



Effects on orientation perception of manipulating the spatio–temporal prior probability of stimuli

Kun Guo ^{*,1}, Angel Nevado ¹, Robert G Robertson, Maribel Pulgarin, Alexander Thiele, Malcolm P Young

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

Received 21 January 2004; received in revised form 21 April 2004

Abstract

Spatial and temporal regularities commonly exist in natural visual scenes. The knowledge of the probability structure of these regularities is likely to be informative for an efficient visual system. Here we explored how manipulating the spatio–temporal prior probability of stimuli affects human orientation perception. Stimulus sequences comprised four collinear bars (predictors) which appeared successively towards the foveal region, followed by a target bar with the same or different orientation. Subjects' orientation perception of the foveal target was biased towards the orientation of the predictors when presented in a highly ordered and predictable sequence. The discrimination thresholds were significantly elevated in proportion to increasing prior probabilities of the predictors. Breaking this sequence, by randomising presentation order or presentation duration, decreased the thresholds. These psychophysical observations are consistent with a Bayesian model, suggesting that a predictable spatio–temporal stimulus structure and an increased probability of collinear trials are associated with the increasing prior expectation of collinear events. Our results suggest that statistical spatio–temporal stimulus regularities are effectively integrated by human visual cortex over a range of spatial and temporal positions, thereby systematically affecting perception.

© 2004 Published by Elsevier Ltd.

Keywords: Orientation discrimination; Predictor; Spatio–temporal regularity; Bayesian inference

1. 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, what we see is not simply an immediate reflection of the physical characteristics of a scene but instead is highly dependent on the processes by which our brain attempts to interpret the scene. Our visual perception, therefore, is the result of an interaction between information entering the eyes and mechanisms that relate the incoming information to previous visual experiences and current expectations (Gilbert, 1998).

Accordingly, it is now increasingly acknowledged that the process of visual perception involves knowledge-rich inferential processes (Barlow, 1989; Friston, 2002; Knill & Richards, 1996; Young, 2000). In this framework, knowledge of the probability structure of the spatio–temporal dynamics of scenes is likely to be informative for an efficient visual system, and the responses of visual neurons and our perceptual sensitivities should be substantially altered by processing scenes for which the system has prior expectations of dynamics or structure. Growing evidence from neurophysiological investigation suggests that neural systems do make use of the probability structure of the real world for the efficient information coding (Abbott & Blum, 1996; Chirimuuta, Clatworthy, & Tolhurst, 2003; Guo et al., 2002; Mehta, Quirk, & Wilson, 2000; Sharma, Dragoi, Tenenbaum, Miller, & Sue, 2003; Yao & Dan, 2001; Zhang, Ginzburg, McNaughton, & Sejnowski, 1998). In computational terms, because scenes and stimuli in the visual world are not all presented with equal probability,

* Corresponding author. Tel.: +44-191-222-5942; fax: +44-191-222-5622.

E-mail address: kun.guo@ncl.ac.uk (K. Guo).

¹ Both authors made the same contribution to this project.

it would be statistically optimal for the brain to compute the Bayesian posterior probability of a scene (perception), given the prior probability of the state and the current input information from the eye ($Posterior \approx Prior \times Likelihood$) (Friston, 2002; Knill & Richards, 1996; Rao, 1999). Indeed, recent psychophysical measurements demonstrated that the human perceptual performance to luminance discrimination (Ciaramitaro, Cameron, & Glimcher, 2001), motion perception (Ascher & Grzywacz, 2000; Hürlimann, Kiper, & Carandini, 2002; Weiss, Simoncelli, & Adelson, 2002), depth perception (Kersten, Mamassian, & Knill, 1997) and object recognition (Liu, Knill, & Kersten, 1995; Rao, 1999) can be influenced by manipulating the probability structure of visual stimuli, as predicted by the Bayesian perspective.

Human observers are extremely effective in judging the orientation of lines (Westheimer, 1990). However, when measuring foveal orientation sensitivity to a target line, previous psychophysical studies have shown that orientation perception can be manipulated by the spatial or temporal arrangement of the target line and surrounding configurations. For example, when surrounded by lines at differing orientation, the perceived orientation of the center line appears to shift in a direction away from that of the surround lines (orientation contrast effect) (Wenderoth & Johnstone, 1988; Westheimer, 1990). The orientation discrimination threshold to the center line can also be markedly raised by the surround lines of a similar orientation (Kapadia, Westheimer, & Gilbert, 2000; Mareschal, Sceniak, & Shapley, 2001), even if they are presented a short time before or after the center line (Wehrhahn, Li, & Westheimer, 1996; Westheimer & Ley, 1997). In this experiment, we employed the orientation discrimination threshold as a test parameter to investigate the effect of manipulation of the spatio-temporal prior probability of a stimulus on human psychophysical performance, and we developed a Bayesian model to account for the psychophysical observations.

2. Materials and methods

Visual stimuli, five short bars (1.5° length, 0.1° width, 24 cd/m^2 luminance), were generated using VSG 2/3 graphics system (Cambridge Research Systems) and displayed on a high frequency non-interlaced gamma-corrected color monitor (110 Hz, 1024×768 pixels resolution, Sony GDM-F500T9) with uniform grey background (6 cd/m^2 luminance). Four co-linear bars (predictors 1–4) had the same diagonal orientation (45°), the orientation of the fifth bar (target) varied randomly from 35° to 55° with 1° steps. A small red fixation point (FP, 0.15° diameter, 7.8 cd/m^2 luminance) was displayed at the location where the target would appear (Fig. 1).

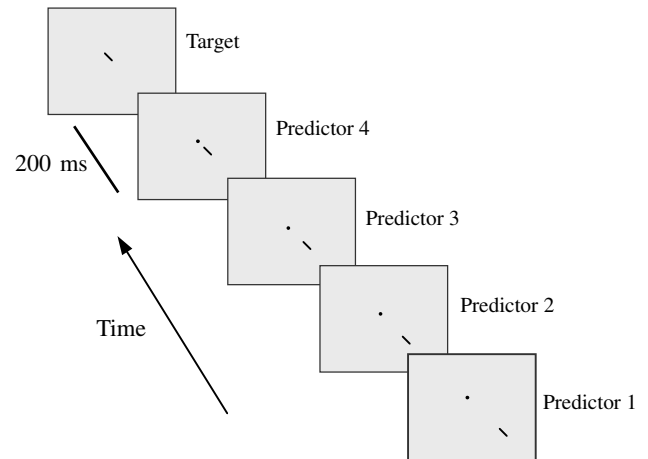


Fig. 1. Stimulus demonstration. Five short bars were presented in a linearly ordered spatio-temporal sequence. Four of them (first, second, third and fourth bars) were predictors with the same orientation (45°), and one of them (fifth bar) was the target with randomly varied orientation ($35\text{--}55^\circ$ with 1° increment). Each bar was presented for 200 ms. A small red fixation point (0.15° diameter) was displayed at the center of the target.

Five volunteers (two of them authors) with normal or corrected-to-normal visual acuity participated as subjects. They were trained for several sessions (normally 100 trials) before the formal test. During the experiments, the observers sat in a quiet, darkened area enclosed by curtains, and viewed the display binocularly. At a viewing distance of 114 cm the monitor subtended a visual angle of $20^\circ \times 15^\circ$. The trial was started by a 350 Hz warning tone lasting 150 ms followed by a delay of 1000 ms. The four predictors and target bar were then presented on the screen in a highly predictable spatial and temporal sequence (predictor 1 \rightarrow predictor 2 \rightarrow predictor 3 \rightarrow predictor 4 \rightarrow target). 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. In a given trial, the orientation of the target was randomly varied $\pm 10^\circ$ around the predictors with 1° steps, so the probability of a physically collinear target was one out of twenty one. The observers were instructed to maintain fixation of the FP throughout the trial, and had to indicate, by pressing one of two keys in a computer keyboard, whether the target had the same or different orientation as the predictors (collinear or non-collinear) after the stimulus presentation. No feedback was given. The chance performance (correct judgment of the target orientation) in a given presentation was 50%. The trial interval was set to 1500 ms. A minimum of 20 trials were presented for each target orientation. During the experiments, the observers were encouraged to have a short break if it was necessary.

The observers' performance (percentage of correct target orientation judgment) was measured as a function

of orientation difference between the predictors and the target. A cumulative normal psychometric function was fitted to the data of each subject by probit analysis (Finney, 1986). The goodness of fit was determined by computing the chi-square (χ^2) statistic and probability associated with the hypothesis that the difference between data and fit was due to noise. The orientation discrimination threshold was determined as the point at which the psychometric function crossed the 75% performance level. To prevent perceptual learning (training) or other long-term effects, results in each test were always obtained as a self-contained series even though sometimes two series included identical patterns. This explains the several occasions in this study where there were differences between threshold values for the same observer.

3. Results

3.1. Psychophysical observation

To establish whether the predictable priors (predictors) can affect the subjects' orientation perception, the orientation discrimination threshold for the target was measured in three different protocols. (1) Normal sequence: predictors and target were illuminated in a predictable spatial and temporal sequence as shown in Fig. 1. Each bar was presented for 200 ms. (2) Random order sequence: predictors and target were illuminated in random spatial and temporal sequence. Each bar was presented for 200 ms. (3) Random duration sequence: predictors and target were illuminated in a predictable spatial sequence, and the target bar was presented for 200 ms, but the presentation time for each predictor was varied randomly between 100 and 500 ms.

The predictors illuminated in a linearly ordered spatial and temporal sequence (condition 1) significantly reduced the subjects' orientation sensitivity to the target (Fig. 2A). Their orientation perception to the target was biased towards the predictors' orientation. Consequently, the highest orientation discrimination threshold ($5.0^\circ \pm 0.21^\circ$, Mean \pm SEM) was observed under this condition (ANOVA, $p < 0.01$). Breaking this regularity inherent in the priors by randomising illumination order (condition 2) or duration (condition 3) reduced the bias of the target orientation judgment. The orientation discrimination thresholds were decreased by 39% ($3.07^\circ \pm 0.22^\circ$) and 20% ($4.0^\circ \pm 0.21^\circ$) respectively. The difference between discrimination thresholds measured under the conditions of random order and random duration sequence also indicates that the spatial regularity of the sequence appears to be more critical than the temporal regularity to influence the orientation perception of the target.

Clearly, subjects' orientation perception was influenced by the regularity of the spatio-temporal priors. In other words, prior expectation of co-linearity under the test condition of normal sequence strongly biased the orientation judgment of the foveal target bar towards that of the predictors. We further investigated whether the probability of the collinear trials can influence the subjects' performance. If the prior expectation can also be influenced by recent visual experience, a high frequency collinear trials should reduce subjects' orientation sensitivity to the target. The tests were arranged in two blocks. In one block, the probability of the target having the same orientation as the predictors (the probability of the collinear trials) was set at 0%. In the second block, the probability was set to 50%. In both blocks, the predictors and the target were presented in a predictable spatial and temporal sequence (normal

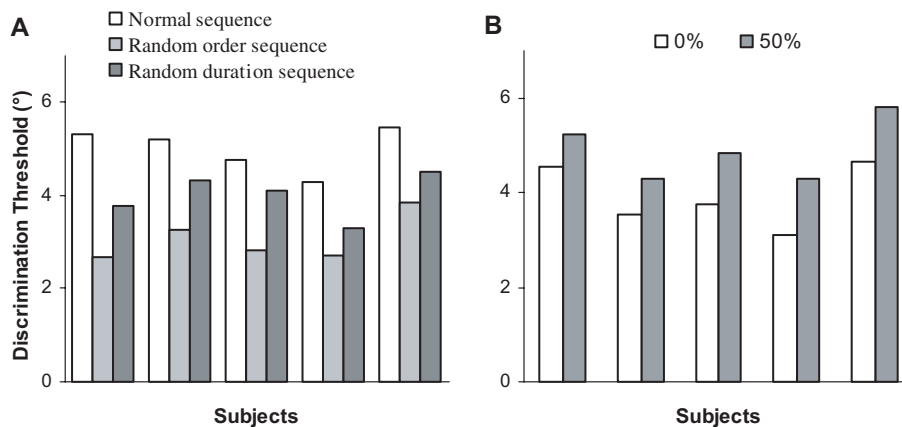


Fig. 2. Effect of spatio-temporal prior probability on orientation discrimination threshold. (A) The threshold was tested under three different stimulus conditions: normal sequence, random order sequence and random duration sequence. The observers' orientation discrimination thresholds for the foveal target were increased significantly when the predictors were presented in normal sequence. Breaking this normal sequence by randomising illuminating sequence or duration, decreased the discrimination thresholds. (B) Orientation discrimination thresholds to the foveal target bar were increased significantly when the probability of the collinear trials was increased from 0% to 50%.

sequence). Given that in block one there never was a collinear stimulus (although the subjects were not informed about this), one would expect to see more “non-collinear” responses if the system was able to sense this difference, i.e. more correct responses and therefore a lower threshold. In block two (normal sequence with 50% collinear trials), the orientation discrimination threshold was calculated only for those trials that are not collinear, so that it is compatible with the threshold calculated from the data sampled in block one (normal sequence with 0% collinear trials). The result showed that decreasing the probability of the collinear trials significantly increased subjects’ orientation sensitivity to the target bar (Fig. 2B), which is consistent with our prediction. The mean discrimination threshold decreased by 20%, from $4.89^\circ \pm 0.29^\circ$ to $3.93^\circ \pm 0.3^\circ$, when the probability of the collinear trials was decreased from 50% to 0% (*t*-test, $p < 0.01$).

During the trial presentation, the subjects were instructed to maintain fixation of the fixation point which was presented at the location where the target bar would appear. Although we did not monitor the subjects’ eye movements, it is unlikely that the differences among the orientation discrimination thresholds measured under different test conditions were due to eye tracking of individual bars and repeated fovea stimulation (i.e. orientation adaptation). In fact, a similar or the same predictor presentation in different test conditions (i.e. normal sequence vs random duration sequence; normal sequence with 0% collinear trials vs normal sequence with 50% collinear trials) yielded statistically different results, which is in disagreement with the difference between test conditions being due to adaptation.

Human foveal orientation discrimination thresholds can be affected by stimulus length and exposure duration (Westheimer, 1990; Westheimer & Ley, 1997). To determine the range of the spatial summation of spatio-temporal priors, we systematically varied the number of predictors between 1 and 8, covering between 1.5° and 12° of the visual field. In this experiment, the predictors were always presented in turn preceding the target at successive positions that approached the target position (normal sequence). For trials containing only one predictor, it was presented in a position immediately adjacent and in a time immediately preceding the target; for trials with eight predictors, the first predictor was presented at the farthest position away from the target in both space and time. Four subjects’ orientation discrimination thresholds to the target were measured as a function of predictor numbers. On average, the discrimination threshold was gradually increased from $3.70^\circ \pm 0.19^\circ$ to $5.28^\circ \pm 0.21^\circ$ when the number of the predictors was increased from 1 to 8 (ANOVA, $p < 0.01$) (Fig. 3A). However, this increasing threshold is not associated linearly with the increasing predictor numbers. The threshold increased rapidly (0.29° per

predictor) when the number of the predictors was changed from 1 to 6. After that, the threshold was close to saturation. This result clearly indicates that human visual cortex can effectively integrate briefly presented orientation signals prior to and distant from the foveal target bar; in our case, this temporal summation covers over 1200 ms and spatial summation extends over 9° (6 predictors).

We also systematically varied the stimulus duration of the predictors and the target (between 100 and 400 ms) to determine the temporal summation of the spatio-temporal priors. In this experiment, four predictors and target were presented in a predictable spatial and temporal sequence (normal sequence). The orientation discrimination thresholds of four subjects for the target were measured as a function of the presentation duration (Fig. 3B). The discrimination thresholds were monotonically decreased with the increasing stimulus exposure duration (ANOVA, $p < 0.01$). The mean threshold was $5.37^\circ \pm 0.35^\circ$ with 100 ms presentation duration, and $3.53^\circ \pm 0.16^\circ$ with 400 ms presentation duration.

3.2. Modeling

We further investigated to what extent a Bayesian model can account for the psychophysical results. In the context of Bayesian inference (i.e. Knill & Richards, 1996), if the system’s prior expectation is that collinear stimuli are more likely than non-collinear ones, the perceived orientation of the target will be biased towards that of the predictors. We expect that the expectation for a collinear target will be higher (sharper prior distribution) when the predictors’ sequence resembles a natural sequence which is well defined and predictable (i.e. normal sequence). Likewise, an increase in the experimental frequency of collinear trials (i.e. 50% col-

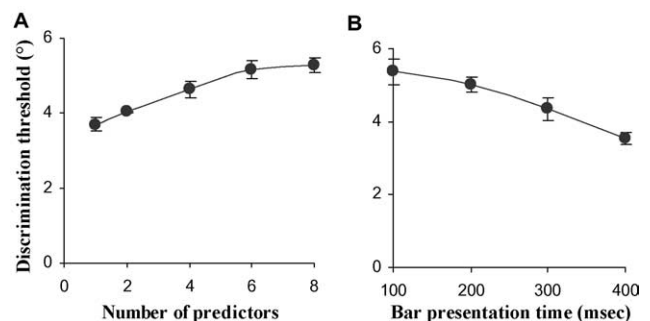


Fig. 3. (A) Orientation discrimination threshold to the target bar as a function of the number of predictors. The thresholds of four observers were increased with increasing predictor numbers. (B) Orientation discrimination threshold to the target bar as a function of the presentation duration of individual predictors. The thresholds of four observers were decreased with increasing presentation duration. Error bars indicate standard error of the mean.

linear trials condition) may produce an analogous increase in the system's prior expectations.

In our model, we consider three orientation differences between the predictors and the target: (1) The real (physical) orientation difference θ , (2) the representation of this orientation difference in the brain θ_r , which is subject to noise, before taking into account the prior probabilities, and (3) the perceived orientation difference θ_p , which is based on θ_r and the priors.

In the Bayesian framework, the probability $\rho(\theta|\theta_r)$ of a real orientation difference θ given the brain representation of that difference θ_r can be expressed as the product of the prior expectation $\rho(\theta)$ times the likelihood of the representation given the real difference $\rho(\theta_r|\theta)$.

$$\rho(\theta|\theta_r) = \frac{1}{k} \rho(\theta) \rho(\theta_r|\theta) \quad (1)$$

where k is a normalisation constant. The system estimates the real orientation difference between predictors and target θ based on the noisy representation of this orientation difference by the brain θ_r and the prior expectations. We assume that this representation, which is defined by the likelihood, is subject to Gaussian noise.

$$\rho(\theta_r|\theta) = G(\theta_r - \theta; \sigma_1) \quad (2)$$

$$G(x; \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (3)$$

where $G(x, \sigma)$ represents a Gaussian function of argument x and standard deviation σ . The confidence in the sensory representation increases as the width of the likelihood σ_1 , the noise, decreases.

Since co-linearity commonly occurs in natural settings (Geisler, Perry, Super, & Gallogly, 2001; Sigman, Cecchi, Gilbert, & Magnasco, 2001), the visual system may assign a higher prior probability to the collinear stimuli. We take the distribution of prior probability as a Gaussian centered at 0° orientation difference:

$$\rho(\theta) = G(\theta - 0; \sigma_p) \quad (4)$$

The smaller the width of the prior σ_p the higher the prior expectation of a collinear target is, and therefore, the closer to collinear the perceived orientation difference θ_p between the predictors and the target will be.

Substituting Eqs. (2) and (4) in Eq. (1) we obtain the probability of a perceived orientation difference given the representation of the orientation difference before taking the priors into account. The perceived orientation difference can be estimated as the maximum or the average of this posterior probability. In our case both values coincide.

$$\theta_p = \frac{\sigma_p^2 \theta_r}{\sigma_p^2 + \sigma_1^2} \quad (5)$$

This expression gives us the estimated perceived orientation difference θ_p for each possible representation of the real difference θ_r . Therefore there is a deterministic relationship between θ_p and θ_r . On the other hand, due to noise, each real difference θ gives rise to a distribution of representations $\rho(\theta_r|\theta)$ (Eq. (2)) and therefore it is also associated with a distribution of perceived orientation differences $\rho(\theta_p|\theta)$:

$$\rho(\theta_p|\theta) = \rho(\theta_r|\theta) \frac{d\theta_r}{d\theta_p} \quad (6)$$

From Eqs. (2) and (4) we have

$$\rho(\theta_p|\theta) = G\left(\theta_p - \frac{\sigma_p^2 \theta}{\sigma_p^2 + \sigma_1^2}; \frac{\sigma_p^2 \sigma_1}{\sigma_p^2 + \sigma_1^2}\right) \quad (7)$$

We now assume that if the perceived orientation difference is smaller than a certain threshold θ_t the subject treats the predictors and the target as collinear. From Eq. (7) we can calculate the fraction of times the subject perceived predictors and target as collinear for each real orientation difference θ

$$\begin{aligned} C(\theta) &= \int_{-\theta_t}^{\theta_t} d\theta_p \rho(\theta_p|\theta) \\ &= \frac{1}{2} \left(\operatorname{erf}\left(\frac{\theta_t - \bar{\theta}_p}{\sqrt{2}\sigma_{\theta_p}}\right) - \operatorname{erf}\left(\frac{-\theta_t - \bar{\theta}_p}{\sqrt{2}\sigma_{\theta_p}}\right) \right) \\ \operatorname{erf}(x) &= \frac{2}{\sqrt{\pi}} \int_0^x \exp(-y^2) dy \end{aligned} \quad (8)$$

$$\bar{\theta} = \frac{\sigma_p^2 \theta}{\sigma_p^2 + \sigma_1^2}$$

$$\sigma_{\theta_p} = \frac{\sigma_p^2 \sigma_1}{\sigma_p^2 + \sigma_1^2}$$

The fraction of collinear responses $C(\theta)$ can be fitted with a least-squares regression to the psychophysical data, by adjusting the three free parameters: likelihood width σ_1 , prior width σ_p and co-linearity threshold θ_t . Fig. 4 shows the psychophysical data for subject YW (fraction of collinear responses versus orientation difference between predictors and target) and the fit of Eq. (8) for five different test conditions: normal sequence, random order sequence, random duration sequence, normal sequence with 0% collinear trials and normal sequence with 50% collinear trials. As explained before, we assume that the differences in responses for different conditions can be explained by differences in prior expectations associated to each type of predictor. Therefore the five fits are constrained to have the same width of likelihood σ_1 and collinearity threshold θ_t while having their own individual prior width σ_p . Clearly, the model fits well with this subject's psychophysical data. The prior width σ_p shows clear differences among test conditions of normal sequence, random order sequence and random duration sequence (Fig. 4A), and between

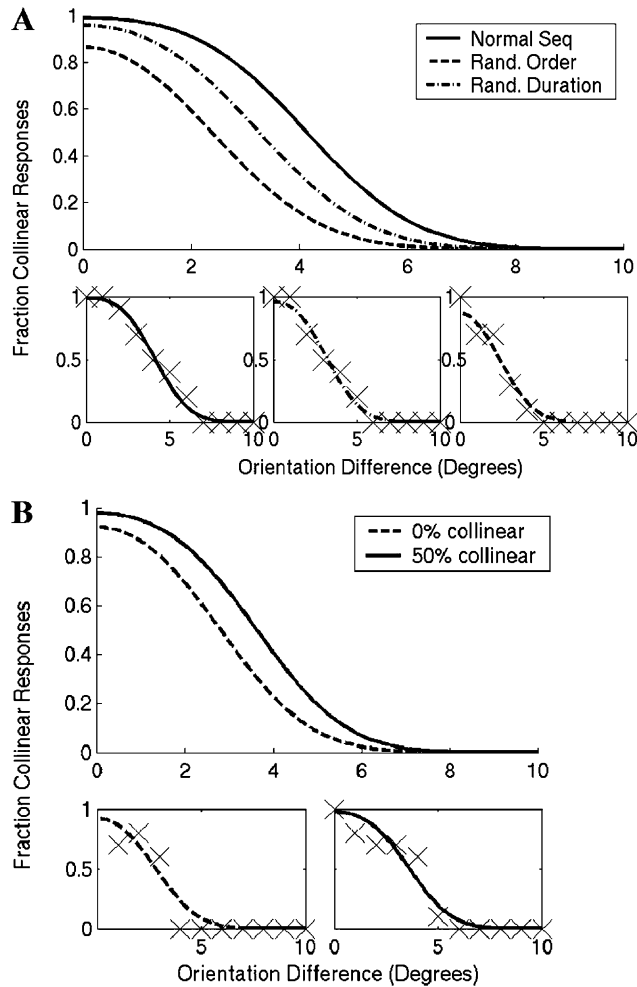


Fig. 4. Psychophysical fraction of collinear responses for different test conditions and model fitting for subject YW (normal sequence, random order sequence and random duration sequence in A; normal sequence with 0% and 50% collinear trials in B). The 'x' represents the fraction of times the subject responded collinear when presented with the orientation difference between the predictors and the target on the x -axis. The solid line is the model's least-squares fit. The width of likelihood and collinear threshold were the same for different test conditions.

test conditions of normal sequence with 0% and 50% collinear trials (Fig. 4B). The fact that the experimental points reach 100% for 0° orientation difference under some test conditions is peculiar to the subject chosen as an example. For other subjects slightly different values were found.

Similar goodness of fit was found for the other subjects. Fig. 5A presents the value of the width of prior σ_p for different test conditions, and Fig. 5B shows the values of the width of likelihood σ_l , co-linearity threshold θ_t , and distance between psychophysical data and fitted model χ^2 . This distance is shown for comparative purposes only, since the error bars are arbitrary. Grey bars represent the parameter values for individual subjects and black bar represents the average

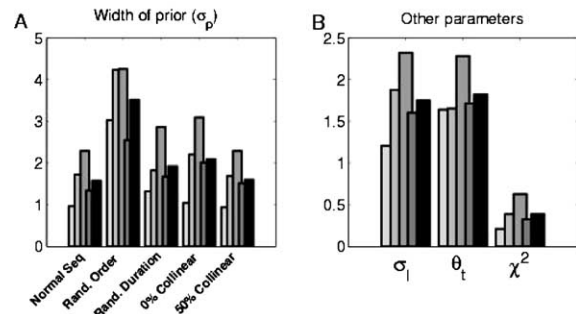


Fig. 5. (A) Width of prior distribution (σ_p) for five experimental conditions: normal sequence, random order sequence, random duration sequence, normal sequence with 0% collinear trials and normal sequence with 50% collinear trials. For each condition, the first four bars (grey bars) correspond to subjects KG, RR, XU and YW respectively, and the fifth bar (black bar) is the parameter averaged across four subjects. (B) Other parameters of the model, fitted simultaneously for all conditions: Width of likelihood σ_l , co-linearity threshold θ_t and fitting error χ^2 .

parameter value across subjects. Compared with the test conditions of random order sequence and random duration sequence, the prior distribution is the sharpest for the normal sequence (average $\sigma_p = 1.59$; Fig. 5A). It is for this condition that subjects more often perceived the stimulus as collinear. For the random duration sequence, the prior distribution takes an intermediate width (average $\sigma_p = 1.93$), and it is the broadest for the random order sequence (average $\sigma_p = 3.51$). The differences between all three conditions were significant (Paired t -test, $p < 0.05$). This is consistent with our expectation that predictor sequences which resemble natural sequences are associated with sharper priors.

Likewise, a comparison of the normal sequence with 0% collinear trials (average $\sigma_p = 2.09$) and 50% collinear trials (average $\sigma_p = 1.61$) shows that the prior distribution is sharper for higher frequency collinear trials (Fig. 5A, Paired t -test, $p < 0.05$). These prior expectations bias perception more strongly towards co-linearity. Again, this is to be expected if we think that prior expectations are influenced by sensory experience.

These results are consistent with the orientation discrimination threshold analysis (Fig. 2), in which subject's orientation perception for the target bar was more often biased towards the orientation of the predictors under the test condition of normal sequence with 50% collinear trials. Clearly, larger discrimination thresholds correspond to sharper prior distributions.

In summary, a simple Bayesian model, which assumes that collinear stimuli are more likely a priori, and that neuronal representations are subject to a certain level of noise, can account for our psychophysical data. A predictable spatio-temporal stimulus structure and an increased frequency of collinear trials were both associated with an increase in the prior expectation of collinear events, and, consequently, with the perception of

the target's orientation being more strongly biased towards that of the predictors.

4. Discussion

It is well known that the discrimination of the orientation of a line in the human fovea can be modulated by the spatial or temporal arrangement of the surrounding configurations (e.g. Wenderoth & Johnstone, 1988; Westheimer, 1990; Westheimer & Ley, 1997). For example, when two co-presented flanking lines are positioned in a collinear arrangement with the central target line, the target appears tilted towards the flanks and the subjects' orientation discrimination threshold to the target is elevated (Kapadia et al., 2000). The orientation discrimination can also be impaired if the foveal target line is followed immediately by a surround mask of various configurations (Wehrhahn et al., 1996), especially if the mask composes of lines with the same orientation as the target (Li, Thier, & Wehrhahn, 2000).

Our experiments demonstrated that human orientation perception can also be modulated by the conditional prior probability of a spatio-temporal predictor stimulus. The orientation discrimination to a foveal target bar was significantly biased towards the orientation of the predictors presented in discrete spatial and temporal sequence. The discrimination threshold was further elevated with increasing probability of the collinear trials and number of predictors. These results suggest that the human visual cortex can effectively integrate orientation signals over a range of spatial and temporal positions, and this integrated perceptual sensitivity can be modulated by spatio-temporal prior probability of stimuli, as suggested by Bayesian inference.

Two relatively simple interpretations, masking effect and spatio-temporal proximity of flanker interaction, may partly account for the differences in orientation discrimination threshold to the target measured under test conditions of normal sequence, random order sequence and random duration sequence (Fig. 2A). Results of psychophysical masking experiments revealed that the presence of adjacent high-contrast collinear flankers (masks) can elevate the contrast discrimination threshold to the center target (Adini & Sagi, 2001; Chen & Tyler, 2002; Foley, 1994; Zenger & Koch, 2001), suggesting a localised spatial and temporal interaction between the flankers and the target. Furthermore, the orientation discrimination threshold to the center target is markedly raised by the collinear flankers co-presented with the target or presented a short time before or after the target (Kapadia et al., 2000; Mareschal et al., 2001; Wehrhahn et al., 1996; Westheimer & Ley, 1997). This contextual effect can extend spatially to 1° (Mareschal et al., 2001) and is most pronounced for a temporal

onset asynchrony of around 100 ms (Westheimer & Ley, 1997). As the predictor bar used in our experiment has a length of 1.5° and a presentation duration of 200 ms, it seems that the final judgment of the target orientation depends on the last predictor presented in a position immediately adjacent and in a time immediately preceding the target (predictor 4 in Fig. 1). Comparing to normal sequence, in random order sequence and random duration sequence there was much less spatio and temporal proximity between the last predictor and the target. Consequently, the masking effect is reduced and the orientation discrimination threshold to the target is less impaired. However, the spatio-temporal proximity of flanker interaction and flanker masking effect can not fully interpret the results related to the experiments of varying the probability of the collinear trials (Fig. 2B) and varying the number of the predictors (Fig. 3A) without involving further processing in the visual system. In these two experiments, for a given trial, there was always a collinear flank presented immediately adjacent and preceding the target. The subject's orientation discrimination threshold for the target, however, was increased with increasing probability of the collinear trials and number of the predictors. Therefore, although the spatio-temporal interaction between the target and the last predictor, presented next to and immediately before the target, can contribute to the target orientation perception, the final elaboration of the target's orientation critically depend on the ability to sense spatio-temporal prior probability of stimulus or stimulus statistics. We have suggested a simple Bayesian model, based on the prior expectation of collinear events, and we show that this model can account parsimoniously for the full pattern of results we see (see further discussion in "Bayesian inference").

4.1. Possible mechanism

Although the process of visuospatial integration has been traditionally ascribed to high-order cortical visual areas, there is growing evidence suggesting that the primary visual cortex (area V1) may play an important role as well. Orientation selective neurons in area V1 provide the basis for the elaboration of the 'orientation' attribute of stimuli, and their orientation selectivity shows remarkable plasticity to complex temporal or spatial stimulus dynamics. Temporally, paired visual stimuli at different orientations can induce a shift in a neuron's orientation tuning (Dragoi, Sharma, & Sur, 2000; Muller, Metha, Krauskopf, & Lennie, 1999), the direction of shift depends on the temporal order of the pair and could be linked to the perceptual level (Yao & Dan, 2001). Spatially, the oriented stimuli co-presented outside the classical receptive field (CRF) have a powerful modulatory influence on neurons' orientation selectivity to stimuli inside the CRF (Gilbert, Das, Ito,

Kapadia, & Westheimer, 1996; Knierim & Van Essen, 1992; Levitt & Lund, 1997; Sillito, Grieve, Jones, Cudeiro, & Davis, 1995). This modulated contextual interaction could likely be the neural basis for the increased human orientation discrimination threshold for similar stimuli (Gilbert & Wiesel, 1990; Kapadia et al., 2000; Li et al., 2000).

Using similar stimuli to the ones presented to human subjects, our electrophysiological recording from monkeys showed that the orientation responses of a substantial population of V1 neurons were significantly modulated by events (predictors) prior to and distant from stimulation (target) of their CRFs (Guo et al., 2002; Pulgarin et al., 2003). Therefore, the final elaboration of the orientation attribute for the target line may depend on the confluence of signals from neurons whose CRFs have the same orientation selectivity but whose locations are arrayed along the direction of the line. Although feedback from extrastriate areas may play a role, the long-range horizontal connection in area V1 which tend to connect cells with similar orientation preferences (Gilbert & Wiesel, 1990; Lamme, Super, & Spekreijse, 1998) could be a strong candidate for the anatomical substrate. As it has the ability to provide both excitatory and inhibitory inputs to their postsynaptic neurons and thus modulate their discharges (McGuire, Gilbert, Rivlin, & Wiesel, 1991), the orientation selective signals over a range of spatial positions can be pooled effectively.

4.2. Bayesian inference

Visual scenes are predictable, since they are both spatially and temporally redundant. In a visual world in which visual stimuli, and their co-appearances, do not appear with equal probability, it would be statistically optimal for the visual system to compute the Bayesian posterior probability of a feature, based on information from the eye, and prior knowledge of the statistical structure of the visual world (Friston, 2002; Knill & Richards, 1996; Rao, 1999; Young, 2000). This Bayesian perspective has been successfully implemented to deal with the dynamical aspects of real sensory inputs, such as luminance discrimination (Ciaramitaro et al., 2001), motion perception (Ascher & Grzywacz, 2000; Hürliemann et al., 2002; Weiss et al., 2002), depth perception (Kersten et al., 1997) and object recognition (Kersten & Yuille, 2003; Liu et al., 1995; Rao, 1999).

The statistical structure of the visual world is reflected in a number of regularities commonly shared by natural scenes and images. One of these regularities is that the co-occurrence of line segments is characterised by a greater probability density for collinear line segments than for non-collinear ones (Geisler et al., 2001; Sigman et al., 2001). Furthermore, during normal vision, the visual system frequently has ‘reason to believe’ that a

particular feature is present 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 structure of the visual world (Young, 2000). These regularities of collinearity and predictable spatio-temporal dynamics of visual scenes should be known to the visual system, and they should be exploited for efficient coding by reflecting themselves in distributions of prior probability.

In this experiment, we demonstrated that the systematic changes in spatio-temporal prior probability of oriented predictors lead to systematic changes in human orientation discrimination to the foveal target bar, providing a psychometric function for the influence of spatio-temporal prior probability on perceptual sensitivity. These psychophysical observations can be predicted by a Bayesian model which assumes that collinear stimuli are more likely a priori and that neuronal representations are subject to a certain level of noise. A predictable spatio-temporal stimulus structure, the predictors presented in normal sequence which more closely resembles a natural scene, such as the trajectory of a moving object, is associated with the increasing prior expectation of collinear events (sharper width of prior, σ_p). Consequently, the subjects’ orientation perception of the target bar is more strongly biased towards that of the predictors. The variability introduced in the random duration sequence, and especially, in the random order sequence, makes the stimulus structure less natural, and therefore, it comes as no surprise that they are associated with a broader prior distribution.

The frequency of co-linearity of the priors and the target are directly manipulated in the 50% collinear trials and 0% collinear trials blocks. Interestingly, manipulating these frequencies experimentally also yields a change in the prior distribution in the expected direction. An increased probability of collinear trials is associated with a sharper prior distribution.

The two effects, change in prior distribution for different spatio-temporal structures and for different collinear frequencies, can be accounted for by the same Bayesian model. Nevertheless, it is possible that these effects are mediated by different mechanisms or brain areas. The first one probably depends on the statistics of natural images and is probably embedded in the system permanently. The relative lower visual areas in the hierarchical visual system, such as area V1 as discussed above, may provide sufficient information for the orientation perception. The second one depends on the recent visual experience and the final perception may involve some higher cortical areas.

Our model is based on three assumptions: (1) collinear targets are more likely; (2) there is noise in the brain representation of the orientation difference; and (3) the differences in responses for different conditions can be explained by differences in prior probability distribu-

tions. Other authors (e.g. Weiss et al., 2002) introduce the noise formally in the input, although they then discuss how neural noise may play a similar role. In our case given the simplicity of our stimulus it seems more natural to conceptually assign the noise to the neural representation.

In the same way, the fact that there is a distribution of responses to a fixed stimulus may arise from different sources. Weiss et al. (2002) place this variability in the decision stage. Mamassian and Landy (1998) propose that responses may follow a probability which matches the posterior probability. In our case the variability arises for the distribution of representations for each particular stimulus, and is associated to the likelihood function. This approach has the attractive feature that it does not introduce new hypotheses or free parameters.

Both the width of the likelihood and the co-linearity threshold are related to the orientation sensitivity in the absence of priors. It is reassuring to find that, in the model, both parameters have similar values.

Acknowledgements

This work is supported by Wellcome Trust, HFSP0 and EU FP5.

References

- Abbott, L. F., & Blum, K. I. (1996). Functional significance of long-term potentiation for sequence learning and prediction. *Cerebral Cortex*, *6*, 406–416.
- Adini, Y., & Sagi, D. (2001). Recurrent networks in human visual cortex: psychophysics evidence. *Journal of the Optical Society of America A*, *18*, 2228–2236.
- Ascher, D., & Grzywacz, N. M. (2000). A Bayesian model for the measurement of visual velocity. *Vision Research*, *40*, 3247–3434.
- Barlow, H. B. (1989). Unsupervised learning. *Neural Computation*, *1*, 295–311.
- Chen, C. C., & Tyler, C. W. (2002). Lateral modulation of contrast discrimination: Flanker orientation effects. *Journal of Vision*, *2*, 520–530.
- Chirumuuta, M., Clatworthy, P. L., & Tolhurst, D. J. (2003). Coding of the contrasts in natural images by visual cortex (V1) neurons: a Bayesian approach. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, *20*, 1253–1260.
- Ciaramitaro, V. M., Cameron, E. L., & Glimcher, P. W. (2001). Stimulus probability directs spatial attention: an enhancement of sensitivity in humans and monkeys. *Vision Research*, *41*, 57–75.
- Dragoi, V., Sharma, J., & Sur, M. (2000). Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron*, *28*, 287–298.
- Finney, D. J. (1986). *Probit analysis* (3rd ed.). Cambridge, UK: Cambridge University Press.
- Foley, J. M. (1994). Human luminance pattern-vision mechanisms: masking experiments require a new model. *Journal of the Optical Society of America A*, *11*, 1710–1719.
- Friston, K. (2002). Functional integration and inference in the brain. *Progress in Neurobiology*, *68*, 113–143.
- Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research*, *41*, 711–724.
- Gilbert, C. D. (1998). Adult cortical dynamics. *Physiological Reviews*, *78*, 467–485.
- Gilbert, C. D., Das, A., Ito, M., Kapadia, M., & Westheimer, G. (1996). Spatial integration and cortical dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, *93*, 615–622.
- Gilbert, C. D., & Wiesel, T. N. (1990). The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Vision Research*, *30*, 1689–1701.
- Guo, K., Robertson, R. G., Thiele, A., Panzeri, S., Mahmoodi, S., & Young, M. P. (2002). V1 neurons use spatio-temporal prior probabilities of stimuli in dynamic scenes in Bayesian inference. *Program No. 557.2. Abstract Viewer/Itinerary Planner*. Washington, DC: Society for Neuroscience, 2002. CD-ROM.
- Hürlimann, F., Kiper, D. C., & Carandini, M. (2002). Testing the Bayesian model of perceived speed. *Vision Research*, *42*, 2253–2257.
- Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (2000). Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *Journal of Neurophysiology*, *84*, 2048–2062.
- Kersten, D., Mamassian, P., & Knill, D. C. (1997). Moving cast shadows induce apparent motion in depth. *Perception*, *26*, 171–192.
- Kersten, D., & Yuille, A. (2003). Bayesian models of object perception. *Current Opinion in Neurobiology*, *13*, 150–158.
- Knierim, J. J., & Essen, D. C. (1992). Van Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology*, *67*, 961–980.
- Knill, D. C., & Richards, W. (1996). *Perception as Bayesian inference*. Cambridge University Press.
- Lamme, V. A., Super, H., & Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Current Opinion in Neurobiology*, *8*, 529–535.
- Levitt, J. B., & Lund, J. S. (1997). Contrast dependence of contextual effects in primate visual cortex. *Nature*, *387*, 73–76.
- Li, W., Thier, P., & Wehrhahn, C. (2000). Contextual influence on orientation discrimination of humans and responses of neurons in V1 of alert monkeys. *Journal of Neurophysiology*, *83*, 941–954.
- Liu, Z., Knill, D. C., & Kersten, D. (1995). Object classification for human and ideal observers. *Vision Research*, *35*, 549–568.
- Mamassian, P., & Landy, M. S. (1998). Observer biases in the 3d interpretation of line drawings. *Vision Research*, *38*, 2817–2832.
- Mareschal, I., Sceniak, M. P., & Shapley, R. M. (2001). Contextual influences on orientation discrimination: binding local and global cues. *Vision Research*, *41*, 1915–1930.
- McGuire, B. A., Gilbert, C. D., Rivlin, P. K., & Wiesel, T. N. (1991). Targets of horizontal connections in macaque primary visual cortex. *Journal of Comparative Neurology*, *305*, 370–392.
- Mehta, M. R., Quirk, M. C., & Wilson, M. A. (2000). Experience-dependent asymmetric shape of hippocampal receptive fields. *Neuron*, *25*, 707–715.
- 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.
- Pulgarin, M., Nevado, A., Guo, K., Robertson, R. G., Thiele, A., & Young, M. P. (2003). Spatio-temporal regularities beyond the classical receptive field affect the information conveyed by the responses of V1 neurons. *Program No. 910.16. Abstract Viewer/Itinerary Planner*. Washington, DC: Society for Neuroscience, 2003. CD-ROM.
- Rao, R. P. N. (1999). An optimal estimation approach to visual perception and learning. *Vision Research*, *39*, 1963–1989.
- Sharma, J., Dragoi, V., Tenenbaum, J., Miller, E. K., & Sue, M. (2003). V1 neurons signal acquisition of an internal representation of stimulus location. *Science*, *300*, 1758–1763.

- Sigman, M., Cecchi, G. A., Gilbert, C. D., & Magnasco, M. O. (2001). On a common circle: natural scenes and Gestalt rules. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 1935–1940.
- Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J., & Davis, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. *Nature*, *378*, 492–496.
- Wehrhahn, C., Li, W., & Westheimer, G. (1996). Patterns that impair discrimination of line orientation in human vision. *Perception*, *25*, 1053–1064.
- Weiss, Y., Simoncelli, E. P., & Adelson, E. (2002). Motion illusion as optimal percepts. *Nature Neuroscience*, *5*, 598–604.
- Wenderoth, P., & Johnstone, S. (1988). The differential-effects of brief exposures and surrounding contours on direct and indirect tilt illusions. *Perception*, *7*, 165–176.
- Westheimer, G. (1990). Simultaneous orientation contrast for lines in the human fovea. *Vision Research*, *30*, 1913–1921.
- Westheimer, G., & Ley, E. J. (1997). Spatial and temporal integration of signals in foveal line orientation. *Journal of Neurophysiology*, *77*, 2677–2684.
- 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. *Spatial Vision*, *13*, 137–146.
- Zenger, B., & Koch, C. (2001). Flanker effects in peripheral contrast discrimination psychophysics and modelling. *Vision Research*, *41*, 3663–3675.
- Zhang, K., Ginzburg, I., McNaughton, B. L., & Sejnowski, T. J. (1998). Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. *Journal of Neurophysiology*, *79*, 1017–1044.