Chapter 5: Separable codes for attention and luminance contrast in the primary visual cortex

Arezoo Pooresmaeili, Jasper Poort, Alexander Thiele, Pieter R. Roelfsema
Summary

Our visual system encodes the basic features of visual stimuli as well as their behavioural relevance. Previous studies suggested that the representation of luminance contrast, one of the most basic low-level features, is confounded with the effects of attention, as these factors jointly determine the firing rate of neurons in the visual cortex. Here we therefore investigate how reliable the luminance contrast of a visual stimulus and the locus of attention are coded in the primary visual cortex (area V1) on a single trial. To this aim, we simultaneously recorded from multiple groups of V1 neurons, and decoded stimulus contrast as well as attention using a linear decoding scheme (support-vector machine). The success of this approach demonstrates that visual attention and stimulus contrast are represented in the visual cortex by largely separable codes.

Introduction

The visual image that we perceive is initially registered by neurons at early processing stages which encode low-level visual features such as color, motion, depth and luminance contrast of local visual elements (Hubel, 1988; Marr, 1982). At the same time, neurons in lower-level areas of the visual cortex contribute to the integration of features into coherent object representations (Roelfsema, 2006). Neurons in the primary visual cortex, for example, enhance their responses if they code the image elements of a larger figure relative to when they code the elements of the background (Lamme, 1995; Lee et al., 1998). This response enhancement is necessary for the perception of the figure, as perception fails without the response enhancement (Supèr et al., 2001). Similarly, if the task demands perceptual grouping of a number of edges into an elongated curve, then the neurons in the primary visual cortex that code these to-be-grouped edges enhance their response (Roelfsema, 2006), while perception errrs if neurons coding the wrong edges enhance their response (Roelfsema and Spekreijse, 2001). At a psychological level of description, attention is directed to the features to be bound into a coherent object representation (Roelfsema, 2006). Thus, neurons in low-level areas of the visual cortex, like area V1, appear to play a dual role: they code low-level image features, and they contribute to the attentive integration of these features into coherent object representations. Here we ask if these two roles are compatible. Firstly, does the modulation of the neuronal responses by visual attention cause distortions in the representation of low-level features? Secondly, can V1 neurons code the locus of attention reliably, in spite of variations in luminance contrast, a low-level feature that influences the activity of all visually responsive neurons?
Previous neurophysiological studies suggested that attention and contrast have comparable effects on neuronal activity in early visual areas such as area V4 (Reynolds et al., 2000) and MT (Martinez-Trujillo and Treue, 2002). These findings inspired a ‘contrast gain model’ of attention suggesting that a shift of attention is equivalent to an increase in stimulus contrast. However, there are a number of potential disadvantages if cortical neurons would strictly adhere to the contrast gain model. Firstly, attention shifts could cause aberrations in the perception of the visual scene, as the apparent contrast of attended visual objects would increase. One psychophysical study (Carrasco et al., 2004) reported precisely such an effect, although other studies suggested that attention has no effect or only a small influence on perceived contrast (Prinzmetal et al., 1997; Schneider, 2006; Tsal et al., 1994). Secondly, if the effect of attention on neuronal activity were equivalent to an increase in stimulus contrast, then high contrast regions of an image should have privileged access to attentional processing, and subjects might even be unable to direct their attention to image regions with a lower contrast. This prediction is neither in accordance with our subjective experience nor with psychophysical data (Einhäuser et al., 2008; Pashler et al., 2004). These results, taken together, suggest that attention and contrast cannot have strictly equivalent effects on neuronal activity. But how does the neuronal code for attention differ from the code for attention? To address this question we will attempt to decode image contrast and locus of attention from the activity of a population of neurons in area V1. We note that a non-separability of the neuronal codes for attention and contrast might have severe consequences as it would be inherited by most, if not all, higher visual areas.

We depart from the observation that not all neurons in area V1 are influenced by attention to the same degree. There are many neurons in area V1 that are hardly influenced by attention (Roberts et al., 2007; Roelfsema et al., 2004; Roelfsema and Spekreijse, 2001; Vidyasagar, 1998) and these cells could code the contrast of a stimulus reliably, irrespective of attention shifts. It might, in addition, be possible to isolate a pure attentional signal from area V1, by comparing the activity of neurons that are and that are not influenced by attention (Fig. 1). Imagine that we record the activity of 2 V1 neurons that have a similar contrast response function when attention is directed elsewhere (blue bars in the lower panels of Fig. 1). Suppose that one of the neurons increases its firing rate if attention is shifted to the receptive field (red bars) while the activity of the second cell is hardly affected by attention shifts. It should then be possible to decode the locus of attention regardless of stimulus contrast by measuring the difference between the two responses. This scheme can be generalized to situations where neurons have different firing rates by computing a weighted difference.
response, and also to larger populations of neurons, as long as there are cells with different magnitudes of the attentional effect. Specifically, we will here use a support vector machine algorithm that is also widely applied in human fMRI studies (Haynes et al., 2007; Haynes and Rees, 2005; Kamitani and Tong, 2005; Kamitani and Tong, 2006; Li et al., 2007; Mourão-Miranda et al., 2007) to determine the optimal linear combination of responses to decode stimulus contrast and the locus of attention (Fig. 1).

**Figure 1. Decoding luminance contrast and attention from neuronal responses in area V1.**

We recorded the responses of neurons at several multi-unit recording in area V1 \( (r_1, r_2, \ldots, r_n) \) in a task that involves shifts of attention and variations in stimulus contrast. We investigate whether we can decode attention and contrast on the basis of two types of responses: (1) strongly modulating responses (SMR) that distinguish between attended and non-attended stimuli and are also influenced by luminance contrast, and (2) weakly modulating responses (WMR) that are mainly sensitive to stimulus contrast and weakly influenced by attention. We will use support vector machines (SVMs) to decode attention and contrast and expect that contrast decoding might be possible by focusing on the weakly modulating responses, while a relatively pure attention signal might be extracted by evaluating the difference between strongly and weakly modulating responses.

To test this decoding scheme we trained two monkeys on a curve tracing task that is associated with attention shifts (Roelfsema et al., 1998) while we varied the contrast of the stimulus. We recorded the spiking activity of V1 neurons with chronically implanted multi-electrode arrays that enabled us to monitor the neuronal activity of many neurons at the same time. We decoded attention and contrast by combining (1) the activity of neurons with a different strength of attentional modulation, and (2) different time windows of the neuronal response. Our results demonstrate that it is indeed possible to decode both the locus of
attention and the luminance contrast from neuronal activity of area V1 neurons, on a single trial.

Results

Figure 2A illustrates the curve-tracing task used to induce shifts of visual attention. The task of the monkeys was to locate a circle at the end of a target curve that was connected to the fixation point while ignoring the other curve that served as distractor. On any trial, the two curves always had the same contrast (selected from a set of eight contrast levels; i.e. 1.7%, 2.6%, 4.3%, 6%, 6.8%, 8.5%, 10.9% and 19.3% Michelson contrast) and the only difference between them was a small contour segment that connected the target curve to the fixation point. Figure 2B shows the performance of the two monkeys as a function of stimulus contrast. Accuracy was high for the higher contrasts and came close to chance levels at the lowermost contrasts. At these low contrast levels, the two curves were difficult to discriminate from the grey background and the monkeys may have directed their attention to the wrong curve on a fraction of trials. Here we will therefore focus on the data from the stimuli for which the monkeys’ performance was better than 85% correct (luminance contrast of 4.3% and higher).
Effects of attention and luminance contrast on the V1 population response

We recorded multi-unit spiking activity (Supèr and Roelfsema, 2005) at 56 recording sites in area V1 of the two monkeys (17 sites in monkey G, 39 in monkey A, Fig. 2C) in the same sessions in which we obtained the behavioral data described above. The use of chronically implanted electrode arrays permitted simultaneous recording of spikes of all the neurons. The receptive fields of the V1 always fell either on the target or on the distracter curve, and we ensured that the contour element that connected the fixation point to one of the curves was located outside the neurons’ receptive fields so that differences between the conditions could be attributed to attention shifts (Roelfsema et al., 1998).

Figure 2. The curve-tracing task and the behavioral performance of the monkeys.

(A) Curve-tracing task. While the monkeys directed their gaze to a fixation point (FP), we presented two curves with circles at their end. One of the curves was connected to the fixation point (target curve, T) while the other one was not (distracter, D). The receptive field (RF) of the V1 neurons under study either fell on the target or on the distracter curve. The monkeys had to make an eye movement to the circle at the end of the target curve after a delay of 500 ms.

(B) Behavioral performance of the two monkeys as function of the luminance contrast of the two curves.

(C) Receptive fields of the recording sites in the two monkeys relative to the two curves.
Figure 3A shows population responses evoked by the target and distracter curve at various levels of luminance contrast. When a segment of one of the curves appeared in the receptive field the neurons first exhibited a transient response, which was followed by a more sustained response caused by the continuous receptive field stimulation during the 500-ms fixation delay. Attention did not have a strong effect on the transient response, while the sustained response evoked by the target curve was stronger than the response evoked by the distracter curve, as described previously (Roelfsema et al., 1998). To examine the encoding of stimulus contrast, we averaged responses across the target and the distracter curve at the various contrast levels (Fig. 3B). The strength of the population responses increased monotonically with the luminance contrast in both windows, although the responses evoked by nearby contrasts differed more during the peak response than in the phase of sustained activity.

**Figure 3** V1 population responses in the curve tracing task.

(A) Neuronal responses evoked by the target (red) and distracter (blue) curve at six contrast levels, averaged across all the recording sites ($N=56$). For our analysis, we measured the neuronal response magnitude in two time windows: an early window that contained the response transient (peak-window, 35-135 ms, grey bar) and a later window when the neuronal activity reached a sustained level (200-500 ms, orange bar). (B) Neuronal responses evoked by curves of different contrasts, averaged across the two attention conditions. (C, D) Attention $d'$ (abscissa) and contrast $d'$ (ordinate) across recording sites in the early (C) and late (D) response windows. The dashed lines show the mean $d'$.
Attention and contrast tuning at individual recording sites

We next explored how well neuronal responses discriminated between attention conditions and luminance contrasts, in two time windows. The first time-window was centered on the transient response (35-135 ms after stimulus onset, grey bar in Fig. 3A) and the second time window on the sustained response phase (200-500 ms, orange bar). We note that these time windows were non-overlapping and had different lengths. We computed attention and contrast $d$-primes ($d'$) in these windows to quantify how well individual recording sites discriminated between attention and contrast conditions on single trials (see Methods). Figure 3C shows the distribution of attention and contrast $d'$ across the population of V1 recording sites (N=56) in the early time window. As expected, the transient neuronal responses discriminated well between contrasts (mean $d'$ = 1.65; comparison between contrasts of 4.3 and 19.3%) but poorly between attention conditions (mean $d'$ = 0.07). In the late time window, the neurons discriminated less well between contrasts (mean $d'$ = 1.29) but the discrimination between target and distractor curve improved (mean $d'$ = 0.41, Fig. 3D). There also were a number of recording sites with attention $d'$ close to zero in the late window, indicating that these neurons were hardly influenced by the shifts of attention. We considered the possibility that the absence of an attentional effect was caused by a weaker response or a lower signal-to-noise ratio (SNR) of the MUA recordings at some of the recording sites. In order to examine this possibility we computed the SNR as the ratio of the peak response to the standard deviation of the spontaneous activity, and compared the SNR between recording sites with the highest and the lowest attention $d$'s (lowest and highest 33% of sites, N=19). The difference between the SNR values in these two groups of recording sites was not significant ($P > 0.1$, U-test), which indicates that the absence of attentional effects at weakly modulating recording sites was not caused by a small SNR. Moreover, the heterogeneity in the effects of attention were reproducible, because recording sites without an effect of attention on one day also tended to lack the attentional effect on other days. The attention $d$'s were significantly correlated across sites between recordings sessions on subsequent days (Monkey A, $\rho = 0.77$; Monkey G, $\rho = 0.88$; both $Ps < 10^{-5}$) and even between sessions more than 6 months apart (Monkey A, $\rho = 0.70$; Monkey G, $\rho = 0.75$; both $Ps < 0.005$).

We aimed to exploit the heterogeneity in the degree of contrast and attention tuning across neurons and time windows to decode attention and contrast at the same time, using two methods. Firstly, we combined the neuronal responses from the two time windows. Figure 4A illustrates this approach for an example recording site. At this recording site, the neurons had weak attentional modulation in the early window ($d'$=0.08), and were strongly modulated
by stimulus contrast (with a $d'$ of 2.0, Fig. 4C). The same neurons had stronger attentional modulation in the late window ($d'=1.6$), while the contrast $d'$ decreased ($d'=1.1$, Fig. 4D). Thus, attention and luminance contrast could be decoded in different time windows for this example recording site.

A second possibility to decode attention and contrast separately is to combine late window activity of neurons with strong and weak attentional modulation. Figure 4B shows the neuronal responses at a recording site with weak attentional modulation (blue cross in Fig. 3C, D). (C, D) Magnitude of responses at recording site 1 in the early (C) and late (D) window evoked by the target (red bars) and distractor curve (blue bars). (E) Neuronal activity evoked at recording site 2 in the sustained response window.

Figure 4 Variation in the effects of attention and contrast across two example recording sites that were recorded at the same time.

(A) Responses of neurons at recording site 1 (blue plus in Fig. 3C, D) exhibited strong attentional modulation. Responses evoked by the target curve are shown in red and responses evoked by the distracter in blue. (B) Neuronal responses at a recording site with weak attentional modulation (blue cross in Fig. 3C, D). (C, D) Magnitude of responses at recording site 1 in the early (C) and late (D) window evoked by the target (red bars) and distractor curve (blue bars). (E) Neuronal activity evoked at recording site 2 in the sustained response window.

A second possibility to decode attention and contrast separately is to combine late window activity of neurons with strong and weak attentional modulation. Figure 4B shows the neuronal responses at a second recording site that was recorded at the same time as the recording site of Fig. 4A and was only weakly modulated by attention (attention $d'=0.14$). The neurons at this second recording site discriminated relatively well between contrasts in the late time window (contrast $d'=1.44$) (Fig. 4E). We conclude that relatively pure attention and contrast signals also exist in the sustained response phase of different neurons. The examples of Fig. 4 were chosen to illustrate the idea of using different time windows and
different neurons, and we will now generalize these ideas to simultaneous recordings from a larger population of recording sites.

**Simultaneous decoding of attention and contrast on single trials**

We wished to determine the linear combination of responses across a larger number of recording sites that would best decode attention and contrast. To this aim, we constructed an “attention decoder” and a “contrast decoder” support vector machine (SVM, see Methods). As input to the decoders, we provided the single-trial activity of all V1 recording sites with receptive fields at nearby locations on the same curve (17 in monkey G and 39 in monkey A, Fig. 2C), in the two time windows. The attention decoder had to classify whether the receptive fields fell on the target or the distracter curve irrespective of stimulus contrast, while the contrast decoder had to classify two contrast levels, 4.3% and 19.3%, irrespective of the locus of attention.

The attention support vector machine was defined by two weights, \(a_{\text{early}}\) and \(a_{\text{late}}\), for every recording site. The neuronal activity in the two windows, \(r_{\text{early}}\) and \(r_{\text{late}}\), is multiplied by the weights and then summed across windows and recording sites, and if this sum exceeds a threshold, \(\theta_{\text{Attention}}\), then the decoder indicates that the curve falling in the RF is attended, while it indicates that the curve is not attended otherwise (Fig. 1):

\[
\begin{align*}
\sum_i w_{\text{early}} \cdot r_{\text{early},i} + \sum_i w_{\text{late}} \cdot r_{\text{late},i} & > \theta_{\text{Attention}} \quad \Rightarrow \text{Attended} \\
\sum_i w_{\text{early}} \cdot r_{\text{early},i} + \sum_i w_{\text{late}} \cdot r_{\text{late},i} & < \theta_{\text{Attention}} \quad \Rightarrow \text{Not attended}
\end{align*}
\]

Figure 5A (left) shows the performance of the attention decoder in the two animals. In monkey A, the mean output of the decoder was stronger when the curve in the receptive fields was attended than when it was not, for every level of contrast. In monkey G the attention decoder also permitted classification with a single threshold, except for the responses evoked by the lowermost contrast (4.3%), where the output of the attended condition was just below the \(\theta_{\text{Attention}}\). We do not know whether this difference between the results in the two animals was caused by the smaller number of recorded neurons in monkey G or by his poorer performance at this low luminance contrast (Fig. 2). We next investigated how much information was present in the population of neurons about the locus of attention on a single trial by applying the leave-one-trial-out cross validation method (see Methods). The leave-one-trial-out accuracy for the attention decoder was 81% in Monkey A and 72% in Monkey...
showing that it is indeed possible to isolate the effects of attention by a linear combination of the activity of a population of neurons in area V1.

When we examined the weights that the SVM had assigned to the two windows, we found that the weights $w_{i,\text{early}}$ of the early window were negative, on average, while the weights $w_{i,\text{late}}$ in the late window were positive (Supplementary Fig. 1), in accordance with our hypothesis that a relatively pure attention signal can be obtained by subtracting neuronal responses that are weakly modulated by attention from responses with a stronger modulation (Fig. 1). Moreover, we observed a strong positive correlation between $w_{i,\text{late}}$ and the neurons’ attention $d’$ in the late window ($\rho=0.92$, $P<10^{-10}$), as well as a negative correlation between $w_{i,\text{early}}$ and the contrast $d’$ in the early window ($\rho=-0.62$, $P<5.10^{-7}$), indicating that the attention SVM assigned larger positive weights to recording sites that were more sensitive to

Figure 5 Output of the attention and contrast decoders.

(A) The output of the attention decoder is shown in the left column for the two monkeys and the output of the contrast decoder in the right column. Red bars show responses evoked by the attended stimuli and blue bars responses evoked by non-attended stimuli as function of luminance contrast. Error bars denote the standard deviation of the decoder output across trials. The green dashed line depicts the attention decoders classification threshold ($\theta_{\text{Attention}}$) (B) Confusion matrix of classification performance for all pair-wise contrast comparisons. The values above the diagonal are from monkey A and the values below the diagonal are from monkey G.
the locus of attention and larger negative weights to sites that coded luminance contrast reliably.

We defined the contrast decoder equivalently, with weights, $w_{i,\text{early}}^c$ and $w_{i,\text{late}}^c$ for the two response windows of every recording site $i$ and a single threshold, $\theta_{\text{Contrast}}$. The SVM was trained with only two contrasts, 4.3% and 19.3%, but tested on neuronal responses evoked by all contrast levels. Figure 5A shows that the output of the contrast decoder increased monotonically with the contrast of the stimulus in both monkeys. Decoding of contrast worked well in monkey G, although there was a small residual influence of attention on the output of the decoder, and it was excellent in monkey A where the decoder output hardly depended on attention.

Figure 5B illustrates how well the contrast SVM trained with only two contrasts distinguished between any pair of contrasts on a single trial by determining the percentage of correctly classified trials (see Methods). The decoder worked well above chance level for the most difficult discriminations (neighboring contrast levels), while the trial-by-trial discrimination became more accurate for larger contrast differences. An analysis of the weights of the contrast SVM in the two time windows revealed that the weights in the early window were stronger than the weights in the late window. Moreover, there was a strong correlation between a recording site’s contrast $d'$ and the strength of the weight to the contrast SVM in both windows (Supplementary Fig. 1D).

**Decoding of attention and contrast in one time window**

The above analysis assumes that neuronal responses in both time windows can be combined without loss of information. It is of importance to know how well attention and contrast can be decoded if the responses in only one of these windows is available. We therefore determined the accuracy of decoding in each of the two windows separately and compared the results to the decoding accuracy when the information in both windows was used (Fig. 6). In this analysis we chose two neighboring contrast levels for the contrast decoder that were relatively difficult to discriminate (10.9% and 19.3%), and compared decoding performance to a baseline performance that could be achieved when we randomly assigned trials to two arbitrary categories (shuffling; circles in Fig. 6), thereby ensuring that the classification accuracies were related to the attention and contrast conditions, and not to random regularities in the data. As expected, we obtained best decoding results by combining the responses in both windows. In the early window, decoding of contrast was good, but decoding of attention was close to the baseline, because neuronal responses were hardly
influenced by attention during this response phase (Fig. 3A, C). In the late window, however, decoding of attention was quite good (83.1% in monkey A and 65.9% in monkey G), and the contrast decoder could still distinguish the responses evoked by stimuli with nearby contrasts relatively well on single trials (79.4% in monkey A and 69.1% in monkey G). These results indicate that both attention and contrast can be decoded during the delayed response phase by combining responses of neurons that are weakly and strongly influenced by attention.

**Discussion**

Here we have demonstrated that it is possible to decode the locus of attention as well as the luminance contrast of a stimulus by combining information from a population of neurons in area V1. We extracted a relatively pure contrast signal by focusing on the neuronal activity during the first transient response, or on the sustained response phase of neurons that are only weakly influenced by attention shifts. Furthermore, we isolated an attention signal by subtracting neuronal responses that were not (or only weakly) modulated by attention from responses that were modulated more strongly (Fig. 1). We thereby showed that it is possible

---

*Figure 6* Comparison of classification performances of the attention (upper row) and contrast decoder (lower) from neuronal responses in the early, late and both time windows. Dashed lines show the baseline classification level when the class labels were randomly assigned to the responses (shuffling). The error bars denote the 95% confidence intervals.
to decode attention as well as contrast with a reasonable accuracy from a relatively small sample of the neuronal activity in area V1, on a single trial.

**Decoding accuracy**

These data illustrate the virtues of techniques that permit the simultaneous recording of neuronal activity at multiple cortical locations. The responses of cortical neurons on an individual trial are noisy, and the usual approach to suppress this noise is to average neuronal activity across multiple trials. However, neurons in higher cortical areas have to decode attention and contrast by combining the responses of different neurons rather than of different trials of the same cell (Churchland et al., 2007; Deadwyler and Hampson, 1997; Nicolelis et al., 1997). The response magnitude of different V1 neurons exhibit correlated fluctuations and these so-called noise correlations can reduce the amount of information that is carried about the visual stimulus on a single trial (Abbott and Dayan, 1999; Shadlen et al., 1996). Conventional methods that allow recording from one neuron at a time cannot estimate the noise correlation, and therefore do not provide accurate estimates of the amount of information present in the responses of a population of neurons.

We used a multivariate classification method (a support vector machine) that automatically takes the redundant information present in the correlations between neuronal responses into account. Decoding of stimulus information with the use of support vector machines is now used commonly in fMRI research (Haynes et al., 2007; Haynes and Rees, 2005; Kamitani and Tong, 2005; Kamitani and Tong, 2006; Li et al., 2007; Mourão-Miranda et al., 2007). These decoding techniques have not yet been used widely with single and multi-unit data (Stark and Abeles, 2007) and the present study is, to the best of our knowledge, the first to apply it for the decoding of information from multi-unit activity in the visual cortex. Due to the high signal-to-noise ratio and temporal resolution of multi-unit activity compared to the fMRI signal, we were able to decode attention and contrast at the same time, and on a much finer time scale (in time windows shorter than 400 ms) than can be achieved with fMRI.

There are a number of reasons why it is likely that our results provide a lower estimate of the amount of information that is present in a small population of neurons. Firstly, we showed recently that noise-correlations in area V1 are caused by relatively global trial-to-trial fluctuations in the overall level of neuronal activity (Poort and Roelfsema, 2008). These noise correlations reduce the amount of information that can be gained about the locus of attention when combining neurons with receptive fields on the same curve, as was the case in the
present study where all the receptive fields fell on nearby contour elements of the same curve. However, the difference between the response strengths evoked by the target and distracter curve is only little influenced by these global fluctuations in neuronal activity, and it is possible to decode attention with an accuracy of about 90% on a single trial with an average of only four recording sites, provided that some of the receptive fields fall on different curves (Poort and Roelfsema, 2008). It is therefore likely that the accuracy of attention decoding in the present study would have increased considerably had we recorded from neurons with receptive fields on different curves. Secondly, we subdivided the neuronal responses into two large time windows, one for the transient response and the other for the sustained response phase. A more fine-grained analysis of the profile of the neuronal responses across time might have increased the amount of information about attention and contrast. Indeed, a fine-grained analysis in time would be sensitive to the latency of neuronal response, and response latency in area V1 (Gawne et al., 1996) and also in area V4 (Williford and Maunsell, 2006) depends on stimulus contrast. Therefore, adding information about response latency on individual trials might have further increased the accuracy of contrast decoding.

We note, however, that the integration of activity from different time windows requires storage of traces of the neuronal activity in previous time windows. This storage could occur, for example, in areas of the premotor and frontal cortex where neurons code the traces of previously presented sensory stimuli (Hernández et al., 2002; Miller and Cohen, 2001). It is unclear, however, whether the information that the V1 peak responses carry about stimulus contrast is stored in this manner, and we therefore also decoded stimulus contrast on the basis of the sustained response phase only. Accuracy of decoding in the late time window was only slightly worse than if both windows were used because there are also many V1 neurons that are sensitive to stimulus contrast in their later responses phase.

**Implications for models of visual attention**

The efficacy of our decoding algorithm relies on the heterogeneity of the effects of attention and contrast on the activity of neurons across recording sites. Neurons at some of the recording sites were mainly tuned to attention during their sustained response while neurons at other recording sites were only tuned to contrast and hardly influenced by shifts of attention (Fig. 4). This heterogeneity is not unexpected and has been observed in many, if not all, previous studies on the representation of attended vs. unattended stimuli (Martinez-Trujillo and Treue, 2002; Reynolds et al., 1999; Roberts et al., 2007; Roelfsema et al., 1998;
Roelfsema et al., 2004; Spitzer et al., 1988; Treue and Maunsell, 1996) as well as in studies on the coding of luminance contrast by neurons in the visual cortex (Albrecht and Hamilton, 1982; Sclar et al., 1990).

Our ability to decode both attention and contrast implies that the neuronal code for attention differs from the code for contrast, and does not support a strict interpretation of the so-called ‘contrast gain’ model of attention (Reynolds et al., 2000; Treue, 2004), which states that the effects of attention on the neuronal responses are equivalent to increases in the effective contrast of a stimulus. If the only effect of attention were an apparent increase of luminance contrast, then it would not have been possible to decode both factors from the activity of a population of neurons. Nevertheless, our findings do not rule out the possibility that for some individual neurons the interaction between attention and contrast is well described by a contrast gain model, although all the cells presumably do not conform to this model. A recent study in area V4 (Williford and Maunsell, 2006) revealed a comparable heterogeneity in the interaction between attention and contrast in area V4, as the interaction of contrast and attention was well described by a contrast-gain model for some neurons, while the activity patterns of other neurons were better described by different models.

The existence of distinct representations of attention and stimulus contrast in the visual cortex has a number of advantages. It permits a veridical representation of the luminance contrast of visual stimuli irrespective of attention shifts. Psychophysical studies that measured if and how attention alters contrast perception have obtained conflicting results. One study reported that attention reduces perceived contrast (Tsal et al., 1994), another that it increases perceived contrast (Carrasco et al., 2004), and yet others that there is hardly any influence of attention on contrast perception (Prinzmetal et al., 1997; Schneider, 2006). Our results show that attention need not alter contrast perception because there are also V1 neurons that are mainly tuned to luminance contrast and are not influenced by attention shifts. The second advantage is that shifts of visual attention need not be dominated by high contrast stimuli. If the codes of attention and contrast differ, attention can also be directed to low contrast stimuli, which is in accordance with human psychophysics (Einhäuser et al., 2008; Pashler et al., 2004). In the contour grouping task of the present study, attention is directed to all contour elements of the target curve so that they are grouped in perception (Roelfsema, 2006). The separation of the neuronal codes for attention and contrast would permit the grouping of contour elements with varying contrasts if they belong to the same curve. Neurons with weak attentional modulation can code the contrast of the contour elements veridically, while the
neurons with stronger modulation can label the set of contour elements that belong to the target curve.

**Experimental Procedures**

Two monkeys participated in the present study. In a first operation, a head holder was implanted and a gold ring was inserted under the conjunctiva of one eye for the measurement of eye position. In a separate operation, arrays of 4x5 or 5x5 electrodes (Cyberkinetics Neurotechnology Systems Inc.) were chronically implanted in area V1. The operations were performed under aseptic conditions and general anaesthesia. Details of the surgical procedures and the postoperative care have been described elsewhere (Roelfsema et al., 1998; Roelfsema et al., 2007). All procedures complied with the US National Institutes of Health Guidelines for the Care and Use of Laboratory Animals and were approved by the institutional animal care and use committee of the Royal Netherlands Academy of Arts and Sciences.

**Behavioural task**

The animals performed a curve-tracing task where they had to locate a circular target that was connected to the fixation point by a curve (target curve) and to ignore another curve that was a distracter (Fig. 2A). A trial started as soon as the monkey’s eye position was within a 1°×1° window centered on the fixation point (0.2° diameter). After an interval of 300 ms, the stimulus appeared on the screen consisting of two curves and two circular targets (0.6° diameter). The background display was grey (luminance 16.3 cd/cm²), the circular targets and the fixation point were darker than the background (luminance 9 cd/cm², 28% Michelson contrast), and the curves were brighter than the background. We displayed both curves with the same luminance contrasts; either 1.7%, 2.6%, 4.3%, 5.9%, 6.8%, 8.5%, 10.9% or 19.2% (Michelson contrast). When the stimulus had been in view for 500 ms, the fixation point disappeared and the monkey had to make an eye movement to the circle on the other end of the target curve. All stimulus conditions were randomly interleaved and presented equally often. We recorded at least 40 correct trials for every stimulus in a recording session.

**Recording and data analysis**

Spiking activity was recorded from the chronically implanted multi-electrode arrays (Cyberkinetics Neurotechnology Systems Inc.) with TDT (Tucker Davis Technologies) multi-channel recording equipment. For the detection of multiunit activity (MUA), the signal was amplified, band-pass filtered (300- 9000 Hz) full-wave rectified, low-pass filtered (< 200 Hz)
and sampled at a rate of 760 Hz. The MUA provides an instantaneous measure of the number and the size of action potentials of neurons in the vicinity of the electrode tip. The population response obtained with this method is expected to be identical to the population response obtained by pooling across single units. We recently compared MUA to single unit data in the curve tracing task, and found that MUA indeed provides a reliable estimate of the average single-unit response (Supèr and Roelfsema, 2005). Recordings with a good signal-to-noise ratio were obtained from ~70% of the electrodes. Receptive field dimensions of neurons at these electrodes were measured by determining the onset and offset of the visual response to a slowly moving light bar, for each of eight movement directions (Supèr and Roelfsema, 2005). The median area of the receptive fields was 0.8 deg$^2$ (range 0.12 deg$^2$ to 3.9 deg$^2$). Receptive field eccentricity ranged from 0.9° to 4.4° with an average of 2.5°.

Peri-stimulus time histograms (PSTHs) were calculated in a time window from 300 ms before stimulus onset to 500 ms thereafter, and normalized to the peak response ($P_e$) after subtraction of the spontaneous activity ($S_p$, the average activity in the 300 ms fixation interval prior to stimulus onset). The peak response ($P_e$) was determined as the maximum of response evoked by the two stimuli with the highest contrast, in a time window from 35 ms to 135 ms after the stimulus onset. We included a recording site in our analysis if the peak response ($P_e - S_p$) exceeded the standard deviation of the spontaneous activity, $\sigma_{Sp}$, by a factor of three.

We used $d'$ to quantify how well the neurons at a recording site discriminated between two stimulus conditions. $d'$ was computed as:

$$d' = \frac{\mu_2 - \mu_1}{\sigma},$$  \hspace{1cm} (2) \\
$$\sigma = \frac{\sigma_1 + \sigma_2}{2}$$  \hspace{1cm} (3)

and $\mu_i$ and $\sigma_i$ represent the mean and the standard deviation of the neuronal responses across trials for the two stimuli. To compute attention $d'$, we first calculated the $d'$ for the discrimination between the target and the distracter curve for every contrast, and then averaged these values across contrasts to obtain a single measure of the attention $d'$. Similarly, the contrast $d'$ was computed for the discrimination between the responses to two contrast levels (4.3% and 19.3%), averaged across the two attention conditions.
Decoding analysis

We used the SVM$_{\text{light}}$ implementation (Joachims et al., 1999) of Support Vector Machines (Vapnik, 1995). The input to the SVM classifiers were $Np$-dimensional vectors of neuronal responses, $r = \{r_1, r_2, \ldots, r_N\}$ where $N$ is the number of trials and $p$ is the number of recording sites. The SVM classifier finds the weights that best separate the responses using a linear discriminant function of the form described by equation 1. To evaluate the accuracy of classification, we used the leave-one-trial-out cross validation method. In this method, the neuronal responses of all trials but one were used to derive the SVM, which was then used to classify the probe trial that was kept separate. This process was repeated for all the trials, and the reported accuracies represent the percentage of the trials that were correctly classified. To calculate the confusion matrix of Fig. 5B, we trained the SVM with the data of two contrast levels (4.3% and 19.3%), and classified the single trial responses of any pair of contrasts. For each decoder, we tested whether the classification accuracies were significantly different from a baseline classification accuracy assessed with the leave-one-trial-out cross-validation method, but after randomly assigning two class labels to the data of all the trials (shuffling). We obtained two distributions of the cross-validated classification accuracies for the original and shuffled data by bootstrapping (N=100), i.e. if there were $t$ trials, we randomly selected $t$ trials with replacement and repeated all calculations. We considered the classification accuracies to be significant if the 95% confidence intervals of these two distributions did not overlap.

Reference List


