Dynamic properties of orientation discrimination assessed by using classification images

Isabelle Mareschal*, Steven C. Dakin, and Peter J. Bex

Institute of Ophthalmology, University College London, Bath Street, London EC1V 9EL, United Kingdom

Edited by George Sperling, University of California, Irvine, CA, and approved January 31, 2006 (received for review August 19, 2005)

Recent physiological studies indicate that the tuning properties of neurons under acute preparation in primary visual cortex can change over time. We used a psychophysical reverse correlation paradigm to examine the potential repercussions of this neuronal property for human observers' ability to discriminate the orientation of targets over time. Observers were required to identify the orientation of a Gabor target presented within dynamic white noise. Frames from the noise movies were pooled to compute dynamic classification images (CIs) associated with the observers' discrimination performance, which then were fit with a weighted difference-of-Gabor function. Best-fitting templates were temporally bandpass, tuned to more oblique orientations than the stimulus but, crucially, did not change over time. The results suggest that the template for orientation discrimination is selected within the first 50 ms of stimulus onset and that, unlike the response of single cells, there is no measurable dynamic component to either orientation or spatial frequency tuning of human orientation discrimination.

psychophysics | vision

Visually responsive neurons are sensitive to structure within a region of space referred to as their "classical receptive field." Although this field has been assumed to be a largely fixed property of the neuron, there is now considerable electrophysiological evidence that neurons' tuning to various visual properties can vary substantially over time (1–5). For example, the orientation tuning of neurons in macaque V1 (measured with an acute preparation) is initially broad, but ≈40–50 ms after stimulus onset, responses become increasingly selective. After a peak in selectivity at ≈60–100 ms, tuning becomes broader (1, 4, 6) and sometimes indicates a radical change in preferred orientation (1, 4). A similar change in tuning has been reported for spatial frequency (SF) (2, 3) and binocular disparity (5). However, these findings are controversial: using intracellular recordings (7) and in awake preparations, both Celebrini *et al.* (8) and Mazer *et al.* (3) have failed to find a dynamic shift in orientation tuning. In light of these differences, we sought to examine whether the changes in tuning that had been observed in the first few hundred milliseconds of stimulus presentation might be reflected in the behavior of human observers.

A number of psychophysical studies have examined the influence of stimulus exposure duration on visual performance. Form and sharpness discrimination (9–11), stereoacuity (12) as well as contrast sensitivity for moving patterns (13) improve with exposure duration. Contrast sensitivity is lower for briefly presented than for steady-state stimuli but in a manner inconsistent with a change in tuning (14). Although there is a clear dependence of sensitivity on duration, the change in tuning (i.e., peak sensitivity and bandwidth) of the underlying detectors has not previously been examined systematically as a function of the time course of the stimulus (i.e., for different epochs within the stimulus presentation).

Here, we examine how visual selectivity changes over time using a "classification image" (CI) or psychophysical reverse-correlation technique. Reverse-correlation paradigms estimate the correlation between the behavior of a sensory system and some attribute of a stochastically varying stimulus (typically the luminance of pixels in white noise). In electrophysiology, for example, the technique consists of presenting rapidly changing one- or two-dimensional luminance noise patterns over a neuron's receptive field (e.g., refs. 1, 15, and 16). Noise that is positively correlated with subregions of a receptive field (i.e., light pixels within an "on" zone) will increase the probability of firing. Noise that is negatively correlated (i.e., light pixels within an "off" zone) will decrease the probability of firing, whereas uncorrelated noise produces no change. The correlation between the noise in a given region before the spike generation and the number of spikes elicited has been used to map out the fine structure of receptive fields in two-dimensional space (1, 15, 16) and disparity (5, 17). More recently, reverse-correlation techniques have been elaborated to include higher-order statistical properties of the noise (18, 19).

An analogous paradigm in psychophysics relies on the same principle of correlation between noise and behavior to yield CIs that in turn can be used to infer underlying "perceptual receptive fields" or "templates": the regions of an image that are important for performing particular visual tasks (20–28, †, ‡). In general, for a stimulus embedded in a white-noise mask, positive correlations between stimulus and noise promote detection of the target, whereas negative correlations inhibit detection. By adding the noise fields associated with correct performance and then subtracting the sum of those noise fields associated with errors, one creates a CI in which the luminance of each pixel encodes whether it has a positive or negative influence on the observer's response. If the human visual system operated as an ideal system, then the CI should, for a detection task, converge on an image of the stimulus. For a discrimination task in white noise, the ideal CI is the arithmetic difference between the two possible targets; i.e., locations that are informative about difference are highlighted. In practice, the CI reflects observer's (nonideal) visual processing, which cannot be wholly attributable to the physics of the stimulus. For example, CIs derived for detecting the curvature of illusory contours show contour completion (29), and CIs derived for discriminating the contrast polarity of stimuli inducing the Craik–Cornsweet illusion show "filling-in" (28). Solomon (25) derived the CIs for either detecting a Gabor patch in noise or discriminating its orientation. The results showed that the phase of the target was irrelevant for the detection task, a finding that he concluded was analogous to phase insensitivity observed for complex cells. Furthermore, the CI for the discrimination task was more oblique than the target, suggesting that observers were basing their decision on the activity of noise-limited channels that showed the greatest differential response to the two target orientations (''off orientation looking'').

In summary, static CIs have been used to identify the human receptive fields used to perform visual tasks. In this work, we adapt the CI paradigm to examine changes in spatio-temporal tuning in

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: CI, classification image; SF, spatial frequency; c/deg, cycles per degree.

*To whom correspondence should be addressed. E-mail: i.mareschal@ucl.ac.uk.

†Xing, J. & Ahumada, A. J. (2002) *J. Vision* **2,** 343a (abstr.).

‡Chen, K. Y., Eckstein, M. & Shimozaki, S. S. (2003) *J. Vision* **3,** abstr. 182.

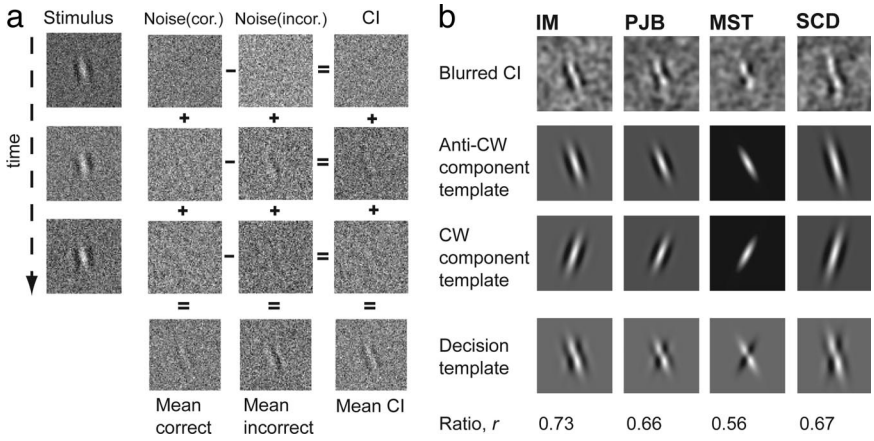

Fig. 1. CI technique and fitting procedure. (*a*) Subjects reported whether a Gabor target, continuously present within 32 frames of dynamic white noise, was tilted clockwise or anticlockwise of vertical. Noise frames were stored according to the observers' response into the ''correct'' (center left column) or ''incorrect'' (center right column) category. The 32 single frame CIs obtained at the end of 10,000 trials (rightmost column) were obtained from the difference between incorrect and correct CIs. Mean CIs for the correct-incorrect categories can be obtained by averaging the corresponding columns. (*b*) Fitting procedure. Because observers are performing a discrimination task, the optimal template for performing this task is the difference between the two possible targets. Consequently, CI data were fit (having first been spatially blurred) with a weighted difference of two component templates (Gabors, one the mirror image of the other). Fits had six parameters, five for the components and one for their weighting. The component-pairs are illustrated in the two rows below the blurred CIs. The decision templates are the differences of the component templates weighted by the *r* value given below. Note that fitted decision templates account well for the appearance of the CIs including the checkered appearance of the template for M.S.T. Most of the decision templates were dominated by a single component indicating that some observers' performance was limited by their ability to detect the target rather than tell what orientation it was (see *Supporting Text* and Fig. 4 for details).

the human visual system over short periods of time and so scrutinize the discrepancy in the electrophysiological literature concerning the temporal evolution of orientation tuning in single neurons.

Subjects were required to identify the orientation of a target stimulus (a Gabor patch) embedded in dynamic random noise, which was maintained at a contrast level near its detection threshold. Over the course of an experiment, we collected the noise samples associated with correct and incorrect responses and used them to compute a dynamic CI, i.e., a frame-by-frame map of which parts of the stimulus the observer uses to perform the task. We then fit the CIs with weighted difference-of-Gabor patterns that allowed us to examine the dynamic evolution of various attributes of the perceptual receptive field (e.g., its orientation tuning).

Results

Fig. 1 illustrates three CI frames associated with correct and incorrect identification of the tilt of an anticlockwise Gabor target embedded in noise. The 32 frames from any one trial are similarly categorized according to psychophysical response and summed, across trials and within response category. Inspecting the mean CIs (Fig. 1*b*, top rows), two features emerge. The first is that CIs are more oblique than the target orientation, a finding reported by Solomon (25) who suggested that discrimination was based on the differential responses of template-pairs, such that more obliquely oriented template pairs are more sensitive to orientation difference than the closest matching template pair. The second feature is that the CIs appear ''Gabor-like'' for some observers (e.g., I.M., 7) but more checkerboard-like for others (particularly the only psychophysically inexperienced observer, M.S.T.). Although checkerboard-like CIs are optimal templates for ''signal discrimination,'' Gabor-like templates are optimal for ''signal detection.'' Because detection must precede discrimination, it may be that our experienced observers were limited by early (i.e., detection-limiting) noise, whereas our inexperienced observer was limited by later (discrimination-limiting) noise. We speculate that one consequence of psychophysical experience may be to reduce late (decision-based) noise.

To examine the temporal dynamics of the decision CIs, we fit combinations of frames (weighted by a temporal-Gaussian function; see *Methods*) within the 32-frame sequence. Fig. 2 shows how parameters of the component templates vary over time (data are from experienced observers: most frames from observer MST were rejected according to our analysis based on the statistical criteria described below, so his data are not shown). In Fig. 2*a*, the estimated SF of the component template is plotted for the three observers in the two different SF conditions [open symbols are at 2.5 cycles per degree (c/deg); filled at 7.0 c/deg]. The shaded gray areas represent the range of SFs over which the template fits were allowed to vary; the dashed lines represent the actual SF of the target. A few points emerge from this plot. First, in both the low and high SF conditions, the estimated template SFs approximate the target SF (slightly lower in the high SF condition). Second, the template SFs do not vary over time. The slopes of the best-fitting lines to the individual observers' data weighted by their error bars were flat (−0.0014, I.M., 2.5 c/deg; 0.0013, I.M., 7 c/deg; 0.0003, P.J.B., 2.5 c/deg; 0.0043, P.J.B., 7 c/deg; −0.0009, S.C.D., 2.5 c/deg; and 0.0014, S.C.D., 7 c/deg) and have not been included on the graphs for visual clarity. Bootstrapped estimates of these slopes were calculated, and the standard deviations (SDs) were on average 0.002. In addition, these values are similar to slopes calculated on a subset of temporally unweighted frames that reached significance and were found to be not significantly different from zero (0.0004, I.M. 2.5 c/deg; 0.0133, I.M., 7c/deg; 0.0007, P.J.B., 2.5 c/deg; 0.001, P.J.B., 7 c/deg; 0.004, M.S.T., 2.5 c/deg). Interestingly, although the stimulus was present throughout, most of the significant frames fell within the first 10–300 ms. Note that there was very little variation on the parameters within the limits that were imposed by the fitting procedure, indicating that constraints were not hindering the fitting procedure.

Fig. 2*b* plots the estimated component template orientation for both SF conditions, with the dashed line corresponding to the actual orientation of the target. The estimated template orientation was more oblique than the target's by ≈10° at both SFs. Again, there was no change in these values over time. Estimates of the component SF bandwidths are plotted in Fig. 2*c*, and component orientation bandwidth estimates are in Fig. 2*d*; both are similar to the actual value of these parameters in the stimulus (dashed line).

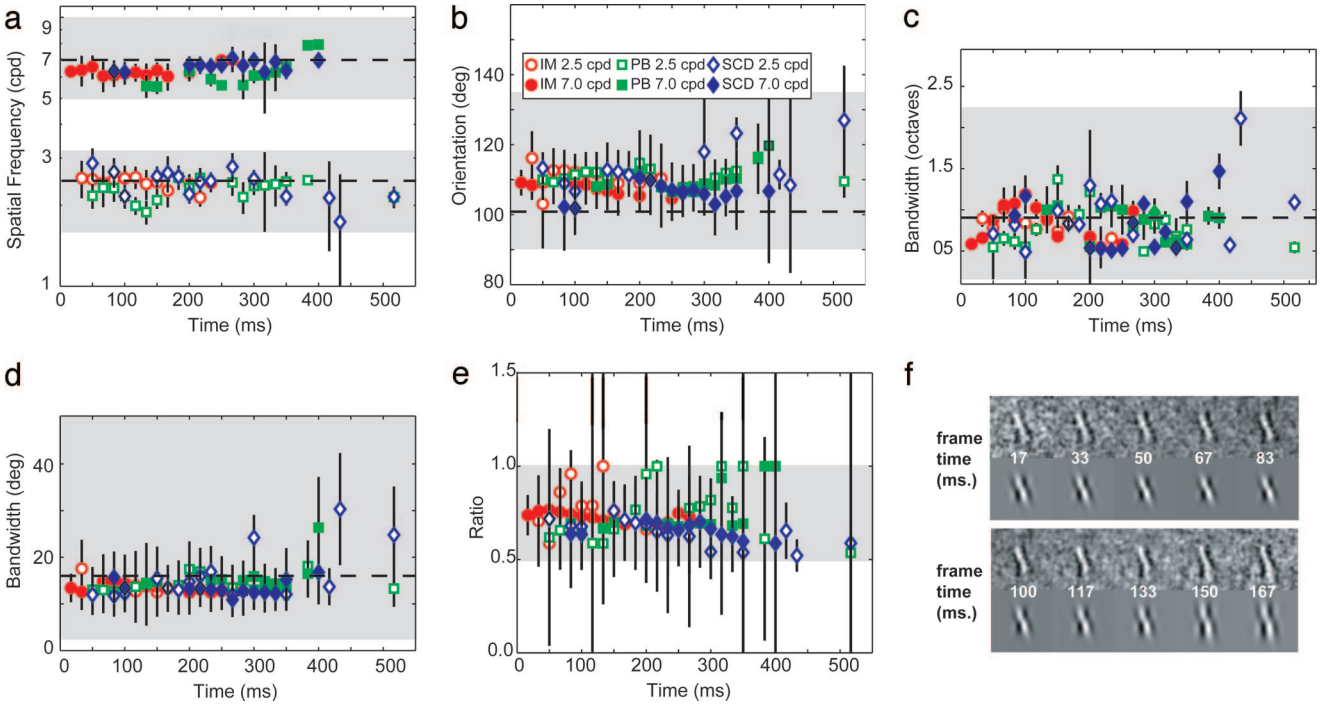

Fig. 2. Changes in component templates over time. Plots show parameters derived from fitting weighted combinations of CIs; only parameters that met our statistical criteria (described in *Methods*) are shown. (*a*) Template SF at 2.5 c/deg (open symbols) and 7.0 c/deg (filled symbols) for three observers. Shaded areas indicate the range over which the fitting procedure could vary; dashed lines are the target SF. Error bars are 95% confidence intervals. (*b*) Template orientation. (*c* and *d*) Template SF-bandwidth (*c*) and template orientation-bandwidth (*d*) derived from best fitting Gaussian functions. (*e*) Relative contribution of the component templates to the decision template. (*f*) Examples of CIs, and their corresponding best-fitting decision templates, from the first 166 ms of stimulus presentation (observer I.M., SF of 7.0 c/deg). The CI remains largely unchanged over this period.

The ideal discrimination template is the weighted sum of the two target components, each weighted by 0.5. Estimates of the ratio in which the component templates contribute to the discrimination template are plotted in Fig. 2*e*. Although the error bars are larger in this plot, there is no systematic change in the ratio over time and the average clusters ≈0.7 for the three observers. Given that these values are neither 0.5 (pure discrimination) nor 1 (pure detection), this result suggests that the CIs obtained on this task reflect a combination of limits on observers' detection and discrimination.

Contrast thresholds (the average contrast level associated with the final 20 response-reversals, i.e., switching from correct-to-incorrect or vice versa, within a run) did not vary between the two different SF conditions for the observers and were, for the 2.5 and 7 c/deg conditions, respectively, 3.9% and 3.7% (I.M.), 6.6% and 6.1% (P.J.B.), 8.0% and 8.6% (S.C.D.), and 3.6% (M.S.T.). Because an ideal observer performing this task produces a contrast threshold of 1.05%, our observers' efficiencies for this task are 7.2% and 8.1% (I.M.), 2.5% and 3.0% (P.J.B.), 1.7% and 1.5% (S.C.D.), and 9.0% (M.S.T.).

Our data do not show any change in time across the different component template parameters. However, a final point can be made regarding the sensitivity of the mechanisms involved in the task throughout the stimulus presentation. As mentioned above, the range of frames producing statistically significant fits did not cover the entire duration of the stimulus presentation. To examine which frames were used to perform the discrimination, we cross-correlated each unweighted frame of the CI sequence to the decision template that best fit the mean CI across all frames (i.e., for I.M. at 7.0 c/deg, the decision template from Fig. 1*b* was correlated with each of the 32 frames of the 7.0 c/deg CI sequence and likewise for the other observers).

The results of this analysis are plotted in Fig. 3. Solid symbols represent statistically significant correlations between the decision templates and individual CI frames. The curves are the Gaussian functions that best fit the correlation coefficient data. Temporal tuning is clearly bandpass, peaking at ≈160 ms, for both SF conditions. Interestingly, subjects use only the initial portion of the stimulus to make a judgment, even though the stimulus is present throughout and it would be more efficient for them to monitor the entire sequence. We expected the functions to plateau after an early peak that corresponded to the time constant of a leaky integrator for contrast sensitivity. We also note that, on average, there was no difference in the peak latency in the high and low SF conditions. This finding is surprising because contrast sensitivity for noise-free stimuli, increases with exposure duration, and the rate of increase is greater at low SFs than high (30), so we would expect the 2.5 c/deg target to reach threshold contrast more quickly than the 7.0 c/deg target.

Single-frame CIs begin to achieve significant correlation with the template at ≈40–50 ms, consistent with the onset of measurable responses in single cells (1, 31). Also, and importantly, the presence of a series of significant frames for all observers means that the CIs did not substantially change over these frames (i.e., that in the raw frames, as for the temporally smoothed ones, the parameters did not vary). Note that the lack of correlation of the later frames to the mean CI was not the result of a change in their tuning characteristics, but rather that these were no longer significantly different from noise.

Discussion

We used a behavioral technique to measure the visual mechanisms that support a simple orientation discrimination task. Decision templates were estimated by using a psychophysical reverse correlation technique. The dynamic properties of templates were examined at 16.6-ms intervals throughout a 531-ms stimulus presentation. We report the following results.

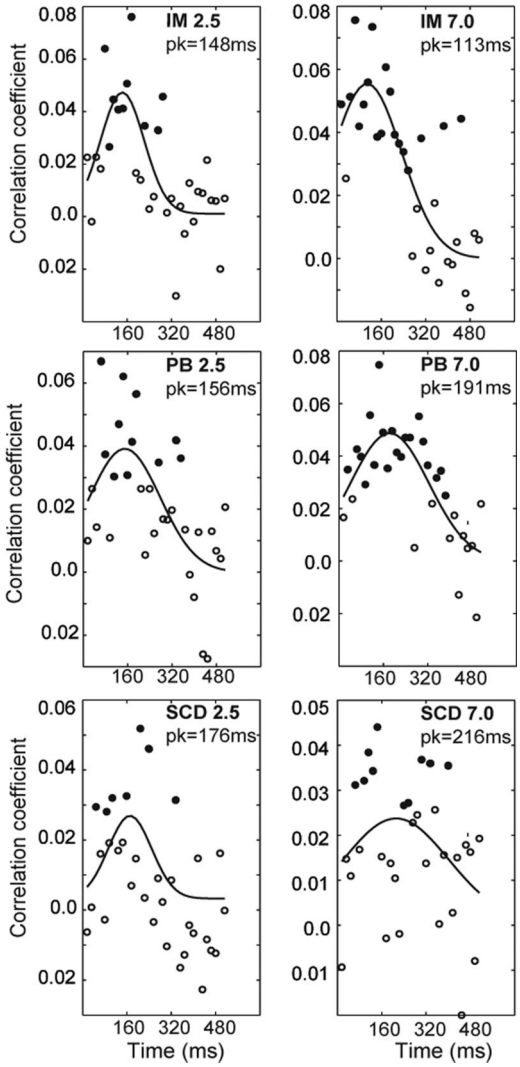

Fig. 3. The cross-correlation between the decision template (fit to CIs averaged across all frames) and the CI on each frame. Filled symbols show significant correlations ($P < 0.05$); open symbols show points that did not reach statistical significance. Solid curves are the Gaussian functions that best fit the data.

1. Psychophysical CIs resemble the weighted combination of two Gabor-like component templates [a result that can be exemplified by using a simple variant on the "hard-threshold" model of Solomon (25); see *Supporting Text* and Fig. 4, which are published as supporting information on the PNAS web site].
2. Component templates were more obliquely oriented than the target.
3. There was no significant change in the tuning of component templates over time.
4. The decision template did not vary systematically over time (i.e., component templates contributed in a constant ratio).
5. The templates were temporally bandpass: observers' discriminations were based on visual information occurring ≈20–300 ms after stimulus onset, regardless of the target SF.

Mismatch in Orientation Between Template and Target. The reliable mismatch in orientation between the templates and the target was present throughout presentation of the target at both SFs tested. We speculate that this result may be attributed to "off-orientation looking," analogous to "off-frequency listening" in auditory psychophysics (32). Off-frequency looking refers to the counterintuitive observation that whereas the most sensitive channel for target detection is the one that is maximally responsive to the particular SF and orientation of the target, the principle of univariance dictates that this channel is not able to discriminate between stimuli that are of equal but opposite distance from its own peak. For our experiment, a channel tuned to vertical would be equally responsive to stimuli that are equidistant from the peak sensitivity of the channel. To discriminate the stimuli, observers must monitor orientation channels on either side of the channel that is optimally sensitive to the mean of the two stimuli. By relying on the channels that are not optimally tuned for the midpoint between the stimuli, the observer's template is biased toward these channels and away from the orientation of the target. Making some reasonable assumptions about the channels (e.g., Gaussian orientation tuning, Poisson noise), Itti *et al.* (33) show that the optimal orientation-offset between channels (i.e., that maximizes the derivative of the log likelihood: known as Fisher information) is equal to $\sqrt{2\sigma_\theta^2}$, where σ^θ is the channel's orientation bandwidth. For our stimuli, this result indicates that channels at 112° and 68° maximize Fisher information. This finding is closely consistent with our results in Fig. 2*b* (note that all data are normalized to anticlockwise-oriented stimuli) where data points fall at ≈111°. Fisher information predicts that there should be no such bias for the discrimination of orientations that differ from one another by an amount ≥45°. We confirmed this prediction in one observer (I.M.) for orientation discrimination of targets that differed by either 80° or 90°. Under these conditions, the CI was no longer more oblique than the target.

Temporal Sensitivity. When we examined the temporal sensitivity of the templates, two findings emerged. First, we found no consistent difference in latency of processing between the low and high SFs across observers, and second, the temporal dependence of the CI was bandpass. We consider these points in turn.

With respect to the lack of effect of SF on processing latency, studies of threshold (30) and suprathreshold (34) contrast sensitivity have consistently shown that the time constant of temporal integration is faster at low than high SFs. In spatial form vision, the notion of coarse-to-fine processing (35) has received psychophysical support from reaction time and visual evoked potential (VEP) studies (36–38). There is also a growing body of physiological data that supports this idea (2, 3, 5). It might therefore be expected that the time constant of the discrimination template for the low SF target would be faster than that for the high SF target. However, we found no consistent difference in the peak temporal tuning of the CIs for low and high SF targets. In a Vernier acuity task using stimuli that were equated for visibility, Waugh and Levi (39) also failed to find a shift in SF tuning over time, and McSorely and Findlay (40) found that square-wave gratings that were presented in fine-to-coarse SF sequence resembled a static square grating more than a coarse-to-fine sequence. Collectively, the results suggest that coarse-to-fine temporal processing is to some extent dependent on the task used to measure it.

In considering the bandpass temporal dependence of the CI, given that the target was present at threshold contrast on each frame, it might be expected that the discrimination template and the CI it produces should be relatively fixed throughout the stimulus presentation. A small effect of forward and backward masking is likely to occur both at the beginning and end of the stimulus, so sensitivity may be lower at these points, and consequently one might expect some loss of a clear CI for the first and last one or two frames of the CI movie. However, in the presence of a mask presented for the same duration as the stimulus, masking should be symmetrical (41), but our results show clear asymmetry. In addition, Foley and Chen (42) measured the strength of masks as a function of the temporal and spatial offsets between mask and stimulus. They found that when the mask and stimulus were spatially overlapping,

backward masking was much weaker than forward masking. If this finding applies to our data (because our mask was always spatially overlapping), we would have expected to see a shallow (if any) fall-off in sensitivity. This expectation was not the case. We therefore conclude that the temporal characteristics of the CIs reflect the tuning of the discrimination template rather than any known temporal properties of contrast sensitivity or masking. A final mention can be made about different response strategies that observers may employ. Although attentional effects might have influenced the overall performance of observers, we found no changes in the contrast thresholds near the end portions of a run (1,000 trials), suggesting that observers were maintaining fairly constant attention levels and also probably similar response strategies. However, the overall influence either of these factors may have on our results cannot be distinguished from our data. Together, we suggest that observers are most sensitive to the initial frames of the stimulus for orientation discrimination, even though integration over the entire movie would be a better strategy. In addition to being defined in the first 50 ms after stimulus onset, it appears that the tuning of the decision template is invariant over the course of the stimulus presentation. This observation is consistent with a recent finding that the orientation tuning for a population of cells does not change over time,[§] suggesting that orientation dynamics may be a characteristic of single cells that is lost in the output of a network of cells.

Temporal and Spatial Resolution of the Technique. A difficulty in relating our results to the tuning properties of single cells lies in the fact that electrophysiological data are recorded with very high temporal precision. We did find it necessary to perform some smoothing across frames of our pooled noise samples to consistently fit the resulting CIs and so derive the data presented in Fig. 2. However, even looking at raw (nonsmoothed) single frames, we did not find large changes in the components underlying the CI. We were able to measure sequences of frames that were statistically significantly correlated to the mean CI (Fig. 3), which suggests that individual frames did not undergo large changes in any of their parameters across single frames. A single frame from our stimuli lasted 16.6 ms, less than the duration of some of the reported changes in orientation lasting 20–50 ms (1, 4). This result is also consistent with an initial analysis performed on nontemporally smoothed frames that also showed no change in the parameters for the considerably fewer frames that could be fit. As mentioned previously, the spatial and temporal smoothing was performed to increase the number of significant frames entering our analysis. Regarding spatial resolution, it is unlikely that the smaller scale changes ($<5°$) in peak orientation that have been reported were too fine to measure (because we were able to reliably measure changes in orientation from one frame to another on the scale of 1–2°). Larger scale change, as well as the inversions in peak orientation preferences (lasting between 20–50 ms for shifts in the peak or orientation inversions to occur) would have been very clearly resolved.

Thus, our failure to observe changes in our CI data suggests that the reported changes in the tuning in single cells might get lost in the activity of a larger population (consistent with population analysis[§]) and has no measurable repercussions for human observers.

Methods

Apparatus and Stimuli. An Apple Macintosh G4 computer running MATLAB (MathWorks, Natick, MA) was used for stimulus generation, experiment control, and recording subjects' responses. The programs controlling the experiment incorporated elements of the PSYCHTOOLBOX (43). Stimuli were displayed on a 22-inch Electron Blue monitor (1,024 × 768 pixel; frame refresh rate 60 Hz; La Cie, Hillboro, OR) driven by the computer's built-in graphics card. We achieved pseudo-12-bit contrast resolution in grayscale by attenuating and combining the RGB outputs from the graphics card and then amplifying and copying the resulting signal to all three guns of the monitor. The display was calibrated by using a photometer and linearized by using look-up tables in software.

Stimuli. The stimuli were 32-frame (16.6 ms per frame), 64 × 64-pixel dynamic white noise movies containing a centrally located Gabor patch (envelope SDs = 6 pixels; sinewave wavelength = 12 pixels; at the viewing distance of 57 cm, 1 pixel subtended 2 arcmins) that was tilted clockwise or counterclockwise of vertical by 11°. We tested two target SFs (2.5 and 7.0 c/deg) in separate runs by varying the viewing distance. The envelope SDs produced an orientation bandwidth of 16° (SD of the best-fitting wrapped Gaussian function) and a SF bandwidth of 0.9 octaves (SD of best-fitting Gaussian, fit on log axes). The noise was drawn from a normal distribution and was fixed at 30% rms contrast, which has been shown to optimize CI generation (25).

Procedure. We used a single-interval two-alternative forced choice orientation identification procedure. Observers fixated a small white square (2 × 2 pixels), which appeared for 49.8 ms and then was extinguished during presentation of the dynamic stimulus. The observers' task was to indicate with a key-press whether the target Gabor was tilted clockwise or anticlockwise of vertical (90°). Auditory feedback followed a response error. The contrast of the target was varied by using a staircase (44) procedure that reduced the contrast by 1/3 dB or increased it by 1dB after a correct or incorrect response respectively. This procedure converges on the stimulus contrast (threshold) eliciting 75% correct discrimination. Observers completed 10,000 trials per condition, in blocks of 1,000 trials. Different conditions were tested in random order, and on a given trial the orientation tested was randomly selected.

Observers. The three authors and one naïve subject served as observers. All wore optical correction as necessary and performed practice trials until their discrimination thresholds stabilized.

CIs. On any given trial, subjects could make one of two possible responses [clockwise (C) or anticlockwise (A) of vertical] for the two possible target configurations [stimulus clockwise (SC) or anticlockwise (SA) of vertical]. This method yields four stimulus–response combinations (denoted C-SC, C-SA, A-SC, and A-SA). Noise images were summed according to whether they elicited a correct or incorrect response. C-SC and A-SC were mirror-reversed (around vertical) before summing; thus, resulting CIs are defined relative to an −11° (i.e., anticlockwise) target. The difference image between the correct and incorrect response noise images gives the "correct response" CI. This process is illustrated schematically in Fig. 1*a*, where the target is oriented anticlockwise throughout the 32 frames (only 3 frames shown for clarity). In our procedure, correct and incorrect noise images were weighted equally. Depending on the observer's response, the 32 noise frames were stored in either C-SA or A-SA categories. If the stimulus had been presented clockwise, the noise sequence would have been stored in either the C-SC or A-SC categories. This process was repeated for the 10,000 trials, and the CI then was calculated by using all of the noise sequences following the above procedure. At the end of 10,000 trials, two 32-frame CIs were obtained: one containing the noise images associated with a correct response (Fig. 1*a Left*) and one containing the noise sequences leading to incorrect responses (Fig. 1*a Center*). Because the noise was dynamic, we were able to analyze individual frames of the CIs (Fig. 1*a Right*) to examine the discrimination template at any given interval or average groups of the individual frames to examine the mean discrimination template for a given epoch. In the example presented in Fig. 1*a*, the target was

[§]Benucci, A., Frazor, R. & Carandini, M. (2005) *J. Vision* **5,** abstr. 84.

presented for 32 frames, and the 32 frames of the CI have been summed to generate a mean CI. For consistency, all CIs were normalized relative to the "anticlockwise oriented correct" CI (i.e., clockwise correct CIs are the same but flipped about the vertical axis of symmetry). The pixel intensities of all CIs were normalized to span the range −1 to +1 for template fitting.

Given the large parameter space, and the limited amount of data, constraints were imposed on the SF, orientation, and two SD parameters of the fitted templates. For SF, the fits were constrained to be within half an octave on either side of the SF of the stimulus; for orientation the fits were allowed to vary from 90 to 135°; the envelope SDs were limited to be within 2 and 16 pixels (i.e., from 1/3 to 2.66 times the actual values associated with the target).

Parameter Fitting Procedure. Fig. 1*b* shows the best fitting templates (using the constrained nonlinear function minimization procedure "fmincon" in MATLAB) to the average CI across all frames, for the four observers, in two different SF conditions. The top row shows four mean CIs that have undergone two-dimensional spatial blurring (Gaussian SD = 1.5 pixels); below each is their corresponding best-fitting component template (termed Anti-CW template for an anticlockwise-oriented stimulus). The following row is the flipped version of the best fitting template (termed CW template), and the bottom row is the decision template formed by taking a weighted difference between the Anti-CW and CW templates (in the ratio given by the value *r* in the lowest row). A ratio of 1 would be a pure Gabor, and a ratio of 0.5 ("ideal observer"; see *Supporting Text* and Fig. 4) would be an equal contribution of the two component templates and would have a checkerboard appearance (see observer MST's decision template CI for an approximation of this statement).

This fitting procedure was applied to the mean and individual frames of the CIs. Because we sought to examine the changes of the template with exposure duration, we initially fit individual frames of the CI sequence but found that few (≈1/3) frames produced a significant fit, given the large number of free parameters on the fit. To increase the number of frames for analysis, we temporally Gaussian-weighted the CI sequence before summing over all frames (SD of 1.5 frames = 24.9 ms). A Gaussian weighting function was centered over a given frame in the CI sequence and then multiplied with the entire sequence. This method assigns weights to all of the frames before their addition, although in effect the influence from frames beyond ±2 frames of the one being calculated is minimal given the function's SD. The center of the Gaussian is shifted by one frame, and the procedure is repeated to yield 32 weighted-average frames. Under this procedure the first and last frames of the sequence have less contribution from neighboring frames because of the asymmetrical nature of the temporal Gaussian at the extremities. We consider limitations of this procedure in *Results* and *Discussion* sections.

We used a bootstrapping procedure to estimate confidence intervals on the fit parameters so derived. First, for a given subject/condition combination, we computed the gray-level variance across trials of every pixel in the CI. We then took the best-fitting decision template as our estimate of the true local mean of the CI and added random variability equal to the estimated gray-level variance to generate a series of noisy templates. Next, we added this value to a noise mask in a proportion similar to the ratio of signal to noise for the at-threshold performance of the subject in the psychophysical experiment. Finally, we fit the resulting noisy template image using the same procedure as described above. By running this procedure repeatedly with 512 independent noise samples, we made a series of pseudo-CIs that then were fit with the six-parameter functions as before. Estimates of the six parameters were stored, and 95% confidence intervals on each parameter were calculated as ±1.96 SDs of the distribution of the bootstrap fits. Frames were rejected from analysis if any one of the four parameters (SF, orientation, and the two envelope SDs) derived using the least-squares fitting procedure fell outside of the 95% confidence intervals of the distributions of corresponding bootstrapped parameters. In addition, a small number of degenerate frames were rejected if the least-squares parameter estimates fell within 2.5% (on either side) of the limits over which the above parameters were allowed to vary.

We thank Joshua Solomon for advice and Miguel Eckstein and Albert Ahumada for helpful reviews. This work was supported by Biotechnology and Biological Sciences Research Council Grant 31/S17766 (to S.C.D. and P.J.B.).

1. Ringach, D. L., Hawken, M. J. & Shapley, R. (1997) *Nature* **387,** 281–284.
2. Bredfeldt, C. E. & Ringach, D. L. (2002) *J. Neurosci.* **22,** 1876–1984.
3. Mazer, J. A., Vinje, W. E., McDermott, J., Schiller, P. H. & Gallant, J. L. (2002) *Proc. Natl. Acad. Sci. USA.* **99,** 1645–1650.
4. Ringach, D. L., Hawken, M. J. & Shapley, R. (2003) *J. Neurophysiol.* **90,** 342–352.
5. Menz, M. D. & Freeman, R. D. (2004) *J. Neurophysiol. 91,* **1782,** 1793.
6. Volgushev, M., Vidyasagar, T. R. & Pei, X. (1995) *Visual Neurosci.* **12,** 621–628.
7. Gillespie, D. C., Lampl, I., Anderson, J. S. & Ferster, D. (2001) *Nat. Neurosci.* **4,** 1014–1019.
8. Celebrini, S., Thorpe, S., Trotter, Y. & Imbert, M. (1993) *Visual Neurosci.* **10,** 811–825.
9. Hood, D. (1973) *Vision Res.* **13,** 759–766.
10. Westheimer, G. (1991) *J. Opt. Soc. Am. A* **52,** 1040–1045.
11. Burr, D. C. & Morgan, M. J. (1997) *Proc. R. Soc. London* **264,** 431–436.
12. Harwerth, R. S., Fredenburg, P. M. & Smith, E. L., III (2003) *Vision Res.* **43**, 505–517.
13. Burr, D. C. & Santoro, L. (2001) *Vision Res.* **43,** 1891–1899.
14. Arend, L. E., Jr., & Lange, R. V. (1977) *Vision Res.* **19,** 195–199.
15. Emerson, R. C., Bergen, J. R. & Adelson, E. H. (1992) *Vision Res.* **32,** 203–218.
16. DeAngelis, G. C., Ohzawa, I. & Freeman, R. D. (1993) *J. Neurophysiol.* **69,** 1118–1135.
17. Ohzawa, I., DeAngelis, G. C. & Freeman, R. D. (1997) *J. Neurophysiol.* **77,** 2879–2909.
18. Sharpee, T. & Rust, N. C. (2004) *Neural Comp.* **16,** 223–250.
19. Simoncelli, E. P., Paninski, L., Pillow, J. & Schwartz, O. (2004) in *The New Cognitive Neurosciences*, ed. Gazzaniga, M. (MIT Press, Cambridge, MA), pp. 1–20.
20. Beard, B. L. & Ahumada, A. J. (1998) *Proc. SPIE Int. Soc. Opt. Eng.* **3299,** 79–85.
21. Neri, P., Parker, A. J. & Blakemore, C. (1999) *Nature* **401,** 695–698.
22. Abbey, C. K. & Eckstein, M. P. (2002) *J. Vision* **2,** 66–78.
23. Ahumada, A. J. (2002) *J. Vision* **2,** 121–131.
24. Neri, P. & Heeger, D. J. (2002) *Nat. Neurosci.* **8,** 812–816.
25. Solomon, J. A. (2002) *J. Vision* **2,** 105–120.
26. Murray, R. F., Bennet, P. J. & Sekuler, A. B. (2002) *J. Vision* **2,** 79–104.
27. Levi, D. M. & Klein, S.A. (2002) *J. Vision* **2,** 46–65.
28. Dakin, S. C. & Bex, P. J. (2003) *Proc. R. Soc. London* **270,** 2341–2348.
29. Gold, J. M., Murray, R. F., Bennett, P. J. & Sekuler, A. B. (2000) *Curr. Biol.* **10,** 663–666.
30. Legge, G. E. (1978) *Vision Res.* **18,** 69–81.
31. Xing, D., Shapley, R. M., Hawken, M. J. & Ringach, D. L. (2005) *J. Neurophysiol.* **94,** 799–812.
32. Patterson, R. D. (1976)). *J. Acoust. Soc. Am.* **59,** 640–654.
33. Itti, L., Koch, C. & Braun, J. (2000) *J. Opt. Soc. Am.* **17,** 1899–1917.
34. Georgeson, M. A. (1987) *Vision Res.* **27,** 765–780.
35. Watt, R. J. (1987) *J. Opt. Soc. Am. A* **4,** 2006–2021.
36. Tolhurst, D. L. (1975) *Vision Res.* **15,** 1143–1149.
37. Mihaylova, M., Stomonyakov, V. & Vassilev, A. (1999) *Vision Res.* **39,** 699–705.
38. Vassilev, A., Mihaylova, M. & Bonnet, C. (2002) *Vision Res.* **42,** 851–864.
39. Waugh, S. J. & Levi, D. M. (2000) *Vision Res.* **40,** 163–171.
40. McSorley, E. & Findlay, J. M. (2002) *Perception* **31,** 955–967.
41. Lu, Z.-L., Jeon, S.-T. & Dosher, B. A. (2004) *Vision Res.* **44,** 1333–1350.
42. Foley, J. M. & Chen, C. C. (1999) *Vision Res.* **39,** 3855–3872.
43. Brainard, D. H. (1997) *Spatial Vis.* **10,** 433–436.
44. Wetherill, G. B. & Levitt, H. (1965) *Math. Stat. Psychol.* **18,** 1–10.