the prediction discrepancy is due to response noise versus model insufficiency. Although powerful, it is important to note that explainable VAFs may sometimes be underestimated when there are insufficient repetitions or unique presentations of a stimulus image ensemble, such that the noise ceiling curve does not approach a well-defined asymptote.

Previous studies using approaches similar to ours with natural images have at best accounted for explainable variances of circa 40% (Carandini et al., 2005), comparable to our results and well below the theoretical ideal of 100%. This might be partly due to effects of nonstationarity and the high degree of response variability of cortical neurons, although the latter should be largely discounted by the noise ceiling analysis. Most importantly, these neurons exhibit nonlinear phenomena that are not adequately captured by an LN model, such as gain control (Heeger, 1992), surround modulation (Tanaka and Ohzawa, 2009), second-order responses (Baker and Mareschal, 2001), or cross-orientation inhibition (Bonds, 1989). By incorporating more elaborate model architectures that can produce these nonlinear response properties, it may be possible to develop more robust models with higher predictive ability.

Many studies have demonstrated that estimated RF models could predict optimal grating responses in A17 (Movshon et al., 1978; Jones and Palmer, 1987; Tadmor and Tolhurst, 1989; DeAngelis et al., 1993b; Gardner et al., 1999; Smyth et al., 2003). Furthermore, other studies have reported that RF estimates generated using stimuli from a specific class better predict responses within than across stimulus classes (Citron et al., 1988; Golomb et al., 1994; Lau et al., 2002; David et al., 2004). Our results showed that white noise and short bar RF estimates perform less well at predicting responses to other broadband stimuli. RF estimates derived from natural stimuli, on the other hand, were much better at predicting responses across stimulus classes.

Future directions

A promising extension to this system identification approach would be to apply nonlinear transformations to the stimulus before the GLM regression (Willmore and Smyth, 2003). Such a basis set transformation (Mitchell, 1997; Bishop, 2006) or “pre-processing” has been used to characterize nonlinear V1/V2 neurons (Theunissen et al., 2001; Chen et al., 2007; Willmore et al., 2010). The basis set could consist of wavelets such as Gabor functions or filters that mimic neural RFs at earlier processing stages. This approach should enable the characterization of many different types of neuronal nonlinearities (e.g., complex cells, nonlinear subunits).

It might be possible, and perhaps advantageous, to design “natural-like” synthetic stimuli that perform as well or even better than natural images. However, natural images are inherently complex, and it is not straightforward to specify or analyze what statistical relationships make them critically different from other random stimuli. But if this problem can be solved, then by replicating important higher-order structural statistics it may be pos-

sible to find combinations of features that can mimic natural images (Portilla and Simoncelli, 2000).

Conclusion

The main goal of this study was to evaluate commonly used synthetic and natural stimuli for use with a newer generation of system identification methods to find RF estimates with strong predictive power that generalize well to other stimulus types. Our results suggest natural images to be a strong candidate for achieving this objective. In addition to the strengths of natural images in RF estimation and in creating robust RF models, they also do not require prior knowledge of the cell’s optimal tuning properties. This makes natural image stimuli, together with regularized GLM methods of system identification, ideal tools for estimating RF models in brain areas beyond striate cortex, as well as for rapidly characterizing the tuning properties of multiple neurons at one time (e.g., multi-electrode recordings, two-photon imaging).

References

- Alonso JM, Martinez LM (1998) Functional connectivity between simple cells and complex cells in cat striate cortex. *Nat Neurosci* 1:395–403.
- Alonso JM, Usrey WM, Reid RC (2001) Rules of connectivity between geniculate cells and simple cells in cat primary visual cortex. *J Neurosci* 21:4002–4015.
- Baker CL Jr, Mareschal I (2001) Processing of second-order stimuli in the visual cortex. *Prog Brain Res* 134:171–191.
- Bishop CM (2006) *Pattern recognition and machine learning*. New York: Springer.
- Bonds AB (1989) Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. *Vis Neurosci* 2:41–55.
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436.
- Carandini M, Demb JB, Mante V, Tolhurst DJ, Dan Y, Olshausen BA, Gallant JL, Rust NC (2005) Do we know what the early visual system does? *J Neurosci* 25:10577–10597.
- Chen X, Han F, Poo MM, Dan Y (2007) Excitatory and suppressive receptive field subunits in awake monkey primary visual cortex (V1). *Proc Natl Acad Sci U S A* 104:19120–19125.
- Citron MC, Emerson RC, Levick WR (1988) Nonlinear measurement and classification of receptive fields in cat retinal ganglion cells. *Ann Biomed Eng* 16:65–77.
- David SV, Gallant JL (2005) Predicting neuronal responses during natural vision. *Network* 16:239–260.
- David SV, Vinje WE, Gallant JL (2004) Natural stimulus statistics alter the receptive field structure of v1 neurons. *J Neurosci* 24:6991–7006.
- DeAngelis GC, Ohzawa I, Freeman RD (1993a) Spatiotemporal organization of simple-cell receptive fields in the cat’s striate cortex. I. General characteristics and postnatal development. *J Neurophysiol* 69:1091–1117.
- DeAngelis GC, Ohzawa I, Freeman RD (1993b) Spatiotemporal organization of simple-cell receptive fields in the cat’s striate cortex. II. Linearity of temporal and spatial summation. *J Neurophysiol* 69:1118–1135.
- DeAngelis GC, Freeman RD, Ohzawa I (1994) Length and width tuning of neurons in the cat’s primary visual cortex. *J Neurophysiol* 71:347–374.
- Drongelen WV (2007) *Signal processing for neuroscientists*. San Diego: Academic.
- Einhäuser W, Kayser C, König P, Körding KP (2002) Learning the invariance properties of complex cells from their responses to natural stimuli. *Eur J Neurosci* 15:475–486.
- Felsen G, Dan Y (2005) A natural approach to studying vision. *Nat Neurosci* 8:1643–1646.
- Felsen G, Touryan J, Han F, Dan Y (2005) Cortical sensitivity to visual features in natural scenes. *PLoS Biol* 3:e342.
- Fernald R, Chase R (1971) An improved method for plotting retinal landmarks and focusing the eyes. *Vis Res* 11:95–96.
- Field DJ (1993) Scale-invariance and self-similar ‘wavelet’ transforms: an analysis of natural scenes and mammalian visual systems. In: *Wavelets, Fractals, and Fourier Transforms* (Farge M, Hunt J, Vascillicos C, eds), pp 151–193. Oxford: Oxford UP.
- Field DJ (1994) What is the goal of sensory coding? *Neural Comput* 6:559–601.

- Frazor RA, Geisler WS (2006) Local luminance and contrast in natural images. *Vision Res* 46:1585–1598.
- Gardner JL, Anzai A, Ohzawa I, Freeman RD (1999) Linear and nonlinear contributions to orientation tuning of simple cells in the cat's striate cortex. *Vis Neurosci* 16:1115–1121.
- Golomb D, Kleinfeld D, Reid RC, Shapley RM, Shraiman BI (1994) On temporal codes and the spatiotemporal response of neurons in the lateral geniculate nucleus. *J Neurophysiol* 72:2990–3003.
- Hagiwara K (2002) Regularization learning, early stopping and biased estimator. *Neurocomputing* 48:937–955.
- Heeger DJ (1992) Normalization of cell responses in cat striate cortex. *Vis Neurosci* 9:181–197.
- Hegd e J, Van Essen DC (2000) Selectivity for complex shapes in primate visual area V2. *J Neurosci* 20:RC61.
- Hida E, Naka K (1982) Spatio-temporal visual receptive fields as revealed by spatio-temporal random noise. *Z Naturforsch C* 37:1048–1049.
- Hsu A, Borst A, Theunissen FE (2004) Quantifying variability in neural responses and its application for the validation of model predictions. *Network* 15:91–109.
- Johnson AP, Baker CL Jr (2004) First- and second-order information in natural images: a filter-based approach to image statistics. *J Opt Soc Am A Opt Image Sci Vis* 21:913–925.
- Jones JP, Palmer LA (1987) The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *J Neurophysiol* 58:1187–1211.
- Lau B, Stanley GB, Dan Y (2002) Computational subunits of visual cortical neurons revealed by artificial neural networks. *Proc Natl Acad Sci U S A* 99:8974–8979.
- Leventhal AG, Thompson KG, Liu D, Zhou Y, Ault SJ (1995) Concomitant sensitivity to orientation, direction, and color of cells in layers 2, 3, and 4 of monkey striate cortex. *J Neurosci* 15:1808–1818.
- Leventhal AG, Wang Y, Pu M, Zhou Y, Ma Y (2003) GABA and its agonists improved visual cortical function in senescent monkeys. *Science* 300:812–815.
- Li G, Talebi V, Yoonessi A, Baker CL Jr (2010) A FPGA real-time model of single and multiple visual cortex neurons. *J Neurosci Methods* 193:62–66.
- Machens CK, Wehr MS, Zador AM (2004) Linearity of cortical receptive fields measured with natural sounds. *J Neurosci* 24:1089–1100.
- Mante V, Frazor RA, Bonin V, Geisler WS, Carandini M (2005) Independence of luminance and contrast in natural scenes and in the early visual system. *Nat Neurosci* 8:1690–1697.
- Mante V, Bonin V, Carandini M (2008) Functional mechanisms shaping lateral geniculate responses to artificial and natural stimuli. *Neuron* 58:625–638.
- Marmarelis PZ, Marmarelis VZ (1978) Analysis of physiological systems: The white-noise approach. New York: Plenum.
- McCullagh P, Nelder JA (1989) Generalized linear models. London: Chapman and Hall.
- Mitchell TM (1997) Machine learning. New York: McGraw Hill.
- Movshon JA, Thompson ID, Tolhurst DJ (1978) Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J Physiol* 283:53–77.
- Nikara T, Bishop PO, Pettigrew JD (1968) Analysis of retinal correspondence by studying receptive fields of binocular single units in cat striate cortex. *Exp Brain Res* 6:353–372.
- Olmos A, Kingdom FAA (2004) A biologically inspired algorithm for the recovery of shading and reflectance images. *Perception* 33:1463–1473.
- Olshausen BA, Field DJ (1996) Natural image statistics and efficient coding. *Network* 7:333–339.
- Pelli DG (1997) The video toolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442.
- Portilla J, Simoncelli EP (2000) A parametric texture model based on joint statistics of complex wavelet coefficients. *Int J Comput Vis* 40:49–71.
- Reid RC, Shapley RM (1992) Spatial structure of cone inputs to receptive fields in primate lateral geniculate nucleus. *Nature* 356:716–718.
- Reid RC, Shapley RM (2002) Space and time maps of cone photoreceptor signals in macaque lateral geniculate nucleus. *J Neurosci* 22:6158–6175.
- Ringach DL, Shapley R (2004) Reverse correlation in neurophysiology. *Cogn Sci* 28:147–166.
- Ringach DL, Sapiro G, Shapley R (1997) A subspace reverse-correlation technique for the study of visual neurons. *Vis Res* 37:2455–2464.
- Ringach DL, Hawken MJ, Shapley R (2002) Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. *J Vis* 2(1):2 12–24.
- Rust NC, Movshon JA (2005) In praise of artifice. *Nat Neurosci* 8:1647–1650.
- Simoncelli EP, Olshausen BA (2001) Natural image statistics and neural representation. *Annu Rev Neurosci* 24:1193–1216.
- Skottun BC, De Valois RL, Grosf DH, Movshon JA, Albrecht DG, Bonds AB (1991) Classifying simple and complex cells on the basis of response modulation. *Vis Res* 31:1079–1086.
- Smyth D, Willmore B, Baker GE, Thompson ID, Tolhurst DJ (2003) The receptive-field organization of simple cells in primary visual cortex of ferrets under natural scene stimulation. *J Neurosci* 23:4746–4759.
- Tadmor Y, Tolhurst DJ (1989) The effect of threshold on the relationship between the receptive-field profile and the spatial-frequency tuning curve in simple cells of the cat's striate cortex. *Vis Neurosci* 3:445–454.
- Tanaka H, Ohzawa I (2009) Surround suppression of V1 neurons mediates orientation-based representation of high-order visual features. *J Neurophysiol* 101:1444–1462.
- Theunissen FE, David SV, Singh NC, Hsu A, Vinje WE, Gallant JL (2001) Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network* 12:289–316.
- Touryan J, Felsen G, Dan Y (2005) Spatial structure of complex cell receptive fields measured with natural images. *Neuron* 45:781–791.
- Tusa RJ, Rosenquist AC, Palmer LA (1979) Retinotopic organization of area 18 and 19 in the cat. *J Comp Neurol* 185:657–678.
- Willmore B, Smyth D (2003) Methods for first-order kernel estimation: simple-cell receptive fields from responses to natural scenes. *Network* 14:553–577.
- Willmore BD, Prenger RJ, Gallant JL (2010) Neural representation of natural images in visual area V2. *J Neurosci* 30:2102–2114.
- Wörg tter F, Eysel UT (1987) Quantitative determination of orientational and directional components in the response of visual cortical cells to moving stimuli. *Biol Cybern* 57:349–355.
- Wu MC, David SV, Gallant JL (2006) Complete functional characterization of sensory neurons by system identification. *Annu Rev Neurosci* 29:477–505.
- Yasui S, Davis W, Naka KI (1979) Spatio-temporal receptive field measurement of retinal neurons by random pattern stimulation and cross correlation. *IEEE Trans Biomed Eng* 26:263–272.
- Yeh CL, Xing D, Williams PE, Shapley RM (2009) Stimulus ensemble and cortical layer determine V1 spatial receptive fields. *Proc Natl Acad Sci U S A* 106:14652–14657.