Comparing the minimum spatial-frequency content for recognizing Chinese and alphabet characters

Hui Wang

Department of Biomedical Engineering,
University of Minnesota, Minneapolis, MN, USA
Present address: Athinoula A. Martinos Center for
Biomedical Imaging, Department of Radiology,
Massachusetts General Hospital and Harvard Medical
School, Charlestown, MA, USA ✉

Gordon E. Legge

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

Visual blur is a common problem that causes difficulty in pattern recognition for normally sighted people under degraded viewing conditions (e.g., near the acuity limit, when defocused, or in fog) and also for people with impaired vision. For reliable identification, the spatial frequency content of an object needs to extend up to or exceed a minimum value in units of cycles per object, referred to as the critical spatial frequency. In this study, we investigated the critical spatial frequency for alphabet and Chinese characters, and examined the effect of pattern complexity. The stimuli were divided into seven categories based on their perimetric complexity, including the lowercase and uppercase alphabet letters, and five groups of Chinese characters. We found that the critical spatial frequency significantly increased with complexity, from 1.01 cycles per character for the simplest group to 2.00 cycles per character for the most complex group of Chinese characters. A second goal of the study was to test a space-bandwidth invariance hypothesis that would represent a tradeoff between the critical spatial frequency and the number of adjacent patterns that can be recognized at one time. We tested this hypothesis by comparing the critical spatial frequencies in cycles per character from the current study and visual-span sizes in number of characters (measured by Wang, He, & Legge, 2014) for sets of characters with different complexities. For the character size (1.2°) we used in the study, we found an invariant product of approximately 10 cycles, which may represent a capacity limitation on visual pattern recognition.

Introduction

Character recognition is a prerequisite for reading and is typically a fast and accurate visual process. It becomes difficult under degraded visual conditions, such as reading small symbols at a long distance or with optical defocus, and is especially difficult in patients with severe low vision. The spatial-frequency properties of letter recognition have been widely explored. Previous studies show that the visual system utilizes a spatial frequency of 1–3 cycles per letter (CPL) for reliable identification (Alexander, Xie, & Derlacki, 1994; Chung, Legge, & Tjan, 2002; Ginsburg, 1978; Gold, Bennett, & Sekuler, 1999; Legge, Pelli, Rubin, & Schleske, 1985; Parish & Sperling, 1991; Solomon & Pelli, 1994), with the optimal spatial frequency depending somewhat on the angular size of letters (Majaj, Pelli, Kurshan, & Palomares, 2002). Kwon and Legge (2011) reported that accurate letter identification is possible with letters containing spatial frequencies only up to 0.9 CPL. These authors applied low pass filters to images of letters and faces and obtained psychometric functions showing recognition performance (percent correct) as a function of the cutoff frequency of the filters. They referred to the minimal spatial-frequency requirement for pattern recognition (with 80% accuracy) as the critical spatial frequency.

Chinese characters differ from alphabetic characters in having a wider range of pattern complexities. Studying Chinese character recognition may elucidate the connection between pattern recognition and pattern complexity. The goal of our study was to determine the critical-frequency requirements for Chinese characters, and to examine the effect of pattern complexity.

Critical cutoff frequencies can be expressed in both retinal spatial frequency (cycles per degree) or image-based spatial-frequency (cycles per character; CPC). In this paper, we will usually refer to spatial frequencies (including cutoff frequencies) in cycles per character. An exception will be our consideration of the effects of the contrast sensitivity function (CSF) in the Discussion.

Previous studies have shown that the acuity limit for recognizing Chinese characters with more strokes requires larger size (Cai, Chi, & You, 2001; Chi, Cai, & You, 2003; Huang & Hsu, 2005). Chinese characters with more strokes also have higher contrast thresholds (Yen & Liu, 1972) and longer response times (Yu & Cao, 1992). However, reports on the spatial frequency properties of Chinese character recognition are scarce. Chen, Yeh, and Lin (2001) adopted the critical-band–masking paradigm used by Solomon and Pelli (1994) to investigate the best central frequencies for Chinese characters. They tested Chinese characters with 3 to 21 strokes, and reported an average spatial frequency of approximately 8 CPC. The study however, did not take the variation of complexities into account, and did not investigate the minimal spatial-frequency requirements for Chinese character recognition.

In this study, we explored the critical spatial-frequency requirements for alphabet and Chinese characters, and examined the effect of complexity on these requirements. As the more complex characters have broader spatial-frequency spectra than the simple characters, they may require higher spatial frequency for character recognition. We divided alphabet characters and Chinese characters into categories, based on ranges of complexity values, using the perimetric complexity metric (Arnoult & Attneave, 1956; Pelli, Burns, Farell, & Moore-Page, 2006). The perimetric complexity of a symbol is defined as its perimeter squared divided by its "ink" area. We showed previously (Wang et al., 2014) that the perimetric complexity metric has high correlation with other complexity metrics, such as the number of strokes, the stroke frequency (Majaj et al., 2002; Zhang, Zhang, Xue, Liu, & Yu, 2007) and the skeleton method (Bernard & Chung, 2011). For each complexity category, we measured recognition performance for sets of 26 characters as a function of the cutoff frequency of low-pass filters.

A second goal of this study was to test an empirical hypothesis of a tradeoff between the critical frequency for character recognition and the visual span for character recognition; we term this the space-bandwidth invariance hypothesis. The visual span is the number of characters that can be recognized without moving the eyes. We have examined the size of the visual span for alphabet letters and Chinese characters, and discovered that the visual span size decreases as complexity increases (Wang et al., 2014). If critical frequencies are found to increase with complexity, it is possible that the product of critical frequency and visual-span size may be constant, representing a form of capacity limitation on visual pattern recognition. In the context of this paper, we refer to the bandwidth of the low-pass filter as the range from zero to the critical frequency. For simplicity, we used the term bandwidth instead of the critical frequency in our hypothesis.

The study of character recognition has important practical implications for reading performance. It is known that a critical frequency is required for uncompromised reading speed in alphabet reading (Kwon & Legge, 2012). Therefore, studying the spatial-frequency requirements for Chinese characters may be relevant to Chinese reading under low-resolution conditions including low vision. It may also have practical applications in designing reading material for difficult viewing conditions.

Methods

Subjects

Six college students (three men, three women) with normal or corrected-to-normal vision participated in the experiments. They were all native Chinese speakers, originally educated in the simplified Chinese script system, and all had more than 10 years education in English. The subjects signed an Internal Review Board (IRB) approved consent form before the experiments.

Stimulus sets

The stimulus characters were lowercase (LL) and uppercase (UL) alphabet letters in the Arial font, and simplified Chinese characters in the Heiti font in which all the strokes have the same width.

The 700 most frequently used Chinese characters (State Language Work Committee, 1992) were divided into five nonoverlapping groups based on their perimetric complexity values (Pelli et al., 2006). Twenty-six characters whose complexity values were close to the mean of the group were selected to form five sets of symbols (C1–C5). Characters with very high or low similarity were excluded from the stimulus sets. A measure of similarity for the characters in each set was computed using a normalized Euclidean distance method (Wang et al., 2014).

To determine whether subjects' familiarity with the characters affected their performance, we included a group of Chinese characters with lower usage frequency in text but comparable in complexity with characters in

LL	a	b	c	d	e
UL	A	B	C	D	E
C1	上	下	之	公	小
C2	生	对	行	去	名
C3	河	怎	低	独	诗
C4	眼	便	候	续	型
C5	情	最	题	量	德
C3'	曲	序	洛	纸	伟

Figure 1. Representative characters from the eight stimulus sets (LL, UL, C1–C5, and C3′). The complexity gradually increases in the first seven rows (from LL to C5). The bottom row (C3′) shows a group with comparable complexity to C3, but lower familiarity.

the group C3. We did this by identifying the next 700 most frequent Chinese characters and divided them into five complexity groups as well, based on the same complexity metric. Twenty-six characters were selected to comprise a comparison group (C3′), which had comparable complexity with C3 but lower frequency and presumably lower familiarity. The pattern complexity in the 1,400 most frequently used characters covers most of the complexity range across all simplified Chinese characters. Remaining characters with even higher complexities are rarely used in ordinary reading. Five representative characters from each stimulus set are shown in Figure 1. Statistics of the perimetric complexity values for each stimulus set are given in Table 1.

Low-pass filtering

A black character was generated on a gray background and stored as a grayscale image. The size of the image was 250×250 pixels, and the size of the characters (height of Chinese characters and x-height of alphabet letters) subtended $1.2°$ visual angle at a viewing distance of 40 cm. The image was blurred through a third order Butterworth low-pass filter (f) given by the following equation:

$$f = \frac{1}{1 + \left(\frac{r}{c}\right)^{2n}} \quad (1)$$

where r is the radius of the components in the frequency domain, c is the radius of the cutoff frequency, and n is the order of the filter. Figure 2A demonstrates the response function of the low-pass filter in the spatial-frequency domain.

To test the recognition accuracy as a function of blurring levels, six cutoff frequencies were selected for each stimulus set while character size remained constant. A demonstration of the characters with and without low-pass filtering is shown in Figure 2. The sets of filter cutoffs used for the eight complexity groups were chosen based on recognition performance in pilot runs. We ensured that the cutoffs were selected so that recognition accuracy spanned a wide range, and the psychometric function exhibited a clear transition from low to high performance accuracy. The cutoffs used for each stimulus set are summarized in Table 2.

Image display

The stimuli were displayed on a 19 in. CRT monitor (refresh rate: 75 Hz, resolution: 1280×960). The luminance of the blurred images on the screen was mapped onto 256 gray levels. The background of the image was set to the gray level 127, corresponding to a mean luminance of 40 cd/m^2. Luminance of the display monitor was made linear using an 8-bit lookup table in conjunction with photometric readings from a Konica Minolta CS-100 Chroma Meter (Konica Minolta Sensing Americas, Inc., Ramsey, NJ). The image luminance values were mapped onto the values stored in the lookup table for the display. The character image was displayed at the center of the screen. The stimulus symbol was created and controlled using MATLAB (MathWorks, Natick, MA) and Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007), running on a Mac Pro computer (Apple, Cupertino, CA).

Procedure

Each subject participated in three test sessions on three days. One session consisted of eight blocks: seven blocks with varied complexity levels (LL, UL, C1–C5),

Group	LL	UL	C1	C2	C3	C4	C5	C3′
Complexity mean (*SD*)	48.6 (11.7)	66.5 (17.9)	98.0 (6.3)	136.9 (2.3)	176.6 (4.3)	216.2 (5.0)	280.1 (33.7)	182.0 (5.2)

Table 1. Perimetric complexity measures for the stimulus sets. *Note*: LL, lowercase letter; UL, uppercase letter; C1–C5, five sets of Chinese characters from the simplest to the most complex; C3′, Chinese character group of comparable complexity with C3 but less familiarity.
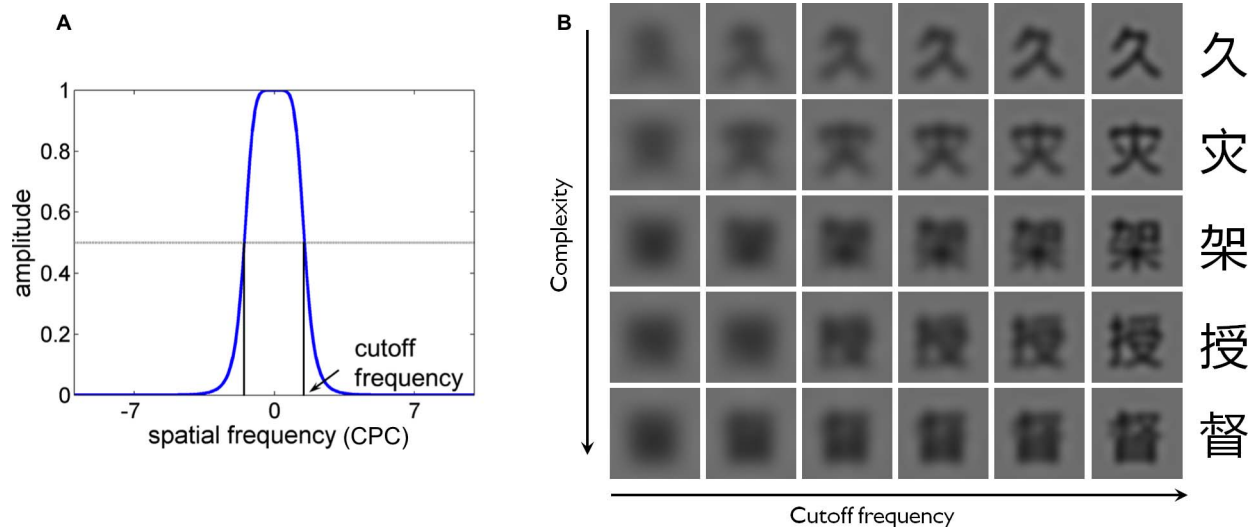
Figure 2. (A) The response function of the third-order Butterworth filter in the spatial frequency domain. The arrow indicates a cutoff frequency of 1.5 cycles per character (CPC) for a 1° letter size. The filter's cutoff is defined as the frequency at half amplitude. (B) Demonstration of low-pass filtered Chinese characters from the five complexity categories. The right column shows the unfiltered character.

and one block with complexity equivalent to C3 but lower character familiarity (C3′). In each block, there were 25 trials for each of six cutoffs forming a total of 150 trials. The stimulus symbol was randomly selected from the 26-character set, and the order of the cutoff frequencies presented was shuffled. The resulting psychometric functions for a given complexity category were therefore based on 450 trials (six cutoff frequencies and 75 trials per cutoff frequency). The orders of the blocks were counterbalanced between sessions and subjects.

The subject was shown the 26 unfiltered symbols on a hard copy page before the start of a block and urged to restrict responses to the stimulus set. During test trials, the subject was directed to fixate on a cross at the center of the screen. In each trial, a character was presented for 200 ms at fixation. After that, the display became uniform at the background level of 40 cd/m^2, and the subject was asked to report the character. The experimenter recorded the responses, and the subject clicked the mouse to start the next trial. A reference page was available, showing the 26 symbols in the current category, if the subject had trouble recalling the characters in the set. Subjects rarely responded with characters outside of the stimulus category (<1% of trials.) The 26 unfiltered characters were tested at the end of every block in order to evaluate the baseline performance for recognition. Performance on the unfiltered stimuli was at the ceiling value of 100%.

A chin rest was used during the test to reduce head movements and to maintain the viewing distance. Practice trials, including all the stimulus sets and the filter cutoffs, were provided at the beginning of the test.

Data analysis

The character recognition accuracy was plotted against the cutoff frequencies for each stimulus set. Cumulative Gaussian functions (Wichmann & Hill, 2001) were used to fit the plots with the least-square criterion. The critical spatial frequency was estimated from the psychometric function, and defined as the cutoff frequency yielding 80% correct responses. It is noted that the guessing level of the psychometric functions is $1/26 = 3.85\%$ for all the groups, because there are 26 stimuli in each complexity set. Figure 3

Group	f1	f2	f3	f4	f5	f6
LL	0.78	1.02	1.27	1.49	1.80	2.16
UL	0.78	1.02	1.27	1.49	1.80	2.16
C1	0.78	1.02	1.27	1.49	1.80	2.16
C2	0.92	1.18	1.42	1.63	1.94	2.34
C3/C3′	1.08	1.32	1.57	1.79	2.1	2.52
C4	1.24	1.44	1.73	1.94	2.28	2.66
C5	1.30	1.54	1.87	2.09	2.46	2.82

Table 2. Butterworth filter cutoff frequencies (in cycles per character; CPC) used for recognition tests with the seven complexity categories. *Note*: LL, lowercase letter; UL, uppercase letter; C1–C5, five sets of Chinese characters from the simplest to the most complex; C3′, Chinese character group of comparable complexity with C3 but less familiarity.
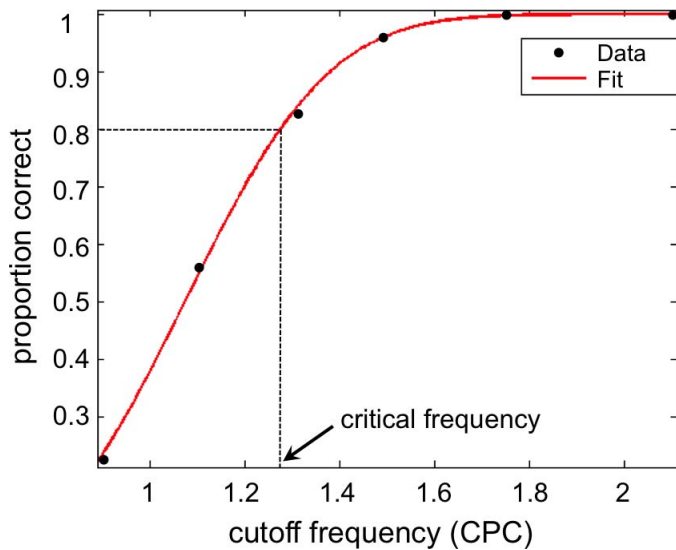
Figure 3. A sample psychometric function showing the recognition accuracy versus cutoff frequency (CPC) for C3 in one subject (black dots), and the cumulative Gaussian fit (red line). The critical spatial frequency is defined as the cutoff frequency yielding 80% correct responses.

demonstrates the data plot and the critical spatial-frequency estimation for stimulus set C3 in one subject. The fitting parameters (mean = alpha, variance = beta) of the underlying Gaussian function represent the x-axis location and the steepness of the psychometric function, respectively. One-way repeat-ed measures ANOVA tests were performed to investigate the effect of pattern complexity on the critical cutoff frequency, and fitting parameters alpha and beta, respectively.

Results

Critical spatial frequencies for alphabet and Chinese characters

Figure 4 shows psychometric functions (percent correct vs. filter cutoff frequency) for the six subjects and the group mean. Each panel shows functions for the seven complexity categories. For high cutoff frequencies, performance was at ceiling (100%). As the cutoff frequency decreased, a value was reached where performance declined rapidly.

As shown in the mean group data as well as the individual data, the filter cutoff frequency at which the response accuracy started to fall shifted to the right on the spatial-frequency axis as the complexity increased. Therefore, reliable identification of more complex characters requires inclusion of higher frequency components. Identification of the lowercase alphabet letters showed the largest tolerance to blur, followed by the uppercase letters, while Chinese character group C5 had the highest spatial-frequency requirement. The slope of the psychometric function was comparable among LL, UL, and C1–C3; however, it was lower in C4 and C5, implying that recognition improvement with higher frequency components is more gradual in complex characters.

We fitted each psychometric function with a cumulative Gaussian curve and estimated the critical spatial frequency for each stimulus set based on a criterion level of 80% correct. We found that the critical cutoffs increased with complexity (Figure 5), from 1.01 CPC for lowercase letters (LL) to 2.00 CPC for the most complex Chinese characters (C5). The critical
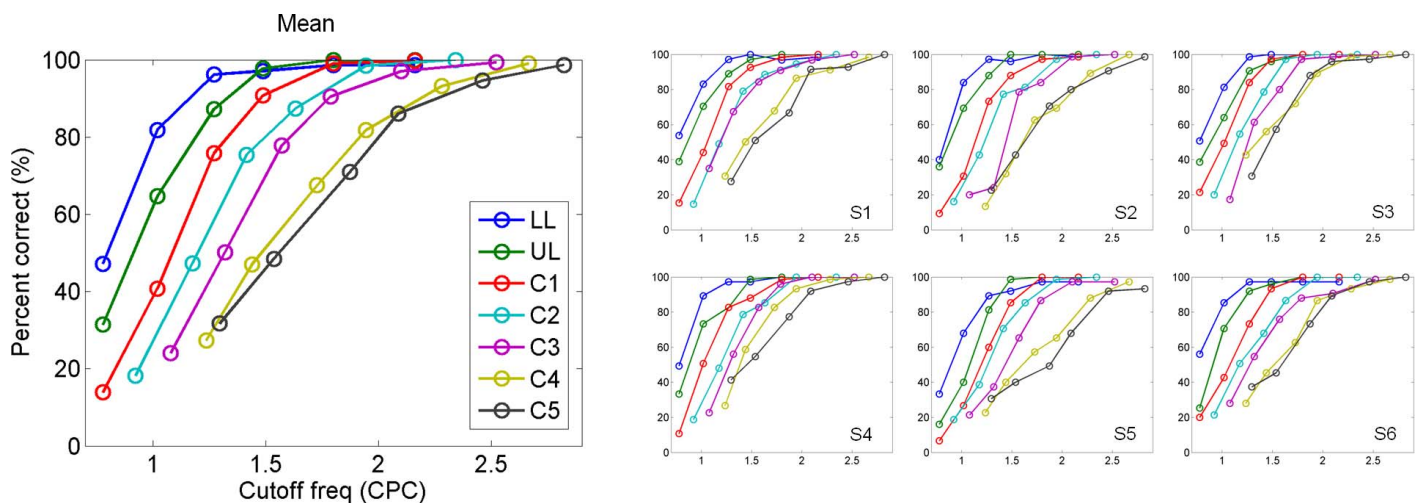

Figure 4. Psychometric functions. Plots of recognition accuracy (percent correct) versus cutoff frequency (cycles per character [CPC]) for the seven complexity groups (left: group mean; right: the individual data).
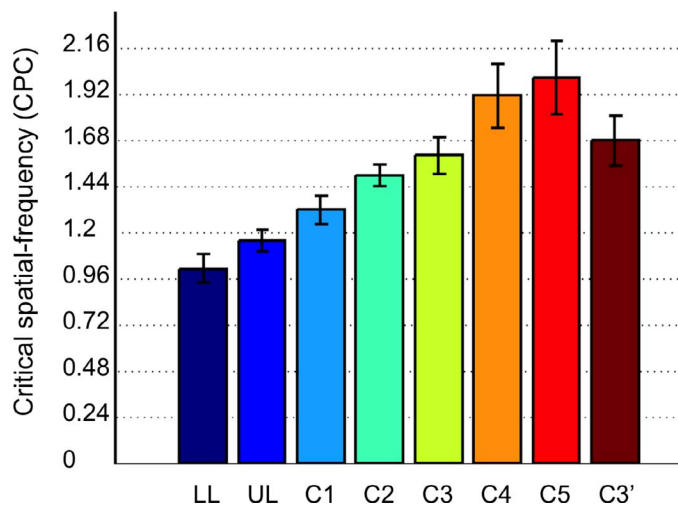
Figure 5. The critical spatial frequency (defined as the cutoff frequency at which recognition accuracy is 80% correct) for the seven complexity groups (LL, UL, C1–C5), and one group with lower character familiarity (C3′). Error bar: standard error (SE).

cutoff frequency for C3′ (1.67 CPC) was similar to that for C3 (1.60 CPC). To investigate the effect of character familiarity, we conducted an analysis of covariance on the critical frequencies of C3 and C3′, while using the perimetric complexity as a covariate because the complexity values for the two groups were not exactly equivalent and thus might confound the statistical test. The result indicated that character familiarity (within the range tested) did not play a significant role in determining the minimal spatial-frequency requirement for character recognition ($F = 2.15$, $p = 0.18$).

The one-way repeated measure ANOVA showed that there was a significant effect of complexity on the critical frequency, $F(6, 41) = 53.77$, $p < 0.001$, indicating that the critical cutoffs were significantly different among complexity groups. Further multiple comparison analysis indicated that the critical cutoff significantly increased between LL, C1, C3, and C5, and between UL, C2, and C4.

The location of the critical cutoffs on the psychometric functions can be influenced by two factors: a lateral shift of the curve along the spatial frequency axis (alpha), and/or a change in steepness of the curve (beta). Parameter estimation indicated that the major effects of complexity on the psychometric functions came from a shift along the spatial-frequency axis. The mean value (alpha) of the underlying Gaussian function gradually increased with complexity. A one-way ANOVA showed that there was a main effect of the complexity on alpha values, $F(6, 41) = 61.63$, $p < 0.001$. Multiple comparisons between groups revealed similar patterns as those for critical cutoffs. The alpha values were significantly different between LL, C1, C3, and C5, and also between UL, C2, and C4. A one-way ANOVA showed that there was also a main effect of the complexity on beta values, $F(6, 41) = 12.95$, $p < 0.001$. However, the beta value changed very little until the complexity levels became high. The beta values of C4 and C5 were significantly greater than the other less complex groups. The critical spatial frequencies and alpha and beta values for each stimulus set are summarized in Table 3 and the significance test results between groups are summarized in Table 4.

Test of the space–bandwidth invariance hypothesis

We used our data on critical frequencies for different complexity groups together with data from Wang et al. (2014) on visual spans for different complexity groups to test our space-bandwidth invariance hypothesis. According to this hypothesis, the product of critical spatial frequency (in units of cycles per character) and size of the visual span (in number of characters) should be invariant across changes in complexity.

Wang et al. (2014) measured visual spans with a trigram paradigm, three characters side by side, in four complexity groups LL, C1, C3, and C5. On each trial, a

Group	Human subject: Mean (SE)			Model observer		
	Critical cutoff (CPC)	Alpha	Beta	Critical cutoff (CPC)	Alpha	Beta
LL	1.01 (0.03)	0.80 (0.02)	0.25 (0.02)	0.91	0.71	0.24
UL	1.16 (0.02)	0.92 (0.03)	0.28 (0.02)	1.17	0.91	0.31
C1	1.31 (0.03)	1.08 (0.02)	0.27 (0.01)	1.44	1.10	0.40
C2	1.48 (0.02)	1.21 (0.01)	0.33 (0.01)	1.60	1.23	0.43
C3	1.60 (0.04)	1.32 (0.03)	0.33 (0.02)	1.64	1.27	0.44
C4	1.91 (0.07)	1.51 (0.04)	0.48 (0.03)	1.87	1.47	0.48
C5	2.00 (0.08)	1.56 (0.04)	0.52 (0.04)	2.01	1.52	0.58
C3′	1.67 (0.05)	1.39 (0.03)	0.34 (0.02)	1.69	1.33	0.43

Table 3. Critical spatial frequency and the fitting parameters for psychometric functions of human subjects and noise-limited human contrast sensitivity function model. *Note*: LL, lowercase letter; UL, uppercase letter; C1–C5, five sets of Chinese characters from the simplest to the most complex; C3′, Chinese character group of comparable complexity with C3 but less familiarity.

	LL	UL	C1	C2	C3	C4	C5
LL		−	*	*	*	*	*
		−	*	*	*	*	*
		−	−	−	−	*	*
UL	−		−	*	*	*	*
	−		−	*	*	*	*
	−		−	−	−	*	*
C1	*	−		−	*	*	*
	*	−		−	*	*	*
	−	−		−	−	*	*
C2	*	*	−		−	*	*
	*	*	−		−	*	*
	−	−	−		−	*	*
C3	*	*	*	−		−	*
	*	*	*	−		−	*
	−	−	−	−		*	*
C4	*	*	*	*	−		−
	*	*	*	*	−		−
	*	*	*	*	*		−
C5	*	*	*	*	*	−	
	*	*	*	*	*	−	
	*	*	*	*	*	−	

Table 4. Multiple comparisons between groups for critical spatial frequency and fitting parameters of the psychometric function (alpha and beta). *Note*: * represents a significant difference and − represents no significant difference. In each cell, Row 1 represents the result for critical spatial frequency, Row 2 for alpha, and Row 3 for beta. *Note*: LL, lowercase letter; UL, uppercase letter; C1–C5, five sets of Chinese characters from the simplest to the most complex; C3′, Chinese character group of comparable complexity with C3 but less familiarity.

trigram was briefly presented at one of 17 locations left or right of fixation, and the subject was asked to report the three characters in the trigram. A visual span profile was constructed, showing percent correct character recognition as a function of character position relative to fixation. The size of the visual span was defined as the width (number of characters) of the profile at a criterion accuracy of 80% correct. It was found that the visual-span size significantly decreased with complexity, from 10.5 characters for LL, to 4.5 characters for C5.

There were three subjects who participated in both the visual span study (Wang et al., 2014) and the current study. We used the two sets of data—visual-span size and critical frequency—from these three subjects to test the space-bandwidth invariance hypothesis. The analysis was based on data from complexity groups LL, C1, C3, and C5 because these were the stimulus categories for which both critical frequencies and visual-span measurements were available. Figure 6A shows the product of the visual-span size and the critical spatial frequency at the four complexity levels for individual subjects and their group mean. The plots were nearly flat across complexity, indicating the presence of an invariant space-bandwidth product. The average product was 9.9 cycles across the four complexity groups. A one-way repeated measure ANOVA showed that the products were not significantly different among the four complexity groups, $F(3, 11) = 0.45$, $p = 0.72$.

Wang et al. (2014) also estimated the size of visual spans for their subjects when mislocation errors were discounted. Mislocations occur when subjects report the correct identity of characters in trigrams, but in the wrong spatial order. When mislocations are discounted, the primary constraint determining the size of the visual span is crowding, and the resulting visual-span size is larger (Wang et al., 2014; He et al., 2015). Figure 6B shows the space-bandwidth results for the three subjects when these modified visual-span sizes were used. In this case, the products yield an average of 11.6
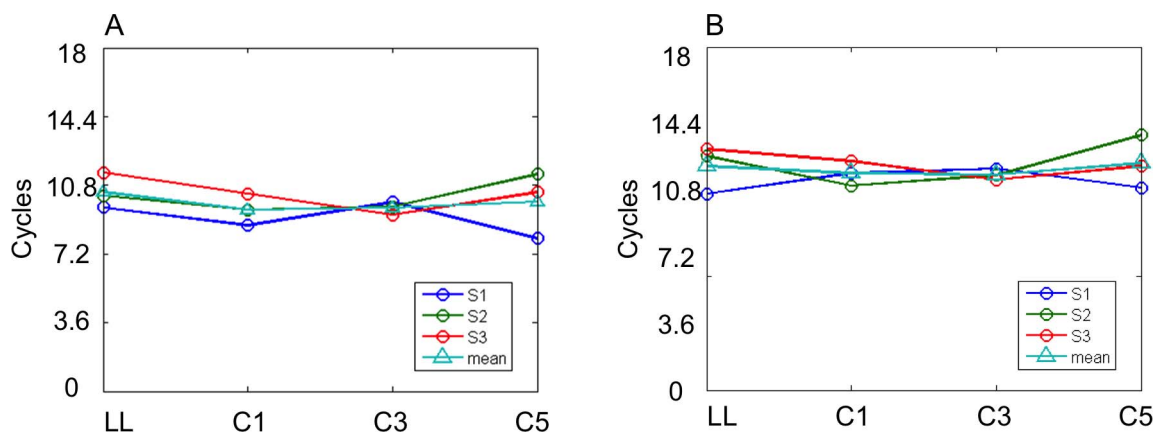

Figure 6. Plots of the space-bandwidth product for three subjects and the group mean at four complexity levels (LL, C1, C3, and C5). Space refers to the number of characters in the visual span and bandwidth refers to the critical cutoff spatial frequency (80% correct criterion) in cycles per character (CPC). (A) The visual-span size was computed without allowing mislocation errors. (B) The visual-span size was computed allowing mislocation errors.
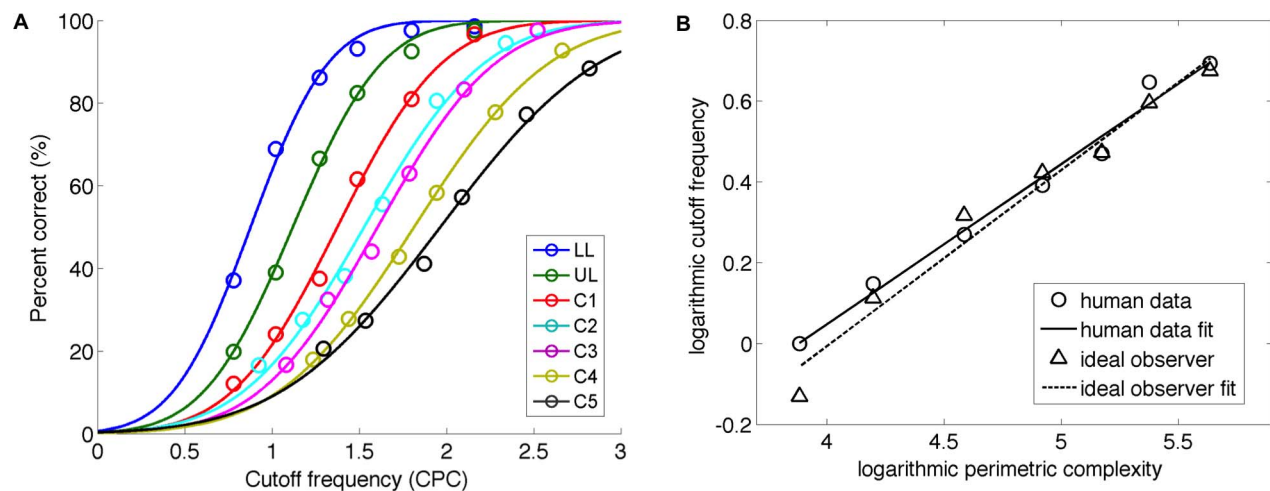
Figure 7. Critical spatial frequency for the seven complexity groups obtained with the noise-limited CSF model. (A) Psychometric functions of the model observer (circle) and cumulative Gaussian fitting (solid line). (B) Comparison of the critical spatial frequencies between human subjects and the model observer. The original data (circle: human subjects; triangle: model observer) and the fitting for logarithmic critical frequency as a function of logarithmic perimetric complexity (solid line: human subjects; dashed line: model observer) are shown.

cycles across the complexity groups. Similar to Figure 6A, we observed nearly flat plots for all three subjects and the mean data. The one-way repeated measure ANOVA indicated that the products were not significantly different among the four complexities, $F(3, 11) = 0.25$, $p = 0.86$.

Noise-limited contrast sensitivity function model for estimating the critical spatial frequency

We used a noise-limited contrast sensitivity function (CSF) model, described by Kwon and Legge (2011), to simulate the recognition process for both alphabet and Chinese characters. Briefly, a low-pass filtered image entered the model as an input, passed through a CSF filter with additive white noise, and then reached an optimal classifier for a decision. The decision rule maximized the recognition accuracy, based on maximum a posteriori probability of the input being a particular target (Green & Swets, 1966; Tanner & Birdsall, 1958). The input images were passed through the same filters and cutoff frequencies as used in the human task. The CSF filter was a linear filter fitted to a human CSF at the fovea (Chung & Tjan, 2009). The filtering was conducted with image frequencies expressed in cycles per degree (rather than cycles per character) so that the CSF filter would weight the spectral components of the stimuli according to the contrast sensitivity at different retinal spatial frequencies for human foveal vision. Gaussian luminance white noise (zero mean) was added to the output of the CSF-filtered stimulus. The standard deviation of the noise

was fixed at 0.45, which was a relative value assuming the image contrast of unfiltered characters was 1. We tuned the standard deviation of noise for each complexity group first to match human performance, and used the mean standard deviation from the seven complexity groups as a fixed noise level to input to the model.

We simulated character recognition for the seven complexity levels (LL, UL, C1–C5) with the goal of determining the model's critical cutoff frequency. For each set of characters in a given complexity category, the performance was tested with six cutoff frequencies for the low-pass filters, the same as the human experiments. There were 700 simulation trials for each cutoff frequency, for a total of 4,200 trials per psychometric function.

The recognition accuracy was plotted as a function of cutoff frequency, and fitted with the same cumulative Gaussian function as used for obtaining the psychometric function of human observers (Figure 7A). We estimated the critical spatial frequency using the same criteria as for the human observers (80% accuracy). The critical cutoff frequency for the model observer increased with complexity, from 0.91 CPC for the simplest group of LL to 1.96 CPC for the most complex group of C5 (Table 3). The results are very similar to the critical cutoff frequencies obtained from human observers: 1.01 CPC for the simplest group of LL to 2 CPC for the most complex group of C5. To compare the performance of the noise-limited CSF model with the human subjects, we fitted the critical spatial frequency as a function of the perimetric complexity for both sets and obtained a log-log slope of 0.40 and 0.44, respectively (Figure 7B). The parameters

(alpha and beta) of the cumulative Gaussian fit for the model observer are shown in Table 3 as a comparison with those obtained for the human observer. Both the mean for the location of the performance curve on the spatial frequency axis and the standard deviation for the slope of the recognition accuracy show an agreement between the model observer and the human observer. Therefore, the noise-limited CSF model appears to predict the impact of blur on the recognition of characters with different complexities.

Discussion

In this study, we measured the minimal spatial-frequency requirements for alphabet and Chinese character recognition by low-pass filtering the character images with different cutoff frequencies. We asked whether the critical spatial frequencies for Chinese characters are different from those for alphabet letters, and how pattern complexity influences the critical spatial-frequency requirements for recognition. We found that the minimum frequencies for recognizing Chinese characters were significantly higher than those for letters, except for the groups of uppercase letters and the simplest Chinese characters C1. As the perimetric complexity of characters increased from the least complex category lowercase letters (LL), to the most complex category of Chinese characters (C5), the critical cutoff frequency doubled, from 1.01 cycles per character (CPC) to 2.00 CPC.

Factors accounting for the dependence of critical spatial frequency on character complexity

The increase of critical spatial frequency with complexity for character recognition is consistent with both a local feature identification theory (Hubel & Wiesel, 1962) and a Fourier component theory (Campbell & Robson, 1968). Characters are represented by an arrangement of oriented lines and curves within a defined area. The more complex the characters are, the higher the density of the features within the character. Assuming the fine features are required to distinguish among the characters, the fine features must remain legible, meaning that the characters are less tolerant to blur. Similarly, the complex characters have a broader spectrum in the spatial-frequency domain. To recover character identity from the blurry images, higher spatial-frequency components would need to be retained. As a consequence, we would expect that more complex characters would require higher critical frequencies.

The difference in minimal spatial-frequency requirements for lowercase and uppercase letters has been found previously. Kwon and Legge (2011) measured the critical frequencies for letter recognition by native English speakers. They reported the critical frequencies of 0.9 CPL and 1.14 CPL for lowercase and uppercase letters, respectively, in central vision. In the current study, the results were similar for bilingual native Chinese speakers: 1.01 CPC for lowercase and 1.16 CPC for uppercase letters.

Another factor that may contribute to the frequency requirements for character recognition is the pattern similarity of characters in the eligible set. The pattern similarity between two images can be defined in terms of the Euclidean distance between the symbols in feature space (such as the grayscale values of pixels). Greater distance indicates less similarity between symbols. Kwon and Legge (2011) invoked an explanation based on pattern similarity to account for the small difference in critical frequencies for lowercase and uppercase alphabet letters. Wang et al. (2014) showed that pattern similarity of Chinese characters increases with complexity. In order to distinguish between more similar patterns, the visual system requires access to the fine features preserved by high-frequency components in the spectra of the characters. Therefore, consideration of pattern similarity plausibly predicts a higher spatial-frequency requirement for identifying more complex characters.

The minimal spatial frequency is a way to examine the spatial resolution requirements for pattern recognition, and may relate to visual acuity. We assume that acuity is limited by fitting the required spatial-frequency content into the contrast sensitivity curve of human vision. In our study, the critical spatial frequencies were measured at much larger size than the acuity limit. However, if the same critical frequencies (in cycles per character) apply at the acuity limit, we would expect that the acuity size of characters would scale in proportion to the critical spatial-frequency requirements. For instance, symbols requiring a critical frequency of 2 CPC should have an acuity size twice that of symbols requiring 1 CPC. If this is the case, the size of acuity characters should increase with their pattern complexity. This expectation is supported by the legibility studies of Chinese characters. Zhang et al. (2007) examined the psychometric functions for recognition of simplified Chinese characters as a function of angular character size. The characters with 2–18 strokes were divided into six complexity levels, based on a stroke frequency metric (i.e., the number of strokes intersected by a line through the letter width). They found that the critical size linearly increased with the stroke frequency by a factor of 1.28 from the simplest to the most complex group. In another study, Huang and Hsu (2005) assessed the minimal size

requirements for recognizing traditional Chinese characters. The characters consisting of 3–27 strokes were divided into five groups based on the number of strokes. The subjects were asked to read a character string under normal reading conditions. Huang and Hsu (2005) estimated the minimal legible sizes for each complexity group, and found a systematic increase with character strokes. The relative enlargement of the acuity sizes reported in the two studies converged to a factor of approximately 1.3 from the 4-stroke group to the 15-stroke group. Watson and Ahumada (2008) described a template model of visual acuity based on an ideal observer limited by optical filtering, neural filtering and noise. Using this model, they predicted that the acuity size for optotypes varied with complexity, and the model showed a good match with human data in low- and medium-complexity optotypes (Watson & Ahumada, 2012). In our study, we found a 1.5 times increase of the critical spatial frequency from the least-complex to the most-complex Chinese characters, which is close to the above-cited scaling factors in acuity size. Therefore, the critical spatial-frequency requirements we have found may apply to acuity limits in recognizing Chinese characters and other complex symbols.

The minimal spatial-frequency requirement and the optimal frequency channel for character recognition

Our study determined the minimum cutoff frequency of a low-pass filter required for character recognition. The results should be distinguished from measures of the optimal frequency band, which is thought to be the frequency channel people used for symbol recognition. The optimal frequency studies typically use a band-limited filter with a noise mask, and optimality is defined as the center frequency that yields the best contrast sensitivity or efficiency for recognition (Gold et al., 1999; Majaj et al., 2002; Solomon & Pelli, 1994). In letter recognition, the optimal frequency is typically found to be 1–3 CPC. Lo (2013), in his doctoral dissertation, used a noise-masking method to estimate the optimal frequencies for Chinese character recognition. One hundred and fifty Chinese characters were selected from a character usage frequency database, and the characters were divided into three complexity groups based on the perimetric complexity metric. He found that the optimal frequencies for the three groups with increased complexity were 4.2, 4.8, and 5.4 CPC, respectively. Rather than focusing on measurements of the optimal frequency, the goal of our study was to determine the minimal frequency requirements for character recognition. In most cases, the optimal frequencies were entirely missing due to the low-pass
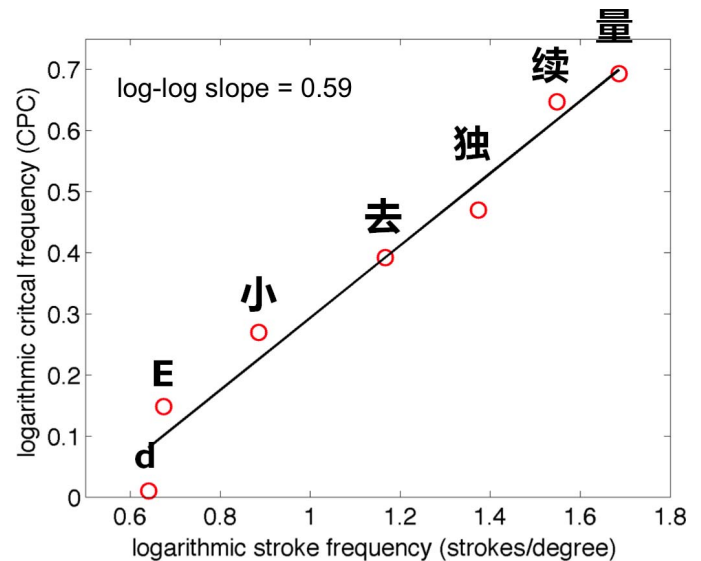

Figure 8. The relationship between critical spatial frequency (CPC) and stroke frequency (strokes per degree) defined by Majaj et al. (2002) for all the complexity groups (LL, UL, and C1–C5). The fitting shows a log-log slope of 0.59 (solid line). Mean data values from the current study are plotted as circles.

filtering. Subjects were forced to rely on information carried by low frequencies to identify the characters. We found that visual perception can recover character identity at surprisingly low frequencies when deprived of the optimal frequencies for recognition.

Majaj et al. (2002) conducted a systematic investigation of the effect of complexity on the optimal frequency channels for character recognition, and reported that for characters with different complexity, the optimal frequency was related to the stroke frequency (strokes per degree) by a power law with exponent 2/3 (i.e., having a log-log slope of 2/3). This means that the optimal frequency for a character, in cycles per character, is 1.59 times higher for characters that are 2 times larger, and for characters with 2 times more strokes for a given size (i.e., greater complexity). To determine whether the critical cutoff frequencies in our study follow a similar law to the optimal frequencies found by Majaj et al. (2002), we plotted the critical spatial frequency as a function of the stroke frequency of our stimuli—the metric used in Majaj et al. (2002) study. We obtained a log-log slope of 0.59 (Figure 8). This result was similar to the log-log slope of 2/3 reported by Majaj et al. (2002). This result indicates that the minimum spatial frequency for character recognition has a similar dependence on character complexity as the optimal frequency.

The minimal spatial-frequency requirements for character recognition are relevant to reading performance. In alphabet reading with low-pass filtered texts, reading speed rapidly drops after the filter frequency decreases below a certain value, referred to as the

critical frequency for reading (Kwon & Legge, 2012; Legge et al., 1985). Similar constraints may apply to Chinese reading. A complicating factor is that character complexity varies substantially within sentences in Chinese text. The critical frequency for reading Chinese text under blur might be determined by the most complex characters in the text; however, it may also depend on the context such as, for instance, the position of the complex characters in the sentence. Future investigations are needed to determine the spatial-frequency requirements for reading Chinese text and their relationship to character recognition.

The space-bandwidth invariance hypothesis

Finally, our data support the hypothesis of a form of space-bandwidth product invariance. In this case, *bandwidth* refers to the range from zero to the minimum cutoff of low-pass frequencies required for a criterion level of character recognition, for which we have used the term *critical frequency* in units of cycles per character (CPC). *Space* refers to the size of the visual span for character recognition; that is, the number of adjacent characters that can be recognized above a criterion level of accuracy without moving the eyes. The *invariance* refers to the empirical observation that as the perimetric complexity of sets of characters increases, the critical frequencies increase and the visual spans decrease, but their product (critical frequency in CPC × visual span in number of characters) remains constant. We acknowledge that our evidence is limited to a particular character size and three bilingual Chinese subjects for whom we have data on both critical frequencies and visual-span size for groups of characters with a wide range of complexity. However, data from the three subjects consistently support a constant space-bandwidth product of 9.9 cycles (or 11.6 cycles when mislocations were discounted in the visual-span measurements) across alphabet and Chinese scripts for the character size we used in the two studies.

This invariance might be expected to hold over a range of character sizes. Majaj et al. (2002) found that the optimal channel frequency (CPC) remained constant across size for band-limited (blurry) characters. Similarly, Legge et al. (1985) found that the minimal spatial frequency (CPC) for reading was size-independent over a wide range. Given that critical frequencies for character recognition are closely related to critical frequencies for reading (Kwon & Legge, 2012), we expect the critical frequencies for character recognition, measured as CPC, to remain constant over a wide range of character sizes. It is also known that the visual-span size (measured as the number of recognizable characters) is invariant over a wide range of character sizes for alphabet characters (Legge et al.,

2007). It seems likely, therefore, that the product of critical frequency and visual span—our *space-bandwidth* product—would also remain constant over a wide range of character sizes.

The invariance of the space-bandwidth product may indicate a capacity limitation on visual pattern recognition in one fixation. Template matching theories of pattern recognition rely on comparisons between the input sensory signals and the stored features in a set of templates. We speculate that there may be an upper bound on the number of features that can be perceived and processed simultaneously. As the complexity of the symbols increases, a higher frequency bandwidth is required, with fewer characters recognized before the maximum feature limit is reached.

It is possible that the space-bandwidth constraint we have observed is related to an information processing constraint early in the visual pathway. Marcelja (1980) pointed out the relevance of Gabor's (1946) theoretical analysis to human vision. Gabor (1946) showed that for a time-varying signal, the product of sample duration and frequency resolution has a lower bound. This is an uncertainty constraint on information. Gabor showed that Gabor functions (Gaussian-windowed sign waves) optimize this tradeoff by meeting the criterion of maximum information transmitted (minimal uncertainty) for a signal. Marcelja (1980) followed by Daugman (1985) extended Gabor's analysis to spatial dimensions and applied the analysis to simple-cell receptive fields in the visual cortex. They argued that sampling of visual signals by Gabor-like functions in receptive fields represent an optimal encoding of visual information for space and bandwidth; therefore, the minimal uncertainty criterion is satisfied by simple cells. The space-bandwidth constraint on pattern recognition of alphabet and Chinese characters in our studies might reflect an optimal encoding mechanism in early visual processing, which satisfies the minimum uncertainty criterion.

Acknowledgments

Commercial relationships: none.
Corresponding author: Hui Wang.

Email: hwang47@mgh.harvard.edu.
Address: Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital and Harvard Medical School, Charlestown, MA, USA.

References

Alexander, K. R., Xie, W., & Derlacki, D. J. (1994). Spatial-frequency characteristics of letter identification. *Journal of the Optical Society of America*, *11*, 2375–2382.

Arnoult, M. D., & Attneave, F. (1956). The quantitative study of shape and pattern perception. *Psychological Bulletin*, *53*, 452–471.

Bernard, J.-B., & Chung, S. T. L. (2011). The dependence of crowding on flanker complexity and target-flanker similarity. *Journal of Vision*, *11*(8):1, 1–16. doi:10.1167/11.8.1. [PubMed] [Article]

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Cai, D., Chi, C.-F., & You, M. (2001). The legibility threshold of Chinese characters in three type styles. *International Journal of Industrial Ergonomics*, *27*, 9–17.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology (London)*, *197*, 551–566.

Chen, I.-P., Yeh, S.-L., & Lin, S. (2001). Critical phase band for Chinese character recognition. *Chinese Journal of Psychology*, *43*, 155–173.

Chi, C. F., Cai, D., & You, M. (2003). Applying image descriptors to the assessment of legibility in Chinese characters. *Ergonomics*, *46*, 825–841.

Chung, S. T. L., Legge, G. E., & Tjan, B. S. (2002). Spatial-frequency characteristics of letter identification in central and peripheral vision. *Vision Research*, *42*, 2137–2152.

Chung, S. T., & Tjan, B. S. (2009). Spatial-frequency and contrast properties of reading in central and peripheral vision. *Journal of Vision*, *9*(9):16, 1–19. doi:10.1167/9.9.16. [PubMed] [Article]

Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A*, *2*, 1160–1169.

Gabor, D. (1946). Theory of communication. *Journal of the Institute of Electrical Engineers*, *93*, 429–457.

Ginsburg, A. P. (1978). Visual information processing based on spatial filters constrained by biological data (Vols. I and II). Ph.D. Thesis. Cambridge University, Aerospace Medical Research Laboratory Report AMRL-TR-78_129. United States Air Force, Wright-Patterson AFB, OH.

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Identification of bandpass filtered letter and faces by human and ideal observers. *Vision Research*, *39*, 3537–3560.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

He, Y., Scholz, J. M., Gage, R., Kallie, C. S., Liu, T. & Legge, G. E. (2015). Comparing the visual spans for faces and letters. *Journal of Vision*, *11*(8):7, 1–13, doi:10.1167/15.8.7. [PubMed] [Article]

Huang, K. C., & Hsu, S. H. (2005). Effects of number of strokes on Chinese character recognition during a normal reading condition. *Perceptual and Motor Skills*, *101*, 845–852.

Hubel, D. G., & Wiesel, T. N. (1962). Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology (London)*, *160*, 106–154.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, *36*, 1–16.

Kwon, M., & Legge, G. E. (2011). Spatial-frequency cutoff requirements for pattern recognition in central and peripheral vision. *Vision Research*, *51*, 1995–2007.

Kwon, M., & Legge, G. E. (2012). Spatial-frequency requirements for reading revisited. *Vision Research*, *62*, 139–147.

Legge, G. E., Cheung, S. H., Yu, D., Chung, S. T., Lee, H. W., & Owens D. P. (2007). The case for the visual span as a sensory bottleneck in reading. *Journal of Vision*, *7*(2):9, 1–15. doi:10.1167/7.2.9. [PubMed] [Article]

Legge, G. E., Pelli, D. G., Rubin, G. S., & Schleske, M. M. (1985). Psychophysics of reading. I. Normal vision. *Vision Research*, *25*, 239–252.

Lo, O. T. (2013). Chinese character recognition—Studies of complexity effect on recognition efficiency, spatial frequency characteristics, crowding and expertise. (Unpublished doctoral dissertation). University of Hong Kong, Hong Kong.

Majaj, N. J., Pelli, D. G., Kurshan, P., & Palomares, M. (2002). The role of spatial frequency channels in letter identification. *Vision Research*, *42*, 1165–1184.

Marcelja, S. (1980). Mathematical description of the

responses of simple cortical cells. *Journal of the Optical Society of America*, *70*, 1297–1300.

Parish, D. H., & Sperling, G. (1991). Object spatial frequencies, retinal. *Vision Research*, *31*, 1399–1416.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, *46*, 4646–4674.

Solomon, J. A., & Pelli, D. G. (1994). The visual filter mediating letter identification. *Nature*, *369*, 395–397.

State Language Work Committee, Bureau of Standards. (1992). *Modern Chinese character frequency table*. Beijing, China: Beijing Language Press.

Tanner, W. P., Jr., & Birdsall, T. G. (1958). Definitions of d' and η as psychophysical measures. *Journal of the Acoustical Society of America*, *30*, 922–928.

Wang, H., He, X., Legge, G. E. (2014). Effect of pattern complexity on the visual span for Chinese and alphabet characters. *Journal of Vision*, *14*(8):6, 1–17. doi:10.1167/14.8.6. [PubMed] [Article]

Watson, A. B., & Ahumada, A. J. (2008). Predicting visual acuity from wavefront aberrations. *Journal of Vision*, *8*(4):17, 1–19. doi:10.1167/8.4.17. [PubMed] [Article]

Watson, A. B., & Ahumada, A. J. (2012). Modeling acuity for optotypes varying in complexity. *Journal of Vision*, *12*(10):19, 1–19. doi:10.1167/12.10.19. [PubMed] [Article]

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling and goodness-of-fit. *Perception and Psychophysics*, *63*, 1293–1313.

Yen, J. S., & Liu, I. M. (1972). Factors affecting recognition thresholds of Chinese characters. *Chinese Journal of Psychology*, *14*, 113–117.

Yu, B., & Cao, H. (1992). A new exploration on the effect of stroke number in the identification of Chinese characters. *Acta Psychologica Sinica*, *24*, 120–126.

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