PERGAMON

Vision Research 38 (1998) 3691–3701

Vision Research

Recognition of band-pass filtered hand-written numerals in foveal and peripheral vision

Risto Näsänen *, Claire O'Leary

Department of Optometry and Vision Sciences, University of Wales, College of Cardiff, PO Box 905, Cardiff CF1 3XF, UK

Received 25 July 1997; received in revised form 15 January 1998

Abstract

The purpose of the present study was to find out what differences between foveal and peripheral pattern recognition remain unexplained by the inhomogeneities of retinal sampling and the optics of the eye. We measured contrast thresholds for pattern recognition at different eccentricities. The effects of retinal sampling were homogenised by using M-scaling of the stimuli, and the effects of the optics of the eye were by-passed either by using strong external noise (signal-to-noise ratio is not affected by optical attenuation) or by computing retinal image contrast by means of the optical modulation transfer function. The stimuli were hand-written numerals filtered to two-octave bands of various centre object spatial frequencies (c/object). The results were described as contrast thresholds and recognition efficiency. At all eccentricities, lowest contrast thresholds and highest recognition efficiencies were found at medium object spatial frequencies. At high object spatial frequencies the peripheral retinal contrast thresholds and recognition efficiencies were nearly as good as at the fovea, but at low object spatial frequencies most of the data showed superiority of the fovea to the periphery. Therefore, at high object spatial frequencies peripheral recognition performance could be explained relatively well by the retinal sampling gradient, or equivalently by the cortical magnification factor, together with the effects of the optics of the eye. Some eccentricity dependent deterioration of recognition at low object spatial frequencies remained unexplained.

Keywords: Peripheral vision; Recognition; Efficiency; Noise; Contrast

1. Introduction

According to the cortical magnification theory, peripheral performance is equal to foveal performance if the cortical magnifications of the stimuli are equal [1–4]. In psychophysical experiments, this equalisation of cortical size can be achieved by enlarging peripheral stimuli in inverse proportion to the cortical magnification factor. This procedure is known as M-scaling. The cortical magnification factor, the linear magnification (mm/°) of the visual field at the striate cortex, is directly proportional to the square root of the density of ganglion cell receptive fields [5–7]. Therefore, M-scaling results in the equalisation of the number of retinal samples of the stimuli at different eccentricities.

The cortical magnification theory would require that the optical and retinal processing characteristics at different eccentricities would be perfectly scaled versions of each other, and that the efficiency of the use of information at later stages be independent of eccentricity. However, the studies of Navarro et al. [8], Williams et al. [9], Curcio et al. [10], and Curcio and Allen, [11] show that the effects of optical factors, and cone and ganglion cell densities on spatial resolution change with different rates as the eccentricity increases. The decline of optical quality is much slower than the increase of sampling interval at the ganglion cell level. At the fovea, spatial resolution is limited by optical factors, while in peripheral vision the density of ganglion cells limits resolution [9]. Therefore, the apparent resolution gradient is not as steep as the decrease of sampling frequency or the magnification factor [12]. It follows that scaling stimuli according to the cortical magnification factor should make the peripheral sensitivity to

* Corresponding author. Present address: University of Helsinki, Institute of Biomedicine, Department of Physiology, PO Box 9, FIN-00014 Helsinki, Finland. E-mail: nasanen@kruuna.helsinki.fi.

0042-6989/98/$19.00
PII: S0042-6989(98)00062-5

high spatial frequencies higher than foveal sensitivity. On the other hand, scaling according to the apparent gradient would lead to non-equal number of retinal samples of stimuli at different eccentricities.

One way to avoid the effects of the optical modulation transfer function is to use strong external noise in the stimuli: the effects of optical attenuation on signal and external noise are similar, leaving the signal-to-noise ratio unchanged. This requires that the effect of external noise is large in comparison to the effect of internal noise, that is, external noise is the dominant noise source limiting performance. Another possibility is to use the foveal and peripheral optical modulation transfer functions to compute the retinal image contrast at each eccentricity.

1.1. The purpose of the study

The purpose of the present study was to find out what differences between foveal and peripheral recognition remain when the number of retinal image samples are equalised by using M-scaling and the effects of the variation of optical attenuation is minimised by using external noise. For patterns presented without noise we also estimated contrast thresholds for retinal images by taking into account the optical modulation transfer function at each eccentricity.

We measured human pattern recognition performance at different eccentricities for M-scaled band-pass filtered hand-written numerals presented with and without spatial noise. To provide some natural variability of stimulus shape and to make the task more complex, we chose hand-written numerals as our stimuli. Ten different samples, that is, shape variations, of each numeral were used. Numerals are patterns that are thoroughly learned by adults, and therefore do not require a lengthy training phase before the experiment. To find out how recognition performance is related to different object spatial frequencies [13,14], the numerals were band-pass filtered to various two-octave spatial frequency bands.

The results were compared with the performance of the ideal observer in the same task and expressed, in addition to contrast thresholds, in human recognition efficiency. Efficiency is a measure that compares the performance of a human observer to that of the ideal observer. The ideal observer uses all available information to maximise its performance. Therefore, efficiency gives a description of how much of the available information human observers are able to use in a visual task and stimulus conditions. Efficiency takes into account the fact that in different tasks and when using different stimuli there may be different amounts of information available.

1.2. Some properties of human character recognition

Legge et al. [15] studied the significance of high spatial frequencies in reading. They measured the effect of low-pass filtering (attenuation of high spatial frequencies) on reading speed. If spatial frequencies above two cycles per letter were removed, reading speed still was the same as for unfiltered letters independently of letter size. Reading speed became increasingly slow when the cut-off frequency was decreased under two cycles per letter. This suggests that high object spatial frequencies have little significance in letter recognition.

Parish and Sperling [13] measured the efficiency of recognising band-pass filtered letters in external spatial noise. They used a two-octave band-pass filter and measured efficiency as a function of the centre frequency of the filtered letters. The maximum efficiency (0.42) was obtained when the centre frequency was 1.5 c/object. Below this, at 0.74 c/object, efficiency was zero. Above 1.5 c/letter, efficiency declined gradually to about 0.1 at 20 c/object. At each object spatial frequency, efficiency was independent of stimulus magnification within a range of 32:1.

Using high-pass and low-pass filtered noise, Solomon and Pelli [14] derived the bandwidth used by human observers in letter recognition. They found that noise effectively masked letter recognition only within a spatial frequency band of about two octaves, which they concluded to represent the bandwidth of the filters used by human observers in letter recognition. Solomon and Pelli reported recognition efficiencies of 0.10–0.13 for their unfiltered letters. Tjan et al. [16] obtained roughly similar efficiency values for letter recognition.

The above studies performed in foveal vision suggest that human observers only use a limited band of spatial frequencies in character recognition, human ability to use letter information is best at relatively low object spatial frequencies, and that the efficiency of character recognition is largely scale invariant.

2. Methods

The stimuli were generated by using a 200 MHz Pentium computer with a high quality 17'' computer monitor (Eizo FlexScan F553-M). The graphics board (Diamond Stealth 64 VRAM PCI) was used at a resolution of 640×480 pixels, and its frame rate was 120 Hz. The pixel size of the display was 0.485×0.485 mm^2, and the average photopic luminance was 50 cd/m^2. The non-linearity of the luminance response of the display was corrected by using its inverse function when the stimuli were computed. The measurements were made in a dark room, where the only light source was the monitor.

The graphics board could produce 256 grey levels. In the absence of noise, this number of grey levels is not sufficient for presenting very low-contrast images. To increase the low contrast information and to reduce the effects of quantization errors on the information contents of the displayed images, we used a quasi-periodic dithering technique, which utilises the Bayer [17] dither matrix. Dithering was only used when the images did not contain noise. A detailed description of the dithering technique is given in Appendix A.

The stimuli were hand-written numerals (0–9). We used ten variations of each numeral, which represented the hand-writing of eight different persons. Two persons wrote two sets of numerals. The size of the digitised numerals was 128 × 128 pixels at the fovea and at the eccentricity of 5°. On the display the size was 6.2 cm. The average numeral height was 3.43 cm, which corresponds to 0.82° of visual angle for viewing at a distance of 240 cm used in the foveal measurements. For a Snellen E symbol this size would correspond to a visual acuity value of 0.1. When the eccentricity was 10 or 20°, the image size was 256 × 256 pixels (12.4 × 12.4 cm^2). The numerals were band-pass filtered using a two-octave logarithmic-exponential Fourier filter

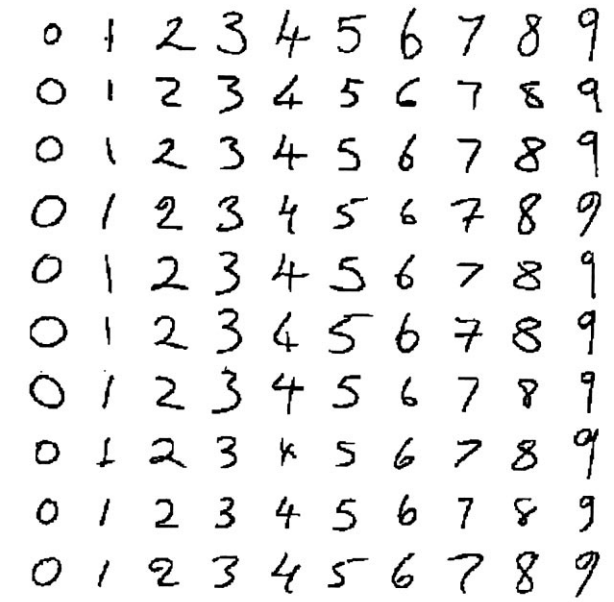
$$H(f) = \exp\{-|\ln(f/f_o)|^3 \ln 2/(b_{1/2} \ln 2)^3\}, \quad (1)$$

where $f_o = \sqrt{(u^2 + v^2)}$, u and v are the spatial frequencies in the horizontal and vertical directions, respectively, f_o is the centre frequency of the filter, and $b_{1/2}$ is half of the filter bandwidth in octaves. Using different centre frequencies ($f_o = 2$, 2.8, 4, 8, 16 and 32 c/image width) of the filter, we produced different pass-bands of the numerals. On a logarithmic frequency scale, the above filter function is symmetric around the centre spatial frequency. Unlike an exponential filter function, such as a Gaussian, the above function has always zero gain at zero spatial frequency. Exponent 3 in the filter function was chosen to make the fall off on both sides of the centre frequency relatively steep and the resulting spatial frequency spectrum well localised in the spatial frequency domain. A filter function with a larger exponent or narrower bandwidth could produce "ringing", repetitive spatial fluctuation of luminance. Because the centre spatial frequency of the resulting images also depends on the spectra of the original images, we computed the centre spatial frequencies (f_c) of the filtered images using the following equation

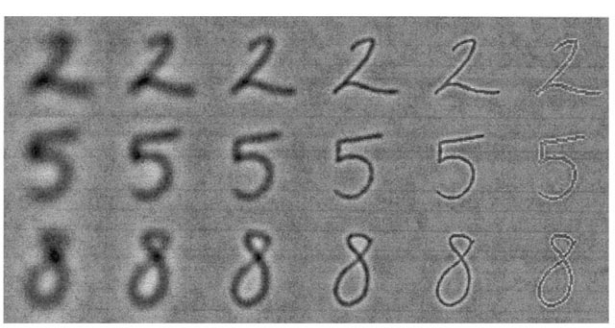
$$f_c = [\Sigma_u \Sigma_v f|F(u,v)|^2]/[\Sigma_u \Sigma_v |F(u,v)|^2], \quad (2)$$

where $|F(u,v)|$ is the Fourier amplitude spectrum of the image, and $f = \sqrt{(u^2 + v^2)}$. Eq. (2) is similar to that used by Parish and Sperling [13]. The average centre spatial frequencies were found to be 1.20, 1.73, 2.46, 4.66, 8.19, and 17.7 c/object height. Examples of the stimuli are shown in Fig. 1.

The stimulus images were presented either with or without white spatial Gaussian noise. The noise was produced by adding a random number to each pixel. The mean and S.D. (RMS contrast) of the Gaussian distribution of the random numbers were zero and 0.2, respectively, except for the highest object spatial frequency (17.7 c/object) for which the S.D. was 0.1. This was because it was difficult to obtain a threshold for the highest object spatial frequency with the higher noise contrast. The spectral density of noise is computed as $N = c_n^2 p^2$, where c_n is noise RMS contrast and p^2 is the pixel area [18]. The spectral density of noise was 94×10^{-6} cm^2 except for the highest object spatial frequency, for which it was 23.5×10^{-6}. Noise spectral density expressed in degrees squared depended on the viewing distance, which varied between 21–240 cm.



A.



B.

Fig. 1. (A) Hand-written numerals collected from eight different persons. Two persons gave two sets of samples of their hand-writing. (B) Some examples of band-pass filtered numerals used in the experiment. The band width was approximately two octaves. The centre spatial frequencies were 1.20, 1.73, 2.46, 4.66, 8.19, and 17.7 c/object height (from left to right). In foveal viewing at a distance of 240 cm, these values corresponded to 1.46, 2.1, 3.0, 5.68, 9.99, and 21.6 c/°. The average size of the numerals was 3.4 cm, which corresponds to 0.82° when viewed at a distance of 240 cm.

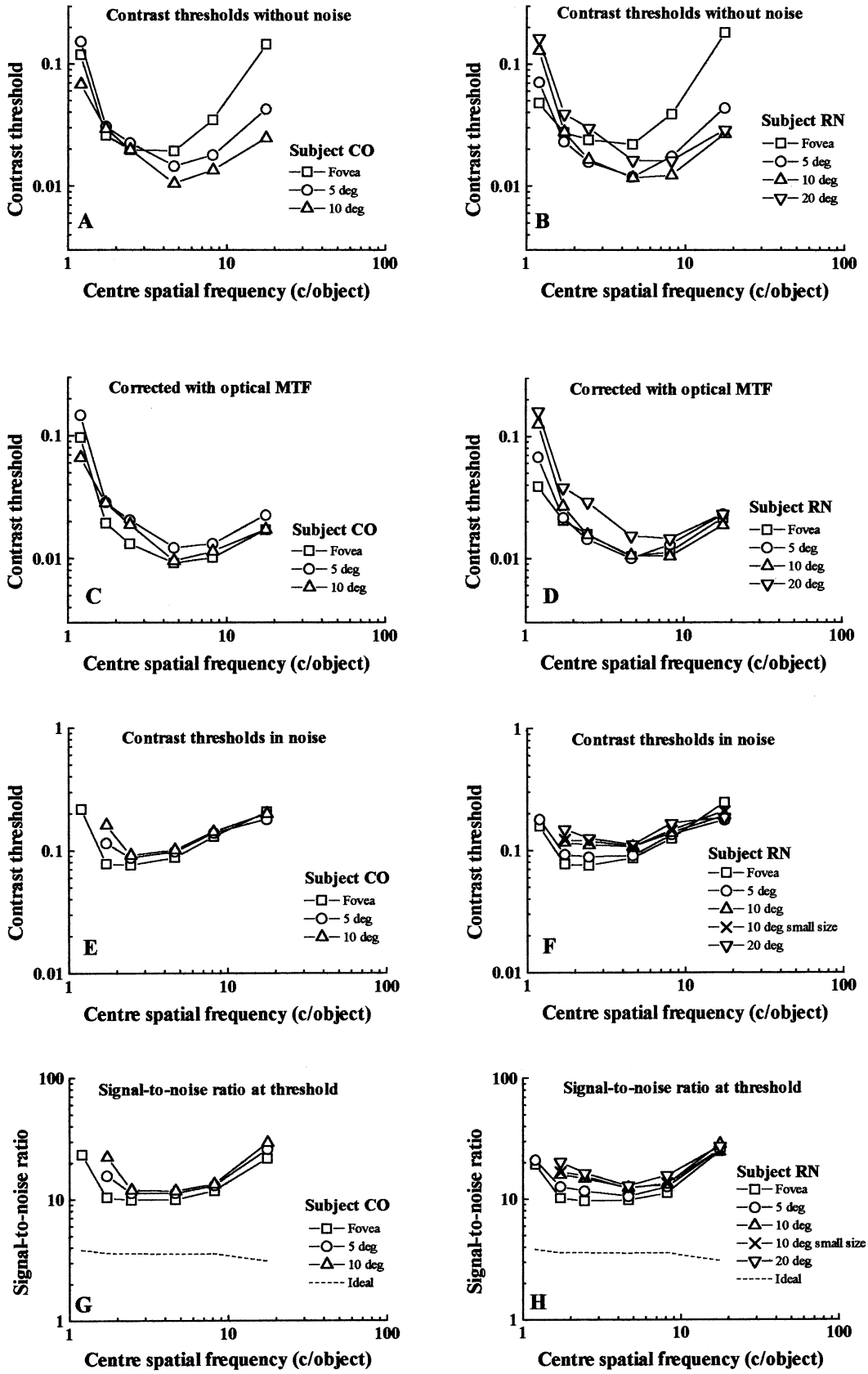


Fig. 2. (*Caption opposite*)

Table 1
Stimulus parameters

Eccentricity (°)	Relative scaling factor	Viewing distance (cm)	Range of spatial frequencies (c/°)	Average pattern height (°)
0	1	240	1.46–21.6	0.820
5	6.5	37	0.225–3.32	5.33
10 large size	12.0	56	0.122–1.8	9.84
10 small size	4.32	39	0.33–5.0	3.54
20	22.6	21	0.0646–0.956	18.5

Contrast energy thresholds were determined at eccentricities 0, 5, 10, and 20° of visual angle of the horizontal meridian. The eccentricity was measured as the angular distance from the fixation point to the centre of the image. In foveal viewing the fixation target was a graphical cross which was swiched off during the stimulus presentation. In peripheral viewing, the fixation target was a small black dot on the face of the monitor. The viewing of the stimuli was binocular with natural pupils of about 4 mm in diameter.

With increasing eccentricity (r) the retinal projections of the stimuli were enlarged (M-scaled) inversely proportionally to the estimate of the cortical magnification factor (M) proposed by Virsu and Hari [19].

$$M^{-1} = 0.0685 + 4.32\ \sin(r) \quad (3)$$

The values of the constants of their estimate have been chosen to agree with human anatomical and psychophysical data [20,21,12,22]. Since the cortical magnification factor can vary meridionally, the above estimate may not be exactly correct for the horizontal meridians used in the present study. However, recognition efficiency should be rather insensitive to errors in magnification if peripheral pattern recognition efficiency is scale invariant as it is at the fovea [13]. To make sure that the choice of the magnification factor estimate did not affect recognition efficiency, one of the observers measured sampling efficiencies at the eccentricity of 10° also with a magnification of the stimulus that was smaller by a factor of 0.36 than that derived from the Virsu–Hari estimate. The viewing distances, stimulus sizes, and spatial frequencies (in c/°) used at each eccentricity are given in Table 1.

Thresholds were determined using a multiple-alternative forced-choice method. The stimuli were presented for 1000 ms. The task of the observer was to indicate to which numeral class the pattern shown belonged. Each numeral class (0–9) consisted of ten variations. Close to the left-hand edge of the screen, there was an array of graphical buttons, one button for each numeral class marked from 0 to 9. To indicate her/his response, the observer pointed and clicked one of the buttons with mouse. This required that the observer first moved her/his fixation from the fixation target to the button array, placed the mouse cursor on the appropriate button, moved fixation back to the fixation target and then pressed the mouse button. The response started a new presentation after a delay of 500 ms. The presentation of the stimulus was indicated by a sound signal. Another sound signal gave feedback about the correctness of the choice of the observer.

After four consecutive correct responses the signal contrast was decreased by a factor of 1.26, and after each incorrect response the contrast was increased by the same factor. A threshold estimate at the probability level of 0.84 of correct answers [23] was obtained as the mean of eight reversals. The number of trials needed for one threshold estimate was 48 on average. Each data point shown in Figs. 2 and (3) represents the arithmetic mean of three threshold estimates. The authors served as observers. Both had normal or corrected to normal vision. There was no systematic training phase before the experiment. However, both subjects (the authors) participated in the collection and processing of the stimulus patterns and, therefore, were familiar with the stimuli before the experiment.

Recognition efficiency (η) e.g. [16] can be defined as the ratio of the contrast energy thresholds for the ideal (E_{ideal}) and human (E_{human}) observers:

$$\eta_{\text{total}} = E_{\text{ideal}}/E_{\text{human}} \quad (4)$$

Contrast energy was computed as follows

$$E = p^2 \Sigma_x \Sigma_y c^2(x, y), \quad (5)$$

Fig. 2. Contrast threshold and signal-to-noise ratio as a function of object centre spatial frequency for band-pass filtered hand-written numerals at different horizontal eccentricities. To compensate for the change of resolution with increasing eccentricity, the retinal sizes of the peripherally viewed patterns were M-scaled, that is, their size was increased in inverse proportion to the cortical magnification factor. (A) and (B) contrast thresholds measured without external noise. (C) and (D) contrast thresholds expressed in retinal image contrast calculated by using the optical MTF's of Navarro et al. [8]. (E) and (F) Contrast thresholds measured in spatial noise. The RMS contrast of noise was 0.2 except for the highest object spatial frequency, for which it was 0.1. (G) and (H) Signal-to-noise ratio at threshold for human and ideal observers. Signal-to-noise ratio is expressed as $[(E\text{-}E_o)/N]^{1/2}$, where E is energy threshold in noise, E_o is energy threshold in the absence of noise, and N is the spectral density of noise. The performance of the ideal observer (Eq. (8)) was obtained with computer simulations.

where $c(x, y)$ is the contrast waveform, and p^2 is the pixel area. Contrast waveform is defined as

$$c(x, y) = (l(x, y) - l_o)/l_o, \quad (6)$$

where $l(x, y)$ is the luminance waveform and l_o is the mean luminance [18].

Sampling efficiency refers to the efficiency of the human pattern recognition mechanism without the effects of internal neural noise and gain factors preceding signal recognition. This can be computed as follows:

$$\eta_{sampling} = E_{ideal}/(E_{human} - E_o), \quad (7)$$

where E_{human} and E_o are human energy thresholds measured with and without external noise, respectively.

The performance of the ideal observer was determined by computer simulations. In the simulations, we used the same threshold estimation algorithm as in the experiments with human observers. The mean of seven thresholds was computed. We used the same ideal observer formulation as Tjan et al. [16] and Braje et al. [24]. The ideal observer maximises the probability of a correct answer by choosing the stimulus class (i) for which a posteriori probability is largest. This is equivalent to finding the maximum of the following:

$$L'_i = \Sigma_j \exp\{-(2\sigma^2)^{-1}\Sigma_x\Sigma_y[s(x, y) - t_{ij}(x, y)]^2\}P(t_{ij}), \quad (8)$$

where $s(x, y)$ is the received signal to be classified, $t_{ij}(x, y)$ is the template for the jth variation in class i, $\Sigma_x\Sigma_y[s(x, y)\text{-}t_{ij}(x, y)]^2$ is the square of the Euclidean distance between the received signal and a template, and σ is the standard deviation of white noise [16]. Templates $t_{ij}(x, y)$ are identical copies of the patterns to be recognised, that is, the templates were filtered in the same way as the stimuli. The probability of the occurrence of each variation of each class ($P(t_{ij})$) was constant. Since there were ten numeral classes and ten variations in each class, $P(t_{ij}) = 1/(10 \times 10) = 0.01$.

3. Results

The contrast threshold data measured in the absence of external noise are shown in Fig. 2A and 2B. For both observers and at each eccentricity contrast thresholds first decreased and then increased with increasing object spatial frequency, which is expressed in c/object height. The contrast threshold curves for different eccentricities were not very similar. At high object spatial frequencies the foveal thresholds were much higher than peripheral thresholds.

Because of M-scaling a constant object spatial frequency corresponds to a decreasing retinal spatial frequency (c/°) as the eccentricity increases. Since the increase of optical degradation with eccentricity is very slow, an M-scaled pattern is less attenuated by optics in the periphery. At high object spatial frequencies the contrast threshold curves of Fig. 2A and 2B should become more similar if we express contrast thresholds according to retinal image contrast, which can be estimated by multiplying the contrast thresholds of Fig. 2A and 2B by the appropriate values of the foveal and peripheral optical modulation transfer functions. For this we used the optical modulation transfer functions measured by Navarro et al. [8] for 4 mm natural pupil.

The estimates of retinal contrast thresholds are shown in Fig. 2C and 2D. At high object spatial frequencies the retinal contrast thresholds were much better superimposed than in Fig. 2A and 2B. The main change was that the foveal thresholds at high object spatial frequencies were clearly lower when expressed in retinal contrast than in external contrast. The effect of the optical modulation transfer function on peripheral contrast thresholds was small. For subject CO the foveal thresholds did not systematically deviate from peripheral thresholds, and the optically corrected M-scaling seems to work relatively well at all spatial frequencies. For subject RN there were systematic differences at the lowest object spatial frequencies. The foveal thresholds were lower than peripheral. This is particularly clear if we compare the data measured at 20° to the foveal data. The average standard error was only 7.4% of the mean, which is much less than the symbol size used in Fig. 2.

Fig. 2E and 2F show the contrast thresholds measured in external noise. At high and medium object spatial frequencies the contrast thresholds for all eccentricities were very similar. There were not such large differences between contrast threshold for different eccentricities as found in the absence of noise (Fig. 2A and 2B). This reflects the fact that the optics of the eye attenuates the contrast of signal and noise in a similar way and, therefore, the signal-to-noise ratio is not affected by optical attenuation.

Fig. 2G and 2H show the data expressed in signal-to-noise ratio at threshold. Signal-to-noise ratio (s/n) was computed as $s/n = \sqrt{[(E - E_o)/N]}$, where E is energy threshold measured in noise (the data of Fig. 2E and 2F expressed in energy thresholds), E_o is energy threshold measured in the absence of noise (the data of Fig. 2A and 2B expressed in energy thresholds), and N is the spectral density of noise. The signal-to-noise ratio functions first decreased and then increased with increasing object spatial frequency. At high object spatial frequencies the signal-to-noise ratio thresholds were relatively similar for all eccentricities, but at low object spatial frequencies the signal-to-noise ratio thresholds increased with eccentricity. The signal-to-noise ratio thresholds are also shown for the ideal observer (the dashed line). The performance of the ideal observer was obtained by using computer simulations. For the ideal observer the signal-to-noise ratio function decreased slightly with object spatial frequency.

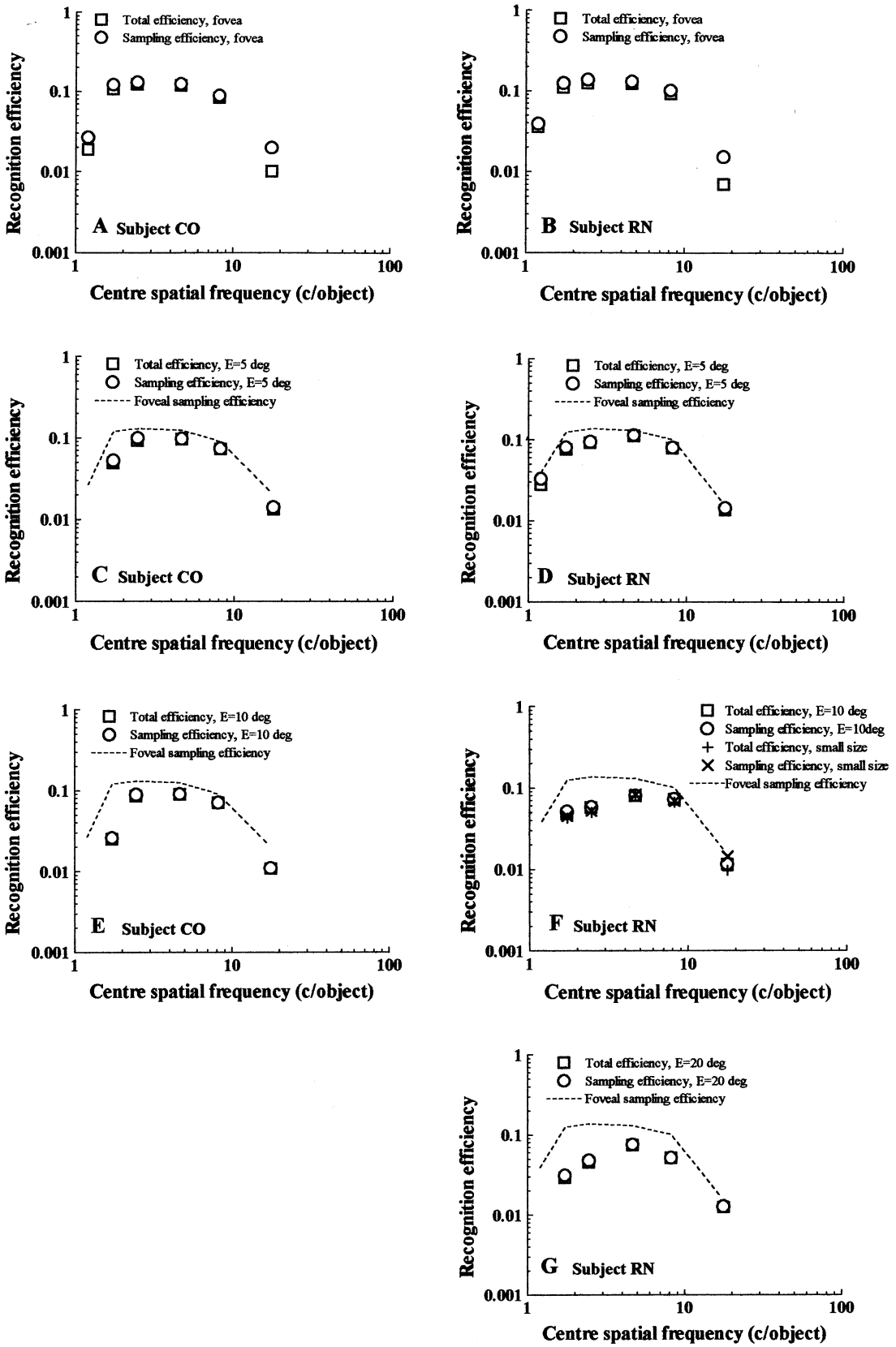


Fig. 3. Recognition efficiency at different eccentricities as a function of centre object spatial frequency. Total and sampling efficiencies were computed from the data of Fig. 2 using Eqs. (4) and (7), respectively.

The foveal efficiencies for the two subjects are shown in Fig. 3A and 3B. Both total and sampling efficiency first increased and then decreased as a function of centre spatial frequency. The highest efficiencies (maximum about 13%) were found at spatial frequencies between 2.5–8 c/object. The total and sampling efficiencies were highly similar. Therefore, internal neural noise played little part in the recognition performance in noise, and the external noise was the dominant noise source limiting performance. The foveal results for the two observers were nearly identical.

In Fig. 3C and 3D, the efficiencies for an eccentricity of 5° are compared with the foveal sampling efficiency function. There appeared to be some decline of recognition efficiency in comparison to the foveal values, particularly in the results of subject CO. For CO it was not possible to determine an energy threshold in noise at the lowest object spatial frequency.

Fig. 3E and 3F show recognition efficiencies measured at an eccentricity of 10°. The peripheral recognition efficiencies at low object spatial frequencies were clearly lower than foveal efficiencies. Both subjects were incapable to measure a threshold for the noisy numerals at the lowest object spatial frequency. For subject RN, additional efficiency functions were determined using a pattern size that was reduced by a factor of 0.36. This change of size did not have any effect on recognition efficiency. Therefore, it seems that, as foveal vision, peripheral vision is scale-invariant in character recognition.

Fig. 3G shows recognition efficiencies at an eccentricity of 20° for subject RN. The trend found at smaller eccentricities continues at this eccentricity; there is a further reduction of recognition efficiency at low object spatial frequencies but little change at high object spatial frequencies.

For the whole data in Fig. 3, the average standard errors were 10 and 13% of the mean for subjects RN and CO, respectively. That is less than half the symbol size used in Fig. 3.

4. Discussion

In all stimulus conditions, contrast thresholds were lowest and recognition efficiencies were highest at medium object spatial frequencies. In the absence of noise the foveal thresholds for high object spatial frequencies were much higher than in the periphery. When the contrast thresholds were transformed to retinal image contrast by applying the optical modulation transfer function, the high object spatial frequency performance became fairly similar at all eccentricities studied.

In external noise the foveal and peripheral contrast thresholds were similar to each other at high object spatial frequencies. At low object spatial frequencies the performance was better at the fovea than in the periphery. Below the eccentricity of 10°, the differences were relatively small. At high object spatial frequencies recognition efficiency was almost similar at all eccentricities studied. However, at low object spatial frequencies recognition efficiency decreased with eccentricity.

4.1. Comparison of the results with other studies

Our result that the human ability to use image information in character recognition is best at medium object spatial frequencies is in qualitative agreement with Parish and Sperling [13]. Our maximum efficiency (about 0.13) was smaller than that (0.42) found by Parish and Sperling [13], but did not deviate much from those found by Solomon and Pelli [14] (0.1–0.13%) and Tjan et al. [16] (0.12–0.16%) for unfiltered letters. The difference in maximum efficiency of the present study and that of Parish and Sperling may be due to differences in the details of the experiments. The letters used by Parish and Sperling were made of horizontal and vertical bars and, therefore, were geometrically simpler and more regular than our hand-written numerals. It may be that simplicity and regularity affects human efficiency. Another difference is that Parish and Sperling used band-pass filtered noise while we used white noise.

Strasburger et al. [25,26] measured contrast thresholds in the absence of noise for unfiltered numerals of a single typeface as a function of pattern size at various eccentricities. At small pattern sizes the thresholds decreased with increasing size and then reached a plateau. Both the point of transition from the decrease to the plateau and the height of the plateau increased with eccentricity. Therefore, it was not possible to equate thresholds across the visual field by horizontal shifts of the curves, that is, by spatial scaling. The foveal study of Solomon and Pelli [14] shows that the recognition of unfiltered characters is based on the utilisation of a band of relatively low object spatial frequencies. If this applies also to peripheral vision, the performance change for low object spatial frequencies can explain the reduction of recognition performance for unfiltered numerals with increasing eccentricity.

4.2. Importance of optical factors

The change of performance with eccentricity is affected both by the optical transfer function of the eye and the eccentricity dependent changes of the neural processing stages. The optical quality of the retinal image deteriorates more slowly than the sampling interval increases at the ganglion cell level. This means that the sensitivity for M-scaled stimuli of high 'cortical' spatial frequency (c/mm) increases with increasing ec-

centricity. Our results are in agreement with this and demonstrate the importance of taking the optical factors into account when peripheral vision is studied.

4.3. The choice of M-factor

There are many estimates of the cortical magnification factor. The one used in this study [19] is fitted to be in agreement with many other estimates [20,21,12,22]. Yet there is a possibility that it is not exactly correct. Errors in the estimate of cortical magnification factor could have affected the results measured in the absence of external noise. However, in spatial noise human foveal efficiency is largely insensitive to scale changes [13]. In the present study we used two magnifications (differing by a factor of 0.36) at the eccentricity of 10°. The recognition performances and efficiencies measured in noise were nearly perfectly similar for these two magnifications. Therefore, it seems that the choice of the magnification factor could not affect the results measured in noise and expressed in terms of efficiency.

4.4. The decrease of performance at low object spatial frequencies

In the literature there are different hypotheses as to why in some cases peripheral discrimination is worse than foveal even when M-scaling is used. These do not seem to be able to account for the present findings, however. Undersampling in peripheral retina e.g. [9] or at some higher stage could, in principle, reduce sampling efficiency, but it cannot explain why efficiency is reduced more at low than at high object spatial frequencies. Uncalibrated disarray of retino-cortical mapping in peripheral vision was proposed by Hess and Field [27] and Hess and McCarthy [28] to account for peripheral positional discrimination. However, positional disarray or jitter would be expected to affect more at high than at low spatial frequencies. Therefore, it cannot explain the present pattern recognition results. Jüttner and Rentschler [29] suggested that in peripheral vision the dimensionality of the internal stimulus representation or feature space used for pattern recognition is reduced. This means that peripherally viewed images would be described internally with a smaller number of features than foveally viewed images. This could reduce sampling efficiency, but it does not explain why low object spatial frequencies are affected more than high.

Our result, that with increasing eccentricity, recognition efficiency remains similar at high object spatial frequencies but decreases at low object spatial frequencies means that the optimal object spatial frequency increases and the useful range of spatial frequencies for recognition becomes narrower. Why such a change occurs is unclear, however.

4.5. Conclusions

The present study shows that the recognition of isolated patterns in peripheral vision does not perfectly obey the cortical magnification theory even when the effects of optics are taken into account. The optical modulation transfer function and the cortical magnification factor explain recognition performance for high object spatial frequencies relatively well, but some eccentricity dependent deterioration of performance at low object spatial frequencies remains unexplained. The selective decrease of sampling efficiency at low object spatial frequencies and the narrowing of the useful bandwidth in peripheral vision may account for the reduced peripheral recognition performance for broad band unfiltered patterns.

Appendix A. Reducing quantization errors by dithering

The dithering algorithm is

$$s_q(x, y) = \mathrm{int}\{s(x, y) + d(x, y)\}, \quad (9)$$

where $s_q(x, y)$ is the displayed quantized signal, $s(x, y)$ is the original signal, $d(x, y)$ is the dither signal and x and y refer to the pixel position. The int{.} operator means rounding to the nearest integer. The dither signal used was based on the Bayer 8×8 dither matrix [17] (Fig. 4). It was used periodically with a period of 8 pixels so that, in each period, the original Bayer matrix or one of its four mirror image transformations in the vertical, horizontal or the two diagonal directions was chosen at random. This technique, which is analogous to the quasi-periodic dithering technique suggested by Allebach and Liu [30] for digital half-toning, makes the resulting error signal more like high spatial frequency noise instead of a deterministic texture. The quasi-periodic dithering has the advantage over the ordered periodic dithering that it is less susceptible to aliasing.

1	33	9	41	3	35	11	43
49	17	57	25	51	19	59	27
13	45	5	37	15	47	7	39
61	29	53	21	63	31	55	23
4	36	12	44	2	34	10	42
52	20	60	28	50	18	58	26
16	48	8	40	14	46	6	38
64	32	56	24	62	30	54	22

Fig. 4. The 8×8 Bayer dither matrix. The matrix was used in 8×8 pixel periods. In each 8×8 block, the above matrix or its mirror-image in vertical, horizontal, or the two diagonal directions was chosen at random. The dither signal ($d(x, y)$) is obtained by transforming the numbers in the matrix ($m(x, y)$) as $d(x, y) = \{[m(x, y) - 1]/n - 0.5(n - 1)/n\}$, where n is the number of elements in the dither matrix.

On the other hand, it is better than simple non-periodic random dither (the value of the dither signal at each position is drawn from a uniform random distribution with a range equal to one quantization step), since it does not produce low spatial frequency noise to the quantized image.

The dither in the displayed images was completely invisible at all viewing distances. Its amplitude was one grey level step, and, close to the mean luminance, its contrast (computed as $(L_{max} - L_{min})/(L_{max} + L_{min})$) was about 0.005. The spatial frequency spectrum of the resulting error signal had a mean spatial frequency of about 6.2 c/cm (0.301 c/pixel) and a bandwidth of about one octave. At a viewing distance of 240 cm, used in the foveal measurements of the main experiments of this paper, the mean spatial frequency of the error signal corresponds to 26 c/°. For our experimental stimuli the mean spatial frequency of the error signal corresponds to 39 c/object.

When averaged over the area of the 8×8 pixels period, dithering increases the number of grey levels by a factor of 64 resulting in a number equal to $64 \times 256 = 16384$ (14 bits). If the viewing distance is long enough, the optics of the eye filters out the high spatial frequency components related to dithering, and there truly is the above number of grey levels in the retinal image of the stimulus. Otherwise, we have an error signal (the intended signal minus the actual signal) related to dither. The error signal, which may be regarded as noise, has an adverse effect on perceived image quality and contrast thresholds only if its effect is large enough in comparison to the effect of the internal noise of the visual system.

The question as to whether dither affects image quality in a detrimental way can be studied experimentally. We compared our dithering method with another equipment utilising the video-attenuation technique of Pelli and Zhang [31]. This technique increases the number of grey levels by combining the information carried by all three colour channels each attenuated in a suitable way. Gabor gratings having a width of 5° and spatial frequencies of 1, 2, 4, and 8 c/° (0.5, 1, 2, and 4 c/cm) were viewed at a distance of 114 cm. The task was to discriminate between vertical and horizontal gratings. The Michelson contrast thresholds obtained with both systems were nearly identical. The thresholds varied between 0.003 and 0.006 depending on spatial frequency. The lowest threshold was measured at 2 c/° and the highest at 1 and 8 c/°. It seems that the dithering technique can be safely used at least for contrasts above 0.003. The lowest contrast threshold measured in the main experiments was about 0.01.

References

[1] Rovamo J, Virsu V, Näsänen R. Cortical magnification factor predicts the photopic contrast sensitivity of peripheral vision. Nature 1978;271:54–6.

[2] Koenderink JJ, Bouman MA, de Mesquita B, Slappendel S. Perimetry of contrast detection thresholds of moving spatial sine wave patterns. III. The target extent as a sensitivity controlling parameter. J Opt Soc Am 1978;68:854–60.

[3] Virsu V, Rovamo J. Visual resolution, contrast sensitivity, and the cortical magnification factor. Exp Brain Res 1979;37:475–94.

[4] Virsu V, Näsänen R, Osmoviita K. Cortical magnification and peripheral vision. J Opt Soc Am A 1987;4:1568–78.

[5] Drasdo N. The neural representation of visual space. Nature 1977;266:554–6.

[6] Drasdo N. Receptive field densities of the ganglion cells of the human retina. Vis Res 1989;29:985–8.

[7] Wässle H, Grünert U, Röhrenbeck J, Boycott BB. Cortical magnification factor and the ganglion cell density of the primate retina. Nature 1989;341:643–6.

[8] Navarro R, Artal P, Williams DR. Modulation transfer of the human eye as a function of retinal eccentricity. J Opt Soc Am A 1993;9:201–12.

[9] Williams DR, Artal P, Navarro R, McMahon MJ, Brainard DH. Off-axis optical quality and retinal sampling in the human eye. Vis Res 1996;36:1103–14.

[10] Curcio CA, Sloan KR, Kalina RE, Hendrickson AE. Human photoreceptor topography. J. Comp Neurol 1990;292:497–523.

[11] Curcio CA, Allen KA. Topography of ganglion cells in human retina. J Comp Neurol 1990;300:5–25.

[12] Drasdo N. Neural substrates and threshold gradients of peripheral vision. In: Cronly–Dillon JR, editor. Vision and visual dysfunction. Basingstoke: Mcmillan. 1991;5:250–64.

[13] Parish DH, Sperling G. Object spatial frequencies, retinal spatial frequencies, and the efficiency of letter discrimination. Vis Res 1991;31:1399–415.

[14] Solomon JA, Pelli DG. The visual filter mediating letter identification. Nature 1994;369:395–7.

[15] Legge GE, Pelli DG, Rubin GS, Schleske MM. Psychophysics of reading-I. Normal vision. Vis Res 1985;25:239–52.

[16] Tjan BS, Braje WL, Legge GE, Kersten D. Human efficiency for reconizing 3-D objects in luminance noise. Vis Res 1995;35:3053–69.

[17] Bayer BE. An optimum method for two-level rendition of continuous-tone pictures. In IEEE International Conference on Communication, Conference Record, Session 26, 1973 pp. 11–15.

[18] Legge GE, Kersten D, Burgess AE. Contrast discrimination in noise. J Opt Soc Am A 1987;4:391–404.

[19] Virsu V, Hari R. Cortical magnification, scale invariance and visual ecology. Vis Res 1996;36:2971–7.

[20] Cowey A, Rolls ET. Human cortical magnification factor and its relation to visual acuity. Exp Brain Res 1974;21:447–54.

[21] Levi DM, Klein SA, Aitsebaomo AP. Vernier acuity, crowding and cortical magnification. Vis Res 1985;25:963–77.

[22] Grüsser OJ. Migraine phosphenes and the retino-cortical magnification factor. Vis Res 1995;35:1125–34.

[23] Wetherill GB, Levitt H. Sequential estimation of points on a psychometric function. Br J Math Stat Psychol 1965;18:1–10.

[24] Braje WL, Tjan BS, Legge GE. Human efficiency for recognizing and detecting low-pass filtered objects. Vis Res 1995;35:2955–66.

[25] Strasburger H, Harvey LO Jr, Rentschler I. Contrast thresholds for identification of numeric characters in direct and eccentric view. Percept Psychophys 1991;49:495–508.
[26] Strasburger H, Rentschler I, Harvey LO Jr. Cortical magnification theory fails to predict visual recognition. Eur J Neurosci 1994;6:1583–8.
[27] Hess RF, Field DJ. Is the increased spatial uncertainty in the normal periphery due to spatial undersampling or uncalibrated disarray? Vis Res 1993;33:2663–70.
[28] Hess RF, McCarthy J. Topological disorder in peripheral vision. Vis Neurosci 1994;11:1033–6.
[29] Jüttner M, Rentschler I. Reduced perceptual dimensionality in extrafoveal vision. Vis Res 1996;36:1007–22.
[30] Allebach JP, Liu B. Random quasiperiodic halftone process. J Opt Soc Am 1976;66:909–17.
[31] Pelli DG, Zhang L. Accurate control of contrast on microcomputer displays. Vis Res 1991;31:1337–50.