Insensitivity of V1 Complex Cell Responses to Small Shifts in the Retinal Image of Complex Patterns

TROELS W. KJAER, TIMOTHY J. GAWNE, JOHN A. HERTZ, AND BARRY J. RICHMOND

1Laboratory of Neuropsychology, National Institute of Mental Health, Bethesda, Maryland 20892-4415; and 2Nordita, DK-2100 Copenhagen, Denmark

Kjaer, Troels W., Timothy J. Gawne, John A. Hertz, and Barry J. Richmond. Insensitivity of V1 complex cell responses to small shifts in the retinal image of complex patterns. J. Neurophysiol. 78: 3187–3197, 1997. An important role for neurons in the early visual system is to convey information about the structure of visual stimuli. However, neuronal responses show substantial variation across presentations of the same stimulus. In awake monkeys, it has been assumed that a great deal of this variation is related to the scatter in eye position (inducing scatter in the retinal position of the stimulus). Here we investigate the implied consequence of this assumption, i.e., that the scatter variation in eye position degrades the decodability of the neural response. We recorded from 50 complex cells in primary visual cortex of fixating monkeys while different complex stimuli were presented. Three types of retinal shifts were considered: natural scatter in the fixation, systematic fixation point shift, and systematic stimulus position shift. The stimulus pattern accounts for >50% of the response variance, always six times that accounted for by the scatter in eye position during fixation. The retinal location of a stimulus had to be shifted by 10–12 min of arc, an amount almost two times larger than the smallest picture element, before the responses changed systematically. Nonetheless, changes of the stimulus at the single pixel level often gave rise to discriminable responses. Thus complex cells convey information about the spatial structure of a stimulus, independent of rigid stimulus displacements on the order of the receptive field size or smaller.

INTRODUCTION

A great deal of the response structure of neurons in primary visual cortex is related to the structure of the visual stimulus that appears on the receptive field (Hubel and Wiesel 1962, 1968; Richmond and Optican 1990; Richmond et al. 1990). However, the spike trains are apparently very noisy in that a substantial proportion of the response variation is not obviously stimulus related. In awake monkeys, it has been assumed that a substantial proportion of the unexplained response variance in single neuronal responses is a consequence of the monkey’s fixation errors (Richmond et al. 1990). The consequence of this assumption would be that the fixation-related response variance significantly interferes with using the neuronal responses to discriminate visual patterns. We study this issue here.

Hubel and Wiesel (1962, 1968) found that for a complex cell, an optimally oriented bar could be placed anywhere in a region having a size on the order of the receptive field and still elicit a maximal response. They suggested that complex cells were generalizing for the optimal stimulus across locations. Motter and Poggio (1990) found that the edge of the receptive fields of striate cortical neurons, as mapped with experiments with neurons that had peripherally located receptive fields, intended fixation rather than the actual gaze. Both of these studies suggest that neuronal responses are insensitive to small changes in eye position. On the other hand, according to Gur and Snodderly (1987), the response strength was smaller and the response variance was larger under normal fixation conditions than when the retinal stimulus position was stabilized by external feedback of the eye position. More recently Gur and coworkers (Gur et al. 1997) have shown that the variation in response to a moving stimulus is smaller during periods when the eye is not moving at all, showing neither drift nor microsaccades.

None of these studies explicitly asked how the fixation errors would affect the stimulus discriminability. Here we study the relation between the eye position and the neuronal response directly, employing as stimuli a subset of the two-dimensional Walsh patterns exhibiting complex two-dimensional structure. These stimuli were chosen so that different subunits within the complex cell receptive field, possibly sensitive to different stimulus features, would be stimulated as the stimulus shifted on the retina with changes in eye position. They were presented so that they covered the entire receptive field. In this way, we minimized any effect of stimulus edge on the response. In the first part of the experiment, eye position shifts are obtained simply from the natural fixation errors the monkey makes. In the second part, larger shifts are induced systematically by shifting either the intended fixation point or the stimulus.

Our results extend the findings of Hubel and Wiesel (1968) that the response to optimally oriented bars was unaffected by small shifts of the bar position in the visual field to our more complex stimuli. We find that the responses do not change in any detectable manner as the stimulus changes position across a substantial proportion, 10–12 min of arc in size, of the excitatory receptive field. For such shifts, the responses can be decoded by subsequent processing stages, without significant loss of information, as if there were no fixation error at all.

METHODS

Rhesus monkeys were taught to sit in a primate chair facing either a video or tangent screen on which visual stimuli appeared. Then they were trained to fixate on a spot that subtended 6 ft on the video screen and 22 ft on the tangent screen. The video screen subtended 17° × 13° of the visual field. The image on the tangent screen was created by projecting the monitor image so that it covered 59° × 45°. This was done so that we could carry out experiments with neurons that had peripherally located receptive fields.
The monkeys initially were trained on a simple visual color-discrimination task. A trial was initiated when the monkey touched a bar. A white fixation spot then appeared. After a random interval lasting 1,200–2,200 ms, the spot became red for 500 ms. If the monkey released the bar while the spot was red, a drop of juice was delivered. The monkey could initiate the next trial after 1,000 ms regardless of the trial outcome. During this training, a second stimulus was turned on and off while the fixation spot was white. This stimulus, which later became the receptive field stimulus, had no behavioral significance. Within a few days of its introduction, this stimulus could be placed next to the fixation point without disturbing the monkey’s fixation. After the monkey reached a performance level of 70% correct in this bar-release task, the task was modified so that the monkey only was required to fixate on the spot without any other behavioral requirements. During the experiments, the stimulus was cycled on for 333 ms and off for 750 ms. Rewards were delivered pseudorandomly during the part of the cycle when the stimulus was absent.

A scleral eye coil, a chronic recording chamber, and a post to stabilize the head were implanted under general anesthesia in a single, sterile surgical procedure carried out in a fully equipped operating suite with veterinary assistance. The monkeys recovered without complication. They were given acetaminophen either alone or with codeine as needed for pain relief during the immediate postoperative period. They were allowed 4 wk to recover before Single neuronal recording experiments were started.

**Eye monitor calibration**

The eye position was monitored using the magnetic search coil method (Judge et al. 1980; Robinson 1963). The eye position was calibrated by systematically shifting the fixation point on a rectangular grid around the center of the screen while the monkey performed the bar-release task described above. The voltages from the horizontal and vertical channels were recorded for several trials of the bar-release task at each position. The average voltage for the series was taken to represent that position. The eye position was calculated by assuming that relation between the voltage and position was linear. During the experiments in which the fixation point was systematically shifted, the voltages were compared with the interpolated voltages and were found to be consistent for the few-degree range over which the fixation point was moved. These findings show that the eye coil was accurate and stable during these experiments, and the variations we saw represented the actual deviations of the eyes from trial to trial. The sensitivity of the eye coil system was such that one unit on the A/D converter was 1.7 min of arc quantizing the eye position to ±0.85 min of arc (Fig. 2). The monkeys were rewarded for keeping their gaze within a rectangular window of ±0.5°. The eye position was sampled every 8 ms.

**Single neuronal recording**

Single neurons were recorded extracellularly from primary visual cortex using platinum-iridium electrodes that penetrated the dura directly while the monkeys fixated. Action potentials were recorded to 1-ms resolution. When parafocal neurons were recorded, the electrode was advanced slowly into the cortex, and the first isolated cell was recorded. After most of the recording sessions, the electrode was advanced until the hissing of layer IVc was heard (Poggio et al. 1977). In all cases, the electrode had to be advanced ≥100 μm before this hissing was heard. These procedures ensured that the parafocal neurons we recorded were supragranular (layer II, III).

The neurons with peripherally located receptive fields were in the calcineum bank of V1, and so they were approached from layer 6. We do not know in which layers these complex cells were. Signals from single neurons were isolated based on the principal components of the neuronal waveform (Abeles and Goldstein 1977; Gawne and Richmond 1993). The spike times were recorded to the nearest millisecond.

**Experimental procedure**

The receptive field was located using white or black bars. The orientations and positions that excited the neuron most strongly were found for both black and white bars. A neuron was characterized as complex if it responded strongly to both the appearance and the disappearance of the optimal stimulus (Richmond et al. 1990).

Then we collected as many responses as possible while the neuron was well isolated using a stimulus set of 16 black-and-white, two-dimensional patterns. Patterned stimuli were chosen because our goal was to measure how well the responses could be used to discriminate among stimuli that can be unambiguously visually discriminated. The stimuli were chosen from the Walsh set, a set that elicits a wide range of responses conveying stimulus-related information in V1 (Richmond et al. 1990). The subset used here spanned a wide range of spatial resolution (Fig. 1.1). This stimulus set was chosen so that subunits of the type described by Hubel and Wiesel (1962) with sensitivities of different spatial resolutions and different spatial locations would be stimulated differentially as the stimulus was shifted on the retina. Because the patterns in this stimulus set contain sharp edges, changes on a hypothetical receptive field subunit can occur abruptly as the position of a stimulus pattern element changes with shifts in eye position.

Our stimuli were centered on the excitatory region. The 16 stimuli had resolution 16 × 16 (Fig. 1A) or their vertically oriented one-dimensional counterparts (Fig. 1B). These patterns were presented in both contrasts, i.e., black and white were interchanged. The stimulus size always was adjusted so that the Walsh stimuli just covered the entire excitatory region (Fig. 1C). Given the elongated excitatory regions, this meant that the stimuli covered a part of the immediate surround. The finest picture elements in the Walsh patterns (in the following referred to as pixels) were 7.8 ± 2.1 (SD) ft the parafoveal neurons and 22.6 ± 3.9 min of arc for the neurons with peripheral receptive fields. On the video screen, the white pixels were 39.2 cd/m² and the black were 0.10 cd/m². For the projected image, white was 13.8 cd/m² and black was 0.37 cd/m². The background was set to the mean luminance of the black and white pixels.

Three different sets of experiments were performed. In the first set, we studied the shift in retinal position of the stimulus that occurred during natural fixations when neither the fixation point nor the stimulus moved. In the second set, the position of the fixation point was jumped between two locations on a trial-to-trial basis while the stimulus remained still. In the third, the stimulus was moved between two locations while the fixation point remained still.

All of the experiments complied with Public Health Service policy on the care and use of laboratory animals and were approved by the National Institute of Mental Health Animal Care and Use Committee.

**Data analysis**

The position of the eye 50 ms after stimulus onset was used in these analyses. This was arbitrary, but it was the time at which the analyses of neuronal responses began. However, as shown in Fig. 2, the eye did not move much during the stimulus presentation; the largest deviations were typically two steps in the A/D converter range, i.e., 3.4 min of arc.

The analysis of the spike trains also started 50 ms after the onset of the stimulus. Data from the full 320-ms period of the stimulus presentations were analyzed.

The neuronal responses were quantified in two ways: the number
of spikes was counted and the responses were decomposed into their principal components. Specifically, each spike train was convolved with a fixed Gaussian kernel having $\sigma = 5.0 \text{ ms}$, an optimal width for neuronal data collected under similar conditions (Heller et al. 1995). The resulting continuous estimate of the firing probability was resampled at 5-ms intervals to obtain a 64-dimensional vector for each spike train. The principal components were extracted from the covariance matrix of these response vectors, and each response was decomposed into a weighted sum of principal components. These weights were used to quantify the responses (Richmond and Optican 1987). By definition, any principal component represents more of the variance in the response than any subsequent one. The first few principal components usually provide very effective data compression (Ahmed and Rao 1975).

We performed two kinds of analysis. First, to search for possible systematic dependence of the neuronal responses on eye position, we performed linear regression of the spike count on the eye position, stimulus by stimulus. Second, we quantified the sensitivity of the response to both stimulus and fixation position using information theory.

Information theory provides a quantitative measure of the discriminability of two or more conditions based on signals that are related to them. The discriminability of the conditions is measured by the mutual (or transmitted) information, given by

$$I(C; R) = \sum_c P(c) \log_2 \frac{P(c|\text{r})}{P(c)}$$

where $C$ is a set of conditions (here stimuli or fixation positions) $c$, and $R$ is the set of signals (here the neuronal responses) $r$. The
FIG. 3. Responses of 1 neuron to 8 stimuli, ordered according to horizontal component \( x \) of fixation shift. This neuron had a typical sensitivity of the response to fixation shift. Optimal orientation of a stationary bar was \(-45^\circ\) (pointing toward 4:30 and 10:30 on a clock). A: rasters show responses where \( x \) is restricted to values within \( \pm 20 \) min of arc of the center of gaze. Graphs (bottom) show the spike counts plotted against fixation shift. Linear regression lines for this figure and Fig. 4 are shown when they are significant at the \( P < 0.01 \) level (none in this figure). B: as in A, but with \( x \) restricted to \( \pm 5 \) min of arc of the center of gaze.

brackets indicate an average over the response distribution. \( P(e) \) is the probability of a condition, given no knowledge about the response, and \( P(e|r) \) is the conditional probability of conditions given the response \( r \). These probabilities were estimated using a neural network model specifically designed for this purpose (Kjaer et al. 1994). We calculated \( I(C;R) \) for two different combinations of condition sets \( C \) and response sets \( R \).

In one class of calculations, we categorized the data into groups (labeled by an index \( f \)) according to fixation position. This divided the response set \( R \) into subsets \( R_f \). We then computed the information \( I(S;R_f) \) transmitted by these responses about the stimulus set \( S \). The dependence of this quantity on fixation position shows how discriminable the stimuli are when the fixation point is displaced by a given amount.

In a second class of calculations, we instead held the stimulus \( s \) constant and computed the information \( I(F;R_s) \) transmitted about the set \( F \) of fixation positions \( f \) by the set of responses \( R_s \) evoked by that stimulus. This quantifies the sensitivity of the responses to fixation position.

RESULTS

Using two monkeys, we recorded from 37 complex cells with receptive fields between 3.6 and 8.4° from the fixation location (median = 4.9°) from the fixation point. In one monkey, we recorded from 13 other complex cells with receptive fields between 28 and 37° from the fixation location (median = 34°).

Influence of natural fixation variation on neuronal responses

We first examined the relation between the eye position and the response for each stimulus for each neuron. When the fixations that were within \( \pm 20 \) min of arc of the center of gaze (the average over trials of the fixation position), the responses to a few stimulus patterns changed as the fixation position changed in either the horizontal or vertical direction. When the fixation window dropped below \( \sim 10 \) min of arc on a side, this dependence on fixation position became so small that it could not be seen by simple inspection. Figure 3A shows data from a typical cell, where the systematic dependence of the response on fixation position was weak. For this cell, a significant linear dependence on the horizontal component of fixation was found for only one pattern. The regression lines are shown for the cases where significance reached the \( P < 0.01 \) level. For this neuron, no such dependence on the vertical component of fixation was found for any pattern. Figure 4A shows data from the cell for which these response changes were strongest. Even in this case, however, it is apparent that the variation of the response...
with stimulus is stronger than that with eye position, a point we will return to later.

To quantify the relation between eye position and neural responses for each stimulus, we performed a linear regression of the eye position onto the response for each pattern for each neuron. Specifically, for each neuron and for each stimulus pattern s, the spike count r was fit to the form $r_s = a_s x + c_s$, or $r_s = a_s x + b_s y + c_s$, where $x$ and $y$ were the horizontal and vertical components of the eye position. This was done for the 28 neurons for which there were enough data to perform both this regression and that involving both eye position and stimulus pattern described below (the latter requires many more data). The results of the regression of eye position onto response, pattern by pattern, are summarized in two ways.

First, only a small percentage of the responses had a significant ($P < 0.01$) relation between eye position and response variance across the stimulus set. When the fixations were within ±20 min of arc of the center of gaze (as in Figs. 3A and 4A), 22% (125/560) of the regressions for the horizontal and 10% (58/560) for the vertical positions reached this level. When the responses were limited to fixations within ±6 min of arc of the center of gaze, the numbers reaching significance fell to 5 and 3%, respectively, and when the responses were further restricted to fixations within ±5 min of arc (as in Figs. 3B and 4B), they fell further to 3 and 2%. Thus when the eye position changes within a region of size $10 \times 10$ min of arc, there is only rarely a significant influence of the eye position on the response.

Second, we looked at the amount of variance that is related to the eye position for each pattern. On average across all responses to a pattern, the amount of variance ($r^2$) that can be accounted for is 7.3% for the horizontal position, 4.7% for vertical position, and 10% for the regression model using both the horizontal and vertical position (Fig. 5A). Thus on average, only a small fraction of the variance in the responses to any one stimulus is related to eye position.

For each neuron, we also identified the particular stimulus with the largest response variance related to the eye position. We found (see, e.g., Fig. 5B) that for many neurons the response to these particular patterns could exhibit substantial variation with eye position.

It could be that these exceptional cases occurred for a neuron’s strong responses (those to near-optimal stimuli) and that no significant effect was seen for most stimuli simply because the responses to them were so weak. Alternatively, one might imagine that, due to some threshold effect, the weak responses were the especially position-sensitive ones. To test for such effects, we plotted, for each pattern, the amount of variance explained by eye position against the mean spike count it evoked (Fig. 6). We found no significant correlation between these quantities. Thus although a few stimulus patterns do elicit strongly position-dependent neuronal responses, they do not seem to be related in any obvi-
ous way to the strength of those responses. Neither could we identify any feature of the spatial structure of these stimuli that appeared to be correlated with position-dependent responses.

Influence of natural fixation variation on stimulus discriminability

The central question of interest to us was: how much do these influences related to eye position affect the discriminability of the stimuli? The data in Figs. 3 and 4 illustrate the principal qualitative feature of our findings: even in the (relatively few) cases where the responses exhibited significant fixation position dependence, their dependence on the stimulus pattern was much stronger (e.g., Fig. 4).

The analyses above did not simultaneously account for the effects of eye position and stimulus pattern. These relations among the responses, eye position, and stimulus pattern, were quantified in two ways. The first analysis was a linear one using analysis of covariance (ANCOVA). The second was a more general technique using information theory.

For the same 28 neurons as used for the analysis above, we carried out an ANCOVA \( r_s = a_s x + b_s y + c_s \), with \( a_s \) and \( b_s \) both either included in the fit, or set to 0. In Fig. 7 we see that the amount of variance accounted for by the stimulus pattern alone (bottom line) is consistently 82–85% of that accounted for by the pattern and eye position (top line). Thus these neurons show far greater sensitivity to differences between the stimuli than to the differences in eye position that arise from the natural scatter across repeated fixations. However, we also see that the amount of variance explained drops as the fixation window size increases, raising the possibility that there is a nonlinear effect related to the increasing size of the fixation errors. It is to account for this possibility that we turned to an information theoretic analysis for the rest of the study.

Influence of systematic fixation spot shifts

In the next part of the experiment, larger eye position shifts were imposed systematically to determine what size shifts are necessary to evoke significant changes in the neuronal responses. For 34 neurons, we moved the fixation point horizontally between two locations, each used for half of the stimulus presentations. For 12 neurons, the fixation spot

---

**FIG. 5.** A: histogram of the average variance explained by eye position on a pattern-by-pattern basis for the 28 neurons analyzed. B: histogram of the maximum variance explained for any pattern for the 28 neurons analyzed.
position was randomized for each stimulus presentation to
one of the two locations, and for the other 22 neurons, the
fixation spot was shifted to the other location every 100
stimulus presentations. The monkey’s gaze shifted systematically with changes in the location of the fixation point (Fig.
8). In this way, we could compare the responses when the
fixations belonged to distributions with different, well-defined mean locations. The separations between these means varied from 4 to 72 min of arc for the cells with parafoveal receptive fields and from 27 to 111 min of arc for those with peripheral receptive fields.

The discriminability of patterns was quantified by
$I(S;R_f)$, the information carried in the responses about the
patterns, and by $I(F;R_f)$, the information carried about the fixation position. The information about the patterns was considerable (parafoveal: 0.62 bits, range 0.07±1.56; peripheral: 0.36 bits, range 0.10–0.65) for stimuli presented at the unshifted position.

The dependence of these quantities on fixation shift size
for the population of parafoveal cells is indicated in Fig.
9, where the calculated values for all cells of the stimulusaveraged $I(F;R_f)$ (Fig. 9A) and of the relative $I(S;R_f)$ (Fig.
9B) are collapsed onto single graphs. There is no significant
information conveyed about the difference in fixation location
when the difference is ≤12 min of arc. Above this threshold, as the separation between the fixation points increases, $(I(F;R_f))$, grows slowly and the $I(S;R_f)$ for shifted relative to nonshifted fixation points falls off, showing that the responses change systematically, albeit weakly, with increasing shift and that these changes reduce the information the cells carry about the stimuli.

For the neurons with peripheral receptive fields, there was no significant information about fixation position for the range of shifts that we studied even though the shifts were sometimes large fractions of the receptive field size.

**Influence of systematic stimulus shifts**

Retinal shifts can be due to displacement of either the stimulus or the eye. In principle, this difference could have an influence on how the neurons treat the shift. Therefore, we held the fixation spot constant and systematically moved the stimulus in one half of the presentations while recording
from 16 neurons. The information conveyed about patterns
presented at the unshifted position was virtually the same [0.64 bits (0.08–1.76)] as that found for the neurons recorded in the paradigm where the fixation point was shifted.

The dependence of the stimulus-averaged information
$(I(F;R_f))$, about stimulus shift and of the information
$I(S;R_f)$ carried about the stimuli for this population is shown as a function of shift size in Fig. 10. There is no apparent difference between these results and those obtained (Fig. 9) with systematic fixation point shifts. There is no information about the stimulus position in the responses when the shift size is <10–12 min of arc. The shift dependence of the information $I(S;R_f)$ (Fig. 10B) conveyed about the stimulus patterns is also qualitatively the same as that found when the fixation point was shifted.

A possible reason for lack of correlation of neuronal re-
sponse and eye position could be that the dominant input to the neuron came from the eye whose position we did not record. Therefore we recorded from three neurons while the eye without the eye coil was covered. The shift imposed on the retinal stimulus location was 7–42 min of arc. The values found for $(I(F;R_f))$ and $I(S;R_f)$ were not significantly different from those found with binocular viewing. Thus although the data are very limited, we do not find any evidence suggesting that variations in disparity could account for the response variance of these neurons. These data are included with the rest in the plots we show.

Finally, we recorded from three neurons while we used a
stimulus set that contained only vertical lines (Fig. 1B). Again we could not detect any change in the results.

**Both types of systematic shifts**

For eight cells (6 with stimulus shifts, 2 with fixation point shifts), recordings were made at more than one shift size. For these cells, we could confirm that the dependence...
lus size was fixed by the experimental paradigm to be of the order of the receptive field size.

Figure 12C shows the ratio of the information about fixation point to that about the stimuli, again plotted as a function of shift for the entire population of parafoveal cells. The small values of the resulting numbers underscore the finding reported above for the (generally smaller) natural fixation shifts; the responses of these neurons are generally much more sensitive to stimulus pattern changes than they are to fixation shifts.

The responses of the cells with peripheral receptive fields were even less sensitive to retinal shifts than those of parafoveal ones. This conclusion holds both with respect to absolute shift size (min of arc) and relative shift size (stimulus pixels). We found no effect of retinal shifts in these data, where the shift was up to four Walsh pixels.

Receptive field size

Our experimental design controlled for relative size effects because the stimulus size always had the same relation to the receptive field, i.e., the size as always set so that the excitatory region was just covered completely (see METHODS). Although it is possible that the sensitivity to position could be related to the absolute rather than the relative receptive field size, we have not been able to identify such a relation. For example, the neurons giving rise to the data in Figs. 3 and 4 had receptive fields of similar size, 61\textsuperscript{1}\textsuperscript{68} min of arc and 67\textsuperscript{1}\textsuperscript{67} min of arc respectively, yet the...
neuron of Fig. 4 was the neuron with the greatest sensitivity to position, whereas the neuron of Fig. 3 was more typical, having little sensitivity.

**Discussion**

Even though the gaze of fixating animals changes by small amounts from instant to instant, the basic properties of primary visual cortical neurons in awake, fixating monkeys are similar to those found in anesthetized ones (Wurtz 1969). However, the responses (in both the anesthetized and awake conditions) show a great deal of variation from one stimulus presentation to the next (Tolhurst et al. 1981). Although there is no evidence that the response variation is greater in awake than in anesthetized animals, it has nevertheless been assumed that a large proportion of the response variation in awake monkeys is due to the variability of fixation. Recent data show that some of the response variance for moving stimuli swept through the receptive field during prolonged fixation is smaller when the responses are only taken from periods in which no drift or microsaccade occurs (Gur et al. 1997).

Because a primary role for visual system neurons is to convey information about the structure of visual stimuli, it was our goal here to study how much the natural scatter in fixation position would affect the discriminability of the stimulus. This entailed measuring how much of the response variance was related to this scatter and assessing how the fixation-related variance would affect the potential decodability of the stimulus given the neuronal response. The findings here show that the variability due to changing the stimulus pattern far outweighs the variability due to the differences in fixation position across natural fixation scatter. The change in eye position had to exceed \([1/5]°\) before these changes in eye position were significantly encoded in the responses or before these eye position changes began to degrade the discriminability of the patterns from the responses. When stimulus position was shifted purposely using either gaze or stimulus displacements, there was no detectable effect on the responses of these neurons when the shift was <10–12 min of arc, even though the neurons were sensitive to changes of features half this size within the stimulus. Thus most of the response variance seen across a set of stimulus presentations is not due to these small shifts in the absolute retinal location of a stimulus.

Quantitative measures of this effect are provided by the ratio of the amount of variance in the responses explained by stimulus pattern to the extra amount explained by eye position, and the related information-theoretic quantity \(I(S;R_e)/I(F;R_s)\). These ratios are typically of the order of 6. Thus nearly all of the systematic variability across the responses is due to the stimulus, not the eye position.

Shift sensitivity can result either from changes in the stimulus pattern within the receptive field or from moving the stimulus in or out of the receptive field. Our experiment, in which the entire receptive field always is stimulated, addresses only the first of these possibilities. Our finding is consistent with that of Hubel and Wiesel (1962, 1968), who
discovered that an optimally oriented bar could be moved about within the receptive field without affecting the response. In both experiments, the total stimulation of the receptive field was held fixed while the light intensity falling on different parts of it was changed, and in both experiments, the response was found to be insensitive to these changes.

Our results extend theirs in an important way. They studied the shift invariance of the response to a single kind of stimulus: an optimally oriented bar. We studied the responses induced by a large stimulus set, a variety of two-dimensional patterns with internal structure on a broad range of spatial scales that covered both the center and near-surround, and found shift invariance over the entire stimulus set.

Moreover, despite this insensitivity to shifts of the entire stimulus pattern, the response is sensitive to spatial structure within a pattern on scales much smaller than the receptive field size. The normal visual acuity at 5° eccentricity in rhesus monkeys is ~20 cycles/°, as measured with sine wave gratings (Merigan and Katz 1990). Our stimulus patterns are based on two-dimensional square waves. Two pixels of these patterns correspond to one sine wave cycle. Consequently we would have expected sensitivity to spatial shifts of 1.5 min of arc, yet we found sensitivity only above 10 min of arc. However, the changes in the internal structure of the stimulus that gave rise to discriminable response differences were smaller than 10 ft (~7 min of arc), leading us to believe that our parafoveal neurons approach (within a factor of 3) the spatial resolution of visual acuity measurements.

Turning to the neurons with peripheral receptive fields (28–37° eccentric), it might be argued that the observed shift insensitivity in these cells could be due to their larger receptive field sizes, i.e., the relative importance of the absolute shift size is less for neurons with large receptive fields (in fact ~3 times larger than those of the parafoveal neurons). In fact for peripheral cells, the natural variation in fixation was small compared with the shift, which is no surprise. However, for the larger, systematically induced shifts, we found that even when the difference in receptive field size is taken into account approximately (by measuring the shift in terms of the pixel size in the patterns, which were set in the experimental procedure to match the size of the receptive field), the sensitivity to shifts was smaller in the periphery than it was close for neurons with receptive fields close to the fovea.

Gur and Snodderly (1987) investigated the shift sensitivity of responses by trying to stabilize the retinal image using feedback of the eye position to the display system. The neuronal responses elicited appeared to be stronger and less variable than those obtained without stabilization. Our results seem inconsistent with theirs. Unfortunately, they did not report what types of cells they recorded from, nor any statistics of neuronal responses, so their results cannot be compared directly with ours. Perhaps they recorded from simple cells, and these neurons are sensitive to shifts to which the complex cells we recorded are insensitive.

It is also possible that their result is related to the stabilization process itself. When external feedback of eye position is used to stabilize an image on the retina, the stabilization is bound to be delayed relative to the eye movement. In Gur and Snodderly’s experiment, the eyetracker must record a change of eye position (5 ms), followed by some electronic delay for constructing the corrected stimulus (we guess 1–2 ms), and finally the average delay for the screen refresh (8.3 ms). Consequently, the retinal image will remain uncorrected for ~15 ms, and a small eye movement would lead to two images appearing on the screen in sequence, one before correction and the other after, 15 ms later. Furthermore, the update cycle of the eyetracker is short relative to the 15-ms feedback lag; this could cause hysteresis. As a result, the image might repeatedly make small shifts on the retina that would appear as motion and lead to increased neuronal discharge. By failing to account for the extra area covered by the adjustments in stimulus position, such movements also would make the receptive field appear smaller.

Therefore, an important control would be to show whether the discriminability of the responses improved with feedback or whether the response to all stimuli was enhanced due to small movements.

Alternatively, our results can be reconciled with theirs if one hypothesizes a drastic increase in retinal position sensitivity, independent of receptive field size, when stimuli are presented very close to the fovea. Our cells were 3.6–8.4° eccentric, whereas theirs were only 1–2° eccentric.

For about half of our cells, a few stimuli (different from cell to cell) did evoke shift-sensitive responses. However, we were unable to relate this occasional sensitivity systematically to response strength, spatial frequency content of the stimulus pattern, absolute receptive field size, or to stimulus element resolution relative to receptive field size.

Other studies have addressed the second possible source of shift sensitivity—shifts of the stimulus in or out of the receptive field. One such experiment was performed by Motter and Poggio (1990), who monitored the onset of the response as the stimulus was moved into the receptive field. They found that receptive field edge, as identified by this onset, did not appear to shift in the way one would naively expect with small spontaneous changes in eye position. They concluded that the receptive field (in particular, its edge) was being moved according to the monkey’s intended fixation position. More recently Gur and Snodderly (1997) performed a similar experiment but reported the opposite conclusion—the response latency was strongly correlated with the eye position, indicating that the receptive field edge was fixed at a specific retinal location.

If correct, the Motter-Poggio result would force a significant revision of the current standard model of complex cells. In that picture, originally proposed by Hubel and Wiesel [and later formalized mathematically by Fukushima (Fukushima et al. 1983) in his neocognitron model], the complex cell output is approximately a logical OR function of the outputs of a set of subunits centered at different locations within the receptive field. The retinal locations of all of these subunits are fixed, so the model predicts unambiguously that the onset of the response in the Motter-Poggio experiment should follow retinal stimulus shifts rigidly. Their claimed results would require circuitry absent from the standard model, perhaps along the lines of the models suggested by Anderson and coworkers (Anderson and Van Essen 1987; Olshausen et al. 1993) to move the entire receptive field (or at least the subunits near the receptive field edge) in response to top-down signals about the monkey’s expectations or intentions.

Our results are not inconsistent with this possibility, but...
they do not require it, because in our experimental design, the stimulus covered the entire receptive field. As our patterns shifted on the retina, changes occurred in the input to all subunits, not just those at the receptive field edge, and any effect that could be attributed to a compensatory shift in the receptive field edge could be explained equally well within the terms of the standard model.

Hubel and Wiesel (1962, 1968) hypothesized that complex cells are concerned predominantly with stimulus orientation. Simple cells detected oriented features at different locations, and complex cells, driven by convergent input from arrays of simple cells with the same orientation tuning but different receptive field centers, signal the presence of oriented features, irrespective of their exact location within the region spanned by the simple cell receptive fields. Our findings, obtained with more complex stimuli, suggest that more complex kinds of features also are processed in this shift-invariant fashion in primary visual cortex.

Detailed characterizations of complex cells often include a description in terms of receptive fields subregions characterized by number, type (on, off, on-off) and degree of interaction (Spitzer and Hochstein 1985; Tolhurst et al. 1981). Our stimulus set was chosen so that sensitivity to shifts across these subunits would be maximized. However, our measurements, designed to detect the influence of retinal stimulus position on stimulus discriminability, do not permit us to infer the nature, e.g., size or location, of these subregions. Experiments with still larger and more diverse stimulus sets will be necessary before a full understanding of complex cell function is achieved.

We thank Dr. Mortimer Mishkin for long-standing support and for very helpful comments on this manuscript.

Present address of T. W. Kjaer: Nordita, Blegdamsvej 17, DK-2100 Copenhagen, Denmark.

Address for reprint requests: B. J. Richmond, Laboratory of Neurophysiology, 49 Convent Dr., Bethesda, MD 20892-4415.

Received 3 January 1997; accepted in final form 9 July 1997.

REFERENCES


