Vision Research 49 (2009) 44–53

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Legibility of Chinese characters in peripheral vision and the top-down influences on crowding

Jun-Yun Zhang [a,1], Ting Zhang [a,1], Feng Xue [b], Lei Liu [c,*], Cong Yu [a,*]

[a] State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China
[b] EENT Hospital, Fudan University, Shanghai, China
[c] School of Optometry, University of Alabama at Birmingham, Birmingham, AL, USA

ARTICLE INFO

Article history:
Received 5 May 2008
Received in revised form 24 September 2008

Keywords:
Chinese characters
Legibility
Crowding
Top-down influences
Peripheral vision

ABSTRACT

Written Chinese is distinct from alphabetic languages because of its enormous number of characters with a great range of spatial complexities (stroke numbers). In this study we investigated the impact of spatial complexity on legibility of Chinese characters as well as associated crowding in peripheral vision. Our results showed that for isolated characters, threshold sizes of complex characters increased faster with retinal eccentricity than did those of simple characters, suggesting possible "within-character" crowding among parts of complex Chinese characters. However, such "within-character" crowding was rendered negligible by strong "between-character" crowding introduced by flankers. When the target and flankers belonged to different complexity groups, the intensity and extent of crowding were greatly reduced, which could be explained by top-down influences as well as lower-level mechanisms. We suggest that crowding can be attributed to multiple mechanisms at different levels of visual processing.

1. Introduction

Most studies of letter legibility use Roman letters. Roman letters are highly stylish visual stimuli that are made of a small number of strokes, have no discernible parts, and are relatively uniform in spatial complexity as a stimulus set. It is less clear how much of our knowledge obtained from such stimuli can be applied to legibility of Chinese characters (CCs) that contain 1 to as many as 52 strokes, and thus have a wide range of spatial complexities. Recently we reported a study on legibility of CCs in foveal vision (Zhang, Zhang, Xue, Liu, & Yu, 2007), in which we measured threshold (acuity) sizes for six groups of frequently used CCs from low to high spatial complexities, and determined the relationship between legibility and optical defocus for Landolt C, Snellen E and three groups of CCs representing low, medium, and high spatial complexities. Our results showed that CC acuity sizes increase steadily with stimulus complexity, though at a slower rate than what would be expected if visual acuity is based on discerning the finest details of the stimuli. Moreover, the acuity size vs. optical defocus functions of the three CCs groups and Snellen E have similar slopes, differing only by a vertical shift (approximately one, two, and three lines above E acuity on an acuity chart, respectively), suggesting the feasibility of using Snellen E acuity, which is the current standard optotype for acuity testing in China, to derive the legibility of CCs in foveal vision. To understand the slower rate of acuity size increase against spatial complexity, we also developed a geometric moment model, in which we propose that human letter recognition performance near the acuity limit can be accounted for by a set of global features described by easy-to-visualize and perceptually meaningful low-order geometric moments (i.e., the ink area, variance, skewness, and kurtosis; manuscript under review).

The current study extends our work to the legibility of CCs, as well as crowding, in peripheral vision. We are particularly interested in two distinct characteristics of CCs that could affect peripheral character legibility and crowding in ways not normally evident when alphabetic stimuli are used. First, the majority of CCs are spatially complicated. Only 4% of CCs are single-body characters (e.g., 个) that have a small number of strokes, and cannot be divided graphically into smaller and meaningful parts. The rest are compounds (e.g., 需) that are made of multiple parts, each of which is a single-body character or its derivation, and are packed in the same square area as the single-body CCs. We suspect that interactions among these parts could interfere with recognition of a complex CC as a whole, and such interactions, or "within-character" crowding, could be magnified in the periphery. If this is indeed the case, acuities of different complexity CC groups may have different spatial scaling functions in the periphery, and thus may not be derived from a standard measurement like E acuity, as we showed previously for foveal vision (Zhang et al., 2007), without

---

* Corresponding authors.
E-mail addresses: liul7788@uab.edu (L. Liu), yucong@bnu.edu.cn (C. Yu).

[1] These authors contributed equally to this work.

0042-6989/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2008.09.021

proper compensations of scaling differences among CC groups. Such a possibility would have important clinical implications in evaluating peripheral vision of patients who read text that contain characters of different spatial complexities.

To address this issue, in the first part of the study, we measured threshold sizes of single CCs of various complexities at different retinal eccentricities. By comparing the slopes of spatial scaling functions for different complexity CCs groups, we revealed an inferiority of complex CCs to simple CCs in the visual periphery, possibly indicating “within-character” crowding among parts of complex CCs. We also measured threshold sizes of flanked CCs in a trigram configuration to assess the impact of ”within-character” crowding on regular “between-character” crowding.

The second distinct characteristic of CCs we are particularly interested in is that, in real-world Chinese text, more than often is a character flanked by characters of different spatial complexities. Such configurations are rarely seen in alphabetic languages because alphabetic letters tend to have similar spatial complexities. In cases where the target and flanking characters have different spatial complexities, some basic stimulus properties, such as the brightness and the spatial frequency contents, are different between the target and flankers. These and other physical stimulus differences including shape, size, polarity, etc., are known to affect crowding by segregating the target and flankers (Chung, Levi, & Legge, 2001; Hess, Dakin, & Kapoor, 2000; Kooi, Toet, Tripathy, & Levi, 1994; Nazir, 1992). Moreover, a Chinese reader knows naturally that the target and flanking characters with very different spatial complexities in a trigram configuration, such as 个需十, are drawn from different stimulus groups, so that he or she will not report a flanking character as the target. There is evidence that such misreporting contributes to crowding (Strasburger, 2005). Therefore, both stimulus differences and high-level top-down influences may affect crowding when the target and flanking characters differ in complexity.

In the second part of this study we assessed the impact of target–flanker complexity contrast on crowding. We also designed experiments to isolate the top-down influence on crowding, using not only CCs but also English Sloan letters. Moreover, after isolation of top-down influences, we were able to manipulate stimulus physical features to identify lower-level mechanisms underlying crowding. On the basis of our results, as well as previously reported findings, we propose an eclectic view that uses multiple mechanisms at multiple processing levels to explain crowding.

2. Methods

2.1. Observers and apparatus

Six observers with normal or corrected-to-normal vision participated in the study. All observers were young (mean age = 23.3 years) native Chinese speakers with college education and at least 6 years of training in reading and writing English. Observers ZJ and ZT were coauthors and were experienced in psychophysical experiments. The others were new to psychophysical observations and were unaware of the purposes of the study. Written informed consent was obtained from all observers prior to the tests.

The stimuli were generated by a Matlab-based WinVis program (Neurometrics Institute, Oakland, CA) and were presented on a 21-in. Sony G520 color monitor (2048 pixel × 1536 pixel, 0.189 mm × 0.189 mm per pixel, 75 Hz frame rate). The minimal and maximal luminance of the monitor was 1.18 and 91 cd/m^2, respectively. Viewing was monocular in a dimly lit room. A head-and-chin rest was used to stabilize the head position.

2.2. Stimuli

The test stimuli (Fig. 1a) consisted of one group of English Sloan letters and four groups of CCs. Each stimulus group contained five letters or characters (with exception in Experiment IV) with similar legibility as determined in a previous study (Zhang et al., 2007). In that study, 500 most frequently used CCs were selected and categorized into six groups according to the number of strokes (CC1–CC6 groups, from 2–4 to 16–18 strokes/character). Then ten characters were selected from each group based on intermediate Euclidean distances of character bitmaps, pronunciation, and spatial configuration. The legibility of these characters, along with ten Sloan letters, was measured in young normal observers using a rigorous psychophysical method. Based on these measurements, five stimuli with the most similar legibility within each group were selected for the use in the current experiments (CC2 and CC5 were not used). Since this article is part of a series of studies of Chinese character acuity, recognition and reading, which uses some or all six CC groups of different complexities, we chose to use these group names to be consistent with other articles. The bitmaps of the Sloan letters and CCs had the same width and height (50 × 50 pixels). The Sloan letter had uniform stroke width equivalent to 1/5 of the letter height. Font type bold Heiti (black font) was used for CCs because the strokes had relatively uniform width and were free of serif. To fit different number of strokes into the same area, stroke widths became gradually thinner as the characters became more complex. For the 50 × 50 pixel bitmaps we used, vertical stroke widths shifted from predominantly 7 pixels in CC1 to predominantly 6 pixels in CC6, and horizontal stroke widths shifted from 5–6 pixels in CC1 to 4–5 pixels in CC6.

The spatial complexity of the stimuli was also described by stroke frequency (Zhang et al., 2007). Each letter or character was sliced at 6 direction/position combinations: horizontal on the upper and lower halves, vertically on the left and right halves, and obliquely at 45° and 135° on the central portion of the stimuli. From each slicing we obtained the average crossed strokes and calculated the maximum of the 6 slicing as the stroke frequency. The average stroke frequency for the Sloan letters was 2.0 strokes/letter. The average stroke frequencies for the six groups of CCs increased monotonically from 2.2 to 5.5 strokes/character (Zhang et al., 2007).

2.3. Procedure

The target was a black Sloan letter or Chinese character presented on a full-screen full-luminance white background. The target was presented either alone or was flanked by two horizontally aligned letters or characters (trigram). The target could be any member of a stimulus group, and the two flankers were always

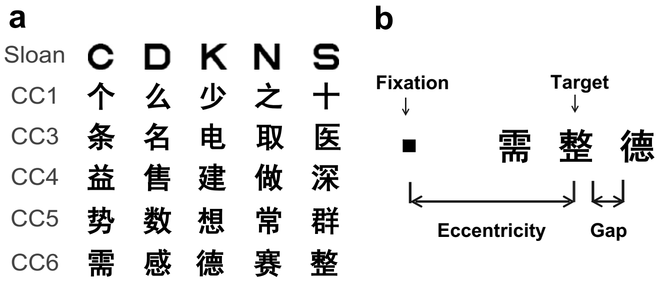

Fig. 1. Stimuli. (a) Sloan letters and four groups of CCs with various spatial complexities used in the study. CC1: 2–4 strokes; CC3: 8–9 strokes; CC4: 11–12 strokes; CC6: 16–18 strokes. (b) Stimulus configuration for peripheral testing when flankers were present. The flanking characters had the same size as the target, and the edge-to-edge gap between target and flanker was one character wide.

different from each other and from the target. The flankers always had the same size as the target, and the edge-to-edge target–flanker gap was one character wide if unspecified (Fig. 1b). The target was presented at 0°, 5°, or 10° retinal eccentricities on the horizontal meridian in the temporal visual field. The viewing distance was 6, 1.6, and 0.8 m for 0°, 5°, and 10° retinal eccentricity, respectively.

In each trial of foveal testing, a 0.1° square fixation was first displayed for 200 ms at the center of the screen accompanied with a beep, which was followed by a 300 ms time gap prior to the onset of the stimulus. The stimulus duration was 200 ms. When flankers were used their display was always synchronized with the target with the same abrupt onset and offset. For peripheral testing, the central fixation was always present, and the observer was asked to fixate at it. At the beginning of each trial, a small square (0.1°) flashed for 200 ms at the target location as a location cue, which was followed by a 300 ms gap prior to the onset of the stimulus. The stimulus was presented for 200 ms. The observer' task was to identify the target from a list of the five members of the target group (the list was printed on paper for observer's reference), and to report the result by pressing a number key. An auditory feedback was provided upon an incorrect response.

The threshold letter size without or with flankers was measured with the method of constant stimuli. In Experiments I and II, which were run together, each experimental session was composed of threshold size measurements with a combination of stimulus group, retinal eccentricity, and flanking conditions. Each threshold measurement was based on five levels of stimulus size with 10 presentations at each level. A typical round of experiments consisted of 30 sessions (5 stimuli groups × 3 eccentricities × 2 flanking conditions), which were run according to a randomly permuted table for each observer and were completed in about two days. Each observer completed 7 rounds of the experiments. All conditions in each sub-experiment of Experiments III and IV could be covered within a 2-h session and were repeated in several days. The percent correct data were fit with a Weibull function: $P = 1 - (1-\gamma)e^{-(x/th)^{\beta}}$, where P was the percent correct, γ was the guessing rate (0.2 in a 5AFC trial), x was the stimulus angular size, β was the slope of the psychometric function, and th was the threshold size for recognition at a 70.6% correct level.

3. Results

3.1. Experiment I: Legibility of Chinese characters in peripheral vision

This experiment measured threshold sizes for four groups of isolated CCs as well as Sloan letters at 0°, 5°, and 10° retinal eccentricities. Individual and mean threshold sizes plotted against eccentricity, along with regression lines (weighted with error bars), were shown in Fig. 2a and b. A repeated measures ANOVA indicated that for all stimulus groups, the threshold sizes increased with the retinal eccentricity linearly ($p < .001$; Fig. 2a and b). The threshold sizes of the more complex CCs (CC4 and CC6) were similar ($p = .978$), and were significantly larger than those of simpler CC1 ($p = .002$) and CC3 ($p = .026$). CC3 threshold sizes were larger than that of CC1 ($p = .032$), and CC1 threshold sizes were larger than that of Sloan letters ($p = .022$). The latter could be explained by the thicker strokes of the Sloan letters (Zhang et al., 2007).

There was a significant interaction between stimulus groups and eccentricities ($p < .001$), suggesting that the increase of threshold sizes with the retinal eccentricity was affected by the stimulus groups. To characterize this interaction, peripheral threshold sizes were normalized by corresponding foveal threshold sizes. The resultant size scaling functions were shown in Fig. 2c, and the function slopes were plotted against stroke frequency in Fig. 2d. These plots showed a systematic increase of scaling function slope from simple to more complex CCs. The slopes of CC6 and CC4 were 24% and 26% greater than that of CC3, respectively, and 56% and 59% greater than that of CC1, respectively. Moreover, when slopes of the scaling functions for four CC groups were plotted against the stimulus complexities (stroke frequencies), the slope of the regression line was significantly different from zero ($p = .002$) (Fig. 2d). These data indicated that the threshold sizes of more complex CCs (CC4 and CC6) increased at a faster rate with the retinal eccentricity than did those of simpler CCs. We interpreted this systematic change of regression slope as evidence for possible interactions among components of more complex CCs, or "within-character" crowding, in the visual periphery (see Section 4).

3.2. Experiment II: Crowding between Chinese characters

A letter is more difficult to identify when it is closely flanked by additional letters (Flom, Heath, & Takahashi, 1963; Stuart & Burian, 1962. See Levi (2008) for a most recent review). Would such crowding between the target and flanker characters be affected by within-target crowding? In this experiment we measured the threshold sizes for flanked Sloan, CC1, CC3, CC4, and CC6 targets at 0°, 5°, and 10° retinal eccentricities. The target and flankers were drawn from the same 5-member stimulus group (Fig. 1a), and the edge-to-edge gap between target and flankers was always one character width (Fig. 1b). This experiment was run together with Experiment 1 on the same observers (see Section 2). Individual data, their averages, and the regression lines are shown in Fig. 3a and b.

As expected, strong crowding was evident in recognition of flanked Sloan letters and CCs in peripheral vision. The slopes of spatial scaling functions were much steeper for flanked targets (Fig. 3c, dashed lines) than for isolated targets (Fig. 3c, solid lines, replotted from Fig. 2c). In the fovea, threshold sizes under the flanker and no-flanker conditions were not significantly different ($p = .591$), consistent with Flom (1991) that foveal crowding did not extend beyond one character width.

The best fitting lines of the threshold size vs. retinal eccentricity functions became steeper with increasing CC complexity (Fig. 3a and b). However, this increase only reflected foveal threshold size differences among the CC groups. When peripheral threshold sizes were normalized by corresponding foveal threshold sizes, the differences among the scaling function slopes of various CC groups were insignificant ($p = .344$; Fig. 3c). When the slopes of the scaling functions for the four CC groups were plotted against stroke frequencies, the slope of the regression line was not significantly different from zero ($p = .679$) (Fig. 3d). These results suggested that when flankers were present, characters of different spatial complexities scaled in a similar manner with retinal eccentricity.

It is important to distinguish the normalized spatial scaling factors by foveal thresholds in our study from Bouma (1970) unnormalized spatial scaling factors. Bouma (1970) reported that the unnormalized scaling factor for critical crowding zone is approximately 0.5 (i.e., half the retinal eccentricity). This factor varied from 0.23 (Sloan) to 0.37 (CC6) in our data when the sizes of the critical zones were calculated in target–flanker center-to-center distance at a 70.6% correct rate (the threshold values were in edge-to-edge gap size in Fig. 3), smaller than Bouma's factor of 0.5. This difference could be due to the different criterions set to define the thresholds (Levi, 2008).

3.3. Experiment III: The effects of target–flanker complexity contrast on crowding

In the introduction we suggested that in normal Chinese text a character is more likely to have neighboring characters with different spatial complexities. Such complexity differences would

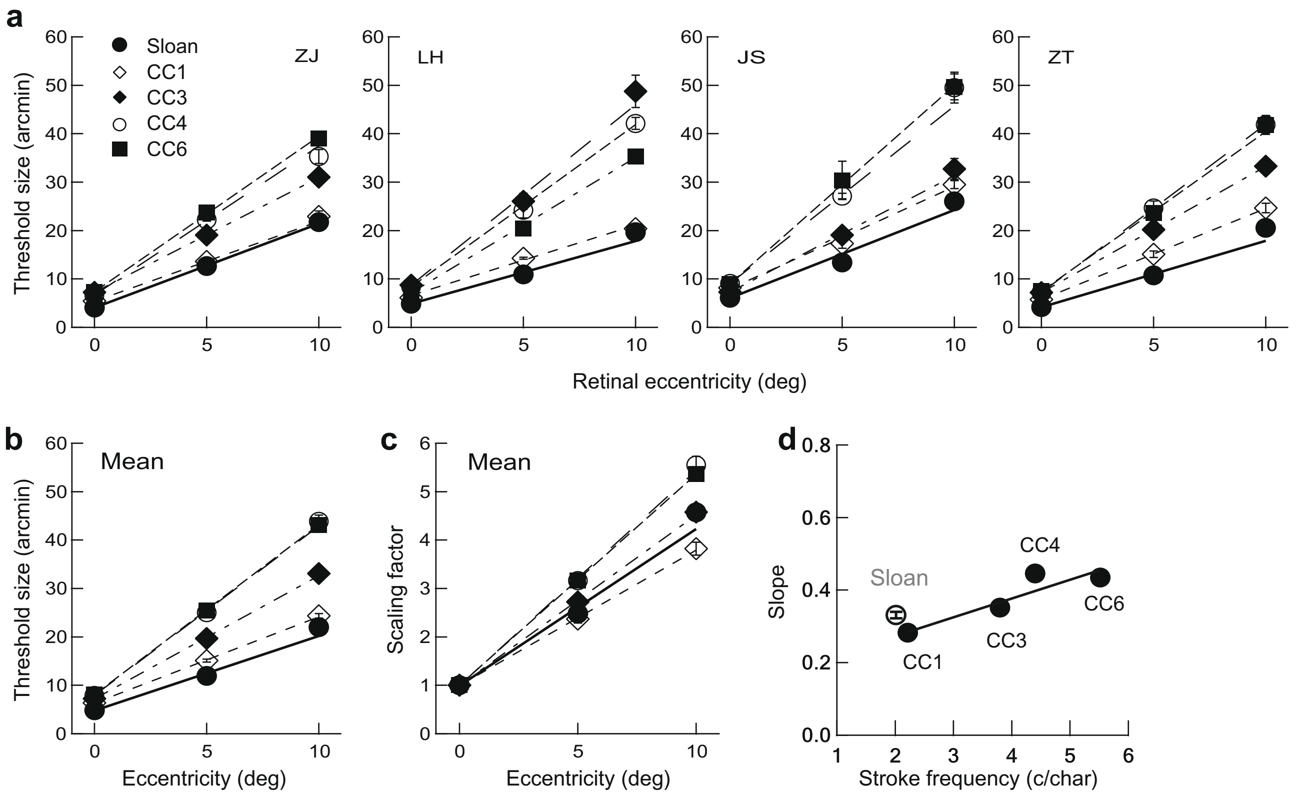

Fig. 2. Legibility for isolated Chinese characters. (a and b) Individual and mean threshold sizes as a function of retinal eccentricity. (c) Scaling factors (normalized threshold sizes by foveal thresholds) as a function of retinal eccentricity. The straight lines represent linear regression. (d) Slopes of the spatial scaling functions against the stroke frequency. The straight line represents linear regression for CC stimulus groups only. Error bars showed one standard deviation.

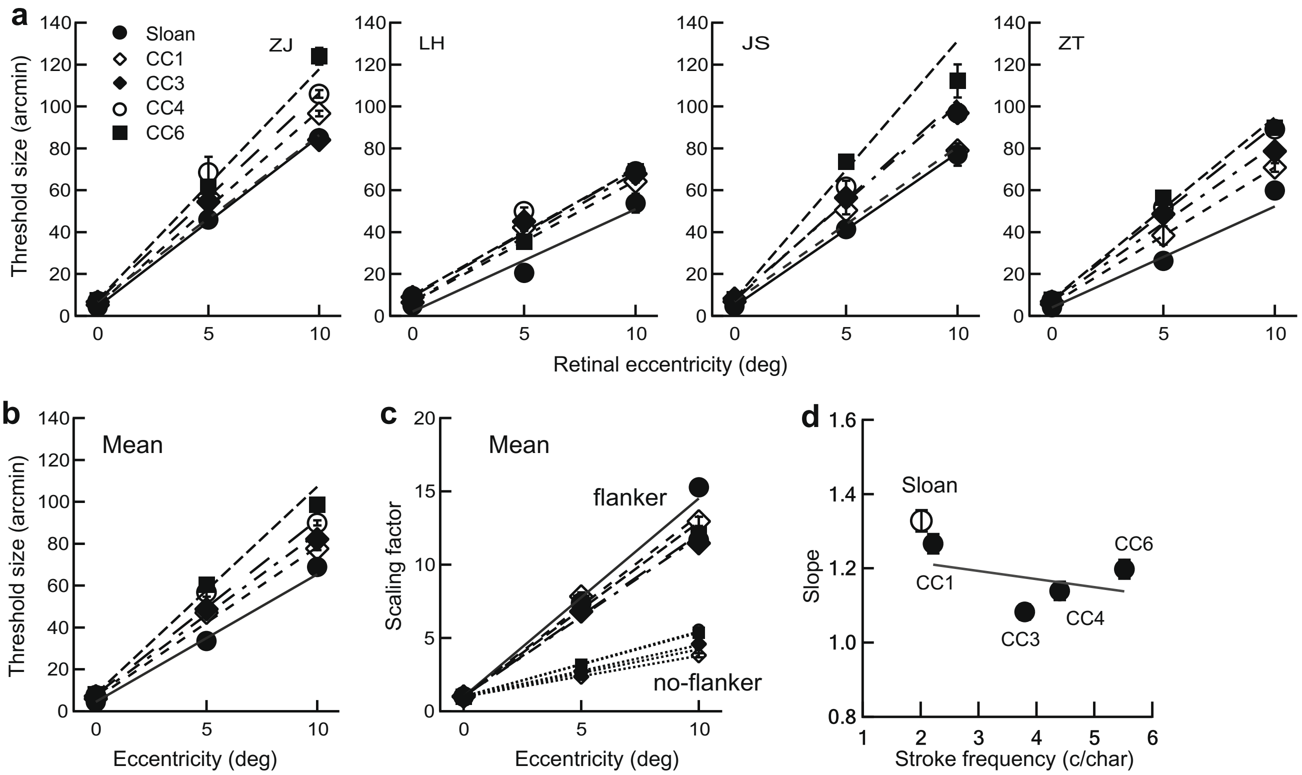

Fig. 3. Crowding with flanked Sloan letters and Chinese characters. (a and b) Individual and mean threshold sizes as a function of retinal eccentricity. (c) Scaling factors (normalized threshold sizes by foveal thresholds) as a function of retinal eccentricity. No-flanker data are replotted from Fig. 2c. Straight lines represent linear regression. (d) Slopes of the size scaling functions as a function of stroke frequency. The dashed line represents linear regression for CC stimulus groups only. Note the y-axis in d and Fig. 2d are in different scales.

introduce low-level brightness and spatial frequency differences between the target and flankers. It would also introduce a top-down influence to segregate the target and flankers, especially when the complexity difference is large. In this experiment, we measured the effects of target–flanker complexity contrast on crowding with CCs. Later in Experiment IV we would isolate the top-down influences on crowding using CCs as well as English Sloan letters as stimuli.

3.3.1. The effect of target–flanker complexity contrast on crowding

To maximize complexity contrast, the least and most complex CCs, CC1 and CC6, were used as target and flanker stimuli. The average stroke frequencies were 2.22 and 5.52 strokes per character for CC1 and CC6 stimuli, respectively. Threshold sizes were measured at 10° retinal eccentricity for CC1 and CC6 targets with three target–flanker complexity contrast conditions: (1) zero complexity contrast: a CC1 or CC6 target with flankers from the same 5-member stimulus group (denoted as "111" and "666" conditions. Digits "1" and "6" stand for CC1 and CC6 characters, respectively, and the left, center, and right digits represent the left flanker, center target, and right flanker, respectively); (2) full complexity contrast: a CC1 target with CC6 flankers ("616" condition) or a CC6 target with CC1 flankers ("161" condition); (3) mixed complexity contrast: a CC1 target with a CC6 flanker and a CC1 flanker ("611/116" conditions) or a CC6 target with a CC1 flanker and a CC6 flanker ("166/661" conditions). Threshold sizes for single CC1 and CC6 without flankers were also measured as baselines (denoted as "1" and "6").

Fig. 4 shows the threshold sizes obtained under various target–flanker complexity contrast conditions. When the target and flankers had full complexity contrasts (616 and 161), crowding was reduced significantly from that at zero complexity contrast (111 and 666) ($p = .001$, repeated measures ANOVA), by 55.5 ± 4.4% for the CC1 target (Fig. 4, gray bars) and 34.0 ± 4.2% for the CC6 target (Fig. 4, black bars). Crowding was reduced more for the CC1 target by the CC6 flankers in the 616 configuration than for the CC6 target by the CC1 flankers in the 161 configuration. This asymmetry could be due to the fact that for the 616 configuration, when the CC1 target was near threshold, the CC6 flankers were most likely below their non-flanker "6" baseline thresholds (Fig. 4). Therefore, the features of these CC6 flankers were not very legible and had less chance to be improperly integrated with features of the CC1 target to produce crowding. However, crowding was not completely eliminated at full complexity contrast. Threshold sizes for 616 and 161 conditions were still significantly larger than "1" and "6" baselines ($p = .002$), which were 29.6 ± 4.0% and 38.7 ± 10.0% larger, respectively.

At mixed complexity contrasts, there was no significant difference whether the same-group flanker was on the left or right side of the target, so the results were averaged. Crowding at mixed complexity contrasts (116/611 and 166/661) was weaker than that at zero complexity contrasts (111 and 666) ($p = .008$ and .021, respectively, Fig. 4), but stronger than that at full complexity contrasts (616 and 161) ($p = .063$ and .021, respectively, Fig. 4).

However, it is worth mentioning that the above estimation of the complexity contrast effects were most conservative, with the assumption that the guessing rate of the center target was unchanged across various flanker conditions. However, letters at the beginning and end of a letter string are known to be more legible than letters in the middle (Wolford & Hollingsworth, 1974), so it was likely that at some character sizes in our experiments, the observers could recognize one or both flankers but not the target. When both flankers were recognized, the target guessing rate was 1/3 under zero complexity contrast conditions (111 and 666) because both flankers were member of the 5-character stimulus group, and 1/5 under full complexity contrast conditions (161 and 616) because both flankers were from a different stimulus group. The higher rates of correct guessing associated with the zero complexity contrasts would have caused underestimation of the threshold sizes for the 111 and 666 conditions, and underestimation of the threshold differences between the zero- and full-complexity contrast conditions.

3.3.2. The effect of target–flanker complexity contrast on critical spacing

Besides the threshold change, crowding is also quantified by its spatial extent or critical spacing (the zone within which flankers interfere with the target recognition). Several studies reported that

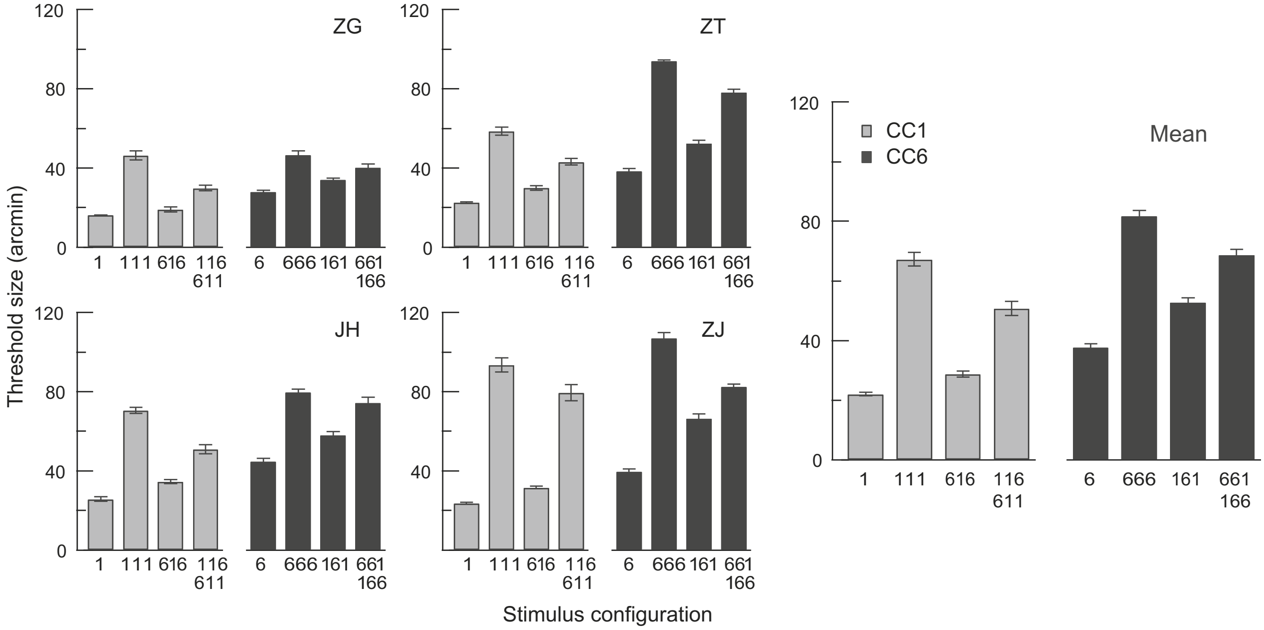

Fig. 4. The effects of target–flanker complexity contrast on crowding. 111 and 666: zero complexity contrasts; 616 and 161: full complexity contrasts; 116/611 and 661/166: mixed complexity contrasts. Digits "1" and "6" stand for CC1 and CC6 stimuli, respectively. The left, center, and right digits represent the left flanker, center target, and right flanker, respectively.

the critical spacing is approximately half the target retinal eccentricity regardless of the target size (Bouma, 1970; Chung et al., 2001; Pelli, Palomares, & Majaj, 2004; Tripathy & Cavanagh, 2002), but the exact value depends on how the spacing is defined (center-to-center or edge-to-edge) and what the criterion is to define the limits of the crowding zone (Levi, 2008).

We measured critical spacing of crowding at zero complexity contrasts (111 and 666) and full complexity contrasts (616 and 161) at 5° and 10° retinal eccentricities for the same four observers. Critical spacing for Sloan letters at zero complexity contrast was also measured for comparison. The sizes of the target and flankers were fixed at 1.2 times each observer's single character threshold sizes (Fig. 4), and the target correct report rate was measured as a function of the target–flanker center-to-center separation. Critical spacing was defined as the center-to-center separation at a 70.6% correct rate. Critical spacing for zero complexity contrast conditions (111, 666 and SSS for Sloan letters) was statistically similar at $1.80 \pm 0.47°$, $2.26 \pm 0.49°$, and $1.85 \pm 0.47°$ at 5° eccentricity (Fig. 5a), respectively, and at $3.17 \pm 0.13°$, $3.24 \pm 0.44°$, and $3.26 \pm 0.17°$ at 10° eccentricity (Fig. 5b), respectively ($p = .462$, repeated measures ANOVA). However, critical spacing was significantly smaller when the target and flankers were at full complexity contrasts ($p = .006$), with an overall reduction of 41.0%. The 616 complexity contrast condition reduced more crowding from the 111 condition (by 49.4%, averaged over 5° and 10° data, Fig. 5a and b, gray bars) than did the 161 complexity contrast condition from the 666 condition (by 32.6%, averaged over 5° and 10° data, Fig. 5a and b, black bars) ($p = .006$). The reductions of critical spacing were similar at 5° and 10° retinal eccentricities ($p = .161$).

3.4. Experiment IV: Top-down and lower-level influences on crowding

Strasburger (2005) reported that under crowding an observer might report the flanking letters as the target, which was supported by our error analysis using the 111 and 666 data in Fig. 4. Specifically, for all stimulus sizes producing less than 60% correct target report rate (mean = 38.6% and 37.8% for 111 and 666 conditions, respectively), the rate that the observers mistakenly reported one of the two flanking characters as the target was significantly higher than the rate reporting the other two unused characters (52.5% vs. 8.9% for the 111 condition and 44.6% vs. 17.6% for the 666 condition; $p < .001$, repeated measures ANOVA). These misreporting rates were calculated against the total number of included trials, not the number of wrong report trials, so the observers even reported the flankers more frequently than the correct target. However, when the target and flankers were drawn from different stimulus groups (i.e., 161 and 616 conditions), the observer would not report the flankers as the target, because he or she knew that the flanking characters were not on the list of reportable characters. Besides stimulus differences (i.e., brightness, spatial frequency) that might have segregated the target and flankers, how much would this top-down influence contribute to crowding reduction in Fig. 4? In this experiment we attempted to isolate this top-down influence on crowding, as well as to study lower-level mechanisms that also affect crowding.

3.4.1. High-level top-down influences

To isolate high-level top-down influences, we compared crowding when the target and flankers were drawn either from the same stimulus group, or from different stimulus groups, while keeping the target–flanker complexity contrast constant. To make this possible, as shown in Fig. 6a, the target in the trigram was always drawn from the five CC1 characters used in above experiments, and the flankers were either drawn from the remaining four characters ("same" flanker condition in Fig. 6), or from five other characters ("diff" flanker condition in Fig. 6). These new characters and the existing five characters had similar number of strokes (2∼4) and similar bitmap Euclidian distances among each other (Zhang et al., 2007). Therefore, the target–flanker complexity contrasts were zero under "same" and "diff" flanker conditions, but the flankers in the "same" condition were reportable characters and the flankers in the "diff" conditions were not. The observers were clearly informed whether the target and flanking characters were from the same stimulus group or from different groups, and the stimuli were listed on paper as a response guide. This design isolated the observer's knowledge of target and flanker identities as a top-down influence on crowding and controlled the impacts of lower-level stimulus factors. We also ran a parallel experiment using Sloan letters following the same procedure. The target was drawn from five Sloan letters (CDKNS) used in above experiments, and the flankers were drawn either from the remaining four letters, or from five other previously unused letters (VROHZ, Fig. 6a).

Fig. 6b showed that when the flankers were drawn from a different stimulus group, crowding was significantly reduced ($p = .007$, repeated measures ANOVA). The mean threshold size was reduced by $27.9 \pm 6.3\%$ for CC1 and $19.5 \pm 5.6\%$ for Sloan letters. There was no significant difference of crowding reduction between CC and Sloan letter stimuli ($p = .221$). These results demonstrated that the observers' knowledge of target and flanker identities as a top-down influence could significantly reduce crowding. However, compared to threshold reduction in the full complexity contrast condition (616) vs. the zero complexity contrast condition (111), which was $55.5 \pm 4.4\%$ (Fig. 4), threshold reduction in the "diff" flanker condition vs. the "same" flanker condition at the current rate of $27.9 \pm 6.3\%$ was less robust. This difference suggested that top-down influences could only account for part of the full complexity contrast effect on crowding, and the remaining effect needed to be attributed to stimulus physical differences that also segregate the target and flankers to reduce crowding (Chung et al., 2001; Hess et al., 2000; Kooi et al., 1994; Nazir, 1992).

Again, the above calculations of thresholds implicitly assumed equal guessing rates of the target in "same" and "diff" flanker conditions. Under the conditions where both flankers were recognizable, the target guessing rates for the "same" and "diff" condition would be 1/3 and 1/5, respectively. So the above estimation of the top-down influences on crowding, which was reflected by the threshold differences between the "same" and "diff" flanker conditions, was most conservative, as discussed in Experiment III.

3.4.2. A close look of the improper feature integration models of crowding

It has been proposed that crowding results from intermediate-level improper integration of target and flanker features when the target and flankers fall into an integration zone (Levi, Hariharan, & Klein, 2002; Pelli et al., 2004). Having quantified the top-down influences on crowding, we were able to manipulate lower-level flanker properties to have a close look of this improper feature integration process. Specifically, we measured crowding with stroke-scrambled CC1 flankers ("strkS" condition, Fig. 6), which scrambled the spatial arrangement of the strokes but retained all legitimate brush strokes (features), and with pixel-scrambled CC1 flankers ("pxlS" condition, Fig. 6), which abolished all legitimate strokes, and compared threshold changes against other flanker conditions.

Like the "diff" flanker conditions, observers would not report the flankers as the target by mistake in the stroke- and pixel-scrambled flanker conditions, so this top-down influence was matched. Moreover, stroke-scrambling broke letter-level processing of flanking characters that would have tied features together, possibly allowing the strokes to be more easily integrated into

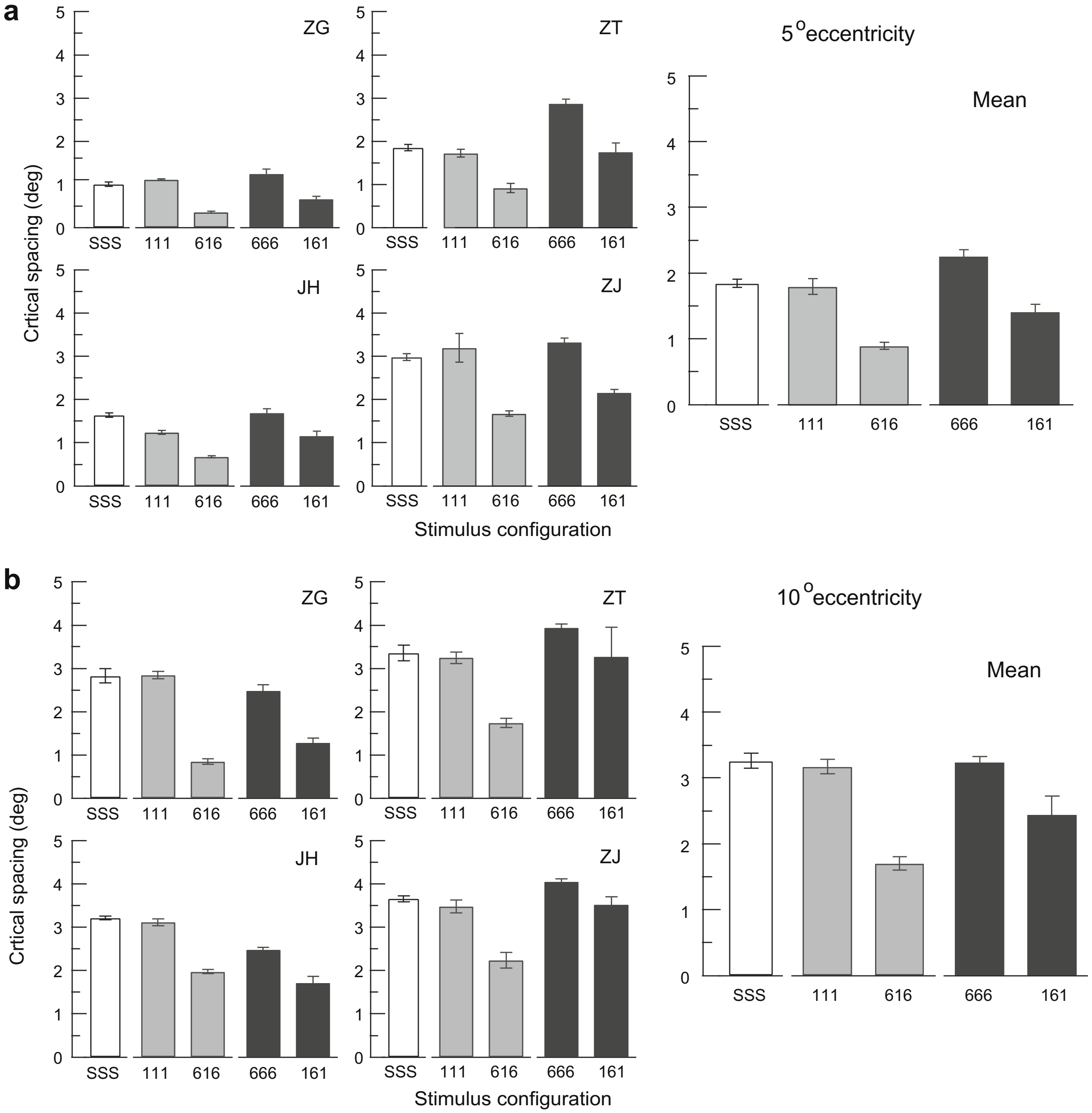

Fig. 5. The effect of target–flanker complexity contrast on critical spacing of crowding. (a) Individual and mean critical spacing for various stimulus configurations at 5° retinal eccentricity. (b) Individual and mean critical spacing at 10° retinal eccentricity. SSS, 111 and 666: zero complexity contrasts; 616 and 161: full complexity contrasts. Letter "S" and digits "1" and "6" stand for Sloan, CC1, and CC6 stimuli, respectively.

the target. Meanwhile, pixel-scrambling destroyed features of the flanking characters, thus discouraged target–flanker feature integration. The results showed that stroke-scrambled flankers ("strkS") raised threshold sizes by 38.4 ± 7.6% compared to those with the unscrambled "diff" flankers (Fig. 6b; $p < .001$, paired *t*-test), suggesting that letter-level grouping of flanker features discouraged target–flanker feature integration. Moreover, after this letter-level feature grouping was disabled by stroke-scrambling of the flankers, the threshold sizes were not significantly different from the "same" flanker condition level ($p = .95$). It is worth mentioning that although the "same" and "strkS" flankers produced similar crowding, crowding by "strkS" flankers was affected by two counteracting processes: a top-down influence that reduced crowding, and a freer target–flanker feature integration due to disabled letter-level feature grouping that facilitated crowding. Such dynamics were not discernible without a baseline reference of top-down impact set by the "diff" flanker condition. On the other hand, pixel-scrambled flankers ("pxlS") nearly wiped out crowding. The threshold sizes were not significantly different from the no-flanker baselines ($p = .086$). This effect was predicted by the feature integration model, because after pixel-scrambling, there were no eligible features in the flankers that could be integrated with the target to produce crowding.

4. Discussion

In this study we demonstrated within-character crowding in recognition of isolated, predominantly complex, CCs in the visual periphery, and showed that such within-character crowding was rendered negligible by much stronger between-character crowding once the target character was flanked by other characters. We also found reduced crowding as a result of spatial complexity contrast

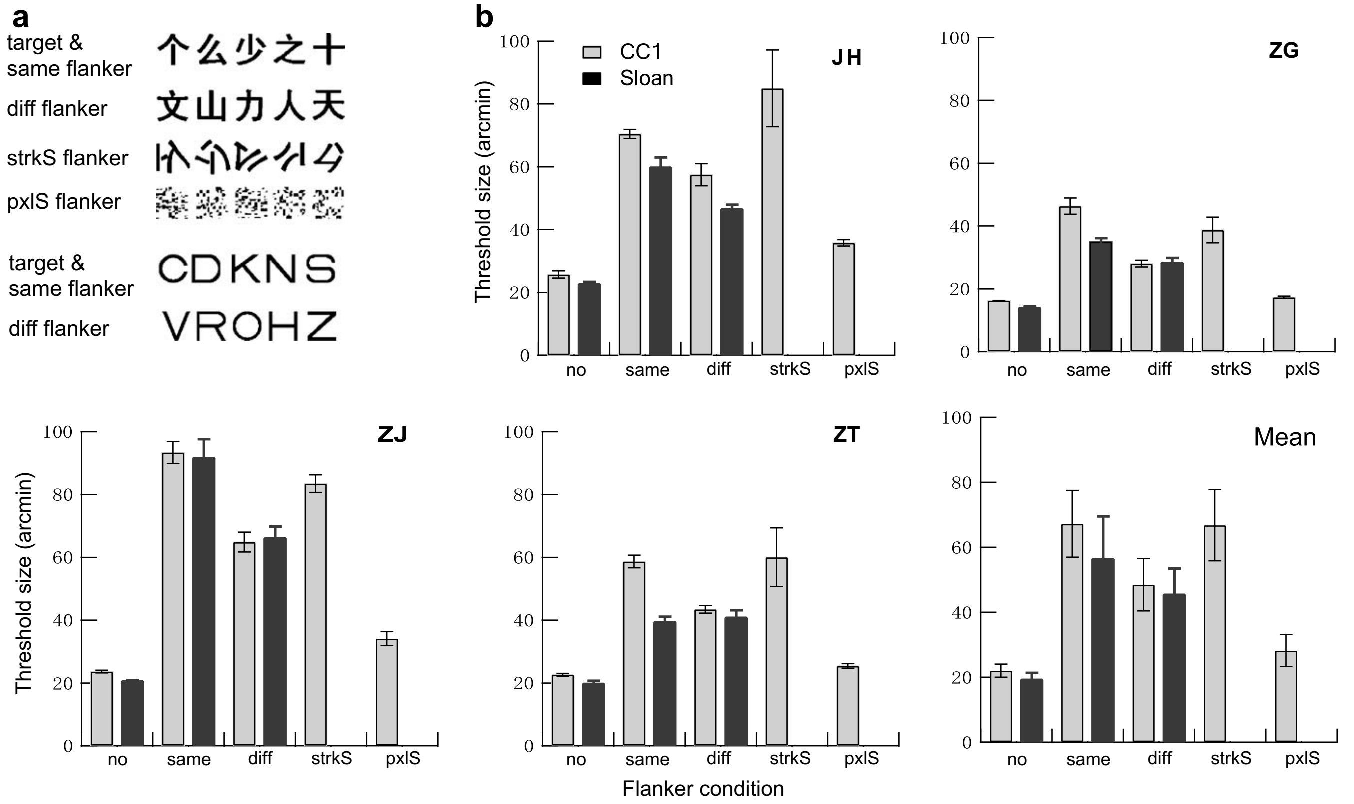

Fig. 6. Top-down and lower-level influences on crowding. (a) CC1 and Sloan letters used as target and different flanker stimuli. (b) Threshold sizes at different flanker conditions. no: no-flanker; same: the target and flankers drawn from the same stimulus group; diff: the target and flankers drawn from different stimulus groups; strkS: stroke-scrambled flankers; pxlS: pixel-scrambled flankers.

between the target and flanking CCs, and assessed the contributions of top-down and lower-level processes to this complexity contrast effect and to crowding in general.

4.1. Within-character crowding and its clinical implications

Our data showed that, as the retinal eccentricity increases, complex CCs have to be enlarged at a more rapid rate than simple CCs to reach equal legibility. Complex characters have more strokes than simple ones, and thus have higher object spatial frequency components (cycles/char, Parish & Sperling, 1991). Would the differences in object spatial frequency account for spatial scaling differences among different CC groups?

It is known that visual acuity varies linearly with retinal eccentricity (Herse & Bedell, 1989; Levi, Klein, & Aitsebaomo, 1985; Ludvigh, 1941; Rovamo & Virsu, 1979). If S_f and S_E are cut-off retinal frequencies in the fovea and at E deg eccentricity, then $S_E = S_f/(1 + E/E_2)$, where E_2 is the eccentricity at which the resolution has changed by a factor of 2. For a character whose height is H deg and whose object frequency is x c/char, its dominant retinal spatial frequency is x/H c/deg. When acuity threshold height is reached at an eccentricity E, the character's retinal frequency $S_E = x/H = S_f/(1 + E/E_2)$, and the threshold character height should vary with eccentricity in a linear fashion: $H = x(1 + E/E_2)/S_f$. At the fovea, the acuity height is $H_0 = x/S_f$. If we normalize each curve by its own foveal acuity height H_0, the normalized acuity height will be $\hat{H} = H/H_0 = 1 + E/E_2$, which is independent of the stimulus object frequency x, and the normalized lines should all be on top of each other. Thus, the differences in object spatial frequency are not responsible for the steeper scaling of complex CCs in Fig. 2c. Rather we hypothesize that the scaling differences might have resulted from interactions among parts of complex CCs, or "within-character" crowding.

Martelli, Majaj, and Pelli (2005) reported that contrast thresholds for recognition of a feature (a mouth or a letter) become higher when the feature is presented within a context (a face or a word) than when it is presented in isolation. This "face and word inferiority effect" appears to occur only in the periphery. Sheedy, Subbaram, Zimmerman, and Hayes (2005) reported a "letter superiority" effect, in that high contrast lowercase letters have 10–20% better foveal acuity than words made of 5–6 lowercase letters. In both cases, parts are more legible when presented alone than when presented within a meaningful whole, which is termed as "internal crowding" by Martelli et al. (2005). Our results revealed a different aspect of the part–whole relationship, in that a compound object made of more than one meaningful part is more difficult to recognize in the visual periphery than an undividable simple object. However, further experiments are required to provide direct evidence for crowding within a compound character. Nevertheless, if such interactions exist, they must occur before the whole is recognized. In comparison, the part or letter superiority effect may occur after the whole is recognized. For this reason, we name the interactions as "within-character" crowding for distinction.

Within-character crowding in the periphery may complicate visual function evaluation of Chinese reading patients. In foveal vision there is a rather simple relationship between the E acuity and legibilities of different complexity CCs (Zhang et al., 2007), which allows inference of foveal visual ability in recognizing different complexity CCs on the basis of one acuity measurement. However, this simple relationship does not apply to the periphery due to within-character crowding. A recent survey in China showed that the prevalence of age-related macular degeneration in the

75+ yr age group is 15–30% (Tian, Zhang, Li, Zhang, & Mu, 2005). Many of these patients may eventually have to rely on peripheral vision for their daily activities, including reading. Their peripheral visual ability will have to be assessed with proper consideration of within-character crowding. On the other hand, in real-world reading materials, CCs are organized in lines with small spacing between them. Our results suggest that within-character crowding may become less important in reading real Chinese text because between-character crowding is likely to dominate (Fig. 3).

4.2. The target–flanker complexity contrast effect and the Bouma's law

Crowding is markedly reduced when the target and flankers are different in spatial complexity (Fig. 4). Such complexity contrast effect may occur only rarely in texts that use alphabets of uniform complexity, but is very common in texts like Chinese and Japanese. Therefore, the effective crowding in such texts may be lower than what predicted from an experiment using targets and flankers of the same complexity.

Bouma (1970) showed that when the center letter of a trigram is presented at an eccentricity E, the critical spacing (the center-to-center spacing between the target and flankers that produced the same acuity as an isolated letter) is roughly $0.5E$. This result has been elevated to the status of a law, which states that the spatial extent of crowding depends only on the retinal eccentricity of the target. Although the exact extent of critical spacing is known to depend on the criterion for threshold (Levi, 2008), once a criterion is set, Bouma's law would predict similar critical spacing for a given eccentricity regardless of the stimulus types and configurations. We found that the center-to-center critical spacing varies from $0.23E$ for Sloan letters to $0.37E$ for CC6 characters, the difference of which could be due to within-character crowding in complex CCs. Furthermore, we found that crowding and critical spacing are significantly reduced in the presence of target–flanker complexity contrast. The changeable critical spacing was also reported by Chung (2007) who demonstrated that critical spacing can be altered through training. These results suggest that retinal eccentricity is not the only variable that determines the spatial extent of crowding. Critical spacing may be influenced by multiple factors, and Bouma's law, as stated in its original form, may be a special case that is valid when stimuli are relatively simple and when the target and flankers share similar spatial complexity.

4.3. The mechanisms underlying crowding

Accumulating evidence from many crowding studies including our current one suggests that crowding may result from two main courses of visual processing. At an intermediate level, Levi et al. (2002) and Pelli et al. (2004) proposed that crowding results from improper integration of target and flanker features in the periphery. The null crowding effect of pixel-scrambled flankers (Fig. 6) is consistent with this account. In addition, the effect of stroke-scrambled flankers (Fig. 6) suggests that target–flanker feature integration is in some measure restricted by letter-level processing. Features are set free for integration with the target when this higher-level letter processing is interrupted, which aggravates crowding. Previous results (Chung et al., 2001; Hess et al., 2000; Kooi et al., 1994; Nazir, 1992) and our current evidence (Fig. 6) also indicated that target–flanker stimulus physical differences help segregate the target and flanker. This stimulus driven target–flanker segregation likely reduces crowding by restricting the target and flanker features to be integrated. This effect is similar to the case in center–surround interaction, in that when the surround and center stimuli are grouped into separate Gestalts, center–surround interaction is greatly weakened (Malania, Herzog, & Westheimer, 2007).

At higher visual processing, our results confirmed Strasburger's report that the observers more likely report a flanking stimulus as the target when a wrong response is made (Strasburger, 2005). The "same" and "diff" flanker effects shown in Fig. 6 indicate that crowding due to this misreporting could be corrected when the observers can separate the target and flanker stimuli through top-down influences. Strasburger explained his finding as dislocated attention to the flanker location. If this is true, the top-down influence could affect crowding by nullifying the positional uncertainty of attention. In addition, the same top-down influence could further facilitate target–flanker segregation initially driven by target–flanker physical differences, a possibility we cannot exclude.

A competing explanation of crowding against the improper feature integration model is that crowding could result from limited attentional resolution in the visual periphery (He, Cavanagh, & Intriligator, 1996; Intriligator & Cavanagh, 2001). The target becomes less legible when flankers are close because the attentional spotlight is not small enough to separate them. Although these two competing models typically make same predictions about crowding (Levi, 2008), the limited attentional resolution model would have difficulty predicting the stroke-scrambling effect since the spatial layout of the trigram stimuli is unchanged. However, our evidence is not necessarily against the attentional resolution model since the latter operates at a higher level of visual processing.

Acknowledgments

This research was supported by a Natural Science Foundation of China Grant (30725018) and a Changjiang Scholar Professorship (C.Y.), and by funds from the Beijing Normal University Project 111 (C.Y. & L.L.). We thank Dennis Levi for his helpful comments on this manuscript.

References

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*(5241), 177–178.

Chung, S. T. (2007). Learning to identify crowded letters: Does it improve reading speed? *Vision Research, 47*(25), 3150–3159.

Chung, S. T., Levi, D. M., & Legge, G. E. (2001). Spatial-frequency and contrast properties of crowding. *Vision Research, 41*(14), 1833–1850.

Flom, M. C. (1991). Contour interaction and the crowding effect. In R. P. Rutstein (Ed.). *Problems in optometry* (Vol. 3, pp. 237–257). Philadelphia, PA: JB Lippincott Company.

Flom, M. C., Heath, G. G., & Takahashi, E. (1963). Contour interaction and visual resolution: Contralateral effects. *Science, 142*, 979–980.

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383*(6598), 334–337.

Herse, P. R., & Bedell, H. E. (1989). Contrast sensitivity for letter and grating targets under various stimulus conditions. *Optometry and Vision Science, 66*(11), 774–781.

Hess, R. F., Dakin, S. C., & Kapoor, N. (2000). The foveal 'crowding' effect: Physics or physiology? *Vision Research, 40*(4), 365–370.

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology, 43*(3), 171–216.

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision, 8*(2), 255–279.

Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research, 48*(5), 635–654.

Levi, D. M., Hariharan, S., & Klein, S. A. (2002). Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *Journal of Vision, 2*(2), 167–177.

Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research, 25*(7), 963–977.

Ludvigh, E. (1941). Extrafoveal visual acuity as measured with Snellen Test-Letters. *American Journal of Ophthalmology*, 303–310.

Malania, M., Herzog, M. H., & Westheimer, G. (2007). Grouping of contextual elements that affect vernier thresholds. *Journal of Vision, 7*(2), 1. 1–7.

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), 58–70.

Nazir, T. A. (1992). Effects of lateral masking and spatial precueing on gap-resolution in central and peripheral vision. *Vision Research, 32*(4), 771–777.

Parish, D. H., & Sperling, G. (1991). Object spatial frequencies, retinal spatial frequencies, noise, and the efficiency of letter discrimination. *Vision Research, 31*(7–8), 1399–1415.

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision, 4*(12), 1136–1169.

Rovamo, J., & Virsu, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental Brain Research, 37*(3), 495–510.

Sheedy, J. E., Subbaram, M. V., Zimmerman, A. B., & Hayes, J. R. (2005). Text legibility and the letter superiority effect. *Human Factors, 47*(4), 797–815.

Strasburger, H. (2005). Unfocused spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision, 5*(11), 1024–1037.

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

Tian, M., Zhang, Y., Li, L., Zhang, C., & Mu, Y. (2005). An epidemiological investigation of age-related macular degeneration. *Journal of Lanzhou University (Medical Sciences), 31*(2), 70–71.

Tripathy, S. P., & Cavanagh, P. (2002). The extent of crowding in peripheral vision does not scale with target size. *Vision Research, 42*(20), 2357–2369.

Wolford, G., & Hollingsworth, S. (1974). Retinal location and string position as important variables in visual information processing. *Perception & Psychophysics, 16*, 437–442.

Zhang, J. Y., Zhang, T., Xue, F., Liu, L., & Yu, C. (2007). Legibility variations of Chinese characters and implications for visual acuity measurement in Chinese reading population. *Investigative Ophthalmology & Visual Science, 48*(5), 2383–2390.