Vision Research 90 (2013) 57–67

Contents lists available at ScienceDirect

Vision Research

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

Accounting for the phase, spatial frequency and orientation demands of the task improves metrics based on the visual Strehl ratio ☆

Laura K. Young [a,*], Gordon D. Love [b], Hannah E. Smithson [a]

[a] Department of Experimental Psychology, University of Oxford, 9 South Parks Road, Oxford OX1 3UD, UK
[b] Centre for Advanced Instrumentation, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

ARTICLE INFO

Article history:
Available online 20 July 2013

Keywords:
Optical distortions
Visual Strehl ratio
Ocular aberrations
Letter identification
Reading
Visual performance metrics

ABSTRACT

Advances in ophthalmic instrumentation have allowed high order aberrations to be measured *in vivo*. These measurements describe the distortions to a plane wavefront entering the eye, but not the effect they have on visual performance. One metric for predicting visual performance from a wavefront measurement uses the visual Strehl ratio, calculated in the optical transfer function (OTF) domain (VSOTF) (Thibos et al., 2004). We considered how well such a metric captures empirical measurements of the effects of defocus, coma and secondary astigmatism on letter identification and on reading. We show that predictions using the visual Strehl ratio can be significantly improved by weighting the OTF by the spatial frequency band that mediates letter identification and further improved by considering the orientation of phase and contrast changes imposed by the aberration. We additionally showed that these altered metrics compare well to a cross-correlation-based metric. We suggest a version of the visual Strehl ratio, $VS_{combined}$, that incorporates primarily those phase disruptions and contrast changes that have been shown independently to affect object recognition processes. This metric compared well to VSOTF for letter identification and was the best predictor of reading performance, having a higher correlation with the data than either the VSOTF or cross-correlation-based metric.

1. Introduction

The first potential limitation to visual performance is that imposed by the optical components of the eye. Imperfections in these components distort the wavefront of light entering the eye and cause degradations to the image that they form on the retina. It is possible to measure the ocular wavefront error but there is no simple link between wavefront error and impairments in visual performance. Our aim in this paper is to compare wavefront-based metrics of visual image quality with performance on two tasks (letter identification and reading). We seek to improve predictions by understanding the type of information captured by the metric and the specific requirements imposed by different visual tasks.

An improved understanding of the link between wavefront error and visual impairment has clear benefits for clinical practice, both in characterising the extent of functional impairment and in considering possibilities of correction. While it is possible to provide a static, open-loop correction of the higher order aberrations of the eye for the purposes of improving vision (Chen et al., 2007; Gao et al., 2009; Jeong & Yoon, 2006; López-Gil et al., 2002, 2003; Marsack et al., 2002, 2007; Marsack, Parker, & Applegate, 2008; Navarro et al., 2000; Netto, Dupps, & Wilson, 2006; Sabesan et al., 2007; Yoon et al., 2004), this is difficult and residual aberrations may still remain. To assess the potential benefit of correcting the wavefront we need to know how to translate a wavefront measurement into real-life performance changes. Also, it is possible to accidentally introduce aberrations (see Applegate & Howland, 1997, for example) and it is important to understand the potential impairments this may cause.

1.1. Predicting visual performance from a wavefront measurement

Current metrics designed to predict visual performance from wavefront measurements fall into two broad categories; those that use an optical quality metric, such as the visual Strehl ratio (Cheng, Thibos, & Bradley, 2003; Thibos et al., 2004), and those that perform a template-matching analysis, such as the cross-correlation method derived by Watson and Ahumada (2008) or the Bayesian model introduced by Nestares, Navarro, and Antona (2003) and further developed by Dalimier and Dainty (2008) and Dalimier et al. (2009). The cross-correlation model implements template-matching explicitly and the Bayesian model implements it by decomposing the image into spatial frequency and orientation

* Corresponding author.
E-mail address: laura.young@psy.ox.ac.uk (L.K. Young).

0042-6989/$ - see front matter © 2013 The Authors. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.visres.2013.06.007

channels. In the next two paragraphs we briefly discuss template matching (via a cross-correlation), and compare this approach to a range of visual Strehl ratio metrics. In later sections, we consider possible improvements to metrics that are based on the visual Strehl ratio.

1.2. Cross-correlation

We have previously employed a template-matching technique based on cross-correlation. In this paper we use this metric for comparison since it has worked well in predicting impairments in reading performance (Young et al., 2011) and the increase in log contrast threshold for letter identification (Young, Love, & Smithson, 2013). Cross-correlation based methods can give predictions of performance for specific stimuli, taking their size and spatial relationships into account. One limitation is that this method gives low values for aberration-induced transformations that change the spatial extent of a letter but maintain geometric similarity. In these conditions spatial overlap is reduced but many letter features (e.g. closed forms and intersections) are preserved and visual performance may not be as impaired as the metric would suggest. A cross-correlation model is likely to perform well for predicting performance outcomes using specific sets of stimuli but it is not the most efficient method for predicting real-life performance since it must necessarily be repeated for every stimulus that an individual will encounter. For text this would mean every letter, symbol and number, in a variety of fonts. Although acuity measures also vary with font, for example, a more efficient, but perhaps less accurate, metric would be one that does not take the specific stimulus into account. The visual Strehl ratio, which we describe in the next section, is one such metric.

1.3. Computational optics and the visual Strehl ratio

For spatially incoherent light, the point spread function (PSF) of an optical system is the squared modulus of the Fourier transform of its complex pupil function, where the wavefront is the phase of that pupil function. Howland and Howland (1977) were the first to describe this wavefront error by a set of orthogonal basis functions called Zernike polynomials, which have been standardised for ophthalmology (ANSI Z80.28, 2010; ISO 24157, 2008; Thibos et al., 2000). In our experiments we have tested the effects of three Zernike modes; defocus $\left(Z_2^0\right)$, coma $\left(Z_3^1\right)$ and secondary astigmatism $\left(Z_4^2\right)$.

The PSF is real-valued and quantifies the appearance of a point source imaged through the system. The OTF, which is the Fourier transform of the PSF, quantifies how spatial frequencies are transmitted by the system and is therefore suitable for analysing the effects of aberrations on extended objects. The OTF is complex and can be split into its magnitude components, (modulation transfer function, MTF) and phase components (phase transfer function, PTF):

$$\mathrm{OTF} = \mathrm{MTF}\; e^{i\mathrm{PTF}}. \tag{1}$$

Alternatively the OTF can be considered in terms of its real and imaginary parts,

$$\mathrm{OTF} = \mathrm{MTF}\mathit{cos}(\mathrm{PTF}) + i\mathrm{MTF}\mathit{sin}(\mathrm{PTF}). \tag{2}$$

Strehl ratio is defined as the ratio of the peak value in the aberrated PSF to the equivalent value in a diffraction-limited PSF. The visual Strehl ratio compares the OTF of a system with its diffraction-limited equivalent. The strength of visual-Strehl-ratio-based metrics is that they also attempt to take visual processing factors into account. This is done by using a frequency-dependent weighting of the OTF according to the neural contrast sensitivity function (NCSF). Thibos et al. (2004) tested 33 metrics for predicting subjective refraction from wavefront measurements of which 10 were calculated using the OTF. Marsack, Thibos, and Applegate (2004) and Cheng, Bradley, and Thibos (2004) also used these same metrics to predict visual performance. All three studies agreed that the best of these metrics for predicting visual performance from the wavefront measurement was the visual Strehl ratio computed in the OTF domain (VSOTF). Other studies have also found good correlation between visual Strehl ratio-based metrics and visual acuity (Buehren & Collins, 2006; Bürhen et al., 2009; Legras & Rouger, 2008; Shi et al., 2011; Tarrant, Roorda, & Wildsoet, 2010) and Ravikumar, Sarver, and Applegate (2012) additionally showed that this correlation is independent of pupil size.

The OTF is a 2-D complex function and the VSOTF reduces this to a single value by only using the real part and integrating over frequency. For a real-valued PSF the negative frequency components of its OTF are complex conjugates of their positive counterparts. Therefore calculating the VSOTF by integrating the imaginary part of the OTF over all frequencies would give a value of zero. It should be noted that for even aberrations, such as defocus, the OTF is entirely real, causing phase shifts of either 0° or 180°. However this is not true for odd aberrations, such as coma, that cause phase shifts between 0° and 180° and so have a non-zero imaginary part. Using the real part of the OTF assumes a cosine-phase weighting on the influence of contrast changes (see Eq. (2)). For even aberrations this produces a weighting of ±1 but for odd aberrations the weighting lies between −1 and 1. Phases of 90° or 270° contribute a weighting of 0 to the real part of the OTF, losing any information about the contrast of these components. Additionally, negative weighting on the real part of the OTF implies that contrast at those frequencies impairs perception beyond the effect of removing contrast, and has the consequence that the VSOTF is not bounded between 0 and 1.

It is perhaps more appropriate to examine phase and contrast separately rather than use the real part of the OTF, which in itself is not a physically measurable or visually relevant quantity. Our intention is not simply to create a real, single-valued metric that incorporates phase, which could be achieved by transforming the (filtered) OTF back to the image domain and calculating a Strehl ratio based on the PSF. Instead we have incorporated phase in the metric in a way that is suggested by psychophysical estimates of the differential effects of particular phase disruptions (Ravikumar, Bradley, & Thibos, 2010). We suggest an alternative to using the real part that uses a linear phase weighting (with 0° phase shifts contributing 1 and 180° phase shifts contributing 0) multiplied by the modulation. This complements the findings of Ravikumar, Bradley, and Thibos (2010) that 180° phase shifts have a large impact on acuity for single letters, letter clusters and faces, and it also allows all modulations at phase shifts up-to 180° to contribute positively to the visual Strehl ratio. Whether a linear weighting on phase angle is appropriate for predicting visual performance is beyond the scope of this paper, although we note that other relationships did not give as a high a correlation with our data. Furthermore, as 180° phase shifts give a weighting of 0 we lose information about the effects of the contrast of these components. However, since 180° phase shifts have a significant negative effect on performance compared to smaller shifts, using this criterion as the zero-weighting value will at least capture something of this known aspect of visual performance.

1.4. The spatial frequency band mediating letter-based tasks

The benefit of the visual Strehl ratio is that it takes neural, as well as optical, effects into account. However, it does not consider the type of stimulus, as the cross-correlation-based metric does, nor the perceptual task performed with that stimulus type. We improve predictions of performance based on the visual Strehl ratio by taking the stimulus and task into account. We do this in a general way by considering how particular classes of object are processed by the visual system.

In the VSOTF calculation the OTF is weighted by the NCSF. The NCSF is the contrast sensitivity function of the visual system ignoring optical factors and has been defined as the observer's contrast sensitivity function (CSF) divided by the MTF of their eye (Campbell & Green, 1965). There is evidence that the human contrast sensitivity function (CSF) is the envelope of multiple overlapping spatial–frequency-tuned channels each of which is selectively sensitive to a narrow range of spatial frequencies (Campbell & Robson, 1968). One might expect that, since letters are broadband stimuli, their identification would be mediated by a broadband channel or by multiple narrow-band channels each passing some signal. Using critical-band masking Solomon and Pelli (1994) showed that, while an ideal observer model would predict a low-pass filter for letter identification, human observers demonstrated that this is in fact mediated by a single band of spatial frequencies centred at 3 cycles per letter with a bandwidth of 1 octave. To quote Oruç and Landy (2009): "The sensitivity profile of this mechanism [of letter identification] is not merely an indication of the stimulus information available to the observer to perform the task". Similar results have been shown by other authors (Alexander, Xie, & Derlacki, 1994; Chung, Legge, & Tjan, 2002; Chung, Levi, & Legge, 2002; Ginsburg, 1980; Majaj et al., 2002; Parish & Sperling, 1991). Majaj et al. (2003) have found evidence that the same spatial frequency band also mediates reading and Chung and Tjan (2009) confirmed that the spatial frequency characteristics of letter identification and reading were closely matched. We propose that to predict performance in letter-based tasks, the OTF should be weighted by the spatial frequency band mediating letter identification, rather than by the NCSF. Importantly we are not simply accounting for the spatial frequencies in the stimulus but rather we are accounting for the mechanism by which the stimulus is identified. This is useful in terms of creating a single-valued metric for predicting visual performance because specific knowledge of exactly which stimulus is presented is not required; we need only to know that the stimulus is a letter and to know its size.

1.5. Orientation effects

We test visual Strehl and cross-correlation metrics ("confusability") against empirical measures of letter recognition and reading performance. Clearly there are likely to be task-specific effects that are difficult to capture in a single metric and we have recently shown that there are task-specific differences between the three types of aberration we have tested (Young, Love, & Smithson, 2013). Specifically we have shown that defocus and secondary astigmatism increasingly impair word recognition with increasing rms amplitude, while coma does not (Young et al., 2011). As coma gave a greater impairment at lower confusability values we have suggested that coma more likely has effects associated with crowding and with eye movements commensurate with the way in which it smears text in one direction, filling in the spaces between letters and words, but does not effect the forms of letters as severely as defocus or secondary astigmatism. Although we did not test for effects of crowding we have shown that differences in the correlation between confusability and impairment for coma compared to defocus and secondary astigmatism disappear in a single-letter identification task.

Crowding (Stuart & Burian, 1962) is a special case of masking in which letter identification is impaired by the close proximity of other letters (see Levi, 2008; Pelli, Palomares, & Majaj, 2004, for a review) and which increases significantly with eccentricity from the fovea (Bouma, 1970). In reading there can be crowding between the component letters of words and between the words themselves (Arditi, Knoblauch, & Grunwald, 1990; Chung, 2004). Letter spacing is important since below a critical separation reading speed is increasingly impaired as letters are presented closer together (Chung, Legge, & Tjan, 2002). Also, it has been shown that letter spacing can have an effect on fixation durations in reading and the planning of eye movements (e.g. Rayner, 1998). Cross-correlation metrics capture the effects aberrations have on single letters (and consequently their component features) but they do not account for changes to letter spacings. To attempt to include this difference between letter identification and reading we additionally modify the calculation of the visual Strehl ratio to more strongly weight phase and contrast changes with a horizontal orientation, since for the text we used in the reading experiment, it is these changes that reduce the white-space between letters and between words.

In summary we will test the ability of single metrics derived from a wavefront measurement to predict visual performance changes. We attempt to improve these predictions by considering:

- The most appropriate way to combine contrast and phase effects.
- The spatial frequency requirements of the task.
- The orientation constraints of the task.

2. Methods

2.1. Experimental data

We compare our predictions of performance to empirical data obtained during two separate tasks: letter recognition (Young, Love, & Smithson, 2013) and reading (Young et al., 2011). The experimental methods are summarised in the following two sections.

2.2. Letter recognition

In a recent experiment (Young, Love, & Smithson, 2013) we studied the effects defocus, coma and secondary astigmatism on single letter identification. Letters (lower case, courier font) with a width of 1° (equal to a Snellen acuity of 20/240) were rendered with either 0.5, 0.6, 0.7, 0.8 or 0.9 μm rms (pupil diameter = 2.5 mm) of one of the three types of aberration via a convolution with the PSF of the aberration. We note that these amplitudes are larger than would typically be measured in the normal population. However, in the context of our experiment they are applied to large letters. We confirmed that the stimuli used in the letter experiment and in the reading experiment were approximately equivalent using a normalised cross correlation. When scaled down further, to letter sizes near to the acuity limit, the equivalent amplitude of aberration is approximately 0.1 μm, which is much closer to amplitudes that might be measured in a normal eye. Stimuli were viewed via a 2.5 mm diameter pinhole to minimise the effects of our subjects' own aberrations on the retinal image and to maintain a constant retinal illumination. Log (Weber) contrast threshold for letter identification was determined by an adaptive staircase procedure (MLPEST using the Matlab Palamedes Toolbox Prins & Kingdom, 2009) using all 26 letters of the alphabet chosen at random on each trial with a probability equivalent to the occurrence of letters in the english language (Jones & Mewhort, 2004). Five subjects completed eight sessions each and for each of these repeated measurements thresholds in the control condition were subtracted before the average impairment was calculated over all subjects.

Letter identification was differently affected by the type of aberration with secondary astigmatism causing the greatest increase in log contrast threshold (factor of 2.61), closely followed by defocus (factor of 2.45) and coma caused the smallest increase (factor of 1.47). We compared these results to the confusability of letters calculated for these particular amplitudes of aberration at the size presented to our subjects. The relationship between this metric and the increase in log contrast threshold could be described by the same curve for all three types of aberration.

2.3. Reading performance

In a separate experiment (Young et al., 2011) we tested subjects' reading performance with text that had a simulated higher-order aberration (either 0.30, 0.35 or 0.4 μm rms (pupil diameter = 3.5 mm) of either defocus, coma or secondary astigmatism). Smaller amplitudes of aberration were used than in the letter identification experiment since the letter size was necessarily smaller (15 min of arc, equivalent to a Snellen acuity of 20/60). A pinhole could not be used in this experiment since subjects were required to move their eyes. However pupil diameter was recorded throughout each trial and was 3.5 mm on average. Eye movements were recorded monocularly (although subjects read binocularly) and the data were analysed for fixation durations on each word and for the number of fixations made in each sentence. Nineteen subjects viewed four repeats of each of the ten experimental conditions (three amplitudes of each of the three types of aberration and a control condition). The average fixation durations were calculated over each sentence and then averaged over the four repeated measurements. The average fixation duration in the control condition was subtracted before averaging impairment over all subjects. Text samples were single-line sentences with a 6-letter target word placed approximately in the middle. For each sentence two target words were generated, either a high lexical frequency word (one that is common in language) or a low lexical frequency word (one that is uncommon in language). Lexical frequency dependent effects on fixation duration are a hallmark of word identification processes (Inhoff & Rayner, 1986; Rayner, 1998; Rayner & Duffy, 1986).

Reading performance was differently affected by the type of aberration with secondary astigmatism causing the greatest increase in the average fixation duration (factor of 1.28), closely followed by defocus (factor of 1.25) and coma caused the smallest increase (factor of 1.05). We compared these results to the confusability of letters calculated for these particular amplitudes of aberration at the size presented to our subjects. The relationship between this metric and the increase in average fixation duration in the presence of secondary astigmatism and defocus could be described by the same curve, however those data obtained in the presence of coma could not. Performance in the presence of coma was worse at a lower confusability value and we have suggested that there are additional effects associated with coma when multiple letters are presented together, as in a word. There was no interaction between lexical frequency and rms amplitude of aberration of coma so we inferred that coma did not effect word recognition per se (Young et al., 2011) and argued that these effects are more likely to be associated with effects on eye movements and/or effects due to crowding. The measures of contrast threshold for letter identification are additionally consistent with this hypothesis since all three types of aberration can be described by the same relationship when letters are presented in isolation. Additionally we noted that spatial interactions between letters may have been significant since coma smears text in one direction.

3. Performance metrics

3.1. Confusability

Cross-correlation-based models have been shown to have a high correlation with acuity measures (Watson & Ahumada, 2008, 2012) and we have shown that this type of metric predicts performance in letter-based tasks reasonably well (Young et al., 2011; Young, Love, & Smithson, 2013). The metric we have used is described fully elsewhere (Young et al., 2011) but we provide a summary here.

In this metric cross-correlations are performed to compare letter stimuli rendered with a specific type and amplitude of aberration. The maximum values of the correlations formed a 26-by-26 matrix that was then normalised such that the elements along the diagonal were equal to one. The columns of this matrix were weighted by the frequencies with which letters occur in language and the mean of the entire matrix was taken to be the confusability value. Matrices were created for cross-correlations between non-aberrated letters, to represent letter identification based on learned templates, and aberrated letters, to represent letter identification based on comparisons between available letter forms. We note that this method is different to that developed by Watson and Ahumada (2008) as we do not include neural noise in our model or perform any Monte Carlo simulations of performance. The values obtained relate to the maximum value of a cross correlation between letter images and not a probability of correct or incorrect responses.

3.2. Visual Strehl ratio

The visual Strehl ratio is the ratio of the integral of the OTF, weighted by the NCSF, to that same integral for a diffraction limited system:

$$\mathrm{VSOTF} = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{OTF}(f_x,f_y)\cdot\mathrm{NCSF}(f_x,f_y)df_x\,df_y}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{OTF}_{\mathrm{DL}}(f_x,f_y)\cdot\mathrm{NCSF}(f_x,f_y)df_x\,df_y}, \tag{3}$$

where $\mathrm{OTF}(f_x,f_y)$ is the OTF calculated via a Fourier transform of a PSF generated for a particular type and amplitude of aberration as a function of vertical (f_y) and horizontal (f_x) spatial frequency and $\mathrm{OTF}_{\mathrm{DL}}(f_x,f_y)$ is that OTF for a diffraction-limited system with the same pupil radius.

The OTF can be broken up into either its real and imaginary parts or into phase (PTF) and amplitude (MTF). We test the visual Strehl ratio computed using the MTF,

$$\mathrm{VSMTF} = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{MTF}(f_x,f_y)\cdot\mathrm{NCSF}(f_x,f_y)df_x\,df_y}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{MTF}_{\mathrm{DL}}(f_x,f_y)\cdot\mathrm{NCSF}(f_x,f_y)df_x\,df_y}, \tag{4}$$

where $\mathrm{MTF}(f_x,f_y)$ is the MTF calculated from the OTF associated with a particular type and amplitude of aberration and $\mathrm{MTF}_{\mathrm{DL}}(f_x,f_y)$ is the diffraction limited equivalent.

As we have discussed, using the real part of the OTF may not be an appropriate method for predicting performance in the presence of odd aberrations such as coma. We have therefore defined another visual Strehl metric that attempts to combine contrast and phase effects in a different manner. We suggest a metric that computes the ratio of in-phase contrast (linearly weighted by phase) to total contrast. This represents the fraction of contrast that has the correct phase:

$$\mathrm{VS}_{combined} = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{MTF}(f_x,f_y)\cdot\left|1-\frac{\mathrm{PTF}(f_x,f_y)}{\pi}\right|\cdot\mathrm{NCSF}(f_x,f_y)df_x\,df_y}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}MTF(f_x,f_y)\cdot\mathrm{NCSF}(f_x,f_y)df_x\,df_y}, \tag{5}$$

where $\mathrm{PTF}(f_x,f_y)$ is the PTF calculated from the OTF associated with a particular type and amplitude of aberration. The PTF is normalised such that is lies between 0 and 1, with a phase shift of 0° corresponding to 1, shifts of 90° and 270° corresponding to 0.5 and a shift of 180° corresponding to 0. This gives a visual Strehl ratio that decreases with increasing phase shifts. The denominator of this equation is the MTF, not the diffraction-limited equivalent and so aberrations that do not cause phase shifts (low amplitudes of defocus for example) will give a visual Strehl ratio of 1. We note that this is a limitation of our new metric. However a simple calculation showed that, as an example, the amplitude of defocus below which phase shifts do not occur corresponded to a quarter of a Diopter for a 3.5 mm pupil up to half of a Diopter for a 2.5 mm pupil. These correspond to the lower limit at which an optometrist would consider a correction aid since patients are typically not troubled by focus errors smaller than this. It is an interesting observation that noticeable degradations to acuity occur at an amount of defocus at which phase shifts begin to occur. Phase is critically important for object recognition since phase shifts change the contours in the image, whereas contrast changes only suppress frequency components. This new metric does not capture effects solely due to contrast changes, such as those caused by Gaussian blur, but for real optical aberrations at noticeable amplitudes phase changes will occur in the image.

In all of these visual Strehl metrics we use NCSF to weight the function inside the double integral. The NCSF is the CSF divided by the MTF of the eye (Campbell & Green, 1965). Mannos and Sakrison (1974) proposed an analytical model of the CSF,

$$\mathrm{CSF}(f) = 2.6(0.0192 + 0.114f)exp[-(0.114f)^{1.1}], \tag{6}$$

where f is the spatial frequency in cycles per degree. This serves as an approximation based on observers' judgments of images and gives similar results to other published measures of the CSF. Observers performed Mannos and Sakrison's task with a pupil size of approximately 3 mm but their ocular aberrations were not measured. The MTF of the eye differs between individuals and without a measure of these observers' MTFs the optical contribution to the CSF can only be approximated. At small pupil sizes (2–3 mm) the aberrations are likely to have been small and closely approximated by a diffraction-limited system (Charman, 1991). We therefore chose to calculate the NCSF by dividing the CSF, as defined above, by a diffraction-limited MTF for a 3 mm pupil diameter.

3.3. Narrowband-limited visual Strehl ratio metrics

If letter identification is mediated by a single narrow-band visual channel then we can weight the OTF by this channel alone to better represent the effects of an aberration on letter identification. Contrast and phase changes at spatial frequencies that are critical for the task will be counted and those that are not likely to mediate the task will not.

Majaj et al. (2002) showed that the frequency channel mediating letter identification had a centre frequency, $f_{channel}$, determined by the stroke frequency, f_{stroke}. They defined the stroke frequency as the number of lines crossing a horizontal slice at half the x-height, averaged over all letters and then divided by the average letter width. In our letter identification task letters had an average width of 1° and we calculated the stroke frequency of courier font letters at this size to be 1.57 strokes per degree. They also determined the relationship between the stroke frequency and the centre frequency of the channel to be

$$\frac{f_{channel}}{10\ \text{cycles/deg}} = \left(\frac{f_{stroke}}{10\ \text{cycles/deg}}\right)^{2/3} \tag{7}$$

for sharp edged letters. Although the letters used in this experiment are not sharp edged, due to the filtering effects of the aberrations, we have used this relationship to determine the centre frequency of the channel mediating identification of these letters. Majaj et al. (2002) showed that the relationship for filtered letters was

$$f_{channel} \propto f_{center}^{2/3}, \tag{8}$$

where f_{center} is the center frequency of the filter. Aberrations do not produce a Gaussian filter profile from which a center frequency can be determined. Although aberrations produce effects similar to low-pass filtering they may also allow high spatial frequencies to pass, causing spurious resolution for example. We have therefore used Eq. (7) rather than making assumptions about the effects of aberrations on the frequency channel mediating letter identification. Using Eq. (7) we calculated the centre frequency of the channel mediating the identification of 1° courier font letters to be 2.91 cycles per degree. Using the same method we calculated the centre frequency of the channel mediating the identification of 15 min. of arc courier font letters (as used for the reading task) to be 7.35 cycles per degree. These values were used to create a Gaussian filter with a bandwidth of 1 octave, $LB(f_x,f_y)$, which was then used to weight the OTF, MTF or PTF (denoted by transfer function, TF, and $\mathrm{TF}_{\mathrm{DL}}$ in the diffraction limited case) in order to calculate the visual Strehl ratio:

$$\mathrm{VSTF} = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{TF}(f_x,f_y)\cdot LB(s,f_x,f_y)df_x\,df_y}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{TF_{DL}}(f_x,f_y)\cdot \mathrm{LB}(s,f_x,f_y)df_x\,df_y}, \tag{9}$$

where s is the stroke frequency of the letters.

3.4. Orientation-limited visual Strehl metrics

We know that there is a task-specific difference in the effects of these aberrations on letters, which we have attributed to the presentation of multiple stimuli together as opposed to single stimuli in isolation. The implication of presenting multiple stimuli is that contrast from neighbouring stimuli can overlap, which can lead to crowding effects and potentially disrupt eye movements. To take this into account we have tested an orientation weighting applied to the visual Strehl metrics:

$$\mathrm{Mask} = |cos(\theta)|, \tag{10}$$

where θ is the angle in the spatial frequency domain (i.e. components are weighted according to their contribution along the horizontal orientation). This mask is used as a weighting factor, analogous to our use of the spatial frequency band mediating letter identification.

4. Results

4.1. Letter identification

Figs. 1 and 2 show the correlations between the prediction metrics and the increase in contrast threshold for letter recognition for 0.5, 0.6, 0.7, 0.8 and 0.9 μm rms (pupil diameter = 2.5 mm) of defocus, coma or secondary astigmatism. The corresponding Spearman's rank correlation coefficients (magnitudes) are given in Table 1 and show that weighting by the spatial frequency band mediating letter identification, rather than the NCSF, consistently increases the correlation. Of the narrow-band-weighted visual Strehl metrics the VSOTF has the highest correlation with the experimental data (ρ = 0.95), which was the same as the best cross-correlation metric, closely followed by $\mathrm{VS}_{\mathrm{combined}}$ (ρ = 0.94). All visual Strehl based metrics gave correlations 0.9 or above when the narrow-band weighting was applied.

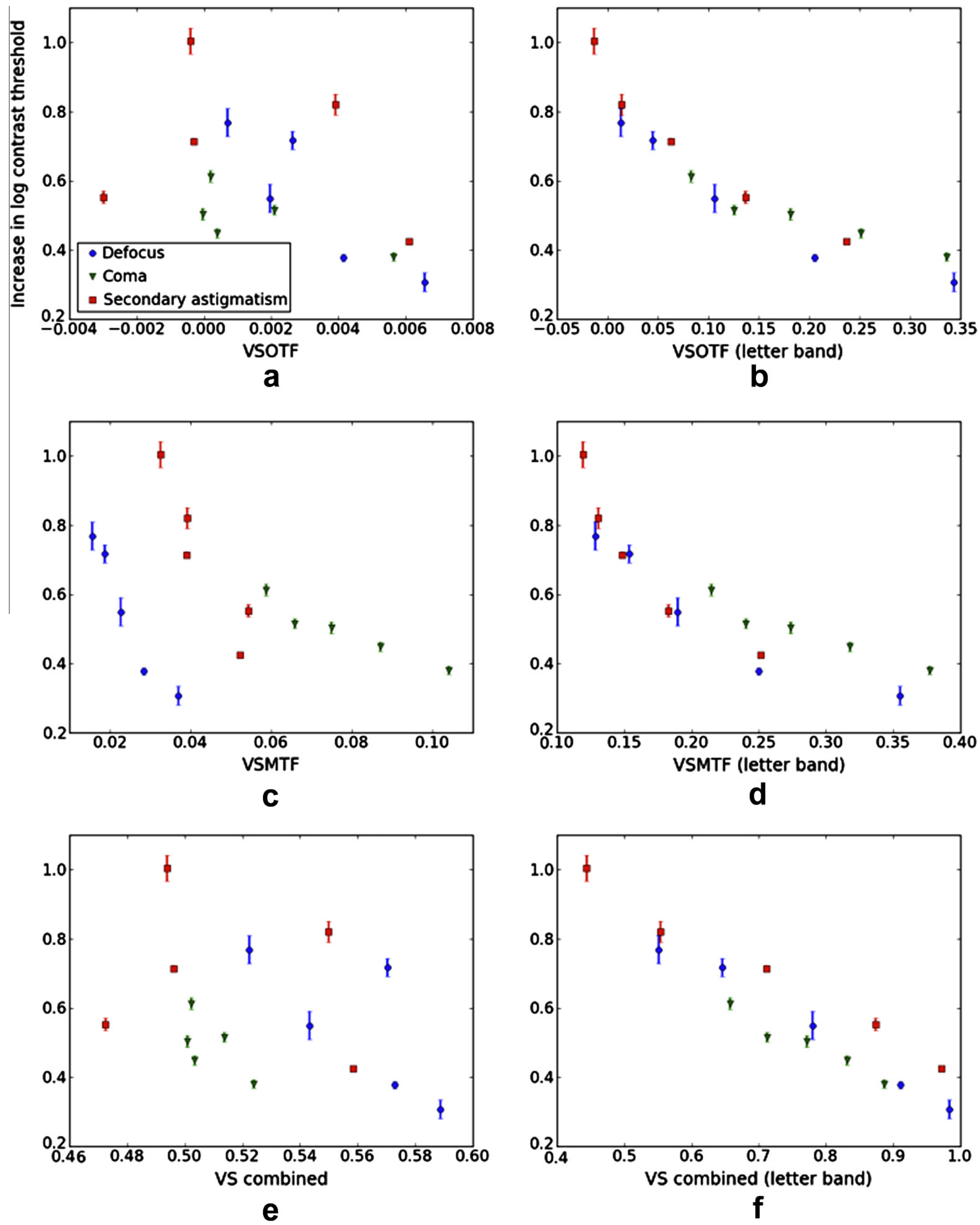

Fig. 1. Comparisons between the measured increase in log contrast threshold for single letter recognition and (a) the VSOTF, (b) the VSOTF calculated using narrow-band weighting, (c) the VSMTF, (d) the VSMTF calculated using the narrow-band weighting, (e) the visual Strehl ratio calculated using phase and contrast components and (f) the visual Strehl ratio as for (e) but using the narrow-band weighting.

4.2. Reading

The correlations between the performance prediction metrics and the increase in fixation duration for 0.3, 0.35 and 0.4 μm rms (pupil diameter = 3.5 mm) of defocus, coma or secondary astigmatism are given in Figs. 3 and 4. Table 2 shows the corresponding Spearman's rank correlation coefficients (magnitudes). These again show that weighting the OTF by the spatial frequency band mediating letter identification, rather than the NCSF, consistently increases the correlation. Of the narrow-band visual Strehl metrics the $VS_{combined}$ had the highest correlation with the experimental data ($\rho = 0.93$) and out-performed the cross-correlation metric. Weighting by the orientation mask additionally improved the VSOTF and $VS_{combined}$ but not the VSMTF. Overall the highest correlation with the experimental data was obtained using the $VS_{combined}$ with the narrow-band weighting and the orientation mask ($\rho = 0.97$).

5. Discussion

In order to understand an ocular wavefront measurement in terms of the implications for visual performance it is important to characterise the effects that aberrations have on real-life tasks (Pepose & Applegate, 2005). Ideally we would like to reduce the wavefront measurement to a single number that represents visual performance taking different visual tasks into account. The VSOTF has been shown to predict acuity well for some aberrations (e.g. Cheng, Bradley, & Thibos, 2004; Thibos et al., 2004) however it can produce negative values. The advantage of visual Strehl metrics is that they take both optical and neural factors into account, without requiring knowledge of the stimulus set used in the visual task. Cross-correlation-based metrics produce excellent predictions of performance in specific tasks (Watson & Ahumada, 2008, 2012; Young et al., 2011; Young, Love, & Smithson, 2013) but they

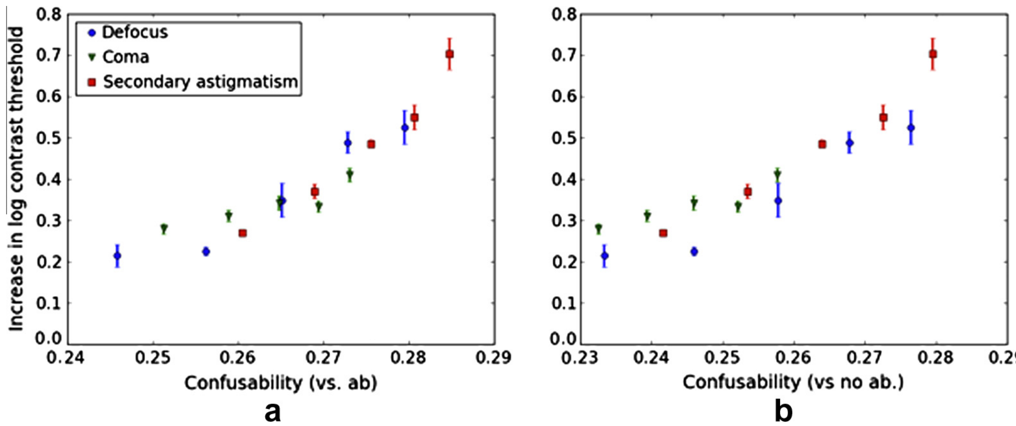

Fig. 2. Comparisons between the measured increase in log contrast threshold and the confusability of letters, calculated from a cross-correlation between letter stimuli. Cross-correlations are performed between aberrated letters and (a) aberrated letters or (b) non-aberrated letters.

Table 1
Magnitude of the Spearman's rank correlation coefficient for each metric tested against the increase in log contrast threshold for letter identification.

Metric	NCSF weighting	Narrow-band weighting
VSOTF	$\rho = 0.59$, $p = 0.022$	$\rho = 0.95$, $p < 0.001$
VSMTF	$\rho = 0.36$, $p = 0.187$	$\rho = 0.90$, $p < 0.001$
$VS_{combined}$	$\rho = 0.46$, $p = 0.084$	$\rho = 0.94$, $p < 0.001$
	vs. aberrated letters	vs. unaberrated letters
Confusability	$\rho = 0.95$, $p < 0.001$	$\rho = 0.92$, $p < 0.001$

require knowledge of exactly which stimuli are presented. We have compared these two types of metric and attempted to produce an improved visual Strehl metric that has the strengths of a cross-correlation-based metric without the need to know which exactly stimuli contributed to the empirical measure of performance. We do this by considering not the spatial representation of specific stimuli but rather the spatial frequency components that mediate the identification of that class of stimulus. In this case the only thing we need to know is the centre frequency and the bandwidth of the frequency selective channel that mediates the task. For letter identification the bandwidth is about 1 octave and the centre frequency is related to the stroke frequency of the letters (Majaj et al., 2002), which is determined by the size of the letters and to a lesser extent the font, since most fonts (at least those without embellishments) have a similar number of strokes per letter.

When using a metric with the narrow-band weighting that is derived from estimates of the channel mediating (unaberrated) letter identification, it is prudent to consider the invariance of that weighting in different aberration conditions. Majaj et al. (2002) showed that when a letter (at a constant size) is filtered the centre frequency of the channel scales less than proportionally with the centre frequency of the object. Additionally, Oruç and Landy (2009) suggested that observers can switch spatial–frequency channels, although they may not switch to the optimal channel. These results lead to a prediction that the addition of aberrations, which act to spatially filter an image, could affect the spatial frequency channel mediating letter identification. However, without explicitly testing this we cannot include shifts of the channel frequency in our analysis. Besides, we wish to produce a metric that can be calculated without knowledge of complex interactions between the type of aberration and the type of stimulus. Effects due to channel shifting would reduce the correlation between an empirical measure of performance and a metric that does not take this into account. This should be kept in mind when conclusions are drawn from our analyses.

The most important result in this paper is that, subject to the limitations described in the previous paragraph, weighting the terms in the integrals of the visual Strehl ratio by the spatial frequency band that is thought to mediate the task significantly improved performance metrics based on the visual Strehl ratio. In the case of letter identification the correlations between the visual Strehl metrics and the increase in log contrast threshold improved from less than 0.6 to values of at least 0.9. These values are highly competitive against the cross-correlation-based metric which gave correlation values of up to 0.95 (comparing between aberrated letters) and 0.92 (comparing aberrated with unaberrated letters). The correlations between reading performance and visual Strehl metrics were similarly improved by weighting the terms in the integrals of the visual Strehl ratio by the spatial frequency band that mediated the task, although correlation values were not as high as those for letter identification. The maximum correlation was obtained for the combined visual Strehl ratio ($\rho = 0.93$). Additionally taking orientation effects into account in the combined visual Strehl ratio ($\rho = 0.97$) gave the best correlation overall. The cross-correlation-based metric in this case only gave a correlation coefficient of 0.85 (comparing between aberrated letters) and 0.70 (comparing aberrated with unaberrated letters).The performance measures we have used are different to those that have been used previously to test visual Strehl metrics. The correlations we have obtained for the unmodified visual Strehl ratios are significantly lower than those obtained using traditional measures of visual acuity, which are typically around 0.8.

One of the aims of this paper was to consider the relative contributions of the phase and contrast effects of aberrations. Rather than incorporating phase by transforming back to the PSF or by only using the real part of the OTF, we have attempted to include phase effects by considering how they impact visual performance separately to contrast effects. Ravikumar, Bradley, and Thibos (2010) found that for letters (and faces) 180° phase reversals with sufficient contrast reduce visual acuity. Phase shifts caused by coma are smaller than 180° and have a smaller impact. This suggests that an equal weighting of phase values (i.e. ignoring contrast) may indeed not be appropriate and additional analyses not reported in detail here confirmed that a visual Strehl metric based only on phase changes did not perform as well even with the narrow-band weighting ($\rho = 0.88$ for letter identification and $\rho = 0.72$ for reading performance). We cannot confirm that the linear relationship employed by the combined visual Strehl ratio is the most appropriate model for phase effects on visual performance, however for our limited data set it produced better correlation coefficients than other relationships we tested.

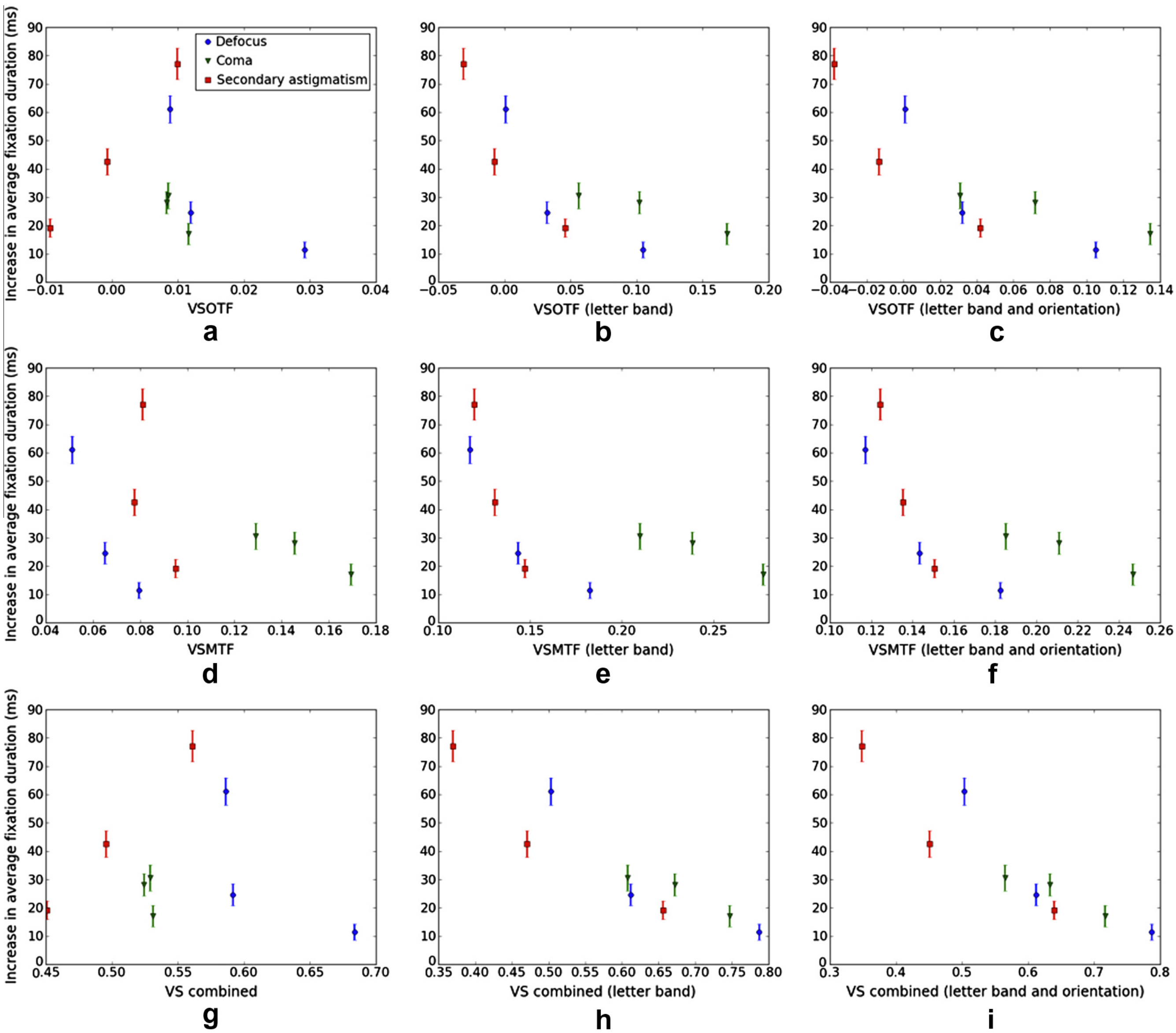

Fig. 3. Comparisons between the measured increase in average fixation duration and (a) the VSOTF, (b) the VSOTF calculated using narrow-band weighting, (c) the VSOTF using the narrow-band weighting and the orientation mask, (d) the VSMTF, (e) the VSMTF calculated using the narrow-band weighting, (f) the VSMTF calculated using the narrow-band weighting and the orientation mask (g) the visual Strehl ratio calculated using phase and contrast components, (h) the visual Strehl ratio as for (g) but using the narrow-band weighting and (i) the visual Strehl ratio as for (g) but using the narrow-band weighting and the orientation mask.

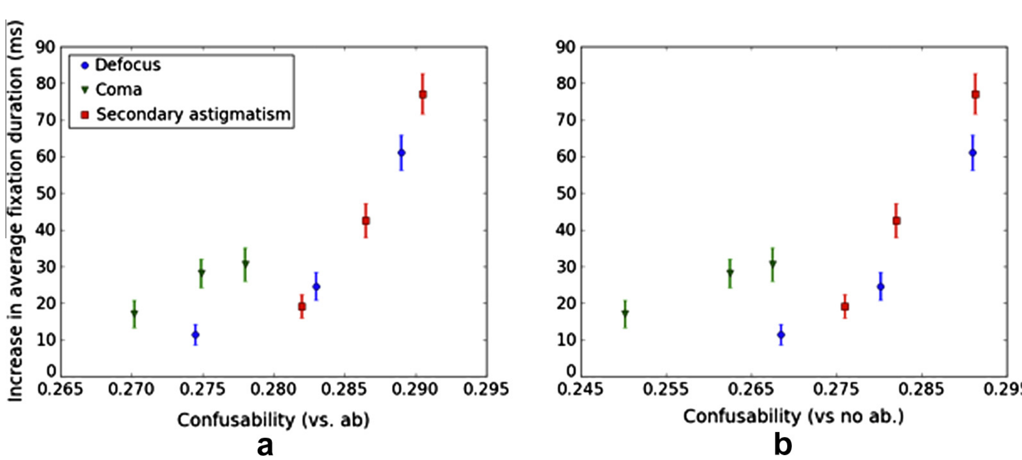

Fig. 4. As for Fig. 2 but compared with the measured increase in average fixation duration.

Table 2
Magnitude of the Spearman's rank correlation coefficient for each metric tested against the increase in average fixation duration.

Metric	NCSF weighting	Narrow-band weighting	Narrow-band weighting and orientation mask
VSOTF	$\rho = 0.33$, $p = 0.381$	$\rho = 0.83$, $p = 0.005$	$\rho = 0.92$, $p = 0.001$
VSMTF	$\rho = 0.35$, $p = 0.356$	$\rho = 0.68$, $p = 0.042$	$\rho = 0.68$, $p = 0.042$
$VS_{combined}$	$\rho = 0.13$, $p = 0.732$	$\rho = 0.93$, $p < 0.001$	$\rho = 0.97$, $p < 0.001$
	vs. aberrated letters	vs. unaberrated letters	
Confusability	$\rho = 0.85$, $p = 0.004$	$\rho = 0.70$, $p = 0.036$	

One potential limitation of our new metric is that the phase weighting and normalisation mean that pure contrast changes (where the PTF is zero for all frequencies) gives a combined visual Strehl ratio of one. To consider the possible benefit of incorporating overall reductions in contrast we performed a multiple regression analysis using $VS_{combined}$ and the VSMTF and this gave correlation values that were marginally improved, except where the orientation weighting had been applied. For letter identification the correlations improved from $\rho = 0.90$ ($p < 0.001$; VSMTF) and $\rho = 0.94$ ($p < 0.001$; $VS_{combined}$) to $\rho = 0.95$ ($p < 0.001$). For reading performance where the visual Strehl is calculated using the narrow-band weighting the correlations improved from $\rho = 0.68$ ($p = 0.042$; VSMTF) and $\rho = 0.93$ ($p < 0.001$; $VS_{combined}$) to $\rho = 0.97$ ($p < 0.001$). For reading performance where the visual Strehl is calculated using the narrow-band and orientation weightings the correlations changed from $\rho = 0.68$ ($p = 0.042$; VSMTF) and $\rho = 0.97$ ($p < 0.001$; $VS_{combined}$) to $\rho = 0.93$ ($p < 0.001$). However, in all cases the contribution of the VSMTF to the regression was not significant ($p = 0.077$ for letter identification, $p = 0.614$ for reading performance and $p = 0.817$ for reading performance with the orientation weighting).

These results seem to imply that while all visual Strehl metrics perform well with the narrow-band weighting, the $VS_{combined}$ is the best predictor of visual performance in letter-based tasks as it has consistently high correlation with measures of reading impairment and with measures of letter-identification impairment. While this metric does not work for contrast-only changes to an image, it is reassuring that for real optical aberrations contrast-only changes occur only at small amplitudes, below those which an optometrist would seek to correct. Importantly this metric takes phase changes and the contrasts at which they occur into account in a logical manner (recall as a counter-example that using the real part of the OTF weights contrast by zero when the phase shift is 90°). Phase is critically important for object recognition as shifts in phase change the contours in the image.

To further probe why this metric performs so well (without knowledge of the source of impairment for a particular task, e.g. effects due to crowding or disruption to eye movement planning) we suggest specifically testing different orientations of aberrations. Visual Strehl metrics are derived from a symmetric function (the OTF is multiplied by the NCSF or the band mediating letter identification, both of which are rotationally symmetric) and so are insensitive to orientation specific effects. We have taken orientation in account by restricting contrast and phase changes to those along the horizontal direction. However, it would be interesting to explicitly test whether this orientation mask produces the same results for text with vertical coma. In this case we would expect spatial interactions between letters to be reduced as contrast is smeared vertically. If orientation is ignored in the visual Strehl ratio metric then the same results would be predicted for either orientation of coma. The cross-correlation metric produces different results for the two orientations of coma. Additionally, if a greater loss of reading performance for a lower confusability value for coma is indeed due to letters being smeared horizontally with the text then repeating the experiment with vertical coma should not produce this effect. We also note that, while not reported here, calculating the visual Strehl ratio by restricting contrast and phase changes to those along the vertical direction increased the difference between the correlation for coma and that for defocus and secondary astigmatism.

As we have previously noted the cross-correlation metric appears to capture functional differences between the three types of aberration (Young et al., 2011; Young, Love, & Smithson, 2013), particularly that coma has a higher than predicted impact on performance in reading. These differences appear in the VSMTF predictions as well (see Fig. 3). If functional effects associated with this paradigm are revealed by cross-correlation metrics (and the VSMTF) then this type of analysis could be important for predicting visual performance with other classes of object, for example faces, which have component features.

We note here that the data we discuss were obtained with larger amplitudes of aberration than are typically measured in the normal population. This is partly due to the large size of the letters used. As discussed in Young, Love, and Smithson (2013) we have scaled the amplitudes of aberration so that the stimuli at the larger letter size used in the letter-recognition task (1°) were similar in appearance to the smaller letter size used in the reading task (0.25°). The near equivalence of the two arrangements was confirmed using a normalised cross-correlation between the stimuli used in each experiment. For letters near to the acuity limit the equivalent amplitude of aberration would be even smaller (around 0.1 μm). This value approaches those typically found in the normal population, and is closer to those found in abnormal eyes. We also note that normal and abnormal eyes typically exhibit many Zernike modes in combination but here we have induced only one type at a time. We selected the modes we tested in part because they exhibit qualitative differences in their effects on the stimulus, particularly as a result of differences in the PTF. The metrics predict performance in the same way for different modes, despite large differences in the effects of different types of aberration, both on the stimulus and on task performance. While it is clear that the multiple Zernike modes that are present in real eyes interact in visual performance space, it is our hope that the characteristics captured by the metrics namely spatial frequency dependent effects in the letter-band and phase and orientation changes might also be those characteristics that impact performance when multiple aberrations are combined. It is important to recognise however that the improvements we show to the visual Strehl metrics are not tested with real-life aberration structures and so the utility of these modified metrics for clinical assessment is still undetermined.

6. Conclusion

We have tested the correlations between four visual Strehl metrics and empirical measurements of letter identification and reading. These metrics were tested with and without a narrow-band weighting that accounts for the spatial frequency requirements of the task. Of these four metrics the newly defined combined VSOTF ($VS_{combined}$) performed consistently well across both visual tasks with this weighting and performed exceptionally well for quantifying reading performance when an additional restriction was placed on the orientation of phase and contrast changes. These findings suggest that by considering the spatially frequency and orientation requirements of the task visual Strehl metrics can perform as well as, and in some cases better than, our cross-correlation-based metric. These improvements to the visual Strehl metric should be tested with other types and amplitudes of aberrations as well as in different types of visual tasks to fully assess their robustness in predicting object recognition performance.

Acknowledgments

We would like to thank Prof. Tom McLeish for an insightful discussion on the optical transfer function and to thank Prof. Simon Liversedge for his advice with the experiment on reading performance. We also wish to thank the reviewers for their suggestions for improving the manuscript. We acknowledge support from the John Fell Fund.

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.
ANSI Z80.28 (2010). *Methods of reporting optical aberrations of eyes*.
Applegate, R. A., & Howland, H. C. (1997). Refractive surgery, optical aberrations and visual performance. *Journal of Refractive Surgery, 13*(3), 295–299.
Arditi, A., Knoblauch, K., & Grunwald, I. (1990). Reading with fixed and variable character pitch. *Journal of the Optical Society of America A, 7*, 2011–2015.
Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*(241), 177–178.
Buehren, T., & Collins, M. J. (2006). Accommodation stimulus–response function and retinal image quality. *Vision Research, 46*(10), 1633–1645.
Bürhen, J., Pesudovs, K., Martin, T., Strenger, A., Yoon, G., & Kohnen, T. (2009). Comparison of optical quality metrics to predict subjective quality of vision after laser in situ keratomileusis. *Journal of Cataract and Refractive Surgery, 35*(5), 846–855.
Campbell, F. W., & Green, D. G. (1965). Optical and retinal factors affecting visual resolution. *Journal of Physiology, 181*, 576–593.
Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, 197*, 551–566.
Charman, W. N. (1991). Wavefront aberration of the eye: A review. *Optometry and Vision Science, 68*(8), 574–583.
Chen, M., Sabesan, R., Ahmad, K., & Yoon, G. (2007). Correcting anterior corneal aberration and variability of lens movements in keratoconic eyes with back-surface customized soft contact lenses. *Optics Letters, 32*(21), 3202–3205.
Cheng, X., Bradley, A., & Thibos, L. N. (2004). Predicting subjective judgment of best focus with objective image quality metrics. *Journal of Vision, 4*, 310–321.
Cheng, X., Thibos, L. N., & Bradley, A. (2003). Estimating visual quality from wavefront aberration measurements. *Journal of Refractive Surgery, 19*(5), S579–S584.
Chung, S. T. L. (2004). Reading speed benefits from increased vertical word spacing in normal peripheral vision. *Optometry and Vision Science, 81*, 525–535.
Chung, S. T. L., Legge, G. E., & Tjan, B. S. (2002). Spatial–frequency characteristics of letter identification in central and peripheral vision. *Vision Research, 42*, 2137–2152.
Chung, S. T. L., Levi, D. M., & Legge, G. E. (2002). Spatial–frequency properties of letter identification in amblyopia. *Vision Research, 42*, 1571–1581.
Chung, S. T. L., & Tjan, B. S. (2009). Spatial–frequency and contrast properties of reading in central and peripheral vision. *Journal of Vision, 9*(9), 1–19 (article no. 16).
Dalimier, E., & Dainty, C. (2008). Use of a customized vision model to analyze the effects of higher-order ocular aberrations and neural filtering on contrast threshold performance. *Journal of the Optical Society of America A, 25*(8), 2078–2087.
Dalimier, E., Pailos, E., Rivera, R., & Navarro, R. (2009). Experimental validation of a Bayesian model of visual acuity. *Journal of Vision, 9*(7), 1–16 (article no. 12).
Gao, H., Dong, X., Yu, X., Deng, Q., & Pan, L. (2009). Correcting higher-order aberrations of the human eye with the use of a phase plate fabricated by a mask-moving technique. *Optical Engineering, 48*(2), 023202.
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.
Howland, H. C., & Howland, B. (1977). A subjective method for the measurement of monochromatic aberrations of the eye. *Journal of the Optical Society of America, 67*(11), 1508–1518.
Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics, 40*, 431–439.
ISO 24157 (2008). *Ophthalmic optics and instruments – Reporting aberrations of the human eye*.
Jeong, T. M., & Yoon, G. (2006). Customized correction of wavefront aberrations in abnormal human eyes by using a phase plate and a customized contact lens. *Journal of the Korean Physical Society, 49*(1), 121–125.
Jones, M. N., & Mewhort, D. J. K. (2004). Case-sensitive letter and bigram frequency counts from large-scale English corpora. *Behaviour Research Methods, Instruments and Computers, 36*, 388–396.
Legras, R., & Rouger, H. (2008). Just-noticeable levels of aberration correction. *Journal of Optometry, 1*, 71–77.
Levi, D. M. (2008). Crowding – An essential bottleneck for object recognition: A mini-review. *Vision Research, 48*, 635–654.
López-Gil, N., Castejón-Monchón, J. F., Benito, A., Marin, J. M., Lo-a-Foe, G., & Marin, G. (2002). Aberration generation by contact lenses with aspheric and asymmetric surfaces. *Journal of Refractive Surgery, 18*(5), S603–S609.
López-Gil, N., Château, N., Castejón-Monchón, J. F., Artal, P., & Benito, A. (2003). Correcting ocular aberrations by soft contact lenses. *The South African Optometrist, 62*, 173–177.
Majaj, N. J., Liang, Y. X., Martelli, M., Berger, T. D., & Pelli, D. G. (2003). Channel for reading. *Journal of Vision, 3*(9), 812.
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.
Mannos, J. L., & Sakrison, D. J. (1974). The effects of a visual fidelity criterion on the encoding of images. *IEEE Transactions in Information Theory, 20*(4), 525–535.
Marsack, J. D., Milner, T., Rylander, G., Leach, N., & Roorda, A. (2002). Applying wavefront sensors and corneal topography to keratoconus. *Biomedical Sciences Instrumentation, 38*, 471–476.
Marsack, J. D., Parker, K. E., & Applegate, R. A. (2008). Performance of wavefront-guided soft lenses in three keratoconus subjects. *Optometry and Vision Science, 85*(12), E1172–E1178.
Marsack, J. D., Parker, K. E., Nui, Y., Pesudovs, K., & Applegate, R. A. (2007). On-eye performance of custom wavefront-guided soft contact lenses in a habitual soft lens-wearing keratoconic patient. *Journal of Refractive Surgery, 23*, 960–964.
Marsack, J. D., Thibos, L. N., & Applegate, R. A. (2004). Metrics of optical quality derived from wave aberrations predict visual performance. *Journal of Vision, 4*, 322–328.
Navarro, R., Esther, M.-B., Bará, S., & Mancebo, T. (2000). Phase plates for wave-aberration compensation in the human eye. *Optics Letters, 25*(4), 236–238.
Nestares, O., Navarro, R., & Antona, B. (2003). Bayesian model of Snellen visual acuity. *Journal of the Optical Society of America A, 20*(7), 1371–1381.
Netto, M. V., Dupps, W., & Wilson, S. E. (2006). Wavefront-guided ablation: Evidence for efficacy compared to traditional ablation. *American Journal of Ophthalmology, 141*, 360–368.
Oruç, I., & Landy, M. S. (2009). Scale dependence and channel switching in letter identification. *Journal of Vision, 9*(9), 1–19 (article no. 4).
Parish, D. H., & Sperling, G. (1991). Object spatial frequencies, retinal spatial frequencies, noise and the efficiency of letter discrimination. *Vision Research, 31*, 1399–1415.
Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision, 4*, 1136–1169.
Pepose, J. S., & Applegate, R. A. (2005). Making sense out of wavefront sensing. *American Journal of Ophthalmology, 139*(2), 335–343.
Prins, N., & Kingdom, F. A. A. (2009). *Palamedes: Matlab routines for analyzing psychophysical data*. <http://www.palamedestoolbox.org>.
Ravikumar, A., Sarver, E. J., & Applegate, R. A. (2012). Change in visual acuity is highly correlated with change in six image quality metrics independent of wavefront error and/or pupil diameter. *Journal of Vision, 12*(10), 1–13 (article no. 11).
Ravikumar, S., Bradley, A., & Thibos, L. N. (2010). Phase changes induced by optical aberrations degrade letter and face acuity. *Journal of Vision, 18*, 1–12.
Rayner, K. (1998). Eye movements in reading: 20 years of research. *Psychological Bulletin, 124*, 372–422.
Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity and lexical ambiguity. *Memory & Cognition, 14*, 191–201.
Sabesan, R., Jeong, T. M., Carvalho, L., Cox, I. G., Williams, D. R., & Yoon, G. (2007). Vision improvement by correcting higher-order aberrations with customized soft contact lenses in keratoconic eyes. *Optics Letters, 32*(8), 1000–1002.
Shi, Y., Wei, X., Ravikumar, A., & Applegate, R. A. (2011). Changes in log visual Strehl predicts changes in visual acuity. *Investigative Ophthalmology and Visual ScienceARVO abstract*, 2814.
Solomon, J. A., & Pelli, D. G. (1994). The visual filter mediating letter identification. *Nature, 369*, 395–397.
Stuart, J. A., & Burian, H. M. (1962). A study of separation difficulty: Its relationship to visual acuity in normal and amblyopic eyes. *American Journal of Ophthalmology, 53*, 471–477.
Tarrant, J., Roorda, A., & Wildsoet, C. F. (2010). Determining the accommodative response from wavefront aberrations. *Journal of Vision, 10*(5), 1–16 (article no. 4).
Thibos, L. N., Applegate, R. A., Schwiegerling, J. T., & Webb, R. (2000). Standards for reporting the optical aberrations of eyes. In *Vision science and its applications* (p. SuC1). Optical Society of America.
Thibos, L. N., Hong, X., Bradley, A., & Applegate, R. A. (2004). Accuracy and precision of objective refraction from wavefront aberrations. *Journal of Vision, 4*, 329–351.

Watson, A. B., & Ahumada, A. J. (2008). Predicting visual acuity from wavefront aberrations. *Journal of Vision, 8*(4), 1–19 (article no. 17).

Watson, A. B., & Ahumada, A. J. (2012). Modelling acuity for optotypes varying in complexity. *Journal of Vision, 12*(10), 1–19 (article no. 19).

Yoon, G., Jeong, T. M., Cox, I. G., & Williams, D. R. (2004). Vision improvement by correcting higher-order aberrations with phase plates in normal eyes. *Journal of Refractive Surgery, 20*, S523–S527.

Young, L. K., Liversedge, S. P., Myers, R. M., Love, G. D., & Smithson, H. E. (2011). Not all aberrations are equal: Reading impairment depends on aberration type and magnitude. *Journal of Vision, 11*(13), 1–19 (article no. 20).

Young, L. K., Love, G. D., & Smithson, H. E. (2013). Different aberrations raise contrast thresholds for single-letter identification in line with their effect on cross-correlation-based confusability. *Journal of Vision, 13*(17), 1–14 (article no. 12). http://www.journalofvision.org/content/13/7/12.