Flank facilitation and contour integration: Different sites

Pi-Chun Huang a,*, Robert F. Hess a, Steven C. Dakin b

a McGill Vision Research, Department of Ophthalmology, McGill University, Montreal, Que., Canada H3A 1A1
b UCL Institute of Ophthalmology, Bath Street, London, UK

Received 10 February 2006; received in revised form 14 April 2006

Abstract

Observers’ ability to integrate features into extended contours, and to exploit the flanking structure to facilitate contrast detection (flank facilitation), exhibit a similar dependence on element spacing and orientation. Here, we investigate whether this reflects the operation of a common cortical mechanism by comparing performance for both tasks under monocular, binocular, dichoptic, and stereoscopic viewing conditions. Our results clearly implicate different cortical sites for flank-facilitated detection and contour integration; the former is a purely monocular phenomenon and must therefore occur at the earliest stages of cortical processing. In contrast, contour integration is a binocular process and occurs after the encoding of relative disparity, suggesting substantial extra-striate involvement. We conclude that the sites, and therefore the mechanisms, underlying these two seemingly related psychophysical phenomena are different. © 2006 Elsevier Ltd. All rights reserved.

Keywords: Lateral spatial facilitation; Contour integration; Paths; Disparity; Dichoptic

1. Introduction

The detectability of a spatially bandpass test-element is dependent on the properties of elements in its local spatial neighbourhood. Neighbouring elements that form a common global alignment can facilitate the detection of a test element; this is termed flank facilitation (Polat, 1999; Polat, Mizobe, Pettet, Kasamatsu, & Norcia, 1998; Polat & Sagi, 1993, 1994a, 1994b; Polat, Sagi, & Norcia, 1997; Woods, Nugent, & Peli, 2002). A typical stimulus configuration is shown in Fig. 1A where the detectability of a central test-element is measured either in the presence or absence of two high-contrast flanking elements of the same orientation. The key determinants of this facilitation are the suprathreshold contrast of the flanks, the flank-to-test distance (i.e. 3–6× spatial period of the target) and the global orientation alignment of the test and flank ensemble (Polat, 1999; Polat et al., 1998; Polat & Sagi, 1993, 1994a, 1994b; Polat et al., 1997; Woods et al., 2002). Although the response of cells in V1 has been shown to be modulated by the presence of spatially aligned stimuli falling outside the classical receptive field (Kasamatsu, Polat, Pettet, & Norcia, 2001; Mizobe, Polat, Pettet, & Kasamatsu, 2001; Polat et al., 1998), it is yet unresolved what mechanism underlies lateral spatial interactions near threshold. One possibility is that detection is mediated by neurons with elongated receptive fields, a notion supported by the finding that the flank facilitation is phase- and contrast-dependent (Solomon, Watson, & Morgan, 1999; Williams & Hess, 1998; Woods et al., 2002). Another explanation involves multiple neurons at different locations in the visual field interacting via the long-range lateral connections known to exist between V1 cells of similar orientation preference (Hirsch & Gilbert, 1991; Ts’o, Gilbert, & Wiesel, 1986; Weliky, Kandler, Fitzpatrick, & Katz, 1995) or by way of feedback from extra-striate sites (Gilbert & Wiesel, 1989; Girard, Hupe, & Bullier, 2001). Contour integration involves the linking of the responses of a number of cells with different orientation preferences to represent a more global feature. Following Field, Hayes, and Hess (1993), there is evidence that the outputs of cells tuned to different orientations are linked or...
facilitated when those orientations are consistent with the presence of simple first order curves. The notion of the “Association Field” (Field et al., 1993) is that this linkage is achieved by excitatory connections along the curve, and inhibitory connections to orientations/positions that are inconsistent with the local contour structure (Hess & Field, 1999; Kovacs, 1996). The typical stimulus configuration for contour integration experiments is illustrated in Fig. 1B where a subset of contour-consistent Gabor, embedded in a field of randomly oriented and positioned distracter elements, are perceived as a distinct perceptual whole. There is debate about the site of such contour integration in the human visual hierarchy. Running along similar lines to the proposed mechanism for flank facilitation described above, one school of thought has proposed that contour integration arises in V1 via either the long-range connections between cells with similar orientation preference (Hirsch & Gilbert, 1991; Ts’o et al., 1986; Weliky et al., 1995) or as the result of extra-striate feedback (Gilbert & Wiesel, 1989; Girard et al., 2001). The opposing view is that, in an analogous way to global motion processing (Movshon, Adelson, Gizzi, & Newsome, 1985; Newsome & Pare, 1988; Salzman, Murasugi, Britten, & Newsome, 1992), the site of contour integration lies in extra-striate cortex (Kiorpes & Bassin, 2003). More specific proposals have also been made for the involvement of cortical area V2, which shows sensitivity for angles (Ito & Komatsu, 2004). Finally, there are arguments both for (Polat, 1999; Polat & Bonneh, 2000) and against (Ito & Komatsu, 2004; Williams & Hess, 1998) the proposition that flank facilitation and contour integration are manifestations of the same underlying cortical process. Certainly, the stimulus conditions are almost identical; for example, consider the triplets along a typical contour that are highlighted in Fig. 1C. In qualitatively similar ways, both effects depend on element orientation and spacing (Field et al., 1993; Polat & Sagi, 1993), that is to say locally aligned orientations and close spacing are optimal. Furthermore, both depend on the local element spatial frequency in similar ways (Dakin & Hess, 1998). On the other hand, there are some possibly important differences between the two effects. For example, one involves an intrinsic contrast metric (i.e. flank facilitation is defined solely in terms of contrast sensitivity) whereas the other does not (Hess, Dakin, & Field, 1998) and also they have different phase dependencies (Solomon et al., 1999; Williams & Hess, 1998); contrast facilitation does not operate for adjacent elements that are 180° out of phase spatially whereas contour integration depends very little in the spatial phase of adjacent elements (Field, Hayes, & Hess, 2000). Furthermore, the notion that both phenomena are mediated by long-range intrinsic connections in V1 is challenged by the proposal that such connections have a fixed size (Hirsch & Gilbert, 1991; Ts’o et al., 1986; Weliky et al., 1995) while both flank facilitation and contour integration exhibit scale-invariance (Field et al., 2004).

Fig. 1. Typical stimulus arrangement for flank facilitation; the global orientation of the two outer supra-threshold flanks facilitate detection of the central, low-contrast test stimulus. (B) Generic stimulus arrangement for contour integration, in which a subset of elements have a common alignment to form an integrated contour. (C) Illustration of the hypothetical formation of elongated collinear fields, through pair-wise interactions, assuming that flank facilitation and contour integration share a common mechanism. (D) Production of stereo images: for each stereo pair, the central Gabor target was fixed-phase but the flanks were horizontally shifted (patch shift) by an amount that produced a 180° phase difference between flanks and target in each stereo pair. This results in the target being presented in the fixation plane and the flanks being both presented either in front or behind of the fixation plane. (E and F) Show a dichoptic image-pair for the contour integration task.
1993; Polat & Sagi, 1993). More fundamentally, it is not clear that the extent of these connections is sufficient to account for the large inter-element distances over which these effects can occur (Kiorpes & Bassin, 2003).

We sought to shed some light on the neuronal basis of these two effects by establishing whether they occur before or after binocular integration, thought to be in V1 (Hubel & Wiesel, 1977) and before or after relative disparity processing, believed to be in V2 (Parker & Cumming, 2001; Thomas, Cumming, & Parker, 2002). The results suggest that flank facilitation and contour integration occur at different sites along the pathway, the former at a monocular cortical site (e.g. layer 4C of V1) whereas the latter occurs after relative disparity has been encoded (e.g. V2).

2. Experiment 1: Flank facilitation

2.1. Method

2.1.1. Apparatus

The stimuli were presented on an Electrohome (Retro III) back-projection CRT monitor (138 × 104 cm). The projector was controlled by a VSG2/5 graphic card (Cambridge Research Systems), which had 15 bits contrast resolution. The projector was gamma corrected. The screen resolution was 1024 × 768 pixels with a frame rate of 120 Hz and mean luminance of 67 cd/m².

2.1.2. Subjects

Three experienced psychophysical observers (RH, PH, and DE) participated in Experiment 1 and Experiment 2. All had normal or corrected-to-normal vision.

2.1.3. Stimuli

The target and flankers were Gabor patches that were defined by the following equation:

\[ L(x, y, \theta) = L_0 + L_0 C \cos(2\pi f (x \sin \theta + y \cos \theta) + \rho) \times \exp[-(x^2 + y^2)/2\sigma^2] \]  

Where \( L_0 \) was mean luminance, \( C \) the contrast of the Gabor, \( f \) the spatial frequency of the carrier, \( \sigma \) the standard deviation of the Gaussian envelope, \( \theta \) the element orientation in degrees and \( \rho \) the phase of the carriers with respect to the center of a Gaussian window. The spatial frequency used was 0.75 cycles/deg, and space constant (\( \sigma \)) was 0.53° at a viewing distance of 208 cm. Consequently, the bandwidth (full-width at half-height) of the Gabor was one octave. The Michelson contrast of the flanks was set to 0.4 and the absolute phase of the target and flanks varied randomly between each trial, but the relative phase of the target and flanks were always the same. The center-flank separation was set to 3\( \lambda \) and the target was presented to the fovea.

Four viewing conditions were employed: monocular, binocular, dichoptic, and stereoscopic. To present different stimuli to the two eyes, we used a frame interlacing technique in conjunction with liquid crystal goggles (Cambridge Research Systems, FE-1 goggles) synchronized to the monitor’s frame rate. To ensure that there was sufficient contrast resolution for the low contrast target presentation while high contrast flanks were presented, the total contrast resolution (8 bits) was divided into two parts, giving 64 steps of resolution for the flanks and 192 steps of resolution for the target. Each image pair contained a black square frame (18° × 18°, and the width of the frame is 2 pixels, equivalent to 0.07°) surrounding the stimuli to aid peripheral fusion.

2.1.4. Procedure

For the flank facilitation conditions, a temporal, two-alternative forced-choice (2AFC) paradigm, without feedback, was used to measure the target contrast-detection threshold. Subjects were required to choose which of two intervals contained the target. Subjects were pre-cued to trial onset by an audible tone, and by a fixation point, present for 250 ms. This was followed by a 1000 ms (temporal square pulse) stimulus presentation, then a 1000 ms inter-stimuli interval (ISI, consisting of homogeneous mean-luminance field). A second similar (pre-cued) stimulus interval followed. A black square frame around the stimulus was present throughout the experiments. The contrast detection of a central Gabor target was measured under six conditions: monocular viewing of either (1) isolated or (2) flanked target-Gabor, (3) dichoptic viewing of flanked target, binocular viewing of (4) isolated or (5) flanked target, (6) stereoscopic viewing conditions. In each run, either conditions 1–3 or conditions 4–6 were blocked together. In subsequent control experiments we provided nonius markers just prior to stimulus presentation to ensure correct ocular alignment and replicated the dichoptic results reported here without nonius markers.

The monocular presentations involved presenting the stimuli to one eye and a uniform field of the same mean luminance to the other eye. The dichoptic presentations involved presenting the central Gabor target to one eye and the two outer flanks to the other eye. The binocular presentations involved showing the identical stimuli to both eyes. The stereo image pairs were produced by horizontally shifting the positions of the flanks so that the flanks were 180° out of phase with the central Gabor target (see Fig. 1D) producing a disparity of 1.3°. Half the presentations involved flanks of crossed disparity relative to the central Gabor target in the fixation plane, and half of the presentations involved flanks of uncrossed disparity. Within any run, the sign of the flank’s disparity was randomly assigned. This arrangement was crucial in that it ensured facilitation could not occur from a purely monocular process because it has been shown that facilitation over these inter-element distances depends on local phase alignment (Williams & Hess, 1998). As a result, the target was presented in the fixation plane and the flanks were randomly presented in front or
behind of the fixation plane. A demonstration is shown in Fig. 1D. The method of constant stimuli was used to determine the contrast threshold. Each data point had at least 50 trials with 5–7 contrast levels. A bootstrapping procedure (Wichmann & Hill, 2001a, 2001b) was used to establish the confidence interval associated with the Weibull fit, which was defined as \( 0.5 + 0.5 \times (1 - \exp(-x/\beta)) \) with \( x, \beta \) free parameters. The detection threshold was determined at 75% correct level.

2.1.5. Results

The results of Experiment 1 are shown in Fig. 2. In Fig. 2A, the thresholds for each condition are shown with error bars indicating ±1 standard deviation (SD) of the threshold estimate (derived from the bootstrapping procedure). In Fig. 2B, the threshold elevation −1 was plotted for each condition to compare the relative amounts of facilitation. The threshold elevation was defined as detection threshold with flanks divided by the detection threshold without flanks. For the threshold elevation value in the monocular and dichoptical conditions, the monocular viewing of a single Gabor patch was used as the baseline condition. For the binocular and stereoscopic conditions, the threshold obtained under binocular viewing of a single Gabor patch was used to derive the threshold elevation value. Consequently, a value below zero indicates facilitation, whereas a value above zero indicates suppression. The error bars show ±1 SD of the elevation estimate. For three subjects, significant threshold facilitation occurs for the single Gabor patch in the presence of the flanks when viewed under binocular and monocular conditions, but not in stereo and dichoptical viewing conditions (at 95% confidence interval). For the statistical analyses, the one sample t-test for each viewing condition across all subjects was used and it showed that the magnitude of facilitation was significant in the monocular and binocular viewing conditions (monocular: \( t_{(23)} = -9.070, p < 0.006 \); binocular: \( t_{(23)} = -4.170, p < 0.026 \) ) but no significant facilitation was observed in either the stereo or dichoptical viewing condition (Stereo, \( t_{(23)} = 0.320, p = 0.390 \); dichoptical, \( t_{(23)} = 0.548, p = 0.319 \)).

We do not believe that the lack of facilitation found in the dichoptical and stereo viewing conditions was because of binocular rivalry because (1) the nature of the dichoptical stimulation (i.e. iso-orientation, same spatial frequency, and different field positions) was not one that would be expected to produce rivalry (Blake, 1977) and (2) none of our subjects reported a rivalrous percept. However, to directly assess whether rivalry between the two eyes images in the dichoptical condition was responsible for the lack of facilitation, we ran a condition on one subject, PCH, with a presentation time (100 ms), short enough to bypass any rivalry (Wolf, 1986). We obtained the same pattern of results (i.e. no facilitation under dichoptical conditions) showing that rivalry could not explain the lack of facilitation we observed (see Fig. 2A, right column).

3. Experiment 2: Contour integration

3.1. Methods

3.1.1. Apparatus

Stimuli were displayed on a Sony Trinitron monitor driven by a VSG 2/4 graphics board (Cambridge Research Systems) with 15 bits contrast resolution, housed in a Pentium PC computer. The frame rate of the display was 100 Hz. The monitor was gamma corrected by software with lookup tables using luminance measurements obtained from a United Detector Technology Optometer (UDT S370) fitted with a 265-photometric sensor. The monitor was viewed in a dimly lit room. The mean luminance of the display was 14.2 cd/m².

3.1.2. Stimuli

The stimuli were square patches (21° × 21°) of pseudo-randomly distributed Gabor elements (Fig. 1B). The elements were defined by Eq. (1). The sinusoidal frequency \( f \) is 0.75 cycles/deg, and the space constant (\( \sigma \)) is 0.58° at viewing distance 60 cm. The contrast of the element was set to 0.5 Michelson contrast. The separation between the elements was \( 3\lambda \) (from center to center).

The observer’s task was to discriminate a “path” from a “no-path” stimulus. No-path stimuli were composed of randomly oriented Gabor elements which were placed within a 14 × 14 grid of equally sized cells. Path stimuli were similar except the orientations and positions of 10 (randomly selected but neighboring) elements in the grid (path elements) were assigned to values consistent with the presence of a continuous contour passing through their corresponding cells. The curvature of this path is controlled by a parameter, \( x \), which determines the angular difference between adjacent elements. To avoid the occurrence of straight paths when \( x \) was 0°, an orientation jitter uniformly distributed between ±10° was added to the path curvature. Finally, to avoid random closure of the paths with high curvature, which can affect detection (Elder & Zucker, 1993; Kovacs & Julesz, 1993), paths which looped back on themselves were discarded and regenerated. We further ensured that at least one path element passed through the central region of the stimulus (defined as a circular region 3° in diameter).

We compared four different viewing conditions by using a frame interlacing technique in conjunction with liquid crystal goggles synchronized to the frame rate. In the monocular viewing condition, the stimulus was presented to one eye and the other eye saw a homogeneous field of the same mean luminance. In the binocular viewing condition, the same stimulus pair was presented to each eye. In the dichoptical viewing condition, alternate path elements were shown to alternate eyes and half of the background elements were selected at random and presented to each eye (see Figs. 1E and F for illustration). For what we refer to as the “half-dichoptical” viewing condition, one eye was covered by an eye patch.
whilst viewing the dichoptic stimulus. This control condition was to assess the extent to which any purely monocular input contributed to contour detection in the dichoptic conditions.

3.1.3. Procedures

A temporal 2AFC paradigm was used to measure the subject’s ability to detect which of two stimulus presentations (background elements + path vs. background...
back was given although performance showed little dependence on it. A small fixation mark appeared in the center of the display during the whole session. Fusion surrounds were used to aid peripheral fusion in the dichoptic viewing condition and the reliability of eye alignment was assessed using dichoptic vernier marks presented just prior to stimulus presentation.

3.1.4. Results

The results from Experiment 2 for three subjects are graphed in Fig. 3. The horizontal axis indicates path-curvature and the vertical axis, the percent correct performance. The error bars show ±1 SD. Note that all subjects exhibit similar performances under conditions of monocular or binocular viewing. Furthermore, although performance declines in the dichoptic viewing condition, it remains significantly above chance across a range of path angles and is consistently superior to the performance we found for the half-dichoptic condition (where half of the stimulus elements were displayed to only one of the eyes). This comparison between dichoptic and half-dichoptic viewing clearly shows that probability summation between purely monocular processes cannot explain the dichoptic performance. If performance in the half dichoptic case is at chance, as it is for path angles of 10°, then any argument based on probability summation to explain the above chance performance in the dichoptic case fails. Thus contour integration of paths can be performed dichoptically and that this cannot be due to the independent activation of purely monocular processes.

4. General discussion

These results suggest that the sites of the mechanisms responsible for flank facilitation and contour integration are different. Flank facilitation is a monocular process; we found no evidence for flank facilitation for dichoptic or stereoscopic presentations. This is a surprising result because the dependence of flank facilitation on flank orientation suggests a cortical origin where the majority of cells are binocular (Hubel & Wiesel, 1977). However, recent neurophysiology suggests that there is a diversity of orientation tuning in all cortical layers (Ringach, Shapley, & Hawkin, 2002) and, on the basis of our finding, we suggest that the site of flank facilitation must be either at the earliest level of cortical processing where the majority of cells are monocular or be restricted to just monocular cells at later cortical levels. In contrast, our subjects could integrate contours composed on dichoptically-presented constituent elements, suggesting that its site involved binocularly-tuned cells. Furthermore, previous results have demonstrated that contours that oscillate between two different depth planes can be detected as well as contours that are restricted to a single plane (Hess & Field, 1995). This suggests that all or some of the mechanisms responsible for contour integration of the type examined here occur after the site of relative disparity encoding, presently

Fig. 3. Results for the contour integration experiment. Percent correct identification performance is plotted against path curvature. Four different viewing conditions are compared: monocular, binocular, dichoptic, and half-dichoptic.

elements alone) contained the path. Each trial consisted of this pair of stimuli presented sequentially for the same duration, namely 1000 ms. Presentations were abrupt with a 500 ms ISI, consisting of a blank screen with the same mean luminance as the stimuli. In each run, the path angle $\alpha$ was set to 0°, 10°, 20°, 30°, 40°, and 50°. After each trial the subject indicated which interval contained the path. Subjects performed 50 trials per session for each experiment with 1–4 sessions for each curvature condition. Feed-
thought to be V2 (Parker & Cumming, 2001; Thomas et al., 2002).

An alternate possibility is that flank facilitation is the result of spatial and temporal uncertainty effects, as mentioned by Williams and Hess (1998) and later investigated by Petrov, Verghese, and McKee (2006). Imagine, for some reason, in our dichoptic and stereo conditions, the normal reduction in uncertainty provided by the flanks was disrupted. In these conditions, the flanks were always perceived in alignment with the central test gabor (i.e. we ensured there was correct ocular alignment, using peripheral fusion markers and in subsequent control experiments, nonius lines), so we can rule out this obvious possibility. Furthermore, we did not observe any steepening in the psychometric slopes between the binocular condition and either the dichoptic or stereo condition. This led us to the conclusion that a reduction in uncertainty was not a suitable explanation for the lack of facilitation in the dichoptic and stereo conditions. We did however observe that psychometric functions were generally shallower for binocular viewing when flanks were present (Petrov et al., 2006).

Another possibility that had been raised is that flank facilitation involves the same underlying mechanism as superimposed facilitation in a typical contrast increment task (Solomon et al., 1999; Williams & Hess, 1998). However it is not at all resolved whether there is or is not facilitation for superimposed stimuli in the dichoptic case. Blake and Levinson (1977); Levi, Harwerth, and Smith (1980); Meese, Georgeson, and Baker (2005) suggest there is, whereas Legge (1979) shows an absence of superimposed facilitation in the dichoptic case. More recent results from Meese et al. (2005) suggest a reduction (i.e. 4 dB dichoptically versus 8 dB binocularly) rather than absence. It is too early to say whether the absence of flank facilitation in the dichoptic condition reported here is a consequence of the reduced facilitation in the superimposed condition.

The finding that contour integration can occur dichoptically but is reduced in sensitivity may suggest multiple sites, specifically, an early site that is purely binocular (i.e. V1) and a later site after the processing of relative disparity (i.e. after V2). This would seem to accord with recent findings that V2 neurons encode relative orientation, a necessary building block in the process of extracting and representing extended contour structure (Ito & Komatsu, 2004). The present result does not however allow us to rule out the possibility that some part of the contour integration process is performed at a purely monocular site as well.

The fact that these two seemingly related psychophysical phenomena have different sites adds further weight to the suggestion that they have different underlying mechanisms and make different contributions to perception. This is not unexpected (though see Polat, 1999; Polat & Bonneh, 2000) because the two phenomena also have different dependencies on the phase (Solomon et al., 1999; Williams & Hess, 1998) and the contrast (Hess et al., 1998) of their components: flank facilitation is a purely threshold phenomenon involving contrast coding, whereas contour integration is a suprathreshold phenomenon that uses a code other than contrast (e.g. target salience).

That our results suggest contour integration occurs at a later stage in the visual pathway (after processing of relative disparity) is in good agreement with recent functional MRI data which have shown that extra-striate brain areas play a dominant role in detecting sparse contours in natural scenes (Dumoulin, Dakin, & Hess, 2004). This does not exclude a significant V1 contribution and indeed the reduced performance for dichoptic stimuli may well reflect this.

Acknowledgments

This work was supported by a CIHR (3108-18) Grant to R.F.H. and a BBSRC Grant to S.C.D. The authors would like to thank the anonymous reviewers’ suggestions.

References


