Crowding alters the spatial distribution of attention modulation in human primary visual cortex

Fang Fang

Department of Psychology and Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, PR China

Sheng He

Department of Psychology, University of Minnesota, Minneapolis MN, USA

Crowding effect is the visibility reduction of a target when presented with neighboring distractors. It has been explained by either lateral inhibition at a pre-attentive level or coarse spatial resolution of attention. To test these theories, high-resolution fMRI was used to measure V1 response to the target in the presence or the absence of the distractors in both attended and unattended conditions. We found the cortical response to the target was not affected by the presence of distractors in the unattended condition. However, the spatial distribution of attention modulation in the target and its surrounding area depended on the crowding configuration. When distractors were placed in the same radial axis as the target, a configuration with a severe crowding effect, significant attention enhancements were observed not only in the target's and the distractors' locations, but also in regions next to the target where even no stimulus was presented. But this spread of attention enhancement did not occur when distractors were placed in the same circumference as the target, a configuration with a weak crowding effect. The pattern of interaction between attention and target-distractor configuration supports that crowding results from coarse spatial resolution of attention.

Introduction

When a target is presented with neighboring distractors, its visibility is reduced (Figure 1A). This phenomenon is known as crowding in psychophysics (Bouma, 1970; Strasburger, Harvey, & Rentschler, 1991) but is referred to as surround suppression in neurophysiology (Cavanaugh, Bair, & Movshon, 2002; Levitt & Lund, 1997). The crowding effect occurs under a wide range of conditions and tasks, including Vernier acuity (Levi, Klein, & Aitsebaomo, 1985; Williams & Essock, 1986), stereoacuity (Westheimer & Truong, 1988), orientation discrimination (Westheimer, Shimamura, & McKee, 1976), contrast discrimination (Wilkinson, Wilson, & Ellernberg, 1997), letter recognition (Pelli, Palomares, & Majaj, 2004), and face recognition (Louie, Bressler, & Whitney, 2007; Martelli, Majaj, & Pelli, 2005). Two competing theories propose different explanations for the crowding effect. One emphasizes that it is due to sensory level lateral inhibition (Banks, Larson, & Prinzmetal, 1979; Levi, Klein, & Yap, 1987) and occurs at a pre-attentive level (Pelli et al., 2004). The other proposes that attention is a key factor (Walley & Weiden, 1973; Wolford & Chambers, 1983) and that crowding could be attributed to coarse spatial resolution of attention (He, Cavanagh, & Intriligator, 1996, 1997; Intriligator & Cavanagh,

2001) or unfocussed spatial attention (Strasburger, 2005; Strasburger et al., 1991). Crowding effect has been an active research topic in visual psychophysics for more than four decades (Bouma, 1970; Stuart & Burian, 1962). However, few neuroimaging studies have been conducted to explore the underlying neural mechanism of the crowding effect and test these theories. The goal of the current study is to investigate two issues, using high-resolution fMRI: (1) How is the cortical response to the target affected by the presentation of distractors when attention is not directed to the target (unattended condition)? (2) How does attention modulate the cortical responses in crowded and non-crowded conditions at the target's and the distractors' locations? With answers to these two questions, we would be at a better position to address some long-standing questions about the crowding effect.

Methods

Participants

A total seven healthy subjects (3 male and 4 female) participated in the experiments, all of whom had extensive experience as subjects in psychophysical and fMRI
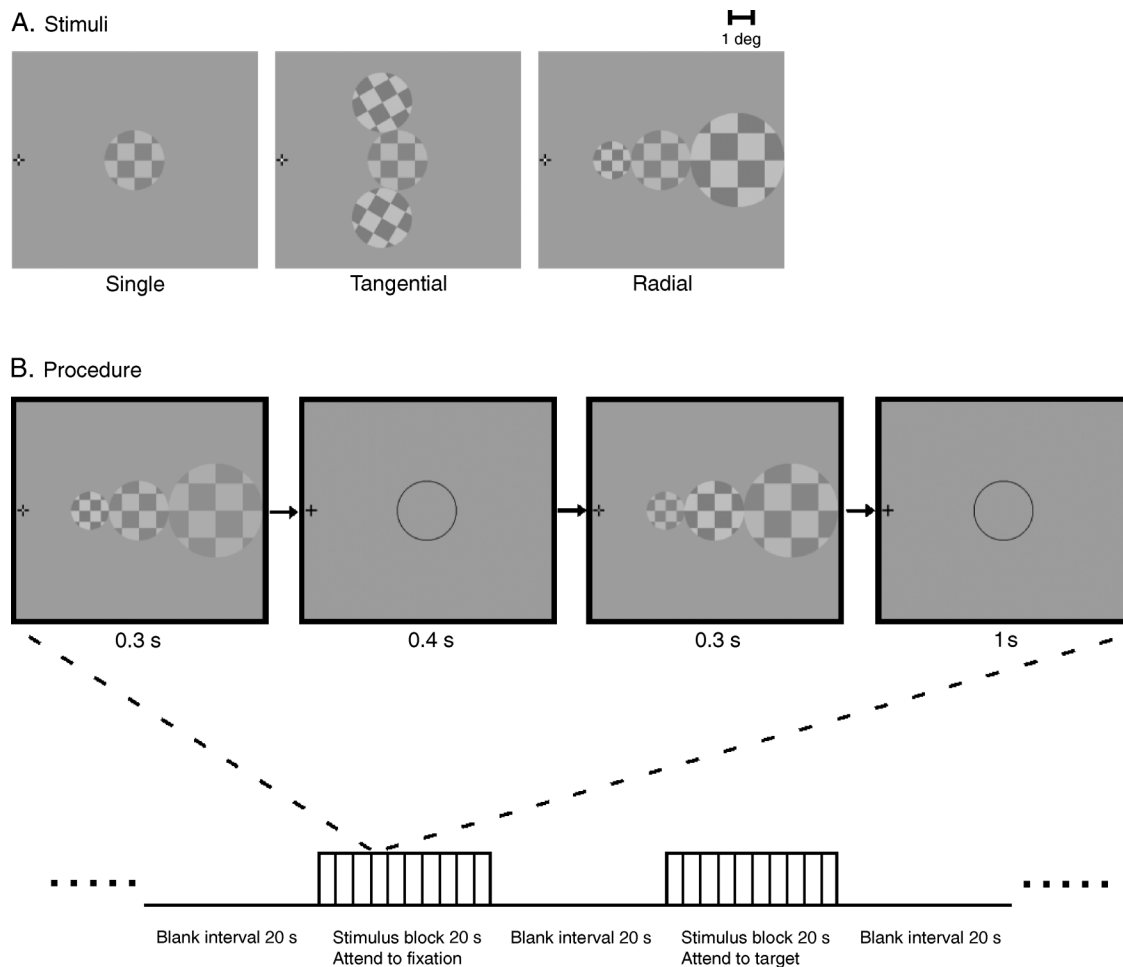
Figure 1. Example stimuli and experimental design. (A) Example stimuli used in the experiment. A target was always positioned on the right horizontal meridian. It was presented either alone or with two distractors positioned either tangentially (above and underneath the target) or radially (left and right of the target). (B) Schematic description of the experiment. Stimulus blocks were interleaved with blank intervals. A stimulus block consisted of ten trials, in which subjects were asked to perform either a luminance discrimination task at the fixation point (attend-to-fixation condition) or a contrast discrimination task to the target (attend-to-target condition).

experiments. They were right-handed, reported normal or corrected-to-normal vision, and had no known neurological or visual disorders. Ages ranged from 23 to 34. They gave written, informed consent in accordance with the procedures and protocols approved by the human subjects review committee of the University of Minnesota.

Stimuli and design

The stimulus elements used in the main experiment were round checkered patches with a mean luminance of 120 cd/m^2 and were centered at 2.88°, 5°, and 8.33° eccentricity in the right visual field. Size and spatial frequency were scaled for cortical magnification (Duncan & Boynton, 2003), with patches subtending 1.65°, 2.59°, and 4.06° of visual angle. There were three stimulus configurations (Figure 1A): single, tangential, and radial. In the single configuration, only a 2.59° size patch was

presented at 5° eccentricity, which was the target in this experiment. In the tangential configuration, two 2.59° size patches were presented immediately above and underneath the target at the same eccentricity (along the circumference). In the radial configuration, a 1.65° and 4.06° size patches were presented immediately to the left and right of the target at 2.88° and 8.33° eccentricity, respectively. These patches were the distractors in this experiment.

There were six 260-s functional scans in the main experiment. Each scan consisted of seven 20-s blank intervals and six 20-s stimulus blocks that were interleaved with each other (Figure 1B). A stimulus block contained ten trials. In each trial, two successive stimuli (single, tangential, or radial configuration) were presented for 0.3 s respectively, with a 0.4-s blank interval between them followed by a 1-s blank interval as response period. From the first stimulus to the second one, the contrasts of all the patches increased or decreased randomly and

independently by 0.08, but under the constraint that their values were between 0.2 and 0.4. The phases of the patches were also independently counterphase randomized. Another difference between the two stimuli in a trial was the luminance difference in a tiny region ($0.12° \times 0.12°$) at the center of the fixation cross. During a stimulus block, subjects were asked to fixate the cross and perform either a luminance discrimination task at the center of the fixation cross (attend-to-fixation condition) or a contrast discrimination task to the target (attend-to-target condition), depending on a task cue presented throughout the preceding 20 s blank interval. The task cue was a slight length increase of either the vertical bar (attend-to-fixation) or the horizontal bar (attend-to-target) of the fixation cross. To help the subjects localize the position of the target, a thin black circle with the same diameter as the target was presented at the target's position during all the blank intervals. This kind of position cue has been shown to have little effect on the crowding effect in the periphery (Strasburger, 2005; Wilkinson et al., 1997). The main experiment had total six combinations of attention conditions (attend-to-fixation and attend-to-target) and stimulus configurations (single, tangential, and radial), which were distributed in a functional scan (one combination per stimulus block) and were counter-balanced within subjects. The stimuli in the attend-to-target condition and the attend-to-fixation conditions were the same.

Retinotopic visual areas were defined by a standard method developed by Engel, Glover, and Wandell (1997) and Sereno et al. (1995). Five 10-Hz counterphase-flickering patches (Figure 2A) were used in a block-design scan to localize five regions of interest (ROI) (central, upper, lower, left, and right) corresponding to the locations of target and the distractors in the main experiment. Except in that they were at full contrast, these five patches for defining the ROIs were the same as those in

the main experiment. The ROI scan consisted of five cycles, and each cycle consisted of six 10-s blocks for presenting five patches and a blank interval. This scan started with a 10-s blank interval. Before the subjects were scanned, they were given a 30-min training session in the psychophysics lab for practicing with the stimuli used in the main experiment.

MRI data acquisition

In the scanner, the stimuli were back-projected via a video projector (60 Hz) onto a translucent screen placed inside the scanner bore. Subjects viewed the stimuli through a mirror located above their eyes. The viewing distance was 92 cm. Functional MRI data were collected using a 3T Siemens Trio scanner with an eight-channel phase-array coil. BOLD signals were measured with a GRAPPA sequence (TE: 35 ms, TR: 2000 ms, FOV: 22×22 cm^2, matrix: 128×128, flip angle: 75, slice thickness: 2 mm, gap: 0 mm, number of slices: 26, slice orientation: axial). fMRI slices covered the occipital lobe. A high-resolution 3D structural data set (3D MPRAGE; $1 \times 1 \times 1$ mm^3 resolution) was collected in the same session before the functional runs. All seven subjects participated in two fMRI sessions for the retinotopic mapping experiment and the main experiment respectively.

MRI data processing and analysis

The anatomical volume for each subject in the retinotopic mapping session was transformed into the AC-PC space and then inflated using BrainVoyager 2000. Functional volumes in all the sessions for each subject were preprocessed which included 3D motion correction using SPM99, linear trend removal, and high-pass (0.015 Hz)
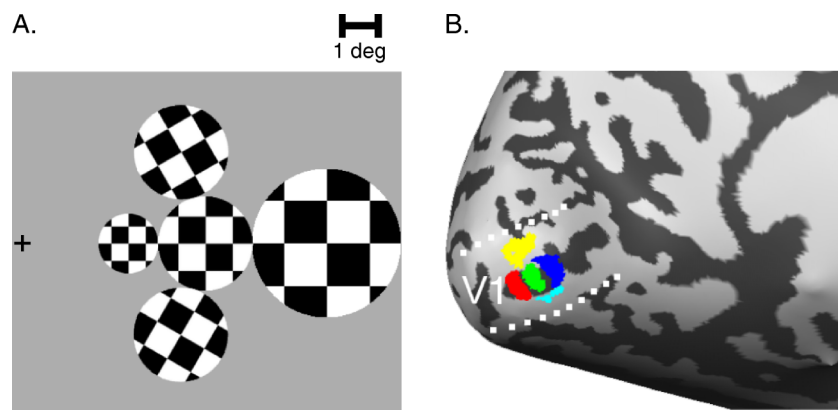

Figure 2. Regions of interest. (A) Five flickering round checkered patches with a full contrast were used to define the ROIs (central, upper, lower, left and right). They occupied the same spatial extents as the target and the distractors. (B) Cortical activations by the five patches are depicted in a representative inflated brain. The red, green, blue, yellow, and light blue areas correspond to the left, central, right, lower, and upper ROIs, respectively. V1 is defined by retinotopic mapping and its boundaries are indicated by the white dashed lines.

(Smith et al., 1999) filtering using BrainVoyager 2000. Head motion within any fMRI session was less than 1.5 mm for all subjects except two female subjects. The data from these two subjects were excluded from further analysis. The images were then aligned to the anatomical volume in the retinotopic mapping session and transformed into the AC-PC space. The first 14 s of BOLD signals was discarded to minimize transient magnetic-saturation effects.

A GLM (general linear model) procedure was used for ROI analysis. The five ROIs in V1 were defined as areas that responded more strongly to the corresponding flickering round patch than blank interval ($p < 10^{-4}$, corrected) and confined by the V1/V2 boundaries defined by the retinotopic mapping experiment. Even with such a high statistical threshold, there were still a few voxels defined to belong to more than one ROI in some subjects. We excluded these voxels from further analyses, which meant that all the ROIs were spatially non-overlapping.

The BOLD signals induced by the stimulus blocks were calculated separately for each ROI and each subject. For each fMRI scan, the time course of MR signal intensity was first extracted by averaging the data across all the voxels within the pre-defined ROI and then normalized by the average of the last two time points of all 20-s blank intervals. The peak response in an ROI was extracted by averaging the response within a 6- to 20-s interval after the start of the stimulus block and then averaged according to different experimental conditions. Attention modulation was defined as the BOLD signal difference between the attend-to-target condition and the attend-to-fixation condition. Paired t-tests were used to evaluate BOLD signal differences between experimental conditions for each of the ROIs.

Eye movement recording

Eye movements were recorded at 60 Hz with a long distance optics module of ASL eye tracker (Applied Science Laboratories, Bedford, Massachusetts) in the 3T magnet during the experiment for four subjects and with an iView X RED eye tracker (SensoMotoric Instruments GmbH, Teltow, Germany) in a psychophysics lab for the other three subjects when they viewed the same stimuli as those in the magnet.

Results

The behavioral result (Figure 3A) shows that adding distractors induced a crowding effect and impaired the contrast discrimination to the target, especially in the radial configuration ($t = 17.962$, $p < 0.001$), although the difference between the single and tangential configurations did
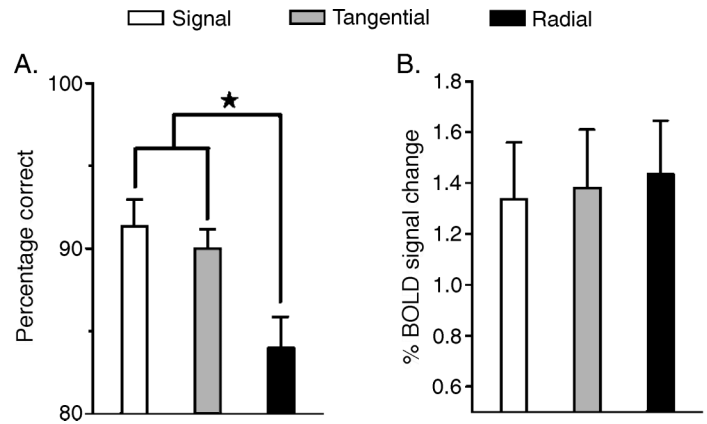

Figure 3. Behavioral and cortical responses to the target in the single, tangential and radial configurations. (A) Performance in the contrast discrimination task. (B) BOLD responses to the target with the luminance discrimination task at the fixation point. Error bars denote 1 *SEM* calculated across subjects.

not reach a significant level. A hallmark of the crowding effect is its anisotropy in the radial or tangential direction. There is much less interference between items when they are arranged tangentially, rather than radially, relative to the eyes' fixation (Toet & Levi, 1992).

For the fMRI data, we first looked at the BOLD signals at the location of the target when subjects directed their attention away from the target and performed a demanding central fixation task (accuracy mean ± *SEM*: 0.83 ± 0.03). There was no significant difference between three stimulus configurations (single, tangential, and radial) in this unattended condition (Figure 3B).

We further investigated the attention modulation in different stimulus configurations. Figure 4 shows the BOLD signals in all ROIs in three stimulus configurations when subjects attended to either the target or the fixation. Attention modulation was defined as the BOLD signal difference between the attend-to-fixation condition and the attend-to-target condition. At the target location, we found that there was significant attention enhancement in all three stimulus configurations (all $t > 3.738$ and $p < 0.02$) but no significant difference between them (Figure 5A). However, we found significant alterations in the spatial distributions of attention modulations between the stimulus configurations. In other words, significant difference in attention modulation was found at the locations around the target. For the single configuration, the BOLD signals were enhanced by attention strongly in the left ROI ($t = 4.347$, $p = 0.012$) and moderately in the right ROI ($t = 1.828$, $p = 0.142$). However, attention had little effect in the upper and lower ROIs, even a weak suppression effect in the upper ROI.

When we compared the spatial distribution of attention modulation in the single configuration with those in the tangential and radial configurations, there were always larger attention enhancements at the locations where
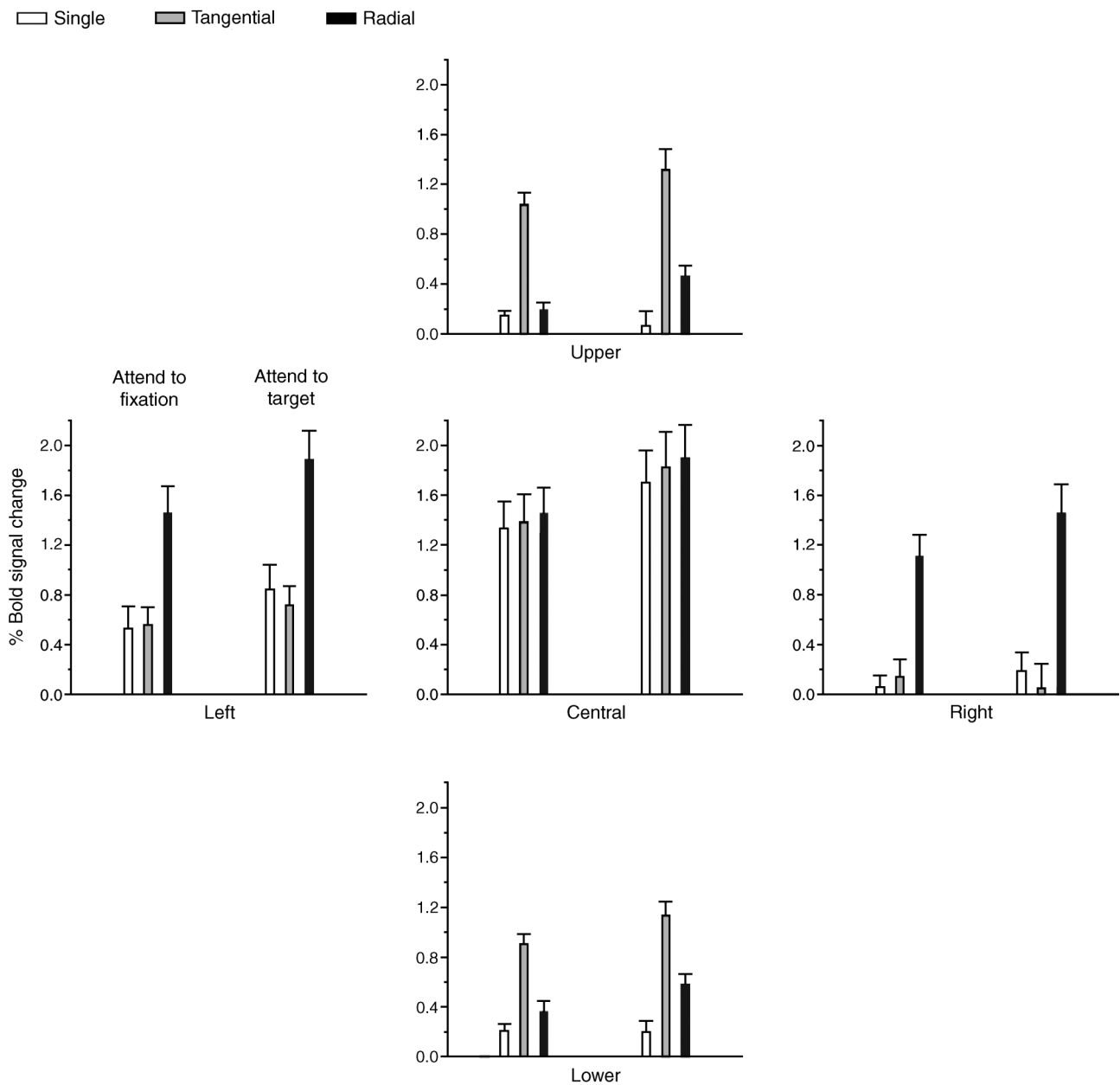
Figure 4. BOLD signals in the left, right, upper, lower, and central ROIs in the single, tangential, and radial configurations when subject attended to either the fixation (left part of a panel) or the target (right part of a panel). Error bars denote 1 *SEM* calculated across subjects.

distractors were present. The enhancements at the locations of distractors are likely to be automatic and stimulus driven (Corbetta & Shulman, 2002). It is more interesting to examine the attention modulation at the "no-distractor present" ROIs next to the target in the tangential and radial configurations. Compared to the single configuration, the attention enhancements in the tangential configuration significantly dropped down in the left ($t = 3.118$, $p = 0.036$) and right ($t = 3.623$, $p = 0.022$) ROIs, even became a bit suppressive in the right ROI (Figure 5A, gray bars). On the other hand, the attention enhancements in the radial configuration significantly boosted up

in the upper ($t = 3.64$, $p = 0.022$) and lower ($t = 3.863$, $p = 0.018$) ROIs (Figure 5A, black bars).

The overall picture of attention modulation across the locations of the target and the distractors is that, relative to the single configuration, attention spread into the neighboring regions in all directions in the radial configuration but was narrowed down to form a ridge along the circumference in the tangential configuration (Figure 4B). This effect was consistently found in all subjects.

Eye movement data demonstrated that subjects could fixate very well. Figure 6 shows the horizontal and vertical eye positions during an fMRI scan averaged
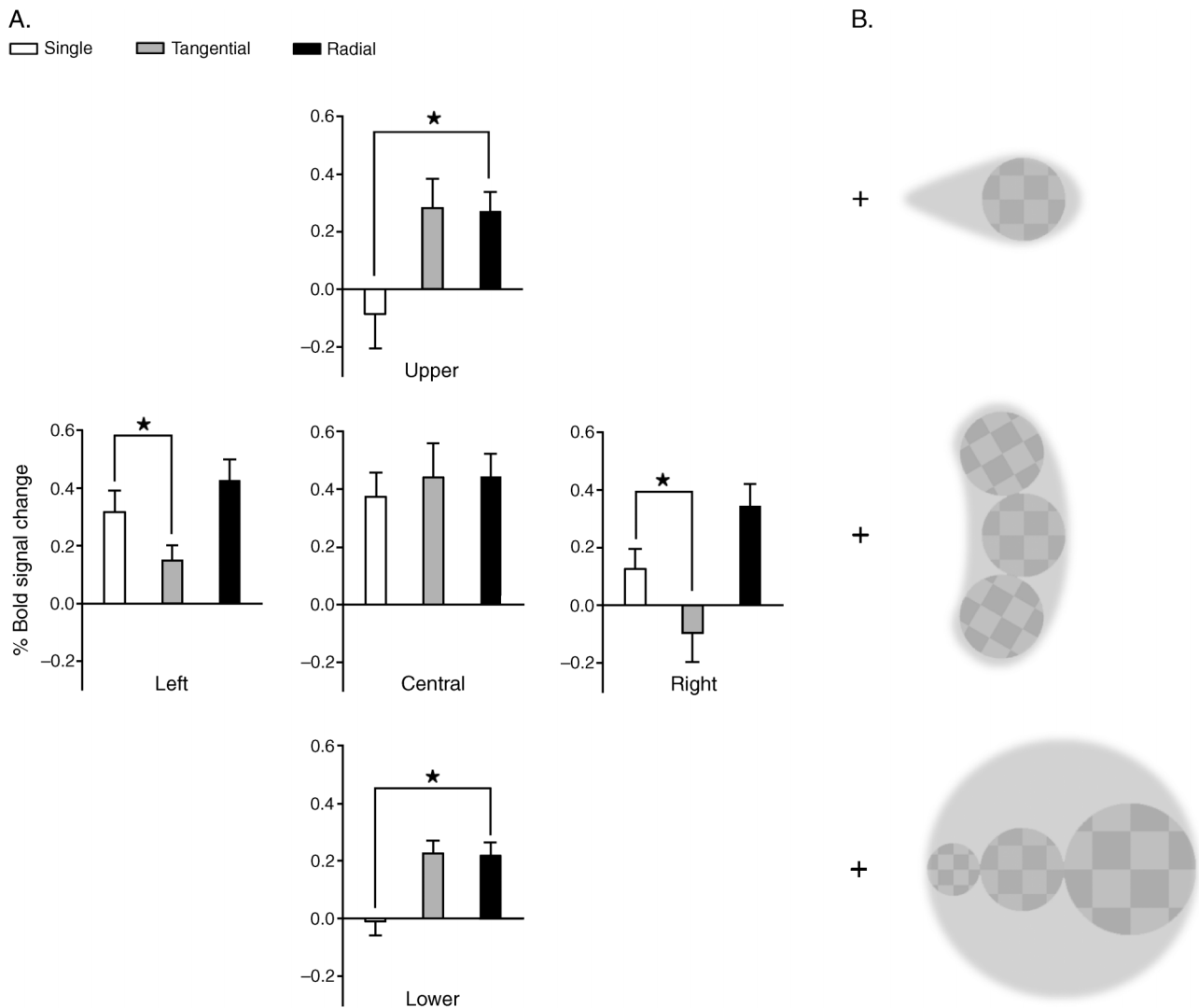
Figure 5. Attention effects at the target location and its surrounding area. (A) Attention modulation in the left, right, upper, lower, and central ROIs in the single, tangential, and radial configurations. Attention modulation was defined as the BOLD signal difference between the attend-to-target condition and the attend-to-fixation condition. Error bars denote 1 *SEM* calculated across subjects. (B) Schematic description of regions showing attention enhancements in the single, tangential, and radial configurations.

across subjects. Their eye movements were small and further statistical analyses confirmed that both horizontal and vertical mean eye positions did not significantly deviate from the fixation point in both attention conditions and for all stimulus configurations (all $t < 2$ and $p > 0.12$). These results suggest that it is unlikely that our results could be significantly confounded by eye movements.

Discussion

With the behavioral and fMRI data, we tried to examine the two distinct theories of the crowding effect. If the crowding effect results from pre-attentive lateral inhibition, we would observe a BOLD signal decrease at the

target location by presenting the distractors, as a previous study has demonstrated (Zenger-Landolt & Heeger, 2003). However, there was no significant difference between three stimulus configurations (single, tangential, and radial) in this unattended condition. The discrepancy between our study and Zenger-Landolt and Heeger's study might be attributed to stimulus differences, the tasks subjects performed and fMRI protocol (e.g., high resolution vs. standard resolution). In another study, Arman, Chung, and Tjan (2006) varied the distance between the target and the distractors to manipulate the strength of the crowding effect. They found that this manipulation did not affect the overall V1 response to the target and the distractors, which is in line with our observation.

In this study, we measured cortical responses not only at the target location, but also at its surrounding area, which
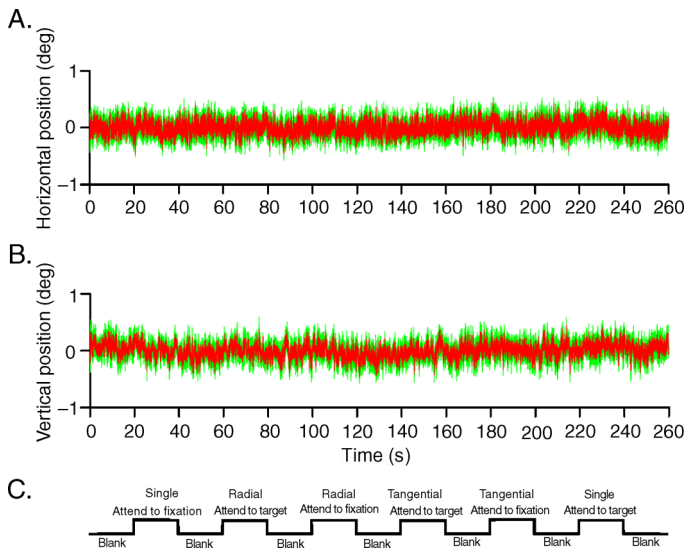
Figure 6. Horizontal (A) and vertical (B) eye positions during an fMRI scan averaged across subjects and their corresponding stimulus protocol (C). Red traces are the measured eye positions and the green shaded regions indicate ±SEM across subjects.

made it possible to investigate the spatial distribution of attention modulation in the crowded and non-crowded conditions. Many researchers have used a "spotlight" metaphor to describe that attention can be restricted to a small part of the visual field and that this beam of attention can move around. Psychophysical measurements of the distribution of attention within the beam suggested that the gradient of attention has a tear-drop shape, orientated along radial lines from fixation (Andersen & Kramer, 1993; LaBerge & Brown, 1989). The schematic description of attention enhancement for the single configuration stimulus (see Figure 4B) supports the claim about the shape of "attention spotlight" acquired from psychophysics.

Although there was no significant difference in attention enhancement at the target location between three stimulus configurations, we found that the spatial distribution of attention modulation was altered by crowding. Compared with the single configuration, attention enhancement spread into the neighboring regions in all directions in the radial configuration. However, this phenomenon did not occur in the tangential configuration. The radial arrangement of target and distractors is pervasive in crowding studies because it is very effective at inducing a profound crowding effect. Previous studies by He et al. (1996, 1997) suggested that spatial resolution of attention could determine the strength of the crowding effect. Intriligator and Cavanagh (2001) found that spatial resolution of attention is coarser in the radial direction than in the tangential direction, which they used to explain the radial/tangential asymmetry in the crowding effect.

Similarly, Strasburger (2005) has argued that crowding is due to spatially imprecise focusing of attention, especially in the radial configuration. Either a coarse resolution of spatial attention or imprecise focusing of attention could lead to the difficulty of localizing a target. Our subjects reported this kind of difficulty in the radial configuration, even with a preceding position cue. The spread of attention enhancement in the radial configuration might reflect a coarse spatial resolution of attention and/or unfocussed attention accompanying the localization difficulty. In other words, the radial configuration made it difficult to restrict spatial attention precisely on the target and consequently attention enhancement had spread into the regions next to the target where even no stimulus was presented. Overall, our data support the attention explanation of the crowding effect.

As mentioned above, the crowding effect occurs under a wide range of conditions and tasks. It should be noted that, although our data favor the attention explanation, it is important to test if our conclusion can be generalized to other conditions in the future. Furthermore, although we did not find the BOLD response to the target can be modulated by the presence of the distractors in the unattended condition, it remains possible that a modulation, if any, can be detected by more sensitive techniques and better designs.

In summary, we found the local cortical response to a peripheral target was not affected by the presence of neighboring distractors in the unattended condition. However, crowding altered the spatial distribution of attention modulation in the target and its surrounding area. Our data showed that in the radial configuration, attention enhancement spread into the target's surrounding area regardless of whether there was a stimulus presented. This pattern of results is more consistent with that the crowding effect arises from coarse spatial resolution of attention and unfocussed spatial attention rather than sensory level lateral inhibition at a pre-attentive level.

Acknowledgments

Commercial relationships: none.
Corresponding author: Fang Fang.
Email: ffang@pku.edu.cn.
Address: Department of Psychology and Key Laboratory of Machine Perception (Ministry of Education), Peking University, 5 Yiheyuan Road, Beijing 100871, PR China.

References

Andersen, G. J., & Kramer, A. F. (1993). Limits of focused attention in three-dimensional space. *Perception & Psychophysics, 53,* 658–667. [PubMed]

Arman, A. C., Chung, S. T. L., & Tjan, B. S. (2006). Neural correlates of letter crowding in the periphery [Abstract]. *Journal of Vision, 6*(6):804, 804a, http://journalofvision.org/6/6/804/, doi:10.1167/6.6.804.

Banks, W. P., Larson, D. W., & Prinzmetal, W. (1979). Asymmetry of visual interference. *Perception & Psychophysics, 25,* 447–456. [PubMed]

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226,* 177–178. [PubMed]

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. *Journal of Neurophysiology, 88,* 2530–2546. [PubMed] [Article]

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews, Neuroscience, 3,* 201–215. [PubMed]

Duncan, R. O., & Boynton, G. M. (2003). Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron, 38,* 659–671. [PubMed] [Article]

Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex, 7,* 181–192. [PubMed] [Article]

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383,* 334–337. [PubMed]

He, S., Cavanagh, P., & Intriligator, J. (1997). Attentional resolution. *Trends in Cognitive Sciences, 1,* 115–121.

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology, 43,* 171–216. [PubMed]

LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review, 96,* 101–124.

Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research, 25,* 963–977. [PubMed]

Levi, D. M., Klein, S. A., & Yap, Y. L. (1987). Positional uncertainty in peripheral and amblyopic vision. *Vision Research, 27,* 581–597. [PubMed]

Levitt, J. B., & Lund, J. S. (1997). Contrast dependence of contextual effects in primate visual cortex. *Nature, 387,* 73–76. [PubMed]

Louie, E. G., Bressler, D. W., & Whitney, D. (2007). Holistic crowding: Selective interference between configural representations of faces in crowded scenes. *Journal of Vision, 7*(2):24, 1–11, http://journalofvision.org/7/2/24/, doi:10.1167/7.2.24. [PubMed] [Article]

Martelli, M., Majaj, N. J., & Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision, 5*(1):6, 58–70, http://journalofvision.org/5/1/6/, doi:10.1167/5.1.6. [PubMed] [Article]

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision, 4*(12):12, 1136–1169, http://journalofvision.org/4/12/12/, doi:10.1167/4.12.12. [PubMed] [Article]

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science, 268,* 889–893. [PubMed]

Smith, A. M., Lewis, B. K., Ruttimann, U. E., Ye, F. Q., Sinnwell, T. M., Yang, Y., et al. (1999). Investigation of low frequency drift in fMRI signal. *Neuroimage, 9,* 526–533. [PubMed]

Strasburger, H. (2005). Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision, 5*(11):8, 1024–1037, http://journalofvision.org/5/11/8/, doi:10.1167/5.11.8. [PubMed] [Article]

Strasburger, H., Harvey, L. O., Jr., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception & Psychophysics, 49,* 495–508. [PubMed]

Stuart, J. A., & Burian, H. M. (1962). A study of separation difficulty: Its relationsip to visual acuity in normal and amblyoipic eyes. *American Journal of Ophthalmology, 53,* 471–477. [PubMed]

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research, 32,* 1349–1357. [PubMed]

Walley, R. E., & Weiden, T. D. (1973). Lateral inhibition and cognitive masking: A neuropsychological theory of attention. *Psychological Review, 80,* 284–302. [PubMed]

Westheimer, G., Shimamura, K., & McKee, S. P. (1976). Interference with line-orienation sensitivity. *Journal of the Optical Society of America A, 66,* 332–338. [PubMed]

Westheimer, G., & Truong, T. T. (1988). Target crowding in foveal and peripheral stereoacuity. *American Journal of Optometry and Physiological Optics, 65,* 395–399. [PubMed]

Wilkinson, F., Wilson, H. R., & Ellemberg, D. (1997). Lateral interactions in peripherally viewed texture arrays. *Journal of the Optical Society of America A, 14,* 2057–2068. [PubMed]

Williams, R. A., & Essock, E. A. (1986). Areas of spatial interaction for a hyperacuity stimulus. *Vision Research, 26,* 349–360. [PubMed]

Wolford, G., & Chambers, L. (1983). Lateral masking as a function of spacing. *Perception & Psychophysics, 33,* 129–138. [PubMed]

Zenger-Landolt, B., & Heeger, D. J. (2003). Response suppression in V1 agrees with psychophysics of surround masking. *Journal of Neuroscience, 23,* 6884–6893. [PubMed] [Article]