

Spatio-temporal prediction and inference by V1 neurons

Kun Guo,* Robert G. Robertson, Maribel Pulgarin,^{†,‡} Angel Nevado,^{†,‡} Stefano Panzeri,[§] Alexander Thiele and Malcolm P. Young

Department of Psychology, Henry Wellcome Building for Neuroecology, University of Newcastle, Newcastle upon Tyne NE2 4HH, UK

Keywords: contextual influences, orientation tuning, primary visual cortex, spatio-temporal regularity

Abstract

In normal vision, visual scenes are predictable, as they are both spatially and temporally redundant. Evidence suggests that the visual system may use the spatio-temporal regularities of the external world, available in the retinal signal, to extract information from the visual environment and better reconstruct current and future stimuli. We studied this by recording neuronal responses of primary visual cortex (area V1) in anaesthetized and paralysed macaques during the presentation of dynamic sequences of bars, in which spatio-temporal regularities and local information were independently manipulated. Most V1 neurons were significantly modulated by events prior to and distant from stimulation of their classical receptive fields (CRFs); many were more strongly tuned to prior and distant events than they were to CRFs bars; and several showed tuning to prior information without any CRF stimulation. Hence, V1 neurons do not simply analyse local contours, but impute local features to the visual world, on the basis of prior knowledge of a visual world in which useful information can be distributed widely in space and time.

Introduction

In our dynamic visual environment, objects and scenes often occur and move in statistically predictable ways. The projection of the visual world onto the retina therefore often reflects measurements of a stream of events, which are spatially and temporally coherent. Consequently, the visual system frequently may have ‘reason to believe’ that a particular feature is presented at a particular location, because of the spatial structure of the current scene, the temporal structure of its evolution over time, and prior knowledge of the spatio-temporal structure of the visual world (Knill & Richards, 1996).

Our recent psychophysical results indeed demonstrate that prior knowledge of the probability structure of these spatial and temporal regularities influences visual processing (Guo *et al.*, 2004b). Using stimulus sequences, which comprised four collinear bars (predictors) appearing successively towards the region of the fovea, followed by a target bar with the same or different orientation (see Fig. 1), subjects’ orientation judgements of the foveal target bar were biased toward the orientation of the predictors. The degree of this bias was correlated to the spatio-temporal prior probability induced by the orientated predictors (i.e. stronger bias for the predictors presented in a highly ordered and predictable sequence, with less bias for the predictors

presented in a randomized order or randomized duration). A Bayesian inference model, which associated predictable spatio-temporal stimulus structure with increased prior expectation of colinear events, accounted well for the human visual processing of spatio-temporal regularities at the perceptual level (Guo *et al.*, 2004b). This finding demonstrates that spatio-temporal regularities of the external world are used to reconstruct the visual scene whereby local information is processed in light of the context within which it occurs.

An important question is whether this computation is implemented by neurons at early stages of the visual system, or whether it is mostly present in specialised late-stage processors? The function of the visual system has traditionally been viewed as hierarchical analysis of the immediate retinal (or feedforward) input (Livingston & Hubel, 1987; Felleman & Van Essen, 1991; Lennie, 1998), whereby neurons at each successive stage act effectively as local filters. Mounting evidence from neuroanatomical studies, however, suggests that visual neurons are embedded in an extensive neural network with feed-forward, lateral and feedback connections (Das & Gilbert, 1995; Angelucci *et al.*, 2002; Chisum & Fitzpatrick, 2004) which should enable them to have access to a wide variety of spatially and temporally dispersed signals on which to base their computations (Young, 2000). Such an integration can occur already in the retina, where neuronal activity is consistent with the anticipation of moving stimuli (Berry *et al.*, 1999). Moreover, a variety of studies have demonstrated that neurons at different cortical stages ‘make sense’ of the information that is present inside their classical receptive field (CRF) by integrating information from larger areas outside their CRF (nCRF), whereby they contribute to perceptual pop-out, contour integration, surface perception, figure-ground segregation, and motion integration (Knierim & van Essen, 1992; Zipser *et al.*, 1996; Levitt & Lund, 1997; Sugita, 1999; Kapadia *et al.*, 2000; Super *et al.*, 2003; Guo *et al.*, 2005; Li *et al.*, 2006; Huang *et al.*, 2007; Thiele, 2007).

Correspondence: Professor Alexander Thiele, as above.

E-mail: alex.thiele@ncl.ac.uk

**Present address:* Department of Psychology, University of Lincoln, Lincoln LN6 7TS, UK.

[†]*Present address:* Department of Psychology, University of York, York YO10 5DD, UK.

[‡]*Present address:* York Neuroimaging Centre, University of York, York YO10 5DG, UK.

[§]*Present address:* Faculty of Life Sciences, University of Manchester, Manchester M60 1QD, UK.

Received 27 March 2007, revised 24 May 2007, accepted 18 June 2007

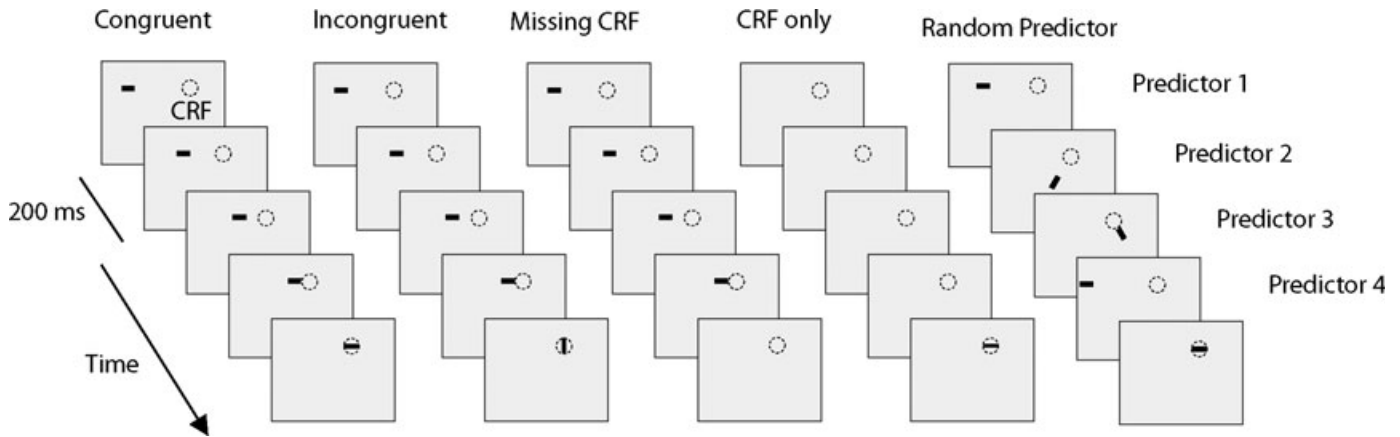


FIG. 1. Stimulus conditions. Five colinear bars ($\pm 60\%$ contrast) were presented in a linearly ordered spatio-temporal sequence. The first four bars (the predictors) were of the same orientation, and the fifth bar was the target. Predictor-target sequences were drawn from the following conditions in a randomized block design. Congruent, orientation of target same as orientation of predictor sequence; incongruent, orientation of target differs from orientation of predictor sequence; missing CRF, predictor sequence only, no target; CRF only, no predictor sequence, target only; Random 'predictors', individual 'predictor' bar presented randomly at each of the 24 predictor positions (an arbitrary example of four of the total 24 possible predictor locations is shown).

We speculated that V1 neurons, by means of nCRF integration, would be informed of spatially and temporally distributed scene properties, and contribute to the computation of spatio-temporal regularity demonstrated in our psychophysical experiment. To examine this, we adopted visual stimuli similar to those employed in our psychophysics experiment and recorded responses of orientation selective neurons in area V1 of two anaesthetized and paralysed macaques. As we expected, when the visual stimulus was presented in a highly predictable spatial and temporal sequence towards a neuron's CRF, a substantial number of V1 neurons responded to the predictable events prior to and distant from the stimulation of their CRFs and transmitted more information about the CRF stimulation, suggesting V1 neurons directly contribute to the coding of spatio-temporal regularities in our visual world.

Materials and methods

Animal preparation

We recorded extracellularly from single units in area V1 of two anaesthetized and paralysed adult rhesus monkeys (*Macaca mulatta*, 5.0 and 8.0 kg). All procedures complied with the 'Principles of laboratory animal care' (NIH publication no. 86-23, revised 1985), UK Home Office regulations on animal experimentation [Animals (Scientific Procedures) Act, 1986], and the European Communities Council Directive 1986 (86/609/EEC). The electrophysiological techniques used for recording have been reported in detail elsewhere (Guo *et al.*, 2004a). Briefly, the animals were initially anaesthetized with ketamine (10 mg/kg). General anaesthesia was achieved using halothane (1-3%) and a mixture of 70% N₂O and 30% O₂, and alfentanil (30 μ g/kg/h). Adequate depth of anaesthesia was verified by repeatedly checking the absence of withdrawal reflexes. During the recording, the animals were paralysed with a continuous infusion of pancuronium bromide (0.1 mg/kg/h) in sterile glucose-saline and artificially ventilated. The rectal temperature (~ 38 °C), expired CO₂ (3.5-4.5%), blood pressure and electrocardiogram were continuously monitored and maintained to ensure appropriate level of anaesthesia. A small craniotomy was positioned over the central visual field representation of area V1.

Pupils were dilated with atropine. The corneas were protected with zero-power contact lenses and artificial pupils (3-mm diameter) were

positioned in front of each eye. The refractive state of each eye was measured with an ophthalmoscope, and additional lenses were used to focus the eyes on the stimulus monitor.

Visual stimuli

Visual stimuli were generated using a VSG 2/3 W graphics system (Cambridge Research Systems, Cambridge, UK) and displayed on a high frequency noninterlaced gamma-corrected colour monitor (110 Hz, Sony GDM-F500T9) with the resolution of 1024 \times 768 pixels. At a viewing distance of 57 cm the monitor subtended a visual angle of 40 \times 30°. The mean luminance of uniform grey background was kept at 6.0 cd/m².

The main visual stimuli comprised five bright (24 cd/m², 60% contrast) or dark (1.5 cd/m², -60% contrast) bars, which were presented successively in a highly predictable spatial and temporal sequence towards the neuron's CRF. The first four bars (the predictors) were of the same orientation and presented outside the neuron's CRF, and the fifth bar was the target presented within the neuron's CRF. The length of each bar was set to the diameter of the neuron's CRF, and the width of the bar was 0.3°. Each bar was presented for 200 ms. There was no spatial and temporal interval between adjacent bars. The bars were flashed in turn in a position immediately adjacent (end-to-end) and in a time immediately preceding the next bar at successive positions. The orientation of the trajectory of the sequence (the predictors) and the target were varied independently and randomly between 0° and 180° in 30° steps. As each predictor moved parallel to its orientation towards the receptive field of the neurons, the differently orientated predictors started in different locations, and were also presented in different spatial locations, surrounding the CRF.

During the recording, the predictor-target sequences were drawn from the following five conditions in a randomized block design (Fig. 1). Congruent - four predictors and one target with the same orientation were presented sequentially for a total of 1000 ms; incongruent - four predictors and one target were presented sequentially for a total of 1000 ms, but the orientation of the target was different from that of the predictor sequence; missing CRF - four predictors were presented sequentially for a total of 800 ms followed by 200 ms blank screen, the target was not presented; CRF only - the

predictor sequence was not presented, the target was presented for 200 ms after 800 ms 'blank' screen; random predictor – individual predictor was presented for 200 ms randomly at one of the 24 possible predictor locations (four predictor spatial locations for each of six predictor sequence orientations). In this condition, a bar was presented at each 'predictor' location at the same orientation as the bars in the congruent sequences, while ensuring that if more than one example of the 'random predictor' condition were drawn in consecutive trials, the presented bars did not accidentally generate part of a predictor sequence. The number of congruent trials was chosen such that it matched the sum of incongruent, missing CRF and CRF only trials in each block. Therefore, a typical block consisted of 42 congruent trials, 30 incongruent trials, six missing CRF trials, six CRF only trials and 24 random predictor trials. A minimum of ten blocks were tested for each analysed neuron (i.e. a minimum of 1080 trials per neuron). In a nutshell, this resulted in ten CRF only, ten missing CRF, 50 incongruent and 70 congruent trials per orientation per neuron.

Recording and data analysis

The activity of V1 neurons was recorded using glass-coated tungsten microelectrodes (1–2 M Ω impedance), and was amplified and sampled through a CED1401 plus digital interface (Cambridge Electronic Design, Cambridge, UK). Spikes were stored with a 0.1-ms interval resolution. Single neuron activity was determined on the basis of the size and shape of the spike waveform, and was confirmed by a spike-sorting program (Spike2, Cambridge Electronic Design, Cambridge, UK) with a template-matching procedure. The approximate laminar position of recorded neurons was determined by the depth of the microelectrode and the characteristic features of layer 4 (such as nonorientation selective, high spontaneous activity and brisk 'on' and 'off' responses). No attempt was made to select neurons from a particular layer of cortex. We intentionally avoided microlesions at the end of individually electrode tracks, as these would have destroyed part of the local network we were specifically interested in. Therefore a detailed reconstruction of layering information was not possible.

Having isolated a neuron and determined the dominant eye (nondominant eye was covered with a black disk during the recording), the CRF was carefully mapped using a sweeping bar and a sinusoidal grating patch moving across the screen with variable length, width and velocity. Within the patch, the optimal grating drifted continuously in the neuron's preferred direction. In this way, we estimated CRF width and breadth profiles (Guo *et al.*, 2004a). To avoid underestimating the size of the CRF, the CRF was further covered with a uniform grey background and an annular window was centred on the CRF. A drifting sinusoidal grating with moderately high contrast (~60%) and the neuron's preferred direction was presented within the window. The outer diameter of the annulus was fixed at 10°, while the inner diameter was adjusted until there was no response in excess of spontaneous activity. The final size of the inner diameter of the annulus was treated as the size of the CRF of the recorded neuron, and corresponds to the annular response field that was used as a conservative CRF size estimate in previous studies (Cavanaugh *et al.*, 2002). A drifting sinusoidal grating with the size of the CRF was then placed at the centre of the CRF. The grating's orientation/direction, spatial and temporal frequency were systematically varied to determine the neuron's preferred tuning characteristics. For all neurons reported in this experiment, their CRF locations were within the central 10° and CRF diameters ranged from 0.6° to 1.5°.

After determining the characteristics of a neuron's CRF, the testing stimuli were presented to the dominant eye for at least ten repetitions

in a randomized order. The interstimulus interval was 1000 ms. The neuron's discharge for the duration of stimulus presentation was averaged and normalized after the recording, and analysed off-line using software developed in Matlab.

Individual experiments lasted 4–5 days, and each neuron needed to be recoded continuously for at least one and half hours to complete the testing. In total we had a full set of data from 34 neurons (six from one macaque and 28 from the other macaque). No obvious difference in neuronal responses was observed between the two animals, over the experimental time or with the fluctuation of anaesthesia level.

To quantify how neurons code the stimulus orientation we calculated the mutual information (Shannon, 1948) which measures the maximum amount of knowledge about the stimulus set – the upper bound of Information – available to an observer who 'reads off', on each trial, the signals carried by the spike train:

$$I = \sum_{\theta, n} P(\theta)P(n|\theta) \log_2 \frac{P(n|\theta)}{P(n)} \quad (Eq.1)$$

where $P(\theta)$ is the probability with which orientation θ is presented, $P(n|\theta)$ is the probability of observing n spikes in a certain time window given orientation θ .

The mutual information can be expressed as a weighted average of the stimulus-specific terms information, which measures the difference between the stimulus response probability $P(n|\theta)$ and the unconditional response probability $P(n)$ and is a measure of which particular stimulus value θ is best discriminable from the observation of neuronal activity (Borst & Theunissen, 1999).

$$i(\theta) = \sum_n P(n|\theta) \log_2 \frac{P(n|\theta)}{P(n)} \quad (Eq.2)$$

We corrected for the systematic error (bias) induced by limited sampling as detailed next. First, we reduced the dimensionality of the response space by limiting the number of possible responses n in the above Eqns (1 and 2) into no more than 12 classes, as follows. If, for a given cell and time window under analysis, the maximum number of spike counts across all trials was 11 or lower, then no response manipulation was needed because the number of possible responses was less or equal to 12. However, for some cells (when considering a 100-ms window) we observed more than 11 spikes in one trial. In such cases we grouped the response spike counts n into 12 approximately equipopulated classes. This grouping provides a better sampling of the response probabilities given that approximately 40 trials per stimulus were available (Panzeri & Treves, 1996; Pola *et al.*, 2003). Second, we applied the analytical sampling corrections of Panzeri & Treves (1996) to the information quantities. This procedure requires the number of trials to be 2–4 times bigger than the number of possible response classes in order to work well and eliminate effectively the sampling bias (Panzeri & Treves, 1996; Pola *et al.*, 2003). Given that 40 trials per stimulus were available for both congruent and incongruent conditions, our information calculation was expected to be reliable in such conditions. However, this information bias correction procedure would not work well in cases (such as the CRF only condition) when only ten trials per stimulus are available. For this reason, we decided not to perform an information analysis of CRF only data and we concentrated on the information about target orientation conveyed during congruent or incongruent stimulation.

Results

Our stimuli were sequences of colinear bars, which occurred in a highly predictable spatial and temporal sequence towards the neuron's CRF (Fig. 1), resulting in an apparent motion sequence. We hypothesized that following sequence onset, the visual system accumulates 'reason to believe' that bars of a specific orientation and location will appear at a specific time in the future. Typically, the fifth bar (target) in a sequence was placed in the neuron's CRF. To test whether neurons in V1 were influenced by the spatio-temporal regularities of the sequence presented outside their CRF, we recorded orientation selective neurons in area V1 of two anaesthetized and paralysed macaques. We employed a factorial design, which studied combinations of orientations for the bar sequences; expected and

incongruent orientations of the target; and omission or presence of the target. If V1 neurons use information from spatio-temporal priors they will (i) be sensitive to the colinearity and timing of the predicting bars; (ii) develop tuned responses in cases where no target is present; and (iii) be sensitive to the mismatch in orientation between the predicting bar sequence and the CRF target.

In the congruent condition, the predictors and the target had the same orientation and were presented in a linearly ordered spatio-temporal sequence (200 ms for each position), orientated towards and through to the CRF. Seventeen out of 34 (50%) neurons responded to predictors presented outside their CRFs (early responses). Figure 2A and B shows an example neuron. When analysing its responses to the CRF target stimulation (time window 800–1000 ms after trial onset) in the CRF only and congruent conditions, this neuron showed clear and

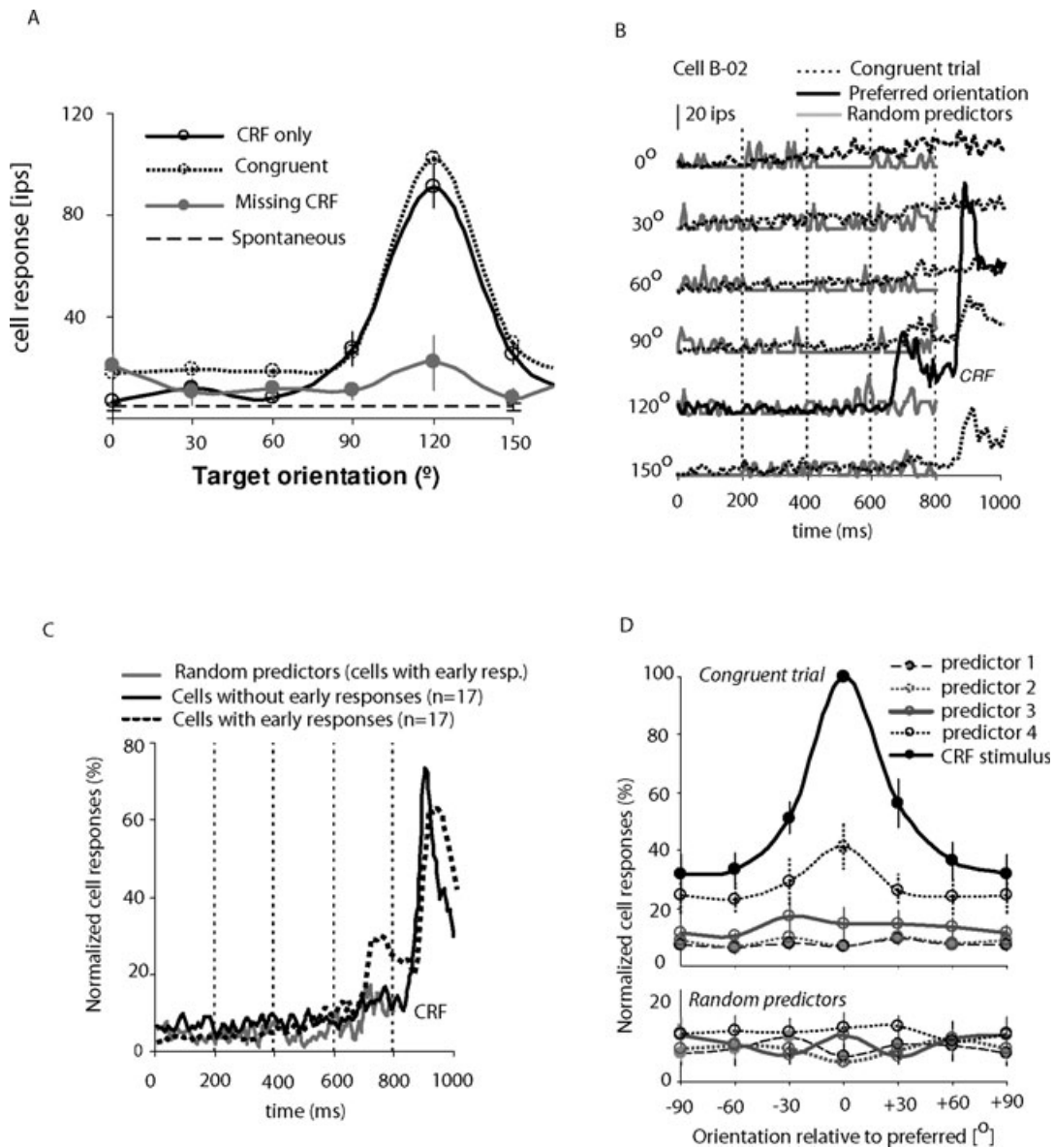


FIG. 2. (A) Orientation tuning responses of cell B-02 to the CRF target bar presented with and without a predictor sequence. (B) Response of cell B-02 to the congruent predictor-target sequence with six different orientations (10-ms bins). The target bar appeared in the CRF during the 800 ms to 1000 ms window. (C) Normalised cell responses of all V1 neurons tested to bars with preferred orientation presented in the congruent condition. The neuron's maximum firing rate to its CRF stimulus was treated as 100% response. (D) The early responses were tuned by the orientation of the predictor sequence (ANOVA, $P < 0.05$), in the same manner as the neuron's response to the orientation of the target bar. Responses to targets presented at the preferred orientation (0°) were treated as 100% response. In C and D, single bars presented randomly at the same positions (random predictors condition) failed to elicit any significant response (*posthoc* test, $P > 0.05$).

same orientation tuning to the target bar presented with and without a predictor sequence (solid and dashed black lines in Fig. 2A; one-way ANOVA, $P < 0.01$). In the random predictor condition, individual predictor bars presented randomly at each of the 24 predictor positions did not elicit neuronal responses in excess of spontaneous activity (one-way ANOVA, $F_{23,230} = 1.41$, $P = 0.09$; grey lines in Fig. 2B, for the purpose of demonstration, the responses to the random predictors were plotted according to the predictors' position in the congruent condition). However, in the congruent condition when the sequence of predictors was presented at this cell's preferred orientation (120°), the cell responded well above its spontaneous firing rate during the appearance of the fourth predictor in a position adjacent to, but entirely outside the CRF (time window 600–800 ms; *posthoc* test Tukey's least significant procedure, $P = 0.0003$; black solid line in Fig. 2B). This early response, approximately 30% of the neuron's maximum firing rate to the preferred CRF target bar, was not due to intrusion by predictor 4 into the otherwise determined CRF [the annular response field (Cavanaugh *et al.*, 2002)], as a bar of the same orientation presented at the same location but in the absence of a predictor sequence (random predictor condition) failed to elicit any change from the spontaneous activity (grey line labelled with 120° in Fig. 2B). We did not observe obvious differences in neuronal responses between the two animals with respect to these early responses; three out of six neurons recorded from one monkey and 14 out of 28 neurons from the other monkey responded to predictors presented outside their CRFs.

Figure 2C shows normalised responses of 34 neurons to the predictor-target sequence presented with each neuron's preferred orientation in the congruent condition. The neuron's maximum firing rate to the CRF stimulus was treated as 100% response. Black solid and dashed lines represent the averaged responses of 17 neurons having and not having significant early responses to the predictors, respectively, grey line represents the averaged responses to the random predictors for those neurons showing early responses. Although early responses could be evoked by several of the predictors, even as far from the CRF as the location of predictor 2, the strongest early response, on average $38\% \pm 8$ (mean \pm SEM) of the neurons' maximum firing rate to the preferred CRF stimulus, was evoked by predictor 4 at the neurons' preferred orientation (Fig. 2C). The early responses were also tuned to the orientation of the predictor sequence (one-way ANOVA, $P < 0.05$), and this tuning reflected the orientation tuning inside the CRF. Population analysis in Fig. 2D shows that the responses evoked by the predictors, especially predictor 4, exhibited the same orientation tuning as the target bar. Most importantly, when the same predictors were presented randomly at these locations, the neuronal responses were indistinguishable from their spontaneous activities (one-way ANOVA, *posthoc* test for individual neurons, $P > 0.05$). We found that in 15 out of 17 neurons (which showed significant early responses), the shape of the orientation tuning curves elicited by predictor 4 and target bar were not significantly different (Kolmogorov–Smirnov test, $P > 0.05$).

In the missing CRF condition, no stimulus appeared in the CRF during the target bar period. However, three neurons still gave robust responses that were significantly above the spontaneous firing rate during the period in which a target would be expected to appear (*posthoc* test for individual neurons, $P < 0.05$). Figure 3A shows such an example. When the target was presented in the absence of a predictor sequence (CRF only condition, black solid line in Fig. 3A), or was presented as part of a sequence (congruent condition, black dashed line in Fig. 3A), the neuron B-16 demonstrated clear orientation tuning, with preferred orientation at 150° (one-way ANOVA, CRF only condition $F_{5,59} = 36.75$, $P = 3.33\text{E-}16$; congru-

ent condition: $F_{5,59} = 41.11$, $P = 5.3\text{E-}18$). During the time window of target presentation in the missing CRF condition (grey line in Fig. 3A), when the orientation of the predictor sequence was 150° , the response of this neuron was 44% of the response measured with a single 150° bar presented within its CRF (one-way ANOVA, $F_{5,59} = 9.02$, $P = 2.8\text{E-}6$). On average this value was $36\% \pm 6$ for the three neurons that showed significant responses in the missing CRF condition. Figure 3B shows the associated peri-stimulus-time-histograms when the predictor-target sequence with the preferred orientation (150°) was presented. Clearly, this neuron showed early responses to the fourth predictor in the congruent and missing CRF conditions (*posthoc* test, $P < 0.01$; red and orange lines in Fig. 3B), but did not show any above spontaneous responses to the same predictor presented in the random predictors condition (*posthoc* test, $P > 0.05$; blue line in Fig. 3B). In the missing CRF condition the neuron's response to the missing CRF stimulus was not due to a sustained early response, as the PSTH in Fig. 3B clearly shows that the response to predictor 4 was phasic, and returned to baseline within ~ 80 ms after response onset. It was also not due to an 'off' response to the fourth predictor, as the offset of an identical predictor bar in the random predictor sequence did not evoke a response (Fig. 3B).

In the incongruent condition, the orientation of the predictors differed from that of the target. If V1 neurons receive no information about the predictors presented outside their CRFs, the response to a target of a given orientation should be constant, no matter what the predictors' orientation. However, the response of some neurons to the same CRF target orientation was modulated strongly according to the orientation of the predictor sequence. Figure 3C demonstrates an example of this effect. This neuron was broadly tuned to the orientation of the target in the CRF only condition with preferred orientation around 120° (one-way ANOVA, $F_{5,59} = 5.8$, $P = 0.0007$). In the incongruent test condition, as the orientation of the predictors approached the preferred orientation of the neuron, the responses to targets at nonpreferred orientations (i.e. 0° , 30° or 60° target orientation) gradually approached the magnitude of the responses to targets at the preferred orientation. Conversely, when the predictor was of the nonpreferred orientation (e.g. 30°), the neuron responded to a CRF stimulus of preferred orientation (120°), as if a nonpreferred orientation had been presented inside the CRF. Thus this neuron was more influenced by the predictor orientation than by the CRF orientation, and this predictor induced tuning was significant (one-way ANOVA, $P < 0.01$). Out of 34 neurons tested, eight showed increased responses to the target when the orientation of the predictors approached the preferred orientation, while the response of two others was decreased.

In some cases we encountered a tuning shift in the preferred orientation of the target induced by the predictors. An example is shown in Fig. 3D. It shows the preferred target orientation for each predictor orientation in the incongruent condition. The tuning of this neuron to the target was significantly shifted towards the orientation of the predictor sequence. This effect was maximal when the orientation of the predictor sequence was within 30° of the preferred orientation of the target determined in the CRF only condition. For ten neurons whose response to a given CRF orientation was manipulated by different predictors' orientation (see Fig. 3C for an example), the maximum tuning shift of the preferred CRF orientation ranged between 4° and 59° with the mean value of $18.95^\circ \pm 6.82$ (Fig. 3E). However, as the bar orientation was manipulated between 0° and 180° in 30° steps in this experiment, the Gaussian fitting used to determine the preferred orientation was relatively less reliable, and this design potentially underestimated the effect of different predictor's

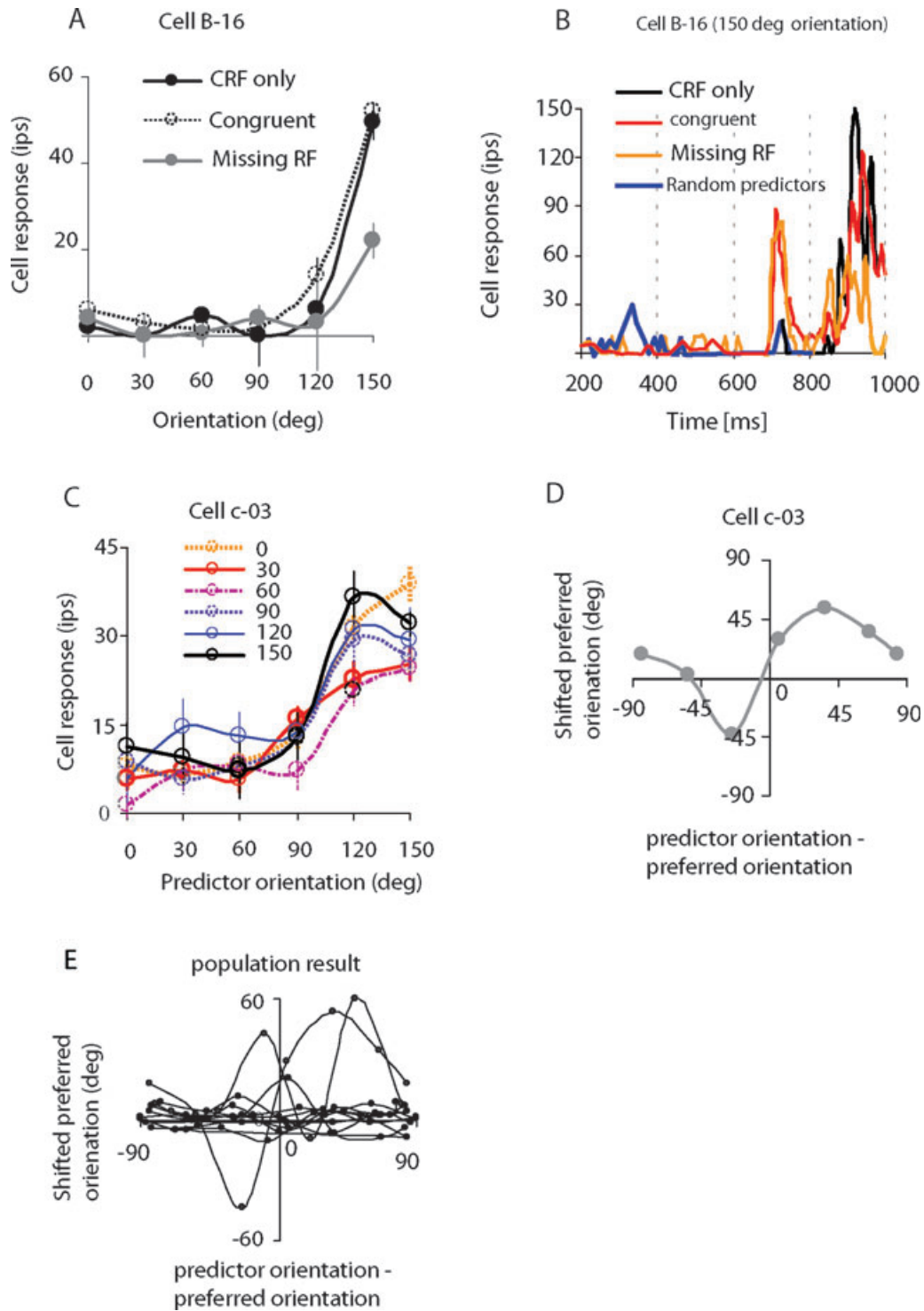


FIG. 3. (A) Orientation tuning curves for cell B-16 under CRF only, congruent and missing CRF conditions. (B) Peri-stimulus-time-histograms (10-ms bins) for the responses compared in Fig. 3A, only responses to the predictor-target sequence with the preferred orientation (150°) are shown. (C) Influence of the orientation of the predictor sequence on the neuronal responses to the target of a given orientation. Each curve represents the mean response at a fixed orientation of the target bar depending on the orientation of the predictor sequence. (D) Shift of preferred target orientation in the presence of predictors as a function of the difference between the orientation of the predictor sequence and the preferred target orientation in the absence of predictors (CRF only condition). The tuning to the target bar in the presence of predictors was estimated by fitting smooth functions to the raw orientation tuning data with an iterative least-squares algorithm (Gaussian fitting). The preferred orientation was obtained from the peak point of the best fitting Gaussian function. (E) Shift of preferred orientation for all cells. Preferred orientation was assessed as described above.

orientation on neuron's orientation tuning to the CRF target. We intend to use more finely grained differences in bar orientation to further examine this tuning shift effect in the near future.

What is the function of this modulation of CRF response introduced by the prior sequence of stimuli? One possibility is that the visual system is able to exploit the regularity of naturally moving visual

stimuli to make the representation of the stimulus in the CRF more accurate and reliable. If this idea is correct, then we would expect the CRF response to provide more information about the CRF stimulation in congruent conditions than in incongruent. To address this question, we employed Information Theoretic analysis techniques to quantify to what extent the predictors affect the accuracy of information encoding on a single trial basis. Information Theoretic quantities measure how well the neuronal response in an individual trial predicts the stimulus, and have the advantage to take properly into account all features of the response affecting the encoding of the visual stimulus, such as response modulation, tuning bandwidth, baseline firing and response variability as well as the mean firing rate reported above. For each of the 34 neurons, we measured the mutual information (the amount of knowledge about all possible presented orientations, Eqn 1) and the stimulus-specific information (a measure of how well each individual orientation can be discriminated, Eqn 2).

We first investigated whether the congruent prediction increased the information about the target orientation with respect to the incongruent condition. To this end, we computed the mutual information (MI) between the stimulus orientation and the spike counts in each trial in a 40-ms sliding window [Fig. 4A–B; 1-ms steps starting at 200 ms before CRF stimulus onset ($t = 0$) until 250 ms thereafter]. The time course of information about target orientation for an example neuron (neuron B-07) and the population result (34 neurons) are plotted in Fig. 4A and B, respectively. For neuron B-07 (Fig. 4A), the information about target orientation available from responses in incongruent trials (grey line) peaked at $t = 67$ ms, reaching a value of 0.28 bits. In the case of congruent trials (black line), the information about target orientation peaked also at $t = 67$ ms after the bar was shown in the CRF reaching a value of 0.42 bits. As here we used a 100-ms window to count the spikes, the congruent stimulation gave a peak information of 4.02 bits per second of observation of the neural response, whereas the incongruent stimulation gave a peak of 2.80 bits/s. Rates of information about visual stimuli obtained from

single V1 neurons in anesthetized macaques are reported in the range 1–15 bits/s (with the majority of data distributed in the range 1–6 bits/s), depending on the stimulus condition used (Heller *et al.*, 1995; Reich *et al.*, 2000; Reich *et al.*, 2001). Thus, our results conform to previously reported information rates, while they go beyond previous findings in demonstrating that information was higher in the congruent than the incongruent case.

For the population result (Fig. 4B), the time courses of congruent and incongruent conditions were significantly different in the time window between 10 and 90 ms after target onset (as shown by the paired t -test values on the bottom of Fig. 4B, where black bars above the light grey line indicate significant difference; $P < 0.05$, uncorrected for multiple comparisons). The information about target orientation reached a peak at ~ 70 ms post-CRF-stimulation in both conditions. At this time point, the gain of information about target orientation obtained from congruent as compared to incongruent trials was 24% (0.26 bits vs. 0.21 bits). From Fig. 4B it seems that there is a difference in information before the CRF stimulus onset between the congruent and the incongruent conditions. However, this difference is due to the fact that predictor bars had the same orientations as the CRF bar in the congruent conditions, while they had different orientations in the incongruent condition. Thus, in the incongruent condition each CRF bar orientation was preceded by all (six different) predictor bar orientations. The information for the prestimulus period was thus calculated from a set of responses to an identical mix of stimuli, which by definition should yield zero information (as was indeed the case in Fig. 4B). The two time-courses are therefore not comparable before target onset. However, the increase of information before CRF stimulation in the congruent condition is testament of tuned responses to nCRF stimulation, which is evident from Fig. 2D.

Mutual information can depend on the size of the window used to calculate it. We were thus interested whether the result of increased information with the congruent condition is consistent for

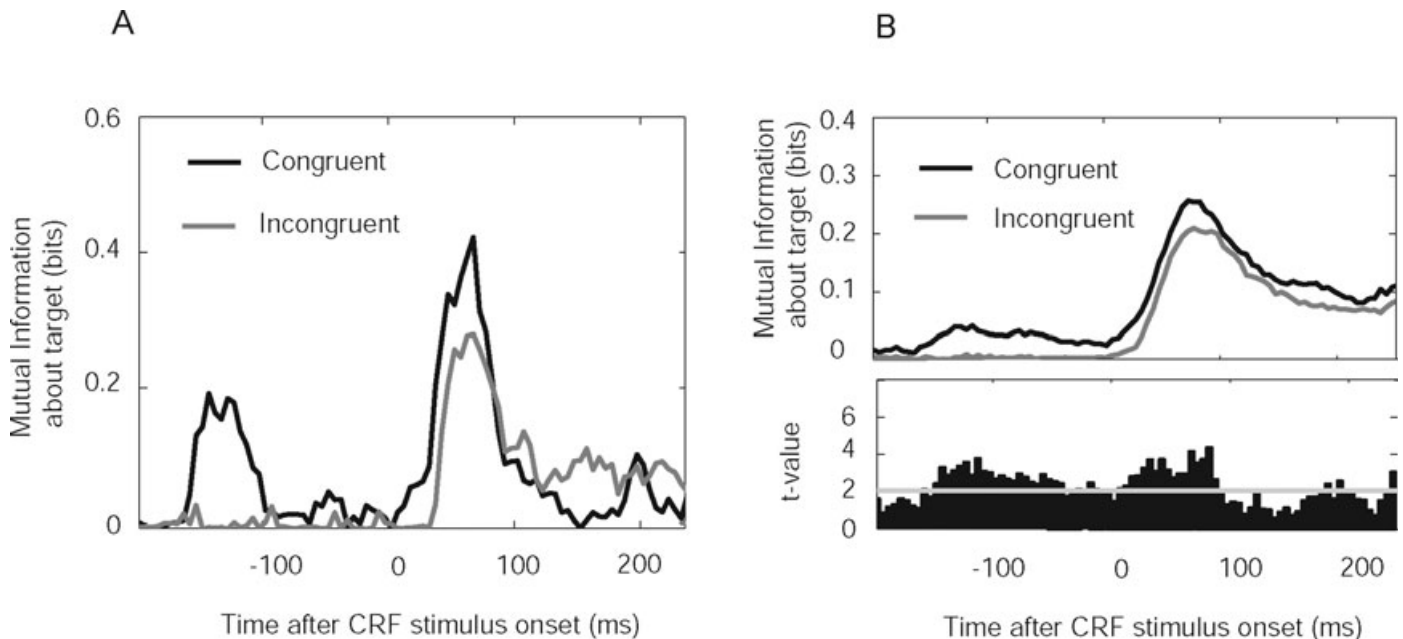


FIG. 4. (A) Mutual information about target orientation in the congruent condition (black line) and incongruent condition (grey line) for a representative neuron (Cell B-07). (B) Mutual information about target orientation in the congruent (black line), incongruent (dark grey line averaged across all neurons ($n = 34$)). The black bars at the bottom of (B) indicate the t -value corresponding to a paired t -test between the congruent and incongruent condition. Values above the grey dashed line are significant at the 5% level. $T = 0$ denotes target onset. Negative information values have been set to zero.

different window widths and calculated Mutual information using windows of 10, 20, 40 and 100 ms. We found that the congruent condition yielded more mutual information for all window width (see Fig. 5).

It would also be of interest to compare the information about the target to the CRF only condition. However, for this condition we only had ten trials for each orientation and cell, and this precluded to calculate whether the information about the CRF only condition was significantly different from the other two conditions (information measures may have been biased due to the relatively small number of trials).

The above analysis of mutual information reported the overall ability of the neuronal responses to discriminate about all presented target orientations. However, it does not tell which orientation is best represented by the neuronal responses. The latter question is addressed by the use of stimulus-specific information (see Materials and methods). For the stimulus-specific information the values reported in Fig. 6 correspond to the response period from 30 to 230 ms after CRF stimulus onset for the congruent and incongruent condition. It corresponds to the period from 200 ms before until CRF stimulus onset for the predictor condition. The population stimulus-specific information in Fig. 6 was calculated from the 40 trials per condition for each cell ($n = 34$). The stimulus-specific information was larger for all orientations in the congruent compared to the incongruent condition ($P = 0.03$, two-way ANOVA).

Neurons also carry information about the orientation of the predictor (rather than the target) particularly when the predictor orientation is of the neuron's preferred orientation (Fig. 6, light grey line). Although information about predictor orientation was much smaller than information about target orientation, it was significant for the neurons' preferred orientation (permutation test, $P < 0.01$; surrogates were created by randomly reassigning orientation labels to trials).

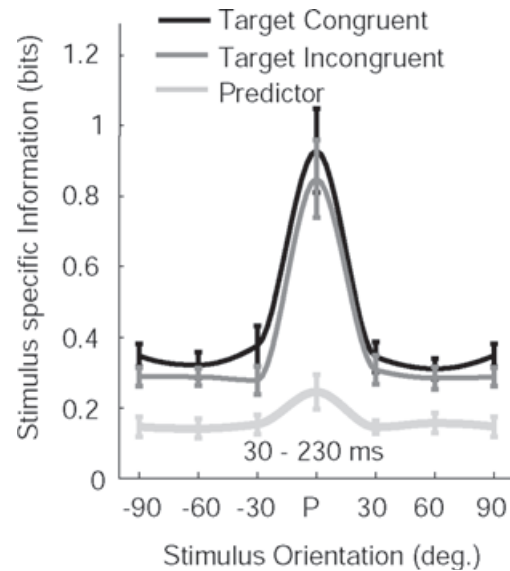


FIG. 6. Stimulus-specific information (Eqn 2) about the orientation of the target in the congruent and incongruent condition (black and dark grey lines), and the predictor (light grey line). Bars denote mean and SEM across all neurons ($n = 34$). Stimulus-specific information was calculated during the time period of 30–230 ms after CRF target onset, and the time –200 to 0 for the predictor information. Stimulus-specific information was calculated for each neuron from all 40 trials per orientation.

Taken together, these results indicate that early responses convey information about the orientation of the predictors, and that information about target orientation is influenced by the prior occurrence of the predictors. The effect of congruence was to make more information available about the target.

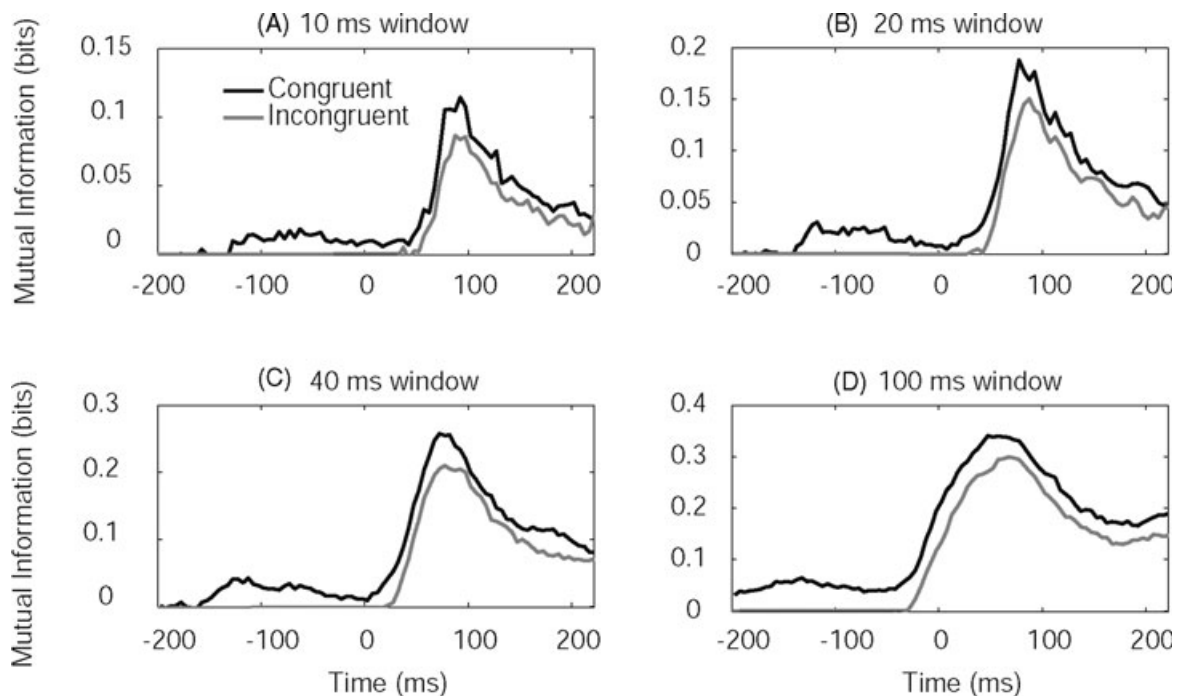


FIG. 5. Mutual information about target orientation in the congruent condition (black line) and incongruent condition (grey line) as a function of window size (A–D). Mutual information was higher in the congruent condition compared to the incongruent condition for window sizes of 10, 20, 40, and 100 ms. In line with previous reports mutual information increased with window size. Negative information values have been set to zero.

Discussion

We found that neuronal responses of V1 neurons were significantly modulated by spatio-temporal information occurring well outside and prior to stimulation of their CRFs. Predictors presented in a linearly ordered sequence caused 'early' orientation tuned responses. Moreover the orientation of predictors could significantly affect the orientation tuning elicited by bars occurring inside the CRF, and could result in tuned predictive responses when in fact the CRF stimulus was omitted. We will first address and rule out the possibility that our results are a simple artefact of CRF underestimation. Thereafter we will discuss our findings in relation to established centre-surround interactions in V1 and argue that our data can be interpreted in light of inferential processes that promote 'meaningful' scene interpretation.

It is unlikely that neuronal responses to the predictors in our study were due to underestimation of the CRF size. Firstly, our method of CRF estimation (annular response field) is usually considered to be a conservative estimate of the CRF size (Cavanaugh *et al.*, 2002). Secondly, the early response was only recorded when the predictors were presented in an ordered and predictable sequence. Predictors presented at exactly the same locations but in a randomized sequence did not elicit responses significantly above spontaneous level. Thirdly, although V1 neurons extend their CRFs into surrounding regions under some stimulus conditions (e.g. artificial scotomas), this process occurs over a period of several minutes (Pettet & Gilbert, 1992), not within the very short periods used in our experiment. Our findings can also not be explained by changes in spatial integration due to changes in luminance contrast (Kapadia *et al.*, 1999; Sceniak *et al.*, 1999; Cavanaugh *et al.*, 2002) as our stimuli were presented at constant high luminance contrast. We would rather argue that the early responses result from changes in CRF size due to spatio-temporal regularities in the CRF's vicinity. These spatio-temporal regularities could result in increased surround summation by which the previously nonresponsive modulatory surround becomes sufficiently influential to drive responses. Indeed spatial integration of V1 neurons varies with the presence or absence of surrounding stimuli (Kapadia *et al.*, 1999). Such a change allows neurons to switch in a context dependent manner from scene segmentation to integration. Such a form of short-term plasticity (in the ms range) would enable neurons to integrate dispersed signals in a meaningful and structured way.

In this study we provided evidence that V1 neurons are sensitive and strongly influenced by the spatio-temporal regularity of stimuli occurring outside their CRFs. This spatio-temporal regularity mimics (in a reductionist setting) the often experienced statistical structure in our dynamic visual world, as objects generally move in a predictable (and smooth) manner, and changes of orientation also occur in a smoothly graded and predictable way (Young, 2000). Without direct stimulation of CRFs, neurons showed orientation selective early responses to the predictors in the congruent condition; some even showed modulated responses during the time window in which a target would be expected to appear, while in effect it was omitted (missing CRF condition). This observation cannot be explained by 'low level' mechanisms such as adaptation or divisive contrast gain control (Heeger, 1992; Wilson & Humanski, 1993; Levitt & Lund, 1997), as the predictors presented at the same location but in a randomized sequence failed to elicit above-spontaneous neuronal activities, or any form of tuned responses. Our finding is also at odds with classically documented centre-surround interaction in V1. In centre-surround interactions the co-presented surrounding stimuli modulate neuronal activity elicited by stimuli placed inside the CRF, but do not drive responses when the surround is presented alone (Albright & Stoner, 2002). In contrast, our predictors presented outside the CRF in the

congruent condition often elicited early responses in the absence of direct CRF stimulation. Moreover, centre-surround interaction mainly suppresses neuronal responses to the CRF stimulation when the copresented high-contrast visual stimuli inside and outside the CRF are colinear (Nelson & Frost, 1978; Levitt & Lund, 1997; Jones *et al.*, 2002; Guo *et al.*, 2005), while cross-orientation facilitation can occur when centre and surround stimuli have different orientations (Jones *et al.*, 2002). Iso-orientation suppression and cross-orientation facilitation are both mechanisms by which scene segmentation is supported. Scene segmentation is fundamental for visual analysis, however, integration is equally important (Thiele, 2007). The increase of selectivity from spatio-temporal predictive pattern found in our study could be interpreted as a mechanism by which integration, rather than segmentation, is mediated.

In nearly one-third (10/34) of our neurons orientation tuning to stimuli presented inside the CRF depended on the orientation of the predictor sequence. For some of them, the preferred orientation to the target bar was significantly biased towards the orientation of the predictor sequence. Previous studies demonstrated that V1 neurons' orientation selectivity shows remarkable plasticity to complex temporal or spatial stimulus dynamics. For example, paired visual stimuli with different orientations presented within the CRF successively can induce a shift in a neuron's orientation tuning (Muller *et al.*, 1999; Dragoi *et al.*, 2000), the direction of shift depending on the temporal order of the pair, an effect that may be linked to the perceptual level (Yao & Dan, 2001). However, our results cannot be explained in terms of adaptation or fatigue of orientation selective neurons, as there was no direct CRF stimulation during the presentation of the predictor sequence in our experiment.

Recent influential ideas argue that analysis of visual scenes benefits substantially from inference (Knill & Richards, 1996; Young, 2000; Friston, 2002). Within the framework of 'vision-as-inference', responses of visual neurons should be highly context-sensitive, and neuronal responses evoked by CRF inputs should change with the context established by prior information and expectations. While there is evidence that neurons at later stages of the visual hierarchy respond to higher cognitive demands based on previous knowledge and experience, such as perceptual filling-in, anticipation and prediction (Duhamel *et al.*, 1992; Thiele & Hoffmann, 1996; Friston, 2002), the degree to which neurons at early stages are susceptible to prior information is less known. It has, however, been established that V1 neurons are sensitive to images containing real world probability structure (Baddeley *et al.*, 1997; Vinje & Gallant, 2000), such that they use context to predict missing information and to generate tuned responses that are only weakly dependent on analysis of the local image region corresponding to their CRFs (Grosf *et al.*, 1993; Sugita, 1999; Komatsu *et al.*, 2000). We think that our findings could be interpreted in the framework of 'vision-as-inference', where visual neurons have access, through their embedding neural network, to information about the distribution of prior probabilities of stimuli. As a result, the output of a neuron critically depends on the interaction between the likelihood function (i.e. CRF visual input) and the prior probability (i.e. extra-CRF information) rather than simply on direct visual input from its CRF (Young *et al.*, 2000; Sharma *et al.*, 2003; Guo *et al.*, 2004b). The source of this information may be the lateral connections within V1, or feedback connections from higher areas. From our current data it is impossible to infer which of the two are more influential.

In light of our results, a natural interpretation is that neurons in area V1 are not only specialized for extracting local features (i.e. orientation), but also represent and encode signals that reflect the statistical structure of the visual world, such as spatio-temporal

regularities, as proposed by the model of 'vision-as-inference'. On this interpretation, a neurophysiologically mapped CRF represents the neuron's projection of an inferred probability onto the world, and not the result of simple analysis of the local features within the CRF.

Acknowledgements

This work was supported by Wellcome Trust, EPSRC, HFSP and EU FP5. We thank Dr Y. Tadmor for technical advice.

Abbreviations

CRF, classical receptive field; nCRF, larger areas outside their CRF.

References

- Albright, T.D. & Stoner, G.R. (2002) Contextual influences on visual processing. *Annu. Rev. Neurosci.*, **25**, 339–379.
- Angelucci, A., Levitt, J.B., Walton, E.J., Hupe, J.M., Bullier, J. & Lund, J.S. (2002) Circuits for local and global signal integration in primary visual cortex. *J. Neurosci.*, **22**, 8633–8646.
- Baddeley, R., Abbott, L.F., Booth, M.C., Sengpiel, F., Freeman, T., Wakeman, E.A. & Rolls, E.T. (1997) Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proc. Biol. Sci.*, **264**, 1775–1783.
- Berry, M.J., 2nd, Brivanlou, I.H., Jordan, T.A. & Meister, M. (1999) Anticipation of moving stimuli by the retina. *Nature*, **398**, 334–338.
- Borst, A. & Theunissen, F.E. (1999) Information theory and neural coding. *Nature Neurosci.*, **2**, 947–957.
- Cavanaugh, J.R., Bair, W. & Movshon, J.A. (2002) Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *J. Neurophysiol.*, **88**, 2530–2546.
- Chisum, H.J. & Fitzpatrick, D. (2004) The contribution of vertical and horizontal connections to the receptive field center and surround in V1. *Neural Networks*, **17**, 681–693.
- Das, A. & Gilbert, C.D. (1995) Long-range horizontal connections and their role in cortical reorganization revealed by optical recording of cat primary visual cortex. *Nature*, **375**, 780–784.
- Dragoi, V., Sharma, J. & Sur, M. (2000) Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron*, **28**, 287–298.
- Duhamel, J.-R., Colby, C.L. & Goldberg, M.E. (1992) The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, **255**, 90–92.
- Felleman, D.J. & Van Essen, D.C. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex*, **1**, 1–47.
- Friston, K. (2002) Functional integration and inference in the brain. *Prog. Neurobiol.*, **68**, 113–143.
- Grosf, D.H., Shapley, R.M. & Hawken, M.J. (1993) Macaque V1 neurons can signal 'illusory' contours. *Nature*, **365**, 550–552.
- Guo, K., Benson, P.J. & Blakemore, C. (2004a) Pattern motion is present in V1 of awake but not anaesthetized monkeys. *Eur. J. Neurosci.*, **19**, 1055–1066.
- Guo, K., Nevado, A., Robertson, R.G., Pulgarin, M., Thiele, A. & Young, M.P. (2004b) Effects on orientation perception of manipulating the spatio-temporal prior probability of stimuli. *Vis. Res.*, **44**, 2349–2358.
- Guo, K., Robertson, R.G., Mahmoodi, S. & Young, M.P. (2005) Centre-surround interactions in response to natural scene stimulation in the primary visual cortex. *Eur. J. Neurosci.*, **21**, 536–548.
- Heeger, D.J. (1992) Normalization of cell responses in cat striate cortex. *Vis. Neurosci.*, **9**, 181–197.
- Heller, J., Hertz, J.A., Kjaer, T.W. & Richmond, B.J. (1995) Information flow and temporal coding in primate pattern vision. *J. Comp. Neurosci.*, **2**, 175–193.
- Huang, X., Albright, T.D. & Stoner, G. (2007) Adaptive surround modulation in cortical area MT. *Neuron*, **53**, 761–770.
- Jones, H.E., Wang, W. & Sillito, A.M. (2002) Spatial organization and magnitude of orientation contrast interactions in primate v1. *J. Neurophysiol.*, **88**, 2796–2808.
- Kapadia, M.K., Westheimer, G. & Gilbert, C.D. (1999) Dynamics of spatial summation in primary visual cortex of alert monkeys. *Proc. Natl Acad. Sci. USA*, **96**, 12073–12078.
- Kapadia, M.K., Westheimer, G. & Gilbert, C.D. (2000) Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *J. Neurophysiol.*, **84**, 2048–2062.
- Knierim, J.J. & van Essen, D.C. (1992) Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J. Neurophysiol.*, **67**, 961–980.
- Knill, D.C. & Richards, W. (1996) *Perception as Bayesian Inference*. Cambridge University Press, New York.
- Komatsu, H., Kinoshita, M. & Murakami, I. (2000) Neural responses in the retinotopic representation of the blind spot in the macaque V1 to stimuli for perceptual filling-in. *J. Neurosci.*, **20**, 9310–9319.
- Lennie, P. (1998) Single units and visual cortical organization. *Perception*, **27**, 889–935.
- Levitt, J.B. & Lund, J.S. (1997) Contrast dependence of contextual effects in primate visual cortex. *Nature*, **387**, 73–76.
- Li, W., Piech, V. & Gilbert, C.D. (2006) Contour saliency in primary visual cortex. *Neuron*, **50**, 951–962.
- Livingston, M.S. & Hubel, D.H. (1987) Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *J. Neurosci.*, **7**, 3416–3468.
- Muller, J.R., Metha, A.B., Krauskopf, J. & Lennie, P. (1999) Rapid adaptation in visual cortex to the structure of images. *Science*, **285**, 1405–1408.
- Nelson, J.I. & Frost, B.J. (1978) Orientation selective inhibition from beyond the classic visual receptive field. *Brain Res.*, **139**, 359–365.
- Panzeri, S. & Treves, A. (1996) Analytical estimates of limited sampling in different information measures. *Network*, **7**, 87–107.
- Pettet, M.W. & Gilbert, C.D. (1992) Dynamic changes in receptive-field size in cat primary visual cortex. *Proc. Natl Acad. Sci. USA*, **89**, 8366–8370.
- Pola, G., Thiele, A., Hoffmann, K.P. & Panzeri, S. (2003) An exact method to quantify the information transmitted by different mechanisms of correlational coding. *Network*, **14**, 35–60.
- Reich, D.S., Mechler, F., Purpura, K.P. & Victor, J.D. (2000) Interspike intervals, receptive fields, and information encoding in primary visual cortex. *J. Neurosci.*, **20**, 1964–1974.
- Reich, D.S., Mechler, F. & Victor, J.D. (2001) Formal and attribute-specific information in primary visual cortex. *J. Neurophysiol.*, **85**, 305–318.
- Sceniak, M.P., Ringach, D.L., Hawken, M.J. & Shapley, R. (1999) Contrast's effect on spatial summation by macaque V1 neurons. *Nature Neurosci.*, **2**, 733–739.
- Shannon, C. (1948) A mathematical theory of communication. *Bell. Sys. Tech. J.*, **27**, 379–423.
- Sharma, J., Dragoi, V., Tenenbaum, J.B., Miller, E.K. & Sur, M. (2003) V1 neurons signal acquisition of an internal representation of stimulus location. *Science*, **300**, 1758–1763.
- Sugita, Y. (1999) Grouping of image fragments in primary visual cortex. *Nature*, **401**, 269–272.
- Super, H., Spekreijse, H. & Lamme, V.A. (2003) Figure-ground activity in primary visual cortex (V1) of the monkey matches the speed of behavioral response. *Neurosci. Lett.*, **344**, 75–78.
- Thiele, A. (2007) Reconstructing the world: Switching from segmentation to integration allows neurons in area MT to make 'sense' of the visual scene. *Neuron*, **53**, 623–625.
- Thiele, A. & Hoffmann, K.P. (1996) Neuronal activity in MST and STPp, but not MT changes systematically with stimulus-independent decisions. *Neuroreport*, **7**, 971–976.
- Vinje, W.E. & Gallant, J.L. (2000) Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, **287**, 1273–1276.
- Wilson, H.R. & Humanski, R. (1993) Spatial frequency adaptation and contrast gain control. *Vis. Res.*, **33**, 1133–1149.
- Yao, H. & Dan, Y. (2001) Stimulus timing-dependent plasticity in cortical processing of orientation. *Neuron*, **32**, 315–323.
- Young, M.P. (2000) The architecture of visual cortex and inferential processes in vision. *Spat. Vis.*, **13**, 137–146.
- Young, M.P., Hilgetag, C.-C. & Scannell, J.W. (2000) On imputing function to structure from the behavioral effects of brain lesions. *Philos Trans R Soc Lond [Biol.]*, **355**, 147–161.
- Zipser, K., Lamme, V.A.F. & Schiller, P.H. (1996) Contextual modulation in primary visual cortex. *J. Neurosci.*, **16**, 7376–7389.