# Visibility and acceptance of discrete-sampling artifacts in visual displays

Christos Kaspiris-Rousellis
Newcastle University

This dissertation is submitted for the degree of
*Doctor of Philosophy*

Biosciences Institute Sep 2020

# Abstract

Digital visual displays are aimed to provide an illusion of a continuous reality through a discrete presentation of visual information. This thesis explored three topics on (i) angular, (ii) spatial, and (iii) temporal sampling characteristics, related to distortion visibility, acceptance, and discomfort. In the first topic, we addressed the issue of optimizing the view density in continuous parallax visualization by replicating the changing views of a 3-D object for a moving observer. We measured the visibility of the related artifacts and evaluated the performance of full-reference visual quality metrics. We found that the state-of-the-art metrics can indirectly characterize artifact visibility and established a quantitative relationship for threshold estimation on varying conditions. The second topic addressed the relation of the contrast sensitivity function (CSF) to adaptation luminance and specifically its asymptotic behavior at high light levels essential to modern high-luminance displays. Using a custom-built system, we measured the CSF at relatively high luminance levels and spatial frequency range, integrating our dataset to the existing research. We found a gradual transition among the linear to DeVries-Rose to Weber regions with steeper slopes for higher frequencies and lower luminance. A further decreasing region was located at low to intermediate frequencies. Following this construct, we adopted a model consisting of central elements in the visual signal processing and proposed an eight-parameter form for the CSF in the luminance domain. The final topic addressed the effects of frame rate on distortion acceptance and its impact on visual discomfort during regular display use. We assessed the perceived symptoms, preference, and task performance under varying conditions. The measurements indicated that for non-demanding everyday tasks, the frame rate could be reasonably reduced without severe effects on the observer; however, this tolerance diminished under more dynamic content. A potential association of discomfort with the blinking activity was also discussed.

## Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 80,000 words and has less than 150 figures.

Christos Kaspiris-Rousellis
Sep 2020

# Acknowledgements

The work in this thesis was partly supported by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 676401, European Training Network on Full Parallax Imaging and Huawei Technologies Co., Ltd. I would like to thank Jenny Read for securing the funding and supervising the projects. The thesis was based on the following manuscripts with the corresponding collaborative contributions:

- Kaspiris-Rousellis C., Simmons A., Read, J. C. A. (2018). Perceptually optimized view density for continuous parallax. Journal of Vision 2018;18(10):512. Abstract

  Kaspiris-Rousellis C., Simmons A., Read, J. C. A. (2019). Visibility and acceptance of angular distortions in full-parallax content. Unpublished manuscript

  - CKR and JR designed the research
  - CKR and AS collected the data

- Kaspiris-Rousellis, C., Fernandez-Alonso, M., and Read, J. C. A. (2019). Extending the human foveal spatial contrast sensitivity function to high luminance range. Proceedings of the European Light Field Imaging Workshop

  Kaspiris-Rousellis, C., Fernandez-Alonso, M., and Read, J. C. A. (2019). The human spatial contrast sensitivity at high luminance. Unpublished manuscript

  - CKR, MFA, and JR designed the research

- Kaspiris-Rousellis, C., Fernandez-Alonso, M., Tong C.C., and Read, J. C. A. (2019). The effect of high frame rates on visual comfort, performance, and blinking activity. Unpublished manuscript

  - CKR, MFA, and JR designed the research
  - CKR and CCT collected the data

# Table of Contents

**List of Figures** vii

**List of Tables** ix

**General introduction** 1

**1 Visibility of angular distortions in full-parallax content** 14
- 1.1 Introduction . . . 14
- 1.2 Methods . . . 18
  - 1.2.1 Participants . . . 18
  - 1.2.2 Apparatus . . . 18
  - 1.2.3 Image sequence simulation . . . 18
  - 1.2.4 Experimental task and conditions . . . 19
  - 1.2.5 The psychometric function and the degradation scale . . . 21
- 1.3 Results . . . 22
  - 1.3.1 Psychophysical data . . . 22
  - 1.3.2 Distortion ranking similarity and metrics performance . . . 24
  - 1.3.3 Model estimates and validation . . . 25
- 1.4 Discussion . . . 27

**2 The human contrast sensitivity function at high luminance** 33
- 2.1 Introduction . . . 33
- 2.2 Measurements of the human CSF at high luminance . . . 35
  - 2.2.1 Participants . . . 35
  - 2.2.2 Apparatus and experimental design . . . 35
  - 2.2.3 CSF characteristics . . . 38
  - 2.2.4 Statistical analysis . . . 39
  - 2.2.5 Results . . . 40
- 2.3 CSF modeling in the luminance domain . . . 50

2.3.1 Pre-processing . . . 52
2.3.2 Segmented regression . . . 52
2.3.3 Extending the CSF to high luminance range . . . 53
2.3.4 Results . . . 55
2.3.5 Barten's CSF formula and model assessment . . . 60
2.4 Discussion . . . 62

**3 The effect of frame rate on visual discomfort and fatigue 68**
3.1 Introduction . . . 68
3.2 Methods . . . 70
3.2.1 Participants . . . 70
3.2.2 Apparatus and experimental design . . . 70
3.2.3 Subjective measures . . . 73
3.2.4 Objective measures . . . 73
3.2.5 Statistical analysis and modeling . . . 75
3.3 Results . . . 76
3.3.1 Task performance . . . 76
3.3.2 Eye blinks . . . 77
3.3.3 Subjective feedback . . . 81
3.4 Discussion . . . 85

**General discussion 91**

**Appendix A Supplementary material for Chapter 1 93**

**Appendix B Supplementary material for Chapter 2 97**

**Appendix C Supplementary material for Chapter 3 101**

**References 109**

# List of Figures


0.1 The human eye and the retina . . . 2
0.2 Receptive fields of ON and OFF cells . . . 3
0.3 Sinusoidal gratings . . . 4
0.4 Spatiotemporal contrast sensitivity function . . . 5
0.5 Sampled motion and the window of visibility . . . 6
0.6 Triangulation-based depth cues . . . 7
0.7 Block diagram of a generic error visibility metric . . . 10

1.1 Two-plane lightfield parameterization in an image acquisition setup . . . 15
1.2 Epipolar plane images and angular sampling in the Fourier domain . . . 16
1.3 Angular view simulation schematic . . . 17
1.4 Image sequence simulation process . . . 19
1.5 The stimulus with an example of the respective distortions . . . 20
1.6 Psychophysical data . . . 23
1.7 Disparity threshold estimates . . . 25
1.8 Metric ranking correlation matrix . . . 26
1.9 Metric performance compared with the psychophysics data . . . 27
1.10 Individual psychometric function fits using the GMSD score . . . 28
1.11 Final model estimates . . . 29
1.12 An example of model validation on different condition subsets . . . 30

2.1 The linear to DeVries-Rose to Weber transition . . . 34
2.2 Projector-based display schematic . . . 36
2.3 CSF stimuli subset . . . 37
2.4 Typical human CSF characteristics . . . 39
2.5 Individual CSFs for each luminance condition . . . 41
2.6 Boxplot of the contrast sensitivity measurements . . . 43
2.7 Boxplot of the main CSF characteristics . . . 46
2.8 Contrast sensitivity estimates at the extended illuminance range . . . 48

2.9 Gain of increase in sensitivity at the extended illuminance range . . . . . 49
2.10 The gain surface in the spatial frequency-luminance domain . . . . . 50
2.11 CSF model diagram . . . . . 51
2.12 CSF model fits . . . . . 56
2.13 Segmented regression slopes . . . . . 57
2.14 Fitted critical illuminance constants . . . . . 60
2.15 Main CSF characteristics as a function of retinal illuminance . . . . . 61

3.1 First-person view of the experimental task . . . . . 72
3.2 Normalized task performance measures . . . . . 77
3.3 Normalized blinking activity measures . . . . . 78
3.4 Estimated effect on the blinking activity . . . . . 81
3.5 Aggregated evaluation observer feedback for experiment I . . . . . 82
3.6 Aggregated evaluation observer feedback in experiment II . . . . . 83
3.7 Preference feedback scaling . . . . . 85
3.8 Hold-type blur example . . . . . 86

A.1 Individual psychometric function fits based on disparity . . . . . 93
A.2 Metrics performance across conditions using the Weibull function . . . . . 94
A.3 Metrics performance across conditions using the cumulative Gaussian function 94
A.4 Metrics performance across conditions using the Logistic function . . . . . 95
A.5 Example of correlation between two metrics . . . . . 95
A.6 Example of correlation between metrics and psychophysics . . . . . 96

B.1 Statistical sensitivity . . . . . 99

C.1 Visual fatigue and discomfort questionnaire in experiment I . . . . . 101
C.2 The frame rate evaluation questionnaire in experiment I . . . . . 102
C.3 The frame rate evaluation questionnaire in experiment II . . . . . 103
C.4 Separate linear regression coefficients of the individual blink rates . . . . . 104
C.5 Statistical sensitivity . . . . . 105
C.6 Visual discomfort and fatigue symptoms questionnaire for experiment I . . . 105
C.7 Individual data on the frame rate evaluation in Experiment I . . . . . 106
C.8 Individual data on frame rate evaluation in Experiment II . . . . . 106
C.9 Alternative preference feedback scaling . . . . . 108

# List of Tables


1.1 Summary of the disparity threshold estimates . . . . . 24

2.1 The mean sensitivity thresholds . . . . . 42
2.2 Statistics on the effect of luminance at each spatial frequency level . . . . . 43
2.3 Pairwise comparisons of sensitivity estimates . . . . . 44
2.4 Mean estimates of the CSF characteristics . . . . . 45
2.5 Statistics of the effect of luminance on the main CSF characteristics . . . . . 46
2.6 Pairwise comparisons of the CSF characteristics . . . . . 47
2.7 Estimated global model parameters . . . . . 58
2.8 Fitted model constants . . . . . 59
2.9 CSF model comparisons . . . . . 63
2.10 Re-estimated parameters of the proposed model among studies . . . . . 64
2.11 Estimated parameters of the proposed model within studies . . . . . 65

3.1 Estimated means of the objective measures . . . . . 79
3.2 Statistics of the frame rate effect on the objective measures . . . . . 80
3.3 Follow-up tests tests for linear and quadratic contrasts . . . . . 80
3.4 Correlation matrix of the evaluation feedback questionnaire . . . . . 84

B.1 Summary table of the meta-analysis data . . . . . 97
B.2 Fitting performance of the descriptive CSF model . . . . . 98
B.3 Correlation matrix of the main CSF characteristics . . . . . 98

C.1 Frame-based performance of the eye state classifier . . . . . 104
C.2 Predicted probabilities of the overall preference scaling model . . . . . 107

# General introduction

In the field of visualization and modern electronic displays, vision science is increasingly gaining attention. Evaluating human responses and limitations is valuable both from the perspective of efficient resource allocation as well as for delivering a higher quality of experience with reduced symptoms of visual fatigue and discomfort. Digital visual displays are discrete systems aimed to provide an illusion of continuous reality, or equivalently to truthfully represent the plenoptic structure that characterizes the rays of light through every point in space (Adelson and Bergen, 1991). However, sampling and reproducing the light rays at every desired location in three-dimensional space and for every angle, time point, and wavelength to an adequate degree is an inherently complex problem. It follows that specifying how dense is dense enough depends on the properties and limitations of the average human observer, e.g., spatiotemporal resolution, binocular viewing, or trichromatic vision, which becomes challenging to model as the input increases in complexity or one generalizes to non-static or multiple observers.

The human eye is the device that allows us to capture the incident light rays in order to construct a visual representation of our surroundings, their shapes, colors, and dimensions under varying illumination conditions. Therefore, specifying the optimal sampling characteristics from a perceptual perspective first requires understanding how the early visual processes encode the intensity distribution of these light rays in the brain. When fixating on an object, the reflected light entering the eye is refracted by the cornea and the crystalline lens and focused on the retina, while an aperture (the pupil) regulates the light entering the eye in response to the light's intensity (left column in Figure 0.1; more details on the optical structure and properties can be found in Atchison and Smith, 2000). The light is then converted into an electrical signal which travels through the optic nerve along the visual pathway to higher cortical areas. The phototransduction process starts in the outer layer of the retina, which contains a class of specialized photoreceptor cells: rods and cones (left column in Figure 0.1). These photoreceptors transmit the information to ganglion cells that are directly signaling to the optic nerve. Their function and organization have been well studied over the past decades (a detailed introduction can be found in Kandel et al., 2012). Rods have high sensitivity to light and thus mediate vision under dim light levels (scotopic conditions; outside the cone

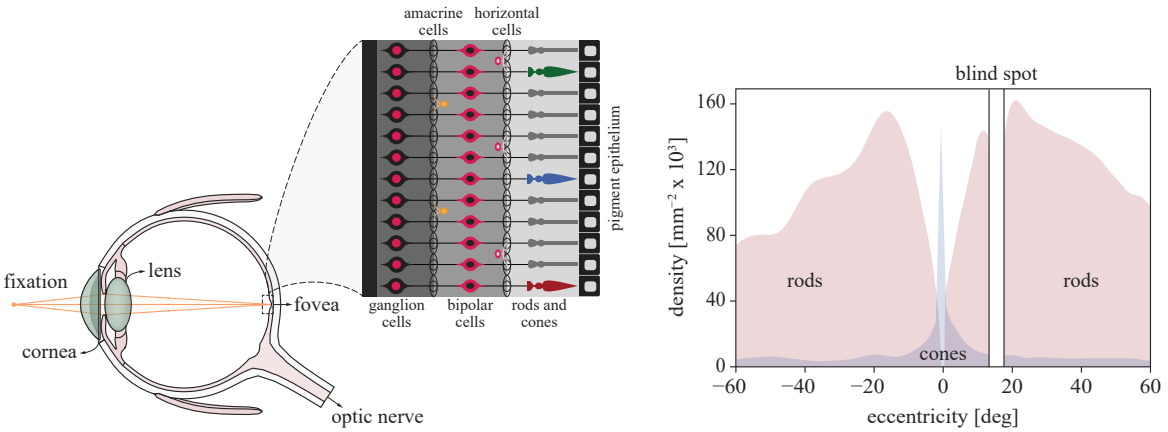


**Figure 0.1 Left**: an illustration of the human eye and the retina (the zoomed-in region in the middle). The reflected light rays reach the retina at the back of the eye after passing through the cornea and the crystalline lens, while the pupil (located between the cornea and the lens) regulates the light entering the eye. The retina consists of a light-sensitive tissue that contains multiple pigmented, cellular, and nerve fiber layers. Cones and rods (easily recognizable from their shapes) populate the outer layer along the horizontal temporal section of the retina. The zoomed-in region depicts the three subtypes of cones in blue (S cones $\sim$ 420 nm), green (M cones $\sim$ 534 nm), and red (L cones $\sim$ 564 nm) according to their peak sensitivity to different wavelengths, which is responsible for the human trichromatic system that enables color vision. Rods (depicted in grey) are more sensitive to wavelengths around 498 nm and can only support achromatic vision. When the light reaches the retinal pigment epithelium, a cascade of events activates the sodium channels in the cell membrane of the photoreceptors. The biochemical message passes from rods and cones to the inner nuclear layers as an electrical signal through an interactive network of horizontal, bipolar, and amacrine cells that integrate and transmit the information to ganglion cells. Within the inner plexiform layer, the ganglion cells then convey the visual information to higher subcortical and cortical areas along the optic nerve. **Right**: the distribution of rods and cones as a function of eccentricity around the fovea (the data were extracted from Osterberg, 1935). Cone density is higher in the fovea and declines rapidly towards the periphery. In contrast, rods are absent in the fovea and populate the periphery of the retina. The blind spot (depicted as a white bar) corresponds to the location of the optic nerve where no photoreceptors are present

operating range). In contrast, cones are responsible for photopic vision (higher light levels where rods saturate); they have faster response and higher spatial resolution. This transition from rod to cone-dominated vision occurs at intermediate light levels (mesopic range), where both types are sensitive. Cones are further classified into three subtypes according to their spectral sensitivities (short, medium, and long wavelengths), which provides the basis for color discrimination. Unlike cones, rods contain a single photosensitive pigment and thus cannot support chromatic vision. The photoreceptors are arranged in a hexagonal grid (retinal mosaic; Curcio and Hendrickson, 1991) within the outer neuroepithelial layers of the retina. Their distribution varies as a function of eccentricity with a higher concentration of cones in

the fovea — the central location where vision is the sharpest (right column in Figure 0.1). Cone density then decreases towards the periphery and falls rapidly outside the fovea, where rods are located. Ganglion cells are within the inner retina, receiving the signals from rods and cones through several synaptic connections. Their activation region is generally divided into ON and OFF depending on whether they are triggered by increments or decrements in light intensity, respectively. ON and OFF subregions are located either in the center or the periphery of the cell's receptive field (ON-center combined with OFF-surround and vice versa); this center-surround antagonism generally reveals higher sensitivity to image gradients than uniform background fields (Figure 0.2). The axons of the cells converge at the optic nerve, and this segregation remains intact up to more advanced cortical areas.

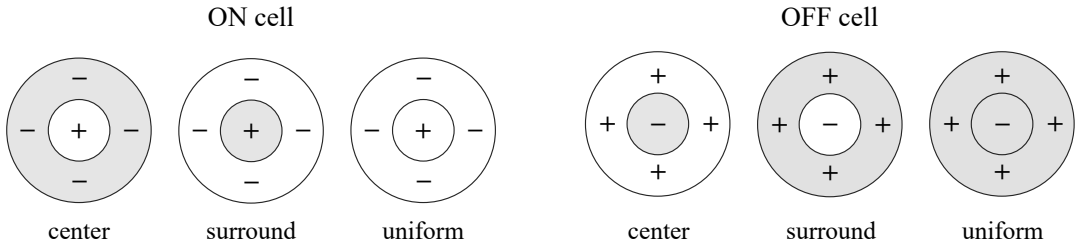


**Figure 0.2** An illustration of center-surround receptive fields. The left and right drawings depict ON-center and OFF-center (OFF-surround and ON-surround) cells, respectively. A spot of light flashed on the center of an ON cell (left first column) will cause excitation (indicated with the plus symbol) and increase its firing rate. Conversely, if the light falls on the same cell's surround (left second column), the response will be inhibited (indicated with the minus symbol). Suppose one uses a uniform stimulation instead that covers both the center and the surround of the cell (left third column). In that case, the surround will be inhibited and the center excited — a mechanism known as lateral inhibition. OFF cells (shown on the right) have the opposite organization (complementary responses and activation)

The human visual system can operate over a massive range of light intensities spanning more than eleven orders of magnitude (see Table 1.4 in Makous, 1998). However, only part of this range can be handled by the pupillary light reflex and the rod and cone subsystems described above; additional adaptation mechanisms are needed to maintain the system's responsiveness at different light levels (before the response reaches its ceiling; see Walraven et al., 1990). One characteristic example is pigment bleaching — photopigment molecules are depleted faster than they can be restored at high light levels reducing the photon catch and allowing the cones to continue to operate. Evidently, adaptation is a process that allows the visual system to respond to the relative luminance variations of the surrounding objects (the surface reflectances in a typical scene vary to a much lesser degree than the range of illumination throughout the day). The different aspects of light and dark adaptation have been previously analyzed using various paradigms (a detailed review can be found in Hood and Finkelstein, 1986). For the thesis

purposes, our primary interest lies in how the visual sensitivity to these relative luminance differences varies with the level of adaptation. In psychophysics, visual sensitivity refers to the reciprocal of the threshold contrast, i.e., the just-noticeable light increment over the background luminance required for the stimulus to be detected (given the Weber contrast definition). Note that for periodic light patterns (e.g., the spatial sine-wave in Figure 0.3), a more appropriate expression is the Michelson contrast, i.e., the amplitude of the luminance modulation over the average luminance (this is also the convention that this thesis follows). A comparison among the various contrast measures can be found in Kukkonen et al. (1993).

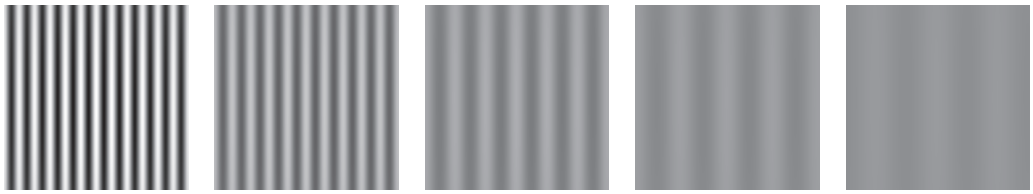

**Figure 0.3** A sinusoidal grating decreasing in spatial frequency and luminance contrast (rightwards). For periodic luminance patterns, contrast is most commonly defined as the difference between the maximum and minimum luminance values over their sum (Michelson contrast)

Early studies typically employed a brief light stimulus presented against a uniform background to measure sensitivity as a function of adaptation luminance (commonly referred to as threshold versus intensity curve; an example can be found in Figure 2.1). Sensitivity is generally known to increase with background luminance before obeying Weber's law, i.e., the change in luminance required to reach the threshold becomes proportional to the background level (sensitivity plateaus; Chapter 2 analyzes this relationship in more detail). However, it is well-known that adaptation luminance is not the only associated factor; several other stimulus characteristics affect sensitivity, e.g., the spatial and temporal frequencies of the luminance pattern (see a detailed summary in Graham, 1989). A widely-used measure for these effects is the contrast sensitivity function (CSF). The spatiotemporal CSF describes the sensitivity of the observer as a function of the spatial and temporal frequencies of the stimulus (Figure 0.4; see Chapter 2 for a detailed analysis of the spatial frequency component). Sensitivity variations in the other pattern dimensions (e.g., mean luminance, orientation, eccentricity, wavelength, and size) are often described through their effect on the CSF; the same also applies to individual factors such as the observer's age or other viewing conditions (e.g., monocular). In other words, the observer's ability to detect or discriminate visual patterns can be generally quantified by their CSF. It follows that combinations of spatial and temporal frequencies outside the area defined by the sensitivity limits are invisible to the human eye.

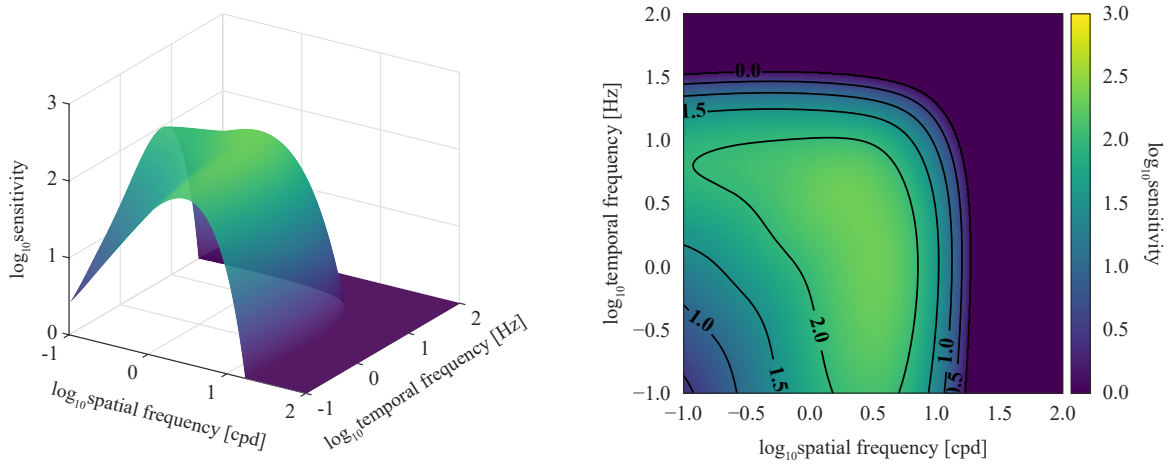


**Figure 0.4** Typical spatiotemporal contrast sensitivity surface (left) and contour (right) following the model described in Kelly (1979). The model was fitted to sensitivity measurements for drifting sinusoidal gratings at various spatial and temporal frequencies. The sensitivity boundary (i.e., where log sensitivity asymptotes to zero) describes the spatial and temporal limits; frequencies outside this range are invisible to the human eye. The cut-off frequencies were measured at the lower end of the photopic range, where contrast sensitivity is generally not expected to have reached a plateau. Note that the spatiotemporal surface measured with counterphase flickering stimuli (e.g., Robson, 1966) has the same shape over a wide frequency range, scaled-down by a factor of two in sensitivity after retinal stabilization (i.e., after compensating for the eye movements of the observer; Kelly, 1979)

Following this brief overview, it becomes apparent why central characteristics such as the spatial and temporal resolution (minimum sampling intervals), the average luminance, or the dynamic range of a visual display are linked with the properties of the human visual system. Classical sampling theory states that the minimum sampling requirement can be determined by the Nyquist limit — the signal can be exactly reconstructed if the sampling rate is at least twice the highest frequency component one wishes to sample (usually after prefiltering the signal to remove higher frequencies). If that requirement cannot be satisfied, then aliasing occurs; spectral aliases begin to overlap, and direct reconstruction is no longer possible (an example can be found in Figure 1.2). Ideally, the highest frequencies that one aims to reconstruct should generally match what human vision can perceive. One typical example is the appearance of sampled motion. Suppose one wishes to depict a vertical line moving smoothly across the screen at constant velocity (left column in Figure 0.5). Its frequency spectrum will be a line passing through the origin with spectral replicas along the temporal direction due to the finite sampling in the time domain (middle and right columns in Figure 0.5). If these replicas fall outside the region of visible frequencies (known as the window of visibility; Watson et al., 1986), one can generally assume that no artifacts will be visible and motion will appear smooth.

Although this is a simplified description of the observer's sensitivity, the general idea is clear; the window's edges define the region of frequencies that one would need to preserve in the visual signal. Naturally, factors that change the spatiotemporal sensitivity limits also affect the window's size and shape (see Watson, 2013a). Note that the necessary sampling rates cannot always be achieved, nor are they always desirable (e.g., when resource efficiency is the primary objective instead). In that case, the sampling requirements can also be optimally specified by considering the acceptability (rather than visibility) of the related visual artifacts.

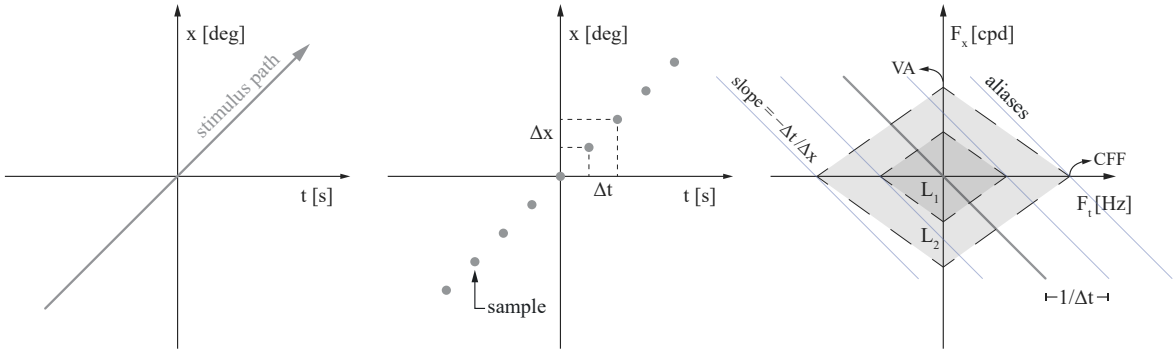


**Figure 0.5** An illustration of the window of visibility for sampled motion. **Left**: the horizontal position of a smoothly moving vertical line over time (gray arrow). **Middle**: a stroboscopic presentation of the stimulus with constant velocity $\Delta x/\Delta t$, where $\Delta t$ is the refresh interval and $\Delta x$ is the displacement. **Right**: the Fourier domain representation of the sampled signal. The horizontal and vertical axes show the temporal $(F_t)$ and spatial $(F_x)$ frequencies, respectively. The temporal sampling produces aliases spaced $1/\Delta t$ apart with a slope of $-\Delta t/\Delta x$ (parallel thin blue lines). The spectrum is otherwise identical to the smoothly moving stimulus (ignoring any subtractive distortions for simplicity). The shaded diamonds (similar to the sensitivity boundary in Figure 0.4 on linear scale) are a simplified representation of the spatiotemporal limits in human vision (referred to as window of visibility) at two different background luminance levels $(L_1 < L_2)$. The diamond corners depict the highest spatial (VA) and temporal (CFF) frequencies that can be perceived (commonly referred to as visual acuity and critical flicker fusion). If the spectral aliases lie outside this window (as shown for the lower luminance), we predict that they will be invisible to the human eye, and motion will appear smooth. In the opposite case (as shown for the higher luminance), the displayed image sequence would be contaminated with visible artifacts. Note that eye movements shear the window in the temporal frequency domain (Girod, 1992)

The visual scene is projected onto a two-dimensional retinal surface, but we perceive a three-dimensional world. Regardless of how dense and broad sampling one can effectively support in the other dimensions, the display will fall short of replicating optical reality if it cannot provide a truthful representation of the scene's depth. Depth cues can be broadly classified into ones based on triangulation, perspective projection, and light transport effects (Banks et al., 2016). The last two types include cues such as linear perspective, relative size, shading, and occlusions, whereas the first comprises binocular disparity, motion parallax, and

retinal blur (Figure 0.6). The primary interest (for the thesis purposes) lies in the first type, which is more closely related to the specific display characteristics and generally far more challenging to reproduce (to the extent that display classification is often based on which of these cues they can support). Binocular disparity describes the difference in the location of a scene object in the left and right retinal images. Motion parallax expresses how the retinal location of this object changes over time as the eyes translate. Blur (in the current context) refers to the defocus of objects located away from the eye's focal plane. Note that all three cues are neurally coupled with extraretinal signals; binocular disparity with eye vergence (see Figure 0.6), motion parallax with smooth eye movements (Nadler et al., 2009), and defocus blur with accommodation — the eye switches its focus (accommodates) to another object in depth by changing the power of the crystalline lens.

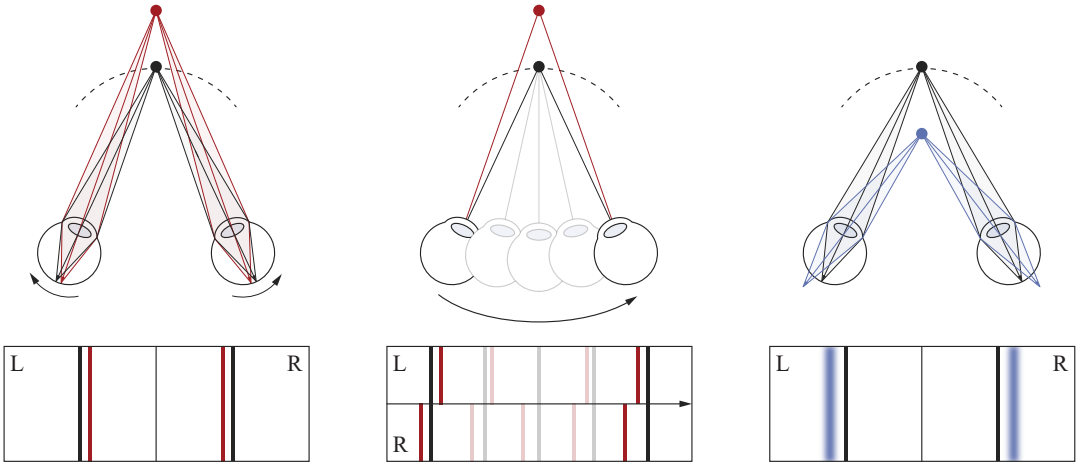


**Figure 0.6** An illustration of triangulation-based depth cues. The bottom panels show a simplified version of the corresponding left (L) and right (R) retinal projections. **Left**: binocular disparity, i.e., the difference in the corresponding locations of a scene object on the retinae (red vertical lines), and vergence, i.e., inward and outward counter-rotational eye movements. The curved dashed line depicts the geometric horopter — the set of points around fixation with zero retinal disparity (black vertical lines). Binocular disparities are commonly described as uncrossed for objects located outside the horopter (red marker in the left column; the eyes need to diverge to fixate on it) and crossed for objects closer to the observer (blue marker in the right column; the eyes need to converge instead). Note that the visual system cannot successfully fuse the two images if the range of disparities exceeds an upper limit on either side of the horopter (a region known as Panum's fusional area). **Middle**: motion parallax cues for an observer translating laterally to the right while maintaining fixation (the image depicts the left eye). The monocular retinal location of the object changes over time in a geometrically similar manner to binocular disparity. Objects at different depths produce different retinal velocities with opposite signs around fixation (objects with crossed and uncrossed disparities appear to move in opposite directions). **Right**: defocus blur and accommodation (collectively referred to as focus cues). Under natural viewing conditions the eye adjusts its accommodative distance to bring the point of fixation into focus (black marker). Objects at different distances (e.g., blue marker) are blurred on the retina

In the disparity domain, a similar but much narrower window of visibility can be generally defined based on the spatiotemporal disparity thresholds (Kane et al., 2014). Note that the observer's sensitivity function can be measured in a similar way to the CSF by using periodic patterns in depth instead of luminance. Thus, one can generally (within the same simplified framework) specify the spatiotemporal disparity variations they need to preserve in the stereoscopic signal (helpful in applications such as disparity compression, e.g., see Didyk et al., 2011). Conventional 3-D display technologies support disparity cues, but there is only a limited depth range relative to the focal distance that allows comfortable viewing (known as the zone of comfort; Shibata et al., 2011). On the display side, the primary source of this discomfort is the vergence-accommodation conflict (Hoffman et al., 2008). The eye accommodates to the physical display surface, but vergence varies depending on where the observer fixates when viewing the scene; this mismatch produces a conflict between disparity and focus cues. Typically, more advanced displays designed to address this issue introduce additional objectives depending on the specific technology they implement. These objectives may include, for instance, specifying the maximum depth plane separation in multiplane displays (the sampling of focal distances; see MacKenzie et al., 2010) or the minimum angular view density in directional displays (the density of discrete views entering the pupil; e.g., see Nakamura et al., 2013) to drive the accommodative response effectively. Furthermore, additional constraints on the disparity range apply to displays offering parallax cues to non-stationary observers. One characteristic example is the shallow depth-of-field (the range where objects appear sharp) in multiview systems. It is known that depicting scenes with a larger depth range requires a higher view density to avoid aliasing artifacts (Chapter 1 demonstrates this relationship in detail); however, the number of discrete views (and thus depth) is already upper-bounded in practice. This problem is often treated using prefiltering (depth-of-field blur; see Zwicker et al., 2006), leading to this effect. It follows that from a perceptual perspective, an ideal display would need to deliver smooth continuous parallax without noticeable image artifacts or cue conflicts.

Given that distortions may appear at various stages during the signal acquisition, transmission, and visualization process, satisfying these criteria becomes quite challenging in practice (particularly for complex natural scenes). As a result, the displayed images often suffer (to a higher or lesser degree) from visible artifacts, and hence subjective visual assessment becomes essential in evaluating and optimizing these processes. However, subjective assessments are usually impractical and tedious, creating a need for objective metrics that can perform similarly to a human observer. It is known that conventional pixel-based signal fidelity measures such as the peak signal-to-noise ratio (PSNR) do not generally correlate well with subjective judgments (e.g, see Wang and Bovik, 2009). This shortcoming led to numerous perceptually-based

quality metrics (i.e., computational models) incorporating how human observers process visual information. It should be noted that fidelity and quality are not interchangeable terms; fidelity is rather one of the various factors affecting the perceived quality (e.g., Winkler, 2001). For our purposes, the primary interest lies in the ability of these models to predict the threshold sampling density for any given scene by gradually increasing the strength and visibility of the related artifacts, and hence this distinction is not made. Quality metrics can be generally classified based on the input they require as reference. Full-reference metrics assume that the (undistorted) reference signal is known and available. Reduced-reference metrics require only some features of the reference signal. No-reference metrics, as the name also suggests, receive only the distorted signal. Naturally, having more information makes it easier for the model to predict the perceived degradation level; however, these requirements cannot always be satisfied in practice (which is why different options exist). Here, our evaluation was limited to full-reference metrics; other types may be included in future work (see general discussion). Any further classification, although not clear-cut, can be usually based on similarities in the underlying approach or the specialization of the respective family of metrics (a detailed introduction can be found in Wang and Bovik, 2006). Note that dedicated video and stereo image and video metrics form distinct subcategories. Although 2-D image metrics can also be applied to videos by pooling the individual scores (see Chapter 1), they cannot directly account for temporal factors such as different presentation rates and protocols. The same applies to stereo image and video pairs in the presence of asymmetric distortions.

In the development of full-reference metrics, one can distinguish two main approaches. The first (bottom-up) is derived from near-threshold psychophysics, modeling the different aspects of the visual system that determine artifact (error) visibility, e.g., the visible differences predictor (VDP; Daly, 1992). The second (top-down) is based on hypotheses about the overall functionality of the visual system, e.g., extracting the image structural information (SSIM; Wang et al., 2004). However, there is no sharp distinction between these (conceptually different) approaches, and indeed many metrics combine elements from both domains. A generic error visibility model is depicted in Figure 0.7. The input images are typically first passed through a preprocessing step. Common preprocessing operations include luminance conversion, color transformation, and low-pass filtering simulating the point-spread function of the eye's optics. The next stage is usually a decomposition of the preprocessed images. Both psychophysical and neurophysiological evidence supports the idea of multiple mechanisms selective to spatial frequency and orientation (e.g., Graham, 1985). Following a similar logic, these processing algorithms usually decompose the images into different frequency and orientation bands. Such decompositions often include the cortex transform (Watson, 1987) and steerable pyramids

(Simoncelli and Freeman, 1995). The differences in the resulting coefficients can then be adjusted (normalized), considering the variations in contrast sensitivity (i.e., the CSF) and masking effects (e.g., contrast masking; Legge and Foley, 1980) — the term masking generally describes how the presence of a signal component (masker) affects the visibility of another. Finally, these differences (errors) are combined into a single distortion measure using a pooling method (e.g., Minkowski summation). Note that the differences can also be mapped to detection probabilities (before pooling) using an appropriate psychometric function (see definition in Subsection 1.2.5). It follows that a visibility metric that claims it can effectively simulate all the relevant visual system components would also need to demonstrate it can predict the visual system response to arbitrary input. In reality, the human visual system is complex and not yet fully understood; thus, any modeling attempt to satisfy this criterion becomes quite challenging in practice (even comprehensive bottom-up models usually need to be based on many simplifying assumptions). Wang et al. (2004) summarized several drawbacks of error visibility methods, such as the fidelity-quality equivalence described above or the generalization to suprathreshold distortions, and proposed an alternative top-down measure based on the structural similarity between the images (currently one of the most widely used general-purpose image metrics). Structural distortions were indirectly defined as errors that cannot be classified as non-structural; non-structural errors are the ones that do not distort the structure of the objects in the scene, such as alterations in contrast and luminance. Naturally, these methods are not limited to structural-based measures; however, an in-depth review is beyond our purposes (a recent survey can be found in Zhai and Min, 2020).


**Figure 0.7** Block diagram of a generic error visibility metric. The reference and distorted images are first passed through preprocessing, followed by a decomposition and error normalization considering the contrast sensitivity function and masking effects. At the final stage, the errors are pooled to provide a single distortion measure. Note that the CSF is sometimes applied before the decomposition

The perceived image quality and naturalness of the depicted scenes are undoubtedly significant parts of the overall visual experience; however, they are not the only factors involved in specifying the optimal (from a human-centric perspective) sampling characteristics. As the discussion moves away from artifact visibility thresholds, aspects such as task performance or visual comfort and fatigue become critical in determining the limits that human observers can tolerate. The vergence-accommodation conflict described above is a characteristic example of how viewing discomfort becomes a central issue in the design of (stereoscopic) displays.

Indeed, the design of common display characteristics was partly influenced by human factors. For instance, it is known that perceptible display flicker induces visual discomfort (e.g., Isensee and Bennett, 1983). It is also known that the refresh rate is a central variable in the design of flicker-free displays (e.g., Watson and Ahumada, 2011). A similar connection can be made between the screen resolution and reduced visual performance (e.g., Ziefle, 1998). Naturally, these factors can be more effectively assessed under ecologically valid conditions.

### Outline and thesis structure

This dissertation investigated current aspects of modern visualization and display technologies from the perspective of the human observer. We explored three topics on (i) angular, (ii) spatial, and (iii) temporal sampling characteristics related to the visibility and acceptance of sampling distortions and visual discomfort. Each topic is presented in a dedicated and self-contained chapter, including the background and objectives, description of the experimental design, analysis of the results, discussion of the findings and their limitations, and potential extensions, with the necessary supplementary material in a separate appendix. The last chapter contains an overview of the main findings, limitations, and future work.

### Summary

A key objective in emerging 3-D technologies, e.g., lightfield displays (Yamaguchi, 2016), lies in their ability to adequately support full-parallax visualization, which would allow the viewer to experience a stereoscopic 3-D scene from multiple continuous perspectives, i.e., the observer can move around the scene and see different aspects of an object or occluded regions of the background. A fundamental problem in visualizing continuous-parallax content is the angular sampling limitations, i.e., the number of discrete perspectives. Trivially, a scene depicting a flat Lambertian surface on the screen plane would only require one view. A scene with a higher depth range would be expected to require more views. Additionally, the velocity with which the viewer moves around the scene might matter; changing rapidly between perspectives might make distortions less noticeable than if they occur widely-spaced in time.

In the first chapter, we investigated the perceptually optimal view density; the just-noticeable step for a human observer to perceive a distortion as they transition from one perspective to another. We addressed this problem by analyzing the artifacts that arise in an image sequence replicating the changing views of a synthetic 3-D object for a moving observer. Using a generic simulation on a passive stereoscopic display, we performed a psychophysical study to measure the visibility and acceptance of these artifacts changing the object's depth range and the virtual

observer's velocity. In a two-interval forced-choice task, participants indicated the "better looking" interval between a reference sequence where we explicitly updated the correct view at every image frame and a distorted one where a sub-sampling factor was modulated to increase the severity of the artifacts. As the view density threshold generally varies across different scenes and conditions, we assessed how well full-reference quality metrics could capture the perceived distortions in the simulated image sequences. We found that the metric scores could indirectly characterize the perceived degradation intensity on a unidimensional scale, allowing us to establish a general method for predicting observer performance.

The most general way of quantifying what human vision can perceive is the contrast sensitivity function (CSF). Recall that the CSF describes how the sensitivity to visual stimuli changes as a function of their spatial and temporal frequencies. It predicts which artifacts will be detectable on a display and what changes to hardware will result in noticeable improvements. Many of the quality metrics described above, were explicitly based on the visual system properties, as quantified by the CSF. One of the main factors affecting contrast sensitivity is adaptation luminance. Generally, increasing the luminance of a display results in higher sensitivity, with a shift of the peak magnitude towards higher frequencies (e.g., Van Nes and Bouman, 1967). Modern displays can now reproduce images at much higher luminance range than was previously feasible; thus, it is essential to understand how sensitivity varies in this extended range, particularly concerning its asymptotic behavior at high-light levels (i.e., the upper limit where increasing luminance further will not result in any improvement). The changes of sensitivity with luminance are often expected to fall into three regions from low to high luminance levels, described as a transition from the linear to DeVries-Rose to Weber range where contrast sensitivity reaches its ceiling (Graham, 1989). The CSF has been widely studied in the past; however, datasets where this extended range could be more thoroughly investigated, particularly the region near saturation at high frequencies, are limited.

In the second chapter, we measured the spatial CSF at relatively high luminance levels and frequencies using a custom-built display setup and supplemented our data with a large number of measurements in the literature. We then explored the effect of luminance on contrast sensitivity and assessed how well the theoretical transition described above approximates the empirical data. Using segmented regression analysis, we verified a gradual transition among the three regions with steeper slopes for higher spatial frequencies and lower luminance levels. A consistent sensitivity decrease following saturation was also found at low to intermediate spatial frequencies across studies. Based on this theoretical construct, we adopted and modified a CSF model (Rovamo et al., 1994, 1997) comprising a low-pass optical modulation transfer function, high-pass lateral inhibition, a matched filter, and three limiting internal noise components

corresponding to each region. We then assessed the model's performance on the full dataset against similar approaches and suggested an eight-parameter form to approximate the contrast sensitivity surface in the spatial frequency-luminance domain.

Another characteristic of modern visual displays is their capability to support higher frame rates that improve motion fidelity by reducing the severity of the related artifacts, e.g., motion blur that arises in hold-type displays. As the optimal frame rate is naturally task-specific, the overall user acceptance and perceptual quality and comfort will vary under different conditions, with more dynamic content demanding a higher frame rate. However, most consumer displays today typically operate at relatively low frame rates (usually 60 Hz), with an even lower rate for portable devices where resources (e.g., battery life) are limited, making them more susceptible to motion artifacts. Since the frame rate modulation is also relevant for moving static images and text around the screen, e.g., when a user scrolls through a document, one question that naturally follows is whether these effects are also evident during typical display use, or observers generally exhibit some tolerance. Currently, little is known about the potential gains in the perceived quality and comfort when users perform regular display tasks.

In the third chapter, we measured the subjective effects of high frame rates (up to 240 Hz) on visual comfort, fatigue, and overall preference and investigated their correlation with objective indicators in task performance and blinking activity during regular display use. In a naturalistic reward-based scan reading task, the participants scrolled through a text document searching for errors at different frame rates while we recorded their eye activity using a binocular eye tracker. We evaluated two scenarios: in the first one, the subjects scrolled through the text at will, whereas, in the second, the text speed was fixed at a relatively high value increasing the task dynamics. The measurements indicated that for the self-paced scrolling condition, the observers performed the same with no significant differences in their overall preference, and no reported increase in visual fatigue and discomfort symptoms for frame rates down to 30 frames/s. During the more dynamic task, the differences were generally evident even at the highest frame rates (given the observer preferences) without any significant decrease in task performance. In both cases, the increase in blinking activity generally correlated with the decrease in the frame rate, with a more substantial effect during the dynamic task. However, the effects were practically small compared with other associated factors. The results suggested that for everyday display tasks, e.g., self-paced scrolling through text at moderate speed, the frame rate can be substantially reduced (even to 30 frames/s or lower) without severe effects on the user. In contrast, this tolerance diminished under the more dynamic content.

# Chapter 1. Visibility of angular distortions in full-parallax content

## 1.1. Introduction

The optimal representation of a three-dimensional scene, regardless of the specific display technology, could be defined as one which provides a veridical free-viewing visual experience by minimizing the required computational and visualization resources. Given this objective, a display would need to replicate the equivalent retinal images that would be created if an observer or multiple observers moved around the natural scene by also eliminating redundant visual information. In other words, one would need to sample and reproduce the rays of light through every point in space, i.e., the lightfield (Gershun, 1939), based on the properties and limitations of the human visual system (HVS). The distribution of these light rays in the most general form can be expressed as a 7-D function, denoted by $P(\theta, \phi, \lambda, t, V_x, V_y, V_z)$, where $\theta$ and $\phi$ are the propagation angles, $\lambda$ is the wavelength, $t$ is the time, and $\mathbf{V}$ is the viewing position in 3-D space (Adelson and Bergen, 1991). It follows that for a monochromatic static feature in the scene, this function can be reduced to four dimensions in free space. This 4-D function can then be represented in various ways, with the two-plane parameterization (Levoy and Hanrahan, 1996) being the most common. In that case, the lightfield is denoted by $L(u, v, s, t)$, where $(u, v)$ and $(s, t)$ are the points where each ray crosses the two planes (see Figure 1.1). This convenient structure allows us to assign one plane to the different scene viewpoints and the other to the image or display plane depending on the context (image acquisition or display). However, due to the discretization, one can only sample this field with a finite spatial and angular resolution (number of perspectives). This chapter explores the perceptually optimal discrete angular view density for continuous parallax; that is, how finely-spaced must these views be so that a human observer does not notice any distortion as they transition from one perspective to another.

Naturally, failure to meet the angular sampling requirements could result in visible artifacts during the image acquisition and rendering process or the display, e.g., abrupt transitions and double or blurry images (see examples in Halle, 1994; Levoy and Hanrahan, 1996; Moller and Travis, 2005). If one ignores the specific display characteristics, the minimum sampling rate can be analytically derived as a function of the scene depth and highest spatial frequency

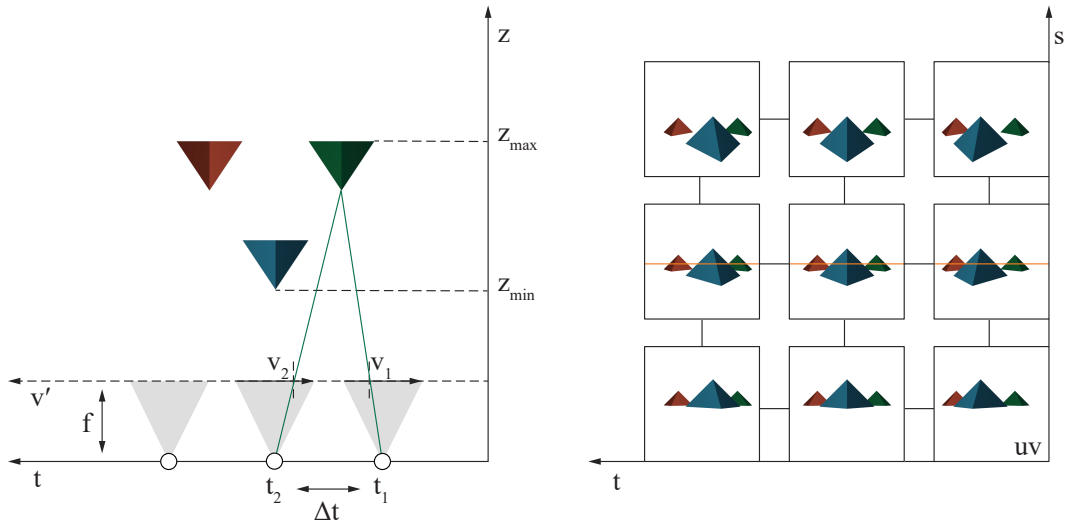


**Figure 1.1 Left**: Top-view of an image acquisition setup using the two-plane parameterization. A point in the 3-D scene at depth $z$ is seen by two pinhole cameras located at $t_1$ and $t_2$ with focal length $f$, at horizontal image coordinates $v_1$ and $v_2$ with disparity $\Delta v = v_2 - v_1$. **Right**: The captured images using a $3 \times 3$ grid on the st-plane. Slicing the views (here the central) along the u-axis and stacking the pixels on the orange line would form an EPI of the 3-D scene, similar to the one depicted in Figure 1.2

following the work in Chai et al. (2000). This approach can be briefly demonstrated if we consider equidistant camera positions along the t-axis described above (horizontal parallax). The captured images can then be combined to form a 3-D volume of different perspectives with the t-axis perpendicular to the uv-plane. This arrangement can be further simplified by slicing the volume along the u-axis (i.e., stacking the image rows for a fixed u across the perspective views). Consequently, the complex scene structure can be directly mapped to features within these slices or epipolar plane images (EPI) as defined by Bolles et al. (1987). In this 2-D representation, any visible point in the scene would appear as a line with a slope determined by its depth and the adjacent views distance. This regular structure of the EPI can then be exploited in the Fourier domain ($F_v$ - $F_t$ frequency axes; see Figure 1.2), where its spectrum is limited by the minimum and maximum depths in the scene forming a "bow-tie" shape with replicas along the $F_v$ and $F_t$ direction due to the finite spatial and angular sampling. As the replicas in $F_t$ direction are spaced $2\pi/\Delta t$ apart, by increasing the sampling interval $\Delta t$ along the $t$-dimension, they will start to overlap resulting in these aliasing artifacts. It follows that a higher depth range requires a higher view density. In any case, a sufficient condition to avoid these artifacts is to limit the disparity among the neighboring camera images to $\pm 1$ pixel. Although this approach is given in the context of signal reconstruction, similar principles would apply to the displayed image sequence if we simulated an observer translating along the camera path. Note that when the observer moves between two sampled perspectives,

due to the eye offset from the intended viewing position, the acceptance of the intermediate views (e.g., the nearest-neighbor perspective) is not solely controlled by these artifacts but could also be affected by geometric distortions. However, we expect these effects to appear at relatively larger mismatches as human observers partially compensate for incorrect viewing positions (Vangorp et al., 2013). Additionally, as each eye receives the correct perspective at different intervals, stereoscopic effects could also emerge, e.g., in the form of temporal binocular disparity variations, asymmetric distortions, or even temporary loss of stereo parallax when both eyes receive the same perspective. The purpose of this study was to assess from an end-user viewpoint the tolerance of a human observer to these artifacts and provide a method to approximate the perceived distortion intensity.

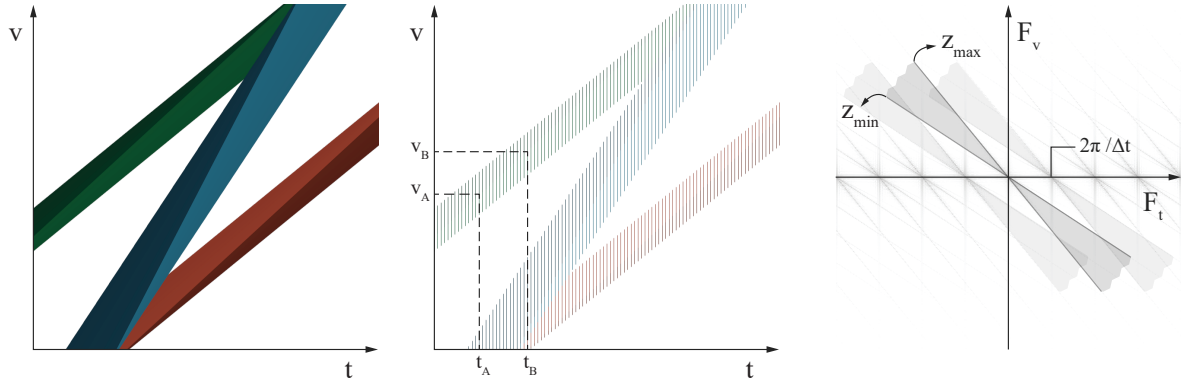


**Figure 1.2 Left**: The central EPI of the continuous lightfield signal for the above scene with three objects placed at two depth levels. **Center**: Sparse sampling along the $t$-dimension. The $v_A$ and $v_B$ indicate the image coordinates of a point seen by two cameras positioned at $t_A$ and $t_B$. **Right**: The resulting spectrum (background) with a diagram depicting the baseband (bounded by lines corresponding to the minimum and maximum depth $z$ of the scene from the two planes) and its aliases (lighter gray) at $2\pi/\Delta t$ distance along the $F_t$ frequency axis

A subjective threshold estimation for all possible conditions and variations would be infeasible. Predicting artifact visibility or acceptance would require modeling both the specific viewing conditions (e.g., display luminance) and the human observer characteristics (e.g., the spatiotemporal contrast sensitivity function), which can become highly challenging as the complexity of the visual input increases. A straightforward "off-the-shelf" method for approximating the perceived degradation by integrating how the HVS processes the visual information (even with a top-down approach based on its overall functionality) is quality metrics. Recall that quality metrics are computational models designed to objectively evaluate the quality of distorted images and videos in agreement with human judgments. As essentially the retinal image sequence quality degrades, one would expect the state-of-the-art metrics to describe human near-threshold performance to some extent, even if they are not primarily

constructed to capture distortions in the angular domain. Previous studies indicated that specific metrics correlate with the perceived quality for compressed lightfield images (Paudyal et al., 2017) and the perceptual scaling of angular specific lightfield distortions when full-reference (i.e., sufficiently dense sequence) is available (Adhikarla et al., 2017).

Following this direction, we generated 3-D image sequences replicating the changing views of a static scene for a moving observer (see Figure 1.3). We used a synthetic stimulus modulating the depth range, the virtual observer's movement speed, and the angular view density. The resulting artifacts included perspective distortions due to incorrect viewing positions, non-smooth motion parallax caused by abrupt view transitions, and blurry images with double contours as a result of inter-perspective crosstalk (referred to here as view blending; a common artifact in autostereoscopic multiview displays). Stereoscopic effects related to the asynchronous correct perspective update in the two eyes and the temporary loss of stereo parallax were also simulated. The analysis was limited to a Lambertian scene with a single 3-D object and horizontal parallax. We measured the perceived degradation using a two-interval forced-choice task on a passive stereoscopic display and evaluated the distortion scaling similarity between the observed data and a wide range of 2-D and stereoscopic image and video quality metrics. We then established a quantitative relationship between the metric scores and the observer performance, which allows for threshold estimation on variable input.

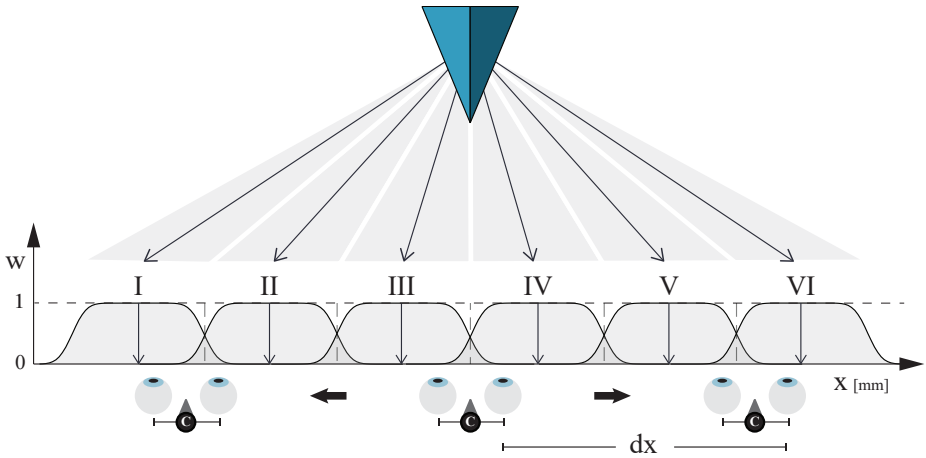


**Figure 1.3** Schematic of the simulated observer where $w$ is the relative weight of the respective views or zones (Latin numerals) at the corresponding eye locations ($X$) and $dx$ is the varying movement step size. The weight $w$ is given from the crosstalk model (referred to here as view blending) as described below. For instance, the observer at the leftmost position sees a sharp image of the first perspective with their left eye, whereas their right eye receives a composite image of the first two perspectives. If view blending is disabled, then the right eye will see a sharp image of the second perspective. In both cases, the retinal images are distorted due to the offset from the correct viewing positions

## 1.2. Methods

### *1.2.1. Participants*

Twelve volunteers were divided into two groups. The first group (n = 7, mean age: 35 years, SD: 13 years, span 21 – 55 years, four males and three females) evaluated the stimulus shown in Figure 1.5 placed at the zero parallax plane. The second group (n = 5, mean age: 33 years, SD: 11 years, span 23 – 50 years, three males and two females) evaluated the same stimulus under the same conditions except for a translation of the 3-D object in depth. They had normal or corrected to normal vision with no history of visual problems. All the observers were able to complete the training session accurately. The author was excluded. The Newcastle University Ethics Committee approved the study, and all participants gave written informed consent.

### *1.2.2. Apparatus*

The stimuli were presented on a 47" passive stereoscopic display (LG 47LD920-ZA) in a dark room at a total resolution of 1920×1080 pixels and a frame rate of 60 Hz, with the left and right views horizontally interleaved. The average luminance of the 3-D object through the polarized glasses was 32 $cd/m^2$, as measured using the Minolta LS-100 meter. Participants were positioned at a viewing distance of 150 cm, measured perpendicularly from the center of the screen to their eyes' midpoint. They viewed the simulated stereoscopic video from a static position using a chin rest to ensure their eyes were aligned correctly. Before each session, the simulation parameters were adjusted to the participant's interocular distance.

### *1.2.3. Image sequence simulation*

The object views were pre-rendered in Blender software (version 2.78, Stichting Blender Foundation, Amsterdam, the Netherlands) using a dense regular horizontal camera grid (Figure 1.4) with a baseline of .5 mm (maximum disparity less than 1 pixel). The array was positioned at 150 cm, focused on a plane passing through the virtual world zero coordinates. The sensor size and the focal length for each camera were set to 35 mm and 29 mm, respectively. Near-zero disparity at the focus plane was achieved by horizontally shifting the camera sensors towards the center (the optical axes remain parallel) using the tools developed by Honauer et al. (2016). The textures were rendered at a resolution of 1920×1080 pixels. The 3-D object, when its geometrical center was placed at the world origin, extended from (-33, -14, -20) to (45, 22, 18) cm, in the horizontal, vertical, and depth axes, respectively. The camera grid extended symmetrically around the central view at a distance of 45 cm covering approximately

a 36 degrees field-of-view relative to the object. The observer's eyes were simulated as two pinhole cameras with off-axis projections coinciding on the zero parallax plane at a distance of 150 cm. The virtual eyes translated on a horizontal path (same as the rendering array) around two identical virtual planes (one plane visible to each eye) placed at zero parallax. The corresponding views (rendered images) at each eye location were then projected on these virtual planes. For simplicity, the movement step size $(dx)$ was set as a multiple of the array baseline value. Before the images were projected on the virtual planes, the textures of the adjacent views were passed through an alpha channel blending function using a super-Lorentzian profile (Li, 2002) to account for display related effects (Figure 1.3; the weight $w$ was given as function of the relative to the zone center and normalized by half its width position $x$: $w(x) = (1+|x|^M)^{-1}$). The parameter $M$ controlled the amount of crosstalk between adjacent perspectives. The output images of the virtual eyes (the image sequence replicating the movement around the object) were then horizontally interleaved and projected on the physical stereoscopic display at the fixed subject's location (Figure 1.4).

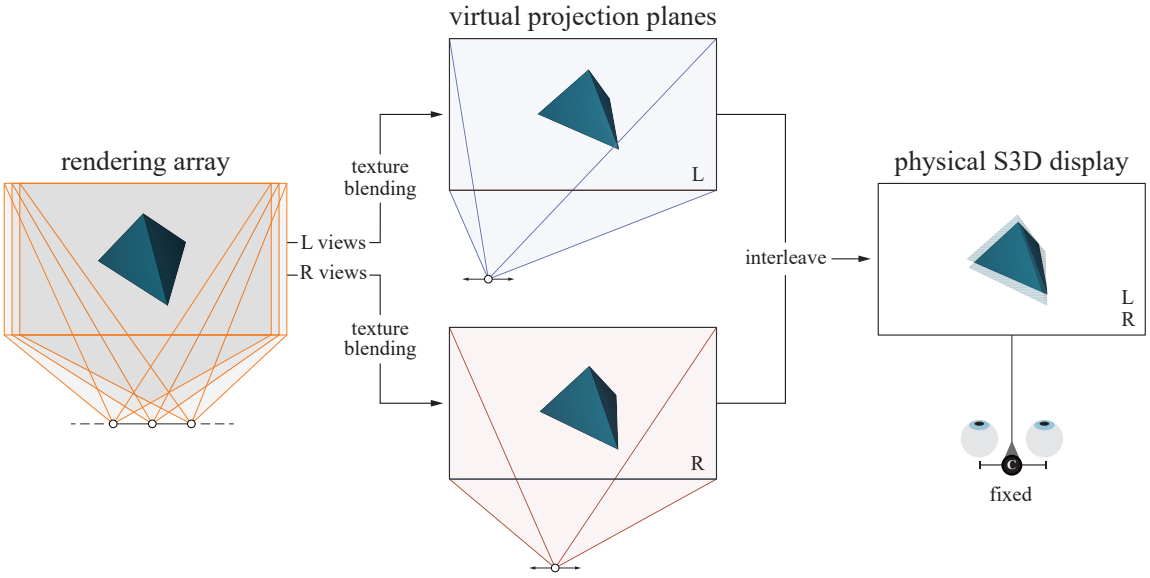


**Figure 1.4** The image sequence generation process for the simulated moving observer

### *1.2.4. Experimental task and conditions*

In a two-interval-forced-choice task (2-IFC), the subjects indicated the "better looking" image sequence between a dense reference condition and a randomly chosen distorted one (Figure 1.5). In the reference condition, we explicitly updated the correct view of the synthetic stimulus at each simulated observer's location. The severity of the distortions was mainly modulated

using a subsampling factor $N$ (distortion level); that is, we updated the correct view at every $N$ steps. The movement step size $(dx)$ was set at 2 mm and 3 mm or equivalently at a velocity of 12 cm/s and 18 cm/s, and eight distortion levels $N \in \{2,3,4,5,7,9,12,16\}$ were evaluated. The view blending was examined as a modified distorted condition ($M = 4$) for the same parameter values as in the sharp view transitions (no blending; $M = 0$). For the first group of participants, the single 3-D object was rendered with its geometrical center positioned on the zero-parallax plane (coinciding with the physical plane of the monitor). For the second group, we explored the effect of moving the object at 15.6 cm off the screen plane and towards the observer. Points further from the zero-parallax plane create larger discontinuities between the successive views in the image sequence, thus resulting in more noticeable distortions (e.g., abrupt transitions). The angular size of the object at its central location relative to the subject was approximately $18\times8$ degrees (horizontal $\times$ vertical angle), increasing to $20\times10$ degrees when we translated the object in depth. The interval (black image) between the two sequences was set at 1 s. Each distorted condition was evaluated five times for a total of 160 trials per observer. The order of the two simulated videos was also randomized. At random intervals, a randomly selected scene with a different object replaced the tested stimulus to maintain participant engagement. The experiment was completed in one visit lasting approximately 1 hour.

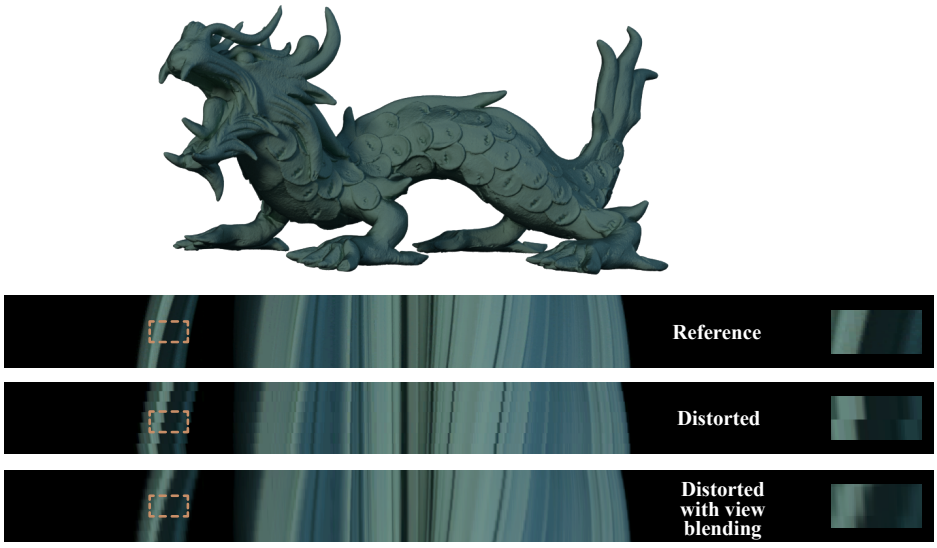


**Figure 1.5** The 3-D model (The Stanford 3D Scanning Repository, 2020) used for generating the experimental stimuli and its central epipolar plane image (i.e., the central image rows from the simulated views stacked together) with an example of the respective distortions. On the right, a zoomed-in version of the image region indicated with the dashed rectangle on the left

### *1.2.5. The psychometric function and the degradation scale*

Given that the sequence degradation is a monotonically increasing function of the subsampling factor $N$, and a correct response can be defined as the selection of the dense sequence in the paired comparison, we can assume that the observer's performance (i.e., the ratio of correct responses) will rise monotonically from an equal probability in the sequence selection to absolute preference for the reference. The general form of this psychometric function $\psi$ (equation 1.1) is commonly given for stimulus intensity $x$ as:

$$\psi(x;\alpha,\beta,\gamma,\lambda) = \gamma + (1-\gamma-\lambda)F(x;\alpha,\beta) \tag{1.1}$$

where $\gamma$ is the guess rate, $\lambda$ is the lapse rate (i.e., the probability of a failure irrespective of the stimulus intensity), and $F(x;\alpha,\beta)$ a function that relates the stimulus intensity $x$ to the underlying sensory mechanism performance, with $\alpha$ and $\beta$ the location and scale parameters (commonly referred to as "threshold" and "slope"). Assuming that the sensory effect corresponding to each sequence is a one-dimensional Gaussian random variable, signal detection theory (Green and Swets, 1966) predicts that performance will rise monotonically from "guessing" (when the difference in the respective means is zero) to perfect (when the distributions do not overlap) as the upper half of a cumulative Gaussian function (their difference is also normally distributed). In practice, our main interest lies in the estimation of the threshold and slope parameters of a well-fitting sigmoid curve; thus, theoretical considerations on the underlying sensory process are not our primary objective. In this study, as it is often the case, the parameter estimates were not considerably affected by the choice of the function; thus, due to slightly better-fitting performance, we selected the Weibull form (equation 1.2).

$$F(x;\alpha,\beta) = 1 - \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right] \tag{1.2}$$

For each condition (i.e., combinations of step size and blending power), we can express the perceptual intensity of the sequence degradation by replacing the factor $N$ with a more relevant quantity defined as the maximum pixel disparity $(d)$ among the input views for the 3-D object's mesh coordinates at depth $z$ (equation 1.3; see the supplement in Honauer et al., 2016).

$$\begin{aligned} d &= (\beta/z - N \times dx \times \text{focal length} \times \text{resolution}) \,/\, (\text{focus distance} \times \text{sensor size}) \\ \beta &= N \times dx \times \text{focal length} \times \text{focus distance} \times \text{resolution} \end{aligned} \tag{1.3}$$

Note that we refer to the disparity in the neighboring camera images and not binocular disparity (see also the general introduction). The probability of selecting the reference stimulus over the

distorted one then extends from .5 (guess rate for a 2-IFC task; the lower asymptote) to 1 - $\lambda$ (the upper asymptote) as a function of increasing pixel disparity.

This distortion intensity measure, however, even if in our design describes the main modulating factor, it is not sufficient to capture different sources of distortion at the same disparity level or their interaction (which might also be complex). Therefore, as our objective was to extract a unidimensional interval scale that can generalize to a different input (including textural information), we suggested using the existing quality metric scores (in the absence of a dedicated model) to approximate the overall perceived sequence degradation intensity (Kaspiris-Rousellis et al., 2018). Explicitly, considering the ease of implementation and consistency with the related studies (see the chapter introduction), we tested the following full-reference 2-D and stereoscopic image and video objective metrics: SSIM (Wang et al., 2004), MS-SSIM (Wang et al., 2003), BVQM (Pinson and Wolf, 2004), VIF (Sheikh et al., 2005), PSNR-HVS (Egiazarian et al., 2006), PSNR-HVS-M (Ponomarenko et al., 2007), CW-SSIM (Sampat et al., 2009), MAD (Larson and Chandler, 2010), IW-SSIM (Wang and Li, 2011), FSIM (Zhang et al., 2011), HDR-VDP-2.2 (Mantiuk et al., 2011), SIQM (Chen et al., 2013), StSD-LC (De Silva et al., 2013), GMSD (Xue et al., 2014), VSI (Zhang et al., 2014), VSPC (Jia and Wang, 2017), and PSNR (considered as the baseline metric). The inclusion of several similar models was intentional, as they would directly indicate which components may improve prediction or the opposite (albeit the latter would be less informative).

The psychometric functions were fitted based on the maximum likelihood using the Nelder-Mead simplex algorithm as implemented in the Palamedes toolbox (Prins and Kingdom, 2018). The lapse rate was fixed at 1% (Klein, 2001) to reduce the number of free parameters. Following the suggestions by Wichmann and Hill (2001), the fitting performance was evaluated in terms of deviance (equation 1.4 for binomial data, where $y$ and $p$ the observed and predicted success rates for $n$ trials at intensity level $i$; see derivation in Collett, 2003).

$$D = 2\sum_{i=1}^{K}\left\{n_i y_i \log\left(\frac{y_i}{p_i}\right) + n_i(1-y_i)\log\left(\frac{1-y_i}{1-p_i}\right)\right\} \tag{1.4}$$

## 1.3. Results

### *1.3.1. Psychophysical data*

Figure 1.6 illustrates the average observer performance for the stimulus shown in Figure 1.5 when its geometrical center was positioned on the zero-parallax plane (left column) or at a distance of 15.6 cm towards the observer (right column). We recorded a correct response when

subjects indicated the reference stimulus (dense image sequence) in the paired comparison. Recall that in the reference condition, we updated the correct view at every simulated position $(N = 1)$. In the distorted sequence, the correct view was updated at successively wider spaced intervals $(N > 1)$. The observer performance was near chance, i.e., artifacts were negligible or invisible, at low $N$ levels but rose to perfect as the $N$ level increased, i.e., artifacts became apparent. As expected, the perceived visual quality was negatively affected (overall better performance) by translating the object in depth. This effect was partially negated when view blending was enabled, i.e., blurrier but smoother view transitions in the simulated movement. In contrast, the movement step size effect was practically negligible (possibly due to the small increase in the design parameters, an interaction between the movement speed and pixel disparities, i.e., larger discontinuities but faster transitions, or both).

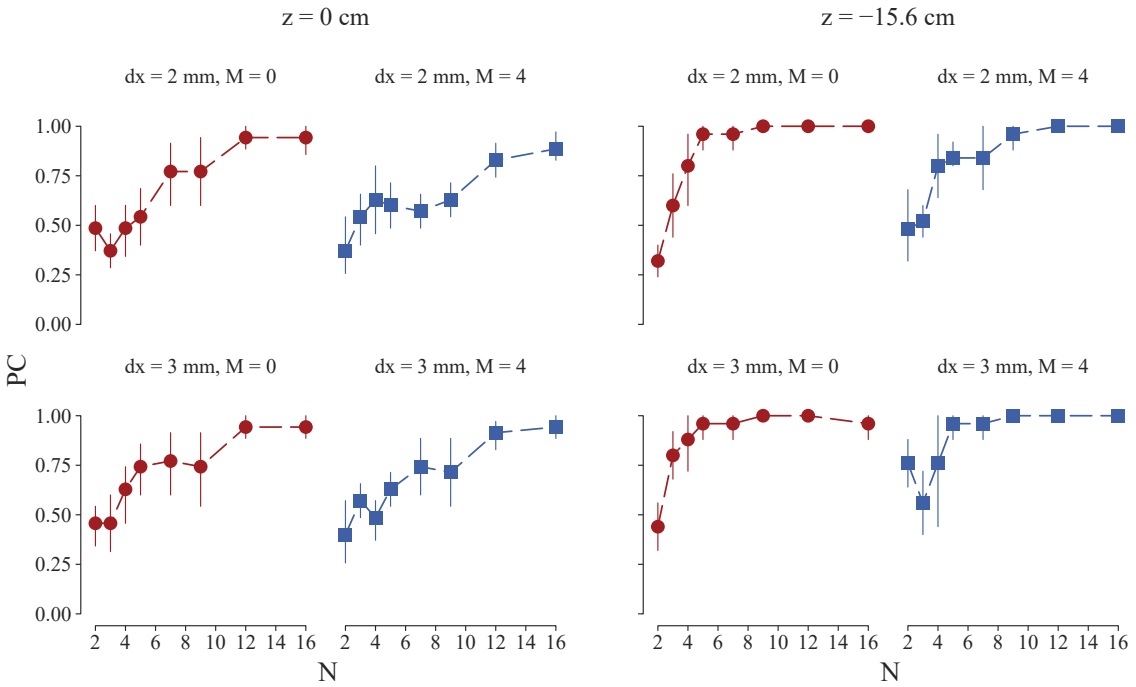


**Figure 1.6** The average proportion of correct responses across observers relative to the distortion level $(N)$, the step size $(dx)$, the view blending power $(M$; zero for sharp view transitions), and the distance $(z)$ from the zero parallax plane (the first group of participants is shown on the left and the second group on the right). A proportion of 1 indicates that the reference condition was always preferred to the distorted one (artifacts evident). The level of chance performance was at .5 (artifacts invisible or negligible). The error bars show the 95% CI by basic bootstrap

Figure 1.7 depicts the threshold estimates of the psychometric fits, assuming that the degradation intensity for each condition can be expressed as a function of the maximum pixel disparity in the updated input views (normalizing the depth and step size). The average threshold across observers ranged roughly between 3 and 5 pixels among the conditions. A summary of

the estimates is given in Table 1.1. The slope parameters were omitted as some of the estimated values tended to infinity (the fitted curve could be approximated by a step function). Figure A.1 shows the individual fits. The results revealed a consistent effect of the view blending on the estimated threshold (a mean overall increase by a factor of 1.29), albeit practically small. If we simultaneously fitted a single psychometric function among the examined conditions for each observer, the probability of obtaining a likelihood-ratio in the Monte-Carlo (MC) samples generated from the restricted model less than the observed value (see a practical guide in Kingdom and Prins, 2016) was relatively small ($p_{LR}$ = .050); thus, assuming the lesser model (the same function parameters across conditions) is not well-supported.

| view blending | dx [mm] | z = 0 cm (n = 7) | | | | | z = -15.6 cm (n = 5) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Mean | SD | Q1 | Mdn | Q3 | Mean | SD | Q1 | Mdn | Q3 |
| M = 0 | 2 | 3.2 | 1.3 | 2.0 | 3.0 | 4.3 | 2.6 | .5 | 2.5 | 2.6 | 2.7 |
| | 3 | 4.4 | 2.4 | 2.8 | 3.2 | 6.1 | 3.6 | 1.5 | 3.0 | 3.0 | 3.5 |
| M = 4 | 2 | 4.8 | 2.0 | 3.7 | 5.0 | 5.2 | 3.5 | 1.0 | 3.4 | 3.4 | 3.8 |
| | 3 | 5.3 | 2.3 | 3.7 | 5.0 | 6.8 | 4.1 | 1.7 | 3.0 | 3.9 | 4.0 |

**Table 1.1** Summary of the disparity threshold estimates for each examined condition across observers

### *1.3.2. Distortion ranking similarity and metrics performance*

Figure 1.8 illustrates how similarly the different metrics ranked the distorted conditions in our study in terms of increasing perceptual degradation (i.e., how they cluster together). For the 2-D image metrics, we calculated the final score (the predicted quality degradation compared with the reference) as the mean value across the sequence image pairs. Generally, the largest discrepancies were among the stereoscopic and video quality metrics, not unexpectedly, as they are tuned to different types of distortions. The metrics performance on the psychophysical data was then assessed both in terms of the distortion ranking similarity and the fitted psychometric function deviance across observers (see Figure 1.9). Note that the correlation coefficients are expected to be affected (presumably decreased) by the upper bound on the proportion of correct responses and the relatively small number of individual repetitions (the actual ranking cannot be fully determined; see Figure A.6). However, they are still informative in terms of relative performance and thus were retained for visualization purposes (e.g., see the BVQM metric's departure from linearity). The metrics that best predicted the measured data (average deviance lower than 26 on the individual responses) clustered around the GMSD and the SSIM

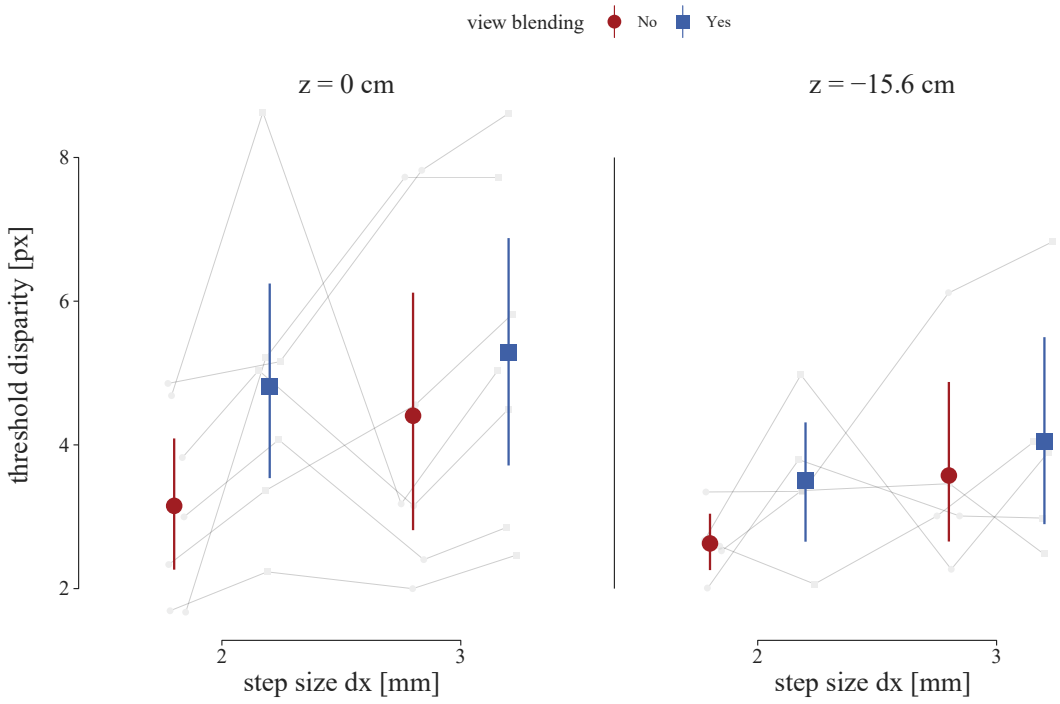


**Figure 1.7** The average estimated thresholds for each condition across observers with the distortion intensity given as the maximum pixel disparity in the updated input views (*Ndx*). The error bars show the 95% CI by basic bootstrap. The thin lines depict the individual observers

(see connected nodes in Figure 1.8), with the former group of 2-D image metrics slightly outperforming the latter. For our purposes, we selected the GMSD metric due also to its consistency across different psychometric functions and the various subsets of experimental conditions (Appendix A). Practically, any other metric in that group except for the IW-SSIM, explicitly the VSPC, the MAD, and the HDR-VDP-2, performed equally well. Note that the parameters of each metric were not tuned to our data.

### *1.3.3. Model estimates and validation*

Figure 1.10 illustrates the individual psychometric function fits with the perceived degradation intensity approximated by the GMSD score. Generally, the metric's scores predicted the difference between the reference condition and each distorted sequence sufficiently. The median threshold estimate across observers was $7.1 \times 10^{-3}$, bias-corrected and accelerated ($BC_a$) $CI_{95\%}$:[5.3, 10.8] $\times 10^{-3}$ by bootstrap. Likewise, the median slope, excluding two of the observers where the parameter value tended to infinity, was 1.9, $BC_a$-$CI_{95\%}$:[1.5, 2.1]. When we constrained the slope to be equal among the observers, the p-value of the likelihood-ratio test for the nested model, based on the MC generated distribution, was $p_{LR}$ = .649, indicating

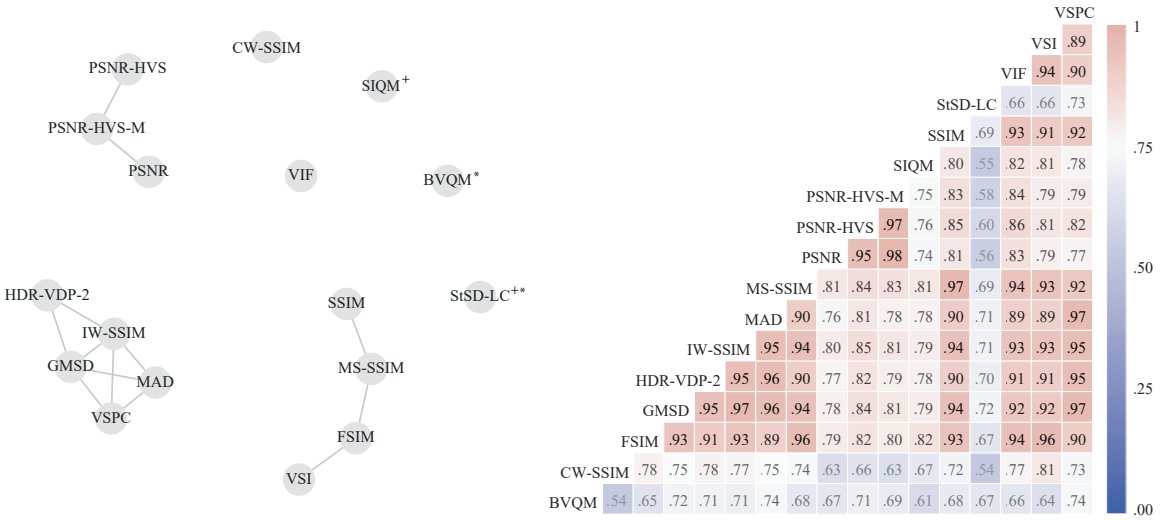


**Figure 1.8 Left**: A Kendall $\tau_b$ correlation network diagram of the distorted conditions ranking among the examined metrics in terms of increasing degradation (edges were drawn for an arbitrary value of .95 and above). The stereoscopic and video metrics are indicated with a cross and an asterisk, respectively. **Right**: The full metric correlation matrix (see examples in Figure A.5)

that it is reasonable to assume the lesser model. The optimal global slope estimate was 2.0, $BC_a$-$CI_{95\%}$:[1.7, 2.9] by parametric bootstrap, with the median threshold across observers approximately the same at $7.1 \times 10^{-3}$, $BC_a$-$CI_{95\%}$:[5.2, 9.9]. Figure 1.11 depicts the threshold and slope estimates for each observer under the final model (constrained slope). In the same figure, we also illustrate the model estimates on different subsets of the examined conditions. Explicitly, we tested fitting the data separately for each step size (dx2 and dx3), view blending (M0 and M4), and object depth (A and B). The maximum absolute difference in the median threshold estimates among the subsets was $1.9 \times 10^{-3}$, found between the two depth range levels (not unexpectedly as they corresponded to different groups of participants). In all cases, the absolute difference relative to the median threshold estimate of the full set did not exceed $1.0 \times 10^{-3}$; however, the slope estimates generally diverged when the number of samples dropped by half (from 32 to 16 per observer in all subsets of the within-subjects conditions). Figure 1.12 provides an example of the observer performance prediction across conditions and participants based on a split (i.e., training and validation on different subsets) between the two depth levels. These results suggest that even with suboptimal distortion levels sampling, we can sufficiently predict the threshold value across the different conditions (within limits).

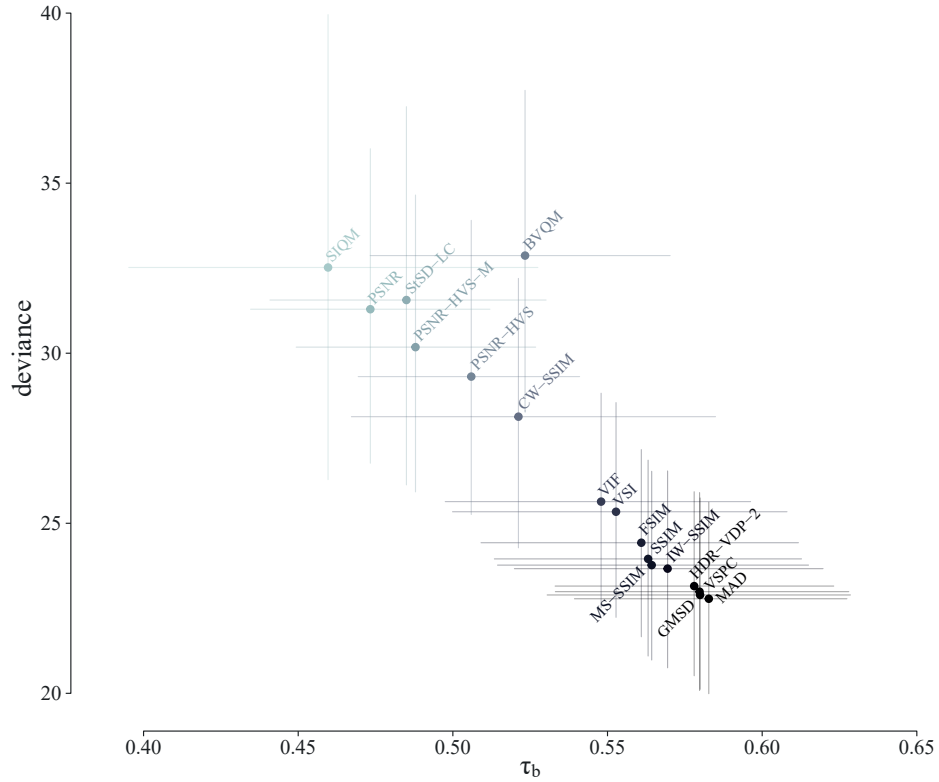


**Figure 1.9** The average metric performance (markers) across observers. The horizontal and vertical axes depict two different measures of metric performance. If a metric captures the perceived degradation intensity, we expect human performance to improve when the metric indicates the quality has degraded. The horizontal axis shows the average Kendall $\tau_b$ correlation between the metric scores and the individual observer responses (higher values indicate a closer agreement; see an example in Figure A.6). Similarly, if a metric captures human near-threshold performance, we would expect to fit a psychometric function predicting the observer responses as a function of the metric scores. The vertical axis shows the average fitted psychometric function deviance on the individual data (lower values indicate a better fit; see an example in Figure 1.10). The error bars depict the 95% CI by basic bootstrap

## 1.4. Discussion

In this chapter, we aimed to define the perceptually optimal discrete angular view density for continuous-parallax content; that is, to specify an upper limit for the visibility and acceptance of the corresponding distortions in the angular domain for a human observer. Using a passive stereoscopic display, we simulated the related artifacts in 3-D videos replicating the changing views of a synthetic object for an observer translating laterally in both directions parallel to the scene (a "look-around" effect). We then performed a psychophysical study to measure the tolerance of human participants to these artifacts, modulating the angular view density, the object depth range, and the virtual observer's velocity. As the view density limit generally

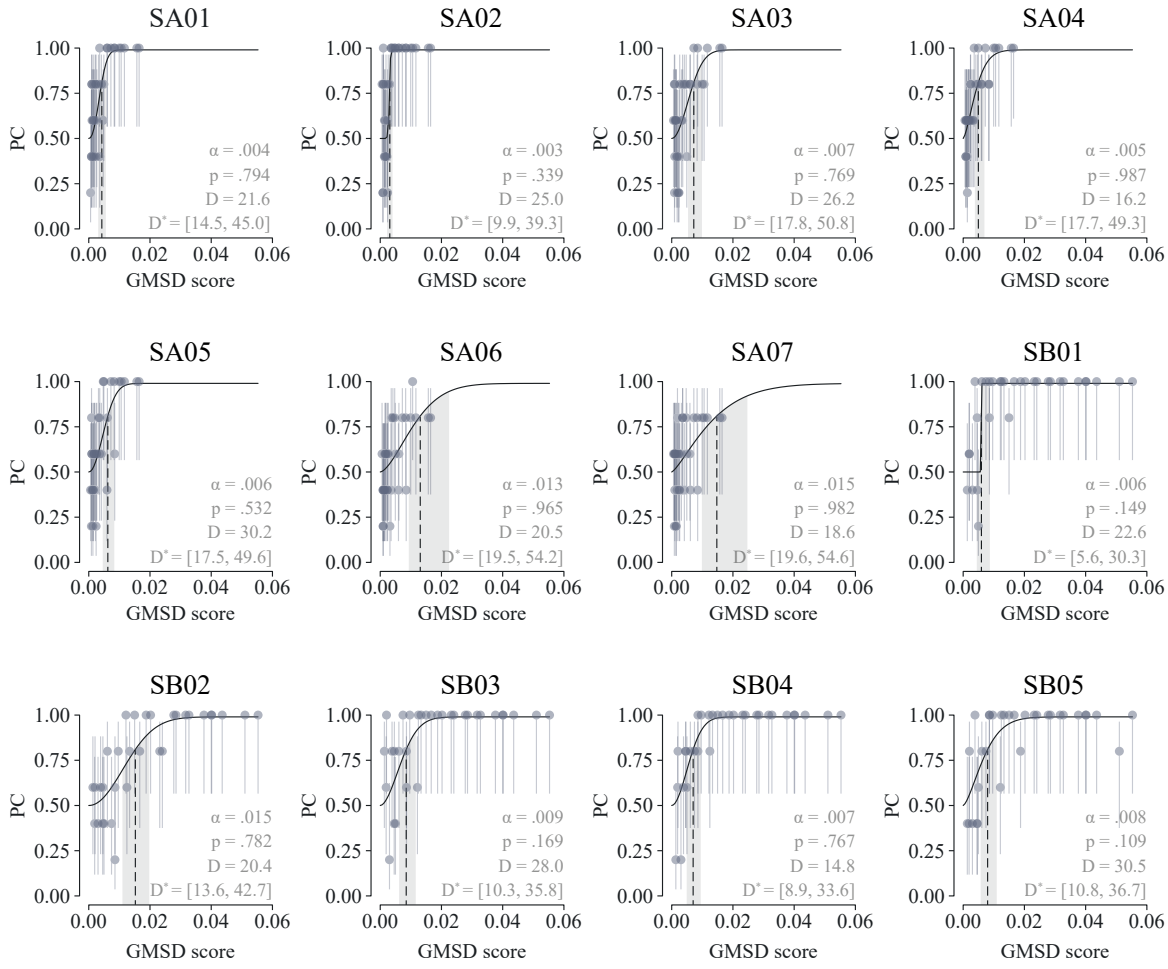


**Figure 1.10** The individual psychometric function fits on the observer performance (markers) using the GMSD score as an approximation of the perceived degradation intensity. The error bars show the Wilson binomial $CI_{95\%}$. The shaded area depicts the $BC_a$-$CI_{95\%}$ for the threshold estimate (dashed line; $\alpha$) by parametric bootstrap. The goodness-of-fit is reported on the bottom right-hand corner (gray text), where D is the observed deviance, $D^*$ the .025 and .975 quantiles of the Monte-Carlo deviance distribution (N = 1e+4 samples), and p the probability of obtaining deviance on the generated samples larger than the observed value. The letters A and B indicate the first and the second group of participants, respectively

depends on the depicted scene, e.g., the depth range, our goal was to establish a general way to estimate these threshold values. After we assessed the distortion ranking similarity between the psychophysics data and a wide range of full-reference quality metrics on the simulated image sequences, we proposed using the metric scores to approximate the perceived degradation intensity on a unidimensional scale and predict near-threshold observer performance. We found this approximation sufficient to describe the observed data, which then allowed us to establish a method to generalize the estimated threshold to different conditions.

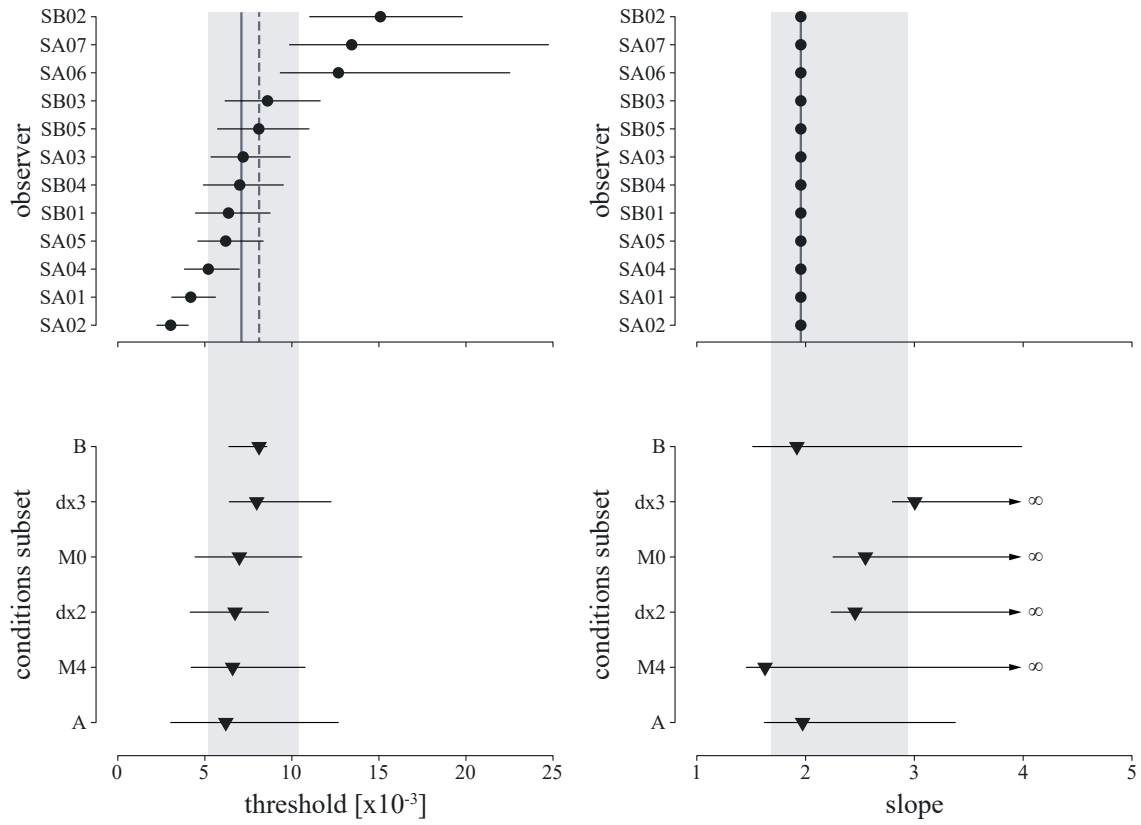


**Figure 1.11 Top**: threshold and slope estimates for each observer under the final model (equal slopes). The vertical solid line shows the median value across observers. The dashed line shows the mean. The shaded area illustrates the $BC_a$-$CI_{95\%}$ for the median threshold (left) and the global slope estimate by parametric bootstrap (PB; right). The error bars depict the PB $BC_a$-$CI_{95\%}$ of the threshold estimates. **Bottom**: the median threshold across observers and the global slope estimate on different subsets of the examined conditions. Explicitly, we tested pooling the data separately for each step size (dx = 2 mm and dx = 3 mm), view blending (M = 0 and M = 4), and depth (A: z = 0 cm and B: z = -15.6 cm). The error bars show the $BC_a$-$CI_{95\%}$ for the median threshold (left) and the global slope estimate by parametric bootstrap (right). The infinity sign indicates a very high upper limit for the slope estimate

In agreement with our results, Adhikarla et al. (2017) found the GMSD and HDR-VDP-2 image quality metrics to perform best on similar distortions in the angular domain (the other ones in our study were the MAD and VSPC). Note that these metrics follow conceptually different approaches extending from a purely bottom-up method modeling the human visual system components (i.e., HDR-VDP) to a simple top-down measure based on the similarity of the image gradients magnitude (i.e., GMSD). Since the latter is also the most parsimonious metric, we can reasonably assume that a central feature in successfully approximating the perceived distortion intensity for the related artifacts generally lies in capturing the local structural and contrast changes. In most cases, the 2-D image metrics outperformed both the

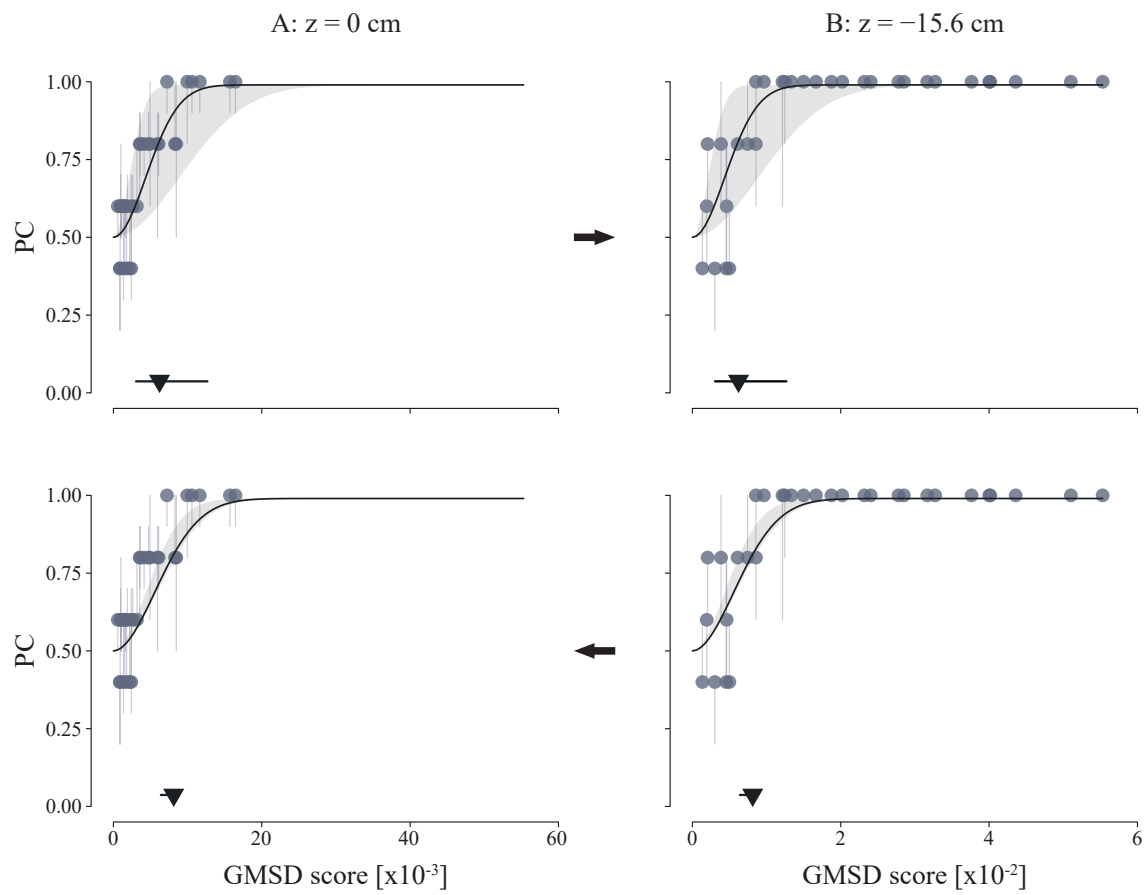


**Figure 1.12** An example of the observer performance prediction across conditions and participants based on a split between the two depth levels (left and right). **Top**: the psychometric function corresponding to the median threshold estimate across observers (solid line; triangle) trained at zero depth (left) and tested at a different depth level and participant group (right). The markers (circles) indicate the mean observer performance. The error bars and the shaded areas show the $BC_a$-$CI_{95\%}$ by bootstrap. **Bottom**: the same procedure in the opposite direction (by switching train and test sets)

stereoscopic and video ones, despite being "blind" to temporal and binocular distortions. If this result was not due to the specific metric selection, we could hypothesize that the corresponding artifacts had either a relatively small effect or were effective well above the view density threshold. When we examined different condition subsets (Appendix A), the stereoscopic and video metrics were most negatively affected at the lowest and the highest depth level, respectively. The former can be more easily explained as the binocular disparities when the object was placed at zero were relatively small and within the stereoscopic comfort zone (Shibata et al., 2011), ranging between .31 and .35 degrees across participants. Presumably, by simulating more excessive binocular disparities or modulating the frame rate instead of the virtual movement step size (or even the step size itself), one could increase the severity of these

effects; thus, balancing the overall metrics performance. In any case, metrics tuned to different types of distortions that can theoretically emerge are still useful in the absence of a dedicated metric as well as if one is interested in suprathreshold distortion scaling. The metric scores can then be potentially fused to improve prediction performance (e.g., Liu et al., 2013).

The optimal view density for the median observer ranged roughly between .9 views/degree to 3.2 views/degree ($M$ = 1.8, $SD$ = .9) over a 36 degrees field of view among the conditions in our synthetic scene. Translating the stimulus in depth increased the relative density estimate by a factor of 2.7 on average across conditions ($SD$ = .3), whereas changing the simulated observer velocity generally produced the same results within the examined range. The absence of view blending (equivalent to a nearest-neighbor interpolation with an added perspective distortion due to the position offset) resulted in a slightly higher view density by a mean factor of 1.3 across conditions ($SD$ = .2), indicating that smoother transitions with moderate crosstalk were preferable to sharp and abrupt parallax shifts. Interestingly, despite the vast differences in the experimental conditions, previous studies investigating the perceived smoothness and naturalness of motion parallax in multiview content reported comparable findings. Speranza et al. (2005) found an average density of approximately 4 views per degree of visual angle for a smooth transition without user interaction. Runde (2000) suggested a minimum limit of 12 views/degree for sampling the eye positions of a moving observer. Takaki et al. (2012) showed that the naturalness of motion parallax improved by increasing the cross-talk among the neighboring views on a super multi-view display (i.e., the view density can be smaller than the pupil diameter), albeit with reduced image quality. Note that our estimates were slightly biased due to discretization effects (we can only approach the threshold at a minimum distance for given scene conditions). From a practical viewpoint, we chose to report the nearest value below the threshold in order to avoid visible distortions; thus, overestimating the density estimate while underestimating some of the differences among the conditions (one condition can be closer to the threshold than the other). In terms of maximum pixel disparity among the input views, the estimated thresholds for the median observer given the discretization restrictions were generally higher than 1 pixel (considered as the limit for a dense sequence), which suggests some observer tolerance up to approximately 5 pixels ($M$ = 3, $SD$ = 1) across conditions. However, the specific density value is of no particular interest as it will naturally vary across different scenes and individuals. Note that the range of the optimal average estimate across participants was .8 to 2.1 views/degree ($M$ = 1.3, $SD$ = .5) and 1.9 to 3.5 views/degree ($M$ = 3.0, $SD$ = .7) for the first and the second group, respectively, which indicates some considerable variability among observers. For practical reasons, due to the relatively small sample size, we followed the more common fit and pool procedure (i.e., estimating the individual thresholds

and then specifying the median or average observer). However, a partial pooling method could potentially provide a preferable solution (see Moscatelli et al., 2012).

The current approach is limited to image sequences where a full-reference is available. For real-world scenes, this is not always possible. Previous attempts using a less distorted sequence as reference were not as successful (Adhikarla et al., 2017). If this is the case, an appropriate reconstruction method could be potentially used to generate a sufficiently dense sequence (e.g., Vagharshakyan et al., 2018). Note that in order to allow a more generic approach and given the available resources, specific display dependencies (e.g., the display type), vertical parallax, and artifacts related to transmission (e.g., compression) or sequence reconstruction were not included. In principle, as long as the displayed images can be simulated, the same procedure applies; however, the metrics' performance could vary. Other extensions, such as non-Lambertian scenes, multiple objects in depth, occlusions, and different conditions for the simulated observer (non-linear paths), could also be investigated. In any case, given that the metric scores can sufficiently approximate the distortion intensity (even if it needs to be re-calibrated or used merely as a prior), the data collection duration can be considerably reduced since it allows us to map the multi-dimensional distortions space to a single variable. Note that this approach is bounded at one just-noticeable difference relative to the reference, which suffices for our purpose. Extending the scale further would also require paired comparisons among the distorted conditions (common in subjective quality assessment), which would allow recovering their relative distance based on a set of assumptions similar to signal detection theory (e.g., case V in Thurstone, 1927; see a practical guide in Perez-Ortiz and Mantiuk, 2017). The metric scores can then be mapped to this extended unified scale similarly.

In summary, this study provides additional evidence that existing state-of-the-art quality metrics can approximate the perceived distortion intensity in simulated full-parallax content even if they are not primarily constructed to capture the related artifacts in the angular domain. Thus, it becomes feasible to describe the observer near-threshold performance as a function of the metric scores and potentially extend it to different conditions (e.g., depth range modulation). Consequently, this procedure allows us to specify the required view density to minimize the visible distortions for the average human observer, generally providing us with a method to optimize or even avoid impractical and extensive subjective evaluations.

# Chapter 2. The human contrast sensitivity function at high luminance

## 2.1. Introduction

The most basic way of characterizing the ability of a human observer to discriminate visual patterns is the contrast sensitivity function (CSF). The CSF reports the sensitivity to visual stimuli as a function of their spatiotemporal frequency. It is an integral part of visual display standardization (e.g., NEMA, 2020) and the central component in HVS-based image quality assessment algorithms for extending performance to higher luminance range (e.g., Daly, 1992; Mantiuk et al., 2011). The latter stems from the fact that one of the primary factors determining the shape of the CSF is adaptation luminance. Generally, an increase in background luminance results in higher peak sensitivity and spatial resolution, while the location of the peak shifts to higher frequencies, additionally changing the CSF shape from low-pass to band-pass (Patel, 1966; Van Nes and Bouman, 1967; van Meeteren and Vos, 1972; De Valois et al., 1974).

The relationship between contrast sensitivity and adaptation luminance is often described as a trilinear transition (Graham, 1989) with each segment corresponding to the dominant noise source that limits visual detection, i.e., early noise or "dark light" (Barlow, 1964), photon shot noise (Rose, 1942; De Vries, 1943), and late neural noise (Pelli, 1990). This theoretical construct is usually referred to as linear to DeVries-Rose to Weber transition, with slopes equal to 1, .5, and 0 in log-log space for each segment, respectively (an example can be found in Figure 2.1). In the case of sine-wave gratings, most commonly employed in CSF measurements, sensitivity was found to demonstrate asymptotic behavior, i.e., to approach the Weber region, at higher luminance for increasing spatial frequency (Van Nes and Bouman, 1967; Hess and Howell, 1988; Rovamo et al., 1994). Several studies provide evidence for a further decreasing, albeit neglected, region for low to intermediate ($\sim$ 8 cpd) spatial frequencies (Depalma and Lowry, 1962; Daitch and Green, 1969; Kelly, 1972; De Valois et al., 1974; Rovamo et al., 1995; Peli et al., 1996; Kim et al., 2013; Silvestre et al., 2018; Bierings et al., 2019). This interesting phenomenon of a decrease in sensitivity with increasing luminance was briefly discussed in García-Pérez and Peli (1997) and Rovamo et al. (1997).

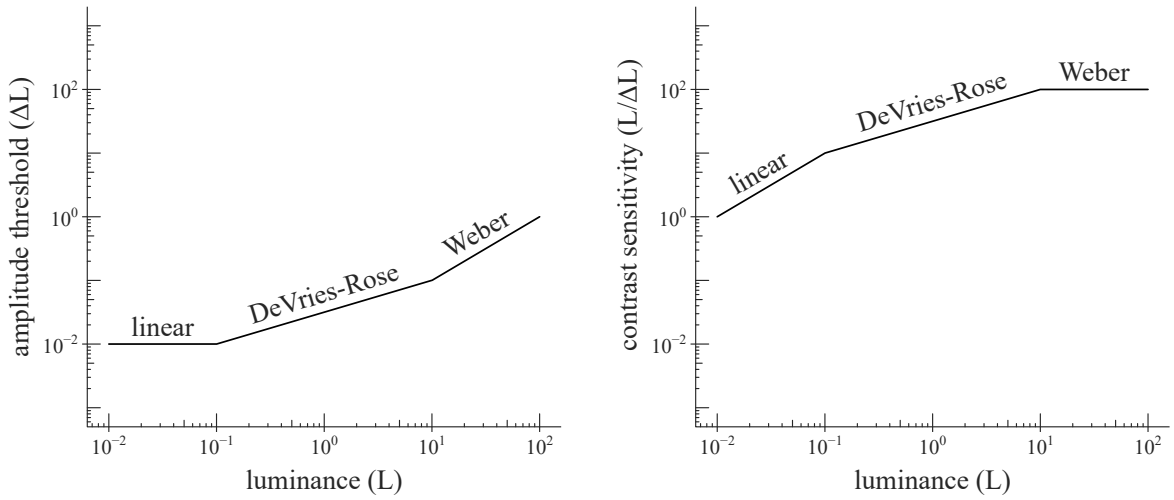


**Figure 2.1** The three linear segments in log-log scale corresponding to the dominant noise source that limits visual detection. At low light levels (linear region), the amplitude threshold is independent of luminance ($\Delta L$ = constant), and contrast sensitivity becomes proportional to the background luminance ($L/\Delta L \sim L$). At intermediate light levels (DeVries-Rose region), the amplitude threshold is proportional to the square root of luminance ($\Delta L \sim \sqrt{L}$), and thus sensitivity is also proportional to the square root of luminance. At higher light levels (Weber region), the amplitude becomes proportional to the background luminance ($\Delta L \sim L$), and thus sensitivity is independent of luminance

The human CSF has been measured extensively over the years. However, only a few studies have systematically investigated relatively high luminance levels with sufficiently broad and dense frequency sampling, and they were limited to one or only a few observers (Depalma and Lowry, 1962; Van Nes and Bouman, 1967; Rovamo et al., 1994; Silvestre et al., 2018). The CSF governs artifact visibility, and as new technology is producing higher luminance displays, it becomes essential to understand how this function behaves in this regime. Here, our purpose was two-fold: first, to assess the asymptotic behavior of the CSF; second, to investigate the validity range of the existing CSF models and provide an additional calibration dataset.

Using a custom-built display system inspired by previous designs on high dynamic range displays (Seetzen et al., 2004), we measured and analyzed the foveal spatial CSF of eleven participants up to a mean background luminance of 1078 cd/m$^2$ at spatial frequencies ranging from 1.5 cpd to 30 cpd. Higher luminance values were then examined for two of the observers through artificial pupil dilation up to an average equivalent display luminance of 8066 cd/m$^2$ (based on a 2 mm pupil diameter). Our dataset was further supplemented by extracting a large number of foveal CSF measurements with varying luminance levels from the literature. We investigated the overall effect of adaptation luminance on contrast sensitivity for sine-wave gratings and examined the validity of the linear to DeVries-Rose to Weber region transition. Based on this theoretical construct, we adopted a CSF model consisting of central elements in

the human visual signal processing and three limiting internal noise components corresponding to each region. We subsequently assessed the model's performance on the measured contrast sensitivities and proposed an eight-parameter form to describe the contrast sensitivity surface in the spatial frequency-luminance domain.

## 2.2. Measurements of the human CSF at high luminance

### *2.2.1. Participants*

Eleven healthy volunteers (mean age: 28.6 years, SD: 4.9 years, span: 20 – 38, three males and eight females) participated in the experiment. The participants had normal or corrected to normal vision with no history of visual problems. The volunteers were screened for normal visual acuity using a Sloan standardized letter chart. The Newcastle University Ethics Committee approved the study, and all participants gave written informed consent.

### *2.2.2. Apparatus and experimental design*

#### *2.2.2.1. Display setup*

The task was performed on a custom-built liquid-crystal display (LCD) system. An 8-bit LCD panel with 1920×1200 pixels resolution and .27 mm pixel pitch was removed from its monitor housing (Dell U2412M), and the backlight source was replaced with a 3LCD projector rated at 2200 lumens (Epson EH-TW6000; used solely as a light source). The projected light was collimated through a Fresnel lens sheet and then redistributed through a diffuser before falling on the back of the panel (a schematic can be found in Figure 2.2). This design achieved a peak luminance of approximately 2160 cd/m$^2$ with a contrast ratio of 860:1 for a 9 cm height viewport. The monitor was calibrated using a Minolta LS-100 luminance meter, and bit-depth was increased to 10 bits using a bit-stealing method as implemented in the Psychophysics Toolbox version 3 (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007). The luminance fall-off from the center of the viewing area to the edges was measured at eight points on each of two concentric circles of 2.25 cm and 4.50 cm radii across the horizontal, vertical, and diagonal directions. The mean values of the relative decrease and their SD were 3.95 ± 2.42% and 6.90 ± 3.48%, respectively. The CIE chromaticity of the background was [x, y] = (.34, .37). The subjects were seated in a dark room at a 280 cm viewing distance from the screen center to their eyes, with their head placed on a chinrest. The display setup was covered and positioned in a second adjacent room that was occluded by two side curtains to minimize stray light. The viewing area was isolated by placing black cardboard with a square aperture equal to the

viewport's height directly in front of the panel. The decrease in mean luminance was achieved using three circular neutral density filters (77 mm) mounted on an optical post assembly placed directly in front of the participant's eyes, with optical densities of .4 and .7 (Shenzhen Neewer Technology Co., Ltd., China), and 1.2 (Hoya, Kenko Tokina Co., Ltd., Japan). The measured shifts in chromaticity were below .01. In order to avoid confounds due to changes in the pupil size as a function of retinal illuminance, the stimuli were viewed monocularly through a 2 mm iris diaphragm attached to a fully adjustable optical trial frame and with the left eye occluded. The diaphragm was positioned at the closest possible distance to the cornea.

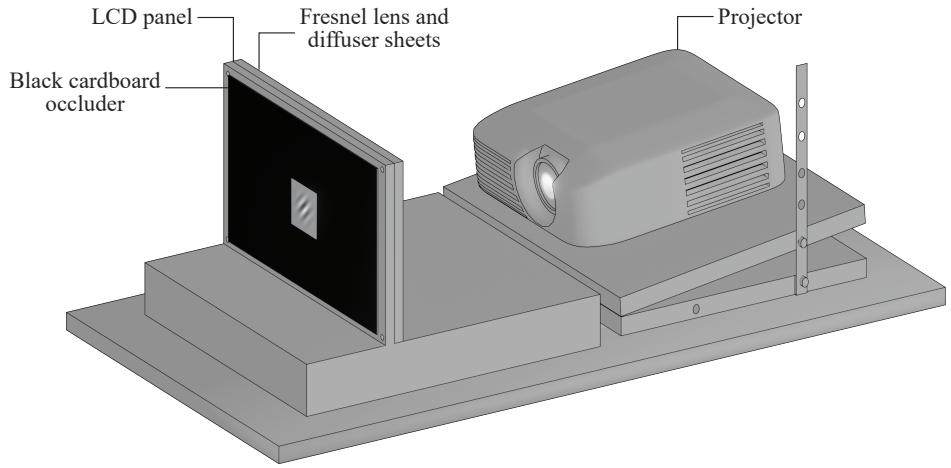


**Figure 2.2** The projector-based display setup schematic (top cover not shown)

#### *2.2.2.2. Stimuli*

The stimuli consisted of static sinusoidal gratings with a 2-D Gaussian envelope, oriented at $\pm$45 degrees from the vertical direction. The images were drawn on a 332x332 pixel grid encompassing a visual angle of 1.8 degrees with the envelope's standard deviation at .3 degrees. The mean background luminance was set to 1078 cd/m$^2$ when viewed with no neutral density filter. A subset of the stimuli is shown in Figure 2.3.

#### *2.2.2.3. Experiment I*

The subjects reported the orientation of a single grating stimulus by a keypress. The orientation ($\pm$45 degrees) and phase of the gratings were randomized. The stimulus presentation time was not limited, and no feedback was provided. A fixation cross was displayed at the center of the window for 500 ms preceding each trial. We examined four luminance conditions, explicitly,

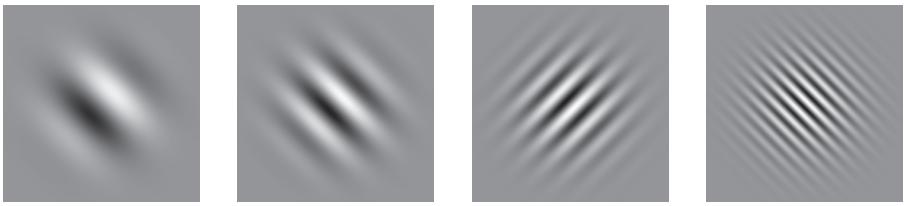

**Figure 2.3** A subset of the stimuli from low (left; 1.5 cpd) to high (right; 9.1 cpd) frequency

1078 cd/m$^2$, 475 cd/m$^2$, 206 cd/m$^2$ and 68 cd/m$^2$ at six spatial frequencies spaced $\log_{10}$-linearly from approximately 1.5 to 30 cpd. Each subject performed four sets of 30 consecutive trials for each spatial frequency and luminance condition at varying contrast levels for a total of 720 trials per set. The frequencies were randomly interleaved within each luminance condition, and the luminance testing order was randomized for each set of trials. The thresholds were estimated using the Ψ Bayesian method with uniform priors as described in Kontsevich and Tyler (1999) and implemented in the Palamedes toolbox (Prins and Kingdom, 2018). The lapse rate was included in the posterior, but the slope and the lapse rate were marginalized in the stimulus level selection (Prins, 2013). The participants completed the experiment in two sessions, each lasting approximately 70-90 min. A training run of 90 or 450 trials based on the participant's experience was performed to familiarize the subjects with the procedure.

#### *2.2.2.4. Experiment II*

In a follow-up experiment, we extended the luminance range above and below the tested levels for two of the participants in experiment I. Higher luminance levels were examined in terms of equivalent retinal illuminance [Td], calculated as the display luminance times the circular pupil area, by artificially dilating the observer's right pupil using eye drops of Cyclopentolate at 1%. The pupil diameter was monitored at regular intervals prior to the session through a photorefractor (PowerRef, Plusoptix GmbH, Nuremberg, Germany) until it was stabilized at 7.4 mm and 8.5 mm for the first (S01) and the second observer (S07), respectively. We measured three retinal illuminance levels with the corresponding display luminance set at 140 cd/m$^2$, 475 cd/m$^2$, and 1078 cd/m$^2$. The lower level was selected based on the available neutral density filters to approximate the maximum retinal illuminance in experiment I, with the purpose of combining the two sets of sensitivity measurements. The matched levels were then used as a step to adjust for the changes in optical aberrations due to the increased pupil size and the lack of accommodation. The selected filter with an optical density of .9 (Hoya, Kenko Tokina Co., Ltd., Japan) displaced the retinal illuminance step point by -7.4% (235 Td) and 1.7% (54

Td) relative to experiment I, for the first and the second observer, respectively. The retinal illuminance values were corrected for the Stiles-Crawford effect (Stiles et al., 1933) by a factor (SC) relative to the pupil diameter (d) as given in Atchison and Smith (2000); equation 2.1, with a $\beta$ coefficient of .12 (Applegate and Lakshminarayanan, 1993).

$$SC(d_{mm}) = \frac{4\left[1-\exp\left(-\beta d_{mm}^2/4\right)\right]}{\beta d_{mm}^2} \tag{2.1}$$

The resulting retinal illuminance levels for the first observer were approximately $2.96 \times 10^3$ Td, $1.00 \times 10^4$ Td, and $2.28 \times 10^4$ Td corresponding to roughly a display luminance of 999 cd/m$^2$, 3388 cd/m$^2$, and 7689 cd/m$^2$ given a 2 mm pupil diameter. Likewise, the levels for the second observer were $3.25 \times 10^3$ Td, $1.10 \times 10^4$ Td, and $2.50 \times 10^4$ Td corresponding to 1096 cd/m$^2$, 3720 cd/m$^2$, and 8442 cd/m$^2$. On the lower end, we examined two additional luminance conditions at 3 cd/m$^2$ and 17 cd/m$^2$ with the artificial 2 mm pupil without cycloplegia by using two neutral density filters with densities of 2.6 and 1.5 (Hoya, Kenko Tokina Co., Ltd., Japan), respectively. The participants completed the experiment in two dedicated sessions for each of the low and high illuminance conditions set, following the same procedure as in the first experiment. For the low luminance sessions, we included an additional adaptation time of five minutes in the dark before the task started.

### *2.2.3. CSF characteristics*

The CSF represents the sensitivity, the reciprocal of the contrast threshold estimate, as a function of the grating stimulus frequency. The shape of the CSF can be generally described by four characteristics: the peak sensitivity ($G_{max}$), the frequency of the peak ($F_{max}$), the grating visibility area under the log-curve (*AULCSF*), and the high-frequency cut-off ($F_c$), i.e., the frequency where sensitivity asymptotes to zero (Figure 2.4). A sensitivity plateau may also be seen at very low frequencies (roughly below 1 cpd); however, it was outside our sampling range. Watson and Ahumada (2005) reviewed several functional forms that can adequately characterize the CSF curve, with only small differences among them in terms of fitting performance. Here, considering mainly the interpretability of the parameters, we adopted an asymmetric form of the log-parabola model (Chung and Legge, 2016) to extract these characteristics. This descriptive model (equation 2.2) uses four parameters to express the log sensitivity $S$ as a function of the log frequency $u$: the location $F_{max}$ and magnitude $G_{max}$ of the peak, and the weights $w_L$ and $w_R$ for the curvature of the left and right branches, respectively.

$$S(u) = \begin{cases} G_{max} - w_L^2 \left[u - log_{10}\left(F_{max}\right)\right]^2, & u < log_{10}\left(F_{max}\right) \\ G_{max} - w_R^2 \left[u - log_{10}\left(F_{max}\right)\right]^2, & u \geq log_{10}\left(F_{max}\right) \end{cases} \tag{2.2}$$

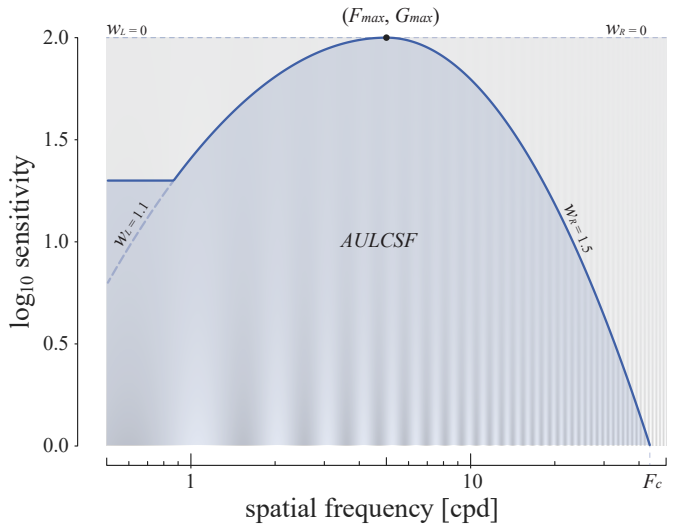


**Figure 2.4** Typical human CSF shape depicted as an asymmetric log-parabola truncated at low frequencies. $F_c$: the highest visible frequency. $F_{max}$ and $G_{max}$: the location and magnitude of the peak. AULCSF: the area under the log curve (shaded area). $w_L$ and $w_R$: the parameters that control the curvature of the left and right branches. The truncation parameter was omitted as it was outside our sampling range. On the background, a Campbell-Robson contrast sensitivity chart demonstrating the CSF shape with a sine-wave increasing in spatial frequency (rightwards) and decreasing in luminance contrast (upwards)

### *2.2.4. Statistical analysis*

A two-way within-subjects ANOVA was conducted on the mean log-sensitivity thresholds in experiment I, with the spatial frequency and luminance as independent predictors. The tests (type III) were performed using the R libraries afex (Singmann et al., 2020) and car (Fox and Weisberg, 2019). The pairwise comparisons among all luminance pairs at each spatial frequency were carried out using the library emmeans (Lenth, 2020) with a multivariate model. The simple effects are reported with the Geisser-Greenhouse's $\hat{\varepsilon}$ adjustment for the degrees of freedom. Likewise, we tested each of the CSF characteristics described above, i.e., the location and magnitude of the peak, the area under the log CSF, and the high-frequency cut-off, using a one-way within-subjects ANOVA with the luminance as an independent predictor.

### 2.2.5. Results

#### 2.2.5.1. Sensitivity estimates

Figure 2.5 illustrates the mean sensitivity estimates for the individual CSFs under the examined luminance conditions. Figure 2.6 shows a summary of the average threshold estimates across observers, along with the significance of the luminance differences. As expected, there was a significant interaction between spatial frequency and luminance, F(15, 150) = 26.55, p < .001, $\eta_p^2$ = .726. Further analysis indicated a significant difference among the luminance conditions for spatial frequencies of 1.5 cpd, 9.1 cpd, 16.5 cpd, and 30.2 cpd (Table 2.2). However, there was no significant effect in our sample on the intermediate frequencies of 2.7 cpd and 5.0 cpd, where sensitivity generally saturated in the examined range. The mean estimates are given in Table 2.1. Sensitivity increased asymptotically with luminance for spatial frequencies of 9.1 cpd and above. These increments were significant for all luminance levels compared with the 68 cd/m$^2$ (given the multiplicity adjustment; Table 2.3). At the highest frequency of 30.2 cpd, significant gains extended above 208 cd/m$^2$ ($\approx$ 616 Td) before sensitivity began to plateau at 475 cd/m$^2$ ($\approx$ 1406 Td). In contrast, at the lowest frequency of 1.5 cpd, the effect was reversed, with sensitivity decreasing significantly at 475 cd/m$^2$ and 1078 cd/m$^2$ ($\approx$ 3191 Td) compared with the 68 cd/m$^2$ ($\approx$ 201 Td). Note that our sample size limits the minimum effect size that we can reliably detect. As the sensitivity generally increased asymptotically in the examined range, increasing the sample size further would result in detecting smaller and smaller effects (excluding the cases where a decrease was found instead). Whether these increments are meaningful in practice also depends on the specific application. The statistical sensitivity is given in Appendix B. Regardless, generalization should be primarily based on the physical model described in the following section.

#### 2.2.5.2. CSF characteristics

Figure 2.7 depicts the main CSF characteristics at each luminance level, as estimated through the best-fitting descriptive model parameters. Naturally, some of these variables are associated (Table B.3); however, the specific luminance effects were of primary interest. The model performance in terms of the RMS error for each CSF curve is summarized in Table B.2. The mean estimates are presented in Table 2.4. The examined luminance range elicited significant changes in the location ($F_{max}$), F(2.74, 27.38) = 3.65, p = .028, $\eta_p^2$ = .267, and the magnitude ($G_{max}$), F(2.39, 23.90) = 13.55, p $<$ .001, $\eta_p^2$ = .575, of the peak sensitivity. Likewise, there was a significant effect of luminance on the extrapolated high-frequency cut-off ($F_c$), F(2.56, 25.61) = 37.20, p $<$ .001, $\eta_p^2$ = .788, and the area under the log curve (AULCSF), F(2.63,

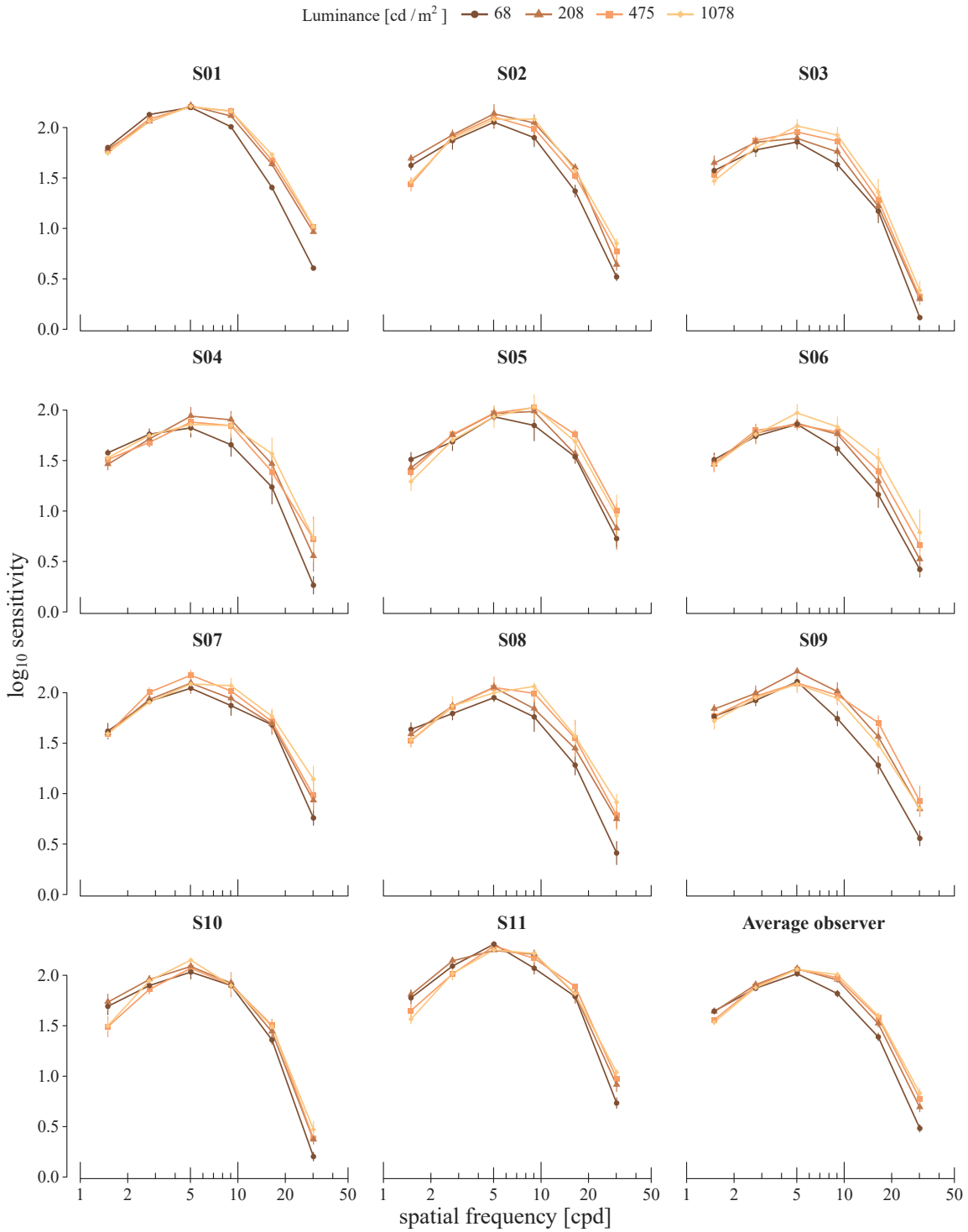


**Figure 2.5** The estimated CSF for each observer and condition. The contrast sensitivity thresholds (vertical axis) at each spatial frequency (horizontal axis) and background luminance condition (color gradient; darker for lower luminance). The markers depict the mean log estimates. The error bars show the standard error. The bottom right facet shows the average CSF across observers

| Spatial frequency | Luminance [cd/m$^2$] | Estimated mean [dB] | SE | df | Lower $CI_{95\%}$ | Upper $CI_{95\%}$ |
|---|---|---|---|---|---|---|
| 1.5 cpd | 68 | 32.88 | .62 | 10 | 31.50 | 34.27 |
| | 208 | 32.79 | .88 | 10 | 30.82 | 34.75 |
| | 475 | 31.14 | .76 | 10 | 29.45 | 32.83 |
| | 1078 | 30.68 | .76 | 10 | 28.99 | 32.37 |
| 2.7 cpd | 68 | 37.43 | .85 | 10 | 35.54 | 39.93 |
| | 208 | 38.13 | .80 | 10 | 36.36 | 39.90 |
| | 475 | 37.84 | .72 | 10 | 36.22 | 39.45 |
| | 1078 | 37.55 | .69 | 10 | 36.01 | 39.09 |
| 5.0 cpd | 68 | 40.32 | .91 | 10 | 38.29 | 42.35 |
| | 208 | 41.30 | .80 | 10 | 39.51 | 43.08 |
| | 475 | 41.19 | .81 | 10 | 39.37 | 43.00 |
| | 1078 | 41.17 | .71 | 10 | 39.58 | 42.76 |
| 9.1 cpd | 68 | 36.37 | .91 | 10 | 34.35 | 38.40 |
| | 208 | 39.08 | .84 | 10 | 37.20 | 40.96 |
| | 475 | 39.52 | .73 | 10 | 37.88 | 41.16 |
| | 1078 | 40.14 | .76 | 10 | 38.44 | 41.83 |
| 16.5 cpd | 68 | 27.78 | 1.23 | 10 | 25.05 | 30.51 |
| | 208 | 30.41 | 1.04 | 10 | 28.10 | 32.72 |
| | 475 | 31.59 | 1.11 | 10 | 29.11 | 34.08 |
| | 1078 | 31.96 | .83 | 10 | 30.10 | 33.82 |
| 30.2 cpd | 68 | 9.66 | 1.33 | 10 | 6.69 | 12.64 |
| | 208 | 13.81 | 1.40 | 10 | 10.70 | 16.92 |
| | 475 | 15.53 | 1.47 | 10 | 12.27 | 18.80 |
| | 1078 | 16.63 | 1.40 | 10 | 13.51 | 19.74 |

**Table 2.1** The mean sensitivity thresholds in dB for each luminance condition and spatial frequency

| Spatial frequency | df | MSE [dB] | F | $\eta_p^2$ | [90% CI] | p-val | $\text{Holm}_6$ p-val |
|---|---|---|---|---|---|---|---|
| 1.5 cpd | (1.63, 16.29) | 2.29 | 11.23 | .529 | [.183, .671] | .001 | **.004** |
| 2.7 cpd | (2.48, 24.80) | .69 | 1.87 | .157 | [.000, .313] | .168 | .168 |
| 5.0 cpd | (2.45, 24.48) | .90 | 3.07 | .235 | [.000, .395] | .056 | .111 |
| 9.1 cpd | (2.09, 20.89) | 1.49 | 29.24 | .745 | [.519, .815] | <.001 | **<.001** |
| 16.5 cpd | (2.73, 27.25) | 2.15 | 20.15 | .668 | [.435, .748] | <.001 | **<.001** |
| 30.2 cpd | (2.55, 25.55) | 1.56 | 77.34 | .886 | [.785, .914] | <.001 | **<.001** |

**Table 2.2** Statistical tests on the effect of luminance at each spatial frequency level. The last two columns show the unadjusted and the Holm corrected p-values for the six tests (the values we used). The reported mean square error was calculated using the sensitivity in dB (the log multiplied by 20). The method for calculating the confidence intervals can be found in Smithson (2001)

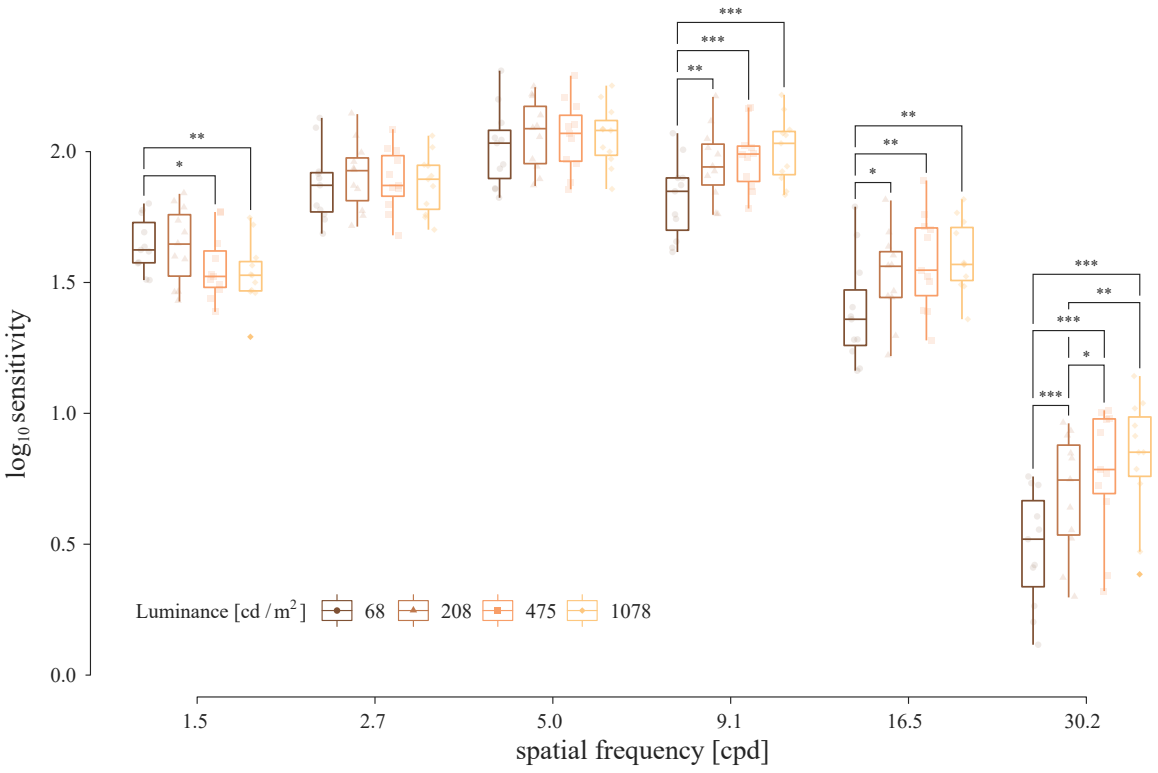


**Figure 2.6** Boxplot of the average log sensitivity estimates for each observer (horizontally jittered markers) per spatial frequency (horizontal axis) and luminance condition (darker gradient for lower luminance). The stars indicate the significant differences (*** $p < .001$, ** $p < .01$, * $p < .05$)

| Spatial frequency | Luminance difference | Estimate [dB] | SE | df | t-ratio | p-val | $\text{Holm}_6$ p-val | $\text{Holm}_{24}$ p-val |
|---|---|---|---|---|---|---|---|---|
| 1.5 cpd | 208 − 68 | -.10 | .40 | 10 | -.25 | .809 | .809 | 1 |
| | 475 − 68 | -1.74 | .41 | 10 | -4.27 | .002 | .008 | **.020** |
| | 1078 − 68 | -2.20 | .44 | 10 | -5.01 | <.001 | .003 | **.008** |
| | 475 − 208 | -1.64 | .60 | 10 | -2.72 | .022 | .065 | .194 |
| | 1078 − 208 | -2.11 | .64 | 10 | -3.28 | .008 | .033 | .091 |
| | 1078 − 475 | -0.46 | .25 | 10 | -1.82 | .099 | .198 | .410 |
| 9.1 cpd | 208 − 68 | 2.71 | .43 | 10 | 6.26 | <.001 | <.001 | **.002** |
| | 475 − 68 | 3.15 | .42 | 10 | 7.43 | <.001 | <.001 | <**.001** |
| | 1078 − 68 | 3.77 | .48 | 10 | 7.83 | <.001 | <.001 | <**.001** |
| | 475 − 208 | .44 | .42 | 10 | 1.04 | .323 | .323 | .968 |
| | 1078 − 208 | 1.06 | .55 | 10 | 1.93 | .082 | .164 | .410 |
| | 1078 − 475 | .62 | .24 | 10 | 2.61 | .026 | .078 | .208 |
| 16.5 cpd | 208 − 68 | 2.63 | .60 | 10 | 4.38 | .001 | .006 | **.018** |
| | 475 − 68 | 3.81 | .64 | 10 | 5.98 | <.001 | <.001 | **.002** |
| | 1078 − 68 | 4.18 | .65 | 10 | 6.46 | <.001 | <.001 | **.001** |
| | 475 − 208 | 1.19 | .49 | 10 | 2.40 | .038 | .075 | .246 |
| | 1078 − 208 | 1.55 | .51 | 10 | 3.01 | .013 | .039 | .131 |
| | 1078 − 475 | .36 | .66 | 10 | .55 | .592 | .592 | 1 |
| 30.2 cpd | 208 − 68 | 4.14 | .56 | 10 | 7.37 | <.001 | <.001 | <**.001** |
| | 475 − 68 | 5.87 | .56 | 10 | 10.50 | <.001 | <.001 | <**.001** |
| | 1078 − 68 | 6.96 | .53 | 10 | 13.26 | <.001 | <.001 | <**.001** |
| | 475 − 208 | 1.72 | .36 | 10 | 4.75 | <.001 | .002 | **.011** |
| | 1078 − 208 | 2.82 | .46 | 10 | 6.12 | <.001 | <.001 | **.002** |
| | 1078 − 475 | 1.10 | .45 | 10 | 2.43 | .035 | .035 | .246 |

**Table 2.3** Pairwise comparisons of the sensitivity estimates among the luminance conditions at each spatial frequency level. The last three columns provide the unadjusted and the Holm corrected p-values for the six and the total twenty four tests (the adjustment we selected) across spatial frequencies

26.27) = 29.88, $p < .001$, $\eta_p^2 = .749$ (Table 2.5). Pairwise comparisons revealed a significantly lower peak and a narrower visibility area at 68 cd/m$^2$, which then saturated in our sample above 208 cd/m$^2$. The highest visible frequency increased significantly with luminance up to 475 cd/m$^2$ before reaching a nearly asymptotic state. Likewise, the location of the peak shifted to a higher spatial frequency with increasing luminance; however, the pairwise differences were not significant after adjusting for the multiplicity (Table 2.6).

| CSF measure | Luminance [cd/m$^2$] | Estimated mean | SE | df | Lower $CI_{95\%}$ | Upper $CI_{95\%}$ |
|---|---|---|---|---|---|---|
| $G_{max}$ [dB] | 68 | 39.78 | .89 | 10 | 37.79 | 41.77 |
| | 208 | 41.12 | .76 | 10 | 39.43 | 42.81 |
| | 475 | 41.28 | .76 | 10 | 39.59 | 42.97 |
| | 1078 | 41.39 | .68 | 10 | 39.87 | 42.92 |
| $F_{max}$ | 68 | 5.59 | .33 | 10 | 4.87 | 6.31 |
| | 208 | 5.96 | .25 | 10 | 5.41 | 6.52 |
| | 475 | 6.26 | .30 | 10 | 5.60 | 6.92 |
| | 1078 | 6.46 | .32 | 10 | 5.76 | 7.17 |
| $F_c$ | 68 | 39.40 | 1.37 | 10 | 36.35 | 42.45 |
| | 208 | 44.53 | 1.98 | 10 | 40.12 | 48.94 |
| | 475 | 46.92 | 1.98 | 10 | 42.51 | 51.33 |
| | 1078 | 48.40 | 2.12 | 10 | 43.66 | 53.13 |
| *AULCSF* | 68 | 43.14 | 1.10 | 10 | 40.68 | 45.60 |
| | 208 | 45.50 | 1.06 | 10 | 43.14 | 47.86 |
| | 475 | 45.83 | 1.00 | 10 | 43.59 | 48.07 |
| | 1078 | 46.10 | .84 | 10 | 44.23 | 47.98 |

**Table 2.4** Mean estimates of the CSF characteristics at each luminance level. $G_{max}$: the peak sensitivity. $F_{max}$: the location of the peak. $F_c$: the highest visible frequency (extrapolated). AULCSF: the area under the log curve reported using the sensitivity in dB (within the measured frequency range)

#### *2.2.5.3. Extended illuminance range*

Figure 2.8 illustrates the adjusted sensitivity at each spatial frequency for the extended retinal illuminance range in the second experiment. Qualitatively, this further increase in illuminance did not produce any noticeable differences. The mean log sensitivity generally plateaued at all spatial frequencies for both observers in the examined range. The gain of increase (Figure 2.9),

| CSF measure | df | MSE | F | $\eta_p^2$ | [90% CI] | p-val | Holm$_4$ p-val |
|---|---|---|---|---|---|---|---|
| $G_{max}$ [dB] | (2.39, 23.90) | .57 | 13.55 | .575 | [.291, .683] | <.001 | <**.001** |
| $F_{max}$ | (2.74, 27.38) | .48 | 3.65 | .267 | [.018, .416] | .028 | <**.028** |
| $F_c$ | (2.56, 25.61) | 5.39 | 37.20 | .788 | [.615, .841] | <.001 | <**.001** |
| *AULCSF* | (2.63, 26.27) | .78 | 29.88 | .749 | [.554, .811] | <.001 | <**.001** |

**Table 2.5** Statistical tests on the effect of luminance on each of the main CSF characteristics. The last two columns show the unadjusted and the Holm corrected p-values for the four tests (the values we used). The reported mean square error for AULCSF was calculated using the sensitivity in dB

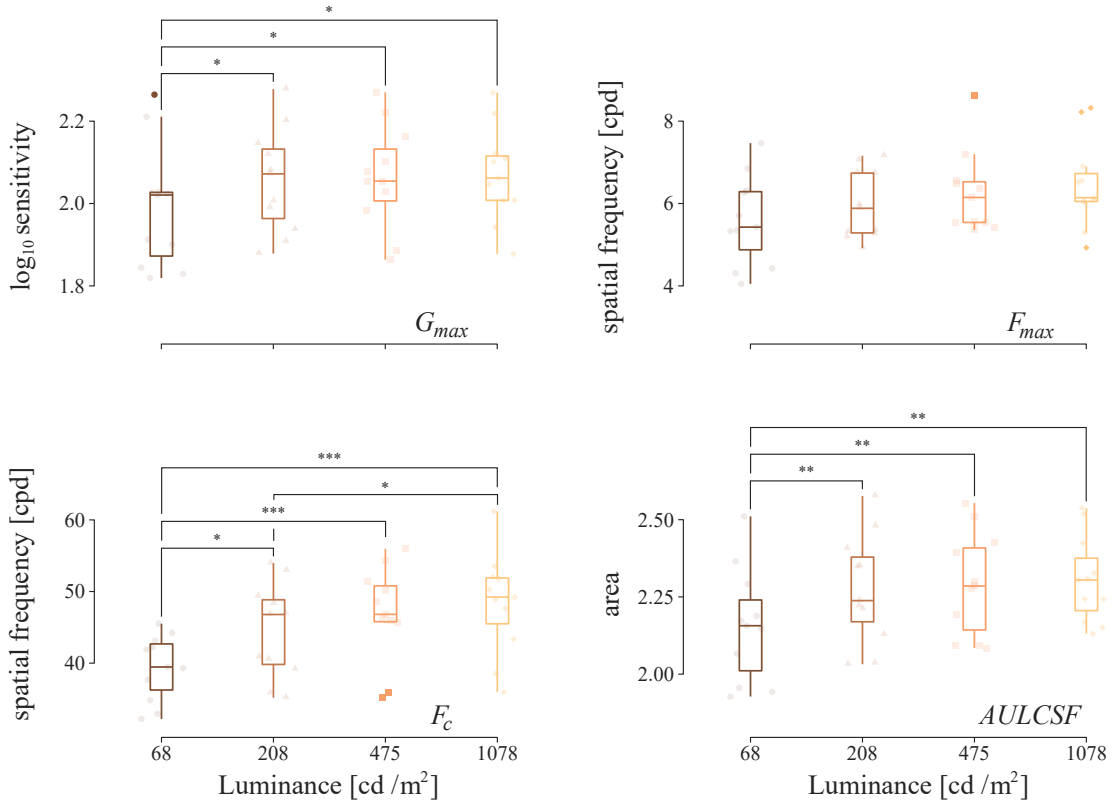


**Figure 2.7** The main CSF characteristics for each luminance condition. **Top left**: the peak sensitivity. **Top right**: the location of the peak. **Bottom left**: the highest visible frequency (extrapolated). **Bottom right**: The area under the log curve (within the measured frequency range). The stars indicate the significant differences among pairs (*** p < .001, ** p < .01, * p < .05)

| CSF measure | Luminance difference | Estimate | SE | df | t-ratio | p-val | $\text{Holm}_6$ p-val | $\text{Holm}_{24}$ p-val |
|---|---|---|---|---|---|---|---|---|
| $G_{max}$ [dB] | 208 − 68 | 1.34 | .25 | 10 | 5.41 | <.001 | .002 | **.030** |
| | 475 − 68 | 1.50 | .32 | 10 | 4.74 | <.001 | .003 | **.047** |
| | 1078 − 68 | 1.61 | .30 | 10 | 5.30 | <.001 | .002 | **.030** |
| | 475 − 208 | .16 | .31 | 10 | .51 | .620 | 1 | 1 |
| | 1078 − 208 | .27 | .32 | 10 | .85 | .413 | 1 | 1 |
| | 1078 − 475 | .11 | .22 | 10 | .53 | .611 | 1 | 1 |
| $F_{max}$ | 208 − 68 | .37 | .25 | 10 | 1.46 | .175 | .524 | 1 |
| | 475 − 68 | .67 | .30 | 10 | 2.26 | .048 | .239 | 1 |
| | 1078 − 68 | .87 | .29 | 10 | 3.04 | .013 | .075 | .908 |
| | 475 − 208 | .30 | .27 | 10 | 1.08 | .307 | .614 | 1 |
| | 1078 − 208 | .50 | .24 | 10 | 2.11 | .061 | .243 | 1 |
| | 1078 − 475 | .21 | .33 | 10 | .63 | .542 | .614 | 1 |
| $F_c$ | 208 − 68 | 5.13 | 1.08 | 10 | 4.73 | <.001 | .003 | **.014** |
| | 475 − 68 | 7.52 | .95 | 10 | 7.94 | <.001 | <.001 | **<.001** |
| | 1078 − 68 | 9.00 | 1.05 | 10 | 8.58 | <.001 | <.001 | **<.001** |
| | 475 − 208 | 2.39 | .82 | 10 | 2.92 | .015 | .031 | .216 |
| | 1078 − 208 | 3.87 | .80 | 10 | 4.83 | <.001 | .003 | **.013** |
| | 1078 − 475 | 1.48 | .73 | 10 | 2.03 | .070 | .070 | .908 |
| *AULCSF* | 208 − 68 | 2.37 | .33 | 10 | 7.24 | <.001 | <.001 | **.002** |
| | 475 − 68 | 2.69 | .35 | 10 | 7.66 | <.001 | <.001 | **.002** |
| | 1078 − 68 | 2.97 | .39 | 10 | 7.64 | <.001 | <.001 | **.002** |
| | 475 − 208 | .33 | .32 | 10 | 1.04 | .325 | .650 | 1 |
| | 1078 − 208 | .60 | .41 | 10 | 1.47 | .173 | .518 | 1 |
| | 1078 − 475 | .27 | .30 | 10 | .90 | .392 | .650 | 1 |

**Table 2.6** Pairwise comparisons of the estimated CSF characteristics among the luminance conditions at each spatial frequency level. $G_{max}$: the peak sensitivity. $F_{max}$: the location of the peak. $F_c$: the highest visible frequency (extrapolated). AULCSF: the area under the log curve reported using the sensitivity in dB. The last three columns provide the unadjusted and the Holm corrected p-values for the six tests and the total twenty four tests (the adjustment we selected)

defined here as the slope of the linear regression through every three consecutive illuminance levels, was higher for increasing spatial frequency and decreasing retinal illuminance. Segments with negative gain, indicating a decrease in sensitivity, were also found at spatial frequencies up to 5.0 cpd. These decrements, although consistent across the two observers, were relatively small with absolute gain values below .1, and appeared to recover with a further increase in illuminance. It is unclear whether this observation was a result of a transition to another state of stability or merely fluctuations around the saturation level.

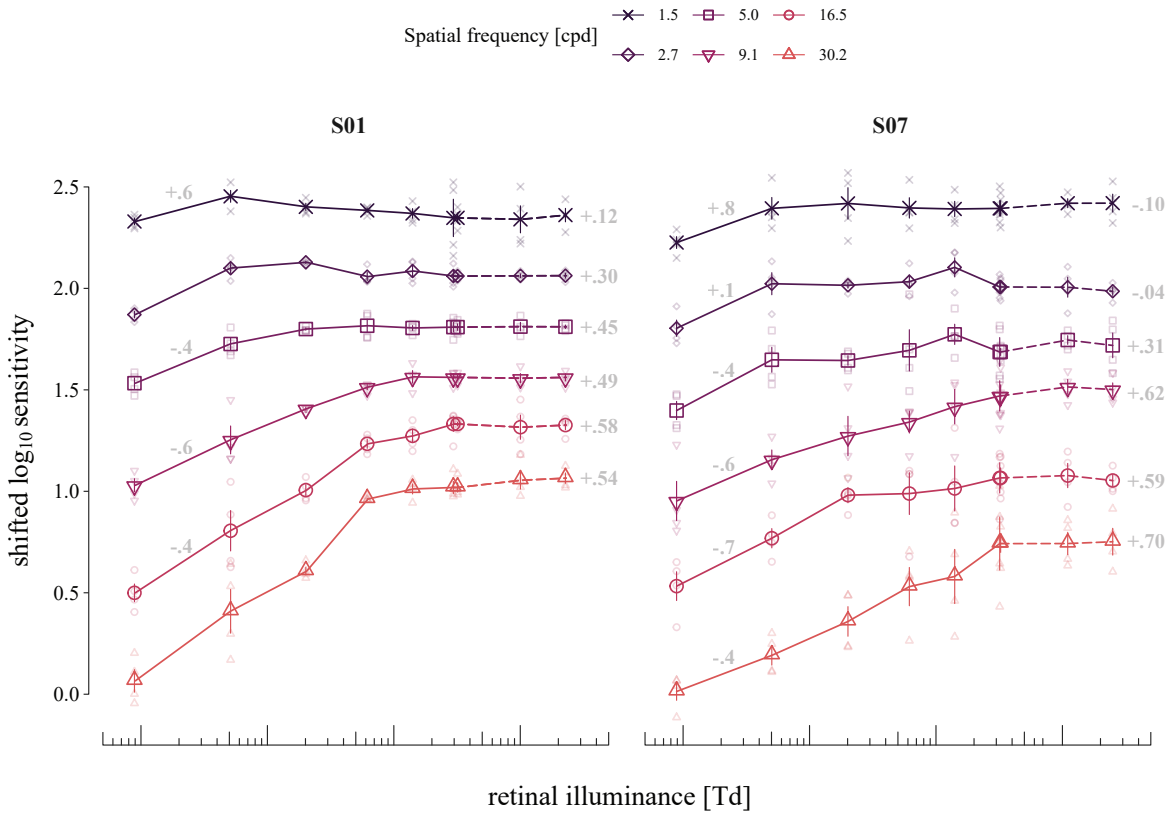


**Figure 2.8** The mean log sensitivity estimates (open markers) as a function of retinal illuminance pooling the data from both experiments. The sensitivities for each spatial frequency (lines) were shifted vertically (light gray numbers on the left side of the curves) from the highest (bottom) to the lowest (top) spatial frequency for visualization purposes. The dotted lines depict the vertically adjusted sensitivities (light gray numbers on the right side) in the extended illuminance range where pupil dilation was used (to line them up with the first experiment). The small markers depict all the raw threshold estimates (also shifted). The error bars show the standard error

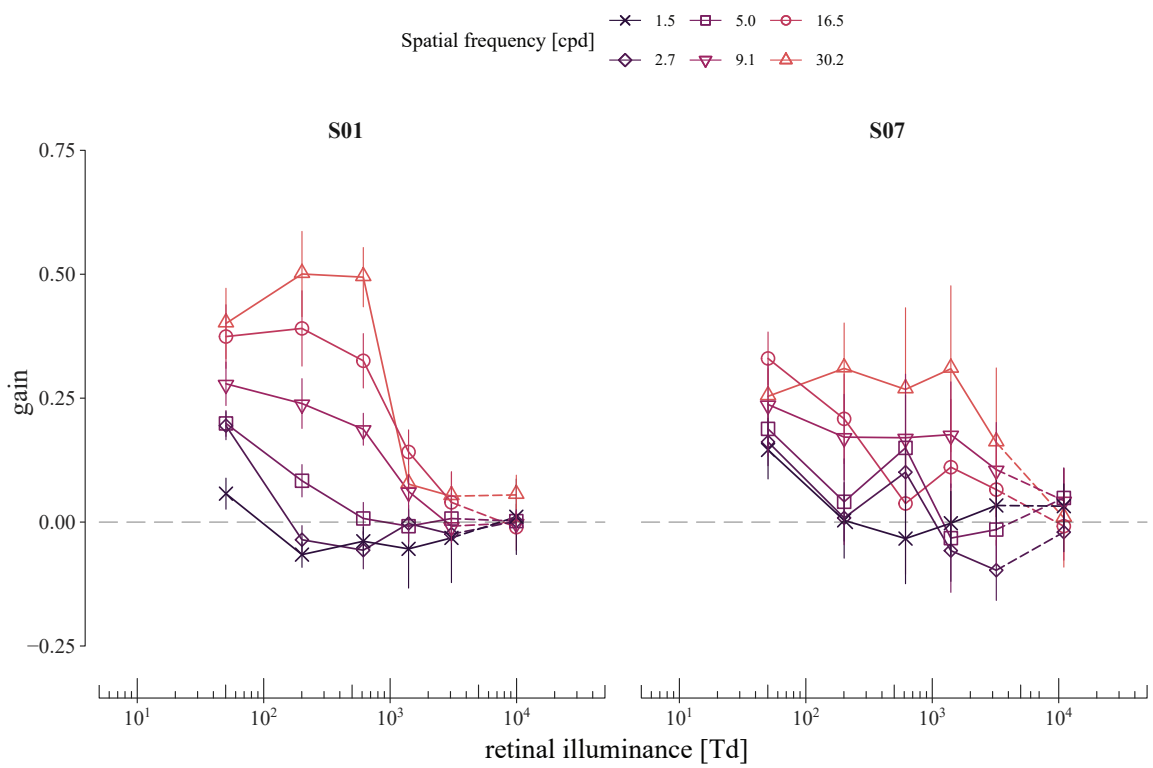


**Figure 2.9** The estimated gain (open markers) for each spatial frequency (lines). The gain was calculated as the slope of the linear regression through every illuminance level and its two neighbors (backward and forward). The boundaries were omitted. The error bars show the regression standard error

#### *2.2.5.4. Sensitivity as a function of illuminance and frequency*

The relationship between the spatial frequencies and the gain of increase with retinal illuminance can be more clearly seen in Figure 2.10. The depicted linearly interpolated mean gain on the illuminance-frequency plane, pooling the data from both experiments, ranged from .46 to -.15. Excluding the negative slope segments, the estimated gains were in fair agreement with a theoretical gradual DeVries-Rose to Weber transition translated to higher illuminance levels with increasing spatial frequency. This lateral shift of the transition point among frequencies was previously described in terms of relative retinal illuminance, i.e., the illuminance divided by spatial frequency squared, which stems from the hypothesis that the transition is determined by the luminous flux collected through circular receptive fields whose size scales as the inverse of the spatial frequency (e.g., Graham, 1989); a theory that was found to reasonably approximate this effect (Mustonen et al., 1993; see parallel orange lines in Figure 2.10).

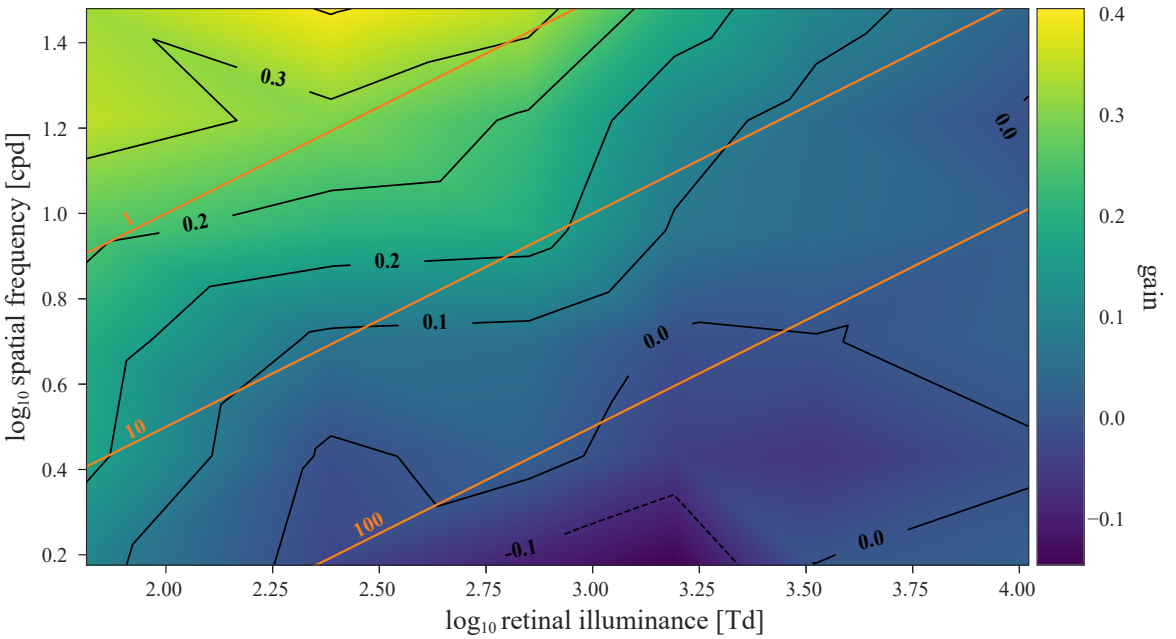


**Figure 2.10** The linearly interpolated average gain of sensitivity increase across observers on the log-log illuminance-frequency plane. The data were pooled from both experiments, including all the observers in the respective illuminance range. The orange parallel isolines depict three indicative values of relative retinal illuminance, i.e., the illuminance divided by spatial frequency squared

## 2.3. CSF modeling in the luminance domain

As described in the chapter introduction, the sensitivity increase in the luminance domain can be approximated by three linear segments in log-log space translated to higher background luminance level for increasing spatial frequency (recall the example in Figure 2.1). According to this approximation, at low light levels where dark noise is dominant (linear region), the amplitude threshold is independent of luminance ($\Delta L$ = constant); thus, contrast sensitivity becomes proportional to the background luminance ($L/\Delta L \sim L$). At intermediate light levels where quantal noise is dominant (DeVries-Rose region), the amplitude threshold is proportional to the square root of luminance ($\Delta L \sim \sqrt{L}$), following the Poisson distributed fluctuations in the number of photons absorbed by the retina. Therefore, sensitivity is also proportional to the square root of the luminance. At higher light levels, after the final transition (Weber region), the amplitude becomes proportional to the background luminance ($\Delta L \sim L$), and thus sensitivity is independent of luminance, obeying Weber's law ($L/\Delta L$ = constant). The suitability of this theoretical construct for describing sine-wave contrast sensitivity as a function of luminance intensity has been questioned over the years (García-Pérez and Peli, 1997; García-Pérez, 2005).

Although this approach constitutes a simplified model of a complex system of contributing adaptation mechanisms, e.g., different types of photoreceptors and their interactions, it has been proven successful in approximating sensitivity variations in the luminance domain (Pelli, 1990; Mustonen et al., 1993; Silvestre et al., 2018) and forms the basis of the most widely-used class of CSF models (Rovamo et al., 1994; Barten, 1999).

In this section, combined with our dataset, we assessed how well this non-linear relationship accounts for the measured data in a wide range of studies in the literature (Depalma and Lowry, 1962; Patel, 1966; Van Nes and Bouman, 1967; De Valois et al., 1974; Banks et al., 1987; Hess and Howell, 1988; Rovamo et al., 1993; Rovamo et al., 1994; Peli et al., 1996; Kim et al., 2013; Shannon et al., 1996; Rasengane et al., 2001; Valero et al., 2004; Silvestre et al., 2018; Bierings et al., 2019) using continuous piecewise linear regression analysis. This regression method is structured to answer whether for a given luminance range: (a) sensitivity is independent of luminance (i.e., a single segment with zero slope), (b) a critical point of change in the slope exists, (c) the linear to DeVries-Rose to Weber law holds (i.e., transition from a slope of 1 to .5 to 0), and (d) a decreasing region is present (i.e., a segment with negative slope). We then adopted a variant of the models derived from this construct (Rovamo et al., 1994) that incorporates basic elements in the human visual signal processing to describe the CSF in the luminance domain. It comprises optical factors, the addition of photon shot noise, lateral inhibition, the addition of late neural noise, and a matched filter (Hauske et al., 1976) with a sampling aperture. The model (Figure 2.11) was adjusted to include an additional early neural noise component (related to the linear segment preceding the DeVries-Rose region) that was found to dominate low luminance intensities and spatial frequencies (Silvestre et al., 2018). Finally, we examine the model's performance on the measured CSFs and evaluate the basic CSF characteristics as a function of luminance, namely, the location and amplitude of the peak, the area under the curve, and the spatial resolution limit.

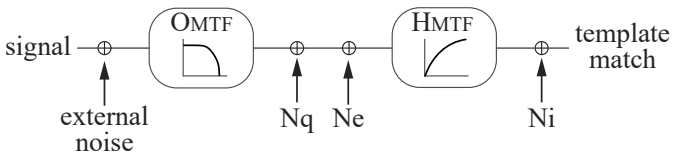


**Figure 2.11** Block diagram of the visual processing model. The stimulus is low-pass filtered by the optical MTF ($O_{MTF}$) of the eye before photon ($N_q$) and early noise ($N_e$) components are added. A high-pass filter ($H_{MTF}$) is then applied due to lateral inhibition, and late internal noise is added ($N_i$) before the signal is interpreted. External noise is included for completeness but assumed negligible

### 2.3.1. *Pre-processing*

Unavailable data were extracted from the published figures using the software by (Rohatgi, 2012). Retinal illuminance values were corrected for the Stiles-Crawford effect (Stiles et al., 1933) following the equation 2.1 given above, with a $\beta$ coefficient of .12. The same formula was used to convert display luminance values to retinal illuminance. Where unavailable, the pupil diameter $d$ was approximated for the corresponding adaptation luminance $L$ using the unified formula in Watson and Yellott (2012) as a function of the corneal flux density $F = LaM$ for an observer of twenty-five years old or the study population mean, where $a$ the adapting field area in $deg^2$ and $M$ a factor equal to .1 for monocular viewing and 1 otherwise. The formula is given as follows:

$$d_{mm}(F, age) = D_{SD}(F) + (age - 28.58)\,[.02132 - .009562 D_{SD}(F)] \tag{2.3}$$

where $D_{SD}$ the adjusted for the number of eyes equation by Stanley and Davies (1995) given as:

$$D_{SD}(F) = 7.75 - 5.75 \left( \frac{(F/846)^{.41}}{(F/846)^{.41} + 2} \right) \tag{2.4}$$

### 2.3.2. *Segmented regression*

Breakpoints were estimated using the iterative method described by Muggeo (2003). Given one breakpoint the model is expressed as:

$$S_i = \alpha + \beta_1 I + U_1 (I - \psi_1)_+ \tag{2.5}$$

where $(I - \psi_1)_+ = (I - \psi_1) \times G(I > \psi_1)$ and $G(\cdot)$ is a step function equal to one when $I > \psi_1$ and zero otherwise, $S_i$ denotes the sensitivity for a spatial frequency $u_i$, $I$ the retinal illuminance, $\alpha$ is the intercept, $\beta_1$ is the slope of the segment before the breakpoint $\psi_1$, and $\beta_2 = U_1 + \beta_1$ is the slope of the segment after the breakpoint. Additional breakpoints can be estimated similarly by adding the appropriate terms to equation (2.5). The existence of a breakpoint against the null hypothesis of a zero change in slope was validated using a two-sided Davies test (Davies, 1987) at .05 significance level with the additional constraint of its confidence interval to lie within the measured luminance boundaries. For the breakpoint estimation, the number of points was increased by N-1, where N is the total number of samples, using a shape-preserving piecewise cubic interpolation (Moler, 2004) on the midpoints between each pair of luminance values.

### 2.3.3. *Extending the CSF to high luminance range*

Following the model derived in Rovamo et al. (1994, 1997), with the assumption that critical illuminance is independent of grating stimulus area and external noise is negligible, the sensitivity-illuminance curve for spatial frequency $u_i$ can be described by:

$$S(u_i, I) = S'_{\max}(u_i)\left[1 + \frac{I_{c_i}}{I} + \left(\frac{I_{d_i}}{I}\right)^2\right]^{-0.5} \tag{2.6}$$

where $S$ is the Michelson contrast sensitivity, $S'_{\max}$ the sensitivity ceiling for a constant grating area, $I$ the retinal illuminance, and $I_{d_i}$ and $I_{c_i}$ the frequency-dependent critical illuminances that mark the transition from the linear to DeVries-Rose and DeVries-Rose to Weber regions, respectively. It should be noted that this form implies a gradual transition, that is qualitatively in better agreement with experimental findings (Rovamo et al., 1997) and satisfies the empirical constraints in García-Pérez (2005). The $S'_{\max}$ is defined as:

$$S'_{\max}(u_i) = S_{max}(u_i)\left[1 + \left(\frac{A_c(u_i)}{A}\right)^{-0.5}\right] \tag{2.7}$$

where $A$ is the stimulus grating area and $A_c(u_i)$ the critical area where spatial integration saturates. Assuming that the latter is independent of retinal illuminance, it can be expressed as:

$$A_c(u_i) = A_0\left[1 + \left(\frac{u_i}{u_{max}}\right)^2\right]^{-1} \tag{2.8}$$

where $A_0$ and $u_{max}$ the upper spatial summation limits for the grating area and the critical spatial frequency, respectively. The $S_{max}$ is then given as:

$$S_{max}(u_i) = K_0 O_{MTF}(u_i) H_{MTF}(u_i)\sqrt{A_c(u_i)} \tag{2.9}$$

where $K_0$ is a constant, $O_{MTF}$ is the low-pass optical modulation transfer function, and $H_{MTF}$ is the high-pass filter due to lateral inhibition. The constant $K_0$ is expressed as:

$$K_0 = \sqrt{\frac{\eta_{max}}{2d'^2 N_i}} \tag{2.10}$$

where $\eta_{max}$ is the maximum efficiency of the local matched filter, $d'$ is a detectability constant (Tanner and Birdsall, 1958) that depends on the task and the threshold level, and $N_i$ is the late

noise. The choice of the human optical MTF formula varies in the literature (Watson, 2013b). Here, for comparison purposes, we adopt a Gaussian form (Barten, 1999) that accounts for both the optical attenuation and retinal sampling factors:

$$O_{MTF}(u_i) = e^{-2\pi^2\sigma(d)^2u_i^2} \tag{2.11}$$

where $\sigma(d)$ is the SD of the line-spread function relative to the pupil diameter $d$ [mm]:

$$\sigma(d) = \sqrt{\sigma_0^2 + (C_{ab}d)^2} \tag{2.12}$$

where $\sigma_0$ can be considered constant for foveal vision and $C_{ab}$ an increment weight for increasing pupil size estimated at 0.08 arcmin/mm (Barten, 1999). In the original model, low-frequency attenuation was found to decrease linearly with increasing spatial frequency. However, this appears to be valid only at a limited frequency range (Donner and Hemilä, 1996). Here, we adopted the following approximation formula for lateral inhibition (Barten, 1999) but allowing for the square exponent to vary; the parameter $u_0$ and is the upper-frequency limit for lateral inhibition and $\nu$ is a free parameter:

$$H_{MTF}(u_i) = \sqrt{1 - e^{-(u_i/u_0)^\nu}} \tag{2.13}$$

The best fit for the parameters $K_0$, $u_0$, $\sigma_0$, $\kappa$, and the vectors $\mathbf{I}_c$ and $\mathbf{I}_d$ was found by simultaneously minimizing the sum of squared errors in log-space for all spatial frequencies with more than two samples in luminance. Where the total number of frequencies was below four, the sensitivity-illuminance curves were estimated using equation (2.6) with $\mathbf{S}'_{\text{max}}$ as a free parameter. The summation parameters $A_0$ and $u_{max}$ were fixed at 320 $deg^2$ and .465 $c/deg$, respectively, as estimated in Rovamo et al. (1994). Although the actual values might differ, any deviation will be reflected in the variability of the fitted parameters among studies. Fitting performance is expressed as the root mean square (RMS) and normalized root mean square (NRMS) errors, as defined in Watson and Ahumada (2005). The penalization was used to adjust for the number of free parameters and is given as:

$$NRMS = \sqrt{\frac{1}{J-N}\sum_{j=1}^{J}\left(t_j - \hat{t}_j\right)^2}$$

where J the number of different stimuli, N the number of the model parameters, $t_j$ the mean thresholds, and $\hat{t}_j$ the model predictions in dB (the $log_{10}$ threshold multiplied by 20).

### 2.3.4. Results

#### 2.3.4.1. CSF model fits

Figure 2.12 illustrates the model fit to the measured contrast sensitivities as a function of retinal illuminance for all spatial frequencies across studies (equations 2.6 – 2.13; the fitted parameters are given below in Table 2.7 and Figure 2.14). Where applicable, the model was fitted to the average observer (eleven studies). Despite the vast differences in the experimental conditions, the data exhibit a qualitatively similar relationship to background luminance. Generally, as the spatial frequency increases, the curve becomes steeper, and the asymptotic region translates to higher luminance. The RMS error for each study is shown at the top left of each panel in Figure 2.12. The total RMS error for all studies combined was 1.19 dB.

#### 2.3.4.2. Segmented regression

Figure 2.13 summarizes the results of the segmented regression on the measured sensitivity data. For visualization purposes, the slopes of each segment are presented as a function of relative retinal illuminance, i.e., the retinal illuminance divided by the spatial frequency squared (Mustonen et al., 1993). Qualitatively, a region where the DeVries-Rose to Weber law holds in a strict sense (the observed values agree with the predicted slopes of .5 and 0 within an arbitrary threshold of $\pm$.05) does seem to exist for specific luminance and frequency conditions, but generally, this range seems restricted (see dashed thin lines in Figure 2.13; panels A2, B2-3, D1, and E3). Instead, the slopes gradually decrease with increasing relative illuminance from a value between .5 and 1 to zero approaching a Weber region. This observation translates as follows: for decreasing/increasing retinal illuminance or increasing/decreasing frequency, the transition towards the contrast sensitivity ceiling becomes steeper/more flat. Negative-slope segments, below an arbitrary threshold value of -.1 ($M$ = -.19, $SD$ = .09), were found in four of the studies (Kelly, 1972; Rovamo et al., 1994; Peli et al., 1996; and the present one) at spatial frequencies between .25 cpd and 8 cpd and starting log relative illuminance between .79 and 3.20 Td deg$^2$. A negative slope was also present in (Depalma and Lowry, 1962; Van Nes and Bouman, 1967; De Valois et al., 1974; Hess and Howell, 1988; Valero et al., 2004; Kim et al., 2013; Silvestre et al., 2018) at roughly the same frequency range that can, however, be considered as negligible ($M$ = -.05, $SD$ = .02). Note that in two of the studies (Shannon et al., 1996; Rasengane et al., 2001) the stimuli were temporally modulated at low temporal frequency (6 Hz) that could diverge the slopes from the ideal DeVries-Rose to Weber transition depending on luminance (Kelly, 1972). However, in both cases, the slopes were in relatively better agreement with this approximation law compared to the rest of the data.

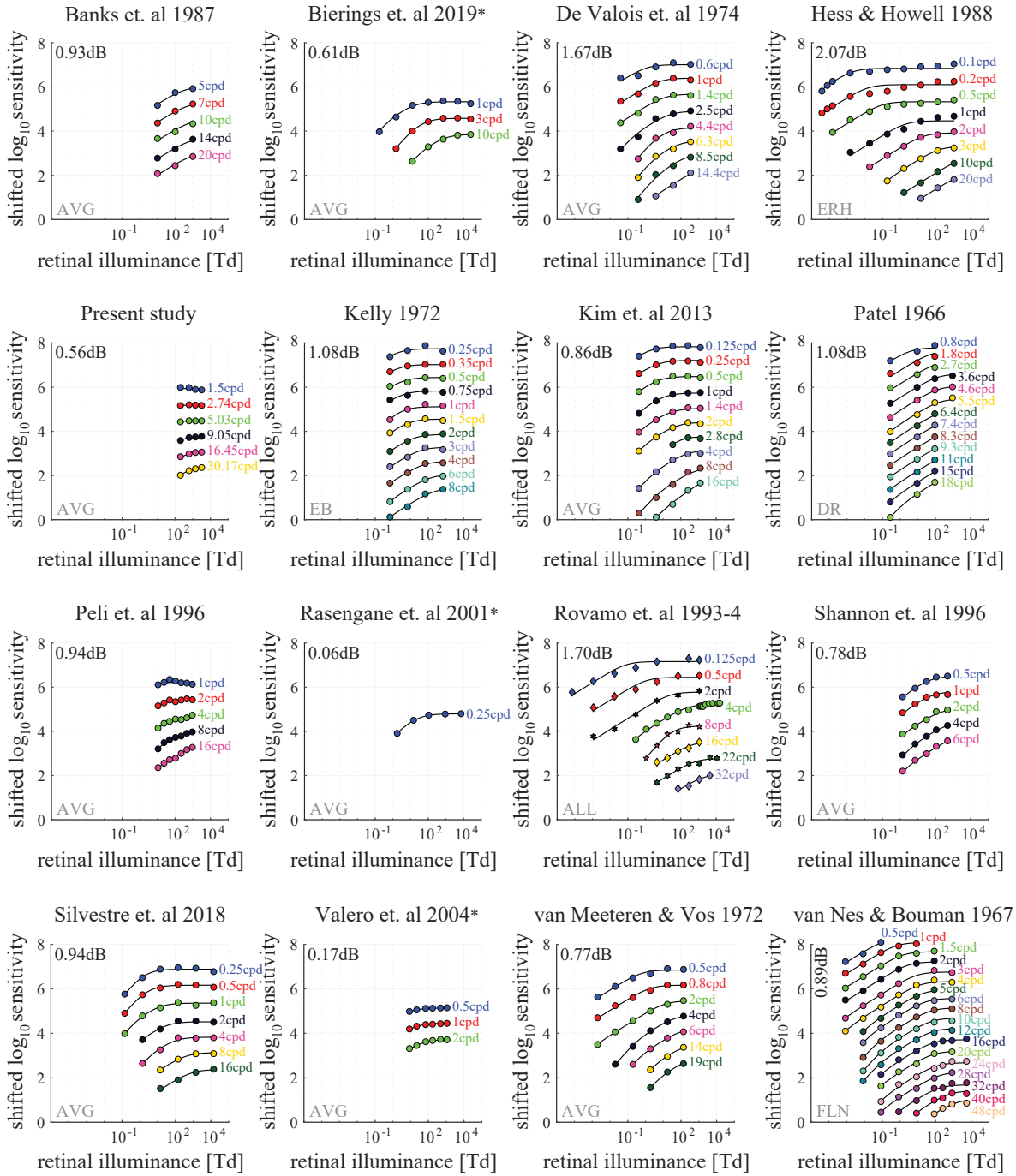


**Figure 2.12** Model fits (black lines) to the measured sensitivity data (markers ≈ .25 log units width) for each spatial frequency as a function of retinal illuminance across studies. Different markers within the same study indicate different observers. Sensitivities were vertically shifted from higher to lower spatial frequency for visualization purposes. The RMS error [dB] is shown at the top left of each panel. The observers' initials are shown at the bottom left, where AVG and ALL indicate the average and all the observers, respectively. The asterisk indicates a model fit with $S'_{max}$ as a free parameter

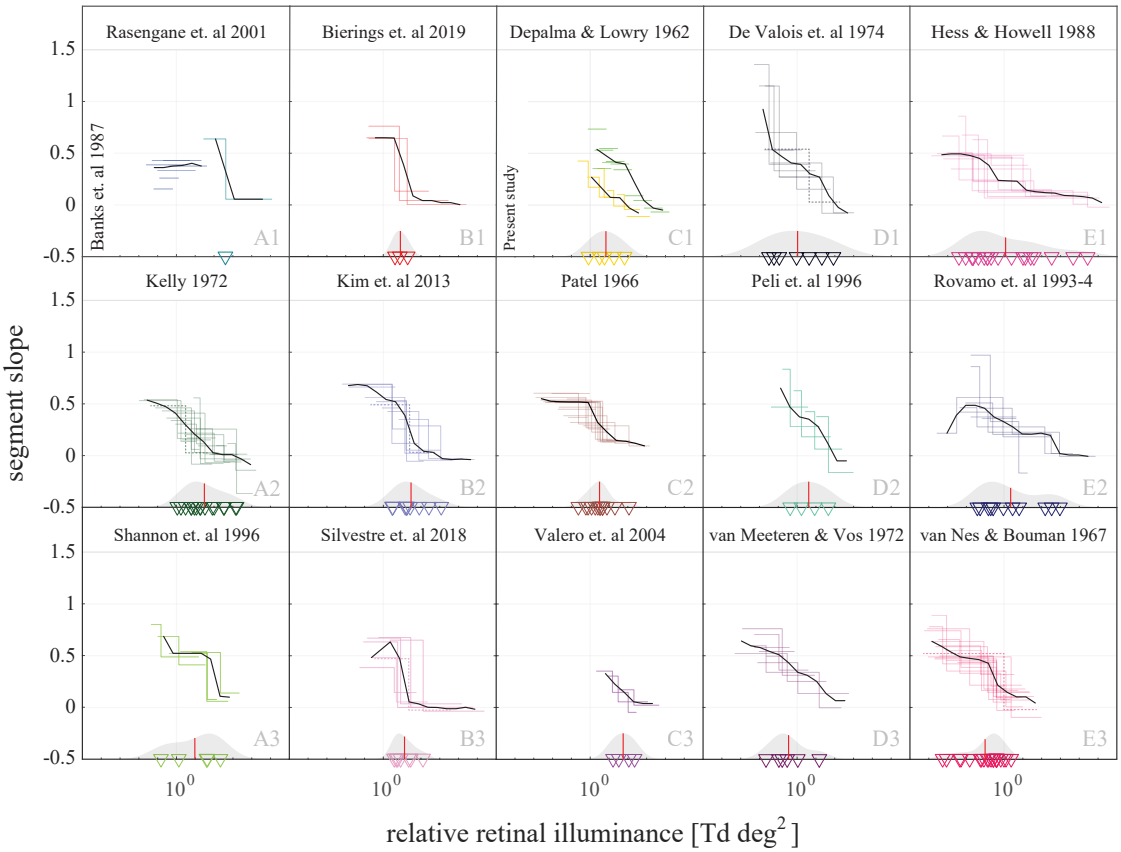


**Figure 2.13** The segmented regression slopes across studies as a function of relative retinal illuminance [Td deg$^2$]. The thin lines depict the slope of the segments for each spatial frequency. The dashed thin lines indicate a transition with slopes .50±.05 and ±.05 (panels A2, B2-3, D1, and E3). The thick black curves show the median slope over a non-overlapping sliding window of .5 log units. The markers depict the estimated breakpoints and the vertical red lines the mean value of their density estimate. The studies in panels A1 and C1 were grouped for visualization purposes (vertical legend for the leftmost curve)

### *2.3.4.3. CSF model estimates and reduction*

The fitted global parameters are given in Table 2.7. The mean estimates were $K_0$ = 343 (*SD* = 248), $u_0$ = 5.4 (*SD* = 3.9) c/deg, $\sigma_0$ = .55 (*SD* = .15) arcmin, and $\nu$ = 2.4 (*SD* = .6). Estimates near the parameter boundaries were excluded in calculating these means. A possible explanation for this discrepancy is discussed below. In another variant of the same class of models, the one from Barten (1999), $u_0$, and $\sigma_0$ were estimated at 7 c/deg and .5 arcmin, respectively, while the parameter $\nu$ was assumed fixed at 2. The estimated critical illuminance vectors are presented in Figure 2.14. The critical illuminance $I_c$ that marks the transition to a Weber region was found to be log-linearly related to spatial frequency, over a wide frequency range. Previous studies indicated that $I_c$ is approximately proportional to the spatial frequency squared (Van

Nes and Bouman, 1967; Mustonen et al., 1993; Rovamo et al., 1994), i.e., a slope of 2. We estimated a mean slope of 1.7 (SD = .6), in good agreement with the above. The relation between the spatial frequency and the critical quantity $I_d$ was less clear, mainly since the luminance-frequency sampling across studies did not allow for reliable estimates. However, at this point, we will assume a first-degree polynomial approximation. Therefore the parameters $I_c$ and $I_d$ as a function of frequency can be expressed as:

$$log_{10}I_c(u) \approx \text{constant}_c + \text{slope}_c\, log_{10}u \tag{2.14}$$

$$log_{10}I_d(u) \approx \text{constant}_d + \text{slope}_d\, log_{10}u \tag{2.15}$$

| Study | $K_0$ | $u_0$ [c/deg] | $\sigma_0$ [arcmin] | $\nu$ |
|---|---|---|---|---|
| Banks et. al 1987 (AVG) | 107.45 | 2.98 | .37 | 2.00 |
| De Valois et. al 1974 (AVG) | 145.10 | 3.67 | <.1 | 2.91 |
| Hess & Howell 1988 (ERH) | 843.90 | 13.64 | .50 | 2.33 |
| Present study (AVG) | 419.81 | 3.56 | .70 | 2.71 |
| Kelly 1972 (EB) | 137.79 | 1.66 | <.1 | 3.92 |
| Kim et. al 2013 (AVG) | 192.86 | 2.65 | <.1 | 2.09 |
| Patel 1966 (DR) | 634.84 | UB | .53 | 2.24 |
| Peli et. al 1996 (AVG) | 76.82 | 2.40 | .62 | 2.54 |
| Rovamo et. al 1994 (JM) | 375.23 | 7.54 | <.1 | 1.95 |
| Shannon et. al 1996 (AVG) | 253.80 | UB | <.1 | 1.30 |
| Silvestre et. al 2018 (AVG) | 177.36 | 3.54 | <.1 | 2.58 |
| van Meeteren & Vos 1972 (AVG) | 389.49 | 7.39 | .74 | 2.30 |
| van Nes & Bouman 1967 (FLN) | 699.34 | 10.45 | .36 | 2.42 |

**Table 2.7** Estimated global parameters. In parenthesis, the observer's initials, where AVG the average observer. The letters UB indicate the upper bound (set at 16 cpd), and the gray color a fixed value

Based on the above results, we explored the effect of reducing the critical illuminance parameters on the total RMS error for all the studies combined. This is an essential step as it not only considerably reduces the total estimated parameters, but it also allows us to extract the CSF surface in the frequency-luminance domain. If we fixed $\text{slope}_c = 2$ in equation (2.14), and fitted $\text{constant}_c$ along with the individual values of $\mathbf{I}_d$, the RMS and NRMS error increased by 1.30 dB and 1.35 dB, respectively. Re-estimating the global model parameters with simultaneous optimization of $\text{constant}_c$ increased the RMS error by .37 dB and .22 dB, respectively. Fitting $\text{slope}_c$ as well, resulted in a smaller increase of .59 dB and .52 dB (optimized estimates: .19 dB

and .03 dB) in RMS and NRMS errors, respectively. Fitting all four terms (proposed model) in equations (2.14) and (2.15) led to an increase of 1.27 dB and 1.06 dB (optimized: .39 dB and .08 dB) for the RMS and NRMS errors, respectively. Increasing the polynomial terms further did not produce any considerable improvement. The fitted values are given in Table 2.8.

| Study | $\text{constant}_c$ | $\text{slope}_c$ | $\text{constant}_d$ | $\text{slope}_d$ |
|---|---|---|---|---|
| Banks et. al 1987 (AVG) | 1.48 | 1.20 | -.10 | .00 |
| Bierings et. al 2019 (AVG) | 1.24 | 1.72 | .55 | 1.75 |
| De Valois et. al 1974 (AVG) | .80 | 1.50 | -2.18 | 3.97 |
| Hess & Howell 1988 (ERH) | .31 | 3.08 | -3.32 | 2.71 |
| Present study (AVG) | -1.22 | 3.16 | -1.09 | 2.38 |
| Kelly 1972 (EB) | .82 | 1.58 | -1.23 | 1.68 |
| Kim et. al 2013 (AVG) | 1.45 | 1.52 | .39 | 1.48 |
| Patel 1966 (DR) | 1.64 | .88 | -1.99 | 2.85 |
| Peli et. al 1996 (AVG) | .58 | 2.07 | .91 | .00 |
| Rovamo et. al 1994 (JM) | .69 | 1.93 | -3.75 | 2.27 |
| Shannon et. al 1996 (AVG) | 2.08 | 1.11 | 1.04 | .75 |
| Silvestre et. al 2018 (AVG) | 1.17 | 1.48 | .64 | 0.94 |
| Valero et. al 2004 (AVG) | 1.34 | 1.20 | -1.03 | .00 |
| van Meeteren & Vos 1972 (AVG) | .50 | 1.51 | -2.79 | 3.83 |
| van Nes & Bouman 1967 (FLN) | -.02 | 1.64 | -3.24 | 3.07 |

**Table 2.8** The fitted constants (not optimized) in the approximation of $I_c$ and $I_d$ (equations 2.14-2.15). In parenthesis, the observer's initials, where 'AVG' the average observer.

#### *2.3.4.4. CSF characteristics*

In Figure 2.15 we examined the relationship of the main CSF characteristics to retinal illuminance across studies (based on the proposed model and the fitted parameters), namely, the location ($F_{max}$) and the magnitude ($G_{max}$) of peak sensitivity, the highest visible frequency ($F_c$), and the area under the log-CSF (AULCSF). The grating area was kept constant at 4 $\text{deg}^2$. Where the pupil size was not fixed, it was estimated as described in the methods section. Datasets with limited spatial frequency sampling (Banks et al., 1987; Rovamo et al., 1994; Shannon et al., 1996; Valero et al., 2004; Bierings et al., 2019) were excluded from further analysis. The peak sensitivity was predicted to asymptote above roughly $10^3$ Td, while the peak frequency gradually increased from about 2 to 10 cpd with increasing illuminance (omitting predicted

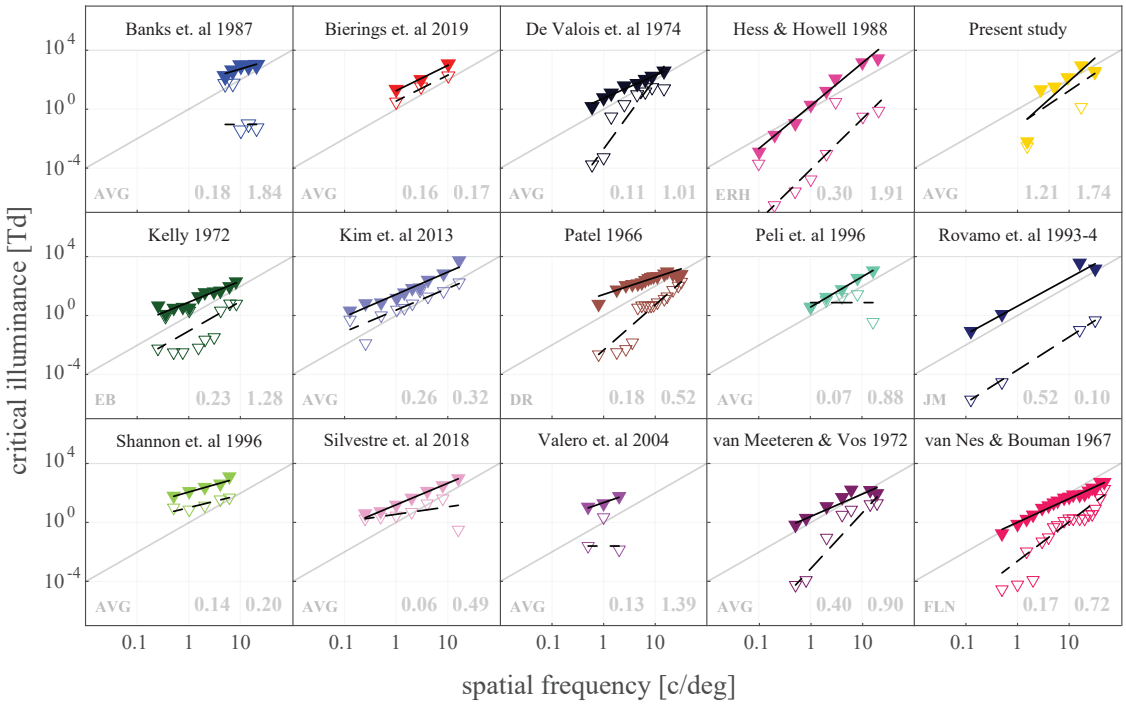


**Figure 2.14** The critical illuminance vectors $I_c$ (filled symbols) and $I_d$ (open symbols) as a function of spatial frequency. The grey line illustrates a critical illuminance proportional to the spatial frequency squared. The solid and the dotted lines depict the log-linear regressions for the $I_c$ and $I_d$ vectors, respectively. On the bottom left the observer's initials, where 'AVG' indicates the average observer. On the bottom right, the RMS errors in log units left for $I_c$ and right for $I_d$

sensitivities below zero). The visibility area and the high-frequency cut-off did not appear to reach a plateau within the tested range. It should be noted that the predicted relationship for the high-frequency cut-off (i.e., visual acuity) does not account for a well-known discontinuity (Hendley, 1948; García-Pérez, 2005) during the transition from scotopic to photopic conditions (approximately between $4.7\times10^{-3}$ and 710 Td; Zele and Cao, 2015).

### *2.3.5. Barten's CSF formula and model assessment*

Barten (1999) produced an influential model of the human CSF. It was found successful in describing previous sensitivity measurements with various grating stimuli in the photopic range and is still generally regarded as the state-of-the-art in the field. Naturally, even though it was similarly derived, this model serves as the reference point for assessing the performance of the proposed approach. Following the suggestion in Barten (1999), we fitted the model (BT P3) with three free parameters: the quantum efficiency of the eye $\eta$, a constant $\sigma_0$, and the signal-to-noise ratio $\kappa$. The constant $\sigma_0$ controls the optical attenuation at high spatial

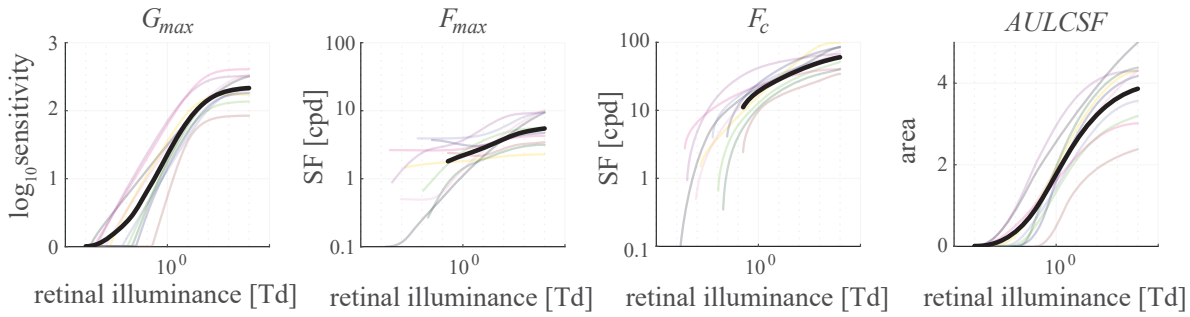


**Figure 2.15** The main CSF characteristics as a function of retinal illuminance. From the left to the right: The peak sensitivity, the spatial frequency of the peak, the high-frequency cut-off where sensitivity asymptotes to zero (extrapolated), and the area under the log-CSF (from .1 cpd to the highest visible frequency). The thin lines depict different studies. The thick black line is the mean. Negative sensitivity values were assumed zero and were omitted from the spatial frequency measures

frequencies, whereas the signal-to-noise ratio $\kappa$ is a multiplication factor. The effective retinal illuminance is modulated by the quantum efficiency $\eta$. A brief model description is provided in Appendix B. Without modifying the model and in order to roughly match the degrees of freedom in the proposed approach, we also extended the number of free parameters to six (BT P6) by allowing the maximum integration area $X_{max}$, the maximum number of cycles $N_{max}$, and the spatial frequency limit for lateral inhibition $u_0$ to vary. The two former variables interact with the field size and the spatial frequency, whereas the latter shifts the location of the low-frequency roll-off. The photon conversion factor $p$ [photons/sec/deg$^2$/Td] was kept constant across studies at $1.2 \times 10^6$ for usual luminance conditions. Although this variable depends on the specific light source, any differences will be absorbed by the estimates of quantum efficiency $\eta$; thus, resulting in equal fitting performance. Note, however, that as we extend to the rod-dominated range (scotopic conditions), due to the shift of the peak spectral sensitivity to shorter wavelength, reduced performance is to be expected if this parameter is not properly adjusted as a function of the background luminance level. In the case of a Gaussian grating envelope, the angular size of the patch was set to two standard deviations. The rest of the parameters were fixed to the typical values for an average observer, foveal vision, and photopic viewing conditions (Table B.4). The model was adjusted to monocular viewing by reducing the sensitivity by a factor of $\sqrt{2}$ (Campbell and Green, 1965; also the adjustment used in Barten, 1999), although the actual value might be higher (Baker et al., 2018). Where applicable, to correct for the well-known oblique effect (see the introduction in Kaspiris-Rousellis et al., 2017) in the diagonally oriented gratings, the maximum number of cycles was reduced by a factor of two as suggested in Barten (2003). Every effort was made to match the experimental conditions of each study, given the available information. However,

as generally, Barten's model is more strict regarding the accuracy of the measured physical quantities, an underestimation of its performance cannot be ruled out (although partly negated by the increase in the number of free parameters). The studies with limited spatial frequency sampling were excluded. The comparison was run with four variants of the modified model presented here. The basic model (RV M0) included the four global parameters ($K_0$, $u_0$, $\sigma_0$, and $v$) along with one critical illuminance estimate $I_c$ for each spatial frequency $u$. This form was extended (RV M2) by introducing the second critical illuminance vector $I_d(u)$. Both models were then tested with a log-linear approximation of their critical illuminance vectors (RV M0 to M1 and RV M2 to M3), as shown in equations 2.14 – 2.15, which also allows for interpolating (or extrapolating within limits) across spatial frequencies. The RMS and NRMS errors for each study are presented in Table 2.9. Across the board, RV M3 gives the best account of the data. The fitted parameters for the proposed model (RV M3) are given in Table 2.10. Where applicable, the variability in the parameter estimates within studies and the corresponding individual RMS errors for the RV M3 are reported in Table 2.11.

## 2.4. Discussion

In this chapter, we analyzed foveal (not strictly cone isolating conditions but generally non-peripheral) contrast sensitivity measurements in the literature that systematically investigated the effect of varying background luminance levels on the CSF. Using a custom-built display system, we supplemented the existing research with a dataset at relatively high luminance for spatial frequencies ranging from 1.5 cpd to 30 cpd and eleven observers to further our understanding of the asymptotic CSF behavior in this regime. We focused on a theoretical construct that describes this relationship, the linear to DeVries-Rose to Weber transition, which manifests as an increase in sensitivity with increasing luminance with slopes of 1 and 0.5 in double log space to a Weber region where sensitivity becomes independent of luminance (i.e., a slope of zero). By using continuous segmented regression analysis, we found that the DeVries-Rose to Weber transition in a strict sense holds only for a limited range across spatial frequencies and luminance conditions. Instead, a curvilinear form with a gradual transition among the three regions appears as a valid approximation for sine-wave gratings. Specifically, when relative retinal illuminance decreased, either by increasing the spatial frequency or decreasing the retinal illuminance, the transition towards the sensitivity ceiling became steeper, in agreement with the empirical constraints derived in García-Pérez (2005). Except for one study (Van Nes and Bouman, 1967), there was insufficient evidence for a Weber region beyond approximately 16 cpd, i.e., increasing luminance continued to increase log sensitivity, and no

| Model | DV 1974 | HH 1988 | Our | KL 1972 | KM 2013 | PT 1966 | PL 1996 | SL 2018 | VMV 1972 | VNB 1967 |
|---|---|---|---|---|---|---|---|---|---|---|
| BT P3 | 3.49 (3.63) | 4.55 (4.68) | 1.36 (1.45) | 3.50 (3.63) | 2.78 (2.87) | 2.04 (2.09) | 2.06 (2.16) | 3.78 (3.94) | 3.07 (3.21) | 2.33 (2.37) |
| BT P6 | 3.11 (3.39) | 3.61 (3.84) | 1.08 (1.25) | 3.18 (3.42) | 2.22 (2.37) | 1.67 (1.75) | 1.31 (1.44) | 2.36 (2.58) | 2.28 (2.50) | 1.81 (1.86) |
| RV M0 | 2.69 (3.12) | 2.10 (2.38) | .60 (.79) | 1.14 (1.40) | 1.60 (1.85) | 1.42 (1.57) | .97 (1.12) | 1.74 (2.08) | 1.77 (2.04) | 1.57 (1.78) |
| RV M1 | 2.89 (3.02) | 2.48 (**2.63**) | **.85** (**.98**) | 1.42 (1.53) | 1.78 (1.86) | 1.66 (1.62) | 1.00 (1.10) | 1.81 (1.99) | 1.93 (2.03) | 1.86 (1.92) |
| RV M2 | 1.67 (2.33) | 2.07 (2.62) | .56 (.97) | 1.08 (1.69) | .86 (1.18) | 1.08 (1.45) | .94 (1.21) | .94 (1.32) | .77 (1.04) | .89 (1.15) |
| RV M3 | **2.07** (**2.23**) | **2.46** (2.67) | **.85** (1.04) | **1.38** (**1.53**) | **1.12** (**1.20**) | **1.36** (**1.35**) | **.97** (**1.10**) | **1.18** (**1.34**) | **1.65** (**1.80**) | **1.46** (**1.53**) |

**Table 2.9** The models' fitting performance in terms of RMS error in dB (the log thresholds multiplied by 20) across studies (columns). In parenthesis, the penalized for the number of parameters NRMS error [dB] (after penalization for the number of parameters). **BT P3**: Barten's suggested model with three free parameters ($\kappa$, $\eta$, and $\sigma_0$). **BT P6**: Barten's model with six free parameters ($\kappa$, $\eta$, $\sigma_0$, $u_0$, $X_{max}$, and $N_{max}$). **RV M0**: The modified Rovamo model with four global parameters ($K_0$, $u_0$, $\sigma_0$, and $\nu$) and one critical illuminance vector with length equal to the number of spatial frequencies: $I_c(u)$. **RV M1**: The modified Rovamo model with six free parameters ($K_0$, $u_0$, $\sigma_0$, $\nu$, and the two critical illuminance coefficients of the log-linear approximation constant$_c$ and slope$_c$). **RV M2**: The modified Rovamo model with four global parameters ($K_0$, $u_0$, $\sigma_0$, and $\nu$) and two critical illuminance vectors with length equal to the number of spatial frequencies: $I_c(u)$ and $I_d(u)$. **RV M3**: The proposed model modification with eight free parameters ($K_0$, $u_0$, $\sigma_0$, $\nu$, and the four critical illuminance coefficients of the log-linear approximation constant$_c$, slope$_c$, constant$_d$, and slope$_d$). In bold, the minimum errors excluding the models RV M0 and RV M2 as they require an additional step for interpolating across spatial frequencies

saturation occurred at the luminance range tested (up to $\approx 10^4$ Td). When we further extended the retinal illuminance for two of the observers in our study, sensitivity asymptoted at all the tested spatial frequencies up to 30 cpd. A decrease in sensitivity with increasing luminance was also present at frequencies between 0.25 cpd and 8 cpd for a large number of studies (Figure 2.13, panels A2, B2-3, C1, C3, D1-2, and E1-3). However, the slope of decrease was relatively small, and the results do not suffice to draw any further conclusions. It should be noted, though, that in most cases, we used the average observer that could cancel out any related individual differences (García-Pérez and Peli, 1997). In our data, we found significant but relatively small (maximum mean difference of .1 log units) decrements in sensitivity across observers above approximately $1.4 \times 10^3$ Td at 1.5 cpd. In agreement with our results, Bierings et al.

| Study | $K_0$ | $u_0$ | $\sigma_0$ | $\nu$ | $\text{const}_c$ | $\text{slop}_c$ | $\text{const}_d$ | $\text{slop}_d$ |
|---|---|---|---|---|---|---|---|---|
| De Valois et. al 1974 (AVG) | 159.05 | 4.88 | <.1 | 2.49 | .69 | 1.62 | -.57 | 2.09 |
| Hess & Howell 1988 (ERH) | 930.04 | UB | <.1 | 2.28 | .31 | 3.01 | -3.69 | <.1 |
| Present study (AVG) | 398.18 | 3.28 | .65 | 2.84 | .69 | 1.74 | -14.63 | 7.93 |
| Kelly 1972 (EB) | 132.21 | 1.57 | <.1 | 4.00 | .99 | 1.24 | -1.36 | 2.56 |
| Kim et. al 2013 (AVG) | 178.89 | 2.16 | .57 | 2.23 | 1.46 | 1.37 | .35 | 1.41 |
| Patel 1966 (DR) | 571.59 | UB | .39 | 2.15 | 1.24 | 1.41 | -3.80 | 4.16 |
| Peli et. al 1996 (AVG) | 74.05 | 2.39 | .40 | 2.44 | .29 | 2.50 | .93 | .53 |
| Silvestre et. al 2018 (AVG) | 179.25 | 3.56 | <.1 | 2.57 | 1.30 | 1.45 | .47 | .35 |
| van Meeteren & Vos 1972 (AVG) | 978.90 | UB | .18 | 2.34 | .46 | 2.52 | -1.55 | 2.88 |
| van Nes & Bouman 1967 (FLN) | 891.88 | UB | .39 | 2.12 | .03 | 1.64 | -1.98 | 1.93 |

**Table 2.10** The re-estimated parameters of the proposed model (RV M3) with simultaneous optimization of the critical illuminance coefficients ($\text{constant}_c$, $\text{slope}_c$, $\text{constant}_d$, and $\text{slope}_d$) across studies. In parenthesis, the observer's initials, where 'AVG' the average observer. A value of 'UB' indicates the upper bound (set at 16 cpd). Studies with limited spatial frequency sampling were omitted

(2019) found a significant sensitivity decrease (roughly .1 log units) for 1 cpd at approximately $3.5 \times 10^4$ Td compared with $4.7 \times 10^3$ Td, with the sensitivity ceiling reached at 710 Td. In the extended range for the two observers in our dataset, the slope of decrease appeared to recover with a further increase in retinal illuminance. However, given the relatively small magnitude of decrease, it was unclear whether this was due to fluctuations around the saturation level or a transition to a second Weber region, as discussed in Rovamo et al. (1997).

Following this theoretical construct, we adopted the model in Rovamo et al. (1994) consisting of a low-pass optical MTF, a high-pass MTF due to lateral inhibition, a local matched filter, and two noise sources (photon shot noise and late proximal noise) that limit visual detection across the luminance domain and generate this observed DeVries-Rose to Weber transition. We modified the model to include an additional early noise component that relates to the linear to DeVries-Rose transition ($I_d$) and was found to have a significant effect at low luminance and spatial frequencies (approximately .16 Td to 1.6 Td at spatial frequencies below 2 cpd; Silvestre et al., 2018). The preliminary segmented regression analysis revealed slopes between 1 and .5 that further supports the existence of a third region, in the context of this model, where contrast sensitivity becomes proportional to the background luminance. The critical illuminance that marks the transition to a Weber region was found to be log-linearly related to the spatial frequency, consistent with the above results. It was roughly proportional to spatial frequency squared, a phenomenon that is usually explained neurophysiologically by

| Study/Observer (RMSE) | $K_0$ | $u_0$ | $\sigma_0$ | $\nu$ | $\text{const}_c$ | $\text{slop}_c$ | $\text{const}_d$ | $\text{slop}_d$ |
|---|---|---|---|---|---|---|---|---|
| Present study | | | | | | | | |
| S01 (1.20 dB) | 537.52 | 3.00 | <.1 | 2.87 | -2.11 | 4.89 | -15.38 | 8.48 |
| S02 (1.48 dB) | 436.53 | 3.40 | .66 | 3.19 | .54 | 1.85 | -8.80 | 4.97 |
| S03 (1.44 dB) | 312.34 | 2.86 | .76 | 2.84 | 1.30 | 1.29 | -9.67 | 5.00 |
| S04 (1.00 dB) | 288.93 | 3.72 | .64 | 2.12 | .12 | 2.04 | 1.18 | 1.05 |
| S05 (1.03 dB) | 404.30 | 4.85 | .61 | 2.67 | .78 | 1.47 | -17.41 | 9.91 |
| S06 (0.89 dB) | 276.16 | 3.08 | .50 | 2.82 | .20 | 2.59 | -11.20 | 6.61 |
| S07 (0.79 dB) | 445.76 | 3.23 | .60 | 3.15 | .99 | 1.28 | -7.05 | 6.26 |
| S08 (1.06 dB) | 408.72 | 3.89 | .45 | 2.53 | -.02 | 2.96 | -6.10 | 2.08 |
| S09 (1.71 dB) | 393.66 | 2.62 | .66 | 2.32 | -4.41 | 1.24 | 1.69 | .69 |
| S10 (1.31 dB) | 400.10 | 3.08 | .80 | 3.03 | .50 | 1.57 | -4.22 | .83 |
| S11 (1.11 dB) | 651.08 | 3.49 | .68 | 3.26 | -.15 | 2.01 | -7.55 | 6.34 |
| Kim et. al 2013 | | | | | | | | |
| FLD (1.88 dB) | 237.08 | 3.82 | <.1 | 2.14 | 1.57 | 1.73 | -1.62 | 3.68 |
| JPT (1.65 dB) | 155.11 | 2.39 | .64 | 2.21 | 1.33 | 1.55 | .06 | 1.73 |
| KJK (1.59 dB) | 234.66 | 2.28 | .58 | 2.13 | 1.36 | 1.09 | .69 | 1.20 |
| PFI (2.10 dB) | 165.36 | 2.05 | .85 | 2.29 | .74 | 1.24 | 1.13 | .54 |
| RFM (1.28 dB) | 193.84 | 1.84 | .66 | 2.27 | 1.42 | 1.16 | .58 | 1.09 |
| TOM (1.30 dB) | 151.45 | 2.34 | <.1 | 1.95 | -6.59 | 8.30 | 1.43 | 1.28 |
| VIB (1.70 dB) | 136.66 | 1.66 | <.1 | 2.27 | -6.43 | 7.82 | 1.62 | .49 |
| Peli et. al 1996 | | | | | | | | |
| AL (2.14 dB) | 126.22 | 1.92 | .63 | 2.88 | -.81 | 2.91 | .92 | .95 |
| EF (0.85 dB) | 66.46 | 3.32 | .79 | 2.55 | .70 | 1.81 | .72 | .41 |
| JI (1.55 dB) | 106.98 | 2.40 | .64 | 2.60 | 1.18 | 1.73 | -11.94 | <.1 |
| RL (1.64 dB) | 41.86 | 2.09 | <.1 | 2.21 | .66 | 2.26 | -6.08 | <.1 |
| Silvestre et. al 2018 | | | | | | | | |
| DS (0.95 dB) | 246.29 | 4.42 | <.1 | 2.55 | .92 | 1.51 | .50 | .85 |
| EM (3.08 dB) | 231.21 | 3.67 | <.1 | 2.78 | 1.74 | 2.14 | .42 | <.1 |
| KL (1.97 dB) | 81.95 | 2.26 | .35 | 2.43 | 1.33 | .35 | .51 | 1.22 |
| MB (1.54 dB) | 269.88 | 4.14 | <.1 | 2.72 | .55 | 1.77 | .67 | .44 |

**Table 2.11** The estimated parameters of the proposed model (RV M3) with simultaneous optimization of the critical illuminance coefficients ($\text{constant}_c$, $\text{slope}_c$, $\text{constant}_d$, and $\text{slope}_d$) within studies

the constant-flux hypothesis (Graham, 1989; Mustonen et al., 1993). Our proposed model with eight parameters, i.e., equations (2.6) – (2.15), resulted in a total RMS error of 2.45 dB (1.57 dB if we re-estimated the global parameters by simultaneous optimization of the log-linear approximation coefficients) for all the studies combined, but further improvements can be made. The fitting performance was comparable with the state-of-the-art model by Barten (1999) tested with six free parameters (see Table 2.9), improving on average the penalized NRMS error by .86 dB (ranging from .2 dB to 1.89 dB) across studies. Note, however, that the results are within the context of extending the CSF in the luminance domain. Regardless, as both models belong to the same family, similar modifications can also be applied to Barten's formula.

Scrutiny of the global parameters revealed an inconsistency in the estimation of $\sigma_0$, which controls the optical attenuation at high spatial frequencies. This discrepancy could be an artifact due to limited frequency range sampling (Kelly, 1972; Shannon et al., 1996), or an overestimation of the pupil diameter. Except for one of these studies, pupil size was estimated from the display luminance (De Valois et al., 1974; Kim et al., 2013), or it was artificially dilated to a high value (Rovamo et al., 1994). The same inconsistency, however, did not persist to all the corresponding individual fits, with the other global individual parameters also generally in agreement. In two of the studies, the parameter $u_0$, the spatial frequency where the lateral inhibition ceases, was at the upper boundary. A possible explanation would be the presence of a local notch on the CSF (Patel, 1966), and limited frequency range sampling (Shannon et al., 1996). Note, though, that when we re-estimated the global parameters by simultaneous optimization of the log-linear approximation coefficients, three more studies converged to the upper boundary, which might indicate the need for additional constraints. The variability in the estimation of the constant $K_0$ was expected due to the vast differences in the experimental conditions. Nevertheless, fixing these parameters still provided us with reasonable fits.

However, other factors might be present. In the derivation of the model, we assumed that the critical area where spatial summation saturates is independent of retinal illuminance. A violation of this assumption would also cause this discrepancy. In fact, in an extension of a similar model, the one by Barten (1999), to scotopic conditions (Jarvis and Wathes, 2012), the spatial integration along with other parameters were adjusted to eccentric viewing that is more appropriate for rod-dominated vision. This approach essentially assumes a discontinuous piecewise function in luminance for the otherwise fixed model parameters, that was found successful in describing contrast sensitivity measurements under scotopic (i.e., rod-dominated) conditions. Inspecting the prediction of our model variant on the main CSF characteristics revealed a limitation at low light levels, i.e., no discontinuity in acuity during the transition

from the scotopic to the photopic range, that limits its application to moderate or higher light levels and foveal vision. A similar two-segment relationship could be investigated here.

Another drawback of this modeling approach is that it does not account for any decrease in sensitivity with increasing luminance, and therefore it is also upper-bounded for low to intermediate frequencies (based on the examined datasets). Incorporating a decreasing term in equation (2.6) is trivial (e.g., by adding the term $(I/I_b)^b$ in the parenthesis, where $I_b$ the transition point and $b$ the slope of decrease). However, this does not appear to be theoretically justified, and high-luminance data where this could be more accurately examined are scarce. Alternatively, given sufficiently dense sampling in luminance and frequencies, it is feasible to extend the CSF only by interpolating the parameters of a mathematical form, e.g., an asymmetric log-parabola (see subsection 2.2.3), in the luminance domain. An advantage of this approach is that by definition it can describe the decreasing sensitivity at lower frequencies while there is still sensitivity increase at the high-end with increasing luminance, and thus can be extended to high light levels and different conditions where this decrease is prominent, e.g., peripheral stimuli (Daitch and Green, 1969). This technique also allows for capturing other CSF features, e.g., low-frequency truncation (Watson and Ahumada, 2005). However, this approach is highly sensitive to measurement noise. We found that combining our model variant for interpolating across luminance with an asymmetric log-parabola form for interpolating across spatial frequencies performed similarly with the same number of free parameters, but could lead to an overestimation of the spatial resolution limit.

Whereas an alternative model derivation incorporating elements from the neurophysiology of vision (e.g., photoreceptor responses, retinal gain controls) would provide a more accurate description, this much more straightforward approach is a reasonable approximation to psychophysically and electrophysiologically (Shannon et al., 1996) measured contrast sensitivity variations in the luminance domain and allows for the extraction of the CSF surface as a function of light level over a wide range of normal viewing conditions.

# Chapter 3. The effect of frame rate on visual discomfort and fatigue

## 3.1. Introduction

The extensive use of electronic visual displays is known to induce ocular and vision-related symptoms of fatigue and discomfort (Bergqvist and Knave, 1994; Ye et al., 2007; Portello et al., 2012; Kim et al., 2016; Larese Filon et al., 2019). Naturally, any prolonged visual task under strenuous conditions, even with hard copies, e.g., sustained near work, would lead to the same results. However, specific display characteristics and their interaction with environmental variables and the user could exacerbate these symptoms, e.g., the effect of ambient light and viewing angle on the appearance of emissive displays. Previous studies found various associated factors related to these properties, e.g., gaze elevation (Menozzi et al., 1994), pixel density (Ziefle, 1998; Wright et al., 1999; Mayr et al., 2017), display luminance and ambient illumination (Benedetto et al., 2014; Kim et al., 2017; Fernandez-Alonso et al., 2020), and screen reflections (Shieh, 2000; Miyake-Kashima et al., 2005).

The current ergonomic guidelines for human-display interaction (ISO 9241-303, 2011) provide a detailed description of display specifications aimed to alleviate these effects. One of the many overall characteristics that manufacturers are required to address is motion fidelity. Motion fidelity is essential not only for depicting dynamic content but also for moving static images or text around the screen, such as when a user scrolls through a document. With sample-and-hold displays becoming the industry standard, e.g., liquid crystal displays, and given their current susceptibility to motion artifacts such as motion blur, the resulting frame sequence degradation has created another potential source of visual discomfort and fatigue. Despite the various techniques that were proposed to reduce these effects, e.g., backlight strobing (Feng, 2006), the effective transition time between two static frames remains the main limitation. Naturally, if not limited by the pixel response time, a higher frame rate results in higher fidelity with better support for faster moving stimuli. Although the highest attainable rate is desirable, when resources are limited, e.g., the battery life in portable devices, this adjustment becomes another variable in the optimization process balancing energy consumption and user satisfaction (e.g., Han et al., 2013; Egilmez et al., 2017).

The "window of visibility" in the spatiotemporal Fourier domain (Watson et al., 1986; Watson, 2013a) successfully described the fundamentals of sampling and motion fidelity, as a function of the stimulus spectrum and the spatiotemporal frequency detection limits in human vision. A sufficiently high frame rate would preserve the signal within this detection window, thus making any artifacts invisible to the human eye. Psychophysical studies have quantified the effect of frame rate on the visibility of several motion artifacts (Wang et al., 2009; Hoffman et al., 2011; Selfridge et al., 2016; Chapiro et al., 2019) and its impact on the perceived motion-image quality (Kuroki et al., 2007; Emoto et al., 2014; Nasiri et al., 2015; Mackin et al., 2019) and task performance (Chen and Thropp, 2007; Claypool and Claypool, 2007; Janzen and Teather, 2014; Kime et al., 2016). Although a link between fidelity and quality could be established under certain conditions (see Chapter 1), this is not always the case (Silverstein and Farrell, 1996). Fidelity is rather one of the several factors determining the perceived quality (see general introduction). Likewise, we may hypothesize that the absence of fidelity is one of the potential factors of visual discomfort and, eventually, fatigue. Currently, little is known about the development of these symptoms as a function of frame rate and motion distortion. Here, our purpose was two-fold: first, to investigate the effect of frame rate variations on fatigue and visual discomfort during regular display use; and second, to explore potential correlation with objective indicators that could yield quantifiable parameters.

The methods of assessing visual fatigue and discomfort can be divided into two classes: subjective and objective measures. Subjective evaluation is usually performed through self-assessment questionnaires on the severity of the most common symptoms, e.g., eye strain, dry eyes, and blurred vision. In contrast, objective assessment includes changes in the critical flicker-fusion threshold and oculometrics, such as blinking patterns, accommodative effects, and pupillometry (Sheppard and Wolffsohn, 2018). The latter approach is further supported by the fact that ocular responses have been widely used in the literature as indicators of physiological and psychological processes related to fatigue such as the state of arousal (Yoss et al., 1970; Häkkänen et al., 1999; Barbato et al., 2000; Caffier et al., 2003; Di Stasi et al., 2013) or vigilance (McIntire et al., 2014; Unsworth and Robison, 2016; Maffei and Angrilli, 2018). Generally, on the intersection of these studies with research on overall fatigue development (Stern et al., 1994a; Stern et al., 1994b; Morris and Miller, 1996; Summala et al., 1999; Van Orden et al., 2000; Kaneko and Sakamoto, 2001; Schleicher et al., 2008; Zargari Marandi et al., 2018) and eye health and comfort (Acosta et al., 1999; Wolkoff, 2005; Himebaugh et al., 2009; Gowrisankaran et al., 2012; Johnston et al., 2013; Portello et al., 2013; Jie et al., 2019) are measures related to blinking activity. It is widely accepted that spontaneous blinks are modulated by a complex interaction of extrinsic, e.g., the state of the ocular surface, and

intrinsic factors, e.g., mental activity (Cruz et al., 2011). Nevertheless, we hypothesized that even in complex interaction, frame rate effects on the observer would be predominantly reflected in their blinking patterns. Blink analysis was also used in studies exploring viewer stress levels in high-frame-rate video content (Tag et al., 2016) and the impact of the visual information presentation rate on fatigue and eye discomfort symptoms (Cardona et al., 2011).

Following this direction, we developed a naturalistic reward-based scan reading task where participants searched for errors in scrolling text while we recorded their eye activity using a portable binocular eye tracker. Four frame rates $\log_2$-linearly spaced between 30 and 240 frames/s were evaluated under two settings: self-paced unconstrained scrolling, and high fixed continuous text velocity. We assessed the perceived visual fatigue and discomfort symptoms, overall observer preference, task performance, and blinking activity. A possible association between the subjective reports and the objective measures was also examined.

## 3.2. Methods

### *3.2.1. Participants*

Twenty-three healthy volunteers (mean age: 31 years, SD: 8 years, span: 21 – 52 years, ten males and thirteen females) participated in experiment I. Fourteen of them were paid in vouchers based on task performance. Nine volunteers (mean age: 25 years, SD: 5 years, span: 21 – 32 years, seven males and two females) participated in experiment II. Six of them completed both experiments with at least a three-month interval between the first and the second experiment (also in order of execution). They had normal or corrected to normal vision with no history of visual problems. Six participants in the first group and one in the second wore contact lenses. The Newcastle University Ethics Committee approved the experimental protocol, and all participants gave written informed consent.

### *3.2.2. Apparatus and experimental design*

The experiments were performed on a 24.5" LCD monitor (Acer Predator XB252Q) in a dark and quiet room. The resolution of the panel was $1920\times1080$ pixels, with a maximum frame rate of 240 Hz. Participants were seated in front of a desk, and the monitor was fixed at 90 degrees. They were instructed to adjust the height of the display and the seat to their preferred position. The average viewing distance was approximately 40 cm, measured perpendicularly from the center of the screen to their eyes' midpoint. The subjects were fitted with a portable binocular eye-tracker (Pupil Labs Core) that recorded their eye activity at 120 Hz. The calibration procedure for the eye-tracker was performed using the Pupil Labs software (Kassner et al.,

2014). Subjects performed a scan reading task searching for scrambled words in scrolling text. They indicated their response with the press of the space bar on the keyboard. If a scrambled word was present inside the visible area, the response was recorded as correct. Participants received one point for locating a scrambled word, and one point was deducted for each false alarm. The task objective was to locate as many text errors as possible within the time frame. Paid participants were informed before the session that their accumulated score would specify their compensation reward amount. The size of the text canvas was set to 540×960 pixels encompassing an average visual angle of approximately 22×40 degrees (horizontal × vertical). The displayed font was Open Sans Semibold with a maximum line height of 25 pixels (approximately 1 visual degree). The light gray background was set to a luminance of 370 cd/m$^2$, as measured using the Minolta LS-100 meter. The dark text luminance level was .5 cd/m$^2$. We examined four conditions of decreasing degradation by modulating the frame update frequency. Explicitly, we emulated a frame rate of 30, 60, 120, and 240 frames/s. We selected four similar legal texts from the European Union law database EUR-Lex (2020), namely the documents with the CELEX unique identification numbers: (A) 12012A/TXT, (B) 32012R0966, (C) 32013R1303, and (D) 32013R1308. The preamble and tabular content were filtered out. Scrambled words were algorithmically generated and were the same for all subjects. The algorithm used a random variable character step size to move along the text and locate the nearest word of above six characters in length. A coin toss decided whether to scramble at this location or move to the next one. The scrambling process consisted of a split between consonants and vowels, a random swap of the consonants, and a merge of the two segments in series with a random order (e.g., European to rnpEuoea). The tedious text content was selected to encourage a quick pace and hence rapid motion, while task engagement was maintained using the reward system. Figure 3.1 shows an example of the experimental recordings.

#### *3.2.2.1. Experiment I*

The subjects were instructed to scroll through the text at their preferred speed using the mouse wheel (Dell Moczul). The mouse wheel sensitivity was set to 27 pixels per scroll event. The experiment consisted of a 5 min training session, followed by one 10 min trial for each condition. The subjects were allowed to recover after each trial without any time restriction. The four frame rates were randomly interleaved and randomly combined with the four texts. The experiment was completed in one visit lasting approximately 1 hour.

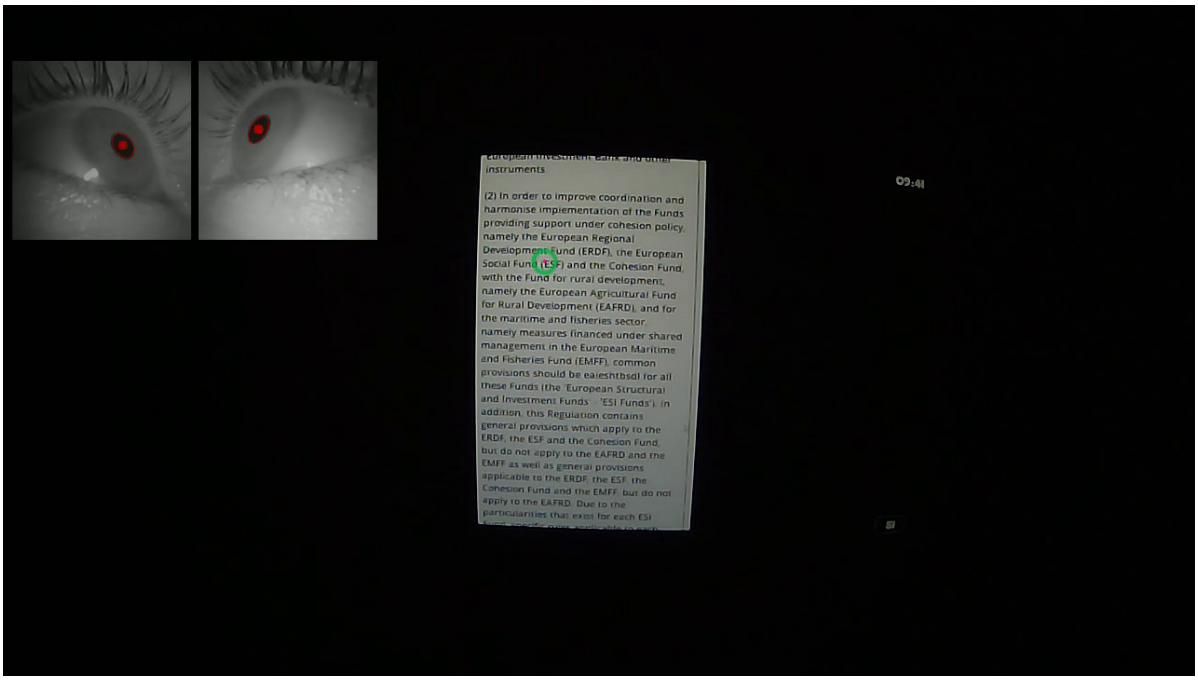


**Figure 3.1** First-person view of the experimental text scrolling task. **Top left**: the binocular eye recordings with the detected pupils shown as ellipses. **Middle**: the text canvas with the estimated gaze depicted as a circle. **Top right**: the elapsed time. **Bottom right**: the accumulated reward points

#### *3.2.2.2. Experiment II*

The second experiment was divided into two parts: speed-adjustment trials and paired comparisons. It consisted of a 5 min training session, followed by an 8 min set of adjustment trials, and an 18 min set of comparisons, with unrestricted rest intervals. In the adjustment trials, the subjects performed the same task on continuously moving text. They were instructed to adjust the text speed to their comfort level using the arrow keys on the keyboard. The initial text velocity was set to 390 pixels/s. They completed two 1 min trials per frame rate for a total of eight trials per subject, with 20 s minimum rest intervals. The conditions were randomly interleaved, and one randomly chosen text was assigned to all conditions. In the paired comparisons, the subjects performed the same task at a fixed text velocity of 390 pixels/s. Each participant completed two 90 s evaluations for each frame rate pair (45 s for each frame rate) for a total of twelve comparisons, with 20 s minimum rest intervals. The pairs were randomly interleaved and randomly combined with the four texts. For every text repetition, in both parts of the experiment, the task continued below the last visible line of the previous trial. The experiment was completed in one visit lasting approximately 50 min.

### 3.2.3. Subjective measures

Subjective feedback on visual fatigue, discomfort, and preference was recorded using the questionnaires in Hoffman et al. (2008). In the first questionnaire (Appendix C.1), the subjects indicated the severity of visual (questions 1, 2, 4, and 5) and musculoskeletal (question 3) symptoms on a five-level scale. The latter question was added by the authors to assess whether the responses were specific to the symptoms. In the second questionnaire (Appendix C.2), the subjects evaluated each pair of conditions in terms of higher visual fatigue and discomfort, and overall preference. They were given the options of strong, mild, or no difference. Participants were instructed not to take into consideration the text content in their evaluation.

In experiment I, the subjects completed the self-assessment of their symptoms before and after each trial. Each evaluation was performed for the last pair on every two consecutive trials. In experiment II, the self-assessment questionnaire was omitted due to the short duration of the trials. The subjects randomly evaluated all the combinations among conditions, from a lower to a higher frame rate and in reverse presentation order. The evaluation questionnaire (Appendix C.3) was supplemented with two questions on "text clarity" (question 4) and "perceived brightness" (question 5). The latter question was added due to comments received from some observers during experiment I, who reported that some conditions appeared brighter.

### 3.2.4. Objective measures

The potential objective indicators of visual fatigue and discomfort were arranged into two categories: ocular responses and task performance. In the first category, we examined the blink rate (BR), the duration of the blinks (BD), and the incomplete blink ratio (ICR). In the second category, we evaluated the individual performance in terms of hit rate (HT), precision (PPV), and adjusted text velocity (MV). The hit rate was defined as the number of hits (detected text errors) over total errors seen, and precision expressed the hits over the sum of hits and false alarms. The MV was calculated as the average scrolling speed during the trial. Note that in the second experiment, the MV corresponded to the preferred constant text speed that we measured during the speed-adjustment task at the beginning of the session (the text velocity was fixed at 390 pixels/s for all the observers in the paired frame rate evaluations). Although the MV can also be used as a predictor for the other responses, it was only treated as a response variable to avoid additional complexities since we did not systematically sample different text velocities. After visually inspecting the data, we set a high cut-off value for the blink duration at 600 ms corresponding to a data exclusion of 1.8%. In the blink rate calculation, multiple blinks in series were treated as one to avoid contamination by local bursts. All the blinks that were

classified as incomplete were excluded from further processing. This exclusion did not alter any of the conclusions as the ratio of incomplete blinks was roughly constant.

#### *3.2.4.1. Blink detection*

Eye blinks were detected using a bidirectional long short-term memory recurrent neural network following the basic architecture in Fogelton and Benesova (2018). The recorded image frames were down-sampled to 24x24 pixels, and optical flow vectors were calculated using the method in Farnebäck (2003). At each time-step, we constructed a feature vector consisting of the Pupil Labs detection algorithm confidence value and the optical flow vectors. The author annotated the eye-activity during the subjects' training trials by inspecting the ocular video recordings at controllable speed using custom software. From the annotated set, we extracted 3957 batches of 256 sample size intervals encompassing the blink events. The data were augmented by approximately 50% with random non-blink intervals and split for each observer and category (i.e., blink and non-blink) in 70/15/15 percent for training, validation, and testing, respectively. In the case of multiple blinks in series, each blink event was isolated, zero-padded, and then masked throughout the network. The model was initialized and trained as described in the original paper (Fogelton and Benesova, 2018). Each frame was classified in one of three classes, i.e., non-blink, incomplete blink, or complete blink. The frame-by-frame classification achieved a weighted average F1 score of .94; a complete report is provided in Table C.1. Practically, the network detected 98.96% of the annotated blinks (total number of events: N = 575) in the test set, with a false alarm rate of 1.49% (N = 269), for a minimum blink duration of 50 ms. In terms of blink completeness, 90.31% (N = 413) of the complete blinks were classified in agreement with the human annotator, with a 9.26% (N = 162) of false positives, for a minimum complete closure duration of 10 ms. A full blink cycle occurs when the upper and lower eyelids come in contact. Here, we adopted a lower threshold of the upper lid covering approximately two-thirds of the iris. The network was implemented using the Keras Python library (Chollet et al., 2015) with TensorFlow (Abadi et al., 2015) as a backend.

The onset and offset of the blink intervals were post adjusted based on the baseline-corrected mean vertical component of the optical flow vectors. Baseline activity was extracted from the non-blink intervals of the annotated trials, and its mean value was subtracted from the signal. The signal was then passed through a median filter of window size three. During a blink, the vertical component exhibits a specific pattern: a positive peak followed by a negative peak corresponding to a downward and upward eyelid movement. Thus, the onset was set to the nearest local minimum or zero-crossing preceding the global maximum of the mean vertical component. Likewise, offset was set to the nearest local maximum or zero-crossing following

the global minimum at a minimum distance of frames (set after visually inspecting the isolated signal segments). If sequential blinks were present, they were visually inspected and then split at the location of the minimum standard deviation of the vertical flow. Compared to the human annotator, there was an average latency of 2 ms (SD = 12 ms) and -43 ms (SD = 34 ms) in the onset and offset detection, respectively. This adjustment resulted in an average decrease of 45 ms (SD = 40 ms) in the blink duration among participants. The process was run for each eye separately, and the signals were downsampled and synced to 60 Hz. The binocular combination was performed by calculating the intersection over union (IoU) of the overlapping blink segments. If the IoU was above a threshold, here set to .2, we adjusted the onset and offset to the midpoints of the difference between the two segments (Fogelton and Benesova, 2016). If a blink was detected or classified as incomplete in only one of the eyes, the corresponding frames were extracted and visually inspected. Finally, blink and non-blink intervals were randomly extracted for each subject and inspected for false positives and negatives.

### *3.2.5. Statistical analysis and modeling*

A within-subjects ANOVA was conducted on each of the objective measures with the frame rate as independent predictor. The tests were performed using the R libraries afex (Singmann et al., 2020) and car (Fox and Weisberg, 2019). The frame rate effects were reported with the Geisser-Greenhouse's $\hat{\varepsilon}$ adjustment for the degrees of freedom. Follow-up tests for linear and quadratic trends were performed using the library emmeans (Lenth, 2020) with a multivariate model. In the second experiment, the data were pooled across repetitions. The statistical sensitivity of our tests is given in the Appendix C.

The subjective feedback on fatigue and discomfort was analyzed using Friedman's test on the self-reported increase in each of the symptoms. The paired comparisons on the preference among the frame update frequencies were fitted with a log-linear Bradley-Terry model (Dittrich et al., 2002b). This modeling approach aims to place the different objects, here the frame rates, on a unified preference interval scale based on the responses for the tested attribute. The estimated non-negative parameters or the objects' worth, i.e., the location on the preference scale, are associated with the probabilities of selecting one object over another. We fitted the model on the overall preference using the R library prefmod (Hatzinger and Dittrich, 2012). The data were grouped into three response categories: preference for the first object, the second, or none, to reduce sparseness. Note that due to the within-subject dependencies, an underestimation of the standard errors is to be expected. Thus, we also examined an alternative model formulation accounting for the dependencies in the response patterns (Dittrich et al., 2002a). We did not find any differences in the object ranking or the significance tests (see

Figure C.9); therefore, we report the results of the former due to better model fits. For convenience, as the estimated standard errors are calculated with reference to a base frame rate, the coefficients were plotted using quasi-standard errors (Firth, 2004). However, the test statistics were performed on the estimated covariance matrix. The remaining feedback questions on the specific symptoms or attributes were analyzed in terms of their repeated measures correlation ($r_{rm}$; Bakdash and Marusich, 2017) with overall preference.

## 3.3. Results

### *3.3.1. Task performance*

Task performance can be quantified as speed, hit rate, and precision. Table 3.1 summarizes the mean estimates for each frame rate and variable. Figure 3.2 depicts the normalized to each observer frame rate differences. Recall that in the first experiment, the participants scrolled through the text (potentially in a "start-stop" way), whereas in the second experiment, we measured the preferred constant scrolling speed at the beginning of the session (the speed-adjustment task; the text velocity was fixed at 390 pixels/s for all the observers in the paired frame rate evaluations where all the other variables were measured).

The mean text scrolling velocity (**MV**; right column in Figure 3.2) was relatively constant in the first experiment with no significant changes among the frame update frequencies, F(2.25, 49.55) = .05, p-corrected = 1, $\eta_p^2$ = .002 (see also Table 3.2). In contrast, during the second experiment's speed-adjustment task, the subjects fixed the MV to higher levels by roughly 18 pixels/s for every two-fold increase in the frame rate, $CI_{95\%}$:[3, 32], before the adjusted velocity began to saturate between 60 frames/s and 120 frames/s. However, the overall effect was practically small and not significant in our sample, F(2.16, 17.26) = 3.41, p-corrected = .267, $\eta_p^2$ = .299. The average MV in the aggregated across conditions data (M = 186, SD = 80 pixels/s) was 52% lower, $CI_{95\%}$:[46%, 59%] by basic bootstrap (boot), than the one we arbitrarily set for the main evaluation task, but 57% higher, boot $CI_{95\%}$: [20%, 69%], than experiment I (M = 119, SD = 65 pixels/s). The latter was expected by design due to the "start-stop" scrolling behavior in the first experiment.

Turning to the error detection performance, the frame rate effect on the first experiment's hit rate (**HT**; left column in Figure 3.2) was practically negligible and not significant, F(2.73, 60.08) = .09, p-corrected = 1, $\eta_p^2$ = .004 (see Table 3.2). In the second experiment, similar to the preferred text speed, the two-fold increase in frequency improved the hit rate by roughly .02, $CI_{95\%}$:[−.01, .06], before performance began to asymptote above 60 frames/s. However, the effect was still practically small and not significant, F(2.50, 20.03) = 2.22, p-corrected =

.502, $\eta_p^2$ = .217. Likewise, precision (**PPV**; middle column in Figure 3.2) did not significantly differ among frame rates in either the first, F(2.55, 56.20) = 1.95, p-corrected = .594, $\eta_p^2$ = .081, or the second experiment, F(2.17, 17.38) = .41, p-corrected = 1, $\eta_p^2$ = .049. The average HT in the aggregated data in experiment II (M = .24, SD = .16) dropped by 58%, boot $CI_{95\%}$:[45%, 70%], compared with experiment I (M = .57, SD = .24). The participants were conservative in both cases, with the average PPV in the second experiment (M = .92, SD = .06) decreasing by 5%, boot $CI_{95\%}$:[3%, 7%], relative to experiment I (Mdn = .97, SD = .04).

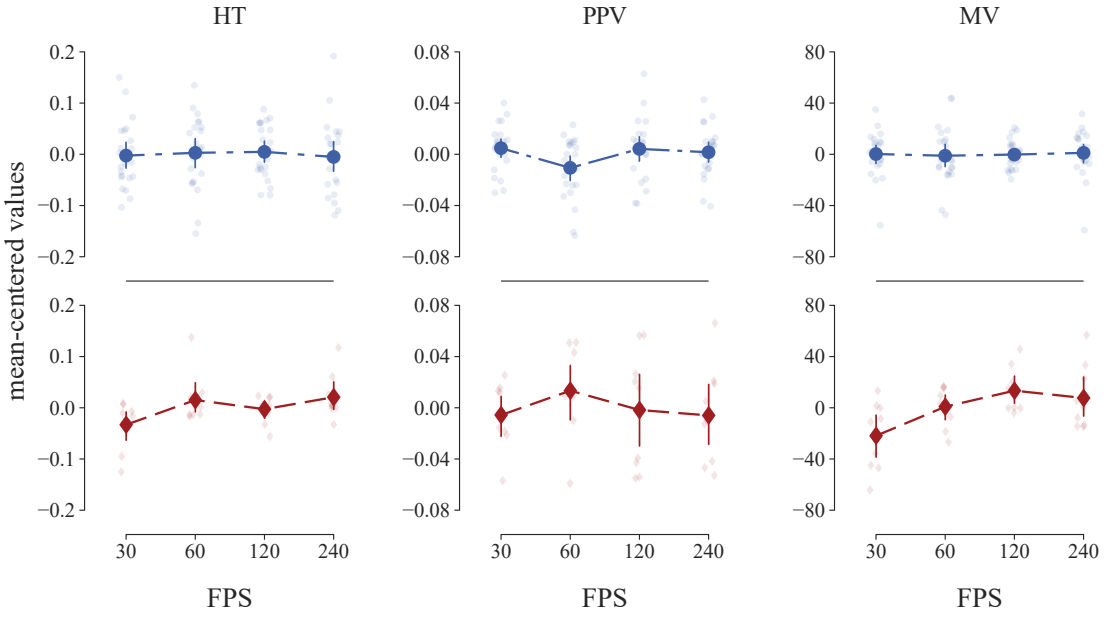


**Figure 3.2** The normalized to each observer frame rate differences in the task performance (mean-centered values for each observer). From left to right, the hit rate (HT), the precision (PPV), and the adjusted text velocity (MV). **Top**: experiment I. **Bottom**: experiment II. The MV in the second experiment corresponds to the speed-adjustment task before the paired frame rate evaluations. The error bars show the bootstrap 95% CI of the average value across observers (large markers). The small transparent markers depict the individual observations (horizontally jittered)

### *3.3.2. Eye blinks*

Table 3.1 summarizes the blinking measures for each frame rate condition. The normalized to each observer differences are illustrated in Figure 3.3. In the first experiment, the blink rate (**BR**; left column in Figure 3.3) decreased approximately linearly (see significant linear trend in Table 3.3) by .66 blinks/min for every two-fold increase in the frame rate. However, the effect was practically negligible and not significant in our sample after adjusting for the multiplicity, F(2.78, 61.15) = 3.19, p-corrected = .197, $\eta_p^2$ = .127. (see Table 3.2). In the second and more demanding experiment, the frame rate effect was considerably larger and significant, F(2.06,

16.44) = 7.07, p-corrected = .035, $\eta_p^2$ = .469, albeit still practically moderate. Consistent with experiment I, the blink rate decreased approximately linearly (the linear contrast was also significant) by 1.45 blinks/min for every two-fold increase in the frame rate, $CI_{95\%}$:[.56, 2.34]. The estimated effects are depicted in Figure 3.4. It is worth noting that a paired t-test on the post-trial blink rate in the first experiment revealed a significant mean relative increase of 21.2 blinks/min during the feedback reports, t(91) = 12.84, p < .001, boot $CI_{95\%}$:[17.9, 24.6].

The effect of frame rate on blink duration (**BD**; middle column in Figure 3.3) was trivial, with no significant differences in either the first, F(2.20, 48.44) = 2.19, corrected-p = .594, $\eta_p^2$ = .090, or the second experiment, F(2.04, 16.30) = .34, corrected-p = 1, $\eta_p^2$ = .040 (Table 3.2). Likewise, any differences in the incomplete blink ratio (**ICR**; right column in Figure 3.3) were practically negligible and non-significant in both the first, F(2.42, 53.26) = .01, corrected-p = .722, $\eta_p^2$ = .062, and the second experiment, F(2.40, 19.21) = .10, corrected-p = 1, $\eta_p^2$ = .013. The average BD in the aggregated data in experiment II (M = 301, SD = 50 ms) increased by roughly 18%, boot $CI_{95\%}$:[10%, 26%], relative to experiment I (M = 255, SD = 62 ms). Likewise, the average ICR in the second experiment (M = .31, SD = .17) increased by approximately 56%, boot $CI_{95\%}$:[23%, 89%], relative to experiment I (M = .20, SD = .18).

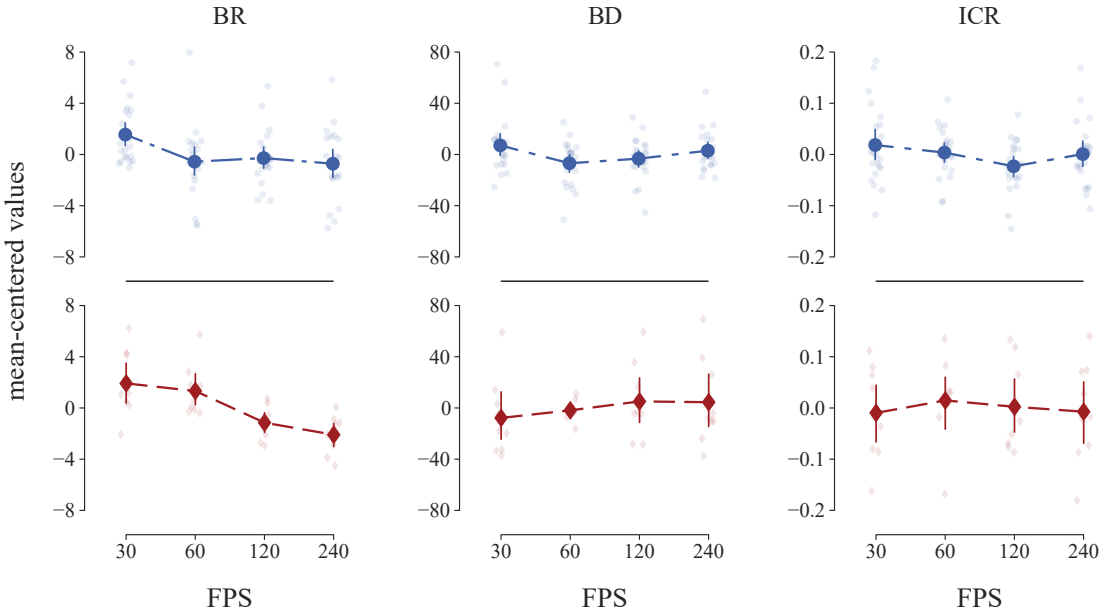


**Figure 3.3** The normalized to each observer frame rate differences in blinking activity. From left to right, the blink rate (BR), the blink duration (BD), and the incomplete blink ratio (ICR). **Top**: experiment I. **Bottom**: experiment II. The error bars show the bootstrap 95% CI of the average value across observers (large markers). The small transparent markers depict the individual observations

| Dependent variable | FPS | Experiment I (df = 22) | | | Experiment II (df = 8) | | |
|---|---|---|---|---|---|---|---|
| | | Mean | SE | 95% CI | Mean | SE | 95% CI |
| **BR** [min$^{-1}$] | 30 | 11.56 | 1.86 | 7.71 – 15.42 | 7.95 | 1.68 | 4.08 – 11.83 |
| | 60 | 9.44 | 1.52 | 6.30 – 12.58 | 7.36 | 1.88 | 3.03 – 11.68 |
| | 120 | 9.73 | 1.66 | 6.29 – 13.16 | 4.88 | 1.22 | 2.06 – 7.70 |
| | 240 | 9.28 | 1.37 | 6.43 – 12.13 | 3.94 | 1.00 | 1.64 – 6.24 |
| **BD** [ms] | 30 | 262 | 14 | 234 – 291 | 293 | 16 | 255 – 331 |
| | 60 | 248 | 12 | 223 – 274 | 299 | 15 | 265 – 334 |
| | 120 | 252 | 12 | 226 – 278 | 306 | 17 | 268 – 344 |
| | 240 | 258 | 14 | 230 – 287 | 306 | 20 | 259 – 352 |
| **ICR** | 30 | .22 | .04 | .13 – .30 | .30 | .04 | .21 – .39 |
| | 60 | .20 | .04 | .13 – .27 | .32 | .05 | .20 – .44 |
| | 120 | .17 | .03 | .11 – .24 | .31 | .07 | .15 – .46 |
| | 240 | .20 | .04 | .11 – .29 | .30 | .07 | .14 – .46 |
| **HT** | 30 | .57 | .05 | .46 – .67 | .21 | .06 | .08 – .34 |
| | 60 | .57 | .05 | .46 – .68 | .26 | .06 | .11 – .41 |
| | 120 | .57 | .05 | .48 – .67 | .24 | .05 | .12 – .36 |
| | 240 | .56 | .05 | .45 – .67 | .26 | .05 | .14 – .38 |
| **PPV** | 30 | .97 | .01 | .96 – .99 | .91 | .02 | .87 – .96 |
| | 60 | .96 | .01 | .94 – .98 | .93 | .02 | .89 – .98 |
| | 120 | .97 | .01 | .96 – .99 | .92 | .02 | .87 – .97 |
| | 240 | .97 | .01 | .95 – .99 | .91 | .02 | .88 – .95 |
| **MV** [pixels/s] | 30 | 119 | 13 | 91 – 146 | 164 | 21 | 116 – 211 |
| | 60 | 117 | 13 | 91 – 144 | 186 | 27 | 125 – 248 |
| | 120 | 118 | 14 | 88 – 148 | 199 | 32 | 126 – 272 |
| | 240 | 120 | 14 | 90 – 149 | 193 | 30 | 124 – 263 |

**Table 3.1** The estimated means of the objective variables and their 95% CI. **BR**: blink rate. **BD**: blink duration. **ICR**: incomplete blink ratio. **MV**: text scrolling velocity. **HT**: hit rate. **PPV**: precision. **FPS**: the frame rate. Note that multiple blinks in series were merged into one for calculating the blink rate

| Dependent variable | df | MSE | F | $\eta_p^2$ | [90% CI] | p-val | $\text{Holm}_3$ p-val | $\text{Holm}_6$ p-val |
|---|---|---|---|---|---|---|---|---|
| **Eye blinks I** | | | | | | | | |
| BR | (2.78, 61.15) | 8.67e+0 | 3.19 | .127 | [.006, .232] | .033 | .099 | .197 |
| BD | (2.20, 48.44) | 5.57e+2 | 2.19 | .090 | [.000, .206] | .119 | .238 | .594 |
| ICR | (2.42, 53.26) | 5.80e−3 | .01 | .062 | [.000, .157] | .241 | .241 | .722 |
| **Task performance I** | | | | | | | | |
| HT | (2.73, 60.08) | 6.01e+1 | .09 | .004 | [.000, .019] | .958 | 1 | 1 |
| PPV | (2.55, 56.20) | 7.20e−4 | 1.95 | .081 | [.000, .181] | .140 | .421 | .594 |
| MV | (2.25, 49.55) | 5.59e+2 | .05 | .002 | [.000, .023] | .966 | 1 | 1 |
| **Eye blinks II** | | | | | | | | |
| BR | (2.06, 16.44) | 6.89e+0 | 7.07 | .469 | [.110, .619] | .006 | .017 | **.035** |
| BD | (2.04, 16.30) | 1.45e+3 | .34 | .040 | [.000, .175] | .723 | 1 | 1 |
| ICR | (2.40, 19.21) | 1.32e−2 | .10 | .013 | [.000, .078] | .930 | 1 | 1 |
| **Task performance II** | | | | | | | | |
| HT | (2.50, 20.03) | 2.85e−3 | 2.22 | .217 | [.000, .385] | .126 | .251 | .502 |
| PPV | (2.17, 17.38) | 2.51e−3 | .41 | .049 | [.000, .184] | .686 | .686 | 1 |
| MV | (2.16, 17.26) | 8.74e+2 | 3.41 | .299 | [.000, .478] | .053 | .160 | .267 |

**Table 3.2** Statistics of the frame rate effect on the objective measures. The last three columns show the unadjusted and the Holm corrected p-values for the three and the total six tests in each experiment (the adjustment we selected). For the calculation of the confidence intervals, see Smithson (2001)

| Dependent variable | Contrast | Est. | SE | t-ratio | Lower $\text{CI}_{95\%}$ | Upper $\text{CI}_{95\%}$ | p-value |
|---|---|---|---|---|---|---|---|
| BR (**I**) | linear | −.66 | .26 | −2.51 | −1.20 | −.11 | **.020** |
| df = 22 | quadratic | .42 | .30 | 1.41 | −.20 | 1.04 | .174 |
| BR (**II**) | linear | −1.45 | .39 | −3.74 | −2.34 | −.56 | **.006** |
| df = 8 | quadratic | −.09 | .34 | −.25 | −.88 | .71 | .810 |

**Table 3.3** Follow-up tests for linear and quadratic contrasts. The contrast coefficients were scaled using the codings [-.3, -.1, .1, .3] for the linear and [.25, -.25, -.25, .25] for the quadratic effects

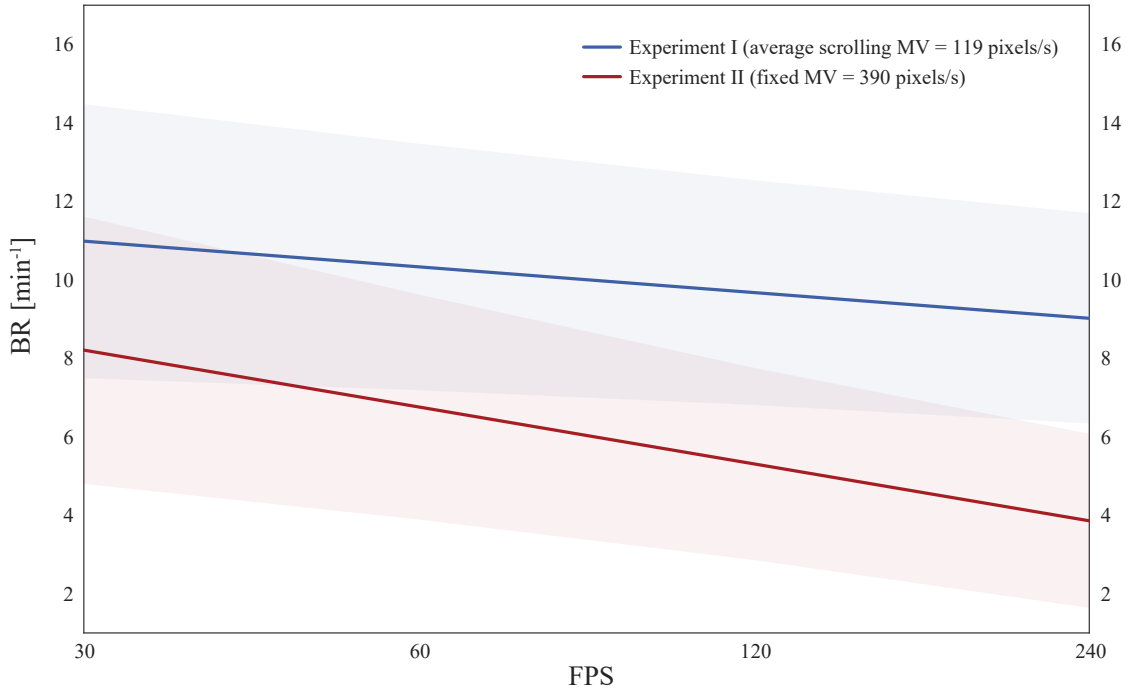


**Figure 3.4** The estimated effect of the frame update frequency on the blink rate (BR; considering also the variability in the observer intercepts). The blue line (above) and the red line (below) depict experiments I and II, respectively. The shaded areas depict the 95% CIs. The coefficients of separate individual linear regressions on the individual data are shown in Figure C.4

### *3.3.3. Subjective feedback*

The self-assessment report on fatigue and discomfort was generally null, i.e., no symptom increase, with zero and negative responses (when corrected for the pre-trial measurement) ranging between 65% and 96% among the feedback questions and frame rates (see Figure C.6). Recall that in experiment I, the participants were asked after each condition to rate five symptoms of discomfort. Neither overall eye tiredness, $\chi^2(3) = 1.42$, $p = .702$, nor eye strain, $\chi^2(3) = 3.19$, $p = .364$, nor vision blurriness, $\chi^2(3) = 3.22$, $p = .359$, were significantly affected by the variations in the frame rate. Likewise, there was no significant effect on headache symptoms, $\chi^2(3) = 4.40$, $p = .221$, or neck and back soreness, $\chi^2(3) = .90$, $p = .827$.

In both experiments, observers were also asked to make direct comparisons between pairs of trials. Recall that in the first experiment, each evaluation was performed for the last pair on every two consecutive trials, whereas, in experiment II, the subjects assessed all the combinations among the conditions (from a lower to a higher frame rate and in reverse presentation order). The aggregated votes in the first and the second experiment are depicted in Figures 3.5 and 3.6, respectively. Qualitatively, the first experiment's feedback was inconclusive without

considerable differences in the paired comparisons. In contrast, the more dynamic second experiment revealed a generally clear preference with higher perceived text clarity for the higher frames rates (bottom and middle right panels dominated by blue bars in Figure 3.6). Participants also generally reported higher fatigue and eye irritation effects for the lower frame rates (top panels dominated by pink bars in Figure 3.6).

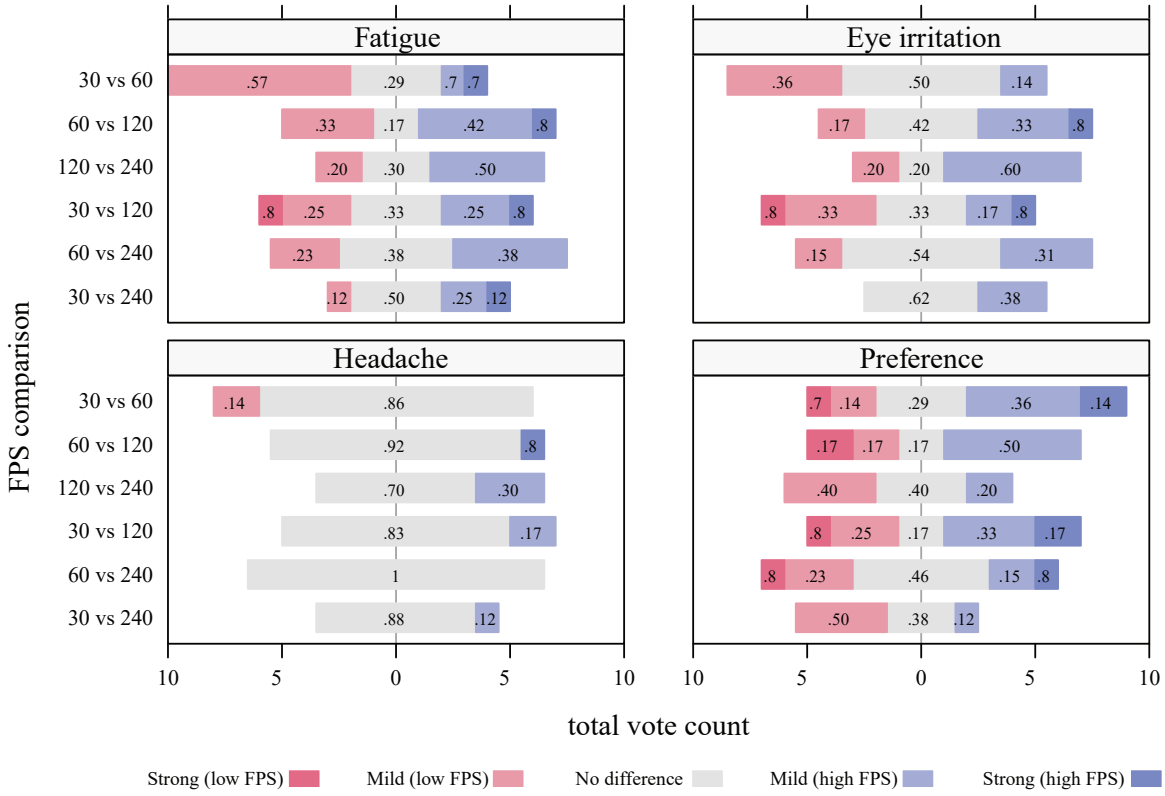


**Figure 3.5** Aggregated votes for the observer evaluation feedback in Experiment I. Note that for the fatigue, eye irritation, and headache questions, a mild or strong selection corresponds to the severity of the symptoms instead of preference, e.g., stronger symptoms in the lower frame rate (see Figure C.2). The data were plotted using the R library HH (Heiberger and Robbins, 2014)

Figure 3.7 illustrates the log model estimates based on the overall observer preferences in the paired evaluations feedback, with the 30 frames/s as the base rate. In experiment I, the coefficients increased with the frame rate up to 120 frames/s before they dropped below the base rate value. However, none of the differences were significant (right panel in Figure 3.7). The estimated worth parameters, given as the exponents of two times the coefficients over the sum of these exponents, were $\pi_{30} = .21$, $\pi_{60} = .28$, $\pi_{120} = .36$, and $\pi_{240} = .16$. The global no preference coefficient ($\gamma = -.10$) revealed a relatively balanced discriminability, and it was not significant (Wald p = .700). Generally, lower negative values indicate a higher tendency

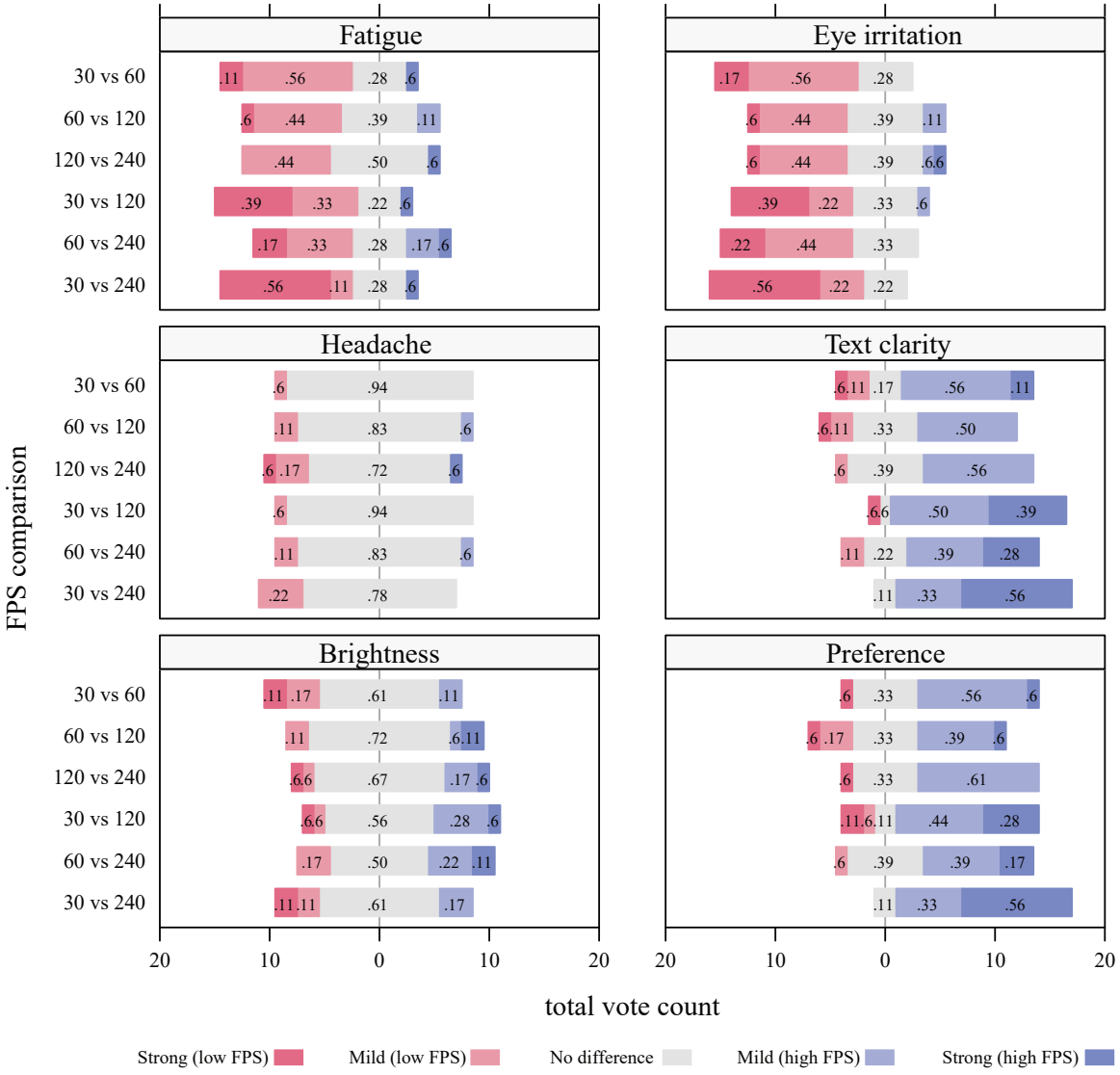


**Figure 3.6** Aggregated votes for the observer evaluation feedback in Experiment II. Note that for the fatigue, eye irritation, and headache questions, a mild or strong selection corresponds to the severity of the symptoms instead of preference, e.g., stronger symptoms during the low frame rate. Likewise, a strong low frame selection for brightness indicates that the lower frame rate was perceived as much brighter. The questionnaire is shown in Figure C.3

| Feedback question | Experiment II | | | | |
|---|---|---|---|---|---|
| | Fatigue | Eye irritation | Headache | Preference | Brightness |
| Text clarity | -.76*** [-.86, -.67] | -.79*** [-.89, -.69] | -.22 [-.32, -.10] | .83*** [.73, .92] | -.06 [-.28, .20] |
| Fatigue | | .78*** [.61, .92] | .30* [.16, .41] | -.72*** [-.84, -.60] | -.04 [-.26, .21] |
| Eye irritation | .81*** [.64, .93] | | .34** [.20, .45] | -.74*** [-.86, -.61] | .06 [-.21, .31] |
| Headache | .47** [.09, .72] | .66*** [.40, .82] | | -.33** [-.45, -.20] | -.08 [-.28, .14] |
| Preference | -.74*** [-.86, -.55] | -.74*** [-.85, -.56] | -.52** [-.70, -.16] | | -.11 [-.29, .09] |
| | Experiment I | | | | |

Notes: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

**Table 3.4** Repeated measures correlation matrix of the evaluation feedback questionnaire. The lower triangular matrix shows the $r_{rm}$ correlation coefficients and their 95% CI by bootstrap (within brackets) for experiment I. Likewise, the upper triangular matrix depicts the coefficients and their 95% CI for experiment II. The stars indicate the Holm corrected p-values given twenty-one tests

towards a decision, whereas higher positive values express a larger indecision effect. The overall preference anticorrelated at the intra-individual level with the perceived fatigue, $r_{rm}(45)$ = -.74, p < .001, eye irritability, $r_{rm}(45)$ = -.74, p < .001, and headache symptoms, $r_{rm}(45)$ = -.52, p = .002. In experiment II, the model coefficients increased for increasing frame rate, with all the differences except for the 60 – 120 frames/s pair being significant in our sample. The estimated worth parameters were $\pi_{30}$ = .01, $\pi_{60}$ = .09, $\pi_{120}$ = .12, and $\pi_{240}$ = .78. The global no preference estimate ($\gamma = .20$) was comparable to the first experiment with a non-decision tendency, but it was also not significant (Wald p = .432). The overall preference anticorrelated with the perceived fatigue, $r_{rm}(98)$ = -.72, p < .001, eye irritability, $r_{rm}(98)$ = -.74, p < .001, and headache symptoms, $r_{rm}(98)$ = -.33, p = .007, and correlated with text clarity $r_{rm}(98)$ = .83, p < .001. Table C.2 shows the predicted probabilities and their relation to the estimated worth parameters for each comparison in the two models. The individual preference data are shown in Figures C.7 and C.8. The full repeated measures correlation coefficients and their 95% CI are given in Table 3.4. The reported p-values for each correlation were adjusted for the total of twenty-one tests using the Holm method.

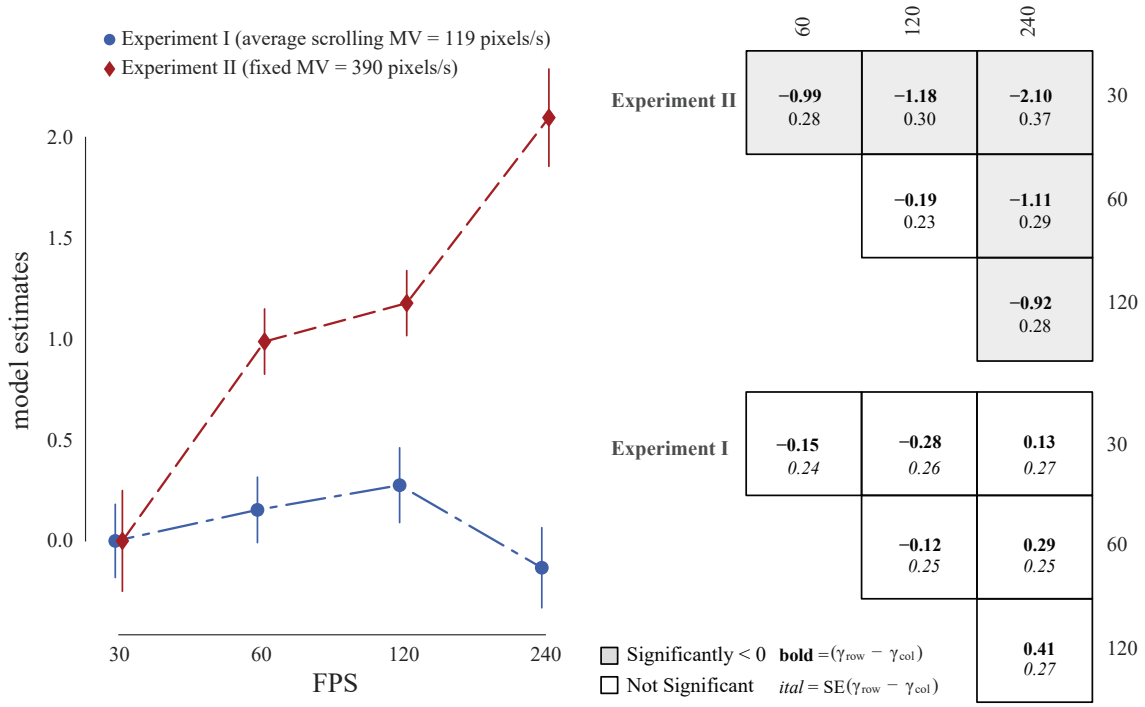


**Figure 3.7** Scaling of the frame rate conditions based on the overall observer preference. **Left**: the log estimates as a function of frame rate (FPS) with the 30 frames/s as the reference, for experiments I (blue circles) and II (red diamonds). The error bars depict the quasi standard errors. **Right**: the difference in the estimates (bold) and the standard error (italics) among the frame rate pairs for experiments I (bottom) and II (top). A negative sign indicates that the column was preferred to the row. The shaded blocks indicate p-values below .05. The data were plotted using the R library factorplot (Armstrong, 2016)

## 3.4. Discussion

This chapter analyzed the effects of frame rate on visual discomfort, fatigue, and overall preference during regular display use, and their association with real-time objective indicators in task performance and blinking activity. The observers performed a typical scan reading task looking for typographical errors in a lengthy text document while we modulated the frame rate. The study comprised two experiments. In the first experiment, the subjects scrolled through the text at their own pace using a mouse wheel, whereas, in the second, the text velocity was fixed at a relatively high value increasing the task difficulty (the preferred fixed text speed was recorded at the beginning of the experiment). Given the exploratory nature of this study and the currently limited knowledge of the specific effects, the purpose of the second approach was to establish a basis for evaluating the results. We hypothesized that if there were an effect due to the varying frame rate, it would be amplified by further increasing degradation and eye-pursuit movements through an accelerated continuous text flow. Smooth eye pursuits are

associated with the most prominent artifact in hold-type displays, hold-type or motion blur (see an example in Figure 3.8), which adversely affects visual annoyance (Kurita, 2001), particularly in the presentation of scrolling text sequences (Tourancheau et al., 2009).

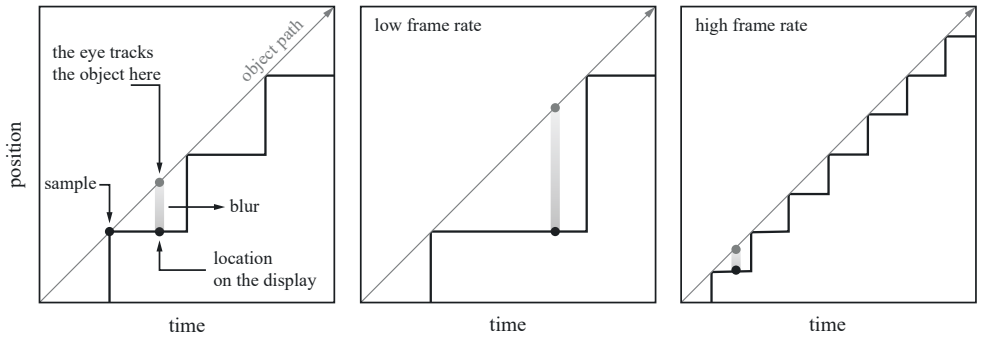


**Figure 3.8** An illustration of hold-type or motion blur for an object translating with constant velocity (in our case the scrolling text). The eye is smoothly tracking the object but the display samples and holds a static frame for the duration of the frame update interval. This positional error results in the object image getting smeared out on the retina. A higher frame rate produces a lower error and thus less blur

The subjective evaluation in the first experiment did not reveal a significantly higher preference for any of the frequencies. Previous studies on user satisfaction with handheld devices operating under different frame rates presented comparable findings. Han et al. (2013) reported that 80% of the users were satisfied scrolling through a webpage at approximately 25 frames/s for a scrolling speed of 45 pixels/s (about half the value of the median observer in our study), increasing to 41 frames/s at 200 pixels/s. Along similar lines, Egilmez et al. (2017) found almost no significant increase in user ratings above 45 frames/s on several applications, surprisingly, including a fast-paced video game, with 30 frames/s becoming more acceptable with less dynamic content. We also found that subjective preference anticorrelated with the perceived difference in fatigue and eye discomfort symptoms. Indeed, one purpose of including these questions was to guide the participant to consider fatigue and eye comfort when selecting their preference, and this association suggests that we achieved this.

Consistent with the paired evaluations, the self-assessment questionnaire did not show any significant increase in the symptoms for a 10 min trial duration at each frame rate condition. Considering the large number of zero responses throughout the first experiment, it follows that this duration was somewhat short or the task not strenuous enough to induce such typically long-timescale effects. Note, though, that it is not unusual to observe a non-significant change in the subjective reports even when objective metrics indicate otherwise (Lin et al., 2008). Regardless, one usually has to demonstrate a substantial effect (detectable even with relatively small sample sizes) to convince that at least a two-fold increase would be necessary within this

context. Based on these results, in the second experiment, we allocated the remaining resources to improve the evaluation feedback reliability by increasing the number of paired comparisons among the conditions (i.e., all the frame rate pairs in both orders of presentation to extract the full comparisons matrix for each observer). Consequently, the symptoms questionnaire was omitted, and the trial duration was reduced to maintain the total time-on-task at roughly the same level. Although an investigation of long-term effects would be equally important, given that the self-reported increase in the symptoms after 40 min on the task was negligible, for practical reasons, we limited our current exploration to short-timescale frame rate effects.

Under these conditions in the second and more demanding experiment, the changes in the overall preference with frame rate became apparent. The subjects reported a higher preference monotonically increasing with the frame rate up to 240 frames/s, with only the 60-120 frames/s pair being relatively close on the estimated scale. Although the sample size was relatively small for this type of analysis, the data were consistent across participants, and the scaling results were generally unambiguous. This considerable variation in the frame rate threshold with changes in the stimulus content and particularly its velocity can also be seen in studies evaluating the perceived quality of video at high frame rates where improvements up to 240 frames/s were previously reported (Kuroki et al., 2007; Emoto et al., 2014). Recall that there was a core difference between the two experiments; here, the text was continuously moving, and the image was updated at the corresponding intervals without the observer's control. Instead, in the first experiment, the image was abruptly updated at the corresponding locations based on the accumulated discrete scroll events and the frame rate. Had we simulated a smooth movement, or the momentum in the user input, e.g., by using a touch-screen, or any other common modifications such as scrolling acceleration, the results would have presumably differed. Note that we can easily create the conditions where a higher frame rate would be needed; however, testing all the possible scenarios was outside the purpose of this study.

Same as in the first experiment, the preference anticorrelated with the perceived fatigue and eye discomfort effects. However, the slightly higher correlation with "text clarity" suggests that the subjects might have responded more generally or that their decision was based on an interaction of factors, including the motion-image quality. We could reasonably assume that a trial of 45 s is rather short to induce actual fatigue symptoms, yet the responses were similar to the perceived eye discomfort and text clarity. This inconsistency, even if it stems from a perceived visual fatigue potential, demonstrates the challenges of decoupling these factors with naive observers and further supports the need for objective indices.

The error detection performance and the adjusted text velocity were not significantly affected by the variations in the frame rate, although a practically small effect was present in the

second experiment (primarily at 30 frames/s where the task demands were higher). As expected, due to the increase in the task difficulty, the overall hit rate dropped considerably from the first to the second experiment, with a higher number of false alarms and lapses. The speed-adjustment trials at the beginning of the second experiment confirmed that the fixed text velocity in all cases was well above the observer's comfort level, approximately two times higher than the average value. This difference was even higher compared with the first experiment, roughly three times above the average observer, due to the unrestricted discontinuous scrolling (i.e., scrolling in a "start-stop" way). Therefore, the pre-set speed was sufficiently low to accomplish the task adequately even at the lowest frame rate, where the hold-type blur was most severe, but high enough to stress the participants. Increasing the text speed further would have eventually impaired task execution at the lower frequencies; however, detecting the threshold performance as a function of frame rate was outside the purpose of this study. Although there was no clear indication of a link between visual fatigue or discomfort with task performance, the examined conditions were too tightly constrained to draw any further conclusions.

The blinking activity generally corresponded to the overall observer preference. In both experiments, the blink rate increased asymptotically with the frame rate, approximately linearly in log scale within the examined range. However, in the first experiment, this increase was practically negligible (maximum mean difference of approximately 2 blinks/min). The corresponding effect in the second experiment was considerably larger (roughly doubled) but still relatively moderate. An increase in blink rate was previously associated with overall fatigue development based on the axiomatic relationship of the latter with increasing time-on-task (Kaneko and Sakamoto, 2001; Zargari Marandi et al., 2018), often combined with decrements in task performance (Morris and Miller, 1996; Van Orden et al., 2000; McIntire et al., 2014). It naturally follows that this behavior is not restricted to visual fatigue, but it could also reflect other processes (Stern et al., 1994b). Nonetheless, the impact of time-on-task effects was negated in our experiments by the short duration of the randomly interleaved trials interrupted by resting periods restricted only to a minimum interval (Stern et al., 1994a). Therefore, it would be misleading to conclude that this increase in blink rate was solely due to higher fatigue and particularly of visual origin. However, visual fatigue (here measured only subjectively) is not always bi-directionally coupled with visual discomfort. Borrowing a definition by Lambooij et al. (2009), visual fatigue refers to a decrement in visual function that may be accompanied by visual discomfort, whereas the nature of visual discomfort is purely subjective. Hence, the latter may also occur during short exposure to an irritating stimulus without necessarily inducing fatigue. In our study, there was some evidence of higher subjective discomfort with decreasing frame rate as expressed through the association with the overall observer preference.

In agreement with our results, Tag et al. (2016) found a decreasing blink count with increasing frame rate in videos with various motion complexities, which they attributed to lower viewer stress levels. It is not illogical to hypothesize that viewer discomfort could lead to moderate psychological stress that, through a consequent increment in dopamine levels (Pruessner, 2004), results in increased blinking activity (Jongkees and Colzato, 2016) — all other things being equal. On the other hand, several studies have linked blink inhibition to ocular surface dehydration that may predispose individuals to visual discomfort symptoms (Wolkoff, 2005; Himebaugh et al., 2009). Given that there was generally no indication of discomfort with increasing frame rate, the latter, if we assume an effect, would have smoothed out the differences by bouncing back the blink count. However, the short trial duration and the frequent resting periods could have also mitigated these symptoms.

Regardless, this complex interaction suggests a mechanism that strategically modulates this behavior to balance these effects while maintaining the continuity of critical visual information, which is naturally interrupted during blinking. Thus, it may also be argued that the blinking frequency was merely adapted to the flow increment of visual information. Previous studies found higher blink inhibition for more dynamic display content (Himebaugh et al., 2009; Cardona et al., 2011), although with vast differences among the tasks in comparison. Whether a maximum lag of 29 ms in the presentation rate (maximum difference among the frame update frequencies) would have been sufficient to affect the blinking behavior and to what extent, to the author's knowledge, is currently unknown. Nevertheless, the overall effect of frame rate on the blinking activity compared with other task-dependent (average blink rate increase of 21 blinks/min during the feedback questionnaire) or individual factors (observer intercepts ranging roughly from 1 to 32 blinks/min at 60 Hz) was practically negligible. Note that this is not to contradict the above argument but to emphasize that these effects were on a different scale in terms of blink inhibition and its link with visual discomfort.

The blink duration and the incomplete blink ratio were not significantly affected by the changes in the frame rate, and they were generally constant across conditions. Based on these results and considering the available resources, we did not explore a further classification for the blink type, e.g., bursts in activity, duration, e.g., extended closure, and completeness, e.g., closure percentage. A more in-depth analysis of the latter, however, could potentially provide a better insight into blink inhibition effects on visual discomfort symptoms (Ousler et al., 2014). It is also worth noting that many participants in our study exhibited binocular asymmetries in their blink completion. Lastly, concerning the effects across the two experiments, the absence of a baseline measurement, the relatively small sample size, and the demographic sample differences on several associated factors do not allow us to draw safe conclusions. However, in

terms of relative measures, there was a noticeable increase in the ratio of incomplete blinks in the second experiment with the more dynamically changing content.

Summarizing, the measurements indicated that for a trial duration of 10 min at relatively low self-paced scrolling speeds, naive observers performed the same with no considerable differences in their overall preference and no reported visual fatigue or discomfort for frame rates ranging from 30 to 240 frames/s. When we increased the task dynamics by changing to a continuous high-speed text flow, the differences among conditions became generally apparent up to the highest frame rate that we tested (240 frames/s) in less than 1 min of exposure. The participants expressed a higher preference for increasing frame rate with evidence suggesting an association with image quality and discomfort, without any practically meaningful reduction in task performance. The blinking activity was higher at the lower frequencies with a more prominent effect in the more dynamic experiment and generally in agreement with the subjective preference estimates. Based on this association, we discussed a possible link between the increase in blinks and moderate visual discomfort along with alternative interpretations. We also noted that the overall effect of frame rate on the blinking activity would likely be dwarfed by other associated factors, making it less useful as an indicator.

In conclusion, the results suggested that during everyday display use, e.g., self-paced scrolling through text at moderate speed and duration, the frame rate can be considerably reduced (even to 30 frames/s or lower) without severe effects on the user. However, this tolerance diminishes under more dynamic content, e.g., by using animated sequences or continuously fast-moving text as in our task (even much higher rates may not be tolerable). Since the optimal frame rate is naturally task-specific, the overall user acceptance and perceived visual comfort would need to be quantified as a function of frame rate for the targeted application and platform and potentially tailored to the individual user. Overall, this study provides evidence of potential gains in adaptive display designs.

# General discussion

The thesis was primarily structured around two general objectives in the field. The first was to describe the visual performance in processing the information from a finite (distorted) version of optical reality. The second was to identify the limitations of human observers and quantify their tolerance to the respective artifacts. Following the summary in the general introduction, a central topic among the studies was the contrast sensitivity function (CSF). Recall that the CSF is the most general way to describe what human observers can detect and an integral part of many visual quality metrics (such as those employed in Chapter 1). It is also a general (and quick) way for a display to adapt to the individual observer and the varying environmental and stimulus demands (especially helpful for portable devices). Two of our studies investigated how the spatial CSF varies with adaptation luminance, mainly focusing on high light levels (the discussion of the methods and results is given in Section 2.4). Recall that sensitivity is generally expected to increase with luminance before reaching a plateau; this creates an apparent trade-off between the level of detail that can be discerned and artifact visibility (greater image detail could also result in more visible artifacts). Furthermore, the presence of the plateau indicates a level where increasing the luminance further would not result in any noticeable improvements. An interesting finding was the decrease (after the peak) in sensitivity with increasing luminance at low to intermediate spatial frequencies. Although not striking, this decrease was also found in some previous studies but was generally neglected. It follows that such a decrease could result in luminance increments that have the opposite of the intended effect; this adds further complexity as lower frequencies reach the critical point at lower luminance levels. Currently, CSF models do not generally account for this effect. The same limitation applies to the model presented here; although we could easily incorporate a decreasing term, this adjustment could not be theoretically justified. An extension of this work would be developing (or modifying) a model that adjusts for this effect while maintaining the same performance across datasets. The same model could then be further modified to handle the drop in performance at the lower end of the luminance range tested (mesopic to scotopic levels). Future work could also include other pattern dimensions such as temporal frequency, eccentricity, chromatic contrast, and stimulus size. The same applies to the individual observer's characteristics (e.g., age) or other viewing conditions such as natural pupils and surround illumination.

The other two main topics were application-specific studies derived from the same overall objectives. Note that an additional purpose of this work was to demonstrate a general approach for treating similar problems. The first of these studies presented a method for estimating the perceptually optimal view density in (arbitrary) continuous-parallax content. Recall that we used the metric scores on the simulated image sequence to extract a unidimensional interval scale and estimate (through fitting an appropriate psychometric function) the near-threshold density for the depicted scene (Section 1.4 discusses this approach and our results in detail). Naturally, future work would need to extend this estimation to multiple arbitrary scenes with varying properties. However, even without an extensive calibration dataset, this mapping offers a practical advantage in reducing the duration of subjective visual assessments (i.e., better sampling of the distortion levels). Further work could directly integrate the predicted distortion intensity into a faster adaptive threshold estimation procedure, such as the Bayesian approach in Chapter 2. Note that the existing metrics were reasonably accurate in estimating the perceived distortion levels for the tested stimuli despite not being explicitly structured to capture the related artifacts; however, a dedicated model could improve prediction performance. Furthermore, the evaluation of no-reference or reduced-reference metrics would benefit applications where the full-reference signal is not known. One interesting result in the present study was that the threshold level was primarily driven by abrupt motion; the presence of slight crosstalk (reduced image sharpness but smoother transitions) generally increased observer tolerance. Naturally, this result would need to be verified by future studies. Future work could also evaluate more display-specific (non-generic) simulations for the displayed images.

The second of these studies explored the effect of high temporal presentation rates on visual comfort and fatigue. The results were generally null in the more realistic task, although the overall negative effect of lower frequencies could be demonstrated under more demanding (and controlled) conditions (see Section 3.4). One of the main difficulties in measuring the perceived symptoms (particularly with naive observers) is decoupling the different factors that may affect the subjective responses (e.g., general preference due to better image quality). A typical way to address this issue is with the use of objective metrics. Researchers have proposed several potential indicators; here, we examined task performance and blinking activity. The suitability of these measures was thoroughly analyzed in Chapter 3. Generally, the main issue was the relatively small effect that we found even when subjective feedback indicated otherwise. Note that it is common practice in vision science to use a relatively small number of participants, which can impact the generality of the results; thus, one should also consider the minimum effect we can reliably detect given the sample size. Future studies could explore alternative indicators and different usage scenarios, including the impact of adaptive designs.

# Appendix A. Supplementary material for Chapter 1

## Individual fits as a function of disparity

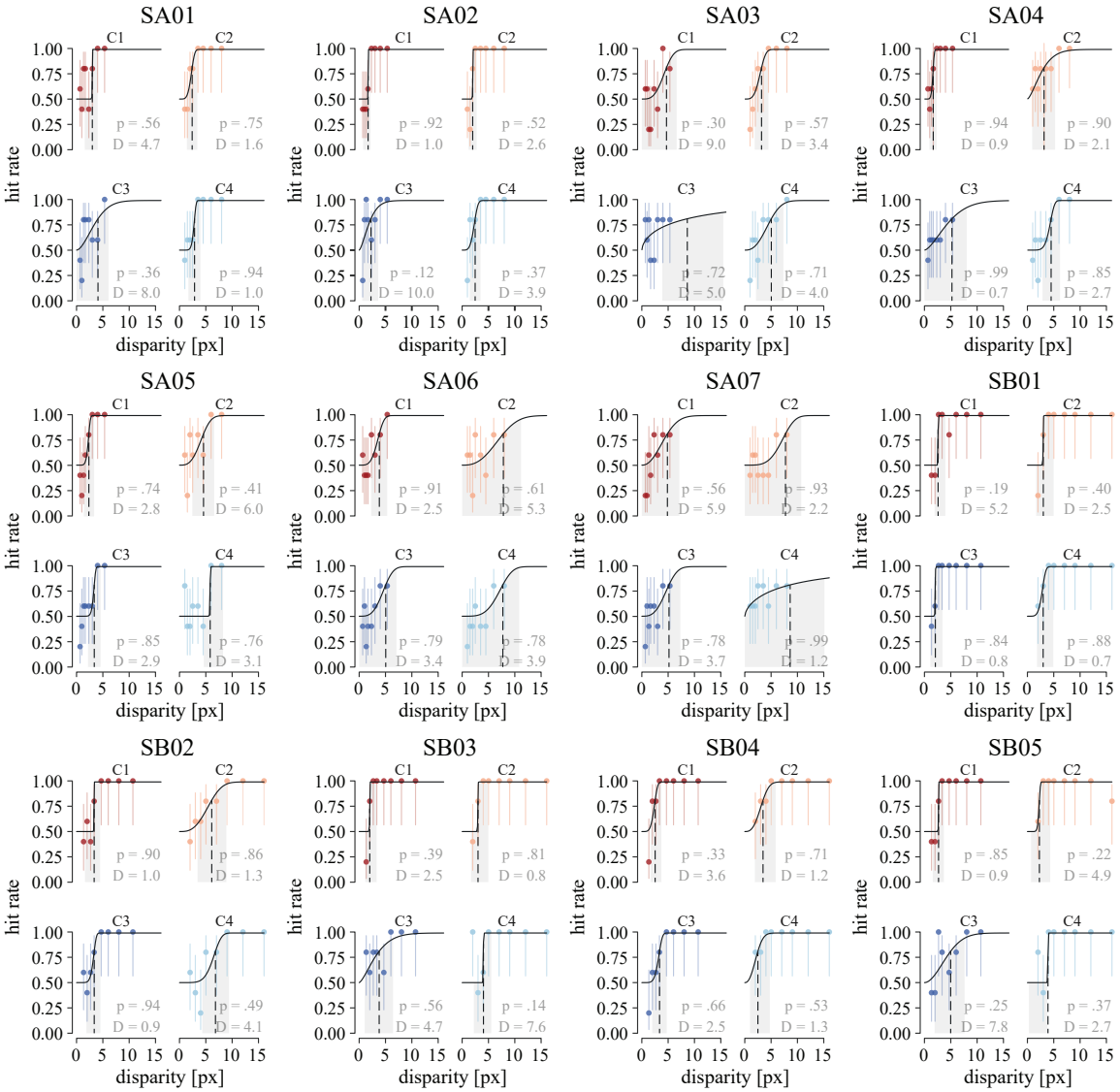


**Figure A.1** The individual psychometric function fits on the observer performance (markers) with the maximum pixel disparity among the input views as the degradation intensity. The error bars show the Wilson binomial 95% CI. The shaded area depicts the 95% CI for the threshold estimate (dashed line) by parametric bootstrap. The goodness-of-fit is shown on the bottom right-hand corner (gray text), where D is the observed deviance and p the probability of obtaining deviance on the Monte-Carlo generated samples (N = 10000) larger than the observed value. Each facet corresponds to different combinations of step size (dx) and view blending (M), where **C1**: dx = 2 mm, M = 0, **C2**: dx = 3 mm, M = 0, **C3**: dx = 2 mm, M = 4, and **C4**: dx = 3 mm, M = 4. The observers with the letter A correspond to the first (depth z = 0 cm) group. Likewise the letter B indicates the second group (z = -15.6 cm)

## Metric performance across conditions

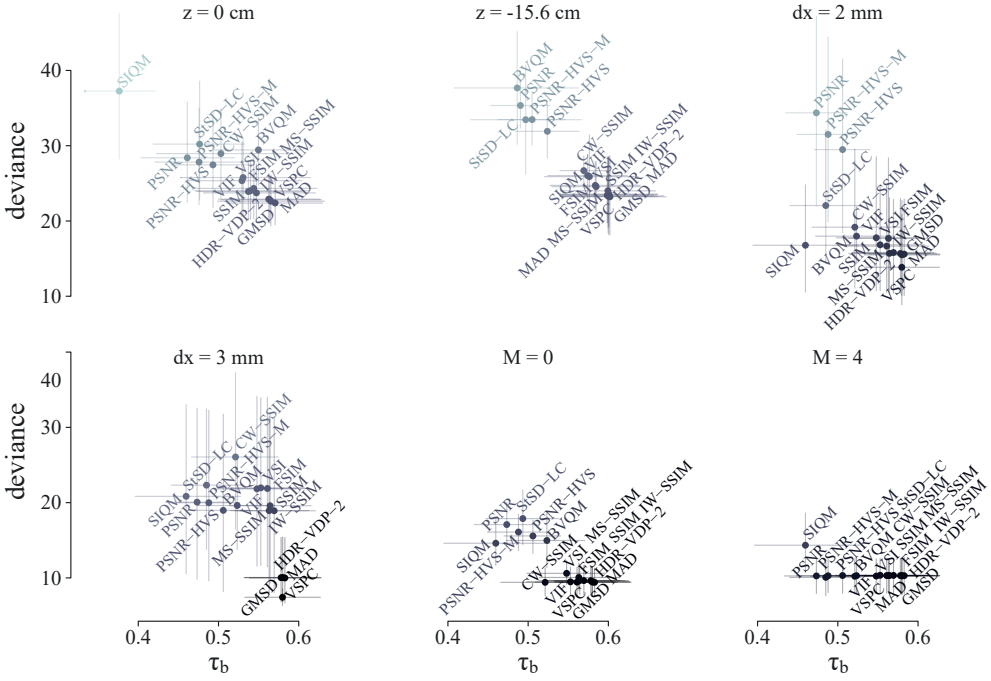


**Figure A.2** The average metrics performance across observers and conditions in terms of the Kendall $\tau_b$ correlation and the fitted **Weibull** function deviance. The error bars depict the 95% CI by basic bootstrap. The text labels were slightly jittered to avoid overlap

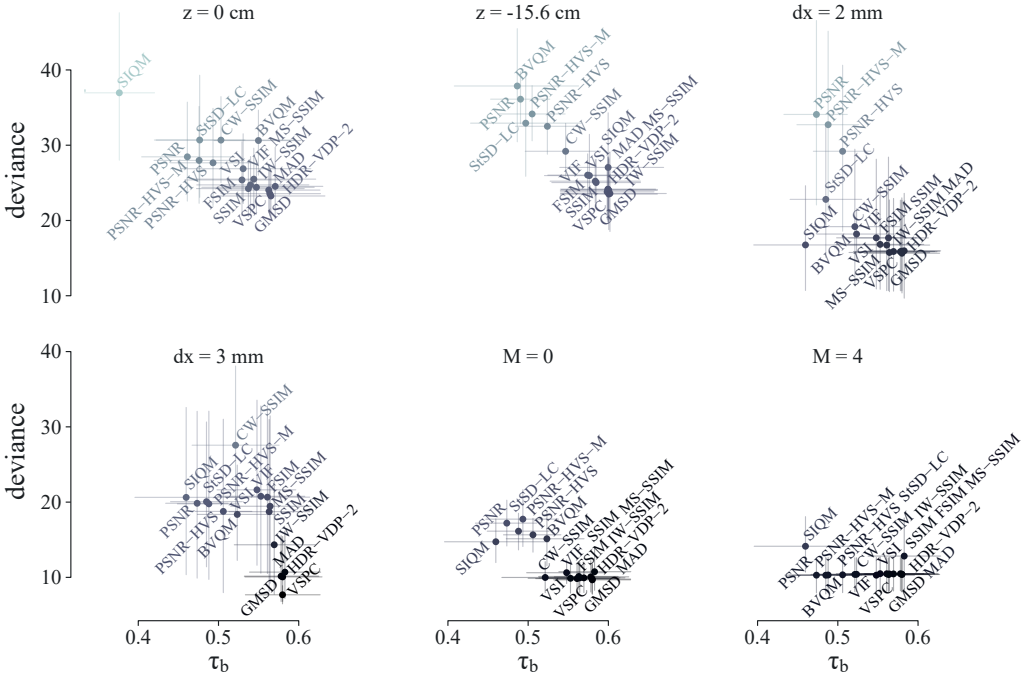


**Figure A.3** The average metrics performance across observers and conditions in terms of the Kendall $\tau_b$ correlation and the fitted cumulative **Gaussian** function deviance. The error bars depict the 95% CI by basic bootstrap. The text labels were slightly jittered to avoid overlap

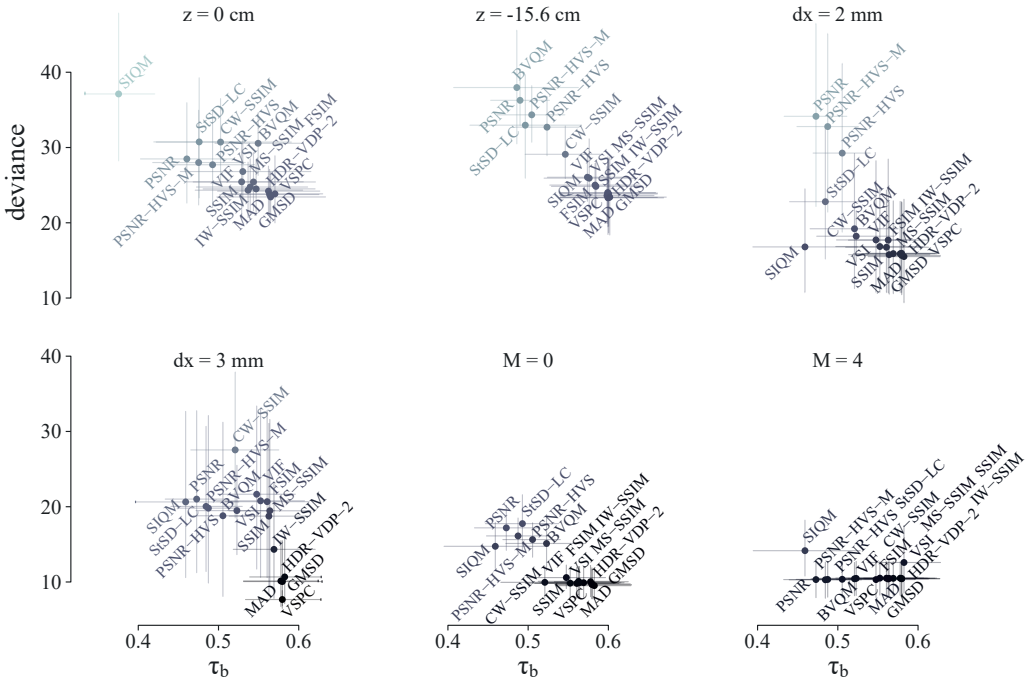


**Figure A.4** The average metrics performance across observers and conditions in terms of the Kendall $\tau_b$ correlation and the fitted **Logistic** function deviance. The error bars depict the 95% CI by basic bootstrap. The text labels were slightly jittered to avoid overlap

## Metric correlation examples

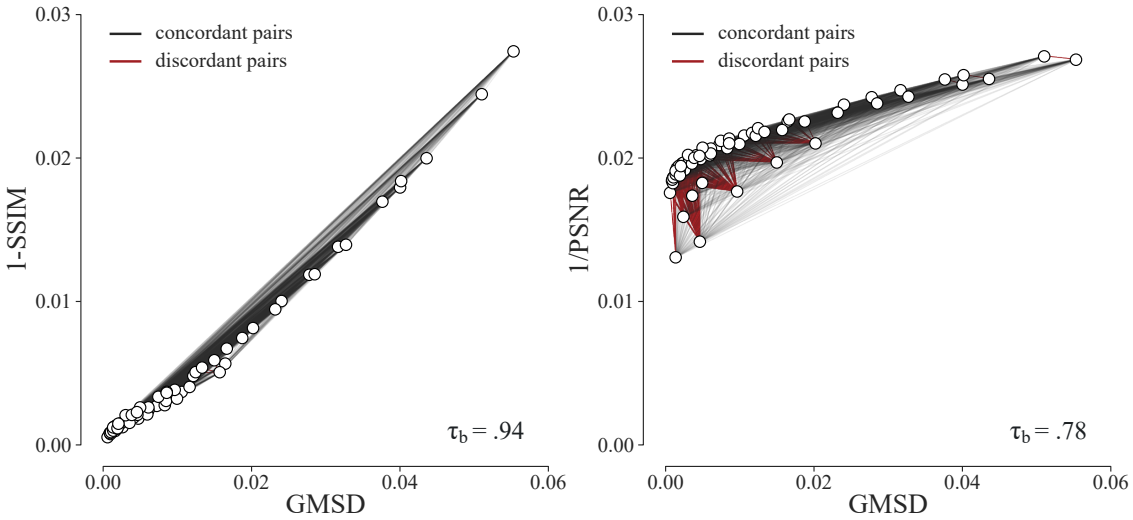


**Figure A.5** An example of the Kendall $\tau_b$ pairwise correlations among metrics as depicted in Figure 1.8 (left panel for high correlation). The markers show the metric scores for the examined distorted conditions in terms of increasing degradation. The concordant and discordant pairs are connected with black and red segments, respectively (more details can be found in Davis and Chen, 2007)

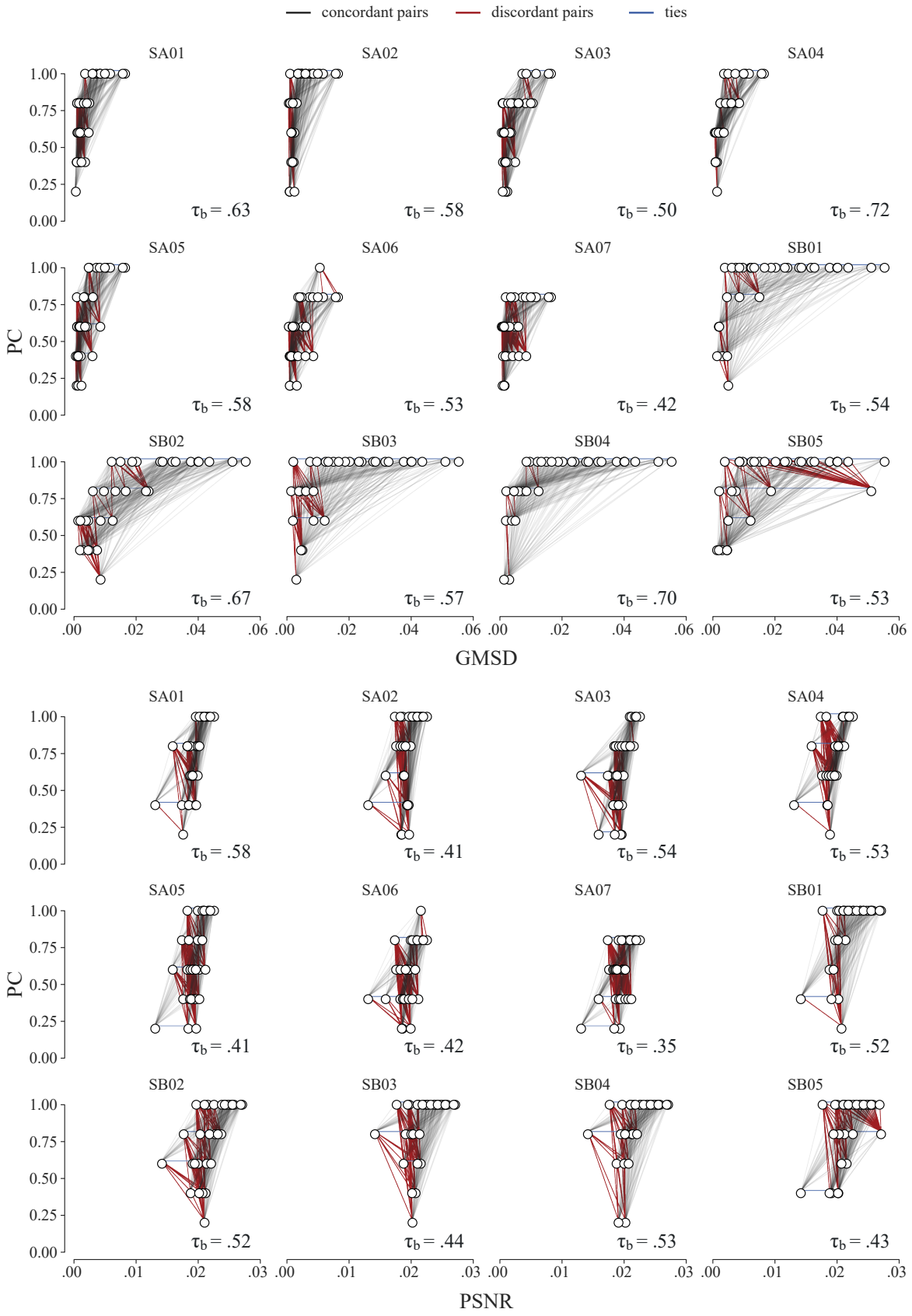


**Figure A.6** An example of the Kendall $\tau_b$ correlation with the proportion of correct responses across observers as illustrated in in Figure 1.9. The concordant and discordant pairs are connected with black and red segments, respectively, whereas ties are depicted with horizontal blue lines

# Appendix B. Supplementary material for Chapter 2

**Basic sampling characteristics of the data included in the meta-analysis**

| Source study | $[N_{obs}, N_{SF}, \text{Mdn}.N_E]$ | $SF$ range [cpd] | approx. max $E$ range [Td] |
|---|---|---|---|
| Banks et. al 1987 (AVG) | [2, 5, 3.0] | [5, 20] | [1.0e+1, 1.0e+3] |
| Bierings et. al 2019 (AVG) | [51, 3, 6.0] | [1, 10] | [1.5e−1, 3.5e+4] |
| De Valois et. al 1974 (AVG) | [5, 8, 4.5] | [.6, 14] | [3.8e−2, 3.5e+2] |
| Hess & Howell 1988 (ERH) | [1, 8, 6.5] | [.1, 20] | [3.8e−5, 1.1e+3] |
| Present study (Exp I) (AVG) | [11, 6, 4.0] | [1.5, 30] | [2.0e+2, 3.2e+3] |
| Present study (Exp II) (S01) | [1, 6, 8.0] | [1.5, 30] | [8.9e+0, 2.3e+4] |
| Present study (Exp II) (S07) | [1, 6, 8.0] | [1.5, 30] | [8.9e+0, 2.5e+4] |
| Kelly 1972 (EB) | [1, 11, 4.0] | [.25, 8] | [6.9e−1, 6.9e+2] |
| Kim et. al 2013 (AVG) | [7, 10, 5.0] | [.125, 16] | [4.3e−1, 1.3e+3] |
| Patel 1966 (DR) | [1, 13, 3.0] | [.8, 18] | [2.8e−1, 9.4e+2] |
| Peli et. al 1996 (AVG) | [4, 5, 7.0] | [1, 16] | [1.1e+1, 9.5e+2] |
| Rasengane et. al 2001 (AVG) | [5, 1, 5.0] | [.25, .25] | [1.8e+0, 7.3e+3] |
| Rovamo et. al 1994 (ALL) | [5, 8, 6.0] | [.125, 32] | [7.0e−5, 1.6e+4] |
| Shannon et. al 1996 (AVG) | [2, 5, 5.0] | [.5, 6] | [1.7e+0, 5.4e+2] |
| Silvestre et. al 2018 (AVG) | [4, 7, 5.0] | [.25, 16] | [1.4e−1, 1.4e+4] |
| Valero et. al 2004 (AVG) | [3, 3, 6.0] | [.5, 2] | [9.1e+0, 1.5e+3] |
| van Meeteren & Vos 1972 (AVG) | [2, 7, 5.0] | [.5, 19] | [2.0e−3, 1.4e+2] |
| van Nes & Bouman 1967 (FLN) | [1, 18, 6.0] | [.5, 48] | [8.5e−4, 5.6e+3] |

**Table B.1** Summary of the basic sampling characteristics of the data included in the meta-analysis. The second column reports the number of individual observers in the estimation of the average sensitivity ($N_{obs}$), the number of spatial frequencies ($N_{SF}$), and the median number of retinal illuminance samples across frequencies (Mdn.$N_E$). The last two columns show the range of the spatial frequencies and the approximate maximum retinal illuminance range tested. The characteristics of experiment II in the present study are only given for comparison purposes as they were omitted from this analysis

### Individual descriptive CSF model fits

| Observer | $RMSE_0$ [dB] | | | | | $RMSE_{ALP}$ [dB] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 68 | 208 | 475 | 1078 | **All** | 68 | 208 | 475 | 1078 | **All** |
| S01 | .59 | .91 | .67 | .55 | .70 | .63 | 1.06 | .82 | .70 | .82 |
| S02 | 1.34 | 2.49 | 2.18 | 2.29 | 2.12 | 1.46 | 2.49 | 2.22 | 2.34 | 2.16 |
| S03 | 2.93 | 2.41 | 2.50 | 2.34 | 2.55 | 2.95 | 2.43 | 2.51 | 2.42 | 2.59 |
| S04 | 3.95 | 3.81 | 3.39 | 3.49 | 3.67 | 3.96 | 3.89 | 3.42 | 3.53 | 3.71 |
| S05 | 4.21 | 1.73 | 3.46 | 3.09 | 3.25 | 4.22 | 1.73 | 3.49 | 3.18 | 3.28 |
| S06 | 4.32 | 2.41 | 3.85 | 2.72 | 3.42 | 4.35 | 2.44 | 3.85 | 2.75 | 3.44 |
| S07 | 2.66 | 2.95 | 2.70 | 2.14 | 2.63 | 2.66 | 2.99 | 2.79 | 2.48 | 2.74 |
| S08 | 2.59 | 3.90 | 2.20 | 3.28 | 3.06 | 2.70 | 3.90 | 2.26 | 3.32 | 3.11 |
| S09 | 2.19 | 3.19 | 2.38 | 2.21 | 2.53 | 2.24 | 3.24 | 2.51 | 2.45 | 2.64 |
| S10 | 2.05 | 2.52 | 2.29 | 1.97 | 2.22 | 2.20 | 2.65 | 2.34 | 1.99 | 2.31 |
| S11 | 1.32 | 1.56 | 1.84 | 1.77 | 1.63 | 1.32 | 1.76 | 1.88 | 2.19 | 1.81 |

**Table B.2** The fitting performance of the asymmetric log-parabola CSF form ($RMSE_{ALP}$) for each observer and luminance level, and for all the levels combined. The $RMSE_0$ corresponds to a model where the threshold is given by the mean log estimate across repetitions at each level. The RMS errors are shown in dB (the log sensitivity multiplied by 20) on the full set of measurements

### Intercorrelations of the main CSF characteristics

| CSF measure | *AULCSF* | $F_c$ | $F_{max}$ |
|---|---|---|---|
| $F_c$ | $.78^{***}$ [.61, .88] | | |
| $F_{max}$ | $.65^{***}$ [.42, .81] | $.41^{*}$ [.07, .69] | |
| $G_{max}$ | $.87^{***}$ [.78, .94] | $.58^{**}$ [.33, .76] | .33 [.04, .59] |

Notes: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

**Table B.3** Repeated measures correlation matrix for the main CSF characteristics. The matrix shows the $r_{rm}$ correlation coefficients and their 95% CI by bootstrap (within brackets) for experiment I. $G_{max}$: the peak sensitivity. $F_{max}$: the location of the peak. $F_c$: the high-frequency cut-off

### Sensitivity analysis

The sensitivity analysis was performed using the R library Superpower (Caldwell and Lakens, 2019). The confidence intervals for the partial eta squared were calculated following the procedure in Smithson (2001). Note that the partial eta squared estimation includes the correlation among the paired measures (see the appendix in Lakens, 2013). In the present sample the mean correlation was .89 (SD = .04).

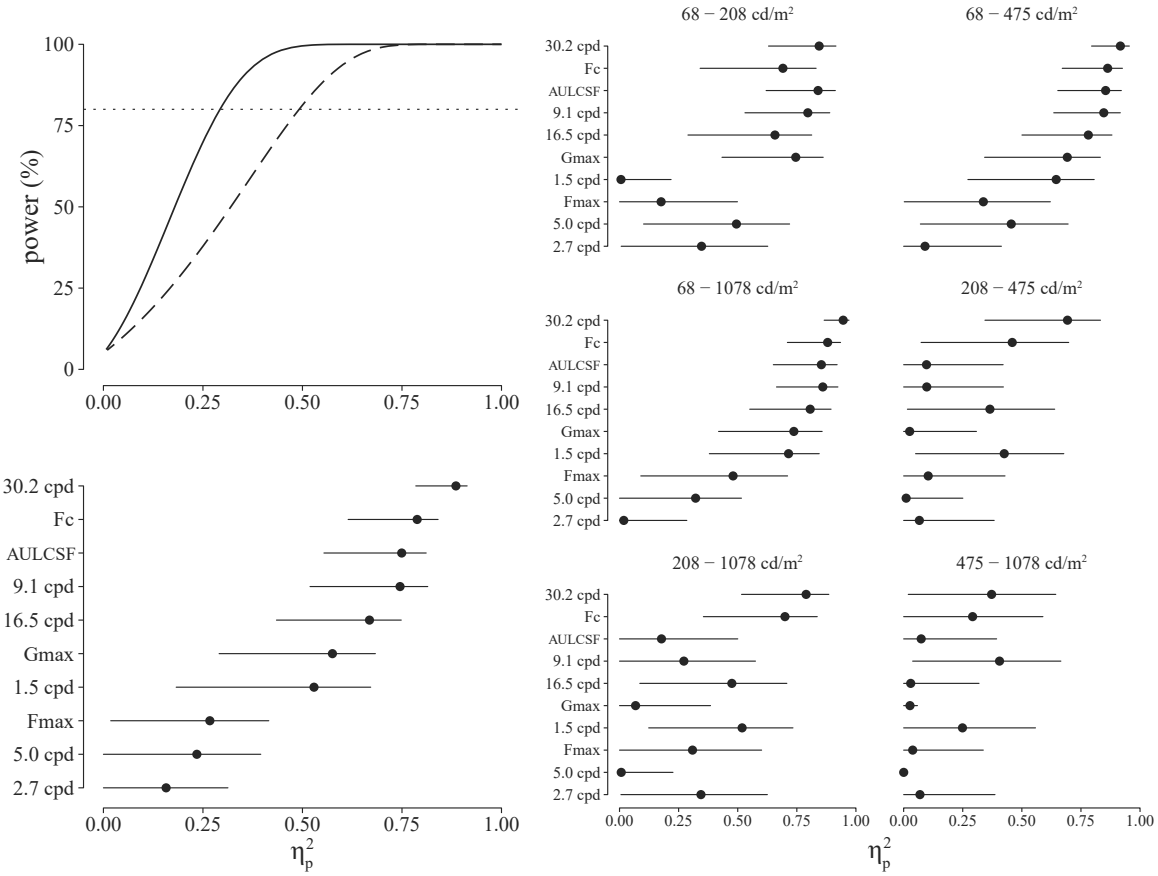


**Figure B.1 Top left**: the estimated power as a function of the partial eta squared measure. The solid line depicts the power of the ANOVAs, and the dashed line the power for the pairwise comparisons (without multiplicity adjustment). The dotted line shows the conventional 80% power level. **Bottom left**: the measured effect size for the tested variables. **Right**: the measured effect size for each pairwise comparison (panels) and variable (we reported the eta squared based on $F = t^2$ when the numerator degrees of freedom equals one; see Table 2 in Friedman, 1982). The error bars show the 90% CI

**Barten's contrast sensitivity model at photopic luminance**

The formula for estimating contrast sensitivity $S$ for binocular viewing of static stimuli as a function of spatial frequency $u$ (roughly with a lower bound of 1 cpd) is expressed with the following equation for retinal illuminance $E$ [Td]:

$$S(u) = \frac{M_{\text{opt}}(u)/k}{\sqrt{\frac{2}{T}\left(\frac{1}{X_0^2} + \frac{1}{X_{max}^2} + \frac{u^2}{N_{max}^2}\right)\left(\frac{1}{\eta p E} + \frac{\Phi_0}{1-e^{(u/u_0)^2}}\right)}} \tag{B.1}$$

where $M_{\text{opt}}(u)$ is the optical modulation transfer function of the eye given as:

$$M_{\text{opt}}(u) = \mathrm{e}^{-2\pi^2\sigma^2\mathrm{u}^2} \tag{B.2}$$

and $\sigma$ is the standard deviation of the eye line spread function that can be expressed as:

$$\sigma = \sqrt{\sigma_0^2 + (C_{\text{ab}} d)^2} \tag{B.3}$$

Barten (1999) provides a detailed interpretation of the parameters and an analytical derivation of the above formulae. The physical quantities and their typical values are given in Table B.4.

| Parameter | Short description | Value/units |
|---|---|---|
| $d$ | pupil diameter | mm |
| $X_0$ | stimulus angular size | degrees |
| $p$ | photon conversion factor | photons/sec/deg$^2$/Td |
| $T$ | integration time of the eye | .1 sec |
| $u_0$ | spatial frequency limit for lateral inhibition | 7 cpd |
| $X_{max}$ | maximum size of the integration area | 12 degrees |
| $N_{max}$ | maximum number of integration cycles | 15 cycles |
| $\Phi_0$ | spectral density of the neural noise | $3 \times 10^8$ sec degrees$^2$ |
| $C_{ab}$ | constant that controls the increase of $\sigma$ with the pupil size | .08 arc min/mm |
| $\sigma_0$ | constant that depends on the eye lens and the cone density | .5 arc min |
| $\kappa$ | signal-to-noise ratio at the detection threshold | 3 |
| $\eta$ | quantum efficiency of the eye | .03 |

**Table B.4** Barten's model parameters and their typical values for an average observer and foveal vision. For usual conditions $p \approx 1.2 \times 10^6$ (depends on the light source spectrum). Note that the $C_{ab}$ and $\sigma_0$ parameters require a conversion to degrees

# Appendix C. Supplementary material for Chapter 3

## Subjective feedback questionnaires

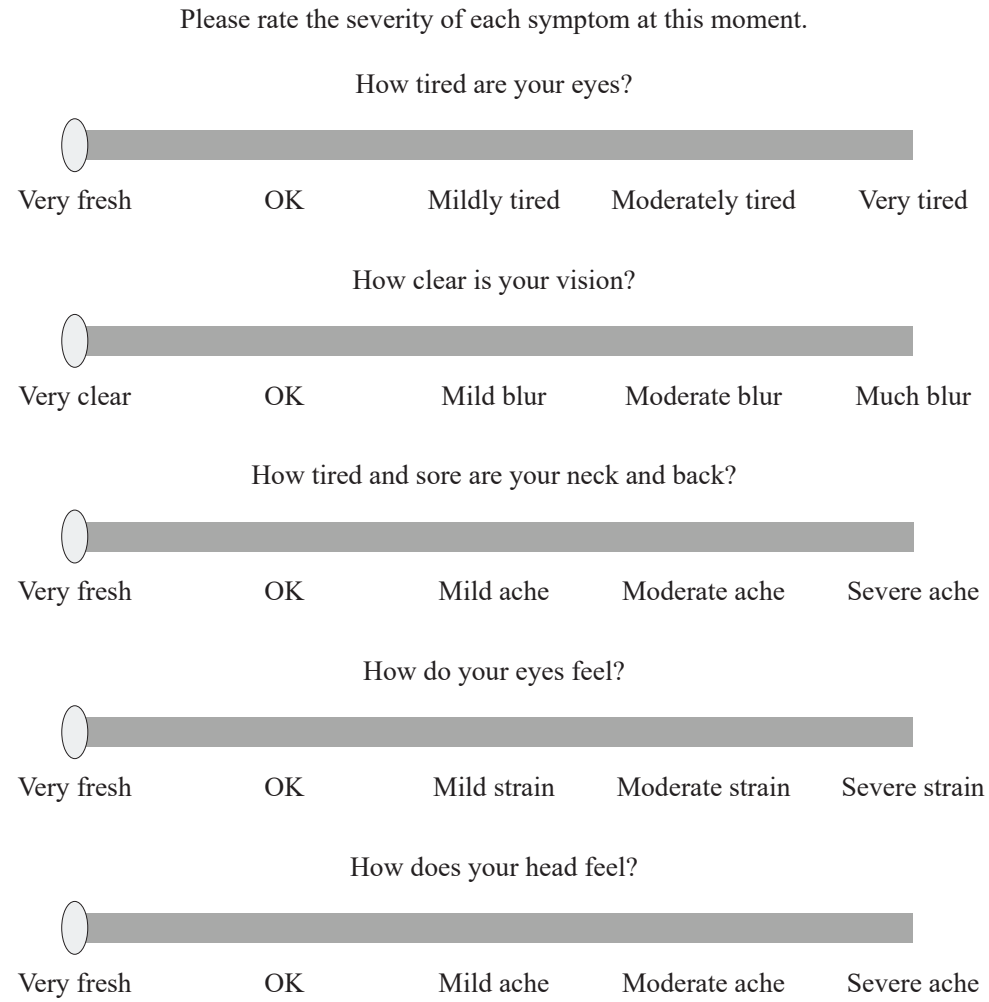


**Figure C.1** The visual fatigue and discomfort symptoms questionnaire in experiment I

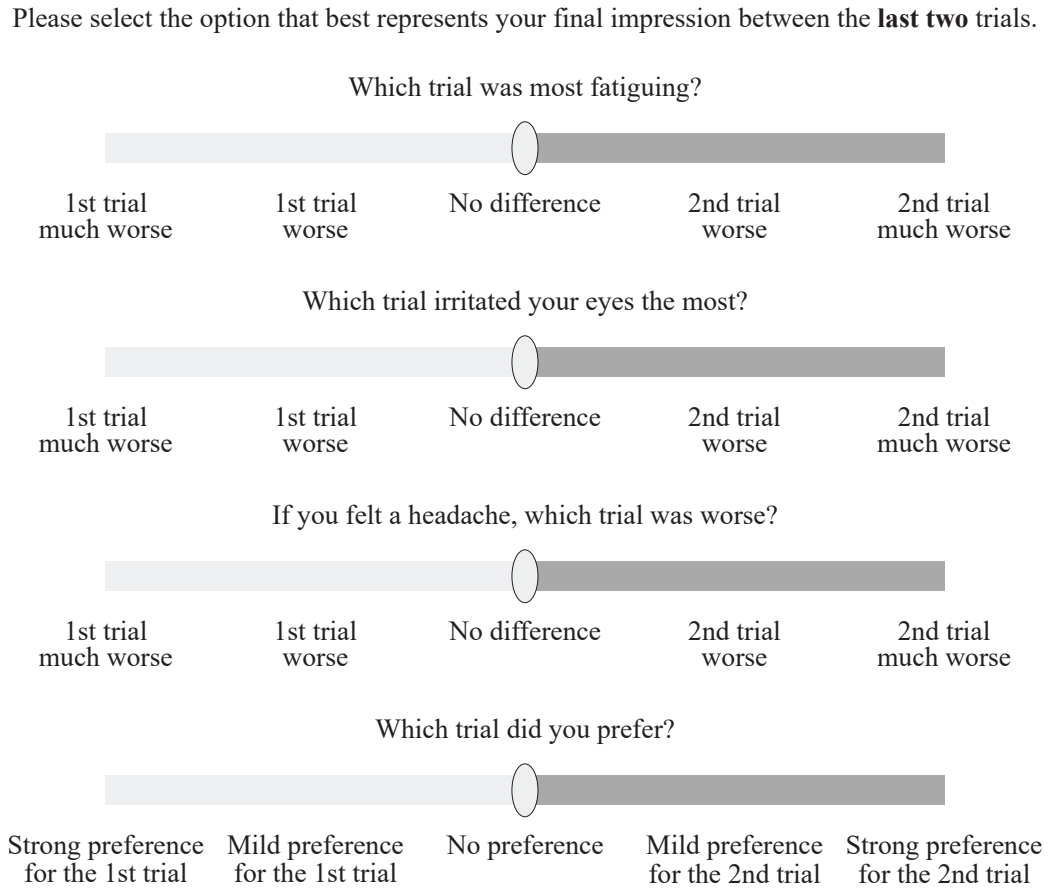


**Figure C.2** The frame rate evaluation questionnaire in Experiment I

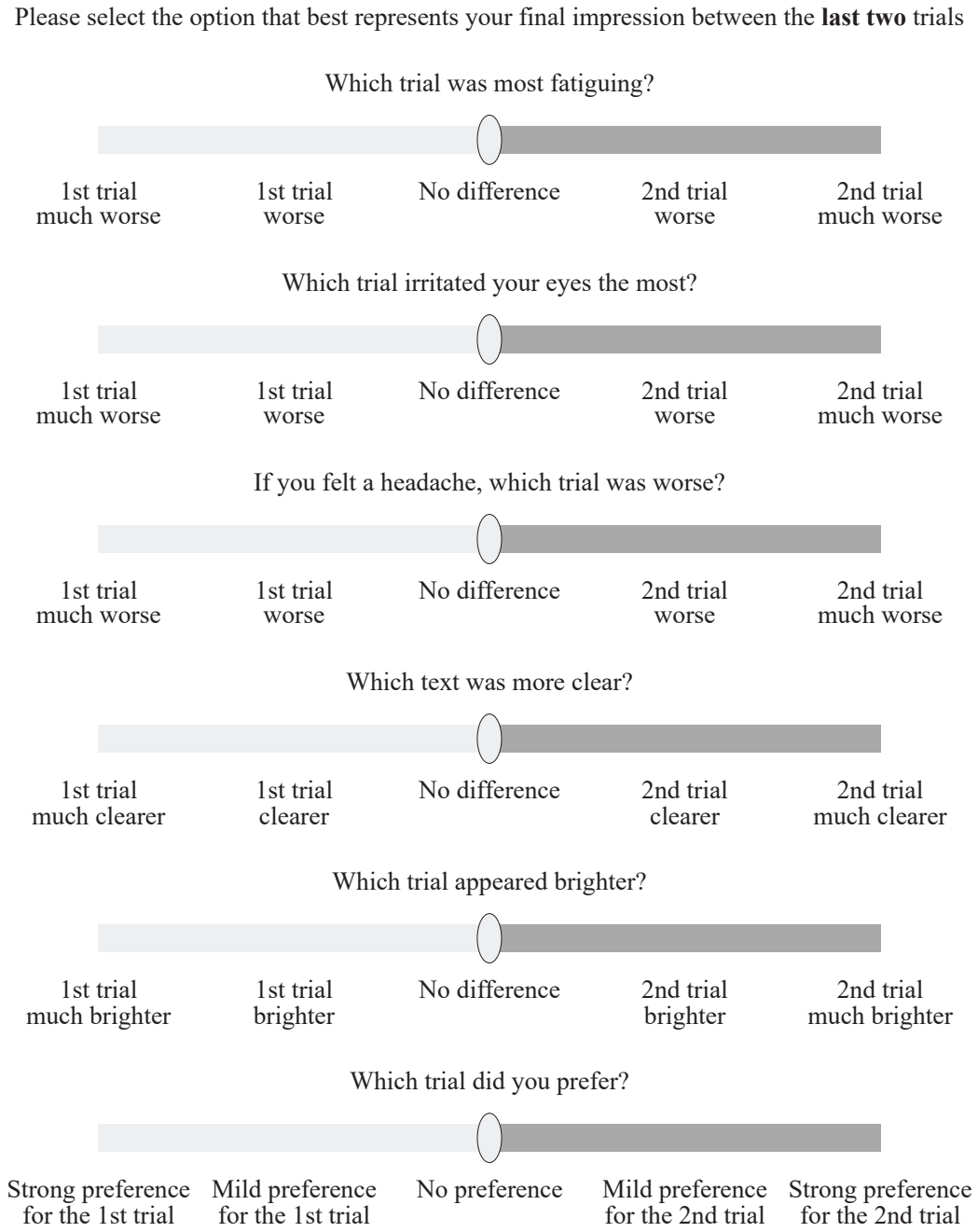


**Figure C.3** The frame rate evaluation questionnaire in Experiment II

### Blink detection classification performance

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| class 0: non-blink | .99 | .94 | .97 | 192195 |
| class 1: incomplete | .62 | .90 | .73 | 19846 |
| class 2: complete | .89 | .86 | .87 | 4023 |
| accuracy | | | .94 | 216064 |
| macro avg | .83 | .90 | .86 | 216064 |
| weighted avg | .95 | .94 | .94 | 216064 |

**Table C.1** Frame-based performance of the eye state classifier

### Supplementary blink rate data

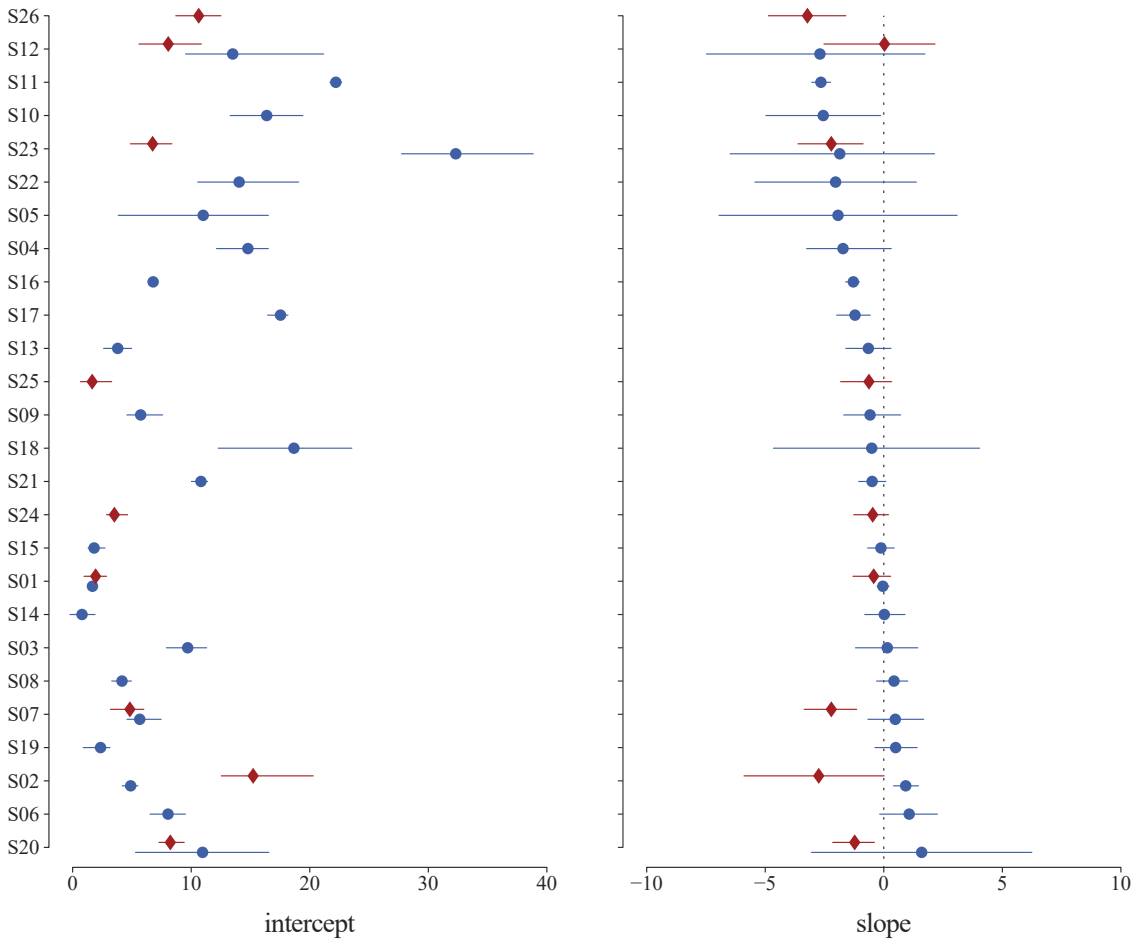


**Figure C.4** The coefficients (blue circles for experiment I; red diamonds for experiment II) of separate individual linear regressions on the blink rate (BR) with predictor the log frequency. The data were centered at 60 Hz. The error bars show the 95% CI by parametric bootstrap

## Sensitivity analysis

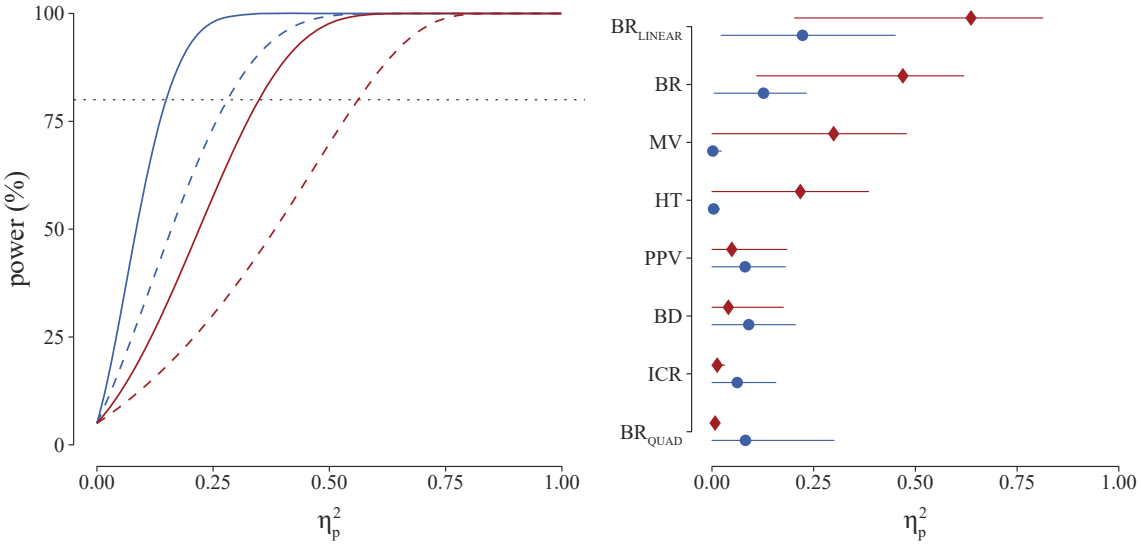


**Figure C.5 Left**: the estimated statistical power as a function of the partial eta squared measure. The solid line depicts the power of the ANOVAs (blue for experiment I and red for experiment II), and the dashed line the power for the follow-up tests (without multiplicity adjustment). The dotted line shows the conventional 80% power level. **Right**: the measured effect size for the tested variables (blue circles for experiment I and red diamonds for experiment II). The error bars show the 90% CI. More details on the sensitivity analysis can be found in Appendix B

## Subjective feedback data

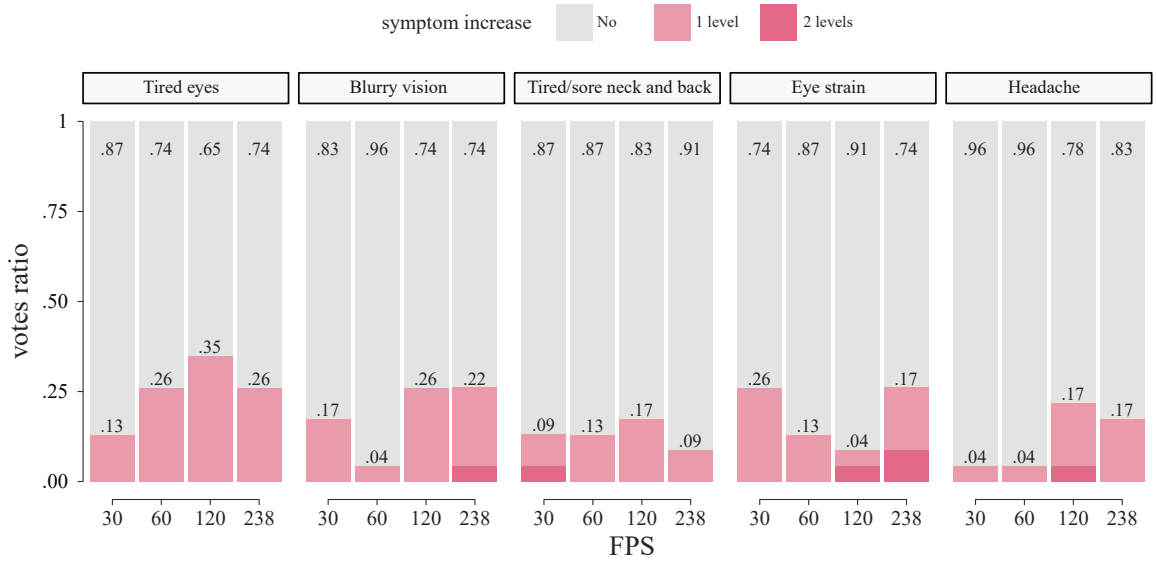


**Figure C.6** The ratio of the self-reported post-trial increase (stacked bars; text labels) of visual discomfort and fatigue symptoms across participants for each frame rate condition in the first experiment. The questionnaire can be found in Figure C.1

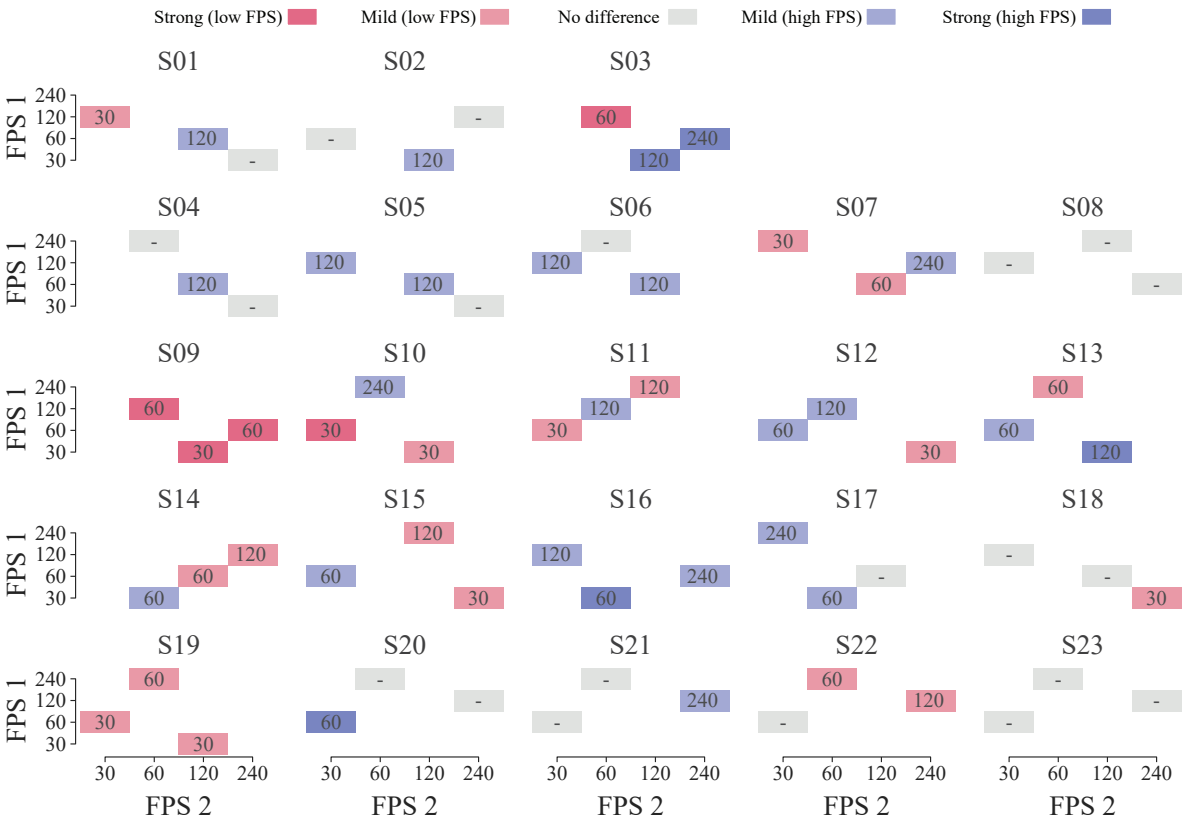


**Figure C.7** Individual preferences of the paired frame rate evaluations in Experiment I

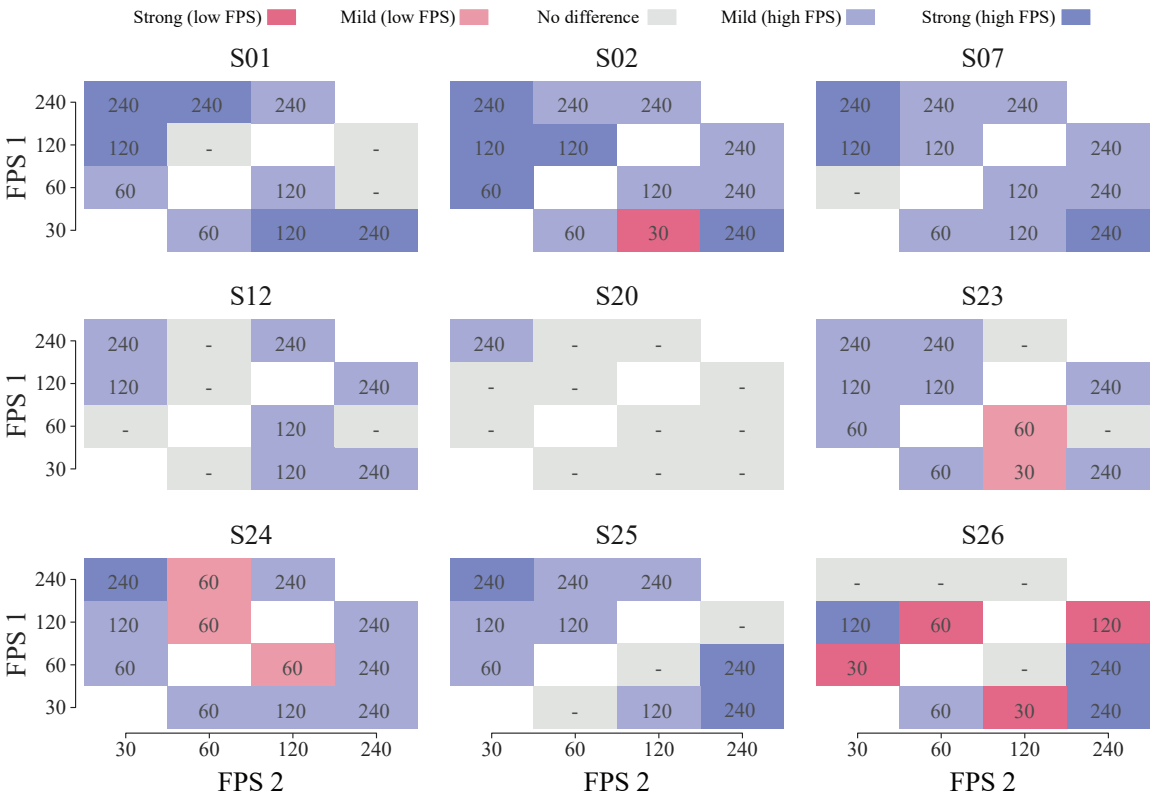


**Figure C.8** Individual preferences of the paired frame rate evaluations in Experiment II

**Supplementary preference scaling data**

The conversion from the worth parameters ($\pi > 0$) to the predicted probabilities of selecting the frequency $i$ to the frequency $j$ ($i \succ j$), or the frequency $j$ to $i$ ($i \prec j$), or none ($i \approx j$) given the no preference estimate $\gamma$, where $i, j \in \{30, 60, 120, 240\}$ frames/s, can be calculated using the Davidson's tie model (equations C.1; Davidson, 1970) with the parameter $\nu = \exp(\gamma)$:

$$
\begin{aligned}
P(i \succ j) &= \frac{\pi_i}{\pi_i + \pi_j + \nu\sqrt{\pi_i \pi_j}} \\
P(i \prec j) &= \frac{\pi_j}{\pi_i + \pi_j + \nu\sqrt{\pi_i \pi_j}} \\
P(i \approx j) &= \frac{\nu\sqrt{\pi_i \pi_j}}{\pi_i + \pi_j + \nu\sqrt{\pi_i \pi_j}}
\end{aligned}
\tag{C.1}
$$

| FPS comparison | | | **Experiment I** | | | **Experiment II** (n = 18) | | |
|---|---|---|---|---|---|---|---|---|
| | | | n | Low | No | High | Low | No | High |
| 30 | vs | 60 | 14 | .29 (.21) | .31 (.29) | .40 (.50) | .09 (.06) | .29 (.33) | .63 (.61) |
| 60 | vs | 120 | 12 | .30 (.33) | .31 (.17) | .39 (.50) | .25 (.22) | .38 (.33) | .37 (.44) |
| 120 | vs | 238 | 10 | .49 (.40) | .29 (.40) | .22 (.20) | .10 (.06) | .30 (.33) | .60 (.61) |
| 30 | vs | 120 | 12 | .25 (.33) | .30 (.17) | .44 (.50) | .06 (.17) | .26 (.11) | .68 (.72) |
| 60 | vs | 238 | 13 | .45 (.31) | .30 (.46) | .25 (.23) | .07 (.06) | .27 (.39) | .66 (.56) |
| 30 | vs | 238 | 8 | .39 (.50) | .31 (.38) | .30 (.12) | .01 (.00) | .13 (.11) | .86 (.89) |

**Table C.2** Predicted probabilities among the frame rate pairs of the overall preference scaling model. **Low**, **High**, and **No** indicate the probabilities of an observer selecting the lower, the higher, or none of the frame rates in comparison, respectively. The gray numbers in the parentheses depict the aggregated observed values. **n**: the total number of votes for each comparison

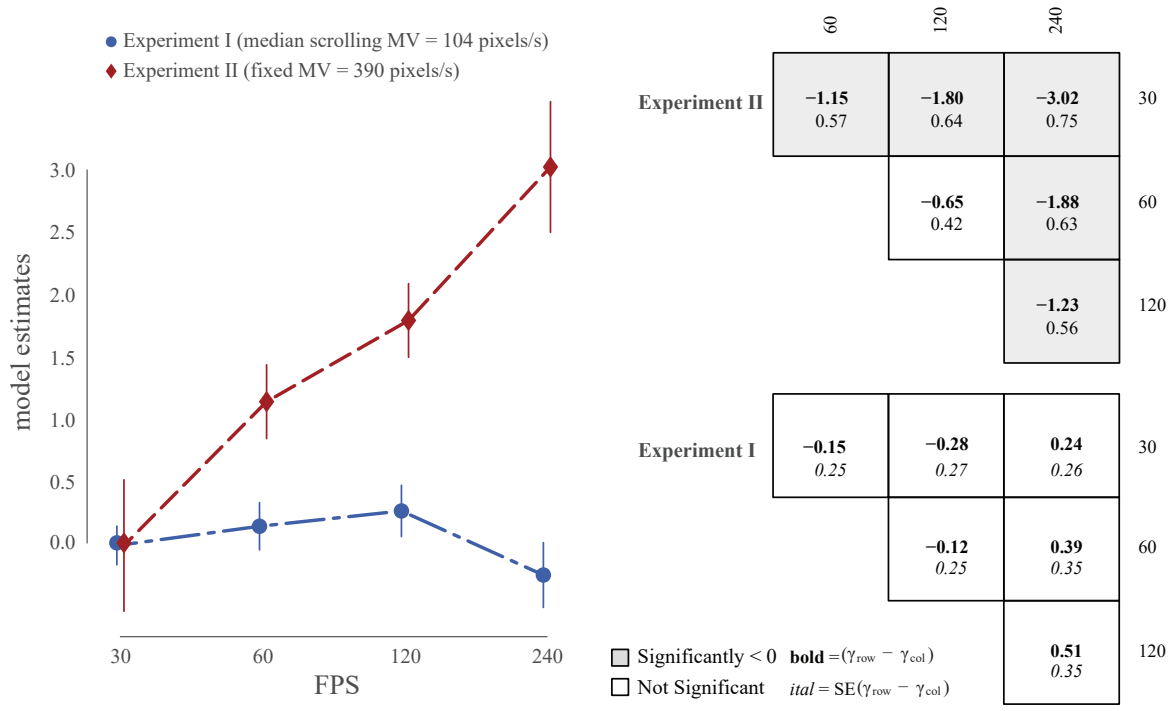


**Figure C.9** Frame rates scaling based on the overall observer preference in the alternative model accounting for the within-subject dependencies. **Left**: The log model estimates (vertical axis) as a function of frame rate (horizontal axis) with the 30 frames/s as the reference value, for experiments I (blue circles) and II (red diamonds). The error bars illustrate the quasi standard errors. **Right**: The difference in the estimates (bold text) and the standard error (italics) among all the frame rate pairs for experiments I (bottom) and II (top). A negative sign indicates that the column is preferred to the row. The shaded blocks indicate p-values below .05

# References


Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org, last accessed: 2020-02-15. 74

Acosta, M., Gallar, J., and Belmonte, C. (1999). The Influence of Eye Solutions on Blinking and Ocular Comfort at Rest and During Work at Video Display Terminals. *Experimental Eye Research*, 68(6):663–669. 69

Adelson, E. H. and Bergen, J. R. (1991). The plenoptic function and the elements of early vision. In Landy, M. S. and Movshon, J. A., editors, *Computational models of visual processing*, pages 3–20. The MIT Press. 1, 14

Adhikarla, V. K., Vinkler, M., Sumin, D., Mantiuk, R. K., Myszkowski, K., Seidel, H.-P., and Didyk, P. (2017). Towards a quality metric for dense light fields. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 17, 29, 32

Applegate, R. A. and Lakshminarayanan, V. (1993). Parametric representation of stiles–crawford functions: normal variation of peak location and directionality. *J. Opt. Soc. Am. A*, 10(7):1611–1623. 38

Armstrong, D. (2016). *factorplot: Graphical Presentation of Simple Contrasts*. R package version 1.1-2. 85

Atchison, D. A. and Smith, G. (2000). *Optics of the Human Eye*. Butterworth-Heinemann. 1, 38

Bakdash, J. Z. and Marusich, L. R. (2017). Repeated Measures Correlation. *Frontiers in Psychology*, 8:456. 76

Baker, D. H., Lygo, F. A., Meese, T. S., and Georgeson, M. A. (2018). Binocular summation revisited: Beyond $\sqrt{2}$. *Psychological Bulletin*, 144(11):1186–1199. 61

Banks, M. S., Geisler, W. S., and Bennett, P. J. (1987). The physical limits of grating visibility. *Vision Research*, 27(11):1915–1924. 51, 59

Banks, M. S., Hoffman, D. M., Kim, J., and Wetzstein, G. (2016). 3D Displays. *Annual Review of Vision Science*, 2(1):397–435. 6

Barbato, G., Ficca, G., Muscettola, G., Fichele, M., Beatrice, M., and Rinaldi, F. (2000). Diurnal variation in spontaneous eye-blink rate. *Psychiatry Research*, 93(2):145–151. 69

Barlow, H. (1964). The physical limits of visual discrimination. In *Photophysiology*, pages 163–202. Elsevier. 33

Barten, P. G. (1999). *Contrast Sensitivity of the Human Eye and Its Effects on Image Quality*. SPIE. 51, 54, 57, 60, 61, 66, 100

Barten, P. G. J. (2003). Formula for the contrast sensitivity of the human eye. In Miyake, Y. and Rasmussen, D. R., editors, *Image Quality and System Performance*, volume 5294, pages 231 – 238. International Society for Optics and Photonics, SPIE. 61

Benedetto, S., Carbone, A., Drai-Zerbib, V., Pedrotti, M., and Baccino, T. (2014). Effects of luminance and illuminance on visual fatigue and arousal during digital reading. *Computers in Human Behavior*, 41:112–119. 68

Bergqvist, U. O. and Knave, B. G. (1994). Eye discomfort and work with visual display terminals. *Scandinavian Journal of Work, Environment & Health*, 20(1):27–33. 68

Bierings, R., Overkempe, T., van Berkel, C., Kuiper, M., and Jansonius, N. (2019). Spatial contrast sensitivity from star- to sunlight in healthy subjects and patients with glaucoma. *Vision Research*, 158(March 2018):31–39. 33, 51, 59, 63

Bolles, R. C., Baker, H. H., and Marimont, D. H. (1987). Epipolar-plane image analysis: An approach to determining structure from motion. *Int J Comput Vision*, 1(1):7–55. 15

Brainard, D. H. (1997). The psychophyics toolbox. *Spatial vision*, 10:433–6. 35

Caffier, P. P., Erdmann, U., and Ullsperger, P. (2003). Experimental evaluation of eye-blink parameters as a drowsiness measure. *European Journal of Applied Physiology*, 89(3):319–325. 69

Caldwell, A. and Lakens, D. (2019). Power analysis with superpower. Last accessed: Jan 2021. 99

Campbell, F. W. and Green, D. G. (1965). Monocular versus Binocular Visual Acuity. *Nature*, 208(5006):191–192. 61

Cardona, G., García, C., Serés, C., Vilaseca, M., and Gispets, J. (2011). Blink Rate, Blink Amplitude, and Tear Film Integrity during Dynamic Visual Display Terminal Tasks. *Current Eye Research*, 36(3):190–197. 70, 89

Chai, J. X., Chan, S. C., Shum, H.-Y., and Tong, X. (2000). Plenoptic sampling. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques - SIGGRAPH '00*, volume 28, pages 307–318, New York, New York, USA. ACM Press. 15

Chapiro, A., Atkins, R., and Daly, S. (2019). A Luminance-aware Model of Judder Perception. *ACM Transactions on Graphics*, 38(5):1–10. 69

Chen, J. Y. C. and Thropp, J. E. (2007). Review of Low Frame Rate Effects on Human Performance. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 37(6):1063–1076. 69

Chen, M. J., Su, C. C., Kwon, D. K., Cormack, L. K., and Bovik, A. C. (2013). Full-reference quality assessment of stereopairs accounting for rivalry. *Signal Processing: Image Communication*, 28(9):1143 – 1155. 22

Chollet, F. et al. (2015). Keras. https://keras.io. Last accessed: 2020-02-15. 74

Chung, S. T. L. and Legge, G. E. (2016). Comparing the shape of contrast sensitivity functions for normal and low vision. *Investigative Ophthalmology and Visual Science*, 57(1):198–207. 38

Claypool, K. T. and Claypool, M. (2007). On frame rate and player performance in first person shooter games. *Multimedia Systems*, 13(1):3–17. 69

Collett, D. (2003). *Modelling binary data*. Chapman & Hall/CRC, Boca Raton, 2nd edition. 22

Cruz, A. A., Garcia, D. M., Pinto, C. T., and Cechetti, S. P. (2011). Spontaneous Eyeblink Activity. *The Ocular Surface*, 9(1):29–41. 70

Curcio, C. A. and Hendrickson, A. E. (1991). Chapter 5 Organization and development of the primate photoreceptor mosaic. *Progress in Retinal Research*, 10:89–120. 2

Daitch, J. and Green, D. (1969). Contrast sensitivity of the human peripheral retina. *Vision Research*, 9(8):947–952. 33, 67

Daly, S. J. (1992). Visible differences predictor: an algorithm for the assessment of image fidelity. In Rogowitz, B. E., editor, *Human Vision, Visual Processing, and Digital Display III*, volume 1666, page 2. 9, 33

Davidson, R. R. (1970). On Extending the Bradley-Terry Model to Accommodate Ties in Paired Comparison Experiments. *Journal of the American Statistical Association*, 65(329):317–328. 107

Davies, R. B. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, 74(1):33–43. 52

Davis, M. K. and Chen, G. (2007). Graphing Kendall's. *Computational Statistics & Data Analysis*, 51(5):2375–2378. 95

De Silva, V., Arachchi, H. K., Ekmekcioglu, E., and Kondoz, A. (2013). Toward an Impairment Metric for Stereoscopic Video: A Full-Reference Video Quality Metric to Assess Compressed Stereoscopic Video. *IEEE Transactions on Image Processing*, 22(9):3392–3404. 22

De Valois, R. L., Morgan, H., and Snodderly, D. M. (1974). Psychophysical studies of monkey Vision-III. Spatial luminance contrast sensitivity tests of macaque and human observers. *Vision Research*, 14(1):75–81. 33, 51, 55, 66

De Vries, H. (1943). The quantum character of light and its bearing upon threshold of vision, the differential sensitivity and visual acuity of the eye. *Physica*, 10(7):553 – 564. 33

Depalma, J. J. and Lowry, E. M. (1962). Sine-Wave Response of the Visual System II Sine-Wave and Square-Wave Contrast Sensitivity*†. *Journal of the Optical Society of America*, 52(3):328. 33, 34, 51, 55

Di Stasi, L. L., Catena, A., Cañas, J. J., Macknik, S. L., and Martinez-Conde, S. (2013). Saccadic velocity as an arousal index in naturalistic tasks. *Neuroscience & Biobehavioral Reviews*, 37(5):968–975. 69

Didyk, P., Ritschel, T., Eisemann, E., Myszkowski, K., and Seidel, H. P. (2011). A perceptual model for disparity. *ACM Transactions on Graphics*, 30(4):1–10. 8

Dittrich, R., Hatzinger, R., and Katzenbeisser, W. (2002a). Modelling dependencies in paired comparison data. *Computational Statistics & Data Analysis*, 40(1):39–57. 75

Dittrich, R., Hatzinger, R., and Katzenbeisser, W. (2002b). Modelling the effect of subject-specific covariates in paired comparison studies with an application to university rankings. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 47(4):511–525. 75

Donner, K. and Hemilä, S. (1996). Modelling the spatio-temporal modulation response of ganglion cells with difference-of-Gaussians receptive fields: Relation to photoreceptor response kinetics. *Visual Neuroscience*, 13(1):173–186. 54

Egiazarian, K., Astola, J., Lukin, V., Battisti, F., and Carli, M. (2006). A new full-reference quality metrics based on HVS. In *CD-ROM Proceedings of the Second International Workshop on Video Processing and Quality Metrics, Scottsdale, USA, 2006*, page 4. 22

Egilmez, B., Schuchhardt, M., Memik, G., Ayoub, R., Soundararajan, N., and Kishinevsky, M. (2017). User-aware Frame Rate Management in Android Smartphones. *ACM Transactions on Embedded Computing Systems*, 16(5s):1–17. 68, 86

Emoto, M., Kusakabe, Y., and Sugawara, M. (2014). High-Frame-Rate Motion Picture Quality and Its Independence of Viewing Distance. *Journal of Display Technology*, 10(8):635–641. 69, 87

EUR-Lex (2020). European union law archive. https://eur-lex.europa.eu/. Last accessed: 15 Feb 2020. 71

Farnebäck, G. (2003). Two-Frame Motion Estimation Based on Polynomial Expansion. In Bigun, J. and Gustavsson, T., editors, *Image Analysis*, pages 363–370, Berlin, Heidelberg. Springer Berlin Heidelberg. 74

Feng, X.-f. (2006). LCD motion-blur analysis, perception, and reduction using synchronized backlight flashing. In Rogowitz, B. E., Pappas, T. N., and Daly, S. J., editors, *Human Vision and Electronic Imaging XI*, volume 6057, pages 213 – 226. International Society for Optics and Photonics, SPIE. 68

Fernandez-Alonso, M., Kaspiris-Rousellis, C., and Read, J. C. A. (2020). The effect of smartphone screen luminance and ambient illuminance on blink rate and visual discomfort. *Unpublished manuscript*. 68

Firth, D. (2004). Quasi-variances. *Biometrika*, 91(1):65–80. 76

Fogelton, A. and Benesova, W. (2016). Eye blink detection based on motion vectors analysis. *Computer Vision and Image Understanding*, 148:23–33. 75

Fogelton, A. and Benesova, W. (2018). Eye blink completeness detection. *Computer Vision and Image Understanding*, 176-177(June):78–85. 74

Fox, J. and Weisberg, S. (2019). *An R Companion to Applied Regression*. Sage, Thousand Oaks CA, third edition. 39, 75

Friedman, H. (1982). Simplified Determinations of Statistical Power, Magnitude of Effect and Research Sample Sizes. *Educational and Psychological Measurement*, 42(2):521–526. 99

García-Pérez, M. A. (2005). Is the DeVries-Rose to Weber Transition Empirically Possible with Sine-Wave Gratings? *The Spanish Journal of Psychology*, 8(2):113–118. 50, 53, 60, 62

García-Pérez, M. A. and Peli, E. (1997). The transition from DeVries-Rose to Weber's laws: Comments on Rovamo, Mustonen and Näsänen (1995). *Vision Research*, 37(18):2573–2576. 33, 50, 63

Gershun, A. (1939). The Light Field. *Journal of Mathematics and Physics*, 18(1-4):51–151. 14

Girod, B. (1992). Psychovisual aspects of image communication. *Signal Processing*, 28(3):239–251. 6

Gowrisankaran, S., Nahar, N. K., Hayes, J. R., and Sheedy, J. E. (2012). Asthenopia and Blink Rate Under Visual and Cognitive Loads. *Optometry and Vision Science*, 89(1):97–104. 69

Graham, N. (1985). Detection and identification of near-threshold visual patterns. *Journal of the Optical Society of America A*, 2(9):1468. 9

Graham, N. (1989). *Visual pattern analyzers*. Number no. 16 in Oxford psychology series. Oxford University Press, New York. 4, 12, 33, 49, 66

Green, D. M. and Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley, New York. 21

Häkkänen, H., LicPsych, Summala, H., Partinen, M., Tiihonen, M., and Silvo, J. (1999). Blink Duration as an Indicator of Driver Sleepiness in Professional Bus Drivers. *Sleep*, 22(6):798–802. 69

Halle, M. W. (1994). Holographic stereograms as discrete imaging systems. In Benton, S. A., editor, *Practical Holography VIII*, volume 2176, pages 73 – 84. International Society for Optics and Photonics, SPIE. 14

Han, H., Yu, J., Zhu, H., Chen, Y., Yang, J., Xue, G., Zhu, Y., and Li, M. (2013). E $^3$: energy-efficient engine for frame rate adaptation on smartphones. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems - SenSys '13*, pages 1–14, Roma, Italy. ACM Press. 68, 86

Hatzinger, R. and Dittrich, R. (2012). prefmod: An R Package for Modeling Preferences Based on Paired Comparisons, Rankings, or Ratings. *Journal of Statistical Software*, 48(10). 75

Hauske, G., Wolf, W., and Lupp, U. (1976). Matched filters in human vision. *Biological Cybernetics*, 22(4):181–188. 51

Heiberger, R. M. and Robbins, N. B. (2014). Design of Diverging Stacked Bar Charts for Likert Scales and Other Applications. *Journal of Statistical Software*, 57(5). 82

Hendley, C. D. (1948). The relation between visual acuity and brightness discrimination. *The Journal of General Physiology*, 31(5):433–457. 60

Hess, R. and Howell, E. (1988). Detection of low spatial frequencies: A single filter or multiple filters? *Ophthalmic and Physiological Optics*, 8(4):378–385. 33, 51, 55

Himebaugh, N. L., Begley, C. G., Bradley, A., and Wilkinson, J. A. (2009). Blinking and Tear Break-Up During Four Visual Tasks:. *Optometry and Vision Science*, 86(2):E106–E114. 69, 89

Hoffman, D. M., Girshick, A. R., Akeley, K., and Banks, M. S. (2008). Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, 8(3):33. 8, 73

Hoffman, D. M., Karasev, V. I., and Banks, M. S. (2011). Temporal presentation protocols in stereoscopic displays: Flicker visibility, perceived motion, and perceived depth. *J. Soc. Inf. Display*, 19(3):255. 69

Honauer, K., Johannsen, O., Kondermann, D., and Goldluecke, B. (2016). A dataset and evaluation methodology for depth estimation on 4d light fields. In *Asian Conference on Computer Vision*. Springer. 18, 21

Hood, D. C. and Finkelstein, M. A. (1986). Sensitivity to light. In Boff, K., Kaufman, L., and Thomas, J. P., editors, *Handbook of Perception and Human Performance, Sensory Processes and Perception*, volume 1. John Wiley & Sons, New York. 3

Isensee, S. H. and Bennett, C. A. (1983). The Perception of Flicker and Glare on Computer CRT Displays. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 25(2):177–184. 11

ISO 9241-303 (2011). Ergonomics of human-system interaction — Part 303: Requirements for electronic visual displays. Standard, International Organization for Standardization. 68

Janzen, B. F. and Teather, R. J. (2014). Is 60 FPS better than 30?: the impact of frame rate and latency on moving target selection. In *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems - CHI EA '14*, pages 1477–1482, Toronto, Ontario, Canada. ACM Press. 69

Jarvis, J. R. and Wathes, C. M. (2012). Mechanistic modeling of vertebrate spatial contrast sensitivity and acuity at low luminance. *Visual neuroscience*, 29(3):169–81. 66

Jia, H. and Wang, T. (2017). Contrast and visual saliency similarity induced index for image quality assessment. *CoRR*, abs/1708.06616. 22

Jie, Y., Sella, R., Feng, J., Gomez, M. L., and Afshari, N. A. (2019). Evaluation of incomplete blinking as a measurement of dry eye disease. *The Ocular Surface*, 17(3):440–446. 69

Johnston, P., Rodriguez, J., Lane, K., Ousler, and Abelson, M. (2013). The interblink interval in normal and dry eye subjects. *Clinical Ophthalmology*, page 253. 69

Jongkees, B. J. and Colzato, L. S. (2016). Spontaneous eye blink rate as predictor of dopamine-related cognitive function—A review. *Neuroscience & Biobehavioral Reviews*, 71:58–82. 89

Kandel, E., Schwartz, J., Jessell, T., Siegelbaum, S., and Hudspeth, A. J. (2012). *Principles of Neural Science, Fifth Edition.* McGraw-Hill. 1

Kane, D., Guan, P., and Banks, M. S. (2014). The Limits of Human Stereopsis in Space and Time. *Journal of Neuroscience*, 34(4):1397–1408. 8

Kaneko, K. and Sakamoto, K. (2001). Spontaneous Blinks as a Criterion of Visual Fatigue during Prolonged Work on Visual Display Terminals. *Perceptual and Motor Skills*, 92(1):234–250. 69, 88

Kaspiris-Rousellis, C., Siettos, C. I., Evdokimidis, I., and Smyrnis, N. (2017). Reaching to virtual targets: The oblique effect reloaded in 3-D. *Neuroscience*, 343:128–139. 61

Kaspiris-Rousellis, C., Simmons, A., and Read, J. (2018). Perceptually optimized view density for continuous parallax. *Journal of Vision*, 18(10):512. 22

Kassner, M., Patera, W., and Bulling, A. (2014). Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. *arXiv:1405.0006 [cs]*. 70

Kelly, D. (1972). Adaptation effects on spatio-temporal sine-wave thresholds. *Vision Research*, 12(1):89–IN1. 33, 55, 66

Kelly, D. H. (1979). Motion and vision II Stabilized spatio-temporal threshold surface. *Journal of the Optical Society of America*, 69(10):1340. 5

Kim, J., Hwang, Y., Kang, S., Kim, M., Kim, T.-S., Kim, J., Seo, J., Ahn, H., Yoon, S., Yun, J. P., Lee, Y. L., Ham, H., Yu, H. G., and Park, S. K. (2016). Association between Exposure to Smartphones and Ocular Health in Adolescents. *Ophthalmic Epidemiology*, 23(4):269–276. 68

Kim, K. J., Mantiuk, R., and Lee, K. H. (2013). Measurements of achromatic and chromatic contrast sensitivity functions for an extended range of adaptation luminance. In Rogowitz, B. E., Pappas, T. N., and de Ridder, H., editors, *Human Vision and Electronic Imaging XVIII*, volume 8651, pages 319 – 332. International Society for Optics and Photonics, SPIE. 33, 51, 55, 66

Kim, S.-R., Lee, S.-H., Jeon, D.-H., Kim, J.-S., and Lee, S.-W. (2017). Optimum display luminance dependence on ambient illuminance. *Optical Engineering*, 56(1):017110. 68

Kime, S., Galluppi, F., Lagorce, X., Benosman, R. B., and Lorenceau, J. (2016). Psychophysical Assessment of Perceptual Performance With Varying Display Frame Rates. *Journal of Display Technology*, 12(11):1372–1382. 69

Kingdom, F. A. A. and Prins, N. (2016). *Psychophysics: a practical introduction*. Elsevier/Academic Press, Amsterdam, 2nd edition. 24

Klein, S. A. (2001). Measuring, estimating, and understanding the psychometric function: A commentary. *Perception & Psychophysics*, 63(8):1421–1455. 22

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., and Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, 36(14):1–16. 35

Kontsevich, L. L. and Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, 39(16):2729 – 2737. 37

Kukkonen, H., Rovamo, J., Tiippana, K., and Näsänen, R. (1993). Michelson contrast, RMS contrast and energy of various spatial stimuli at threshold. *Vision Research*, 33(10):1431–1436. 4

Kurita, T. (2001). 35.1: Moving Picture Quality Improvement for Hold-type AM-LCDs. *SID Symposium Digest of Technical Papers*, 32(1):986. 86

Kuroki, Y., Nishi, T., Kobayashi, S., Oyaizu, H., and Yoshimura, S. (2007). A psychophysical study of improvements in motion-image quality by using high frame rates. *Journal of the Society for Information Display*, 15(1):61. 69, 87

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4. 99

Lambooij, M., Fortuin, M., Ijsselsteijn, W. A., and Heynderickx, I. (2009). Measuring visual discomfort associated with 3D displays. In Woods, A. J., Holliman, N. S., and Merritt, J. O., editors, *Stereoscopic Displays and Applications XX*, volume 7237, pages 169 – 180. International Society for Optics and Photonics, SPIE. 88

Larese Filon, F., Drusian, A., Ronchese, F., and Negro, C. (2019). Video Display Operator Complaints: A 10-Year Follow-Up of Visual Fatigue and Refractive Disorders. *International Journal of Environmental Research and Public Health*, 16(14):2501. 68

Larson, E. C. and Chandler, D. (2010). Most apparent distortion: Full-reference image quality assessment and the role of strategy. *J. Electronic Imaging*, 19:011006. 22

Legge, G. E. and Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America*, 70(12):1458. 10

Lenth, R. (2020). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.4.6. 39, 75

Levoy, M. and Hanrahan, P. (1996). Light field rendering. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques - SIGGRAPH '96*, pages 31–42. ACM Press. 14

Li, Y. (2002). Light beams with flat-topped profiles. *Optics Letters*, 27(12):1007. 19

Lin, Y. H., Chen, C. Y., Lu, S. Y., and Lin, Y. C. (2008). Visual fatigue during VDT work: Effects of time-based and environment-based conditions. *Displays*, 29(5):487–492. 86

Liu, T.-J., Lin, W., and Kuo, C.-C. J. (2013). Image Quality Assessment Using Multi-Method Fusion. *IEEE Transactions on Image Processing*, 22(5):1793–1807. 31

MacKenzie, K. J., Hoffman, D. M., and Watt, S. J. (2010). Accommodation to multiple-focal-plane displays: Implications for improving stereoscopic displays and for accommodation control. *Journal of Vision*, 10(8):22–22. 8

Mackin, A., Zhang, F., and Bull, D. R. (2019). A Study of High Frame Rate Video Formats. *IEEE Transactions on Multimedia*, 21(6):1499–1512. 69

Maffei, A. and Angrilli, A. (2018). Spontaneous eye blink rate: An index of dopaminergic component of sustained attention and fatigue. *International Journal of Psychophysiology*, 123:58–63. 69

Makous, W. (1998). Optics and photometry. In Carpenter, R. and Robson, J., editors, *Vision Research: A Practical Guide to Laboratory Methods*. Oxford University Press. 3

Mantiuk, R., Kim, K. J., Rempel, A. G., and Heidrich, W. (2011). HDR-VDP-2. In *ACM SIGGRAPH 2011 papers on - SIGGRAPH '11*, page 1, New York, New York, USA. ACM Press. 22, 33

Mayr, S., Köpper, M., and Buchner, A. (2017). Effects of high pixel density on reading comprehension, proofreading performance, mood state, and physical discomfort. *Displays*, 48:41–49. 68

McIntire, L. K., McKinley, R. A., Goodyear, C., and McIntire, J. P. (2014). Detection of vigilance performance using eye blinks. *Applied Ergonomics*, 45(2):354–362. 69, 88

Menozzi, M., Buol, A. v., Krueger, H., and Miège, C. (1994). Direction of gaze and comfort: discovering the relation for the ergonomic optimization of visual tasks. *Ophthalmic and Physiological Optics*, 14(4):393–399. 68

Miyake-Kashima, M., Dogru, M., Nojima, T., Murase, M., Matsumoto, Y., and Tsubota, K. (2005). The Effect of Antireflection Film Use on Blink Rate and Asthenopic Symptoms during Visual Display Terminal Work:. *Cornea*, 24(5):567–570. 68

Moler, C. B. (2004). *Numerical Computing with Matlab*, chapter 3. Interpolation, pages 93–116. Society for Industrial and Applied Mathematics. 52

Moller, C. and Travis, A. (2005). Correcting interperspective aliasing in autostereoscopic displays. *IEEE Transactions on Visualization and Computer Graphics*, 11(2):228–236. 14

Morris, T. and Miller, J. (1996). Electrooculographic and performance indices of fatigue during simulated flight. *Biological Psychology*, 42(3):343–360. 69, 88

Moscatelli, A., Mezzetti, M., and Lacquaniti, F. (2012). Modeling psychophysical data at the population-level: The generalized linear mixed model. *Journal of Vision*, 12(11):26–26. 32

Muggeo, V. M. R. (2003). Estimating regression models with unknown break-points. *Statistics in Medicine*, 22(19):3055–3071. 52

Mustonen, J., Rovamo, J., and Näsänen, R. (1993). The effects of grating area and spatial frequency on contrast sensitivity as a function of light level. *Vision Research*, 33(15):2065–2072. 49, 51, 55, 58, 66

Nadler, J. W., Nawrot, M., Angelaki, D. E., and DeAngelis, G. C. (2009). MT Neurons Combine Visual Motion with a Smooth Eye Movement Signal to Code Depth-Sign from Motion Parallax. *Neuron*, 63(4):523–532. 7

Nakamura, J., Tanaka, K., and Takaki, Y. (2013). Increase in Depth of Field of Eyes Using Reduced-View Super Multi-View Displays. *Applied Physics Express*, 6(2):022501. 8

Nasiri, R. M., Wang, J., Rehman, A., Wang, S., and Wang, Z. (2015). Perceptual quality assessment of high frame rate video. In *2015 IEEE 17th International Workshop on Multimedia Signal Processing (MMSP)*, pages 1–6, Xiamen, China. IEEE. 69

NEMA (2020). Ps 3.14, digital imaging and communications in medicine (dicom) standard, part 14: Grayscale standard display function. 33

Osterberg, G. (1935). Topography of the layer of rods and cones in the human retina. *Acta Ophthalmologica*, Supplement 6:1–103. 2

Ousler, G., Abelson, M. B., Johnston, P. R., Rodriguez, J., Lane, K., and Smith, L. M. (2014). Blink patterns and lid-contact times in dry-eye and normal subjects. *Clinical Ophthalmology*, page 869. 89

Patel, A. S. (1966). Spatial Resolution by the Human Visual System The Effect of Mean Retinal Illuminance*. *Journal of the Optical Society of America*, 56(5):689. 33, 51, 66

Paudyal, P., Battisti, F., Sjöström, M., Olsson, R., and Carli, M. (2017). Towards the Perceptual Quality Evaluation of Compressed Light Field Images. *IEEE Transactions on Broadcasting*, 63(3):507–522. 17

Peli, E., Arend, L., and Labianca, A. T. (1996). Contrast perception across changes in luminance and spatial frequency. *Journal of the Optical Society of America. A, Optics, image science, and vision*, 13(10):1953–9. 33, 51, 55

Pelli, D. (1990). The quantum efficiency of vision. In Blakemore, C., editor, *Vision: Coding and efficiency*, pages 3–24. Cambridge University Press. 33, 51

Pelli, D. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Multisensory research*, 10(4):437–442. 35

Perez-Ortiz, M. and Mantiuk, R. K. (2017). A practical guide and software for analysing pairwise comparison experiments. *arXiv:1712.03686 [cs, stat]*. 32

Pinson, M. H. and Wolf, S. (2004). A new standardized method for objectively measuring video quality. *IEEE Transactions on Broadcasting*, 50(3):312–322. 22

Ponomarenko, N., Silvestri, F., Egiazarian, K., Carli, M., Astola, J., and Lukin, V. (2007). On between-coefficient contrast masking of dct basis functions. In *CD-ROM proceedings of Third International Workshop on Video Processing and Quality Metrics for Consumer Electronics VPQM-07, January, 2007*, page 4. 22

Portello, J. K., Rosenfield, M., Bababekova, Y., Estrada, J. M., and Leon, A. (2012). Computer-related visual symptoms in office workers: Visual symptoms in office workers. *Ophthalmic and Physiological Optics*, 32(5):375–382. 68

Portello, J. K., Rosenfield, M., and Chu, C. A. (2013). Blink Rate, Incomplete Blinks and Computer Vision Syndrome:. *Optometry and Vision Science*, 90(5):482–487. 69

Prins, N. (2013). The psi-marginal adaptive method: How to give nuisance parameters the attention they deserve (no more, no less). *Journal of Vision*, 13(7):3–3. 37

Prins, N. and Kingdom, F. A. A. (2018). Applying the Model-Comparison Approach to Test Specific Research Hypotheses in Psychophysical Research Using the Palamedes Toolbox. *Frontiers in Psychology*, 9:1–14. 22, 37

Pruessner, J. C. (2004). Dopamine Release in Response to a Psychological Stress in Humans and Its Relationship to Early Life Maternal Care: A Positron Emission Tomography Study Using [11C]Raclopride. *Journal of Neuroscience*, 24(11):2825–2831. 89

Rasengane, T. A., Palmer, J., and Teller, D. Y. (2001). Infant light adaptation shows Weber's law at photopic illuminances. *Vision Research*, 41(3):359–373. 51, 55

Robson, J. G. (1966). Spatial and Temporal Contrast-Sensitivity Functions of the Visual System. *Journal of the Optical Society of America*, 56(8):1141. 5

Rohatgi, A. (2012). Web plot digitalizer: Html5 based online tool to extract numerical data from plot images. version 4.2. https://automeris.io/WebPlotDigitizer/index.html. Last accessed: July 2019. 52

Rose, A. (1942). The relative sensitivities of television pickup tubes, photographic film, and the human eye. *Proceedings of the IRE*, 30(6):293–300. 33

Rovamo, J., Kukkonen, H., Thppana, K., and Näsänen, R. (1993). Effects of luminance and exposure time on contrast sensitivity in spatial noise. *Vision Research*, 33(8):1123–1129. 51

Rovamo, J., Mustonen, J., and Näsänen, R. (1994). Modelling contrast sensitivity as a function of retinal illuminance and grating area. *Vision Research*, 34(10):1301–1314. 12, 33, 34, 51, 53, 54, 55, 58, 59, 64, 66

Rovamo, J., Mustonen, J., and Näsänen, R. (1995). Neural modulation transfer function of the human visual system at various eccentricities. *Vision Research*, 35(6):767–774. 33

Rovamo, J., Näsänen, R., and Mustonen, J. (1997). Transition from DeVries-Rose to Weber's law: Reply to García-Pérez and Peli (1997). *Vision Research*, 37(18):2576–2578. 12, 33, 53, 64

Runde, D. (2000). How to realize a natural image reproduction using stereoscopic displays with motion parallax. *IEEE Transactions on Circuits and Systems for Video Technology*, 10(3):376–386. 31

Sampat, M. P., Wang, Z., Gupta, S., Bovik, A. C., and Markey, M. K. (2009). Complex wavelet structural similarity: A new image similarity index. *IEEE Transactions on Image Processing*, 18(11):2385–2401. 22

Schleicher, R., Galley, N., Briest, S., and Galley, L. (2008). Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? *Ergonomics*, 51(7):982–1010. 69

Seetzen, H., Heidrich, W., Stuerzlinger, W., Ward, G., Whitehead, L., Trentacoste, M., Ghosh, A., and Vorozcovs, A. (2004). High dynamic range display systems. *ACM Transactions on Graphics*, 23(3):760. 34

Selfridge, R., Noland, K. C., and Hansard, M. (2016). Visibility of Motion Blur and Strobing Artefacts in Video at 100 Frames per Second. In *Proceedings of the 13th European Conference on Visual Media Production (CVMP 2016) - CVMP 2016*, pages 1–10, London, United Kingdom. ACM Press. 69

Shannon, E., Skoczenski, A. M., and Banks, M. S. (1996). Retinal illuminance and contrast sensitivity in human infants. *Vision Research*, 36(1):67–76. 51, 55, 59, 66, 67

Sheikh, H., Bovik, A., and de Veciana, G. (2005). An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Transactions on Image Processing*, 14(12):2117–2128. 22

Sheppard, A. L. and Wolffsohn, J. S. (2018). Digital eye strain: prevalence, measurement and amelioration. *BMJ Open Ophthalmology*, 3(1):e000146. 69

Shibata, T., Kim, J., Hoffman, D. M., and Banks, M. S. (2011). The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of Vision*, 11(8):11–11. 8, 30

Shieh, K.-K. (2000). Effects of reflection and polarity on LCD viewing distance. *International Journal of Industrial Ergonomics*, 25(3):275–282. 68

Silverstein, D. and Farrell, J. (1996). The relationship between image fidelity and image quality. In *Proceedings of 3rd IEEE International Conference on Image Processing*, volume 1, pages 881–884, Lausanne, Switzerland. IEEE. 69

Silvestre, D., Arleo, A., and Allard, R. (2018). Internal noise sources limiting contrast sensitivity. *Scientific reports*, 8(1):2596. 33, 34, 51, 55, 64

Simoncelli, E. and Freeman, W. (1995). The steerable pyramid: a flexible architecture for multi-scale derivative computation. In *Proceedings., International Conference on Image Processing*, volume 3, pages 444–447, Washington, DC, USA. 10

Singmann, H., Bolker, B., Westfall, J., Aust, F., and Ben-Shachar, M. S. (2020). *afex: Analysis of Factorial Experiments*. http://afex.singmann.science/, https://github.com/singmann/afex. 39, 75

Smithson, M. (2001). Correct Confidence Intervals for Various Regression Effect Sizes and Parameters: The Importance of Noncentral Distributions in Computing Intervals. *Educational and Psychological Measurement*, 61(4):605–632. 43, 80, 99

Speranza, F., Tam, W. J., Martin, T., Stelmach, L., and Ahn, C.-H. (2005). Perceived smoothness of viewpoint transition in multi-viewpoint stereoscopic displays. In Woods, A. J., Bolas, M. T., Merritt, J. O., and McDowall, I. E., editors, *Stereoscopic Displays and Virtual Reality Systems XII*, volume 5664, pages 72 – 82. International Society for Optics and Photonics, SPIE. 31

Stanley, P. A. and Davies, A. K. (1995). The effect of field of view size on steady-state pupil diameter. *Ophthalmic and Physiological Optics*, 15(6):601–603. 52

Stern, J. A., Boyer, D., and Schroeder, D. J. (1994a). Blink rate as a measure of fatigue: A review. Technical Report DOT/FAA/AM-94/17, Office of Aviation Medicine, Washington, DC 20591. 69, 88

Stern, J. A., Boyer, D., Schroeder, D., Touchstone, M., and Stoliarov, N. (1994b). Blinks, Saccades, and Fixation Pauses During Vigilance Task Performance: I. Time on Task. Technical Report DOT/FAA/AM-94/26, Office of Aviation Medicine, Washington, DC 20591. 69, 88

Stiles, W. S., Crawford, B. H., and Parsons, J. H. (1933). The luminous efficiency of rays entering the eye pupil at different points. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 112(778):428–450. 38, 52

Summala, H., Hakkanen, H., Mikkola, T., and Sinkkonen, J. (1999). Task effects on fatigue symptoms in overnight driving. *Ergonomics*, 42(6):798–806. 69

Tag, B., Shimizu, J., Zhang, C., Kunze, K., Ohta, N., and Sugiura, K. (2016). In the Eye of the Beholder: The Impact of Frame Rate on Human Eye Blink. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '16*, pages 2321–2327, San Jose, California, USA. ACM Press. 70, 89

Takaki, Y., Urano, Y., and Nishio, H. (2012). Motion-parallax smoothness of short-, medium-, and long-distance 3D image presentation using multi-view displays. *Opt. Express*, 20(24):27180. 31

Tanner, W. P. and Birdsall, T. G. (1958). Definitions of $d'$ and $\eta$ as Psychophysical Measures. *The Journal of the Acoustical Society of America*, 30(10):922–928. 53

The Stanford 3D Scanning Repository (2020). Scan data of a dragon sculpture constructed from wood pulp resin. http://graphics.stanford.edu/data/3Dscanrep/. Last accessed: May 2020. 20

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, 34(4):273–286. 32

Tourancheau, S., Andrén, B., Brunnström, K., and Callet, P. L. (2009). 61.3: Visual Annoyance and User Acceptance of LCD Motion-Blur. *SID Symposium Digest of Technical Papers*, 40(1):919. 86

Unsworth, N. and Robison, M. K. (2016). Pupillary correlates of lapses of sustained attention. *Cognitive, Affective, & Behavioral Neuroscience*, 16(4):601–615. 69

Vagharshakyan, S., Bregovic, R., and Gotchev, A. (2018). Light Field Reconstruction Using Shearlet Transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(1):133–147. 32

Valero, E. M., Nieves, J. L., Hernández-Andrés, J., and García, J. A. (2004). Changes in contrast thresholds with mean luminance for chromatic and luminance gratings: A reexamination of the transition from the DeVries-Rose to Weber regions. *Color Research & Application*, 29(3):177–182. 51, 55, 59

van Meeteren, A. and Vos, J. (1972). Resolution and contrast sensitivity at low luminances. *Vision Research*, 12(5):825–IN2. 33

Van Nes, F. L. and Bouman, M. A. (1967). Spatial Modulation Transfer in the Human Eye. *Journal of the Optical Society of America*, 57(3):401. 12, 33, 34, 51, 55, 57, 62

Van Orden, K. F., Jung, T.-P., and Makeig, S. (2000). Combined eye activity measures accurately estimate changes in sustained visual task performance. *Biological Psychology*, 52(3):221–240. 69, 88

Vangorp, P., Richardt, C., Cooper, E. A., Chaurasia, G., Banks, M. S., and Drettakis, G. (2013). Perception of perspective distortions in image-based rendering. *ACM Trans. Graph.*, 32(4):1. 16

Walraven, J., Enroth-Cugell, C., Hood, D. C., MacLeod, D. I. A., and Schnapf, J. L. (1990). The control of visual sensitivity: receptoral and postreceptoral processes. In Spillmann, L. and Werner, J. S., editors, *Visual perception: the neurophysiological foundations*. Academic Press, San Diego. 3

Wang, Y.-T., Liu, Y.-N., Chien, S.-Y., Yang, Y.-H., Chen, Y.-L., and Yeh, S.-L. (2009). P-36: A Psychophysical Analysis on Perceptual Limitation of Motion Image Artifact Reduction Using 120Hz Displays. *SID Symposium Digest of Technical Papers*, 40(1):1223. 69

Wang, Z. and Bovik, A. (2009). Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures. *IEEE Signal Processing Magazine*, 26(1):98–117. 8

Wang, Z., Bovik, A., Sheikh, H., and Simoncelli, E. (2004). Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, 13(4):600–612. 9, 10, 22

Wang, Z. and Bovik, A. C. (2006). Modern Image Quality Assessment. *Synthesis Lectures on Image, Video, and Multimedia Processing*, 2(1). 9

Wang, Z. and Li, Q. (2011). Information content weighting for perceptual image quality assessment. *IEEE Transactions on Image Processing*, 20(5):1185–1198. 22

Wang, Z., Simoncelli, E., and Bovik, A. (2003). Multiscale structural similarity for image quality assessment. In *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, volume 2, pages 1398–1402. IEEE. 22

Watson, A. B. (1987). The cortex transform: Rapid computation of simulated neural images. *Computer Vision, Graphics, and Image Processing*, 39(3):311–327. 9

Watson, A. B. and Ahumada, A. J. (2005). A standard model for foveal detection of spatial contrast. *Journal of Vision*, 5(9):6. 38, 54, 67

Watson, A. B. and Ahumada, A. J. (2011). 64.3: Flicker Visibility: A Perceptual Metric for Display Flicker. *SID Symposium Digest of Technical Papers*, 42(1):957–959. 11

Watson, A. B., Ahumada, A. J., and Farrell, J. E. (1986). Window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays. *Journal of the Optical Society of America A*, 3(3):300. 5, 69

Watson, A. B. and Yellott, J. I. (2012). A unified formula for light-adapted pupil size. *Journal of Vision*, 12(10):12–12. 52

Watson, A. B. . (2013a). High Frame Rates and Human Vision: A View through the Window of Visibility. *SMPTE Motion Imaging Journal*, 122(2):18–32. 6, 69

Watson, A. B. . (2013b). A formula for the mean human optical modulation transfer function as a function of pupil size. *Journal of Vision*, 13(6):18. 54

Wichmann, F. A. and Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, 63(8):1293–1313. 22

Winkler, S. (2001). Visual fidelity and perceived quality: toward comprehensive metrics. In Rogowitz, B. E. and Pappas, T. N., editors, *Human Vision and Electronic Imaging VI*, volume 4299, pages 114 – 125. International Society for Optics and Photonics, SPIE. 9

Wolkoff, P. (2005). Eye complaints in the office environment: precorneal tear film integrity influenced by eye blinking efficiency. *Occupational and Environmental Medicine*, 62(1):4–12. 69, 89

Wright, S. L., Bailey, I. L., Tuan, K.-M., and Wacker, R. T. (1999). Resolution and legibility: A comparison of TFT-LCDs and CRTs. *Journal of the Society for Information Display*, 7(4):253. 68

Xue, W., Zhang, L., Mou, X., and Bovik, A. C. (2014). Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. *IEEE Transactions on Image Processing*, 23(2):668–695. 22

Yamaguchi, M. (2016). Light-field and holographic three-dimensional displays [Invited]. *J. Opt. Soc. Am. A*, 33(12):2348. 11

Ye, Z., Abe, Y., Kusano, Y., Takamura, N., Eida, K., Takemoto, T.-I., and Aoyagi, K. (2007). The Influence of Visual Display Terminal Use on the Physical and Mental Conditions of Administrative Staff in Japan. *Journal of Physiological Anthropology*, 26(2):69–73. 68

Yoss, R. E., Moyer, N. J., and Hollenhorst, R. W. (1970). Pupil size and spontaneous pupillary waves associated with alertness, drowsiness, and sleep. *Neurology*, 20(6):545–545. 69

Zargari Marandi, R., Madeleine, P., Omland, Ø., Vuillerme, N., and Samani, A. (2018). Eye movement characteristics reflected fatigue development in both young and elderly individuals. *Scientific Reports*, 8(1):13148. 69, 88

Zele, A. J. and Cao, D. (2015). Vision under mesopic and scotopic illumination. *Frontiers in Psychology*, 5. 60

Zhai, G. and Min, X. (2020). Perceptual image quality assessment: a survey. *Science China Information Sciences*, 63(11):211301. 10

Zhang, L., Shen, Y., and Li, H. (2014). VSI: A Visual Saliency-Induced Index for Perceptual Image Quality Assessment. *IEEE Transactions on Image Processing*, 23(10):4270–4281. 22

Zhang, L., Zhang, L., Mou, X., and Zhang, D. (2011). Fsim: A feature similarity index for image quality assessment. *IEEE Transactions on Image Processing*, 20(8):2378–2386. 22

Ziefle, M. (1998). Effects of Display Resolution on Visual Performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40(4):554–568. 11, 68

Zwicker, M., Matusik, W., Durand, F., and Pfister, H. (2006). Antialiasing for automultiscopic 3d displays. In *Symposium on Rendering*, EGSR '06, page 73–82, Goslar, DEU. Eurographics Association. 8