

Temporal dynamics of directional selectivity in human vision

Peter Neri

School of Optometry & Helen Wills Neuroscience Institute,
University of California at Berkeley, Berkeley, CA, USA



Dennis Levi

School of Optometry & Helen Wills Neuroscience Institute,
University of California at Berkeley, Berkeley, CA, USA



We used psychophysical reverse correlation to determine how directional signals are integrated across a time window of 300 ms. Directional tuning was time independent within the resolution of our measurements, as demonstrated by the fact that the perceptual filter was almost perfectly separable in its temporal and directional dimensions. The amplitude of the filter peaked very early (30–60 ms) and then quickly decreased to almost zero, after which it increased slightly again. We successfully modeled this bimodal behavior using a simple circuit where each directional filter normalizes its own output, with the normalizing signal delayed by ~100 ms.

Keywords: noise image classification, delayed normalization, separability

Citation: Neri, P., & Levi, D. (2008). Temporal dynamics of directional selectivity in human vision. *Journal of Vision*, 8(1):22, 1–11, <http://journalofvision.org/8/1/22/>, doi:10.1167/8.1.22.

Introduction

Some aspects of neuronal responses take time to develop. For example, it has been shown that spatial frequency tuning becomes increasingly sharper over a period of ~200 ms after stimulus onset in V1 neurons (Bredfeldt & Ringach, 2002; Mazer, Vinje, McDermott, Schiller, & Gallant, 2002). In MT, some neurons take 50–100 ms to develop plaid-selective responses (Pack & Born, 2001; Smith, Majaj, & Movshon, 2005). Moreover, their directional selectivity undergoes fast changes depending on stimulus history (Perge, Borghuis, Bours, Lankheet, & van Wezel, 2005), a result that has also been demonstrated for orientation tuning in primary visual cortex (Dragoi, Rivadulla, & Sur, 2001; Dragoi, Sharma, Miller, & Sur, 2002; Dragoi, Sharma, & Sur, 2000; Felsen et al., 2002; Müller, Metha, Krauskopf, & Lennie, 1999) and for tangential cells in the fly lobula plate (Neri, 2007).

Our goal was to study directional tuning in human vision and to determine whether it undergoes substantial changes during the first two to three hundred milliseconds after stimulus onset. We derived time-varying directional tuning curves using motion noise that varied across time, similar to Neri and Heeger (2002) and Mareschal, Dakin, and Bex (2006) (see also Tadin, Lappin, & Blake, 2006). The task involved discriminating the direction of a circular field of moving dots, and human responses were analyzed using noise image classification (Ahumada, 2002). This technique exploits the correlation between trial-by-trial fluctuations in the noise samples and individual responses by the observers, allowing retrieval of

the perceptual filter that underlies their behavioral choices (for a recent review, see Neri & Levi, 2006).

We found that directional selectivity was characterized by a tuning curve with a half-width at half-height of 30–50 deg, broadly consistent with electrophysiological studies in MT (Cook & Maunsell, 2004). This degree of tuning emerged very early and was invariant over time, at least within the resolution of our measurements. We verified this invariance by computing the expected direction–time surface filter that is obtained by multiplying the directional tuning curve averaged across time with the temporal impulse response averaged across direction (the assumption of separability) and by subtracting this prediction from the experimental filter. The resulting surface was virtually flat.

One interesting and unexpected feature of our data is that the amplitude of the filter peaked at the direction of the target as expected, reached its maximum value very early, then decreased, and finally showed a tendency to increase again. We were able to model this temporal evolution of amplitude modulation by using a minimal model involving linear directional filters followed by a self-normalization stage.

Methods

Stimuli

The stimulus consisted of a sequence of 10 frames like the one shown in Figure 1A, each lasting 30 ms (total

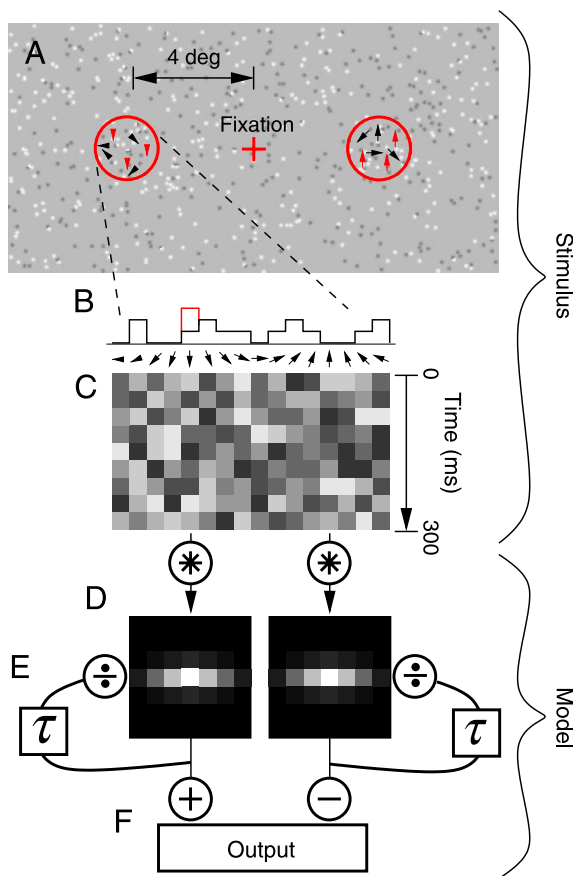


Figure 1. (A) One frame of the stimulus. Dots moved only within the two circular regions to the sides of fixation (the two red circles were actually present in the stimulus). A fixed percentage of the dots (signal dots, indicated by red arrows) moved \downarrow on one side of fixation (left in the example) and \uparrow on the other side. Subjects had to indicate the side of the stimulus containing \downarrow motion (or \uparrow for one of the 3 subjects). (B) Black line shows number of dots moving in the direction indicated by the x axis on one motion frame (two successive frames) and within the left circular region. Red shows additional signal dots. (C) The stimulus contained 9 motion frames, defining a surface with dimensions “direction” (x axis) and “time” (y axis). (D) Model filters used to extract the response at directions \downarrow (left) and \uparrow (right). (E) The output from each filter normalized itself after a delay τ of 90 ms. (F) The outputs from the two filters for opposite directions were subtracted from each other to generate the final response to the circled region on the left. The model processed the circled region on the right in a similar way and selected the region with largest output as target.

stimulus duration was 300 ms). The dots were anti-aliased Gaussian blobs (standard deviation = 3.75 arcmin) with peak luminance of 74 (bright) or 0 (dark) cd/m^2 on a background luminance of 37 cd/m^2 (monitor was gamma-corrected). For regions where blobs overlapped, the pixel was assigned the luminance of the blob with largest absolute value at that position. Dots outside the two

regions circled in red (diameter was 2.1 deg, distance of centre from fixation was 4.3 deg) were always present on the screen and never moved. Because the stimulus consisted of 10 frames, it contained 9 motion impulses. Half the noise dots moved from frame 1 to frame 2, the remaining half from frame 2 to 3 and so forth, meaning that on each frame half the noise dots had terminated their 2-frame motion while the other half was about to move (we adopted this configuration to avoid concurrent motion jumps for the whole display). At the end of each 2-frame motion, each dot was randomly relocated to a new starting position for the next 2-frame motion. Dot velocity was 4.1 deg/s. Within each region, a fixed percentage of the dots were signal dots, while the remaining dots were noise dots. Signal dots moved in opposite vertical directions within the two regions: If signal dots moved upward within the region on the right (as in Figure 1A), signal dots within the region on the left moved downward, and vice versa. Each noise dot could take any of 16 directions (nearby directions differing by 22.5 deg) with uniform probability. There were 16 dots in total for each motion impulse (total of $16 \times 9 = 144$) in each circled region. Of these, 2 (S1 and S3) or 6 (S2) were signal dots.

Task

On each trial, observers were asked to press one of two buttons to indicate which of the two possible signal configurations were presented: either (1) signal dots moved upward within the region on the left and signal dots moved downward within the region on the right or (2) signal dots moved downward within the region on the left and signal dots moved upward within the region on the right. Their response triggered the next presentation. This task was worded in the following terms: press 1 for left or 2 for right to indicate the side where you saw predominantly downward motion, being aware that the other side always contains upward motion (for one subject (S1) the target direction was upward). Observers were initially familiarized with the task by presenting them with 20–30 trials containing 100% signal dots. They learnt the task very quickly, after which we proceeded to reduce the % of signal dots until their performance was around threshold (75% correct responses). The data shown here were collected at a fixed threshold number of signal dots (detailed earlier). Percent correct for the task was 0.64 (S1), 0.78 (S2), and 0.77 (S3). Number of collected trials was 11,150 (S1), 10,100 (S2), and 5,000 (S3). Feedback (correct/incorrect) was provided after each response. S1 and S2 were naive regarding the methodology and purpose of the study and S3 was one of the authors (P. N.). For double-pass experiments, we ran blocks of 100 trials for which the second half of the block (50 trials) consisted of a random permutation of the trials presented during the first half. Number of collected trials was 3,100 (S1), 1,000 (S2), and 400 (S3).

Derivation of perceptual filters

The stimulus presented within each circled region on each trial i can be described as $N_i(d, t)$, the number of noise dots that moved in direction d at time t . There were 16 possible directions and 9 time points, defining the dimensionality of the surface. If $N_i^T(d, t)$ refers to the region containing signal dots moving in the target (T) direction and $N_i^{NT}(d, t)$ refers to the other region (where signal dots moved in the opposite nontarget direction), then $\Delta N_i^c = N_i^T - N_i^{NT}$, where $c = 0$ if the observer responded incorrectly on trial i and $c = 1$ if the response was correct. The perceptual filter is computed as $PF(d, t) = \langle \Delta N_i^1 \rangle - \langle \Delta N_i^0 \rangle$ where $\langle \rangle$ is used to indicate the mean across trials i (see Abbey & Eckstein, 2002). It should be pointed out that the distribution of the number of noise dots at each direction was not Gaussian, but Poisson with mean value smaller than 1 (a consequence of the uniform probability for assigning direction to the noise dots). Although this departure from a circular distribution may have introduced distortions in the recovery of the underlying filter if the system conformed to the linear (static)–nonlinear model (as is typically assumed in reverse correlation studies; Chichilnisky, 2001), our conclusions rely only marginally on this assumption because they are based on a full implementation of an observer model and associated classification images, and the model is not of the simple linear followed by static nonlinearity type anyway.

Subtraction of separable prediction and separability index

Surfaces in Figures 3G and 4 (right column) were obtained by subtracting from $PF(d, t)$ the expression PF^*

$(d, t) = (\sum_d PF(d, t) \cdot \sum_t PF(d, t)) / \sum_{d,t} PF(d, t)$. It is easy to verify that, for $PF(d, t) = f(d) \cdot g(t)$, $PF^*(d, t) = PF(d, t)$. The separability index was computed as $s_i^2 / \sum_i s_i^2$ where s_i is the i th singular value, and singular values were obtained by direct singular value decomposition of the classification image surface (for further details on this method, the reader is referred to Grunewald & Skoumbourdis, 2004; Mazer et al., 2002; Peña & Konishi, 2001).

Modeling

The stimulus within each circled region in Figure 1A is $S_i^T(d, t) = N_i(d, t) + T(d, t)$ or $S_i^{NT}(d, t) = N_i(d, t) + NT(d, t)$ (depending on whether it contained target or nontarget). $T(d, t) = \delta(d - d_T) \cdot s$, where d_T is the target direction (e.g., downward) and s is the number of signal dots ($T(d, t)$ is independent of time t , because signal dots were present at every motion transition throughout stimulus presentation). Similarly for $NT(d, t)$, we convolved S with a directionally selective receptive field defined by a two-dimensional Gaussian function $RF(d, t) = Gauss(\sigma_d, \sigma_t)$ with $\sigma_d = 40$ deg (broadly consistent with, e.g., Cook & Maunsell, 2004; Treue & Martínez Trujillo, 1999) and $\sigma_t = 1$ frame duration = 30 ms (broadly consistent with Bair & Movshon, 2004) to obtain $o_i^T(d, t) = S_i^T * RF$, and similarly for o_i^{NT} . We only used the two outputs at preferred and antipreferred directions d_T and d_{NT} for each convolution, i.e., $pref_i^T(t) = o_i^T(d_T, t)$ and $antipref_i^T(t) = o_i^T(d_{NT}, t)$. Similarly for $pref_i^{NT}$ and $antipref_i^{NT}$. The final response to the circled stimulus region containing signal dots moving in the target direction was $r_i^T = \langle pref_i^T - antipref_i^T \rangle_t$ (this is a scalar) and similarly for the response to the other region r_i^{NT} . Finally, on each trial i , the model selected the region

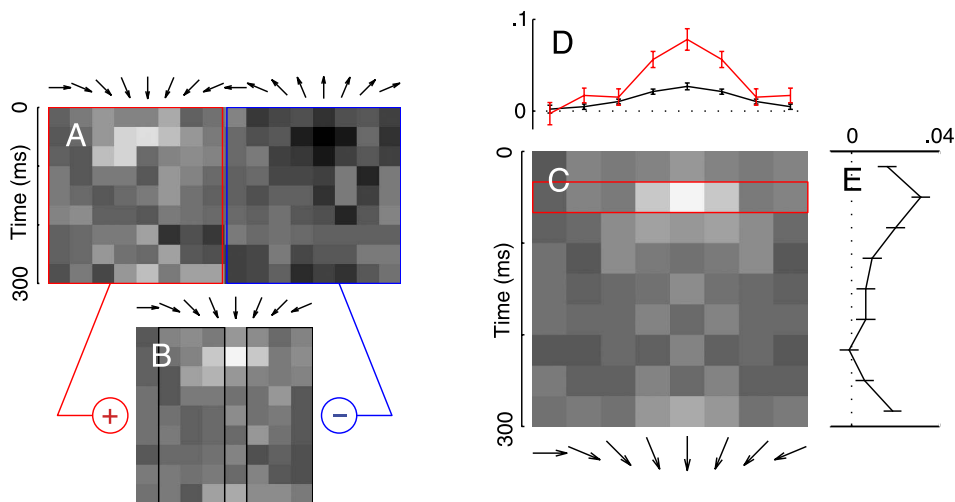


Figure 2. (A) Average classification image (across 3 subjects). Positive values are bright, negative values are dark. (B) Subtraction between the region outlined in red and that outlined in blue in panel A (assumption of symmetry for upward/downward). (C) Obtained from panel B by combining the two regions outlined in black in panel B under the assumption of symmetry around \downarrow . (D) Black line shows panel C averaged across time, red line shows only the values at the motion frame indicated by the red outline in panel C. Panel E shows panel C averaged across direction. Error bars in panels D and E show ± 1 SEM.

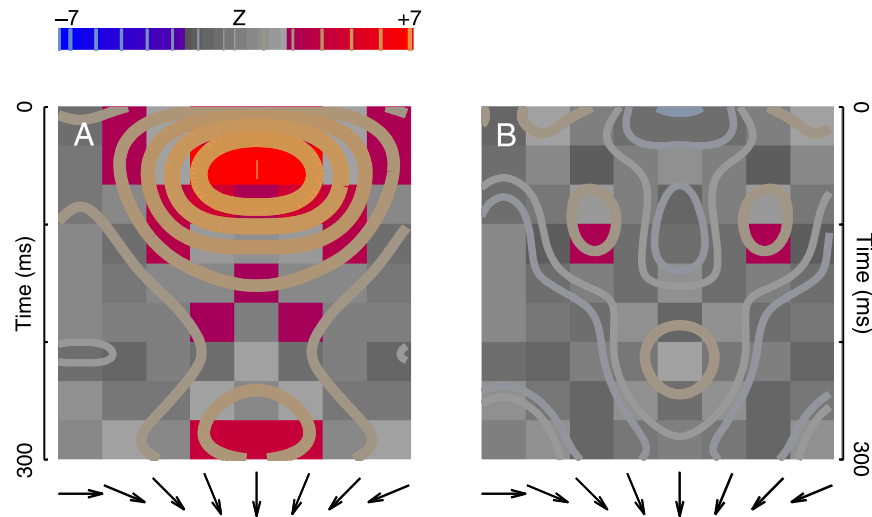


Figure 3. Panel A was derived from panel C in Figure 2 but is plotted here as a Z score map (colored for $|Z| > 2$, blue for negative and red for positive). Contours show smoothed (with 2-D Gaussian of 1 pixel standard deviation) and spline-interpolated surfaces, with color saturation (ranging from blue for negative to red for positive) and line thickness (proportional to absolute value) reflecting modulation intensity. Panel B plotted using the same conventions, but showing the difference between panel C in Figure 2 and the prediction obtained by assuming separability of direction and time (see Methods).

associated with the largest response between r_i^T and r_i^{NT} as being the target region (i.e., it responded correctly if $r_i^T > r_i^{NT}$, and incorrectly otherwise). Before being temporally averaged and combined into r , filter outputs pref and antipref self-normalized their value as follows:

$$\text{pref}(t) = \text{pref}(t) / [\text{pref}(t - \Delta t) + k], \quad (1)$$

with $\Delta t = 3$ frames (90 ms) and $k = 2$ (this expression for normalization is similar to that used by Schwartz & Simoncelli, 2001). These parameter values were selected following pilot simulations. We used the same self-normalization for antipreferred. We challenged the model with 25 K trials taken directly from the data used with psychophysical observers. The black symbols in Figure 6 were obtained by simulating the response of a linear amplifier model (Murray, Bennett, & Sekuler, 2005) that uses the empirical classification image as a template. If $F(d, t)$ is the empirical classification image, the response corresponding to $S_i^T(d, t)$ on trial i was simply obtained by template matching $r_i^T = \sum F(d, t) \cdot S_i^T(d, t)$, and similarly for $S_i^{NT}(d, t)$ to obtain r_i^{NT} . The linear amplifier model chooses the stimulus region associated with largest r as target (correct when $r_i^T > r_i^{NT}$).

Results

General structure of the perceptual filter

Figure 2A shows the perceptual filter (averaged across the 3 subjects) as a function of both direction of motion

(on the x axis) and time (on the y axis). As expected, there are positive modulations in the region of the target direction (\downarrow) and negative modulations in the region of the opposite direction. This is shown by the grey levels within red and blue outlines, respectively. We verified that the measured modulations within these two regions were symmetric within the precision of our measurements, which is not surprising given the nature of the task. We therefore combined the two regions into one region for the purpose of subsequent analysis, as shown in Figure 2B where the region in A indicated by the blue outline has been subtracted from the region indicated by the red outline. We further increased the signal-to-noise ratio of our measurements by symmetrically averaging the two regions indicated by the black outlines in B, based on the assumption that off-target directions would have symmetrical effects. This assumption is not only reasonable, but we also verified that the data conformed to it with the possible exception of subject 2 (S2 in Figure 4) who showed a slight bias for a direction immediately clockwise to target direction (however this bias does not affect the conclusions of this study). Symmetrical averaging of B results in C, which shows a clear positive peak at the target direction between 30 and 60 ms.

Panel D plots the average across time for the surface in C (black line), panel E the average across direction. The directional tuning curve in D shows that, as expected, the perceptual filter peaks at target direction. The half-width at half-height for this function is 36 ± 7 deg (mean \pm SEM). The temporal impulse response function in E shows that the perceptual filter reaches its maximum peak very early, then decreases, and finally increases again. These features of the perceptual filter are also visible in the Z score map (Figure 3A), where a statistically significant

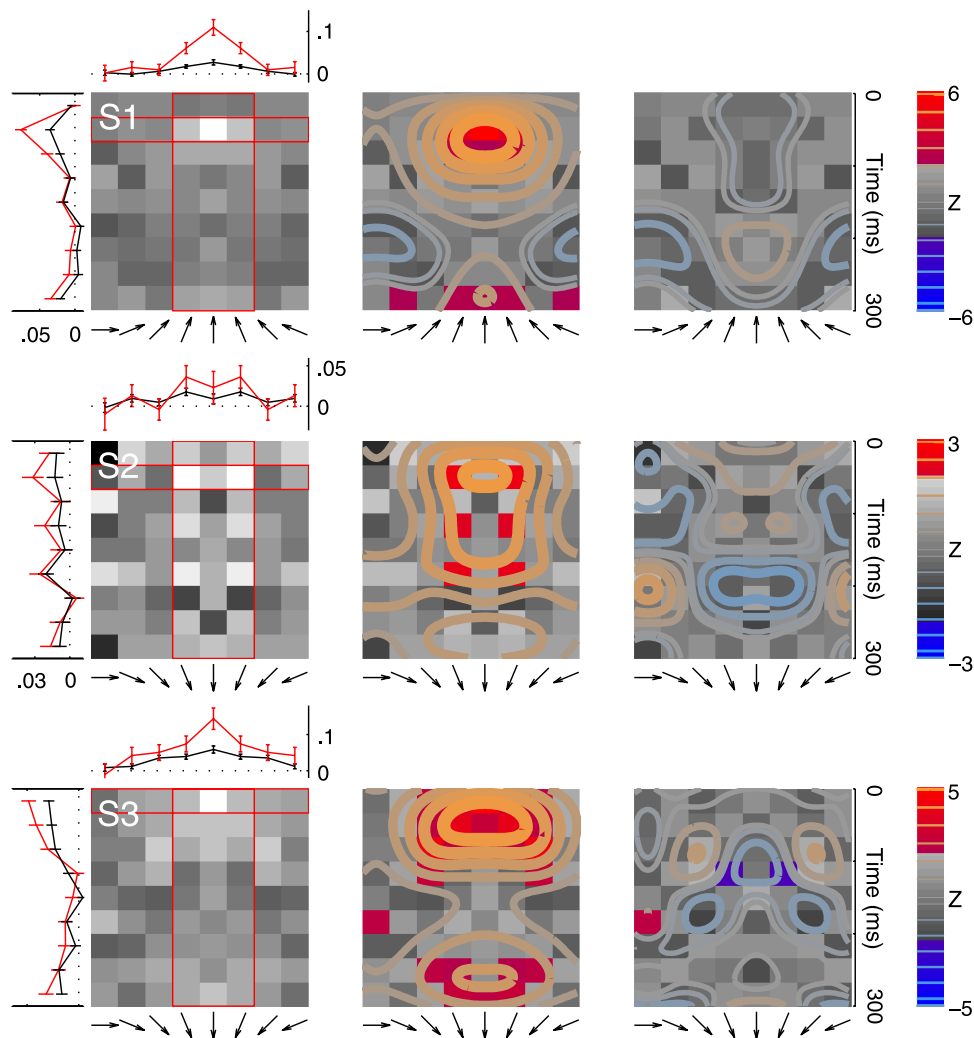


Figure 4. Plots of panels C–E from Figure 2 and the two plots from Figure 3 are shown here for individual subjects. An additional temporal profile is shown in red for pixels pooled within the red vertical rectangular outline rather than across the entire surface (as is the case for the black line). Notice that the directions on the x axes are different for S1 because this subject was assigned \uparrow as target direction rather than \downarrow for S2–S3. The poorer quality of the classification image for S2 is reflected in its lower Z score range compared to the other two subjects (see Z legend to the right).

modulation is observed both within the early and within the very late region of the surface. This late modulation was unexpected (see Discussion).

Separability of directional and temporal responses

An obvious and important question is whether directional tuning varies with time. This question can also be phrased in the following terms: Is the surface in Figure 2C separable across direction and time? In other words, is it well accounted for by simply multiplying D and E for every combination of direction and temporal event? We computed the difference between the measured perceptual filter and the equivalent surface that is obtained by multiplying the two marginal averages (D and E; for details, see Methods). A Z-score map of this difference is plotted in

Figure 3B, which shows very little modulation (there is only 1 pixel (symmetrically on both sides) that meets $|Z| > 2$ (colored in Figure 3B), and none of the subjects shows a similar pattern). This is a clear indication that the perceptual filter is well described by a temporal modulation in the amplitude of a fixed directional tuning curve. We also attempted to measure half-width at half-height at different time points throughout the surface and were unable to find consistent changes across time (not shown).

As an additional test of the independence between directional and temporal responses, we computed a separability index based on singular value decomposition (for details, see Methods). This index ranges from 0.125 (not at all separable) to 1 (entirely separable). The value for the average human classification image was (mean \pm SEM) 0.87 ± 0.04 (for individual subjects (S1–S3) 0.73 ± 0.07 , 0.65 ± 0.08 , 0.78 ± 0.06), confirming our earlier conclusion based on Figure 3B.

Data for individual subjects

Figure 4 plots the data of Figures 2C–2E and Figures 3A and 3B separately for each of the 3 observers. S1 and S3 are highly consistent with the average perceptual filter detailed in Figures 2 and 3. More specifically, both subjects show the same temporal modulation pattern with two positive regions, one early and one late. S2 generated a perceptual filter of very poor quality (see Z score scale to the right), so it is hard to establish whether the pattern is consistent with the other two subjects. It certainly showed a tendency towards larger modulation in the earlier part of the filter, and there is a hint of late positive modulation (see contour plot) even though it did not reach statistical significance. Overall, S2's data do not contradict the other two subjects, and it shows the same coarse trend, but the quality of this subject's data set was not sufficient to retrieve a detailed description of the perceptual filter. For this subject, we similarly attribute the lack of statistically significant modulations at target direction (downward) to the poor quality of the corresponding perceptual filter. The directional slice across the surface (red trace for S2 in Figure 4) shows that there was a positive modulation at target direction, but the associated error was large and the resulting Z score not significant.

Modeling

We modeled the shape of the perceptual filter using simple computational tools (see Methods). The general structure of the model is outlined in Figures 1D–1F. The stimulus (shown in C) is initially filtered by a brief directionally tuned operator (D). Only the two outputs

from operators centered on target (\downarrow) and opposite-to-target (\uparrow) directions were used. Each output normalized itself after a delay of 90 ms (E). This normalizing signal functions as a servo-mechanism, but the presence of a temporal delay leads to an oscillatory pattern that mimics the bimodal behavior observed in the data. This is demonstrated by the simulation in Figure 5A, where the empirical data (left) are juxtaposed with the simulated perceptual filter (right), which shows two positive peaks at early and late points within the surface (the model surface (right) was obtained by simulating a psychophysical experiment where the model acted as observer). We selected a stimulus signal-to-noise-ratio (SNR) that generated a model performance of 70% correct responses, in close agreement with human performance (average across the 3 subjects was 73%). The corresponding model SNR was 10% of the human SNR (this translates into a noninteger number of noise dots; for the model, we could use noninteger numbers of noise dots because the input to our simulations did not consist of a stimulus image (for which the number of dots must take an integer value), but of a distribution like in Figure 1C).

The 10-fold difference in threshold SNR between model and human performance is very large. What could account for this discrepancy? A good candidate could be late internal noise. Human observers do not behave deterministically, i.e., there is a degree of intrinsic randomness to their responses that is independent of the stimulus. This fact can be demonstrated by performing a double-pass experiment, in which the same physical stimulus is presented twice. If the observer behaved deterministically, the same response should be generated on both passes of an identical stimulus. This is not always the case due to observer internal noise (Burgess & Colborne, 1988).

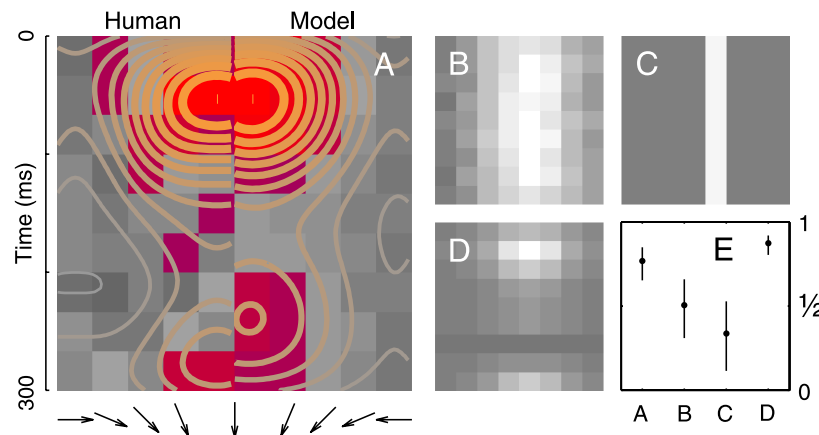


Figure 5. Results of model simulations. (A) Left half of the surface is taken from Figure 3A, right side shows the classification image corresponding to the simulations using the model outlined in Figures 1D–1F (for details, see Methods). Panel B shows the classification image obtained when the delayed normalization stage is removed from the model. Panel C shows the optimal signal-matched template. Panel D shows the separable prediction for the average human classification image in Figure 2C (for details, see Methods). Panel E plots the pixel-by-pixel correlation coefficient between the psychophysical classification image (left half of the surface in panel A) and each one of the four artificial classification images detailed above and depicted through panels A–D (the respective correlation coefficients are indicated by the same letters on the x axis in panel E).

We wished to establish whether internal noise could account quantitatively for the abovementioned discrepancy between our noiseless model simulations and human psychophysical data. We ran a series of double-pass experiments (see [Methods](#)) and determined human–human consistency (the % of trials on which the human observer is consistent with his/her own responses between two identical stimulus presentations) to be (mean \pm SEM) 0.61 ± 0.01 (S1), 0.70 ± 0.01 (S2), and 0.80 ± 0.02 (S3). Human–human consistency can be used to establish upper and lower bounds on the best achievable human–model consistency, i.e., the % of trials on which the model response matches the human response (for details, see Neri & Levi, 2006; for related methods in the electrophysiological literature, see Haag & Borst, 1997; Neri, 2006; van Hateren & Snippe, 2001). This optimal range is indicated by the grey area in [Figure 6](#). For two out of the 3 observers (S1–S2), the human–model consistency achieved by our model is satisfactory (points fall within the grey area). For one observer (S3), internal noise is not sufficient to explain the discrepancy between model responses and human responses. We do not have a ready explanation for this small discrepancy.

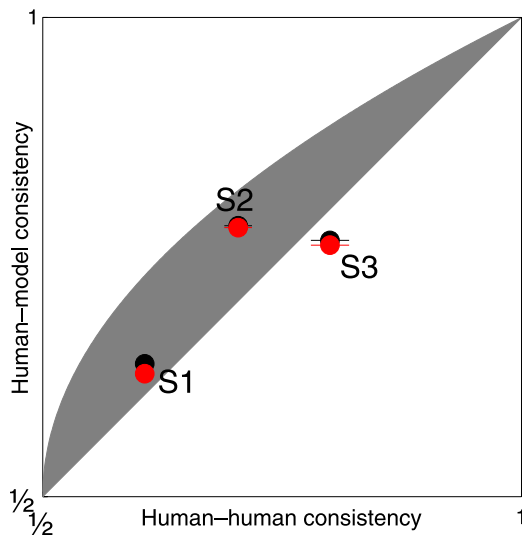


Figure 6. Human–model consistency (the % of trials for which the model response matches the human response) is plotted on the y axis against human–human consistency (the % of trials for which the human observer gives the same response to two identical presentations of the same stimulus). Given human–human consistency C_{hh} , the best achievable human–model consistency C_{hm} may lie anywhere between $C_{hh} \leq C_{hm} \leq \frac{1+\sqrt{2C_{hh}-1}}{2}$ for a 2AFC task (for details on how this formula was derived and its correct interpretation, see Neri & Levi, 2006). This range is shown in grey. Red symbols refer to the delayed self-normalization model, black symbols to the response sequence obtained by template-matching the stimuli with the empirically derived classification images of each subject (see [Methods](#)).

The human–model consistency plotted on the y axis in [Figure 6](#) was obtained by matching the human responses with model responses to identical stimuli, i.e., at the SNR used with human observers. For those SNR regimes, our model performed correctly on every trial or nearly so. This means that the result in [Figure 6](#) is not specific to our model but applies to any model that correctly detects the target on every trial. For comparison, we report human–model consistency for a linear amplifier model (Murray et al., 2005; Neri & Levi, 2006) that uses the classification image itself as template for matching (black symbols in [Figure 6](#); see [Methods](#)). It can be seen that this model is equally (or possibly slightly more) consistent with the human responses. In summary, [Figure 6](#) only allows us to conclude that the high performance of our model is not inconsistent with the human data when one takes into account internal noise, but the results in [Figure 6](#) do not provide specific evidence in favor of our model over many other potential candidates (e.g., an ideal template that only integrates motion energy at the direction of the target for the entire duration of the stimulus, which also correctly detects the target on every trial).

More specific evidence in support of our model is provided by [Figure 5E](#), where we show that it is superior to a few others in its ability to generate classification images that resemble the human data. As a measure of how closely the simulated classification image matches the average human classification image, we simply took the pixel-by-pixel correlation between the two. The correlation for the model discussed so far (classification image shown in [Figure 5A](#)) is 0.77 (plotted in [Figure 5E](#) at abscissa label “A” with 95% confidence interval). For comparison, we consider the classification images generated by (1) the same model but without the delayed normalization stage (classification image in [Figure 5B](#)), (2) a linear amplifier model that uses the optimal signal-matched template (classification image in [Figure 5C](#)), and (3) a linear amplifier model that uses the separable prediction for the human classification image as template ([Figure 5D](#)).

The corresponding correlation coefficients are plotted in [Figure 5E](#), which shows that our model provides a better account for the empirical classification image than both models 1 and 2 and is close to the separable model. The high accuracy of the separable model is not surprising given the result in [Figure 3B](#), where we showed that the human classification image is separable. However, the separable model is simply a descriptive, not a mechanistic model. The self-normalizing model has the advantage that it is mechanistic (i.e., it potentially informs us about the mechanisms underlying the observed data) and at the same time accounts for the classification image almost as well as the separable one. The importance of the delayed normalization stage is emphasized by the drop in correlation that is observed when this stage is removed from the model (for models A and B, compare the two correlation coefficients in [Figure 5E](#)).

- area MT of macaque brain. *Nature*, 409, 1040–1042. [[PubMed](#)]
- Peña, J. L., & Konishi, M. (2001). Auditory spatial receptive fields created by multiplication. *Science*, 292, 249–252. [[PubMed](#)]
- Perge, J. A., Borghuis, B. G., Bours, R. J., Lankheet, M. J. M., & van Wezel, R. J. (2005). Temporal dynamics of direction tuning in motion-sensitive macaque area MT. *Journal of Neurophysiology*, 93, 2104–2116. [[PubMed](#)] [[Article](#)]
- Ringach, D. L. (1998). Tuning of orientation detectors in human vision. *Vision Research*, 38, 963–972. [[PubMed](#)]
- Ringach, D. L., Hawken, M. J., & Shapley, R. (1997). Dynamics of orientation tuning in macaque primary visual cortex. *Nature*, 387, 281–284. [[PubMed](#)]
- Schwartz, O., & Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control. *Nature Neuroscience*, 4, 819–825. [[PubMed](#)] [[Article](#)]
- Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Research*, 38, 743–761. [[PubMed](#)]
- Smith, M. A., Majaj, N. J., & Movshon, J. A. (2005). Dynamics of motion signaling by neurons in macaque area MT. *Nature Neuroscience*, 8, 220–228. [[PubMed](#)]
- Srinivasan, M. V., Laughlin, S. B., & Dubs, A. (1982). Predictive coding: A fresh view of inhibition in the retina. *Proceedings of the Royal Society of London B: Biological Sciences*, 216, 427–459. [[PubMed](#)]
- Tadin, D., Lappin, J. S., & Blake, R. (2006). Fine temporal properties of center-surround interactions in motion revealed by reverse correlation. *Journal of Neuroscience*, 26, 2614–2622. [[PubMed](#)] [[Article](#)]
- Treue, S., & Martínez Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, 399, 575–579. [[PubMed](#)]
- Vajda, I., Borghuis, B. G., van de Grind, W. A., & Lankheet, M. J. (2006). Temporal interactions in direction-selective complex cells of area 18 and the posteromedial lateral suprasylvian cortex (PMLS) of the cat. *Visual Neuroscience*, 23, 233–246. [[PubMed](#)]
- Van Hateren, J. H., & Snippe, H. P. (2001). Information theoretic evaluation of parametric models of gain control in blowfly photoreceptor cells. *Vision Research*, 41, 1851–1865. [[PubMed](#)]