PERGAMON

Vision Research 42 (2002) 2137–2152

Vision Research

www.elsevier.com/locate/visres

# Spatial-frequency characteristics of letter identification in central and peripheral vision

Susana T.L. Chung [a,*], Gordon E. Legge [b], Bosco S. Tjan [c]

[a] College of Optometry, University of Houston, Houston, TX 77204, USA
[b] Department of Psychology, University of Minnesota, Minneapolis, MN 55455, USA
[c] Department of Psychology, University of Southern California, Los Angeles, CA 90089, USA

Received 23 November 2001

**Abstract**

Spatial-frequency characteristics of letter identification are much better understood in the fovea than in the periphery. The purpose of this study was to compare the spatial-frequency characteristics of letter identification in central and peripheral vision. We measured contrast thresholds for identifying single, Times-Roman lower-case letters that were spatially band-pass filtered. Each of the 26 letters was digitally filtered with a set of nine cosine log filters, with peak object spatial frequencies ranging from 0.63 to 10 c/letter, in half-octave steps. Bandwidth of the filters was 1 octave. Three observers with normal vision were each tested monocularly at the fovea, and at 5° and 10° in the inferior visual field. Letter sizes were 0.2, 0.4 and 0.6 log units larger than high contrast, unfiltered acuity letters. Plots of contrast sensitivity for letter identification vs. frequency of the band-pass filters exhibit spatial tuning. In general, the spatial-frequency characteristics of letter identification are fundamentally identical between central and peripheral vision. These characteristics include the scaling of the peak frequency of the spatial-tuning functions with letter size and the bandwidth of the tuning functions. The only difference between the fovea and the periphery is that for the same physical letter size, peak sensitivity of the spatial-tuning functions occurs at a higher retinal frequency at the fovea than in the periphery. To test whether or not the contrast sensitivity function (CSF) can account for the differences in the spatial-frequency characteristics of letter identification between central and peripheral vision, we incorporated a human CSF into an ideal-observer model, and tested the performance of this ideal-observer on the same letter identification task used with the human observers. Data from this CSF-ideal-observer resemble closely those of human observers, suggesting that the spatial-frequency characteristics of human letter identification can be accounted for by the CSF and the letter-identity information, without invoking selection among narrow-band spatial-frequency channels.


*Keywords:* Letter identification; Peripheral vision; Spatial-frequency channel; Contrast sensitivity

## 1. Introduction

Reading in peripheral vision is slow and inefficient, even when character size is not a limiting factor (Latham & Whitaker, 1996; Chung, Mansfield, & Legge, 1998), and when oculomotor demands are minimized with rapid serial visual presentation (RSVP) (e.g. Rubin & Turano, 1994; Latham & Whitaker, 1996; Chung et al., 1998). For instance, using RSVP, maximum reading speed decreases from 862 wpm at the fovea to 143 wpm at 20° eccentricity (Chung et al., 1998). Millions of people who lose their central vision due to diseases such as age-related macular degeneration have to rely on their peripheral vision to read. Therefore, the understanding of why reading is slower in peripheral vision is of utmost importance to visual rehabilitation of these people.

Letters are the fundamental building blocks of text. Differences in the spatial coding of letters in central and peripheral vision may help us understand why reading in peripheral vision is slow. The primary goal of this study was to compare the early visual coding of letters, specifically, the spatial-frequency characteristics of letter identification, in central and peripheral vision.

Data from prior studies in central vision are consistent with the view that a single narrow-band spatial-frequency channel mediates the identification of

* Corresponding author.
*E-mail address:* schung@optometry.uh.edu (S.T.L. Chung).

0042-6989/02/$ - see front matter 
PII: S0042-6989(02)00092-5

letters of a given size. Based on these studies, the spatial representation of letters can be summarized by three characteristics: (1) the spatial frequency of the most sensitive channel for letter identification, to be termed the "channel frequency"; (2) the dependence of this channel frequency on letter size and (3) the bandwidth of the channel. In general, the channel frequency falls within a range of object spatial frequency [1] of 1–3 c/letter (Ginsburg, 1980; Parish & Sperling, 1991; Solomon & Pelli, 1994; Alexander, Xie, & Derlacki, 1994; Majaj, Pelli, Kurshan, & Palomares, 2002). Earlier studies (Parish & Sperling, 1991; Solomon & Pelli, 1994) have suggested that channel frequency, expressed as retinal frequency, scales with letter size, or equivalently, channel frequency, expressed as object frequency, is constant for different letter sizes. The bandwidth of these channels, representing the frequency-selectivity of the channels, is found to be about 2 octaves (Solomon & Pelli, 1994). One purpose of our study was to re-examine these findings and to ask whether the notion of narrow-band channels is necessary to explain them. Here, the notion of narrow-band channels refers to any functionally isolatable band-limited spatial-frequency mechanisms subserving letter *identification* (discriminating one letter from the rest of the alphabet-set) rather than merely detecting the presence of a letter.

Compared with central vision, letter identification in peripheral vision is poorly understood. Anderson and Thibos (1999) found that the critical band of frequencies that is required for discriminating the orientation of the letter E at 30° eccentricity lies in the range 0.9–2.2 c/letter. However, discriminating the orientation of a letter E is not the same as identifying letters. Also, they did not measure the same letter discrimination performance at the fovea, thus, we cannot compare peripheral and foveal performance. Näsänen and O'Leary (1998) compared the contrast thresholds for recognizing band-pass filtered hand-written numerals between central and peripheral vision (up to 20° eccentricity). They found that contrast sensitivity peaks at about 5 c/letter, and recognition efficiency peaks at 2.5–8 c/letter, for foveal and peripheral vision alike. To our knowledge, the spatial-frequency tuning properties of letter identification in peripheral vision have not been studied. A second purpose of our study was to compare spatial-frequency tuning properties of letter identification in central and peripheral vision.

There are a priori reasons why the spatial-frequency characteristics of letter identification might be different between central and peripheral vision. To an observer, the most useful band of spatial frequencies for letter identification depends on two factors—the distribution of letter-identity information [2] across the spatial-frequency spectrum, and the observer's contrast sensitivity as a function of spatial frequency. Because letter-identity information is distributed across a range of object frequencies, its availability will be in part, limited by the shape of the contrast sensitivity function (CSF), which is well known to differ between central and peripheral vision. This interaction between the CSF and letter-identity information may result in (1) a reliance on different channel frequencies between central and peripheral vision, and (2) a dependence of channel frequency on letter size.

To compare the spatial-frequency characteristics of letter identification in central and peripheral vision, we measured the contrast thresholds required for identifying single, band-pass filtered letters at the fovea, and at 5° and 10° eccentricity in the inferior visual field, for three letter sizes at each eccentricity. Based on these data, we constructed spatial-tuning functions for identifying letters of various sizes, from which we derived the frequency at which peak sensitivity occurs, and the bandwidth of the tuning functions. Data collected for the three letter sizes at each eccentricity enabled us to examine the dependence of these tuning functions on letter size in central and peripheral vision. To test whether or not the CSF can account for the differences in the spatial-frequency characteristics of letter identification between central and peripheral vision, we incorporated a human CSF into an ideal-observer model, and tested the performance of this ideal-observer on the same letter identification task as in the human observers.

## 2. Methods

Contrast thresholds for identifying single, band-pass filtered letters were measured as a function of letter size at three retinal eccentricities: 0° (foveal), 5° and 10° in the inferior visual field. Letter stimuli were 26 lower-case letters of the Times-Roman alphabet (from "a" to "z"). For comparison, we also measured contrast thresholds for identifying unfiltered letters. We used all 26 lowercase letters instead of a limited set of letters because of our long-term interest in the relationship between letter recognition and reading. Inevitably, the use of 26 instead of a limited set of letters could introduce more

[1] Object spatial frequency refers to the specification of spatial frequency with respect to the physical size of the letters; whereas retinal spatial frequency, usually expressed in c/deg, relates to the angular subtense of the letters. The size of the letters is defined as the height of the letter x.

[2] Letter-identity information refers to the information contained in the physical stimuli that allows an observer to tell one letter apart from the rest of the letters. It can be determined by measuring an ideal-observer's contrast sensitivity for identifying letters at different spatial-frequency bands (Braje, Tjan, & Legge, 1995). For more details about letter-identity information, refer to Appendix B.

measurement variability in our data. During each trial, one of the 26 letters was randomly selected and presented in the middle of an Apple high-resolution monochrome monitor for a duration of 150 ms. The task of the observer was to identify the letter. An audio tone denoted each correct response. The host computer was a Macintosh 7200/90 and custom-written software was used to run the experiment. We used a 3 down-1 up staircase procedure to track the contrast threshold corresponding to a 79% observed correct probability on the psychometric function. Six reversals were used in each staircase and the average of the last four reversals was taken as the threshold for that block of trials. Each datum reported in the study represents the average of 4–6 independent measures of threshold for the same condition.

We adopted a nominal contrast definition to define the contrast of filtered images. We assigned a contrast of 100% to all filtered letters that were derived from a 100% contrast unfiltered letter through filtering, and without rescaling. In other words, even though a letter filtered with a high spatial-frequency band-pass filter contains less energy than one filtered with a low spatial-frequency filter, these two filtered letters are still considered to have the same 100% nominal letter contrast. For example, if letters filtered with the 2.5 c/letter band-pass filter need to be attenuated by a factor of 20 to reach threshold, then the nominal threshold contrast would be 0.05 of that of their unfiltered parent letters, or, 5%. In essence, we were measuring the contrast required of an unfiltered letter for a certain band of spatial frequencies within it to reach threshold.

To determine contrast thresholds as a function of letter sizes, we first measured the smallest letter size that could be reliably identified using unfiltered letters at the testing eccentricity. The same 3 down-1 up staircase procedure was used to track the "letter acuity". We then presented letter sizes at 0.2, 0.4 and 0.6 log units above this letter acuity at each of the three eccentricities. The sequence of testing was randomized within and between observers.

### 2.1. Stimuli

The 26 lower-case Times-Roman letters, with x-height ranging from 16 to 100 pixels, were each generated as a single letter on a background of $512 \times 512$ pixels. Each letter was positioned such that it was centered on the background horizontally, and that its baseline aligned with that of the letter x, when the letter x was centered on the background. The letters were each spatially filtered with a set of nine raised cosine log filters (Peli, 1990; Alexander et al., 1994; Chung, Levi, & Legge, 2001), with peak object spatial frequencies of 0.63, 0.88, 1.25, 1.77, 2.5, 3.54, 5, 7.07 and 10 c/letter, respectively. Fig. 1 shows a set of these filtered images of the letter "s". Spatial-filtering was accomplished using the HIPS software (Landy, Cohen, & Sperling, 1984). Each filter has a bandwidth (full-width at half-height) of 1-octave and is radially symmetrical in the log-frequency domain. The equation of the filter is given by

$$\text{Amplitude at radial frequency } fr = \frac{1.0 + \cos\left(\pi^* \frac{\log(fr)-\log(ctr)}{\log(cut)-\log(ctr)}\right)}{2}$$

where *ctr* represents the spatial frequency corresponding to the peak amplitude of the filter (center frequency) and *cut* represents the frequency at which the amplitude of the filter drops to zero (cut-off frequency).

To create a digitally filtered image file, we first assigned the stimulus file a color value of 1 (white) for the

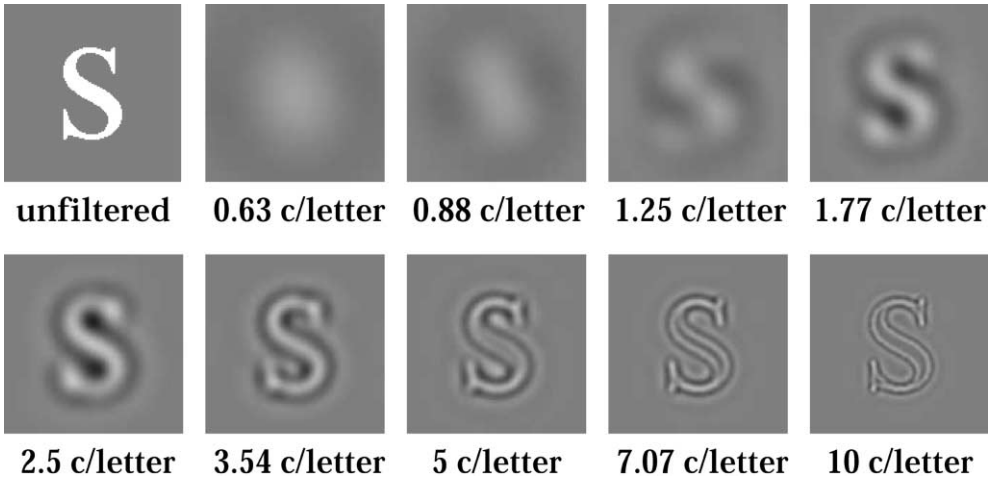


Fig. 1. Samples of the filtered letters. The spatial frequency given underneath each letter sample represents the peak object frequency of the band-pass filter.

letter and 0 (black) for the background. The binary image was then Fourier transformed and multiplied with one of the nine cosine log filters, rendered in the frequency domain. An inverse Fourier transform was then performed on the product which resulted in the final filtered image, with color values falling within a range of ±127.5 (without rescaling). To present the filtered letters on the monochrome monitor, we mapped the color values of the letters to the 256 gray levels of a gamma-corrected monochrome monitor, with each gray level linearly related to some luminance value in cd/m$^2$. The DC (color value of zero) of the filtered images was always mapped to the gray level of 127, equivalent to the mean luminance (96 cd/m$^2$) of the monitor. The calibration and mapping of each color value to luminance was controlled by an ISR video attenuator and VideoToolbox software, as described by Pelli and Zhang (1991).

### 2.2. Observers

Three observers with normal vision, one of the authors and two paid observers unaware of the purpose of the study, participated in the experiments. All had (corrected) acuity of 20/15 or better in both eyes. The observers were either emmetropic, or wore contact lenses to correct for their refractive errors. We did not correct for peripheral refractive errors because changes in peripheral refractive errors and the appearance of significant off-axis astigmatism do not generally develop within the central 20° from the fovea (Millodot & Lamont, 1974). Written informed consent was obtained from each observer after the procedures of the experiment were explained, and before the commencement of data collection. All of the observers had prior experience with psychophysical experiments. Viewing was monocular, with the eye not being tested occluded.

### 2.3. Control experiment

We used the lower visual field as the retinal locus for peripheral testing in the main experiment because it is the region commonly used in studies examining peripheral vision in relation to reading (e.g., Chaparro & Young, 1993; Higgins, Arditi, & Knoblauch, 1996; Latham & Whitaker, 1996; Chung et al., 1998). Consequently, the use of the inferior visual field facilitates comparison of our data with those in the literature. To ascertain that our results obtained in the inferior visual field are not specific to the quadrant of the visual field, we also measured the contrast threshold for identifying filtered letters at 5° and 10° in the *nasal* visual field of the left eye in two of the three observers. As in the main experiment, three letter sizes were tested, each normalized to the acuity measured at the given eccentricity in the nasal visual field.

### 2.4. Ideal-observer

We implemented an ideal-observer model to determine if the characteristics of letter identification can be accounted for by the forms of the CSF and the spatial-frequency spectra of letters. To do so, we filtered all our letter stimuli using a human CSF. We then added white noise to the output of the CSF-filtered stimuli, and determined the thresholds for the ideal-observer to identify the letters. Details of the implementation of this ideal-observer model are given in Appendix A. We refer to this model as the *CSF-ideal-observer*. For comparison and for our analysis, we also measured the ideal-observer's contrast thresholds for identifying letters in the *absence* of a human CSF (equivalent to a flat CSF). The function thus obtained, which we shall call the letter sensitivity function (LSF), represents the letter-identity information distributed across spatial frequency. This LSF takes into account the inter-letter differences, and thus we did not need to assume that the letters are equally detectable or that the power spectra of the letters are the same.

The CSF-ideal-observer used empirical CSFs from our human observer SC, obtained at the fovea and at 10° in the lower visual field (Fig. 2). These CSFs were measured using a sinewave grating detection task with a staircase procedure, controlled by the PSYCHO software supplied by the Visual Stimulus Generator (Cambridge Research Systems Ltd., Cambridge, UK). The absolute sensitivities and the general shape of our CSFs are similar to those reported in previous studies (De Valois, Morgan, & Snodderly, 1974; Virsu & Rovamo, 1979; Banks, Sekuler, & Anderson, 1991).

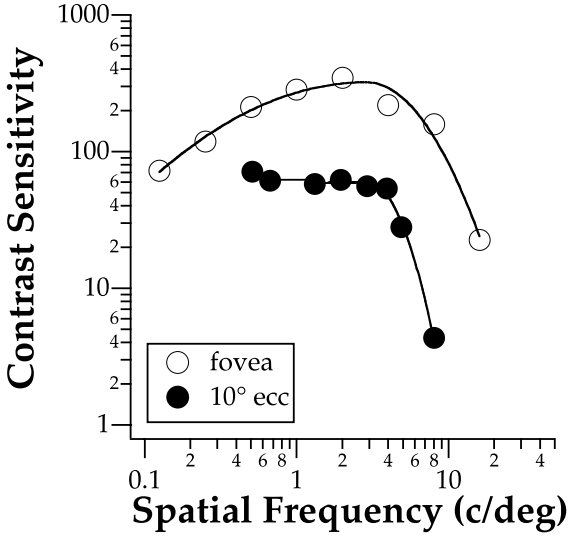


Fig. 2. CSFs of observer SC, obtained at the fovea (unfilled circles) and 10° eccentricity (filled circles). Solid lines represent the asymmetrical parabolic functions fitted to each data-set.

## 3. Results

Relative contrast sensitivity is plotted as a function of the center frequency of the band-pass filter for the three retinal eccentricities and the three human observers in Fig. 3, with letter size as a parameter. For some conditions, thresholds for identifying letters were very high at some high and low filter frequencies. Consequently, we were not able to obtain a measurement and thus there are no data plotted at these frequencies. [3] Relative contrast sensitivity is derived from the ratio of contrast thresholds between filtered and unfiltered letters to facilitate visualization of data across observers and testing conditions. For instance, a band with a relative contrast sensitivity of 0.5 means that the nominal threshold contrast of this band was twice as high as that of an unfiltered letter. In general, the relative contrast sensitivity vs. spatial-frequency plot demonstrates spatial-tuning characteristics. To derive the spatial frequency at which peak contrast sensitivity occurs, which we referred to as the peak tuning frequency, we fitted a parabolic function to each data-set, on log–log axes. The equation of the parabolic function is given by

$$\log(\text{relative contrast sensitivity}) = \log(\text{amplitude}) - \frac{4}{\log(2)\sigma^2}(\log(sf) - \log(sf_p))^2$$

where amplitude represents the full-height of the function, $sf$ is spatial frequency, $sf_p$ is the peak tuning frequency and $\sigma$ is the bandwidth of the function in octaves.

Fig. 3 shows that there is a progressive shift in peak tuning frequency toward higher object spatial frequency from the fovea to 5° and 10° eccentricity (repeated measures ANOVA: $F_{(\mathrm{df}=2,4)} = 31.5$, Greenhouse-Geisser adjusted $p = 0.011$). Averaged across letter sizes, [4] the mean peak tuning frequency (±1 SD) shifts from 2.00 ± 0.38 c/letter at the fovea to 2.63 ± 0.43 c/letter at 10° eccentricity, corresponding to a factor of 1.32 (0.4 octaves). However, the frequency-selectivity of the spatial-tuning functions, as represented by the bandwidth of these functions, does not change with eccentricity (repeated measures ANOVA: $F_{(\mathrm{df}=2,4)} = 0.93$, Greenhouse-Geisser adjusted $p = 0.453$, see Fig. 5). Averaged across letter sizes, the mean bandwidths are 2.36 ± 0.37, 2.55 ± 0.25 and 2.67 ± 0.32 octaves at the fovea, 5° and 10° eccentricity, respectively.

With respect to the effect of letter size, Fig. 3 shows that the peak tuning frequency progressively shifts toward lower frequency when the letter size becomes smaller (repeated measures ANOVA: $F_{(\mathrm{df}=2,4)} = 23.01$, Greenhouse-Geisser adjusted $p = 0.028$). Averaged across the three eccentricities, the mean peak tuning frequency (±1 SD) shifts from 2.73 ± 0.33 for letter size of 0.6 log units above the acuity to 1.91 ± 0.29 c/letter for letter size of 0.2 log units above the acuity, representing a factor of 1.43 (0.52 octaves). In addition to this shift in spatial frequency, the bandwidth of the spatial-tuning functions becomes slightly narrower with increase in letter size (repeated measures ANOVA: $F_{(\mathrm{df}=2,4)} = 17.86$, Greenhouse-Geisser adjusted $p = 0.048$, see Fig. 5). Averaged across eccentricity, the mean bandwidth decreases from 2.77 ± 0.33 octaves for the smallest letter size, to 2.40 ± 0.28 octaves for the largest letter size.

Fig. 4 summarizes the dependence of peak tuning frequency on letter size and eccentricity. Peak tuning frequency, converted to retinal frequency in c/deg, is plotted as a function of letter size, expressed as nominal letter frequency [5] in Fig. 4A. Data plotted were pooled across the three observers and were shown for the three eccentricities. We fitted a power function (straight line on log–log axes) to the data-set of each eccentricity. The exponents of these power functions (slopes of the straight lines) are almost the same at each eccentricity (range: 0.56–0.63). However, the foveal data-set appears to have a vertical offset (displaced upward) from the data obtained at 5° and 10° eccentricity. This is not to be confused with our finding that peak tuning frequency is lower at the fovea than in the periphery, when frequency is expressed as object frequency in c/letter, and letter sizes are expressed as multiples above the acuity (Fig. 3). We will return to the significance of the exponents of the power functions in Section 4.

Similarly, we summarize the dependence of bandwidth on letter size (expressed as nominal letter frequency), and eccentricity in Fig. 5A. Bandwidth remains virtually unchanged for the three eccentricities examined. At the fovea, there is a slight trend of a decrease in bandwidth for larger letters, although as we will see in Section 4, this trend may be of little significance.

Fig. 6 compares the spatial-tuning functions for letter identification between nasal (control experiment) and inferior (main experiment) visual fields. Clearly, the spatial-tuning functions are remarkably similar for any given condition. This suggests that as long as the letter

[3] In contrary, the CSF-ideal-observer was able to identify letters at all filter frequencies. This is because the CSF-ideal-observer could make use of information present in the letter stimuli that human observers were not able to use, e.g. the coarse luminance cues at low frequencies.

[4] In this paper, unless otherwise specified, letter size refers to the letter size normalized to the acuity at each eccentricity. In other words, the *physical* letter sizes tested in the periphery were larger than those tested in the fovea.

[5] Nominal letter frequency is a measurement based on the physical letter size, assuming that a letter subtending 5 arc min in height has an equivalent nominal letter frequency of 30 c/deg. It should not be confused with the spatial frequency of the band-pass filtered letters.

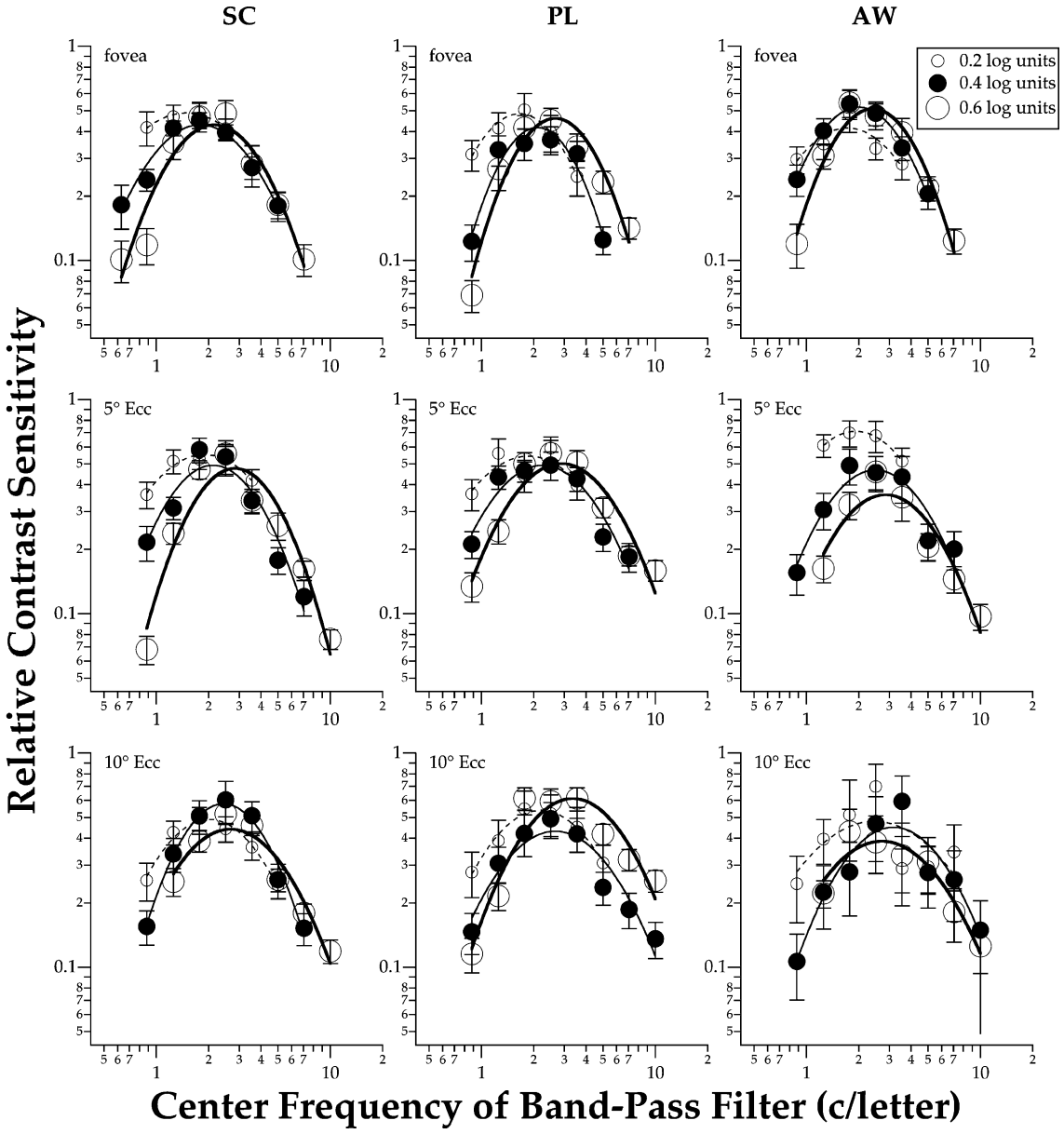


Fig. 3. Relative contrast sensitivity is plotted as a function of the center frequency of the band-pass filter (c/letter), for the three human observers, each tested at three retinal eccentricities. Each panel includes data obtained for three letter sizes, denoted by the size of the symbols. The fitted curve is a parabolic function (see text for details). Error bars represent ±1 s.e.m.

size is scaled with respect to the local acuity, then the spatial-frequency characteristics of letter identification are governed by retinal eccentricity (distance from the fovea) but not the direction in the visual field.

### 3.1. CSF-ideal-observer

Contrast thresholds for letter identification obtained from the CSF-ideal-observer model are plotted as a function of the center frequency of the band-pass filter in Fig. 7, for various letter sizes and for two retinal eccentricities (fovea and 10° eccentricity). Like the human observers' data, plots of relative contrast sensitivity vs. spatial frequency exhibit spatial-tuning characteristics, at both the fovea and 10° eccentricity, and for all letter sizes. However, there is a progressive change in the shape of these tuning functions with letter size. Specifically, when letter size becomes smaller, the peak of the tuning functions shift progressively toward lower object spatial frequency. Also, the tuning functions seem to become more asymmetrical in shape, with less attenuation for frequencies lower than the peak tuning frequency, and more attenuation for frequencies higher than the peak tuning frequency.

The relationship between peak tuning frequency and nominal letter frequency as determined from the CSF-

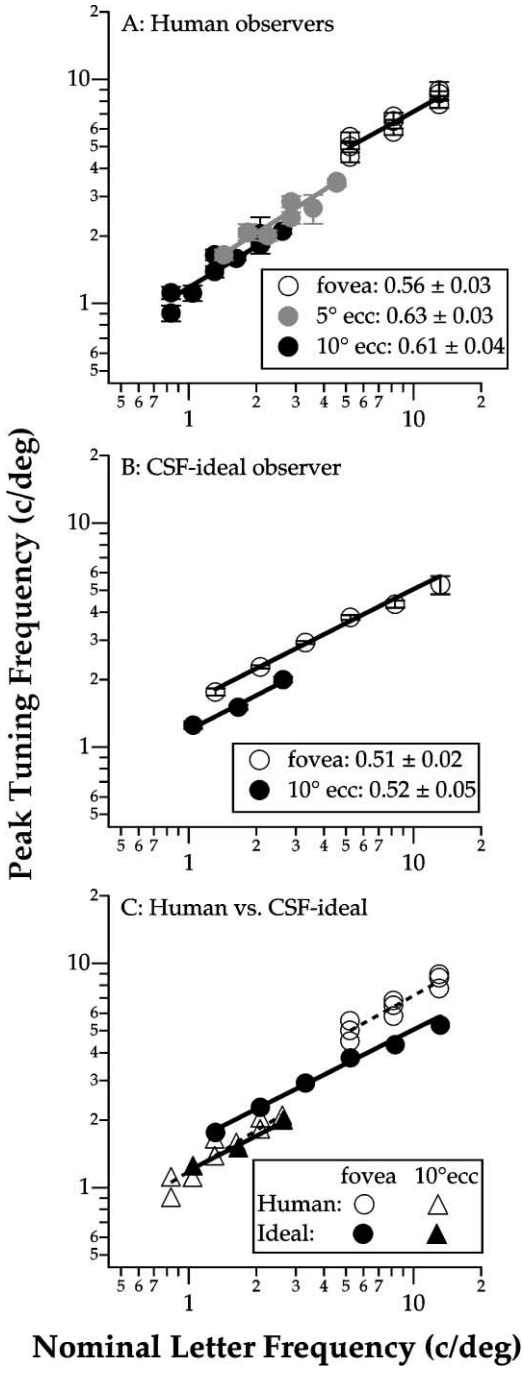


Fig. 4. Peak tuning frequency (c/deg) is plotted as a function of nominal letter frequency (c/deg), for (A) human observers, (B) the CSF-ideal-observer and (C) a comparison between human and the CSF-ideal-observer. Data plotted for the human observers are pooled from the three observers. Data for the three eccentricities are coded by different colored symbols. The solid lines are the power-functions fitted to each data-set, the exponents of which (±SE) are given in the legend. Error bars represent ±1 s.e.m.

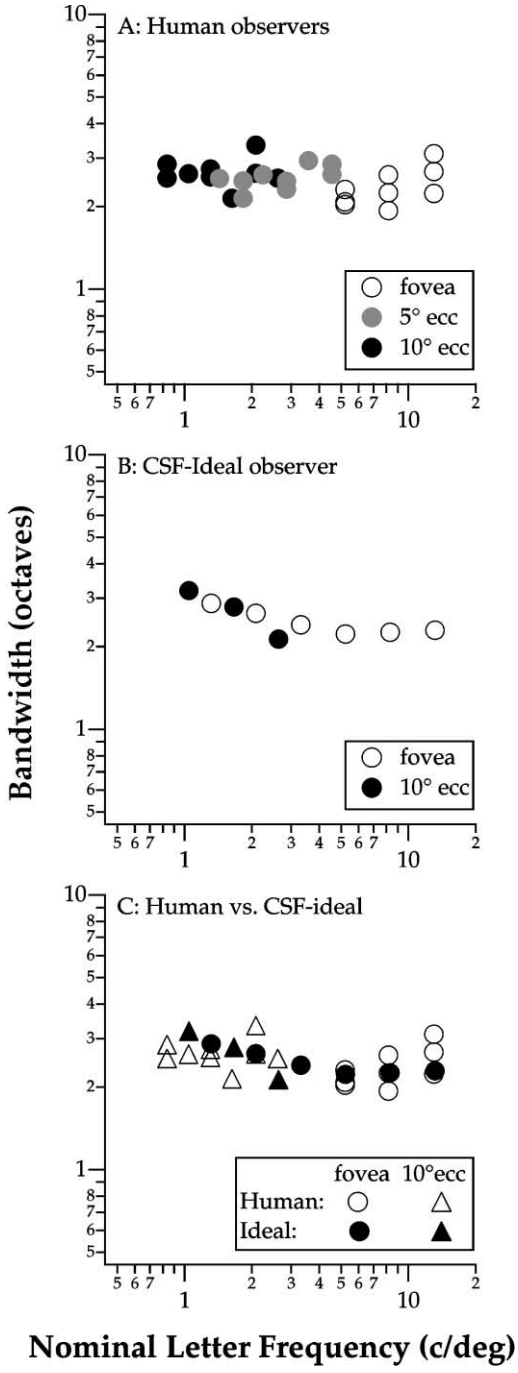


Fig. 5. Bandwidth (octaves) is plotted as a function of nominal letter frequency (c/deg), for (A) human observers, (B) the CSF-ideal-observer and (C) a comparison between human and the CSF-ideal-observer. Data plotted for the human observers are pooled from the three observers. Data for the three eccentricities are coded by different colored symbols.

ideal-observer is shown in Fig. 4B, for the fovea and 10° eccentricity. Like the human data, there is a small vertical offset between the foveal data and the data obtained at 10° eccentricity. The exponents of the power functions are 0.45 and 0.52 for the fovea and 10° eccentricity, respectively.

For the CSF-ideal-observer, the bandwidths are essentially identical for the smallest three letter sizes tested at the fovea, but increase with the larger three letter sizes (Fig. 5B). The three smallest letter sizes are also the sizes

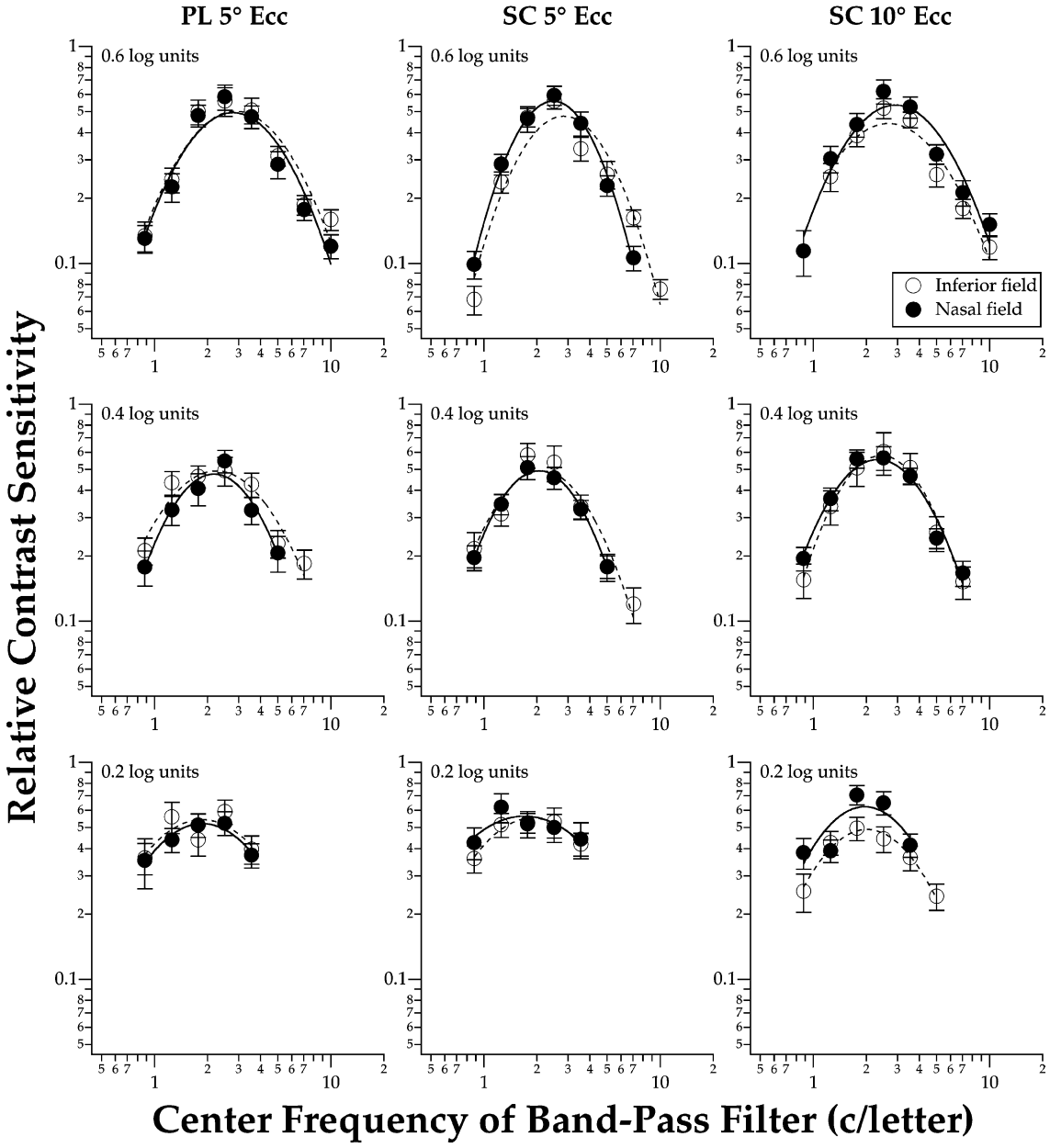


Fig. 6. Plots of relative contrast sensitivity vs. center frequency of the band-pass filter are compared between inferior (original, plotted as unfilled symbols) and nasal (filled symbols) fields, for two observers and two eccentricities. Error bars represent ±1 s.e.m.

used in testing the human observers, and indeed, the bandwidths for these sizes are similar for the CSF-ideal and human observers (Fig. 5C). At 10° eccentricity, the CSF-ideal-observer's bandwidths show an increase with letter size; however, the values of these bandwidths are still within the range obtained from the human observers (Fig. 5C).

## 4. Discussion

The primary goal of this study was to compare the spatial-frequency characteristics of letter identification in central and peripheral vision. In general, the spatial frequency at which peak sensitivity occurs, and the frequency-selectivity of the spatial-tuning functions, are virtually identical between central and peripheral vision, when the CSF is taken into account. For an increase in retinal eccentricity from the fovea to 10°, we observed the following in our human observers: (1) peak tuning frequency (c/deg) increases with nominal letter frequency (linearly related to 1/letter size) according to a power function with an exponent $<1$, but a value that is similar across the three eccentricities tested; (2) bandwidth of the tuning functions, representing frequency-selectivity, remains similar across letter sizes and eccentricities. Based

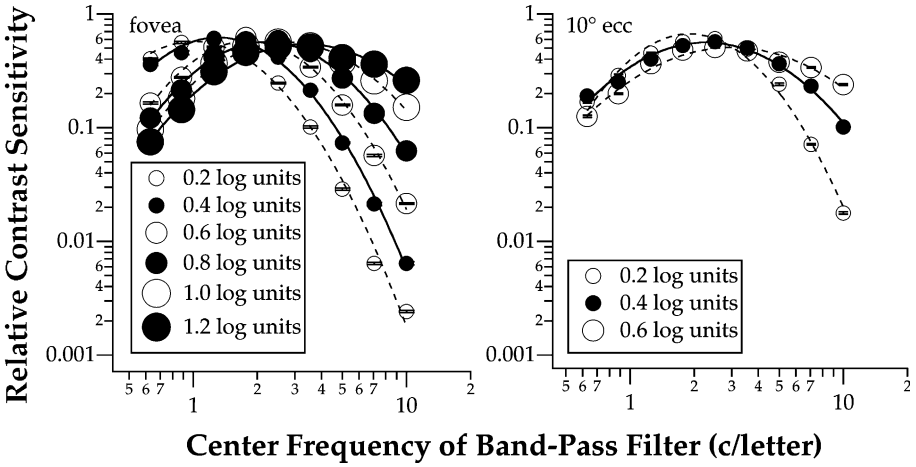


Fig. 7. Relative contrast sensitivity is plotted as a function of the center frequency of the band-pass filter (c/letter), for the ideal-observer, at the fovea and 10° eccentricity. Each panel includes data obtained for different letter sizes, denoted by the size of the symbols. The fitted curve is a parabolic function (see text for details). Error bars represent ±1 s.e.m.

on these two findings, we postulate that the spatial-tuning function for letter identification is at least in part, governed by the angular size of the letter, irrespective of eccentricity. However, there is one major difference between the fovea and the periphery—for the same physical letter size, the peak sensitivity of the tuning function occurs at a higher retinal frequency (c/deg) at the fovea than in the periphery.

To account for the similarities and differences in central and peripheral vision, we asked if letter-identity information alone, combined with the human's spatial resolution as represented by the CSF, can provide a basis for our findings (see Appendix B for a formal analysis). In the following sections, we will compare our human foveal and peripheral data relative to those obtained from the CSF-ideal-observer.

### 4.1. Peak tuning frequency

Fig. 7 shows that the tuning functions exhibited by the CSF-ideal-observer to identify letters at a given eccentricity changes with letter size. When peak tuning frequency is expressed as object frequency in c/letter, there is a shift toward higher frequency as letter size increases. Equivalently, we can express the same shift in retinal frequency, as shown in Fig. 4B. A comparison with the human data (Fig. 4A) reveals a good quantitative agreement in the exponent of the power functions obtained at the same eccentricity. This suggests that the shift in tuning frequency observed in our human observers can be accounted for by the CSF and how the letter-identity information distributes across spatial frequencies, and that it is probably unnecessary to hypothesize any channel-selection process within the human observer.

However, there is a notable difference in the power functions between the human and CSF-ideal-observer—at the fovea, the two power functions are separated vertically (Fig. 4C). Based on the power-function fit, we estimated that at the fovea, peak sensitivity of the tuning functions occurs at a frequency that is approximately 27% (0.34 octaves) higher in human observers than in the ideal-observer. Could the error associated with the estimation of the human CSF account for this difference? Based on Eq. (B.5a) in Appendix B, a 10% error in the estimation of the low spatial-frequency roll-off of the CSF ($\Delta\beta_{csf}$) will cause a 4% error in the estimation of the exponent of the power-function fit. Similarly, a 10% error in the estimation of the high spatial-frequency fall-off of the CSF ($\Delta\alpha_{csf}$) will cause a 1.6% error in the estimation of the exponent of the power-function fit. These errors are too small to explain the 27% difference in peak tuning frequency between the foveal data of the human and the CSF-ideal-observer.

The reason why the peak tuning frequency is higher in human observers than in the CSF-ideal-observer is still unclear. One possibility is that in the fovea, the visual system may rely on broadband features such as edges for pattern perception, because these features are often insensitive to changes in image parameters such as luminance and contrast, thus maintaining perceptual constancy. With respect to our filtered stimuli, these broadband features are missing, but we speculate that the visual system may be able to construct them internally, as in the case of illusory contour.

### 4.2. Size variance

Using critical-band noise masking paradigm to study the properties of letter identification, Solomon and Pelli (1994) suggested that letter identification is mediated by channels that scale with letter size, such that when channel frequency and letter size are both expressed as retinal frequency (c/deg), a change in letter size will lead

to a corresponding change in channel frequency of the same magnitude, i.e., a power function with an exponent of 1.0. This indicates *size invariance*, in agreement with Parish and Sperling (1991). However, using the same paradigm, Majaj et al. (2002) reported an exponent of ≈0.7, for a larger range of letter sizes and for a number of different fonts and alphabets. Consistent with the value reported by Majaj et al., the exponents we obtained from our CSF-ideal, as well as human observers are in the range of 0.5–0.7. These values are much lower than an exponent of 1.0, implying that the shift in the peak tuning frequency (or "channel frequency" as used by Solomon and Pelli (1994)) is less than the corresponding change in letter size and that letter identification is *not* size invariant.

Note that Solomon and Pelli (1994) claimed size invariance in the fovea only for letters between 0.5° and 15.8° in size. The largest letter we used in the fovea, for our human observers, was 0.48°. Therefore, it is possible that the reason we did not find scale invariance in our human observers is due to the smaller letter sizes we used (although we also did not find scale invariance for the CSF-ideal-observer, with letter sizes that fall within the range used by Solomon and Pelli). To test whether or not the lack of scale invariance in our data is due to letter size, we collected additional foveal data from observer SC with letter sizes that extended to 1.9°. The power-function fitted to this set of data yields an exponent of 0.65, a value that is highly consistent with those obtained for the original set of data (Fig. 8A). This result suggests that the lack of scale invariance in our data is not restricted to a range of small letter sizes.

What explains our observed effect of size? Because there is excellent agreement between CSF-ideal and human observers in the exponents of the plots in Fig. 4A and B, the explanation need not go beyond the two factors that limit performance for the CSF-ideal-observer—the letter-identity information in the spatial-frequency spectrum of the stimulus and the CSF. In Appendix B, we demonstrate how a change in letter size of a certain magnitude leads to a change in peak tuning frequency of a smaller magnitude, thus yielding an exponent of <1 in the power-function fit.

### 4.3. Tuning bandwidth

The bandwidth of the tuning functions exhibited by the CSF-ideal-observer in identifying letters averages 2.25 octaves at the fovea and 2.7 octaves at 10° eccentricity, for the three letter sizes for which we have human data. These values compare favorably with the human data which average 2.36 octaves at the fovea and 2.67 octaves at 10° eccentricity (see also Fig. 5).

At the fovea, our human data show that there is a slight decrease in bandwidth when the letter size increases, as shown by the ANOVA results. This contradicts the result of the CSF-ideal-observer, in which the bandwidth increases from 2.29 to 2.87 octaves, when letter size increases from 0.2 to 1.2 log units above acuity. In Appendix B, we showed that bandwidth should indeed increase with letter size, for letter size ranges between 0.23 and 2.7 log units above acuity. We speculate that the slight decrease in bandwidth with increased letter size as observed in the human foveal data is due to the small range of letter size, coupled with the variability between and within subjects. Indeed, when larger letter sizes were tested, we show that the bandwidth of the tuning function in the fovea does not decrease with increased letter size (Fig. 8B).

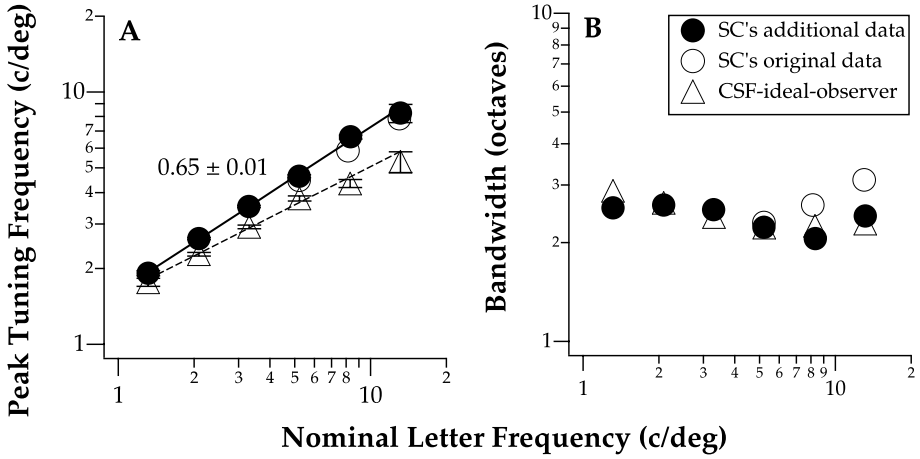


Fig. 8. Peak tuning frequency in c/deg (A) and bandwidth in octaves (B) are plotted as a function of nominal letter frequency in c/deg for the additional foveal data collected from observer SC. These additional data are plotted as filled circles. For comparison, the original data from the same observer (unfilled circles) and data from the CSF-ideal-observer (unfilled triangles) are included. The exponent of the power-function (±SE) fitted to the set of additional data is given in A.

## 5. Conclusions

Our findings suggest that the spatial-frequency tuning characteristics of letter identification are fundamentally identical between central and peripheral vision, when the CSF and the distribution of letter-identity information across spatial frequencies are taken into account. Both spatial-frequency-selectivity (tuning bandwidth) and the effect of letter size can be explained without invoking active selection among narrow-band spatial-frequency channels internal to an observer. The only discrepancy between human data and the CSF-ideal-observer is that human observer's spatial frequency tuning at the fovea is higher than the predicted value by one third of an octave. The cause of this discrepancy is still not known.

## Acknowledgements

This study was supported by research grants EY12810 (STLC) and EY02934 (GEL) from the National Institutes of Health. We thank Andrew Luebker for technical assistance.

## Appendix A

For the CSF-ideal-observer model, we assumed that a linear filter and an additive noise source were situated between the stimulus and an ideal-observer (Fig. 9). The linear filter had a modulation transfer function with a shape identical to a human CSF. The noise was assumed to be a white (Gaussian) luminance noise. White noise was added after the CSF filter so that the signal-to-noise ratio across spatial frequencies followed the shape of the CSF. The exact level of this internal noise is inconsequential for our analysis and was arbitrarily chosen to have a standard deviation of 1.0 and a mean of 0.

Our ideal-observer model is an optimal classifier for identifying the CSF-filtered stimuli. It is optimal in the sense that no other recognition mechanism, biological or otherwise, can exceed its average accuracy level. Thus, it is one that makes use of all the task-relevant information in the stimuli, or in this case, in the CSF-filtered images.

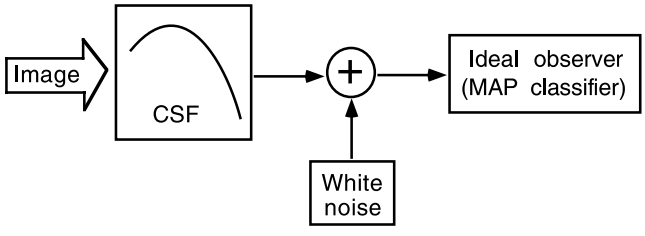


Fig. 9. A schematic cartoon illustrating the CSF-ideal-observer.

The ideal-observer's knowledge about the stimuli includes noiseless image templates (CSF-filtered letters). Furthermore, it knows the contrast of the stimuli and the statistics (mean and standard deviation) of the noise. The decision rule for such an ideal-observer in a more general case has been derived elsewhere (Tjan, Braje, Legge, & Kersten, 1995). For completeness, we briefly restate the derivation specific to the application in this paper.

To be maximally correct on average, one must select the response letter $L$ that is most probable given the stimulus $S$ (CSF-filtered letter plus noise). That is, select $L$ such that the posterior probability $\Pr(L|S)$ is at its maximum. By Bayes rule, we can write

$$\Pr(L|S) = \Pr(S|L)\Pr(L)/\Pr(S)$$

The prior probability $\Pr(L)$ is a constant (all letters occurred equally often in our experiments), and $\Pr(S)$ does not depend on $L$; therefore, to maximize the posterior probability $\Pr(L|S)$ with respect to $L$ is the same as maximizing the likelihood $\Pr(S|L)$. Let $L$ also denote the corresponding noiseless template of the letter $L$. Under Gaussian luminance noise, the likelihood of $L$ can be computed as follows:

$$\Pr(S|L) = C\exp\left(\frac{-(L-S)^2}{2\sigma^2}\right)$$

where $C$ is a normalization constant independent of $L$, and $\sigma$ is the standard deviation of the luminance noise. Because the exponential function is monotonic, maximizing the likelihood $\Pr(S|L)$ is the same as minimizing the Eclidean distance between a noiseless template $L$ and the noisy stimulus $S$. In sum, the optimal decision rule for the CSF-ideal-observer is simply:

$$\text{Choose the response } L \text{ that minimizes } (L-S)^2 \quad \text{(A.1)}$$

The templates $L$ of CSF-filtered letters were kept in the ideal-observer's memory and were also used to generate the noisy stimuli presented to the ideal-observer. They were created as follows. A CSF, denoted as $G(f_r)$, specifies the gain as a function of spatial frequency measured at the retina. To digitally filter an input image $I$ by the CSF, a scaling factor $k$ relating retinal frequency $f_r$ to object frequency $f_o$ (in cycles per pixel) was first determined. Given letter size in pixels $p$ and in visual angle $\alpha$, we have $k = p/\alpha$ and $f_r = kf_o$. The CSF-filtered image $S$ was generated from the stimulus $I$ as follows:

$$S = \mathscr{F}^{-1}[\mathscr{F}[I-bg](f_o)\cdot G(kf_o)] + bg$$

where $\mathscr{F}$ and $\mathscr{F}^{-1}$ represent the forward and inverse fast-Fourier transforms, respectively, and $bg$ is the background luminance.

Numerical simulation based on the optimal decision rule (A.1) was used to determine the CSF-ideal-observer model's contrast thresholds at 79% recognition. The

appendix in Tjan and Legge (1998) provides an efficient implementation of such a simulation.

The model's contrast thresholds were converted to relative sensitivities as in the analysis of the human data. Contrast threshold for the model was measured at the stimulus level before the CSF filter. While contrast threshold obviously depends on the amount of internal noise, relative sensitivity, which is a ratio of contrast thresholds, is independent of the noise. This is because an ideal-observer's accuracy level is determined by the signal-to-noise ratio (Tjan et al., 1995). Thus, for a fixed accuracy level, the ideal-observer's contrast threshold is linearly proportional to the standard deviation of the noise. Expressing contrast-threshold ratio as a relative sensitivity measurement factors out any effect due to the (arbitrary but fixed) internal noise level.

## Appendix B

Our CSF-ideal-observer model behaves *as if* it were using a single narrow-band channel for identifying letters of a given size, even though the model uses no such channels. Both the peak frequency and, to a much lesser extent, the bandwidth of this "phantom" channel, vary as a function of letter size. Variation in eccentricity also leads to a change in peak frequency of the phantom channel and bandwidth due to the difference in the CSF. In this appendix, we will analytically derive the performance of the CSF-ideal-observer model. By doing so, we can understand why the model appears to be using a narrowly tuned channel for letter identification, and why the apparent channel frequency and bandwidth vary with letter size. Without loss of generality, we shall consider only the foveal conditions. Replacing the fovea CSF by the CSF at a given eccentricity will lead to results for that eccentricity. Equations corresponding to key results of our analysis have their equation numbers underlined.

Performance of the CSF-ideal-observer model is limited by two factors—the front-end CSF and the letter-identity information distributed across the spatial-frequency spectrum. Distribution of letter-identity information can be measured using an ideal-observer without the CSF front-end (or equivalently by setting the CSF in the CSF-ideal-observer model to a unit gain for all spatial frequencies). Specifically, we measured the ideal-observer's contrast sensitivity for letter identification as a function of spatial frequency. We call this the *ideal letter sensitivity function*, or ideal LSF. The ordinate, though in units of contrast sensitivity, is linearly related to units of bits of information regarding letter identity transmitted by the stimulus per contrast units per octave. The procedure for measuring the ideal LSF is identical to that used to obtain the tuning functions in Fig. 7 but with a flat CSF. The method section of the main text and Appendix A provide all the necessary details for this procedure. Fig. 10 shows the ideal LSF for the lower-case Times-Roman letters used in this paper, plotting (normalized) contrast sensitivity as a function of object spatial frequency in c/letter. We used a bi-parabolic function of the following form (on log–log coordinates) to describe the shape of the ideal LSF:

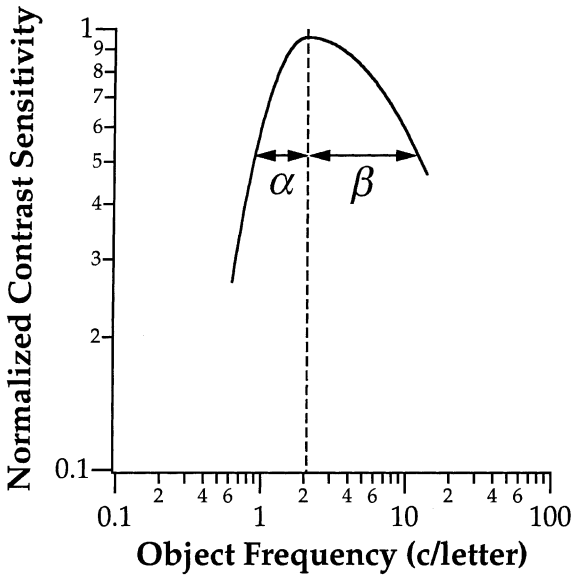


Fig. 10. An ideal LSF for lower-case Times-Roman letters, representing the distribution of letter-identity information across spatial frequency. The function was obtained by measuring the contrast threshold for discriminating the 26 letters, as a function of object spatial frequency. Measurements were obtained from numerical simulations of an ideal observer with a flat CSF. The curve represents a bi-parabola fit to the data according to Eq. (B.1), with peak spatial frequency ($f_p$) = 2.07 c/letter and half-bandwidths ($\alpha, \beta$) = (1.26, 2.71) octaves.

$$\mathrm{LG}(\mathrm{Lf}) = [K] - \frac{1}{\log(2)} \begin{cases} (\mathrm{Lf} - \mathrm{Lf_p})^2/\alpha^2 & \text{if } \mathrm{Lf} \leqslant \mathrm{Lf_p} \\ (\mathrm{Lf} - \mathrm{Lf_p})^2/\beta^2 & \text{if } \mathrm{Lf} > \mathrm{Lf_p} \end{cases} \quad \text{(B.1)}$$

where LG is the log contrast sensitivity, Lf the log spatial frequency, $K$ the log maximum contrast sensitivity, and $\mathrm{Lf_p}$ the log peak spatial frequency. $\alpha$ represents the half-bandwidth (half-width at half-maximum) for the limb of the function where spatial frequencies are lower than the peak spatial frequency, while $\beta$ is the half-bandwidth for the limb of the function where spatial frequencies are higher than the peak spatial frequency. The bandwidth of the bi-parabolic function is $\alpha + \beta$ in octave units. Because the ideal-observer does not have any limitation in spatial resolution, the ideal LSF is invariant to letter size. This is why in Fig. 10, spatial frequency is expressed in units of c/letter and not c/deg. If spatial frequency is expressed in c/deg, then a change in letter size will shift the function horizontally in

log–log coordinates without altering its shape. Therefore, the peak spatial frequency $f_p$, when expressed in c/deg units, is linearly related to 1/size and thus to the nominal letter frequency.

Although in the main text, the performance of the CSF-ideal-observer model (Fig. 7) were obtained with numerical simulation (described in Appendix A), these results can be expressed analytically in terms of CSF and ideal LSF. Specifically, the contrast sensitivity of the CSF-ideal-observer model at a given spatial frequency is simply the product of the CSF and ideal LSF at that frequency. [6] That is, the measured letter channel of the model takes the following form:

*B.1. "Channel" frequency vs. letter size (or nominal letter frequency)*

The spatial frequency with peak sensitivity for letter identification (i.e., the peak frequency of the phantom channel) can be determined by calculating the derivative of Eq. (B.2) with respect to Lf, setting the derivative to zero, and solving for Lf. Eq. (B.2) shows that the tuning function of the phantom channel consists of segments of three parabolic functions in a log–log plot. The form of the middle segment depends on whether the peak frequency of the ideal LSF ($\mathrm{Llsf_p}$) is higher (for smaller letters) or lower (for larger letters) than the peak frequency of the CSF ($\mathrm{Lcsf_p}$). For both cases, it can be verified that the peak channel frequency, $\mathrm{Lf_p}$, lies in the middle segment, as long as $\alpha_{lsf}$, $\beta_{lsf}$, $\alpha_{csf}$ and $\beta_{csf} > 0$.

$$
\begin{aligned}
\log(\text{channel\_gain}) &= \mathrm{LG(Lf)} = \log(\mathrm{CSF}\times\text{ideal\_LSF}) = \log(\mathrm{CSF}) + \log(\text{ideal\_LSF}) \\
&= K - \frac{1}{\log(2)}\left(\begin{cases} (\mathrm{Lf}-\mathrm{Lcsf_p})^2/\alpha_{csf}^2 & \text{if } \mathrm{Lf}\leqslant\mathrm{Lcsf_p} \\ (\mathrm{Lf}-\mathrm{Lcsf_p})^2/\beta_{csf}^2 & \text{if } \mathrm{Lcsf_p}<\mathrm{Lf}\end{cases} + \begin{cases} (\mathrm{Lf}-\mathrm{Llsf_p})^2/\alpha_{lsf}^2 & \text{if } \mathrm{Lf}\leqslant\mathrm{Llsf_p} \\ (\mathrm{Lf}-\mathrm{Llsf_p})^2/\beta_{lsf}^2 & \text{if } \mathrm{Llsf_p}<\mathrm{Lf}\end{cases}\right) \\
&= K - \frac{1}{\log(2)}\left(\begin{array}{l} \text{for } \mathrm{Lcsf_p}\leqslant\mathrm{Llsf_p} \text{ (small letters)}: \\ \quad\begin{cases} (\mathrm{Lf}-\mathrm{Lcsf_p})^2/\alpha_{csf}^2 + (\mathrm{Lf}-\mathrm{Llsf_p})^2/\alpha_{lsf}^2, & \mathrm{Lf}\leqslant\mathrm{Lcsf_p} \\ (\mathrm{Lf}-\mathrm{Lcsf_p})^2/\beta_{csf}^2 + (\mathrm{Lf}-\mathrm{Llsf_p})^2/\alpha_{lsf}^2, & \mathrm{Lcsf_p}<\mathrm{Lf}\leqslant\mathrm{Llsf_p} \\ (\mathrm{Lf}-\mathrm{Lcsf_p})^2/\beta_{csf}^2 + (\mathrm{Lf}-\mathrm{Llsf_p})^2/\beta_{lsf}^2, & \mathrm{Llsf_p}<\mathrm{Lf}\end{cases} \\ \text{for } \mathrm{Llsf_p}<\mathrm{Lcsf_p} \text{ (large letters)}: \\ \quad\begin{cases} (\mathrm{Lf}-\mathrm{Lcsf_p})^2/\alpha_{csf}^2 + (\mathrm{Lf}-\mathrm{Llsf_p})^2/\alpha_{lsf}^2, & \mathrm{Lf}\leqslant\mathrm{Llsf_p} \\ (\mathrm{Lf}-\mathrm{Lcsf_p})^2/\alpha_{csf}^2 + (\mathrm{Lf}-\mathrm{Llsf_p})^2/\beta_{lsf}^2, & \mathrm{Llsf_p}<\mathrm{Lf}\leqslant\mathrm{Lcsf_p} \\ (\mathrm{Lf}-\mathrm{Lcsf_p})^2/\beta_{csf}^2 + (\mathrm{Lf}-\mathrm{Llsf_p})^2/\beta_{lsf}^2, & \mathrm{Lcsf_p}<\mathrm{Lf}\end{cases}\end{array}\right)
\end{aligned}
\tag{B.2}
$$

where $K$ is the sum of the maximum log amplitudes of the CSF and the ideal LSF, $\mathrm{Llsf_p}$, $\alpha_{lsf}$, $\beta_{lsf}$ are the peak spatial frequency in c/deg, and the half-bandwidths of the ideal LSF respectively. Same naming conventions apply to the parameters of the CSF.

Fig. 11 illustrates the relationship between the CSF, ideal LSF, letter size, and the observed letter channel as specified by Eq. (B.2). Note that the position of the ideal LSF relative to the CSF is determined by the letter size, which in turns affects the peak spatial frequency of the letter channel.

Consider the case of $\mathrm{Lcsf_p} < \mathrm{Llsf_p}$, within the range of $\mathrm{Lcsf_p}\leqslant\mathrm{Lf}\leqslant\mathrm{Llsf_p}$, we have

$$
\begin{aligned}
\frac{\mathrm{d}(\log(\text{channel\_gain}))}{\mathrm{dLf}} &= \frac{\mathrm{d}}{\mathrm{dLf}}\Big(K - \Big((\mathrm{Lf}-\mathrm{Lcsf_p})^2/\beta_{csf}^2 \\
&\quad + (\mathrm{Lf}-\mathrm{Llsf_p})^2/\alpha_{lsf}^2\Big)/\log(2)\Big) \\
&= -2\big((1/\alpha_{lsf}^2 + 1/\beta_{csf}^2)\mathrm{Lf} \\
&\quad - (\mathrm{Llsf_p}/\alpha_{lsf}^2 + \mathrm{Lcsf_p}/\beta_{csf}^2)\big)/\log(2)
\end{aligned}
\tag{B.3}
$$

Set Eq. (B.3) to zero and solve for Lf, we obtained the peak channel frequency as

$$
\begin{aligned}
&\text{channel frequency} \\
&= \mathrm{Lf_p} = (\mathrm{Llsf_p}/\alpha_{lsf}^2 + \mathrm{Lcsf_p}/\beta_{csf}^2)/(1/\alpha_{lsf}^2 + 1/\beta_{csf}^2) \\
&= \left(\frac{\beta_{csf}^2}{\alpha_{lsf}^2+\beta_{csf}^2}\right)\mathrm{Llsf_p} + \left(\frac{\alpha_{lsf}^2}{\alpha_{lsf}^2+\beta_{csf}^2}\right)\mathrm{Lcsf_p}
\end{aligned}
\qquad \text{(B.4a)}
$$

[6] In the main text, we measured contrast sensitivity relative to an observer's contrast threshold for identifying unfiltered letters at the given eccentricity and stimulus size. The purpose of using this "relative contrast sensitivity" measure was to facilitate direct comparison across observers, letter sizes, and eccentricities. In log–log plots, the difference between relative and absolute contrast sensitivity is merely a vertical shift, which does not affect the peak channel frequency or the channel bandwidth. For simplicity, we shall refer to contrast sensitivity without qualification. The numerical results will be exact if contrast sensitivity values were either all relative, or all absolute; otherwise a constant scaling factor will be introduced, but it will not have any impact on the derived channel frequency or bandwidth.

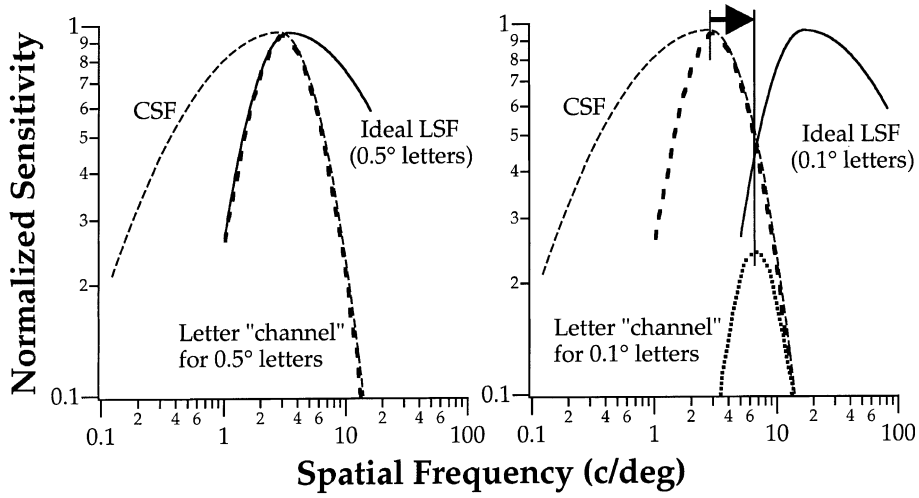


Fig. 11. The observed "spatial-frequency channel" for letter identification by the CSF-ideal-observer model as determined by the interaction between the CSF and the distribution of letter-identity information (ideal LSF). Given a CSF and the ideal LSF, the phantom channel is given by Eq. (B.2). The horizontal position of the ideal LSF, when spatial frequency is expressed in c/deg, is determined by letter size.

A similar derivation can be made for the case of $\mathrm{Llsf_p} < \mathrm{Lcsf_p}$, where

$$\mathrm{Lf_p} = \left(\frac{\alpha_{csf}^2}{\beta_{lsf}^2 + \alpha_{csf}^2}\right)\mathrm{Llsf_p} + \left(\frac{\beta_{lsf}^2}{\beta_{lsf}^2 + \alpha_{csf}^2}\right)\mathrm{Lcsf_p} \quad \text{(B.4b)}$$

Since $\mathrm{Llsf_p}$ differs from the log nominal letter frequency by only a constant, the exponent of the power-function fit between peak tuning frequency and nominal letter frequency is therefore

$$\text{log–log slope} = \begin{cases} \dfrac{\beta_{csf}^2}{\alpha_{lsf}^2 + \beta_{csf}^2} & \text{if } \mathrm{Lcsf_p} < \mathrm{Llsf_p} \\ & \text{(small letters)} \\ \dfrac{\alpha_{csf}^2}{\beta_{lsf}^2 + \alpha_{csf}^2} & \text{if } \mathrm{Llsf_p} < \mathrm{Lcsf_p} \\ & \text{(large letters)} \end{cases} \quad \text{(B.5a)}$$

For the foveal CSF in Fig. 2, the half-bandwidths are $\alpha_{csf} = 3.07$, and $\beta_{csf} = 1.28$. The half-bandwidths for the ideal LSF (Fig. 10) are $\alpha_{lsf} = 1.26$, and $\beta_{lsf} = 2.71$. The exponent of the power function of peak tuning frequency vs. nominal letter frequency therefore lies between 0.50 for small letters and 0.56 for large letters, in excellent agreement with the results of the CSF-ideal-observer model and our human data. It is worth noting that the foveal CSF and the ideal LSF are approximately mirror images of each other, i.e. $\alpha_{csf} \approx \beta_{lsf}$ and $\alpha_{lsf} \approx \beta_{csf}$. This means that regardless of letter size, the exponent is around 0.5, significantly <1 (a value of 1 indicates scale invariance).

The errors associated with estimating the exponent, given the errors in estimating the CSF, can be calculated by taking the partial derivatives of Eq. (B.5a), with respect to $\alpha_{csf}$ and $\beta_{csf}$, leading to the following:

$$\Delta(\text{log–log slope}) = \begin{cases} \left(\frac{2\beta_{csf}\alpha_{lsf}^2}{(\alpha_{lsf}^2+\beta_{csf}^2)^2}\right)(\Delta\beta_{csf}) & \text{if } \mathrm{Lcsf_p} < \mathrm{Llsf_p} \\ & \text{(small letters)} \\ \left(\frac{2\alpha_{csf}\beta_{lsf}^2}{(\beta_{lsf}^2+\alpha_{csf}^2)^2}\right)(\Delta\alpha_{csf}) & \text{if } \mathrm{Llsf_p} < \mathrm{Lcsf_p} \\ & \text{(large letters)} \end{cases} \quad \text{(B.5b)}$$

### *B.2. Bandwidth vs. letter size*

As shown in Fig. 10, the bandwidth of the ideal LSF alone is about 4 octaves. Once the human CSF is taken into account, the resultant phantom channel for letter identification becomes more narrowly tuned (2–3 octaves). In principle, the tuning bandwidth of the phantom channel can be derived by setting Eq. (B.2) to $\log(\text{peak channel gain}) - \log(2)$ and solving for Lf. However, because the tuning function is a piecewise parabolic function, the analytic solution for the tuning bandwidth can appear quite complicated. Nonetheless, we shall proceed with the derivation and show for the most part the relationship between bandwidth and letter size is quite simple.

Consider first the case where $\mathrm{Lcsf_p} < \mathrm{Llsf_p}$. Rewrite Eq. (B.2) and label the three parabolic pieces as $P1$, $P2$ and $P3$, we have

$$\mathrm{LG(Lf)} = \begin{cases} K_1 - \frac{1}{\log(2)}(\mathrm{Lf} - \omega_1)^2/\gamma_1^2 \ (P1), & \mathrm{Lf} \leqslant \mathrm{Lcsf_p} \\ K_2 - \frac{1}{\log(2)}(\mathrm{Lf} - \omega_2)^2/\gamma_2^2 \ (P2), & \mathrm{Lcsf_p} < \mathrm{Lf} \leqslant \mathrm{Llsf_p} \\ K_3 - \frac{1}{\log(2)}(\mathrm{Lf} - \omega_3)^2/\gamma_3^2 \ (P3), & \mathrm{Llsf_p} < \mathrm{Lf} \end{cases} \quad \text{(B.6)}$$

where

$$\omega_1 = \frac{1}{\alpha_{csf}^2 + \alpha_{lsf}^2}\left(\alpha_{csf}^2 \mathrm{Llsf_p} + \alpha_{lsf}^2 \mathrm{Lcsf_p}\right)$$

$$\omega_2 = \frac{1}{\beta_{csf}^2 + \alpha_{lsf}^2}\left(\beta_{csf}^2 \mathrm{Llsf_p} + \alpha_{lsf}^2 \mathrm{Lcsf_p}\right)$$

$$\omega_3 = \frac{1}{\beta_{csf}^2 + \beta_{lsf}^2}\left(\beta_{csf}^2 \mathrm{Llsf_p} + \beta_{lsf}^2 \mathrm{Lcsf_p}\right)$$

$$\gamma_1 = \frac{\alpha_{csf}\alpha_{lsf}}{\sqrt{\alpha_{csf}^2 + \alpha_{lsf}^2}}; \quad \gamma_2 = \frac{\beta_{csf}\alpha_{lsf}}{\sqrt{\beta_{csf}^2 + \alpha_{lsf}^2}}; \quad \gamma_3 = \frac{\beta_{csf}\beta_{lsf}}{\sqrt{\beta_{csf}^2 + \beta_{lsf}^2}}$$

$$K_1 = K - \frac{1}{\log(2)\left(\alpha_{csf}^2 + \alpha_{lsf}^2\right)}\left(\mathrm{Llsf_p} - \mathrm{Lcsf_p}\right)^2$$

$$K_2 = K - \frac{1}{\log(2)\left(\beta_{csf}^2 + \alpha_{lsf}^2\right)}\left(\mathrm{Llsf_p} - \mathrm{Lcsf_p}\right)^2$$

$$K_3 = K - \frac{1}{\log(2)\left(\beta_{csf}^2 + \beta_{lsf}^2\right)}\left(\mathrm{Llsf_p} - \mathrm{Lcsf_p}\right)^2$$

While it is clear that the peak of the tuning function is the peak of the second parabola ($P2$) with a log amplitude of $K_2$, it is not generally true that the line $\mathrm{LG} = K_2 - \log(2)$, which cuts the tuning function at half-amplitude, intersects $P2$. Obviously, this depends on whether the roots of $P2 = K_2 - \log(2)$ are in-between $\mathrm{Lcsf_p}$ and $\mathrm{Llsf_p}$, in which case the bandwidth will simply be $2\gamma_2$ in octaves. From Eq. (B.6), the roots of $P2 = K_2 - \log(2)$ are

$$\omega_2 \pm \log(2)\gamma_2 = \frac{1}{\beta_{csf}^2 + \alpha_{lsf}^2}\left(\beta_{csf}^2 \mathrm{Llsf_p} + \alpha_{lsf}^2 \mathrm{Lcsf_p}\right) \pm \frac{\log(2)\beta_{csf}\alpha_{lsf}}{\sqrt{\beta_{csf}^2 + \alpha_{lsf}^2}} \tag{B.7}$$

Since $\alpha_{lsf}$ and $\beta_{csf}$ are non-negative, the conditions for the roots of $P2$ to be in-between $\mathrm{Lcsf_p}$ and $\mathrm{Llsf_p}$ are therefore

$$\begin{cases} \frac{1}{\beta_{csf}^2+\alpha_{lsf}^2}\left(\beta_{csf}^2 \mathrm{Llsf_p} + \alpha_{lsf}^2 \mathrm{Lcsf_p}\right) - \frac{\log(2)\beta_{csf}\alpha_{lsf}}{\sqrt{\beta_{csf}^2+\alpha_{lsf}^2}} \geqslant \mathrm{Lcsf_p} \\ \frac{1}{\beta_{csf}^2+\alpha_{lsf}^2}\left(\beta_{csf}^2 \mathrm{Llsf_p} + \alpha_{lsf}^2 \mathrm{Lcsf_p}\right) + \frac{\log(2)\beta_{csf}\alpha_{lsf}}{\sqrt{\beta_{csf}^2+\alpha_{lsf}^2}} \leqslant \mathrm{Llsf_p} \end{cases}$$

$$\Rightarrow \begin{cases} \mathrm{Llsf_p} \geqslant \mathrm{Lcsf_p} + \log(2)\frac{\alpha_{lsf}}{\beta_{csf}}\sqrt{\beta_{csf}^2 + \alpha_{lsf}^2} \\ \mathrm{Llsf_p} \geqslant \mathrm{Lcsf_p} + \log(2)\frac{\beta_{csf}}{\alpha_{lsf}}\sqrt{\beta_{csf}^2 + \alpha_{lsf}^2} \end{cases}$$

$$\Rightarrow \mathrm{Llsf_p} \geqslant \mathrm{Lcsf_p} + \log(2)\max\left(\frac{\alpha_{lsf}}{\beta_{csf}}, \frac{\beta_{csf}}{\alpha_{lsf}}\right)\sqrt{\beta_{csf}^2 + \alpha_{lsf}^2} \tag{B.8}$$

This means that if letter size is adequately small such that $\mathrm{Llsf_p}$ is sufficiently large, then the half-amplitude points (roots of $P2 = K_2 - \log(2)$) will both be within the range of $P2$. As a result, the tuning bandwidth is

$$\text{bandwidth} = 2\gamma_2 = \frac{2\beta_{csf}\alpha_{lsf}}{\sqrt{\beta_{csf}^2 + \alpha_{lsf}^2}}, \quad \text{if } \mathrm{Llsf_p} \geqslant \mathrm{Lcsf_p} + \log(2)\max\left(\frac{\alpha_{lsf}}{\beta_{csf}}, \frac{\beta_{csf}}{\alpha_{lsf}}\right)\sqrt{\beta_{csf}^2 + \alpha_{lsf}^2} \tag{B.9a}$$

A similar consideration of the case $\mathrm{Llsf_p} < \mathrm{Lcsf_p}$ shows that for sufficiently large letters, the bandwidth is

$$\text{bandwidth} = \frac{2\alpha_{csf}\beta_{lsf}}{\sqrt{\alpha_{csf}^2 + \beta_{lsf}^2}}, \quad \text{if } \mathrm{Llsf_p} \leqslant \mathrm{Lcsf_p} - \log(2)\max\left(\frac{\beta_{lsf}}{\alpha_{csf}}, \frac{\alpha_{csf}}{\beta_{lsf}}\right)\sqrt{\alpha_{csf}^2 + \beta_{lsf}^2} \tag{B.9b}$$

Note that in these cases where the letter size is either sufficiently small or large, the tuning bandwidth of the phantom letter channel is independent of letter size. It will be left as an exercise for the readers to show that the transition between these two bandwidth plateaus is approximately linearly related to $\mathrm{Llsf_p}$ as follows:

$$\text{bandwidth} = \sqrt{a\mathrm{Llsf_p}^2 + b\mathrm{Llsf_p} + c} \tag{B.10}$$

For subject SC from whom we obtained the CSFs, acuity letter size at the fovea was 0.12°. $\mathrm{Llsf_p}$, with $\mathrm{lsf_p}$ measured in c/letter, is 0.316 log units (Fig. 10). The relationship between $\mathrm{Llsf_p}$, with $\mathrm{lsf_p}$ measured in c/deg, and letter sizes ($s$) in log units above acuity is therefore $\mathrm{Llsf_p} = 0.316 - (\log(0.12) + s)$. For letter sizes of 0.2, 0.4, and 0.6 log units above acuity, the corresponding $\mathrm{Llsf_p}$ were 1.04, 0.84, and 0.64 respectively. Substituting $\mathrm{Lcsf_p} = 0.46$ log units (from Fig. 2 of the main text), $\alpha_{lsf} = 1.26$, and $\beta_{csf} = 1.28$ to Eq. (B.9a), we found that for Eq. (B.9a) to hold, $\mathrm{Llsf_p}$ must be no less than 1.01. Thus, only the condition with letter size at 0.2 log units above acuity reached the regime where bandwidth is independent of letter size. For the rest of the letter sizes, the bandwidth should vary between the two bounds defined by Eqs. (B.9a) and (B.9b), namely, between 1.80 octaves for very small letters (<0.23 log units above acuity or 0.20°), and 4.06 octaves for very large letters (>2.17 log units above acuity or 17.9°). All measured bandwidths for the fovea condition were within and near the lower end of this range (Fig. 5 of the main text).

## References

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

Anderson, R. S., & Thibos, L. N. (1999). Sampling limits and critical bandwidth for letter discrimination in peripheral vision. *Journal of the Optical Society of America A, 16*, 2334–2342.

Banks, M. S., Sekuler, A. B., & Anderson, S. J. (1991). Peripheral spatial vision: limits imposed by optics, photoreceptors, and receptor pooling. *Journal of the Optical Society of America A, 8*, 1775–1787.

Braje, W. L., Tjan, B. S., & Legge, G. E. (1995). Human efficiency for recognizing and detecting low-pass filtered objects. *Vision Research, 35*, 2955–2966.

Chaparro, A., & Young, R. S. (1993). Reading with rods: the superiority of central vision for rapid reading. *Investigative Ophthalmology and Visual Science, 34*, 2341–2347.

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

Chung, S. T. L., Mansfield, J. S., & Legge, G. E. (1998). Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision. *Vision Research, 38*, 2949–2962.

De Valois, R. L., Morgan, H., & Snodderly, D. M. (1974). Psychophysical studies of monkey vision. 3. Spatial luminance contrast sensitivity tests of macaque and human observers. *Vision Research, 14*, 7581.

Ginsburg, A. P. (1980). Specifying relevant spatial information for image evaluation and display design: an explanation of how we see certain objects. *Proceedings of the SID, 21*, 219–227.

Higgins, K. E., Arditi, A., & Knoblauch, K. (1996). Detection and identification of mirror-image letter pairs in central and peripheral vision. *Vision Research, 36*, 331–337.

Landy, M. S., Cohen, Y., & Sperling, G. (1984). HIPS: Image processing under UNIX. Software and applications. *Behavior Research Methods, Instruments, & Computers, 16*, 199–216.

Latham, K., & Whitaker, D. (1996). A comparison of word recognition and reading performance in foveal and peripheral vision. *Vision Research, 36*, 2665–2674.

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.

Millodot, M., & Lamont, A. (1974). Refraction of the periphery of the eye. *Journal of the Optical Society of America, 64*, 110–111.

Näsänen, R., & O'Leary, C. (1998). Recognition of band-pass filtered hand-written numerals in foveal and peripheral vision. *Vision Research, 38*, 3691–3701.

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

Peli, E. (1990). Contrast in complex images. *Journal of the Optical Society of America A, 7*, 2032–2040.

Pelli, D. G., & Zhang, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research, 31*, 1337–1350.

Rubin, G. S., & Turano, K. (1994). Low vision reading with sequential word presentation. *Vision Research, 34*, 1723–1733.

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

Tjan, B. S., Braje, W. L., Legge, G. E., & Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research, 35*, 3053–3069.

Tjan, B. S., & Legge, G. E. (1998). The viewpoint complexity of an object-recognition task. *Vision Research, 38*, 2335–2350.

Virsu, V., & Rovamo, J. (1979). Visual resolution, contrast sensitivity, and the cortical magnification factor. *Experimental Brain Research, 37*, 475–494.