castleCSF — A Contrast Sensitivity Function of Color, Area, Spatio-Temporal Frequency, Luminance and Eccentricity

Maliha Ashraf

Department of Computer Science and Technology
University of Cambridge, Cambridge, UK

Rafał K. Mantiuk

Department of Computer Science and Technology
University of Cambridge, Cambridge, UK

Alexandre Chapiro

Applied Perception Science Group
Meta, Sunnyvale, USA

Sophie Wuerger

Department of Psychology
University of Liverpool, Liverpool, UK

The Contrast Sensitivity Function (CSF) is a fundamental visual model explaining our ability to detect small contrast patterns. CSFs found many applications in engineering where they can be used to optimize a design for perceptual limits. To serve such a purpose, CSF must explain possibly a complete set of stimulus parameters, such as spatial and temporal frequency, luminance, and others. Although numerous contrast sensitivity measurements can be found in the literature, none fully explains the complete space of stimulus parameters. Therefore, in this work, we first collect and consolidate contrast sensitivity measurements from 18 studies, which explain the sensitivity variation across the parameters of interest. Then, we build an analytical contrast sensitivity model that explains the data from all those studies. The proposed castleCSF model explains the sensitivity as the function of spatial and temporal frequencies, an arbitrary contrast modulation direction in the color space, mean luminance and chromaticity of the background, eccentricity, and stimulus area. The proposed model uses the same set of parameters to explain the data from 18 studies with an error of 3.59 dB. The consolidated contrast sensitivity data and the code for the model are publicly available at https://github.com/gfxdisp/castleCSF/.

Keywords: contrast sensitivity functions, color vision, computational modeling

Introduction

A contrast sensitivity function (CSF) is a visual model that explains what is the smallest contrast of a simple stimulus (a sinusoidal grating, a Gabor patch, or a disk) that can be detected on a uniform background (Robson, 1966; Kelly, 1979; Barten, 1999). These types of detection characteristics are fundamental performance measures of the visual system and have been thoroughly measured over several decades using psychophysical methods. Despite the abundance of data, few models summarize these measurements across more than three of the many parameters influencing the visibility of the stimuli. Creating such a model is the goal of our work. Our model explains the available psychophysical data across the most relevant stimulus parameters: the chromaticity of the background, the chromatic direction of modulation, area, spatial and temporal frequency, luminance, and eccentricity. The first letters of those dimensions form the acronym of our model — castleCSF. To facilitate multiple applications and re-

search on sensitivity, we made our model and the data available at https://github.com/gfxdisp/castleCSF.

Although contrast sensitivity explains visual performance only for a limited class of artificial stimuli, it is a building block of more complex models of the visual system, which can generalize to more comprehensive stimuli, including images or videos. For example, while contrast masking has different characteristics depending on the luminance, spatial frequency, and color modulation direction of the patterns, masking characteristics can be unified when contrast is normalized by the sensitivity predicted by the CSF (Daly, 1993; Cass, Clifford, Alais, & Spehar, 2009). While the perceived magnitude of suprathreshold contrast varies across luminance, the deviation from contrast constancy can once again be explained using contrast sensitivity data (Kulikowski, 1976; Peli, 1995). The detection of more complex patterns can be estimated using energy models, which also rely on the CSF (Watson, 2000). Finally, visible differences on complex backgrounds can be explained by multi-channel visual difference predictors (Daly, 1993; Mantiuk, Kim, Rempel, & Heidrich, 2011), which incorporate models of both visual masking and contrast sensitivity.

CSFs are also commonly used to control and optimize display applications, and in particular virtual and augmented reality (VR/AR) headmounted systems. In this work, we focus on the dimensions of the CSF that are the most relevant in the context of color video applications, in particular for AR/VR. Notably, as these displays typically cover a significantly larger field-of-view than traditional displays, much of the content lies outside the user's fovea, and despite presenting very high physical pixel density and resolution, the effective pixel per visual degree (ppd) resolution is often lower than conventional displays. Perceptually-informed rendering algorithms, like contrast-aware (Tursun et al., 2019) or attention-aware (Krajancich, Kellnhofer, & Wetzstein, 2023) foveated rendering, and foveated image reconstruction (Kaplanyan et al., 2019) are proposed to achieve high perceptual content quality while conforming to the limited power and compute budget. Due to their head-tracked nature, AR/VR often requires ultra-high refresh rates and low persistence values, resulting in heavier computational and bandwidth costs. Rendering modalities where a subset of frames are simplified can be used to alleviate this issue (Denes, Maruszczyk, Ash, & Mantiuk, 2019). Recent work by Duinkharjav et al., 2022 explicitly models foveated color discrimination to generate chromatic distortions that can minimize display power usage. A comprehensive model of chroma-aware contrast sensitivity is especially important to guide these kinds of AR/VR algorithm development and is provided by castleCSF.

Factors affecting contrast sensitivity

Contrast sensitivity functions describe the visibility of gratings through Fourier analysis (Campbell & Robson, 1968) — explaining the detection characteristic as a function of spatial and temporal frequency of the visual stimulus. This is motivated by the existence of visual channels, which are tuned to bands of spatial and temporal frequencies (Kulikowski & Robson, 1999). However other parameters of the stimulus have no lesser influence on the sensitivity. The sensitivity is affected by luminance and chromaticity of the background (Blackwell, 1946; Xu, Ye, Mantiuk, & Luo, 2022), the size of the stimulus (Rovamo, Luntinen, & Näsänen, 1993), the angle from the fixation point (eccentricity) (Virsu & Rovamo, 1979; Wright & Johnston, 1983) and the position in the peripheral visual field (nasal, temporal, superior or inferior) (Anderson, Mullen, & Hess, 1991), the orientation of a grating (Campbell, Kulikowski, & Levinson, 1966), and the direction along which the contrast is modulated in a three-dimensional color space (Mullen, 1985; Wuerger et al., 2020). It is also recognized that the sensitivity decreases with age, mostly due to optical factors, such as the increase in the wavelength-dependent opacity (Pokorny, Smith, & Lutze, 1987) or the inability of the pupil to dilate at low light (senile miosis) (Watson & Yellott, 2012), but also due to neural factors. The sensitivity is reduced at shorter viewing distances due to the error in accommodation (Hernández, Doménech, Seguí, & Illueca, 1996). When a field of different luminance surrounds a detected pattern, the sensitivity is reduced (Yi et al., 2022) by glare and local luminance adaptation (Vangorp, Myszkowski, Graf, & Mantiuk, 2015). Binocular sensitivity is higher than monocular sensitivity and can be predicted from monocular sensitivity by quadratic summation ($S_{\text{binocular}} = \sqrt{S_L^2 + S_R^2}$) (Legge, 1984).

Measurements of contrast sensitivity

Measurements are most often conducted via non-invasive psychophysical methods (Campbell & Green, 1965; Pointer & Hess, 1989), where participants are shown stimuli of different contrast, and their performance in a detection task is measured. Alternative approaches use electrophysiological methods such as electroretinog-

raphy (ERG) (Hood & Birch, 1990) and visual evoked potential (VEP) recordings (Norcia, Tyler, & Hamer, 1990), where the electrical response to the stimulus that originates in the retina or the visual cortex is measured directly. The characterization of optical characteristics using techniques such as optical coherence tomography (OCT) (Adam, Shrier, Ding, Glazman, & Bodis-Wollner, 2013) and retinal imaging (Ortiz, Jiménez, Pérez-Ocón, Castro, & González-Anera, 2010) can also provide insights into contrast sensitivity. Data from all these different kinds of studies is often impossible to compare directly due to the differences in methodology and control conditions. Even among psychophysical studies, there is no clear consensus regarding standard test conditions or the type of stimuli used to measure the visual system's responses.

Because of the very large space of possible stimulus parameters, existing contrast sensitivity studies typically measure the variation along 2–3 selected parameters to maintain a feasible length of a study. It follows that creating a CSF that models a larger set of dimensions requires combining data from multiple studies. This poses several challenges. First, each measurement often uses slightly different stimuli (e.g. Gabors, stimuli with sharp edges, annulus rings, different shapes of the aperture, etc.), as well as different protocols and viewing conditions (monocular vs. binocular, natural vs. artificial pupil). To account for these variations we allow for small relative changes in the sensitivity between the datasets. Moreover, many dimensions of the modeled CSF show strong interactions. For example, the spatial and temporal frequency dimensions of the CSF are not separable and cannot be modeled as a combination of independent spatial and temporal CSFs. To create a comprehensive model, it is necessary to fit a model to multiple datasets at the same time. This approach is different from what is found in previous work modeling multidimensional CSFs (Barten, 1999; Watson & Ahumada, 2016), in which datasets were fitted one at a time. For the sake of simplicity, castleCSF also does not attempt to faithfully reproduce early vision mechanisms, such as those modeled in the ISETBIO package (Brainard et al., 2015). We intend to create a practical model, which can summarize and predict a large body of data from the literature.

Modeling contrast sensitivity

Several relevant CSF models are listed in Table 1 and are discussed in this section. Kelly was one of the first to introduce a model of the interaction of spatiotemporal mechanisms (Kelly, 1979). Notably, Kelly pointed out that the spatiotemporal CSF can alternatively be represented as the function of spatial frequency and retinal velocity, which can be more relevant for some applications. His model was based on measurements with retinal images stabilized via eye tracking, which is not representative of typical image or video content consumption use cases. We found Kelly's data to be too distinct from comparable measurements with non-stabilized stimuli to be used in our model. Daly, 1998 extended Kelly's model, which was later fitted to data for non-stabilized stimuli (Laird, Rosen, Pelz, Montag, & Daly, 2006). All these models factor only achromatic stimuli and do not consider other stimulus properties such as luminance or stimulus area.

A comprehensive CSF model can be found in the Visual Difference Predictor (VDP) work of Daly, 1993, though details on how the model was fitted to data are omitted. This method models the effect of spatial frequency, luminance, area, and eccentricity on the contrast sensitivity of static achromatic stimuli. Barten, 1999, proposed a comprehensive CSF, explicitly modeling the optical transfer function, photoreceptor and neural noise, and lateral inhibition. The original model included the effect of spatial frequency, area, and luminance (Barten, 1992). The full model also includes extensions to the parafoveal and temporal domain (Barten, 1999). Moreover, Bozorgian, Pedersen, & Thomas, 2022 have shown that Barten's model could be modified to improve predictions for peripheral contrast sensitivity and validated this with a dataset from Thibos, Still, & Bradley, 1996. Barten's model parameters were fitted individually to multiple datasets from the literature, demonstrating its ability to explain a wide range of measurements. Despite this, in our experiments with Barten's CSF, we found it unable to explain more recent datasets measured at high luminance levels (Wuerger et al., 2020), in particular, the loss of sensitivity at high luminance for low frequency achromatic Gabor patterns. This model also lacks the chromatic component necessary to represent sensitivity to color modulations.

One of the few works that model chromatic stimuli was proposed by Rovamo et al., 1999. The authors modeled the low-pass behavior of spatial contrast sensitivity functions for isoluminant stimuli, as these do not generate lateral inhibition. The detection mechanism was modeled as the combination of optical and neural modulation transfer functions, both affected by noise. The model explained the physiological mechanisms that lead to differences

Table 1: This table shows a selection of popular CSF models from the literature. Notes: (1) for stabilized retinal images. (2) only modeled for the three selected color directions. (3) excluding the region in which spatial frequency is below about 10 cpd, and temporal frequency is below 10 Hz.

Model	Spatial freq.	Temp. freq.	Luminance	Background chromaticity	Area	Eccentricity	Color modulation directions
Kelly's spatiotemporal CSF (Kelly, 1979)	✓	✓[1]	✗	✗	✗	✗	✗
Barten's CSF (Barten, 1999)	✓	✓	✓	✗	✓	✓	✗
VDP CSF (Daly, 1993)	✓	✗	✓	✗	✓	✓	✗
Rovamo et al. CSF (Rovamo, Kankaanpää, & Kukkonen, 1999)	✓	✗	✗	✗	✗	✗	✓[2]
Pyramid of visibility (a) (Watson & Ahumada, 2016; Watson, 2018)	✓[3]	✓[3]	✓	✗	✓	✓	✗
Chromatic pyramid of visibility (b) (Watson, 2021)	✓[3]	✓[3]	✓	✗	✗	✗	✓
Wuerger et al. spatio-chromatic CSF (Wuerger et al., 2020)	✓	✗	✓	✗	✓	✗	✓[2]
Mantiuk et al. spatio-chromatic CSF (Mantiuk et al., 2020)	✓	✗	✓	✓	✓	✗	✓
stelaCSF (Mantiuk, Ashraf, & Chapiro, 2022)	✓	✓	✓	✗	✓	✓	✗
castleCSF (ours)	✓	✓	✓	✓	✓	✓	✓

between the detection of achromatic and chromatic gratings and produced good fits to Mullen's contrast sensitivity data for isoluminant stimuli (Mullen, 1985). However, the model cannot be easily generalized to an arbitrary color modulation direction, and therefore, cannot be used to predict data from many of our datasets. In addition to these chromatic CSF models, some recent studies suggested that spatio-chromatic CSFs can be inferred from the features of deep-neural networks trained on low- and middle-levels tasks, such as image-denoising, autoencoding, edge detection and object recognition (Akbarinia, Morgenstern, & Gegenfurtner, 2023; Li, Gomez-Villa, Bertalmío, & Malo, 2022). To demonstrate that, the authors train a classifier to discriminate contrast and use it as a decision criterion in a two-alternative-choice contrast detection experiment. The results, however, show only a correlation with human data and cannot be used to give accurate predictions of human sensitivity.

The pyramid of visibility is a popular simplified model of contrast sensitivity. The original model (Watson & Ahumada, 2016) accounts for spatial and temporal frequencies, and background luminance, while the extended version (Watson, 2018) adds stimulus area and retinal eccentricity parameters. The pyramid of visibil-

ity establishes linear and log-linear relationships between contrast sensitivity and the studied parameters of the stimulus. This linear behavior of the CSF is only observed at high spatial and temporal frequencies (above around 10 cpd or 10 Hz). A chromatic extension of the pyramid of visibility, which models the effect of spatial frequency, temporal frequency, and luminance for stimuli with modulations in both achromatic and chromatic space, has also been proposed (Watson, 2021). This model works by estimating the projections of the stimulus contrast modulation for hypothetical achromatic, red-green, and yellow-violet opponent mechanisms of the visual system. The sensitivity of each of these three mechanisms is summed together to predict the overall sensitivity for a given stimulus.

castleCSF builds and expands upon our previous models of contrast sensitivity. Wuerger et al., 2020 presented a new high-dynamic-range CSF dataset, and a CSF modeled as a function of spatial frequency, mean luminance, and stimulus size for the three cardinal chromatic directions of the visual system. Mantiuk et al., 2020 extended this work by proposing two spatio-chromatic CSFs, which could additionally account for the chromaticity of the background, and any modulation direction in a color space. The

model was fitted to the combined data from five spatio-chromatic datasets. In another work, we proposed stelaCSF (Mantiuk et al., 2022), which was limited to achromatic contrast sensitivity but introduced temporal frequency, and eccentricity. stelaCSF was fitted to 11 datasets from the literature. This work combines the models and approaches used in stelaCSF (Mantiuk et al., 2022) and the post-receptoral spatio-chromatic CSF (Mantiuk et al., 2020). Although Mantiuk et al., 2020 showed slightly better predictions for the cone-contrast variant of the model, we selected the post-receptoral variant as it let us better isolate achromatic and chromatic mechanisms. We employ a data-driven, physiologically-inspired approach to model contrast sensitivity as a function of 6 different stimulus properties, which is the superset of all parameters of the previous models (refer to Table 1). We do not model the effect of orientation and the effect of the peripheral visual field (nasal, temporal, superior, or inferior) because of the lack of available datasets that are compatible with the datasets included in this study.

Combined contrast sensitivity dataset

Our first goal is to create a comprehensive spatiotemporal chromatic contrast sensitivity model that accounts for the parameters of the stimuli (Gabors), which are the most relevant for AR/VR applications: spatial and temporal frequency, background luminance and chromaticity, the direction of contrast modulation in a color space, area, and eccentricity. As described in *"Measurements of contrast sensitivity"* above, gathering a single dataset that covers all these variables is impractical, so instead we opt to combine datasets from multiple sources.

Contrast sensitivity has been extensively studied over many decades, and almost every aspect of contrast detection has been measured. The main problem is that all those measurements were done independently, each measuring a different slice of a multi-dimensional space that defines the detection stimuli. Our goal is to select a set of measurements that covers all stimuli dimensions of interest, standardize the representation (e.g. use the same contrast definition and color space), and gather in a format that could be used to fit a single model that explains all the data.

Contrast units and color space

As the measure of sensitivity, we use the inverse of cone contrast, where cone contrast is expressed as:

$$C = \frac{1}{\sqrt{3}} \sqrt{\left(\frac{\Delta \mathcal{L}}{\mathcal{L}_0}\right)^2 + \left(\frac{\Delta \mathcal{M}}{\mathcal{M}_0}\right)^2 + \left(\frac{\Delta \mathcal{S}}{\mathcal{S}_0}\right)^2}, \tag{1}$$

where $\Delta \mathcal{L}$, $\Delta \mathcal{M}$, and $\Delta \mathcal{S}$ are the differential cone responses for the stimuli and $\mathcal{L}_0$, $\mathcal{M}_0$, and $\mathcal{S}_0$ are the cone responses for the corresponding background. $\mathcal{L}$, $\mathcal{M}$, and $\mathcal{S}$ cone responses are given by

$$\mathcal{L} = 0.689903 \int_\lambda l_2(\lambda) \, E(\lambda) \, d\lambda, \tag{2}$$

$$\mathcal{M} = 0.348322 \int_\lambda m_2(\lambda) \, E(\lambda) \, d\lambda, \tag{3}$$

$$\mathcal{S} = 0.0371597 \int_\lambda s_2(\lambda) \, E(\lambda) \, d\lambda, \tag{4}$$

where l_2, m_2 and s_2 are 2° CIE 2006 cone fundamentals (CIE, 2006)[1] and E is the measured spectral radiance emitted from the display.

Most achromatic contrast sensitivity studies report background luminance (or retinal illuminance) but do not report the spectral composition or chromaticity of the background. For those, we assumed $\mathcal{L}$, $\mathcal{M}$, and $\mathcal{S}$ cone responses that corresponded to the D65 illuminant and the modulation along the first ($\mathcal{L}+\mathcal{M}$) dimension in the DKL color space (Derrington, Krauskopf, & Lennie, 1984). For chromatic datasets, we either received the spectral emission characteristic from the authors or assumed typical emission spectra of the display used in the study (e.g. emission spectrum of a CRT monitor).

A common practice when measuring chromatic contrast sensitivity is to isolate the chromatic mechanisms of each individual using a heterochromatic flicker paradigm (Wagner & Boynton, 1972). Since our goal is to create a model of an average observer, which is suitable for general applications, we excluded the datasets that measured different color modulation directions for each observer.

Datasets

We selected 18 datasets to train and test our model. Table 2 lists 9 datasets with only achromatic stimuli and the achromatic portions of 4 datasets that contained both achromatic and chromatic stimuli. Table 3 lists 5 chromatic datasets, along with the chromatic portions of the 4 mixed datasets. Notably, Table 2 contains all the datasets used in recent work by Mantiuk et al., 2022,

---

[1]Tabulated cone fundamentals can be found at `http://cvrl.ucl.ac.uk/`

Table 2: The achromatic datasets used in our study. Asterisks (∗) mark cases where the data was obtained directly from the authors. In the remaining cases, the data was scanned from plots in the relevant publication. All listed datasets measured contrast sensitivity using contrast detection tasks. For datasets with more than one observer where individual data was provided, the expected inter-observer variability (standard deviation) within each condition in dB is shown along with the number of observers.

Dataset Name	Spatial freq. cpd	Temporal freq. Hz	Luminance cd/m^2	Eccentricity deg (visual field)	Area deg^2	Stimulus	Inter-observer Variability dB
Robson, 1966	0.5 - 30	0.5 - 32	20	0	6.25	Grating with rectangular aperture	✗
Virsu & Rovamo, 1979	0.5 - 16	✗	10	0, 5, 10, 15, 20, 25, 30 (nasal)	19.65	Grating with circular aperture	✗
Virsu, Rovamo, & Laurinen, 1982	1 - 22.6	1, 18	10	0, 1.5, 4, 7.5, 14, 30 (nasal)	1.57	Grating with semi-circular aperture	✗
Wright & Johnston, 1983	0.25, 2, 6, 9	0, 0.25, 8, 16	100	0 - 12 (superior)	0.75, 2.35, 36	Grating with rectangular aperture	0.4748 (N = 2)
Rovamo et al., 1993	0.125, 0.25, 0.5, 1, 2, 4, 8, 16, 32	✗	50	0	0.003 - 980	Grating with rectangular aperture	✗
Snowden, Hess, & Waugh, 1995	0.25, 1, 2, 4, 5, 10, 20	0.8 - 55.7	0.02 - 870	0	0.25, 1, 4.01	Gabor patch	✗
Modelfest (Watson, 2000) (∗)	1.12 - 30	✗	30	0	0.003 - 0.78	Gabor patch	3.7464 (N = 16)
Colorfest (Wuerger, Watson, & Ahumada Jr, 2002) (∗)	1.2 - 30	✗	40	0	0.78	Gabor patch	2.6553 (N = 3)
Laird et al., 2006	4, 8, 16	9.2 - 31.4	60	0	4.75	Gabor patch	✗
Hansen, Pracejus, & Gegenfurtner, 2009	✗	✗	100	10, 20, 30, 40, 50	50.27	Discs	✗
HDR-VDP CSF (Mantiuk et al., 2011; K. J. Kim, Mantiuk, & Lee, 2013) (∗)	0.125 - 32	✗	0.002, 0.02, 0.2, 2, 20, 150	0	0.07, 0.78, 7.06	Fixed cycles Gabor patch	2.7233 (N = 10)
HDR CSF (Wuerger et al., 2020) (∗)	0.5, 1, 2, 4, 6, 12, 24	✗	0.02, 0.2, 2, 20, 200, 2000, 7000, 10000	0	0.005 - 12.5	Fixed cycles Gabor patch	2.7873 (N = 23)
HDR disc CSF (Ashraf, Mantiuk, & Chapiro, 2023)	✗	✗	0.02, 0.2, 2, 20, 200	0	0.02, 0.2, 3.14	Discs	2.4710 (N = 12)

along with the recently collected HDR disc CSF dataset by Ashraf et al., 2023, but excluding the data from Anderson et al., 1991. We excluded that dataset as it used a different detection criterion than all other datasets (discrimination rather than detection) and its measurements were not comparable. We also avoided the datasets that were collected for an artificial pupil, as we are interested in visual performance for natural viewing. If multiple datasets in the literature covered a similar range of contrast sensitivity parameters, we selected those that were collected for a larger and more representative group of observers.

Aperture

The type of stimulus used for each dataset is listed in the last column of each table (Tables 2 and 3). Most datasets measured contrast thresholds using Gabor patches. Conversely, some datasets used sine gratings with rectangular or circular apertures

instead of smooth Gaussian windows. To standardize these modalities, given a circular aperture with a diameter d, we assumed the stimulus to be equivalent to a Gabor patch with $\sigma = d/2$. For rectangular apertures with the given area a, we assume the stimuli to be equivalent to a disc aperture with the same area ($\pi\sigma^2 = a$). Although those relationships may not fully explain the differences between Gaussian and other apertures, any inaccuracies are compensated for when fitting and allowing for a sensitivity shift in each dataset (see Eq. (32)).

Datasets that were directly obtained from the authors are marked with asterisks. The remaining datasets were scanned from the corresponding papers using the WebPlotDigitizer tool (Rohatgi, 2022). To ensure this scanning method had sufficiently high accuracy and repeatability we performed an experiment where a plot for which the ground truth data was available (achromatic CSF at 200 cd/m^2 from Figure 5 in Wuerger et al. (2020)) was scanned 5

Table 3: The chromatic datasets used in our study. Asterisks (∗) mark cases where the data was obtained directly from the authors. In the remaining cases, the data was scanned from plots in the relevant publication. All listed datasets measured contrast sensitivity using contrast detection tasks. For datasets with more than one observer where individual data was provided, the mean inter-observer variability in dB (standard deviation within each condition) is shown along with the number of observers.

Dataset Name	Chromatic modulation	Spatial freq. cpd	Temporal freq. Hz	Luminance cd/m^2	Eccentricity deg (visual field)	Area deg^2	Stimulus	Inter-observer variability dB
Van der Horst & Bouman, 1969	Yellow-blue and Red-Bluish Green over grey (equal-energy point) background, Red-green over yellow background	0.7 -18	0, 0.5 - 9.5	0.009 - 70	0	14.84	Grating with rect-angular aperture	✗
Colorfest (Wuerger et al., 2002) (∗)	Green-Red, Yellow Green-Violet, Greenish-Pink, Yellow-Blue, Dark Green-Light Pink, Dark Yellow-Light Blue over D65 background	1.2 - 30	✗	40	0	0.78	Gabor patch	4.1975 ($N = 3$)
Hansen et al., 2009	Pinkish Red and Violet over grey background	✗	✗	100	10, 20, 30, 40, 50	50.27	Discs	✗
K. J. Kim et al., 2013	Green-Red, Yellow Green-Violet, Dark Green-Light Pink, Dark Yellow-Light Blue over D65 background	0.25, 0.5, 1, 2, 4	✗	0.02, 0.2, 2, 20, 150	0	7, 28	Gabor patch	3.1301 ($N = 7$)
Lucassen, Lambooij, Sekulovski, & Vogels, 2018	4 color directions modulated in u'v' color space 45° apart from each other over 3 different backgrounds (grey, light yellow and dark yellow)	0.15, 0.3, 0.15, 1.5, 3, 5	✗	108	0	280	Gabor patch	✗
Kong, Pérez, Vogels, Sekulovski, & Heynderickx, 2018 (∗)	0°, 45°, 90°, 135° in u'v' color space over 9 different backgrounds	✗	2, 4, 8, 10, 15, 20, 25	35	0	78.5	Discs	6.6586 ($N = 3$)
HDR CSF (Wuerger et al., 2020) (∗)	Red-Green and Yellow-Violet over D65 background	0.125, 0.25, 0.5, 1, 2, 4, 6, 12, 24	✗	0.02, 0.2, 2, 20, 200, 2000, 7000, 10000	0	0.005 - 50	Fixed cycles Gabor patch	3.0668 ($N = 23$)
HDR disc CSF (Ashraf et al., 2023) (∗)	Pinkish Red and Violet over D65 background	✗	✗	0.02, 0.2, 2, 20, 200	0	0.02, 0.2, 3.14	Discs	4.2842 ($N = 10$)
Five Centers (Xu, Luo, & Sekulovski, 2020) (∗)	6 color directions over red, white, cyan, blue and yellow backgrounds	0.06, 0.12, 0.24, 0.48, 0.96, 1.92, 3.84	✗	8.8, 14.1, 24, 50, 72	0	272	Gabor patches	✗

times. The mean standard deviation of the log sensitivity values was found to be very low (0.00534 dB). The RMSE error between the mean of the 5 scans and the ground truth data was also modest at 0.0117 dB, or 0.135% of the original data. This leads us to conclude that the scanning tool offers a robust method for gathering data.

Contrast Sensitivity Function

In this section, we explain how the proposed castleCSF model can predict the detection threshold for any background color, specified in cone response units ($\mathcal{L}_0$, $\mathcal{M}_0$, $\mathcal{S}_0$), and for any color modulation direction, specified as increments of cone responses ($\Delta\mathcal{L}$, $\Delta\mathcal{M}$, $\Delta\mathcal{S}$). A high-level diagram and description of the model's workflow is shown in Figure 1.

To isolate the three color mechanisms, one achromatic and two chromatic, we transform the increments to the color-opponent DKL space (Derrington et al., 1984) assuming a D65 grey background using the transformation matrix:

$$\begin{bmatrix} \Delta D_{\text{Ach}} \\ \Delta D_{\text{RG}} \\ \Delta D_{\text{YV}} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -2.3112 & 0 \\ -1 & -1 & 50.9875 \end{bmatrix} \cdot \begin{bmatrix} \Delta\mathcal{L} \\ \Delta\mathcal{M} \\ \Delta\mathcal{S} \end{bmatrix} \cdot \quad (5)$$

To compute contrast, we divide the color-opponent increments by the background luminance:

$$\Delta C_{\text{Ach}} = \frac{\Delta D_{\text{Ach}}}{Y}, \quad \Delta C_{\text{RG}} = \frac{\Delta D_{\text{RG}}}{Y}, \quad \Delta C_{\text{YV}} = \frac{\Delta D_{\text{YV}}}{Y}, \quad (6)$$

where $Y = \mathcal{L}_0 + \mathcal{M}_0$. The contrast values of the three mechanisms

are weighted by the sensitivity functions S_{Ach}, S_{RG} and S_{YV}, and pooled together into a contrast energy:

$$E = \sqrt{\sum_{c \in \{\text{Ach,RG,YV}\}} (S_c(\rho, \omega, Y, a, e) \, \Delta C_c)^2} \,. \tag{7}$$

$S_c(\cdot)$ are the sensitivity functions of each mechanism, which we describe in the next section.

We assume that the contrast is detected when the contrast energy $E = 1$. Such a threshold can be found analytically by introducing a constant t:

$$\Delta C_{\text{Ach}} = t \, \Delta \widehat{C}_{\text{Ach}} \,, \quad \Delta C_{\text{RG}} = t \, \Delta \widehat{C}_{\text{RG}} \,, \quad \Delta C_{\text{YV}} = t \, \Delta \widehat{C}_{\text{YV}} \,, \tag{8}$$

where $(\Delta \widehat{C}_{\text{Ach}}, \Delta \widehat{C}_{\text{RG}}, \Delta \widehat{C}_{\text{YV}})$ is the contrast at the detection threshold. If we introduce these expressions to Eq. (7), we get:

$$E = t \sqrt{\sum_{c \in \{\text{Ach,RG,YV}\}} (S_c(\rho, \omega, Y, a, e) \, \Delta \widehat{C}_c)^2} = t \, \widehat{E} \,, \tag{9}$$

where $\widehat{E}$ is the contrast energy at the threshold. Since we assumed that the contrast energy should be 1 at the detection threshold, we have $\widehat{E} = 1$ and hence $t = E$. This shows us that the detection
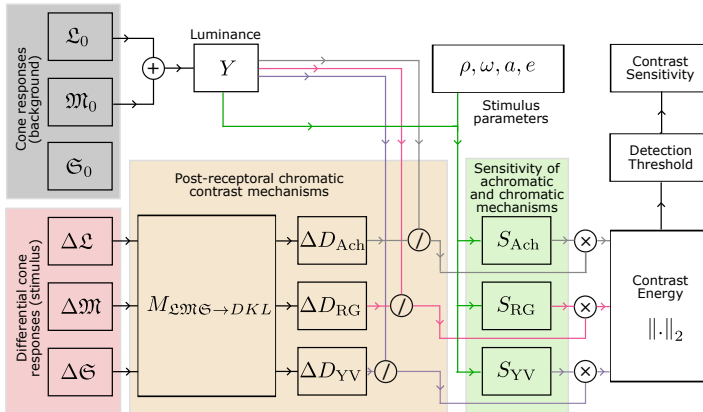
Figure 1: Contrast encoding and the main processing stages of castleCSF. The color of the background is represented as cone responses ($\mathcal{L}_0$, $\mathcal{M}_0$, $\mathcal{S}_0$) and the direction in the color space as the increments of cone responses ($\Delta \mathcal{L}$, $\Delta \mathcal{M}$, $\Delta \mathcal{S}$). The direction is transformed into the DKL color space, and the contrast is computed by dividing by luminance, where luminance is the sum of $\mathcal{L}_0$ and $\mathcal{M}_0$. The contrast for the three cardinal color directions is then modulated by the sensitivity of each mechanism (S_{Ach}, S_{RG} and S_{YV}) and pooled to obtain the contrast energy. We use this energy to find the detection threshold, and finally compute the sensitivity value.

threshold expressed as an increment in the color opponent space can be computed from Eq. (8) as:

$$\Delta \widehat{C}_{\text{Ach}} = \frac{\Delta C_{\text{Ach}}}{E} \,, \quad \Delta \widehat{C}_{\text{RG}} = \frac{\Delta C_{\text{RG}}}{E} \,, \quad \Delta \widehat{C}_{\text{YV}} = \frac{\Delta C_{\text{YV}}}{E} \,. \tag{10}$$

Assuming a local linearity of the transformation around the background color ($\mathcal{L}_0$, $\mathcal{M}_0$, $\mathcal{S}_0$), the same approximately holds for the increment thresholds in the cone response space:

$$\Delta \widehat{\mathcal{L}} \approx \frac{\Delta \mathcal{L}}{E} \,, \quad \Delta \widehat{\mathcal{M}} \approx \frac{\Delta \mathcal{M}}{E} \,, \quad \Delta \widehat{\mathcal{S}} \approx \frac{\Delta \mathcal{S}}{E} \,. \tag{11}$$

Once we know the cone response increments at the threshold, we can calculate the sensitivity as the inverse of the cone contrast from Eq. (1):

$$S = \sqrt{\frac{3}{\left(\frac{\Delta \mathcal{L}}{\mathcal{L}_0 E}\right)^2 + \left(\frac{\Delta \mathcal{M}}{\mathcal{M}_0 E}\right)^2 + \left(\frac{\Delta \mathcal{S}}{\mathcal{S}_0 E}\right)^2}} \,, \tag{12}$$

where E is the energy from Eq. (7).

In the following sections, we explain how the sensitivities of the three mechanisms, $S_{\text{Ach}}(\cdot)$, $S_{\text{RG}}(\cdot)$ and $S_{\text{YV}}(\cdot)$, are modeled.

Mechanism sensitivity

We model the sensitivity of each mechanism, $S_{\text{Ach}}(\cdot)$, $S_{\text{RG}}(\cdot)$ and $S_{\text{YV}}(\cdot)$, by considering the effects of spatial frequency, temporal frequency, luminance, area, and eccentricity. The following sections explain how each factor is modeled.

Temporal frequency

A series of early work by de Lange introduced the concept of distinct temporal channels in the human visual system by reporting varying integration time constants for achromatic and chromatic stimuli at different flicker rates (de Lange, 1958a, 1958b). The site of these mechanisms (retinal, LGN, cortical, etc.), their inputs (from parvo or magnocellular pathways), and the number of these temporal channels is still the subject of an ongoing debate. Results from psychophysical studies have suggested the presence of multiple temporally-tuned channels (Hess & Snowden, 1992; Kelly, 1983; Metha & Mullen, 1996; Robson, 1966). For near-threshold achromatic stimuli, three temporal channels have been proposed with two motion and one flicker detection channel (King-Smith & Kulikowski, 1975; Mandler & Makous, 1984). We opted for the two-channel (low-pass/sustained and band-pass/transient) model supported by studies such as Tolhurst, 1973 and Anderson & Burr, 1985 for the achromatic mechanism because of its relative

simplicity, and because the potential third high temporal frequency channel likely only detects low spatial frequencies (Hess & Snowden, 1992). Some temporal CSF data suggests that the peak of the achromatic transient channel is dependent on luminance (de Lange, 1958a; Snowden et al., 1995), and shifts towards lower temporal frequencies as the luminance level decreases. For chromatic mechanisms, there is evidence of one slow sustained chromatic channel and a faster transient channel (Gegenfurtner & Hawken, 1996; Cropper & Wuerger, 2005; Cass et al., 2009; Metha & Mullen, 1996). We chose not to include the chromatic transient channels in our model as their magnitudes are relatively smaller than the chromatic sustained channels (Cass et al., 2009), and thus contribute little to the overall sensitivity response. It has also been suggested that the higher (temporal) frequency chromatic stimuli might be mediated by luminance mechanisms instead of a dedicated chromatic transient channel (Dobkins, Gunther, & Peterzell, 2000).

The impulse responses of the temporal filters associated with the sustained ($\mathcal{S}$) and transient ($\mathcal{T}$) channels can be well approximated by the generalized exponential functions (Mantiuk et al., 2022):

$$R_{\mathcal{S}}^c(\omega) = \exp\left(-\frac{\omega^{\beta_{\mathcal{S}}^c}}{\sigma_{\mathcal{S}}^c}\right), \quad \forall c \in \{\text{Ach}, \text{RG}, \text{YV}\}, \tag{13}$$

and:

$$R_{\mathcal{T}}^{\text{Ach}}(\omega, Y) = \exp\left(-\frac{\left|\omega^{\beta_{\mathcal{T}}^{\text{Ach}}} - (\omega_0(Y))^{\beta_{\mathcal{T}}^{\text{Ach}}}\right|^2}{\sigma_{\mathcal{T}}^{\text{Ach}}}\right), \tag{14}$$

where ω is the temporal frequency in Hz, and $\beta_{\mathcal{S}}^c$, $\sigma_{\mathcal{S}}^c$, $\beta_{\mathcal{T}}^{\text{Ach}}$, and $\sigma_{\mathcal{T}}^{\text{Ach}}$ are the parameters of the model. ω_0 is the peak temporal frequency of the transient channel which was fixed at 5 Hz in stelaCSF (Mantiuk et al., 2022). In castleCSF, we model this as a parameter dependent on the mean luminance level based on data from de Lange, 1958a and Snowden et al., 1995. We found that the following linear relationship between the transient channel peak frequency and log luminance level explained the data well:

$$\omega_0(Y) = m_{\omega}^{\text{Ach}} \log_{10} Y + c_{\omega}^{\text{Ach}} \quad [\text{Hz}], \tag{15}$$

where m_{ω}^{Ach} and c_{ω}^{Ach} are the fitted parameters. This assumption that at higher luminance levels, the visual system is more attuned to higher temporal frequencies is also consistent with the information flow theory of early vision proposed by Van Hateren, 1993 where they showed that the signal-to-noise ratio (SNR) increases

as a function of log light intensity and that the peak of the neural filter (analogous to the transient channel in our model) shifts towards higher frequencies with increasing SNR. Ferry-Porter's law, which states that the critical flicker fusion frequency (CFF, the temporal frequency at which flicker can no longer be perceived) increases linearly with the logarithm of the luminance level (Ferry, 1892; Porter, 1902), also supports our model because increasing peak temporal frequency would also result in a higher cut-off frequency or CFF. The original works by Ferry and Porter do not suggest any saturation point for luminance level beyond which the CFF would no longer increase but the series of studies by Hecht et al. (Hecht & Verrijp, 1933; Hecht & Shlaer, 1936; Hecht & Smith, 1936) and more recently by Chapiro, Matsuda, Ashraf, & Mantiuk, 2023 have shown that CFF reaches a saturation point depending on the size and eccentricity of the stimulus. This would imply a sigmoidal instead of linear relationship in Eq. (15). However, given limited data, we opted for a simpler model.

The responses of the temporal channels in our model are shown in Figure 2. All the sustained channels are low-pass while the achromatic transient channel is band-pass with the peak shifting towards higher temporal frequencies as the luminance increases.
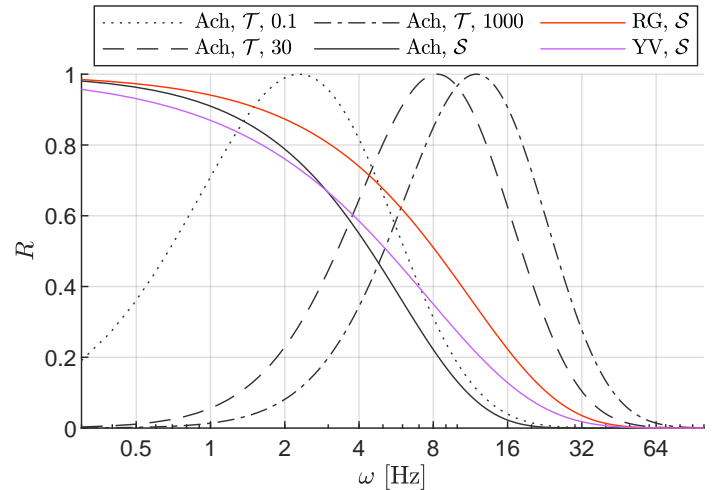

Figure 2: Responses of the temporal channels. The three sustained ($\mathcal{S}$) channels are low-pass while the achromatic transient ($\mathcal{T}$) channel is bandpass. The response of this bandpass channel depends on luminance — shifts towards higher frequencies as the mean luminance increases.

Spatial frequency

The spatial frequency response of the achromatic mechanism (S_{Ach} in Figure 1) is modeled to be band-pass. The assumption that the spatial frequency response for achromatic static gratings is

band-pass is well-supported in visual perception literature. Studies such as those by Blakemore & Campbell, 1969 demonstrate that human visual sensitivity is highest at intermediate spatial frequencies and falls off at both low and high spatial frequencies. The magnitude and the location of the peak sensitivity vary depending on the other properties (luminance, temporal frequency, size, etc.) of the stimulus. At the low spatial frequency end, the sensitivity fall-off is caused by lateral inhibition, but this inhibitory response is not strong when the low spatial frequency stimulus is temporally-modulated (Tolhurst, 1973; Donner & Hemilä, 1996) and results in a low-pass response. We model the spatial frequency responses of both the sustained and temporal achromatic channels as log-parabolas (Ahumada & Peterson, 1992):

$$l_{\mathfrak{t}}^{c'}(\rho, Y) = 10^{-\dfrac{(\log_{10}\rho - \log_{10}\rho_{m,\mathfrak{t}}^{c}(Y))^2}{2^{k_{b,\mathfrak{t}}^{c}}}}, \tag{16}$$

where $c \in \{\mathrm{Ach}, \mathrm{RG}, \mathrm{YV}\}$ is the index of the mechanism, $\mathfrak{t} \in \{\mathcal{S}, \mathcal{T}\}$ is the temporal channel (sustained or transient), $\rho_{m,\mathfrak{t}}^{c}$ controls the position of the peak of the parabola as a function of luminance (described in section *Luminance*), and $k_{b,\mathfrak{t}}^{c}$ controls the bandwidth of the log-parabola. For achromatic spatial contrast sensitivity, the parabola is truncated on the low spatial frequency side of the envelope:

$$l_{\mathfrak{t}}^{\mathrm{Ach}}(\rho) = \begin{cases} 1 - k_{a,\mathfrak{t}}^{\mathrm{Ach}} & \text{if } \rho < \rho_{m,\mathfrak{t}}^{\mathrm{Ach}} \text{ and } l_{\mathfrak{t}}^{\mathrm{Ach}'} < 1 - k_{a,\mathfrak{t}}^{c} \\ l_{\mathfrak{t}}^{\mathrm{Ach}'}(\rho) & \text{otherwise} \end{cases} \tag{17}$$

where, $\mathfrak{t} \in \{\mathcal{S}, \mathcal{T}\}$, and $k_{a,\mathfrak{t}}^{\mathrm{Ach}}$ is a function parameter that controls the drop in achromatic spatial contrast sensitivity at low frequencies. Our model fits coincide with the findings from the literature that the sustained channel has a band-pass shape with the peak shifting with different stimuli properties. The fitted value of the peak spatial frequency parameter for the transient channel ($k_{\rho,\mathcal{T}}^{\mathrm{Ach}}$ in Table 5) has a very small value, which results in a low-pass shape.

The spatial frequency responses of the two chromatic mechanisms are approximately low-pass, that is, the sensitivity decreases with increasing spatial frequency. This is because of the much weaker low-spatial frequency inhibition in the chromatic detection pathways (Kelly, 1983; Metha & Mullen, 1996). The roll-off and cut-off frequency of these responses depends on the other stimulus properties. The low-pass response of the spatio-chromatic CSF can also be modeled as a truncated log-parabola $l_{\mathfrak{t}}^{c'}$ with the sensitivity for lower spatial frequencies leveling off (instead of decreasing

like the achromatic CSF) at the peak frequency. This function is represented as:

$$l_{\mathcal{S}}^{c}(\rho) = \begin{cases} 1 & \text{if } \rho < \rho_{m,\mathcal{S}}^{c}(Y) \\ l_{\mathcal{S}}^{c'}(\rho) & \text{otherwise} \end{cases} \tag{18}$$

where $c \in \{\mathrm{RG}, \mathrm{YV}\}$. The definition of the parameters is the same as that for the achromatic channel.

Area

The stimuli become easier to detect (the sensitivity is increased) as their size increases. The contrast sensitivity increases with stimulus size for achromatic and chromatic CSFs up to a certain value, known as the critical area a_x. The value of this critical area and the rate of increase of sensitivity up until the critical area depends on the spatial frequency of the stimuli and is modeled in Rovamo et al., 1993 as

$$a_{x,\mathfrak{t}}^{c} = \frac{a_{0,\mathfrak{t}}^{c}}{1 + (\rho/\rho_{0,\mathfrak{t}}^{c})^2} \quad [\deg^2], \tag{19}$$

where $c \in \{\mathrm{Ach}, \mathrm{RG}, \mathrm{YV}\}$ and $\mathfrak{t} \in \{\mathcal{S}, \mathcal{T}\}$. $a_{x,\mathfrak{t}}^{c}$ is the maximum value of the critical area and $\rho_{0,\mathfrak{t}}^{c}$ is the minimum value of spatial frequency, beyond which the value of the critical area is no longer constant and begins to decrease with increasing spatial frequency, for each colour and temporal channel. Figure 3a shows the change in the critical area for the four modeled channels. The plot shows that the detection mechanism integrates over a smaller area at high frequencies. The product of a function of this critical area $a_{x,\mathfrak{t}}^{c}$ and spatial frequency is the critical number of cycles over which our visual system can integrate. Figure 3b shows the change in the critical number of cycles as a function of the spatial frequency.

The size-dependent sensitivity response follows the model of Rovamo et al., 1993 and is expressed as:

$$S_{\mathrm{area},\mathfrak{t}}^{c}(\rho, a) = \rho \sqrt{\frac{a_{x,\mathfrak{t}}^{c}}{1 + \frac{a_{x,\mathfrak{t}}^{c}}{a}}} \quad [\mathrm{cyc}]. \tag{20}$$

The effect of size is independent of other factors in our model. The combined effect of spatial frequency and stimulus size on the contrast sensitivity is shown in Figure 4. The log-parabola band-pass shape of the achromatic sustained channel is more pronounced for the stimuli of constant size (rather than the constant number of cycles) where at low spatial frequencies fewer cycles are presented.

Luminance

In dim light, contrast sensitivity increases in proportion to the square root of retinal illuminance, according to the DeVries-Rose law, but in bright light contrast sensitivity follows Weber's law and is independent of illuminance (Blackwell, 1946; Rovamo, Mustonen, & Näsänen, 1995). K. J. Kim et al., 2013 have shown similar results for chromatic stimuli, with sensitivity saturating from approximately 50 cd/m². The recent sensitivity data from (Wuerger et al., 2020), measured up to 10 000 cd/m², shows that the assumption of sensitivity becoming constant at high luminances is not entirely accurate for achromatic stimuli as the sensitivity at very
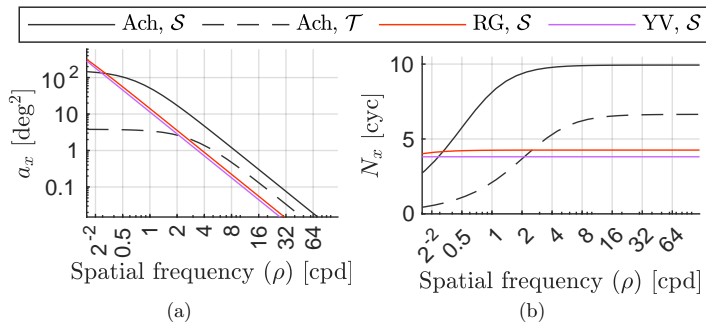
Figure 3: (a) Change in the critical area with spatial frequency. The critical area parameter for the achromatic sustained channel saturates at about 0.5 cpd and for the achromatic transient channel saturates at about 2 cpd. For the red-green and yellow-violet channels, this saturation point lies at a much smaller spatial frequency (b) Critical number of cycles changing with the spatial frequency. The value of the critical number of cycles becomes constant with increasing spatial frequency. For the red-green and yellow-violet channels, our model predicts the critical number of cycles to be constant for approximately the whole range of spatial frequencies detectable by the human visual system.
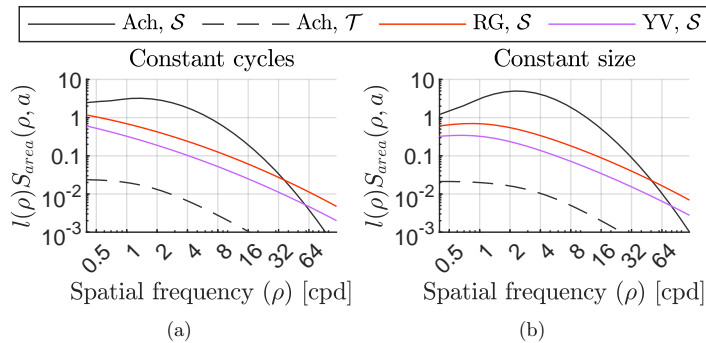
Figure 4: Joint effect of spatial frequency and area on sensitivity (a) Contrast sensitivity of stimuli with 4 visible cycles at each spatial frequency (b) Contrast sensitivity of stimuli with a constant radius of 2° at each spatial frequency.

high luminance, above 1000 cd/m², starts to drop, especially for low frequencies. Following Mantiuk et al., 2022, we model the change of the luminance-dependent sensitivity for achromatic sustained mechanism as an initial increase followed by a plateau and eventual decrease:

$$S_{m,\mathcal{S}}^{\text{Ach}}(Y) = k_{s1,\mathcal{S}}^{\text{Ach}} \left(1 + \frac{k_{s2,\mathcal{S}}^{\text{Ach}}}{Y} \right)^{-k_{s3,\mathcal{S}}^{\text{Ach}}} \left(1 - \frac{1 + k_{s4,\mathcal{S}}^{\text{Ach}}}{Y} \right)^{-k_{s5,\mathcal{S}}^{\text{Ach}}}.$$

(21)

The luminance-dependence of chromatic sensitivity is modeled as an increase and then saturation of sensitivity with increasing luminance:

$$S_{m,\mathcal{S}}^{c}(Y) = k_{s1,\mathcal{S}}^{c} \left(1 + \frac{k_{s2,\mathcal{S}}^{c}}{Y} \right)^{k_{s3,\mathcal{S}}^{c}}, \quad \forall c \in \{\text{RG}, \text{YV}\}. \quad (22)$$

For the achromatic transient channel, we again followed the approach from Mantiuk et al., 2022 and modeled the change in sensitivity as a linear function of luminance:

$$S_{m,\mathcal{T}}^{\text{Ach}}(Y) = k_{s2,\mathcal{T}}^{\text{Ach}} Y^{k_{s1,\mathcal{T}}^{\text{Ach}}}. \quad (23)$$

This relationship is supported by the studies from Swanson, Ueno, Smith, & Pokorny, 1987 and Snowden et al., 1995, showing the linear increase of sensitivity to flickering stimuli with luminance. For the luminance ranges tested in the aforementioned studies, the sensitivity did not saturate with luminance as was the case for static achromatic and chromatic stimuli.

Luminance also causes the shift of the peak sensitivity — as the luminance is reduced, the sensitivity drops but this reduction is stronger for high frequencies. These effects have different characteristics for the sustained and transient achromatic, and two chromatic mechanisms. The effect on the luminance-dependent frequency shift in the spatial log-parabola response (Eq. (16)) is modeled as:

$$\rho_{m,\mathcal{S}}^{\text{Ach}}(Y) = k_{\rho1,\mathcal{S}}^{\text{Ach}} \left(1 + \frac{k_{\rho2,\mathcal{S}}^{\text{Ach}}}{Y} \right)^{-k_{\rho3,\mathcal{S}}^{\text{Ach}}} \quad [\text{cpd}],$$

$$\rho_{m,\mathcal{T}}^{\text{Ach}}(Y) = k_{\rho,\mathcal{T}}^{\text{Ach}} \quad [\text{cpd}], \quad (24)$$

$$\rho_{m,\mathcal{S}}^{c}(Y) = k_{\rho,\mathcal{S}}^{c} \quad [\text{cpd}], \quad \forall c \in \{\text{RG}, \text{YV}\},$$

where Y is luminance in cd/m², $k_{...}^{c}$ are the parameters of the model, $S_{m,\mathcal{S}/\mathcal{T}}^{c}$ is the luminance-dependent sensitivity, and $\rho_{m,\mathcal{S}/\mathcal{T}}^{c}$ is responsible for the spatial frequency shift in Eqs. (17) – (18).

We model the effect of luminance rather than retinal illuminance (in Trolands) because the former is more readily available in engineering applications, in which the pupil size (required for calculating retinal illuminance) is often unknown or difficult to measure. Figure 5 show the values of the luminance-dependent parameters in Eqs. (21) – (24). These parameters only show the effect of luminance and do not include the shift in the peak sensitivity and peak spatial frequency induced by the stimulus size. The achromatic sustained response in Figure 5a shows the decline in sensitivity for luminances above $1000 \, \text{cd/m}^2$. The sensitivity of the achromatic transient channel increases with increasing luminance while both the chromatic channels transition to Weber's region above $20 \, \text{cd/m}^2$. The log-parabola spatial frequency parameter for the achromatic sustained channel increases with luminance up to about $500 \, \text{cd/m}^2$ as shown in Figure 5b and then remains constant with further increase in luminance. An increase in this parameter translates to the shifting of the spatial contrast sensitivity curve towards higher spatial frequencies. The peak spatial frequency parameters for the achromatic transient channel and both chromatic sustained channels are constants and thus independent of luminance.
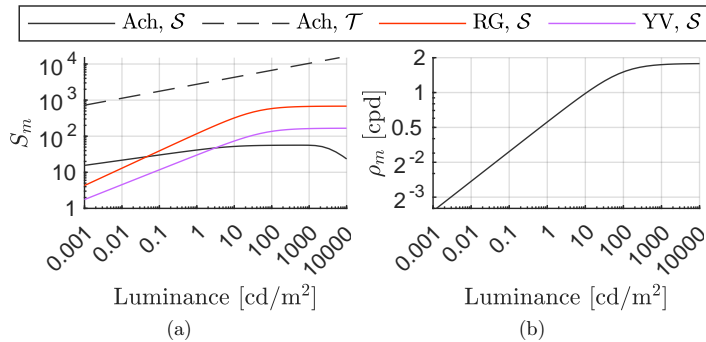
Figure 5: (a) Luminance-dependent shift in sensitivity. The responses of the three sustained channels increase with luminance and then saturate, with the achromatic sustained channel response showing a decline at very high luminances. The transient channel response shows a linear relationship with luminance in log-log space. (b) Luminance-dependent spatial frequency shift for the achromatic sustained channel. The value of the ρ_m parameter for the remaining channels is independent of luminance.

Eccentricity

Contrast sensitivity decreases with retinal eccentricity for achromatic (Robson & Graham, 1981) and chromatic CSFs (Anderson et al., 1991) along with other visual performance metrics such as visual acuity (Anstis, 1974), vernier acuity (Levi, Klein, & Ait-

sebaomo, 1985), crowding (Coates, Chin, & Chung, 2013), etc. The performance differences between the foveal and peripheral regions of the visual field can be explained by the differences in the neural configuration of photoreceptors in different regions of the retina. The receptive field size increases with eccentricity (Hubel & Wiesel, 1962) resulting in a functional increase in spatial summation with eccentricity (Wilson, 1970; Johnson, Keltner, & Balestrery, 1978). The rate of this change depends on the color and the spatial frequency of the stimuli. If all other properties of the stimuli are kept constant, the visual system's sensitivity to a low-frequency stimulus will decrease at a lower rate as compared to a high-frequency stimulus when the stimuli are moved from the fovea to the periphery across the retina (Virsu & Rovamo, 1979; Virsu et al., 1982). We have modeled the sensitivity drop with respect to retinal eccentricity as a log-linear function of spatial frequency and eccentricity, following (Watson, 2018):

$$S_{\text{ecc}}^c(e, \rho) = 10^{-(\hat{k}_{\text{e}1}^c \, \rho \, e + \hat{k}_{\text{e}2}^c \, e)} \, , \qquad (25)$$

where $c \in \{\text{Ach}, \text{RG}, \text{YV}\}$.

Finding a suitable peripheral chromatic contrast sensitivity dataset to validate our hypotheses proved to be challenging. The dataset from Hansen et al., 2009 primarily drove the optimization of the eccentricity-dependent parameters for the chromatic channel in our model. We did not include any datasets that used flicker photometry to isolate the chromatic mechanism as we could not parse their data in LMS cone contrasts (Noorlander, Koenderink, Den Olden, & Edens, 1983; Anderson et al., 1991; Newton & Eskew, 2003; Mullen et al., 2002; Mullen, Sakurai, & Chu, 2005). Moreover, Mullen et al., 2002 and Mullen et al., 2005 used very thin sinusoidal grating strips to avoid detection from neighboring peripheral receptive fields. The equivalent Gabor assumption (see Section *Aperture*) did not result in good fits for these datasets.

The decrease in sensitivity is non-uniform across the visual field, with a slower decrease in the nasal direction (Anderson et al., 1991). To model this effect, the $\hat{k}_{\text{e}1}^c$ and $\hat{k}_{\text{e}2}^c$ eccentricity response parameters from Eq. (25) are calculated as weighted means of sensitivity drop in the nasal ($k_{\text{e}i,\text{nasal}}^c$) and other directions ($k_{\text{e}i}^c$).

$$\hat{k}_{\text{e}i}^c = \alpha k_{\text{e}i}^c + (1 - \alpha) k_{\text{e}i,\text{nasal}}^c \quad \text{where} \quad i = 1, 2 \quad \text{and}$$
$$\alpha = \min \left\{ 1, \left| \frac{\theta - 180}{90} \right| \right\} \, . \qquad (26)$$

θ is the orientation in the visual field in deg. $\theta = 0$ corresponds to the temporal and $\theta = 180$ to the nasal directions; i.e. it is an

angular coordinate for the right eye. Figure 6 shows the reduction in sensitivity as stimulus position on the retina moves from the fovea to the periphery. This decrease in sensitivity is steeper for higher spatial frequencies which implies that our visual system is less sensitive to stimuli with finer details in the peripheral region. The slopes of eccentricity-dependent change in sensitivity also differ for the three different color channels. This is expected as the cone density function for different classes of photoreceptors (Curcio, Sloan, Kalina, & Hendrickson, 1990) as well as the sensitivity of the post-receptoral pathways for different opponent color channels (Mullen et al., 2002, 2005) change at different rates in the periphery.

Another possible approach to model the effect of sensitivity could be to include eccentricity as a parameter in the critical area function in Eq. (19), as there is evidence of the critical area value increasing as a function of retinal eccentricity (Wilson, 1970; Johnson et al., 1978). However, we opted to use the simpler model from (Watson, 2018) based on the current availability of datasets. We did not find any suitable datasets that comprehensively covered the joint effects of eccentricity, area, and temporal frequency. Consequently, our model employs a simpler structure and maintains separability of eccentricity-dependence to avoid overfitting data that does not capture the joint effects of stimuli properties.
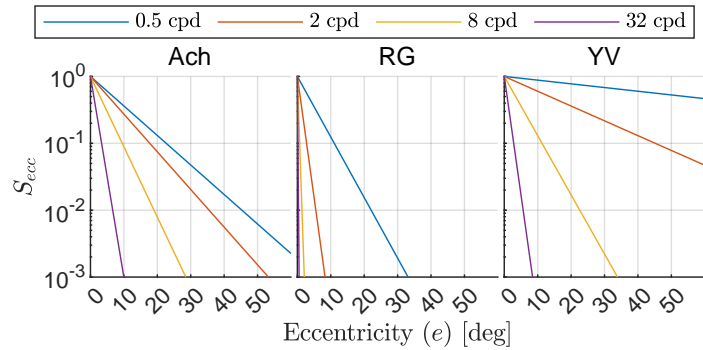
Figure 6: Eccentricity-dependent changes in sensitivity. The spatial frequency and the retinal orientation (axes: nasal, temporal, superior, inferior) jointly affect the rate at which the sensitivity decreases with retinal eccentricity in each color channel.

Combined model

The final sensitivity of our model is given by Eq. (12), which requires computing the contrast energy according to Eq. (9). The contrast energy equation relies on per-mechanism sensitivity functions, given as:

$$S_{\text{Ach}}(\rho, \omega, Y, a, e) = S_{\text{ecc}}^{\text{Ach}}(e, \rho) R_{\mathcal{S}}^{\text{Ach}}(\omega) S_{\text{sal},\mathcal{S}}^{\text{Ach}}(\rho, a, Y) +$$
$$S_{\text{ecc}}^{\text{Ach}}(e, \rho) R_{\mathcal{T}}^{\text{Ach}}(\omega) S_{\text{sal},\mathcal{T}}^{\text{Ach}}(\rho, a, Y) , \tag{27}$$

$$S_{\text{RG}}(\rho, \omega, Y, a, e) = S_{\text{ecc}}^{\text{RG}}(e, \rho) R_{\mathcal{S}}^{\text{RG}}(\omega) S_{\text{sal},\mathcal{S}}^{\text{RG}}(\rho, a, Y) , \tag{28}$$

$$S_{\text{YV}}(\rho, \omega, Y, a, e) = S_{\text{ecc}}^{\text{YV}}(e, \rho) R_{\mathcal{S}}^{\text{YV}}(\omega) S_{\text{sal},\mathcal{S}}^{\text{YV}}(\rho, a, Y) , \tag{29}$$

where $S_{\text{ecc}}^{c}(e, \rho)$ is given in Eq. (25), $R_{\mathcal{S}}^{c}(\omega)$ are the temporal filters from Eq. (13) and Eq. (14) and the combined effects of spatial frequency, stimulus area, and luminance is modelled as the product of individual sensitivities:

$$S_{\text{sal},\mathfrak{t}}^{c}(\rho, a, Y) = S_{\text{m},\mathfrak{t}}^{c}(Y) \, S_{\text{area}}^{c}(\rho, a) \, l_{\mathfrak{t}}^{c}(\rho) , \tag{30}$$

where $c \in \{\text{Ach}, \text{RG}, \text{YV}\}$, and $\mathfrak{t} \in \{\mathcal{S}, \mathcal{T}\}$. $S_{\text{m},\mathfrak{t}}^{c}$ is the luminance-dependent change in the peak sensitivity (Eqs. (21) – (23)), S_{area}^{c} is the function of critical area (Eq. (20)), and $l_{\mathfrak{t}}^{c}$ is the truncated log-parabola representing the contrast sensitivity function envelope across spatial frequencies (Eqs. (17) – (18)).

Figure 7 shows how all the individual components of our model presented in Eqs. (16) – (29) combine to predict S_{Ach}, S_{RG} and S_{YV} which feed into our contrast encoding model (Figure 1) to predict our visual system's sensitivity to stimuli modulated along any color direction.

Extension for edge stimuli

Since some of the stimuli were discs instead of Gabor patches, an extended version of our model was needed to predict them. Edge contrast sensitivity has been shown to be an indicator of the most sensitive contrast vision channel (Levi & Harwerth, 1982; Verbaken & Johnston, 1986). In other words, the peak of the contrast sensitivity envelope (across spatial frequencies) is proportional to the edge sensitivity of the visual system. Since a disc forms a circular edge, we can combine the peak-sensitivity assumption with a multiple-detectors model from (Ashraf et al., 2023) to predict the disc sensitivity as:

$$S_{\text{disc}}^{c}(\omega, Y, a, e) = (2\pi \sqrt{a})^{1/\beta} \max_{\rho}(S_c(\rho, \omega, Y, a_{\text{disc}}, e)) , \tag{31}$$

where, $c \in \{\text{Ach}, \text{RG}, \text{YV}\}$, S_c is the contrast sensitivity of the equivalent Gabor patches from Eqs. (27) – (29), $\beta = 3.01142$, and $a_{\text{disc}} = 2.42437$. The values of these fixed parameters are taken from (Ashraf et al., 2023).

Model training

18 datasets, listed in Tables 2–3 were used to train our CSF model. For the datasets that contained measurements from multiple observers, the mean values over all observers were used for training. All data points with sensitivity values less than 1 were removed, as these values represent contrast thresholds greater than 1. This is only possible when asymmetric contrast modulation is employed. A per-dataset adjustment factor, s_d was applied to the sensitivity measurements from each dataset. This was necessary, as the absolute magnitude of the sensitivity depended on the measurement conditions, psychophysical procedure, specific stimuli used, etc. This is compensated via a vertical shift for the whole dataset in the log sensitivity scale, allowing for the integration of diverse datasets into a single robust and generalizable model. The loss function of the optimization procedure minimized the difference between the model predictions and the adjusted (via s_d factor) data points from all the datasets, as well as minimized the base 10 log-

arithmic value of s_d (so the multiplier is close to 1 and prevents overfitting to any specific dataset characteristics).

$$\mathcal{L} = \sum_d \sum_i \left(\log_{10} S_{i,d} - s_d \log_{10} \tilde{S}_{i,d} \right)^2 + \frac{\lambda}{D} \sum_d (\log_{10} s_d)^2 \tag{32}$$

where $d = 1, \ldots, D$ represents the datasets, $S_{i,d}$ and $\tilde{S}_{i,d}$ are the reference and predicted sensitivity values for the stimulus i in dataset d. We found a suitable value of the regularization parameter ($\lambda = 0.01$) by trial and error. In all our experiments, we fix $s_d = 1$ for the reference HDR CSF dataset. We fit all models using a quasi-Newton method implemented in Matlab's *fminunc* function. To avoid local minima and implausible model parameters, the optimization was initialized using the parameters from stelaCSF (Mantiuk et al., 2022) and the post-receptoral spatio-chromatic CSF (Mantiuk et al., 2020). The parameters resulting from fitting castleCSF to all available datasets are reported in the Appendix Tables 5–7.

Comparison with other CSF models

We compare the predictions of castleCSF with several popular models from the literature, listed in Table 1. We excluded from this comparison the model of Rovamo et al., 1999 as it could not be tested on our datasets (it only works only for select isoluminant directions).

One challenge of evaluating CSF models is that contrast sensitivity data cannot be easily split into datasets used for testing and training. This is because each dataset typically contains uniformly spaced samples across a few select slices of the multi-dimensional space of stimulus parameters. If a model is trained on one dataset and tested on another, each containing different slices of the parameter space, the error measure is not representative of the entire space. Instead of doing this, we perform a 5-fold cross-validation within each dataset (leave-one-out, 5 splits) and use all datasets for both training and testing. The same split of test/train data was used for each compared CSF model. The error is reported as root-mean-square-error (RMSE), represented in dB units:

$$\mathcal{E} = 20 \sqrt{\frac{1}{N} \sum_d \sum_i \left(\log_{10} S_{i,d} - s_d \log_{10} \tilde{S}_{i,d} \right)^2} \quad [\text{dB}], \tag{33}$$

where N is the total number of data points. The errors in Table 4 are reported as means and standard deviations across all five folds. The number of trainable parameters of each model is also listed
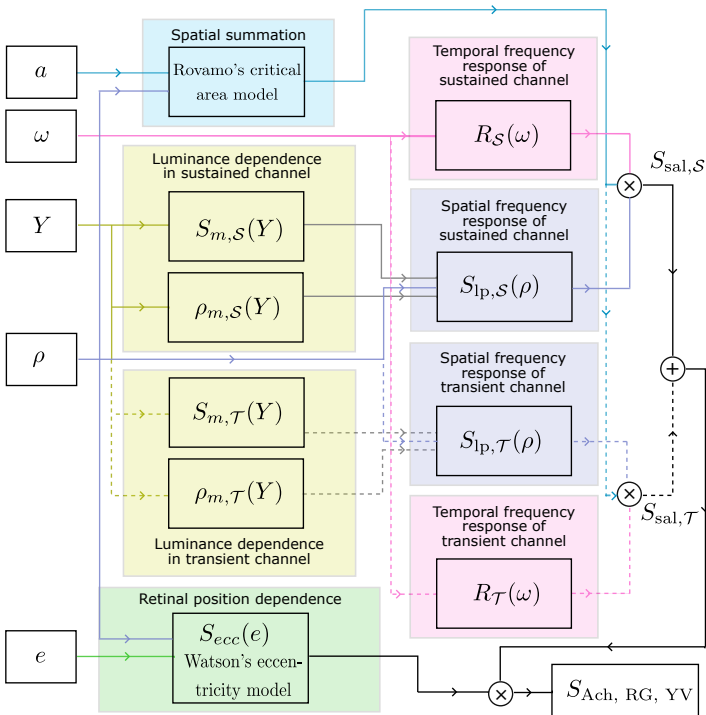

Figure 7: Spatiotemporal luminance, area, and eccentricity-dependent contrast sensitivity for individual opponent color mechanisms (S_{Ach}, S_{RG} and S_{YV}). The blocks represent the different components of the model described in Eqs. (16) – (29). The blocks where the inputs and outputs are dashed lines represent the transient channels and are only relevant for S_{Ach}.

Table 4: *Prediction errors of contrast sensitivity models, tested on the same subset of data points. The errors are reported in the dB units (see Eq. (33)) as the mean and standard deviation across five-folds (leave-one-our cross-validation). Each column shows the result for the subset of the dataset, containing variation along selected parameters of the stimuli (selected to match the parameters supported by CSFs). The value of N is the number of data points included in each subset (testing and training parts). The predictions are reported for 5-fold cross-validation (see text). The symbols denote: ρ — spatial frequency; Y — luminance; a — area; ω — temporal frequency; e — eccentricity, c — modulation along chromatic color directions., c_{DKL} — modulation along the cardinal color direction (achromatic, red-green, yellow-violet) only. Asterisks ($*$) mark the cases where low spatial and temporal frequencies are removed from the datasets.*

Model	No. of model parameters	Achromatic only			Achromatic and chromatic					
		ρ, Y, a	ρ, Y, a, ω	ρ, Y, a, ω, e	ρ, Y, a, c_{DKL}	ρ, Y, a, c	ρ, Y, c	ρ, Y, ω, c	ρ^*, Y, ω^*, c	ρ, Y, a, ω, e, c
N		231	579	779	485	1185	1055	1937	1368	2287
Barten's CSF (Barten, 1999)	13	3.5±0.48	3.9±0.35	4.5±0.31	✗	✗	✗	✗	✗	✗
VDP CSF (Daly, 1993)	6	5.4±0.54	✗	11±0.49	✗	✗	✗	✗	✗	✗
Wuerger et al. spatio-chromatic CSF (Wuerger et al., 2020)	27	3.5±0.71	✗	✗	4.5±0.54	✗	✗	✗	✗	✗
Mantiuk et al. spatio-chromatic CSF (Mantiuk et al., 2020)	31	**2.7 ± 0.32**	✗	✗	**3.1 ± 0.25**	**3.2 ± 0.21**	**3 ± 0.089**	✗	✗	✗
stelaCSF (Mantiuk et al., 2022)	21	2.8±0.47	4.4±0.37	4.4±0.41	✗	✗	✗	✗	✗	✗
Chromatic pyramid of visibility (Watson, 2021)	15	✗	✗	✗	✗	✗	5.3±0.26	5.3 ± 0.093	4.8±0.36	✗
castleCSF	53	2.8±0.27	**3.5 ± 0.35**	**3.3 ± 0.23**	3.2±0.22	3.3±0.22	3.2±0.14	**3.5 ± 0.16**	**3.5 ± 0.17**	**3.6 ± 0.12**

in the table. Note, the parameters of the final version shown in Tables 5–7 are reported for the model fitted to all available data.

Because most existing CSFs model fewer stimulus parameters than castleCSF, we tested other models on the subsets of our complete dataset, shown in the columns of Table 4. For each subset, we selected only the datasets that test the dimensions modeled by the CSFs being tested and fitted each model on that subset. In particular, all the models can predict contrast sensitivity changes along spatial frequency and luminance, and so these properties were included in every comparison. The CSFs that did not model temporal frequency were tested only on static stimuli (0 Hz). Similarly, the CSFs that did not model the effect of eccentricity were tested on foveal (0° retinal eccentricity) stimuli only. In the comparisons where the effect of stimulus size was not tested, we kept only either the fixed size or the fixed cycles subsets from each dataset. The comparisons with chromatic datasets consisted of either stimulus modulated only along the cardinal color directions (c_{DKL}) or modulations along any arbitrary color directions over any background color (c). The number of data points (N) included in each comparison is listed in the header of the table.

The Chromatic Pyramid of Visibility (Watson, 2021) required special treatment as this model is not intended to predict sensitivity for low spatial and temporal frequencies. For that reason, we report

the results of the test with and without (column with ρ^*, ω^* Table 4) low-frequency stimuli. We excluded the stimuli for which $\rho < 4$ cpd and $\omega < 4$ Hz unless it was isoluminant chromatic stimuli, in which case we excluded stimuli for which $\rho < 1$ cpd and $\omega < 1$ Hz.

Instead of proposing a complex physiologically-inspired model, we could fit a neural network to the data, which may provide an even better fit. We experimented with that idea and fitted a Multi-Layer Perceptron (MLP) using the same training/test data split as other models. The MLP had 6 layers, 120 neurons each (the architecture with the lowest validation loss). When fitted to all datasets, the error was larger than for castleCSF (5.8 dB), but more importantly, the MLP could not predict plausible trends or extrapolate the data. The MLP predictions can be found at the project page[2]. This failure of general function approximations can be explained by the scarcity and non-uniform distribution of contrast sensitivity data. Physiologically inspired models introduce regularization that helps them converge to plausible solutions.

castleCSF produces the lowest prediction errors for the majority of the comparisons in Table 4. The prediction error is within the range expected from inter-observer variation (between 3 to 3.9 dB or approximately half an octave) (Dakin & Turnbull, 2016). As

---

[2]Project page: `https://www.cl.cam.ac.uk/research/rainbow/projects/castleCSF`
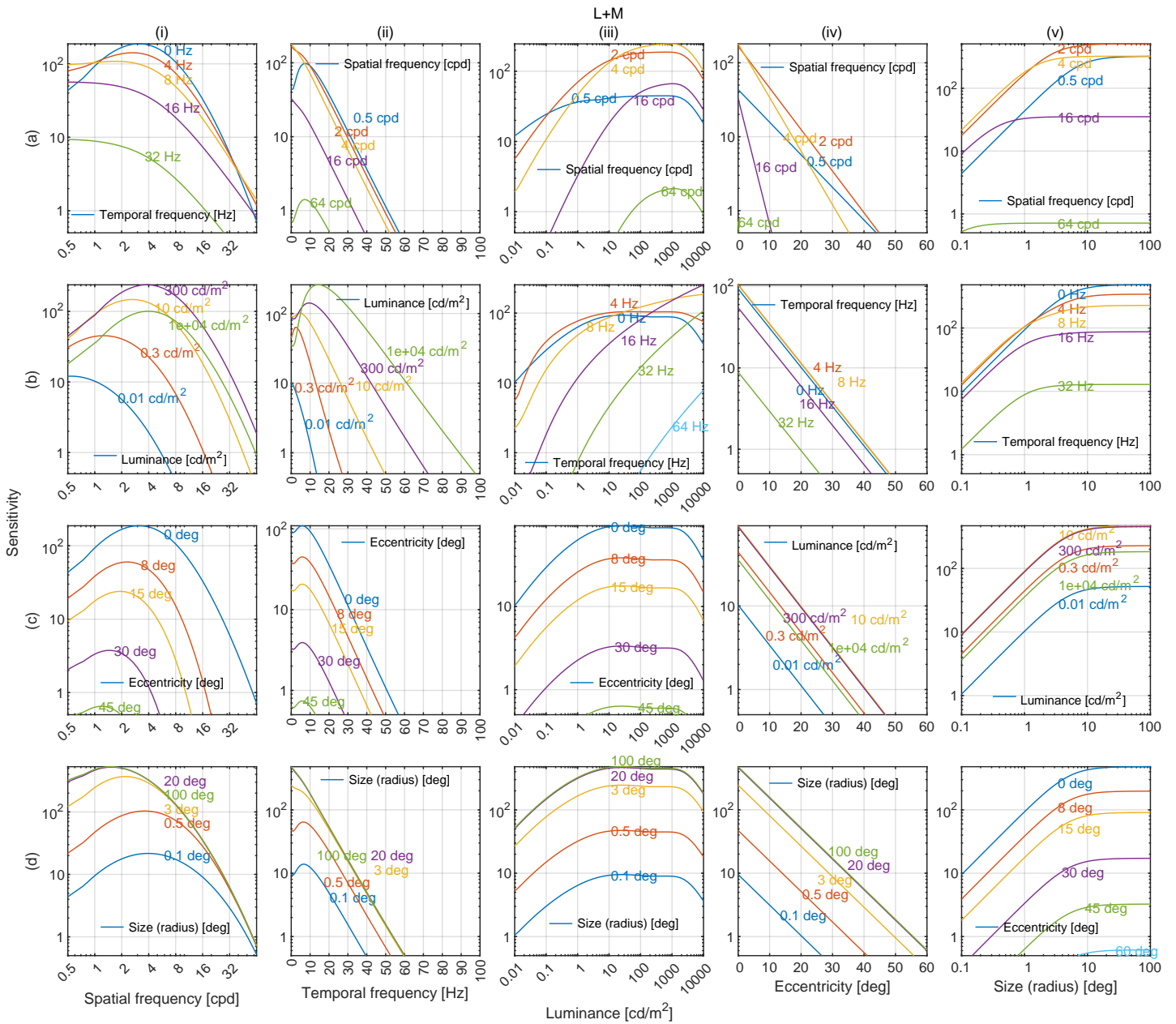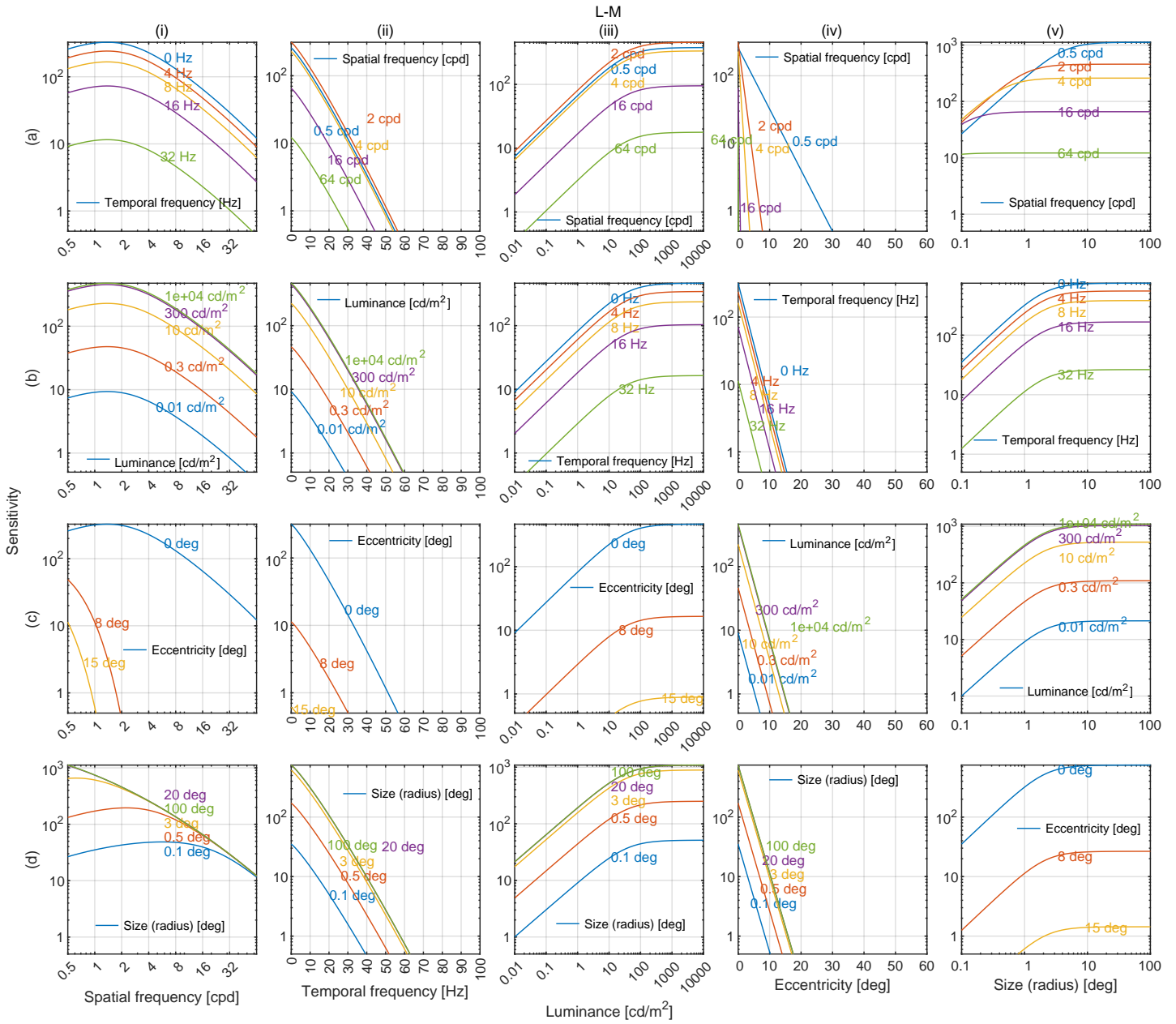
Figure 8: *Achromatic contrast sensitivity predictions by castleCSF along 5 dimensions of stimuli. The unit vector of the incremental cone response* $(\Delta\mathcal{L}, \Delta\mathcal{M}, \Delta\mathcal{S})$ *of the stimuli is:* $(0.6612, 0.3388, 0)$. *The contrast sensitivity response predictions are shown in the 5 columns with spatial frequency, temporal frequency, luminance, eccentricity, and size as independent variables respectively. In each column, the four plots show the sensitivity predictions with each of the remaining 4 model dimensions as the second independent variable. The remaining parameters are fixed as spatial frequency = 1 cpd, temporal frequency = 0 Hz, luminance = 30 cd/m², eccentricity = 0°, stimulus size (radius) = 1°. The rows and columns are notated as a-b and i-v respectively for ease of reference in the discussion.*

a reference, the expected inter-observer variation for the selected datasets is listed in the rightmost column of Table 2. Notably, most CSF models result in significantly larger errors, even when tested on a much more restricted portion of the dataset. A notable exception is Mantiuk et al., 2020's post-receptoral spatio-chromatic

model, which is used as a component of castleCSF. However, this model does not cover temporal frequency and eccentricity. Overall, no existing CSF can account for the same number of stimuli parameters and provide the same accuracy of prediction produced by castleCSF.
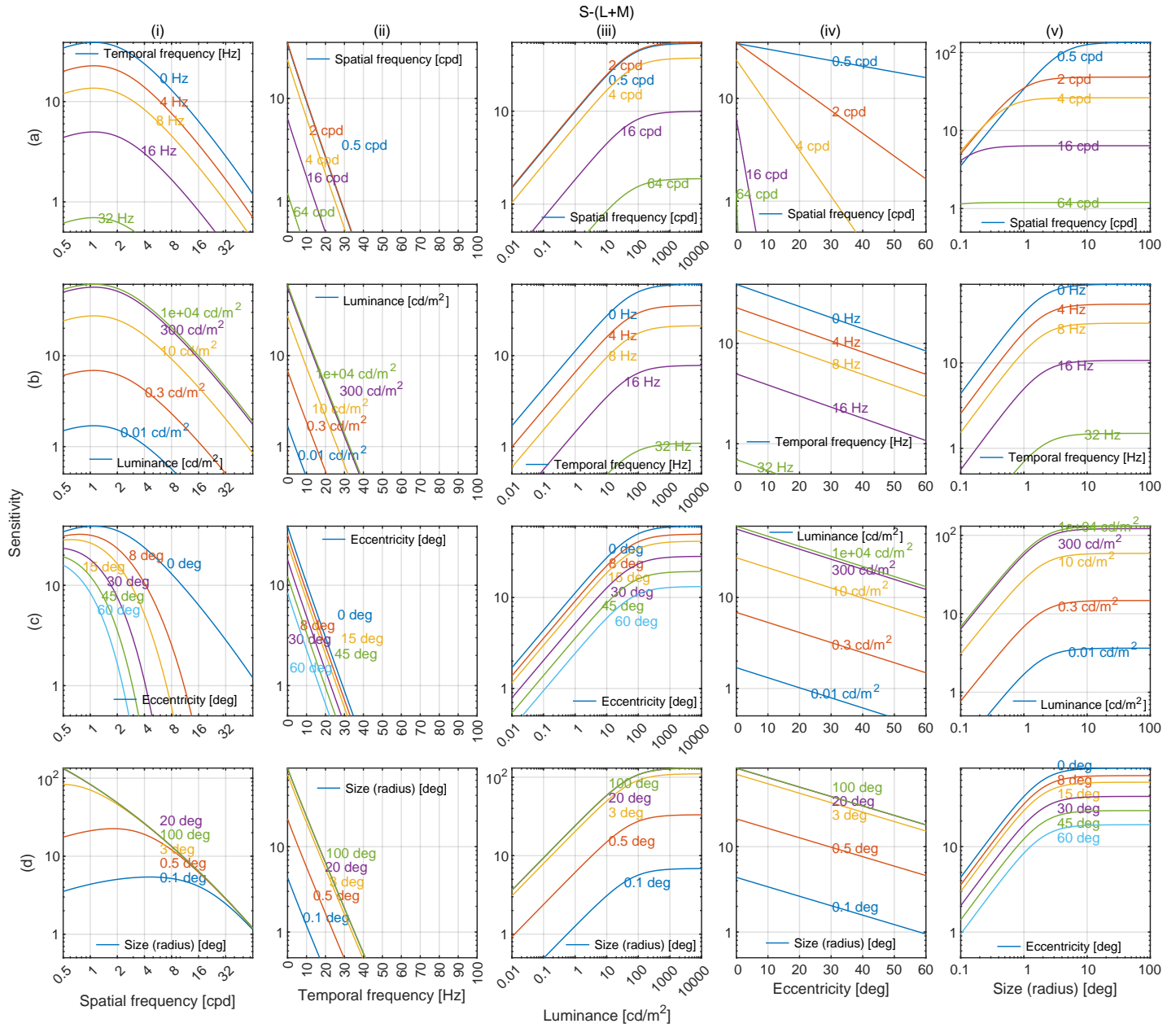
Figure 9: *Red-green contrast sensitivity predictions by castleCSF along 5 dimensions of stimuli. The unit vector of the incremental cone response* ($\Delta\mathcal{L}$, $\Delta\mathcal{M}$, $\Delta\mathcal{S}$) *of the stimuli is:* (0.3388, -0.3388, 0). *The description of the plots is the same as in Figure* 8.

Model predictions

Cross-sections of castleCSF's predictions for all combinations of modeled parameters, when trained on the full dataset, are shown in Figures 8 (achromatic), 9 (red-green), and 10 (yellow-violet). For each subplot, the vertical axis depicts sensitivity, and each column models changes along spatial frequency, temporal frequency, luminance, eccentricity, and stimulus size respectively as the independent variable. Within each column, rows show the joint effect of the respective independent variable and each of the remaining variables on contrast sensitivity.

Aggregate prediction errors for different 2D cross-sections of the 5-dimensional parameter space are shown in Figure 11, along with the sample density of our combined dataset. The upper two rows in the figure represent predictions of achromatic data ($N = 779$), while the lower two show the remaining chromatic data points ($N = 1508$). The color bar in the plots represents the prediction error in decibels (dB) calculated using the error function in Eq. (33).

Figure 10: *Yellow-violet contrast sensitivity predictions by castleCSF along 5 dimensions of stimuli. The unit vector of the incremental cone response* $(\Delta\mathcal{L}, \Delta\mathcal{M}, \Delta\mathcal{S})$ *of the stimuli is:* $(0.0115, 0, 0.0115)$. *The description of the plots is the same as in Figure 8.*

Discussion

In this section, we discuss castleCSF predictions of individual datasets and how they relate to known psychophysical evidence. The predictions reported in this section pertain to the model fitted to the complete dataset, with the parameters listed in Tables 5–7 in the Appendix, using a per-dataset adjustment (s_d in Eq. (32)). The colored numbers shown in the Figures. 12-27 denote the fitting error for the corresponding subset of conditions, reported in dB units.

Sensitivity attenuation at high spatial frequencies

Our model predicts that contrast sensitivity decreases logarithmically with spatial frequency for both achromatic and chromatic mechanisms in the high spatial frequency region, as depicted in the plots in column (i) of Figures 8, 9, and 10. This attenuation can be observed for achromatic datasets in Figures 12a, 12b (C1),
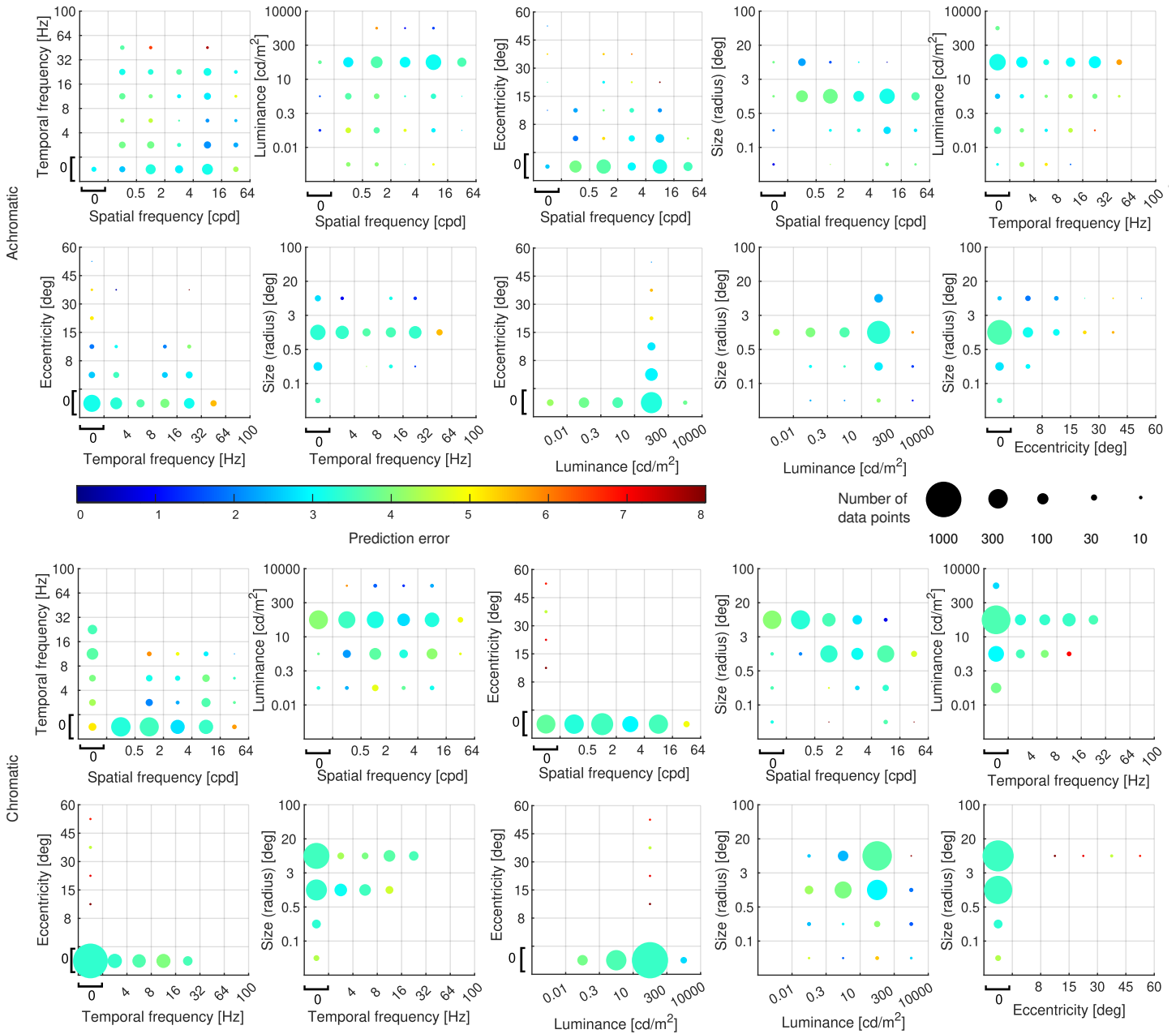
Figure 11: castleCSF prediction errors along with data points density for different 2D projections of the 5-dimensional parameter space. Top two rows: achromatic data. Lower two rows: chromatic data. The upper and lower boundaries of the bins are indicated by the corresponding tick labels on the axes. For discs, static and foveal stimuli, the whole bin represents spatial frequency = 0 cpd, temporal frequency = 0 Hz or eccentricity = 0° respectively. The color bar in the plots represents the prediction error in decibels (dB) calculated using the error function in Eq. (33). The size of the circles represents the number of data points for each combination of parameter values.

13a, 14a, 15, and 16(L+M). The higher frequency decrement can also be observed in chromatic datasets in Figures 12b (C2-C6), 16 (L-M, and S-(L+M)), and 17-20.

For fixed-cycle stimuli and foveal vision, the high-spatial frequency attenuation can be solely explained by the optics of the eye (Banks, Geisler, & Bennett, 1987; Campbell & Green, 1965;

Rovamo et al., 1993). Moreover, it has been shown that the slope of the MTF of the human eye changes with different pupil sizes (Watson, 2013; Deeley, Drasdo, & Charman, 1991; Van Meeteren, 1974), which in turn contract with increasing luminance levels. It follows that for fixed-cycle stimuli in natural viewing conditions, the slope of the CSF at high spatial frequencies should change with

the pupil diameter and, therefore, with luminance. Unfortunately, we do not have a dataset that could demonstrate this effect (fixed-size HDR-CSF was only measured up to 6 cpd). As a result, this effect is not captured by our model, which predicts this slope to be independent of luminance level.

It is well known that we are less sensitive to high spatial frequencies in the chromatic (isoluminant) color directions, owing to the lower resolving power of chromatic pathways (Mullen, 1985). Given that, it may come as a surprise that the ColorFest data shown in Figure 12b and HDR-CSF in Figure 16 show higher sensitivity for chromatic L-M and C2 data than for the achromatic L+M and C1 data — in particular at high spatial frequencies. This can be attributed to two issues: First, the cone contrast units, which define the sensitivity values, use an arbitrary scale for each color direction, and therefore the sensitivity values for the L+M direction cannot be directly compared to those for the L-M one, or any other. We choose the scale where the sum of CIE, 2006 L and M cone responses is equal to the Stockman and Sharpe luminous efficiency function (Sharpe, Stockman, Jagla, & Jägle, 2005) and the

resulting DKL responses reflect the length in this contrast space (Capilla, Malo, Luque, & Artigas, 1998). In such a scale of cone fundamentals, the red-green mechanism has the strongest response (Chaparro, Stromeyer, Huang, Kronauer, & Eskew, 1993; Cole, Hine, & McIlhagga, 1993; Eskew, McLellan, & Giulianini, 1999; Y. J. Kim, Reynaud, Hess, & Mullen, 2017), that is, a small change in cone contrast space would produce a large response in the red-green opponent color direction quantitatively compared to that of the achromatic or yellow-violet early vision mechanism. As a result, the perceived magnitude of contrast differs across the three dimensions of the DKL space (Switkes & Crognale, 1999). Second, neither of the two experiments attempted to isolate the chromatic mechanism of each participant (e.g. via the use of heterochromatic flicker). Therefore, the stimuli in both datasets likely contain contrast that could be detected by achromatic mechanism. For exam-
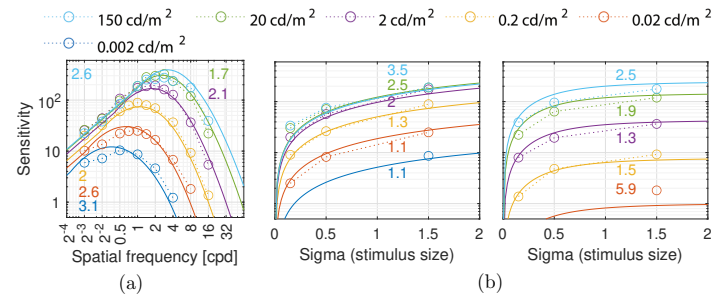

Figure 13: Predictions for the HDR-VDP CSF (Mantiuk et al., 2011) dataset. The stimuli in the sub-plot (a) were vertically oriented achromatic static Gabor patches with a Gaussian envelope of $\sigma = 1.5°$ shown at different luminance levels. The two sub-plots (b) show measurements from similar stimuli of different sizes with spatial frequencies of 1 and 8 cpd. The viewing mode was binocular with natural pupils.
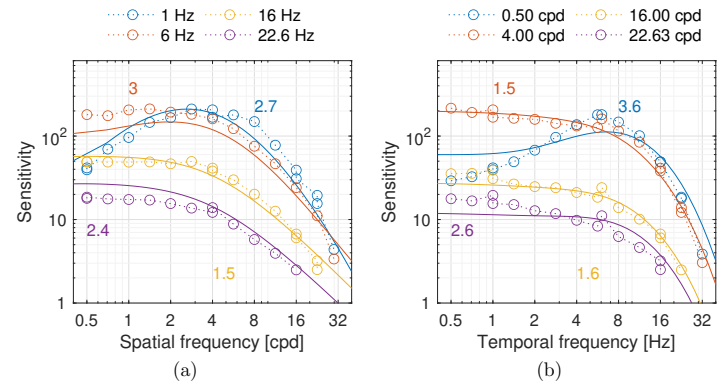

Figure 12: Predictions for (a) the Modelfest (Watson, 2000) and (b) Colorfest Wuerger et al., 2002 dataset. (a) Modelfest: The stimuli were horizontally oriented achromatic static Gabor patches with either constant size (Gaussian envelope of $\sigma = 0.5°$) or constant cycles (2 visible cycles approximately) tested at 30 cd/m². The viewing mode was binocular with natural pupils. (b) Colorfest: The stimuli were static horizontal Gabor patches of fixed size ($\sigma = 0.5°$). The color modulations were C1: black-white, C2: reddish-greenish, C3: yellowish green-violet, C4: greenish-Pink, C5: yellow-blue, C6: dark green-light pink, C7: dark yellow-light blue. C1, C2, and C3 were approximately the cardinal color directions of the human visual system. The background was D65 grey at 40 cd/m². The viewing mode was binocular with natural pupils. The colored numbers shown in the plot denote the prediction error, in dB, per subset of conditions.

Figure 14: Predictions for the Robson, 1966 dataset. The stimuli were achromatic sinusoidal gratings with rectangular apertures viewed at 20 cd/m². The area of the gratings was fixed at 6.25 $\deg^2$.

ple, in Figure 12b(C2), the spatial cut-off frequency is predicted to be at least 32 cpd, while Mullen, 1985 estimates the cut-off for an isolated red-green mechanism to be about 12 cpd.

Lateral inhibition of low spatial frequencies

Contrast sensitivity is reduced for static achromatic patterns at low spatial frequency. This effect is attributed to lateral inhibition (Barten, 1999) — the mechanism that helps us adapt to and perceive scenes spanning very large ranges of luminance. This results in a band-pass shape of the CSF for the achromatic modulation direction. When temporal frequency is increased, the effect of
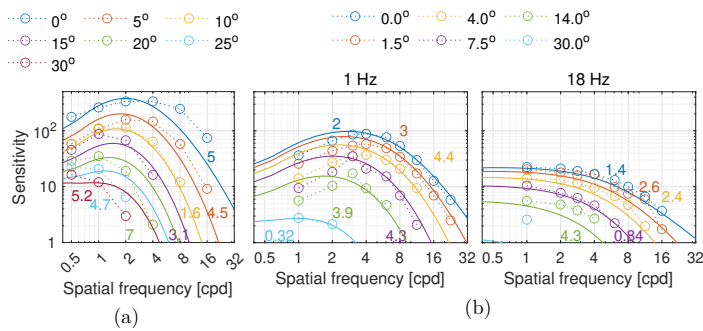
Figure 15: *Predictions for (a) the Virsu & Rovamo, 1979 and (b) Virsu et al., 1982 dataset. (a) The stimuli were static vertical achromatic sinusoidal gratings with circular apertures viewed at 10 cd/m². The size of the aperture was fixed at a 2.5° radius. (b) The stimuli were horizontal achromatic sinusoidal gratings with semi-circular apertures viewed at 10 cd/m². The area of the gratings was fixed at 1.57 deg². The stimuli at different retinal eccentricities were moved along the horizontal retinal meridian in the nasal visual field in both datasets and the viewing mode was monocular with natural pupils.*
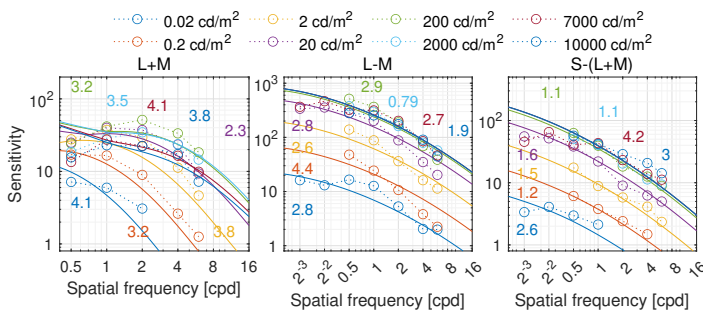

Figure 16: *Predictions for the HDR CSF Wuerger et al., 2020 dataset. The stimuli were static vertical Gabor patches of fixed cycles (2 visible cycles approximately). The three color modulations represent the cardinal color directions of the human visual system. The viewing mode was binocular with natural pupils.*

lateral inhibition is reduced and the CSF gradually becomes low-pass, as seen in Figure 14a and Figure 15b. Donner & Hemilä, 1996 modeled this band-pass to low-pass spatio-temporal response as a Difference-of-Gaussian (DOG) receptive field model and concluded that the spatial and temporal integration mechanisms are not separable. castleCSF models this effect by gradually shifting the response from a (band-pass) sustained to a (low-pass) transient channel.

There is no evidence of lateral inhibition in the isolated chro-
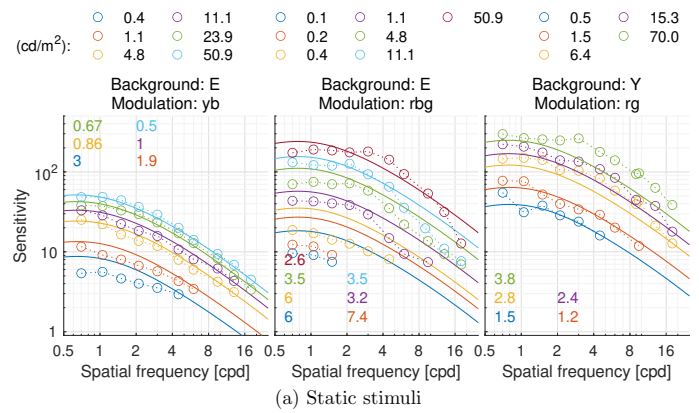

Figure 17: *Predictions for Van der Horst & Bouman, 1969 dataset. (a) Static stimuli at different luminance levels. (b) Top: Temporally-modulated stimuli at high luminance level, Bottom: temporally-modulated stimuli at low luminance level. The stimuli were Gabor patches of fixed size (14.84 deg²). The three columns show results from the three color modulations tested: yellow-to-blue and red-to-bluish-green over grey (E: equal-energy point) background, and red-to-green over yellow (Y) background. The stimuli were viewed monocularly through an artificial pupil of 2 mm diameter.*
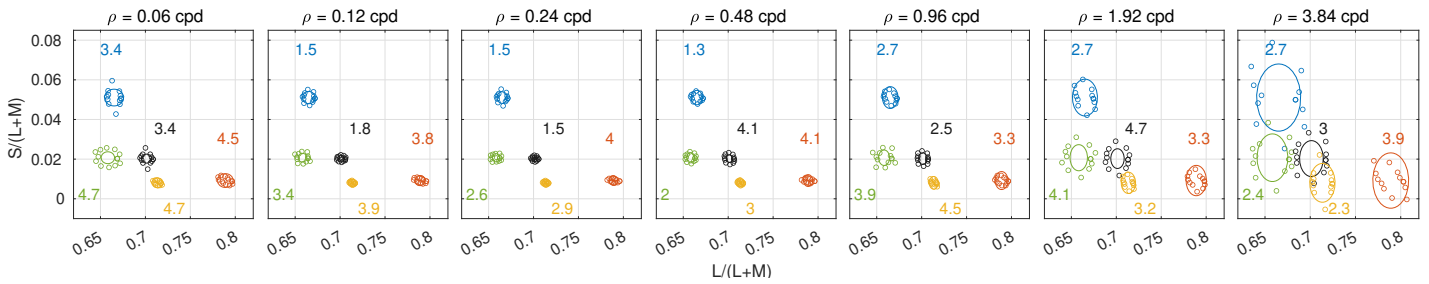
Figure 18: *Predictions for five centers (Xu et al., 2020) dataset. Larger ellipse sizes correspond to lower sensitivity and vice versa. The stimuli were large static Gabor patches with Gaussian envelope of $\sigma = 9.3°$ modulated over 6 color directions in u'v' space on grey (72 cd/m²), green (24 cd/m²), red (14.1 cd/m²), blue (8.8 cd/m²) and yellow (50 cd/m²) backgrounds. The viewing mode was binocular with natural pupils.*

matic mechanism (Kelly, 1983; Mullen, 1985; Metha & Mullen, 1996). However, since most of our datasets were measured for the cardinal color directions with an expected intrusion of achromatic contrast, we can see a small amount of attenuation for lower frequencies in Figure 16 (L-M, S-(L+M)) and in Figures. 19-20.

As the eccentricity of the stimulus increases, the peak of the CSF's band-pass shape shifts towards lower frequencies, as can be seen in Figure 15. This shift appears to be much smaller for the smaller flickering patterns used for the data in Figure 15b than for larger static patterns used for the data in Figure 15a. castleCSF does not currently model this difference.

Temporal response

The temporal response of the visual system is mostly determined by the photoreceptors (Hood & Birch, 1993), which restrict the highest temporal frequency that can be detected. This frequency is typically modelled as the critical flicker fusion frequency
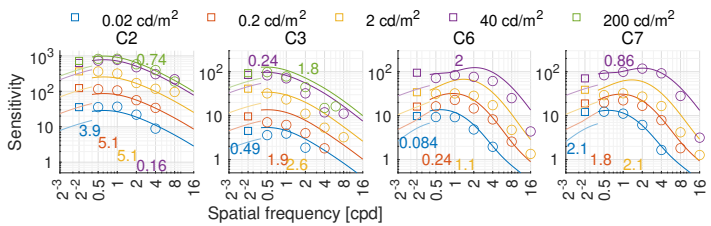
(CFF), corresponding to the temporal frequency at which the sensitivity curve crosses the $S = 1$ line. We did not include in our datasets any CFF measurements — we found modelling this type of data to be highly problematic as it is typically reported for flickering disks rather than Gabors. While the generalization of Gabors and disks shown in Eq. (31) works well for static disks, we found it does not seem to extend well to flickering disks. However, our two-channel model provides a good explanation of the available sensitivity data for achromatic patterns, shown in Figure 14b, 21, 22, and for chromatic patterns, shown in Figure 23. It should be noted that the band-pass shape and reduced sensitivity at low spa-
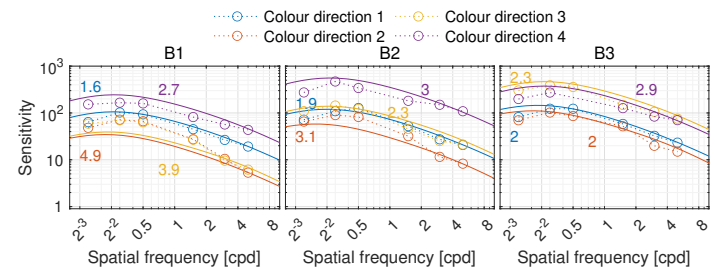

Figure 20: *Predictions for the Lucassen et al., 2018 dataset. The stimuli were large static Gabor patches with Gaussian envelope of $\sigma = 9.45°$ modulated along 4 color directions in u'v' space on B1 (2600K, orange appearance), B2 (3700K, yellow appearance), and B3 (5600, cool/daylight white appearance) backgrounds. The stimuli were shown at 108 cd/m². The psychophysical task was orientation discrimination.*

Figure 19: *Predictions for the K. J. Kim et al., 2013 dataset. The stimuli were horizontal static Gabor patches with a Gaussian envelope of $\sigma = 3°$ for the 0.25 cpd stimuli and $\sigma = 1.5°$ for all other spatial frequencies shown at different luminance levels. The color modulations were C2: reddish-greenish, C3: yellowish green-violet, C6: dark green-light pink, and C7: dark yellow-light blue. C1, C2, and C3 were approximately the cardinal color directions of the human visual system. The background was D65 grey. The viewing mode was binocular with natural pupils.*
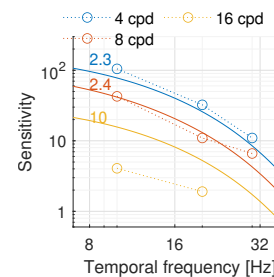

Figure 21: *Predictions for the Laird et al., 2006 dataset. The stimuli were achromatic vertical Gabor patches of fixed size (Gaussian envelope of $\sigma = 2.46°$) displayed at 60 cd/m², which were viewed binocularly with natural pupils.*
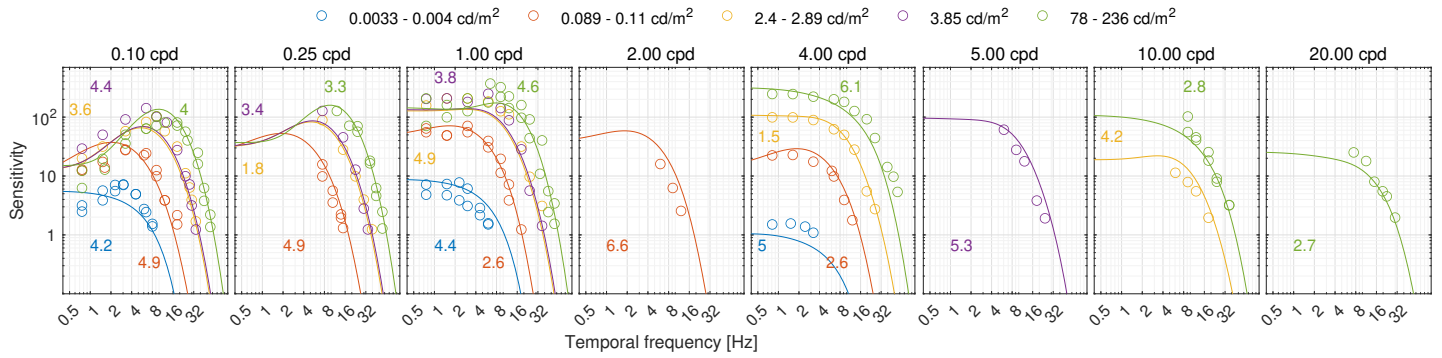
Figure 22: *Predictions for the Snowden et al., 1995 dataset. The stimuli are achromatic Gabor patches modulated in both space and temporal domains. The Gaussian envelope size of the lower spatial frequencies up to 5 cpd is $\sigma = 1.6°$, while that of 10 cpd and 20 cpd is $0.8°$ and $0.4°$ respectively. The lowest frequency (0.1 cpd) shown in this plot was actually a 0 cpd Gaussian blob but we found that our model could predict this by assuming it as a Gabor patch with half a sinusoidal cycle visible. The viewing mode was binocular with natural pupils.*

tial and temporal frequencies is due to lateral inhibition, present in the sustained channel.

Effect of mean luminance

The luminance-dependent achromatic contrast sensitivity response can be generally divided into three regions; 1) the low-luminance region following the DeVries-Rose law, 2) the mid-to-high luminance region following Weber's law, and 3) the very high luminance region in which sensitivity is reduced. It should be noted that the categorization of the DeVries-Rose and Weber regions are merely approximations, and the measured responses are the combined result of several visual mechanisms and adaptations at work. García-Pérez & Peli, 1997 and Rovamo, Näsänen, & Mustonen, 1997 have discussed some nuances of using the Weber law behavior as a proxy for different physiological mechanisms.

The three different regions can be observed in plots (a-d)(iii) in Figure 8. The chromatic contrast sensitivity functions present only the DeVries-Rose and the Weber region as shown in plots (a-d)(iii) in Figures. 9-10. The linear log sensitivity response with respect to log luminance obeys the DeVries-Rose law, which states that the incremental threshold contrast is proportional to the square root of the background intensity (De Vries, 1943; Rose, 1948). In the log-log scale, this square root relationship becomes linear, with an approximate slope of 0.5. With a further increase in mean luminance, this slope starts becoming flatter in accordance with Weber's law, implying that contrast becomes independent of mean luminance (Weber, 1831). The point of transition between the DeVries-Rose region and the Weber region depends on the spatial and temporal frequencies of the stimuli. Looking at the predic-

tions in plot (a)(iii) in Figure 8, this transition happens at lower luminances for lower spatial frequencies, and at higher luminances for higher spatial frequencies. This spatial frequency-dependent transition is also shown in the work of Rovamo et al., 1995 and is predicted well by our model, as demonstrated in analysis over datasets from HDR-VDP CSF (Mantiuk et al., 2011) (Figure 13a)
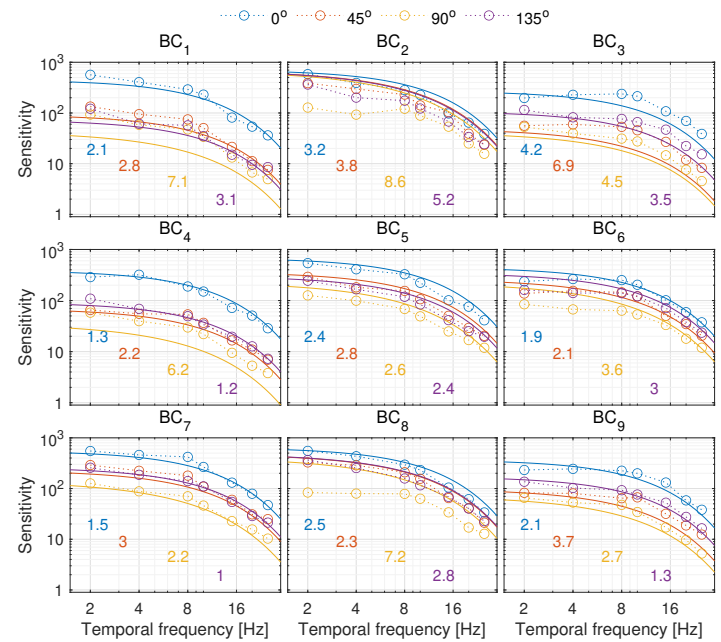

Figure 23: *Predictions from Kong et al., 2018 dataset. The stimuli were static discs modulated along 4 color directions in u'v' space on 9 different backgrounds. The approximate colors of the nine backgrounds are jade green, violet, rose red, rust orange, cornflower blue, magenta, light pink, dull violet, and dull rose red respectively. The radius of the discs was $5°$ and the mean luminance was $35\,cd/m^2$. The viewing model was binocular with natural pupils.*

and HDR-CSF (Wuerger et al., 2020) (Figure 16).

For very high luminance levels (above $2\,000\,\text{cd/m}^2$), our model predicts a decrease in contrast sensitivity which deviates from Weber's law. This prediction is mainly based on the HDR-CSF dataset (Figure 16), which shows that achromatic contrast sensitivity decreases for luminance levels above $2000\,\text{cd/m}^2$. This reduction in contrast sensitivity cannot be explained by photopigment bleaching (Rovamo et al., 1997), as bleaching would proportionally reduce the response to both the background and the stimulus and thus cannot cause a reduction of retinal contrast, nor can it explain the lack of sensitivity reduction in the chromatic L-M and S-(L+M) modulation directions as the effect of bleaching would only be dependent on the luminance level and would reduce both achromatic and chromatic contrast sensitivity for the same light level. This effect cannot also be explained by increased diffraction due to the contraction of the pupil at high luminance levels, as the effect is most salient at low frequencies which are least affected by diffraction. We speculate this reduction of sensitivity may be caused by the lateral inhibition mechanism, as it can be observed mostly at low frequencies for the achromatic mechanism, where this type of inhibition is expected to be strongest. We do not have any data to support this hypothesis, but Xin & Bloomfield, 1999 have shown that the extent of lateral inhibition is dependent on the mean luminance level, with the maximum coupling between horizontal cells (and consequently the strength of contrast enhancement by lateral inhibition) occurs at moderate ambient light levels and extreme dark or light adaptation causes this effect to be weakened.

Foveal and peripheral contrast sensitivity

Contrast sensitivity decreases with eccentricity as shown in plots (a-d)(iv) in Figures 8, 9, and 10. The slopes of these linearly decreasing functions are dependent on the spatial frequency. The rate of sensitivity decrement with retinal eccentricity is slower for lower spatial frequency stimuli compared to higher spatial frequency ones, as predicted in plots (a)(iv) for the achromatic (Figure 8) and yellow-violet (Figure 10) channels. In the case of the red-green channel (Figure 9), our model does not predict any spatial frequency dependence on the slope of sensitivity reduction with respect to retinal eccentricity. The predictions for achromatic stimuli are in line with the data from Virsu & Rovamo, 1979 (Figure 15a) and Virsu et al., 1982 (Figure 15b). At lower spatial frequencies, the difference between their data measured at differ-

ent eccentricities is smaller, depicting a lower rate of sensitivity decline. The dataset from Wright & Johnston, 1983 also shows similar findings (Figure 24) for four different spatial and temporal frequencies. The predictions for the two chromatic channels in our model are entirely driven by the Hansen et al., 2009 dataset (Figure 25). Our model could well predict the achromatic and yellow-violet stimuli from this dataset but was less accurate in predicting the red-green stimuli at higher retinal eccentricities. This could be explained by the different distribution of cells picking up red-green versus yellow-violet signals across the periphery (Mullen et al., 2002). The sensitivity of the red-green opponent mechanism also declines at a much higher rate compared to the achromatic and yellow-violet mechanisms (Mullen et al., 2002, 2005). More contrast sensitivity data for chromatic stimuli measured in the periph-
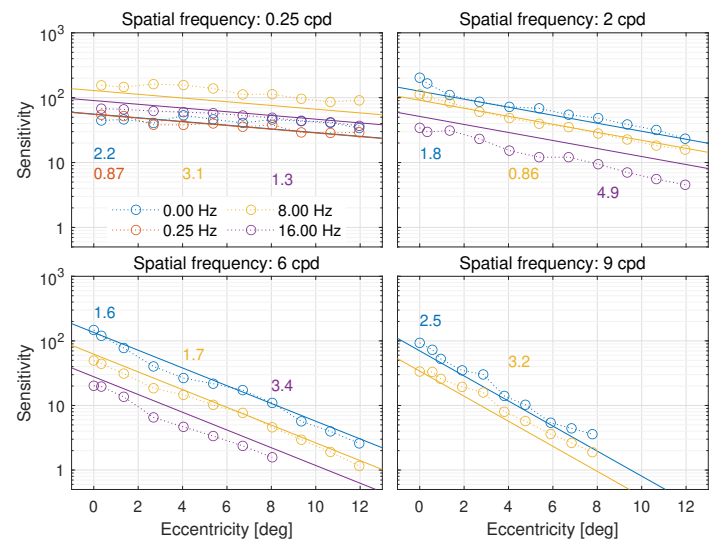

Figure 24: Predictions for the Wright & Johnston, 1983 dataset. The stimuli were vertical achromatic sinusoidal gratings with rectangular apertures tapered at the edges viewed at $100\,\text{cd/m}^2$. The area of the gratings was $36\,\text{deg}^2$ for 0.25 cpd, $2.345\,\text{deg}^2$ for 2 and 6 cpd stimuli, and $0.75\,\text{deg}^2$ for 9 cpd stimuli. The stimuli at different retinal eccentricities were shown along the vertical retinal meridian in the superior visual field. The viewing mode was binocular with natural pupils.
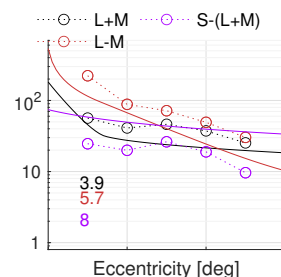

Figure 25: Predictions for the Hansen et al., 2009 dataset. The stimuli were $4°$ diameter discs modulated along the three cardinal color directions in the DKL colorspace.

ery is required to correctly model these differences. Another interesting direction would be to add the effect of photoreceptors other than cones to predict chromatic contrast sensitivity in the periphery. Though Weale, 1953 ruled out the contribution of rods in chromatic sensitivity in the periphery, Horiguchi, Winawer, Dougherty, & Wandell, 2013 have shown that peripheral contrast thresholds are better explained by a four-receptor model for high-luminance stimuli and speculated that this fourth receptor could be driven by melanopsin signals. More data is needed to correctly test these hypotheses in our model.

Spatial summation

Contrast sensitivity increases monotonically with stimulus size until a critical area is reached, following Riccó's law (Riccó, Annibale, 1877). The sensitivity remains constant for stimuli larger than this critical area as shown in plots (a-d)(v) in Figures 8, 9, and 10. The value of the critical area decreases with the spatial frequency of the stimulus for all three chromatic directions, as predicted in plots (a)(v) in Figures 8, 9, and 10. This is in agreement with the data from HDR-VDP CSF, shown in Figure 13b, and from Rovamo et al., 1993, shown in Figure 26. The HDR disc CSF dataset, shown in Figure 27 was measured for smaller sizes and thus the saturation at the critical area is not observed for these measurements. Although we lack a single dataset that would test the effect of size on the detection of chromatic patterns, the union of all datasets provides a good sampling of chromatic patterns of different sizes, as shown in Figure 11.

Note that we have a separate spatial summation model for the disc stimuli, where we integrate over the circumference of the disk (Eq. (31)). This model does not account for the likely saturation of

such a summation, as we do not have data to model this effect.

Our model assumes that the effect of stimulus size is independent of temporal frequency, luminance, and eccentricity, as shown in plots (b-d)(v) in Figures 8, 9, and 10. The independence of spatial summation from luminance agrees with the data from HDR-VDP CSF as shown in Figure 13b. For both spatial frequencies, we can see that the shape of the curves formed by the data points does not change between luminance levels.
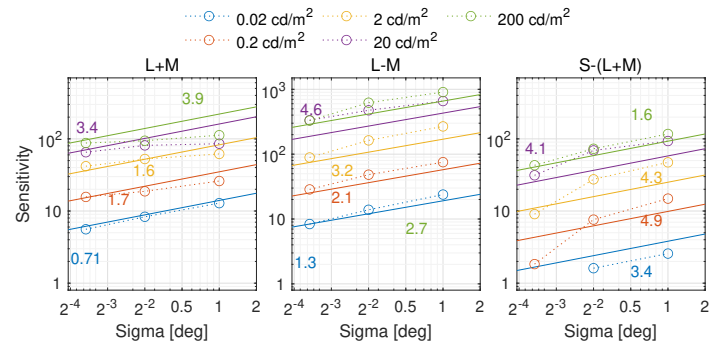
Figure 27: Predictions for HDR disc CSF (Ashraf et al., 2023) dataset. The stimuli were static discs modulated from a grey background to either white (L+M), pinkish red (L-M), or violet (S-(L+M)). The viewing mode was binocular with natural pupils.

Applications of the CSF

This section highlights engineering applications of castleCSF and shows an example in which it is used to assess the visibility of distortions caused by chroma subsampling.

One of the main goals of creating castleCSF was to use it as a core component of the Color Video Visual Difference Predictor (ColorVideoVDP[3]). This visual difference metric is used to compare a distorted image or video to its reference and predict the quality degradation due to display or coding distortions. Figure 28 shows an example in which the metric predicts the quality drop due to chroma subsampling, a computational technique commonly used in video and image compression. The effect of subsampling of both chroma channels in the CIELAB color space, the predicted visual difference maps, and the overall quality drop scaled in the Just Objectionable Difference units (Perez-Ortiz et al., 2019) are presented. We can observe that chroma subsampling results in different levels of visible distortions across the image, with isoluminant areas, which lack luminance contrast, being the most affected.
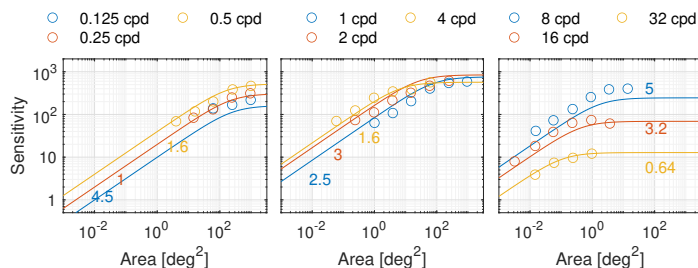

Figure 26: Predictions for the Rovamo et al., 1993 dataset. The stimuli were achromatic static vertical cosine gratings of different spatial frequencies with square apertures of varying areas displayed at 50 cd/m². The viewing mode was binocular with natural pupils.

---

[3]Code available at https://github.com/gfxdisp/ColorVideoVDP

(a) Reference image with isoluminant text

(b) Just Objectionable Difference (JOD) predictions

Difference in JOD scale

(c) Heatmap scale

(d) 2x chroma subsampled image

(e) Heatmap of predicted small perceivable distortions

(f) 4x chroma subsampled image
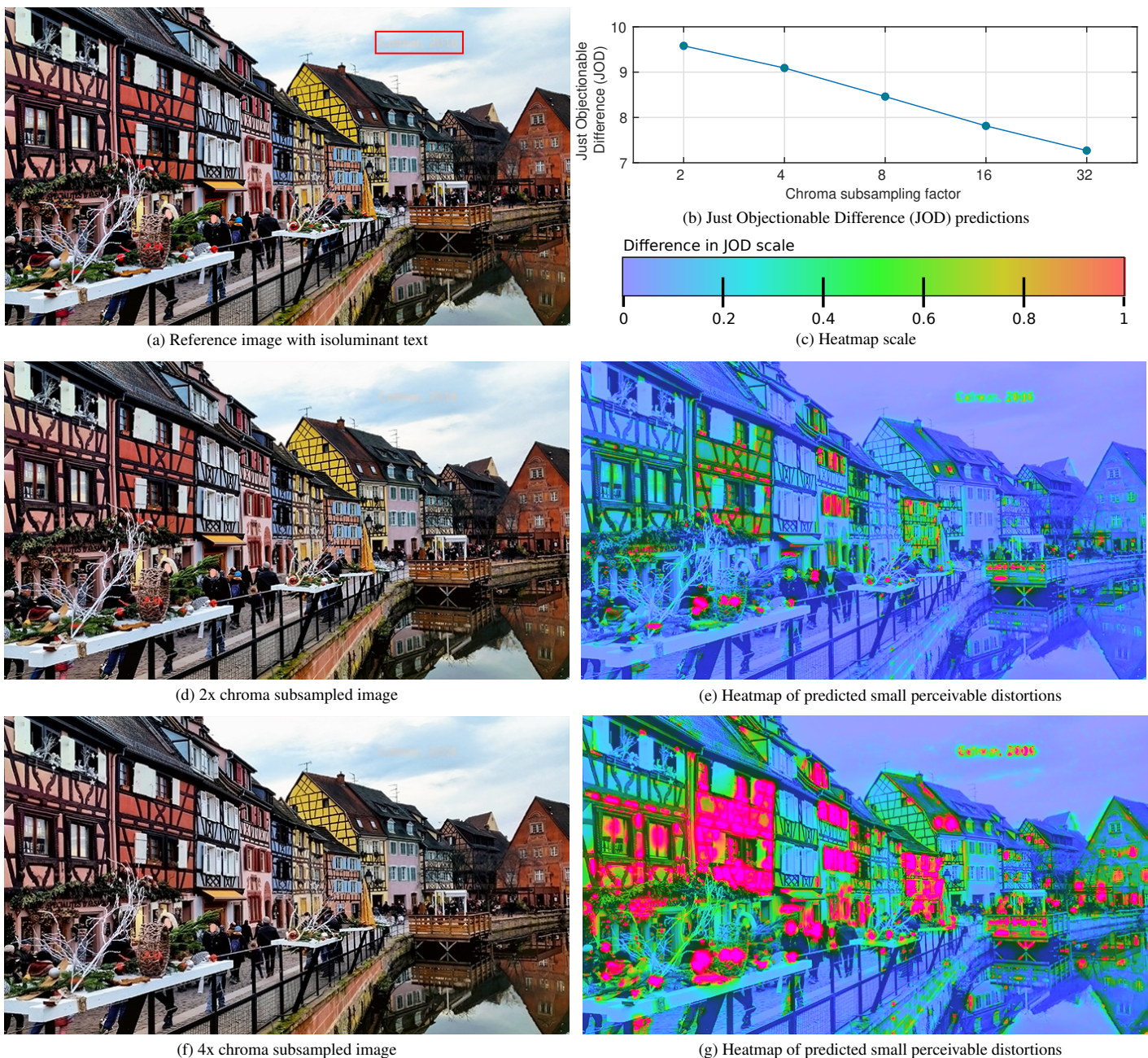
(g) Heatmap of predicted small perceivable distortions

Figure 28: *castleCSF is a fundamental building block of the ColorVideoVDP metric that can predict visible differences in complex images and videos. This figure shows a prediction of the visibility of chroma subsampling. (a) Reference image, (b) Plot of the metric's prediction scaled in Just Objectionable Differences (JODs) (higher values denote higher quality) against the chroma subsampling factors, showing that stronger subsampling leads to higher perceived distortion, (d and f) Chroma-subsampled versions at decreasing spatial resolutions of the chroma channels (1/2, and 1/4 respectively), (e and g) Heatmaps of predicted small (blue to red) perceivable distortions, for each subsampled image, the scale of the distortions can be seen in (c). Note that distortions in areas without luminance contrast (such as the text isoluminant to the sky, marked with a red box in the reference image) are especially prone to distortion from chroma subsampling. For visual assessment, these images should be viewed from a distance of 40 cm, with each image width adjusted to approximately 9 cm.*

Modeling the spatio-chromatic CSF, as done by castleCSF, is critical for this application.

CSF models have found many more applications in engineering. Contrast detection models were used to derive the DICOM grayscale function used in medical monitors (3.14-2003, 2003), and later again to define the encoding used for high-dynamic-range video standards (Mantiuk, Krawczyk, Myszkowski, & Seidel, 2004; Miller, Nezamabadi, & Daly, 2013; ITU, 2017). CSFs are used to characterize the performance of electronic displays (SID, 2022), including the visibility of flicker and display non-uniformity. Contrast sensitivity models are also employed to optimize image coding (Zeng, Daly, & Lei, 2002), match contrast visibility when reproducing high dynamic range images (Ward-Larson, Rushmeier, & Piatko, 1997; Mantiuk, Daly, & Kerofsky, 2008; Tariq et al., 2023), assess image quality (Zhang & Wandell, 1997; Chandler, 2010; Aydin, Čadík, Myszkowski, & Seidel, 2010; Mantiuk et al., 2011; Haun & Peli, 2013), or optimize rendering in real-time computer graphics (Luebke & Hallen, 2001; Jindal, Wolski, Myszkowski, & Mantiuk, 2021). Overall, given the range of applications, it is desirable to build a general model that can explain contrast sensitivity for the large range of conditions encountered in these and similar applications.

The CSF is also useful for applications in optometry, for example, the optical and neural limits of vision can be inferred by contrast sensitivity measurements (Amesbury & Schallhorn, 2003). Spatial contrast sensitivity also provides a direct measure of visual performance (Owsley & Sloane, 1987), which is very useful for clinical evaluations and for use as a diagnostic tool for the detection of visual disorders (American Academy of Opthamology, 1990).

Conclusions

castleCSF is a practical model, intended to summarize and predict the average observer detection data for a wide range of values explored in the literature. It is mainly intended for engineering applications, in which similar models (Barten, 1999) found an ample range of use. The distinct features of castleCSF are that it explains a larger number of stimulus parameters than other models and that a single model (with the same set of parameters) can predict a broad range of measurements from the literature. While our approach could potentially mask subtle differences between datasets that occur under specific conditions and are unique to individual datasets, we believe that the benefits of incorporating a large number of datasets outweigh the limitations considering our model's broad application.

Our work also helps identify the gaps in the existing contrast sensitivity literature as shown in Figure 11, and can help direct future research to address these limitations. The missing datasets include edge CSF data outside the fovea, high luminance temporal CSF data for chromatic stimuli, low luminance parafoveal data, low spatial frequency chromatic temporal flicker, and data for chromatic temporal flicker in the periphery. In the case of chromatic parafoveal datasets, there is an additional challenge of representing the stimulus in cone contrast space, as the standard cone fundamentals are designed for foveal stimuli. Models of suitable modifications representing the changes in cone densities (Volbrecht, Shrago, Schefrin, & Werner, 2000; Curcio et al., 1990; Moreland & Cruz, 1959) and strength of the cone responses (Stabell & Stabell, 1996; Sakurai & Mullen, 2006) in the periphery exist. We aim to test different peripheral cone contrast metrics and the resulting peripheral opponent-color contrast channels in future iterations of this work.

Acknowledgments

References

3.14-2003. (2003). *DICOM Part 14: Grayscale Standard Display Function* (Tech. Rep.). National Electrical Manufacturers Association.

Adam, C. R., Shrier, E., Ding, Y., Glazman, S., & Bodis-Wollner, I. (2013). Correlation of Inner Retinal Thickness Evaluated by Spectral-Domain Optical Coherence Tomography and Contrast Sensitivity in Parkinson Disease. *Journal of neuro-ophthalmology*, *33*(2), 137–142. ,

doi:10.1097/WNO.0b013e31828c4e1a

Ahumada, A., & Peterson, H. A. (1992). Luminance-Model-Based DCT Quantization for Color Image Compression. In *SPIE 1666, Human vision, visual processing, and digital display III*.

Akbarinia, A., Morgenstern, Y., & Gegenfurtner, K. R. (2023). Contrast Sensitivity Function in Deep Networks. *Neural Networks*, *164*, 228–244. , doi:10.1016/j.neunet.2023.04.032

American Academy of Opthamology. (1990). Contrast Sensitivity and Glare Testing in the Evaluation of Anterior Segment Disease. *Ophthalmology*, *97*(9).

Amesbury, E. C., & Schallhorn, S. C. (2003). Contrast Sensitivity and Limits of Vision. *International ophthalmology clinics*, *43*(2), 31–42. , doi:10.1097/00004397-200343020-00006

Anderson, S. J., & Burr, D. C. (1985). Spatial and Temporal Selectivity of the Human Motion Detection System. *Vision Research*, *25*(8), 1147–1154. , doi:10.1016/0042-6989(85)90104-X

Anderson, S. J., Mullen, K. T., & Hess, R. F. (1991). Human Peripheral Spatial Resolution for Achromatic and Chromatic Stimuli: Limits Imposed by Optical and Retinal Factors. *The Journal of Physiology*, *442*, 47–64. , doi:10.1113/jphysiol.1991.sp018781

Anstis, S. M. (1974). Chart Demonstrating Variations in Acuity with Retinal Position. *Vision research*, *14*(7), 589–592.

Ashraf, M., Mantiuk, R. K., & Chapiro, A. (2023). Modelling Contrast Sensitivity of Discs. *Electronic Imaging*, *35*(10), 246–1.

Aydin, T. O., Čadík, M., Myszkowski, K., & Seidel, H.-P. (2010). Video Quality Assessment for Computer Graphics Applications. *ACM Transactions on Graphics*, *29*(6), 1. , doi:10.1145/1882261.1866187

Banks, M. S., Geisler, W. S., & Bennett, P. J. (1987). The Physical Limits of Grating Visibility. *Vision research*, *27*(11), 1915–1924. , doi:10.1016/0042-6989(87)90057-5

Barten, P. G. J. (1992). Physical Model for the Contrast Sensitivity of the Human Eye. In *Human vision, visual processing, and digital display III* (Vol. 1666, pp. 57–72). SPIE.

Barten, P. G. J. (1999). *Contrast Sensitivity of the Human Eye and Its Effects on Image Quality*. Bellingham, WA: SPIE Optical Engineering Press.

Blackwell, H. R. (1946). Contrast Thresholds of the Human Eye. *Journal of the Optical Society of America*, *36*(11), 624. , doi:10.1364/JOSA.36.000624

Blakemore, C., & Campbell, F. W. (1969). On the Existence of Neurones in the Human Visual System Selectively Sensitive to the Orientation and Size of Retinal Images. *The Journal of physiology*, *203*(1), 237–260. , doi:10.1113/jphysiol.1969.sp008862

Bozorgian, A., Pedersen, M., & Thomas, J.-B. (2022). Modification and Evaluation of the Peripheral Contrast Sensitivity Function Models. *JOSA A*, *39*(9), 1650–1658. , doi:10.1364/JOSAA.445234

Brainard, D., Jiang, H., Cottaris, N. P., Rieke, F., Chichilnisky, E., Farrell, J. E., et al. (2015). ISETBIO: Computational Tools for Modeling Early Human Vision. In *Imaging and Applied Optics 2015* (p. IT4A.4). Arlington, Virginia: OSA.

Campbell, F. W., & Green, D. G. (1965). Optical and Retinal Factors Affecting Visual Resolution. *The Journal of Physiology*, *181*, 576–593. , doi:10.1113/jphysiol.1965.sp007784

Campbell, F. W., Kulikowski, J. J., & Levinson, J. (1966). The Effect of Orientation on the Visual Resolution of Gratings. *The Journal of Physiology*, *187*(2), 427–436. , doi:10.1113/jphysiol.1966.sp008100

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier Analysis to the Visibility of Gratings. *The Journal of Physiology*, *197*(3), 551–566. , doi:10.1113/jphysiol.1968.sp008574

Capilla, P., Malo, J., Luque, MJ., & Artigas, JM. (1998). Colour Representation Spaces at Different Physiological Levels: A Comparative Analysis. *Journal of optics*, *29*(5), 324. , doi:10.1088/0150-536X/29/5/003

Cass, J., Clifford, C. W. G., Alais, D., & Spehar, B. (2009). Temporal Structure of Chromatic Channels Revealed through Masking. *Journal of Vision*, *9*(5), 17–17. , doi:10.1167/9.5.17

Chandler, D. M. (2010). Most Apparent Distortion: Full-Reference Image Quality Assessment and the Role of Strategy. *Journal of Electronic Imaging*, *19*(1), 011006. , doi:10.1117/1.3267105

Chaparro, A., Stromeyer, C., Huang, E., Kronauer, R., & Eskew, R. (1993). Colour Is What the Eye Sees Best. *Nature*, *361*, 348–350. , doi:10.1038/361348a0

Chapiro, A., Matsuda, N., Ashraf, M., & Mantiuk, R. (2023). Critical Flicker Frequency (CFF) at High Luminance Levels. *Electronic Imaging*, *35*, 1–5. , doi:10.2352/EI.2023.35.10.HVEI-223

CIE. (2006). *Fundamental Chromaticity Diagram with Psychological Axes - Part 1* (Tech. Rep.). Central Bureau of the Commission Internationale de l' Éclairage.

Coates, D. R., Chin, J. M., & Chung, S. T. (2013). Factors Affecting Crowded Acuity: Eccentricity and Contrast. *Optometry and vision science: official publication of the American Academy of Optometry*, *90*(7). , doi:10.1097/OPX.0b013e31829908a4

Cole, G. R., Hine, T., & McIlhagga, W. (1993). Detection Mechanisms in L-, M-, and S-cone Contrast Space. *Josa a*, *10*(1), 38–51. , doi:10.1364/JOSAA.10.000038

Cropper, S. J., & Wuerger, S. (2005). The Perception of Motion in Chromatic Stimuli. *Behavioral and Cognitive Neuroscience Reviews*, *4*(3), 192–217. , doi:10.1177/1534582305285120

Curcio, C. A., Sloan, K. R., Kalina, R. E., & Hendrickson, A. E. (1990). Human Photoreceptor Topography. *Journal of comparative neurology*, *292*(4), 497–523. , doi:10.1002/cne.902920402

Dakin, S. C., & Turnbull, P. R. (2016). Similar Contrast Sensitivity Functions Measured Using Psychophysics and Optokinetic Nystagmus. *Scientific Reports*, *6*(1), 34514. , doi:10.1038/srep34514

Daly, S. J. (1993). Visible Differences Predictor: An Algorithm for the Assessment of Image Fidelity. In A. B. Watson (Ed.), *Digital images and human vision* (Vol. 1666, pp. 179–206). MIT Press.

Daly, S. J. (1998). Engineering Observations from Spatiovelocity and Spatiotemporal Visual Models. In B. E. Rogowitz & T. N. Pappas (Eds.), *Human vision and electronic imaging* (Vol. 3299, pp. 180–191).

de Lange, D. H. (1958a). Research into the Dynamic Nature of the Human Fovea → Cortex Systems with Intermittent and Modulated Light. I. Attenuation Characteristics with White and Colored Light. *JOSA*, *48*(11), 777–784. , doi:10.1364/JOSA.48.000777

de Lange, D. H. (1958b). Research into the Dynamic Nature of the Human Fovea → Cortex Systems with Intermittent and Modulated Light. II. Phase Shift in Brightness and Delay in Color Perception. *JOSA*, *48*(11), 784–789. , doi:10.1364/JOSA.48.000784

Deeley, R. J., Drasdo, N., & Charman, W. N. (1991). A Simple Parametric Model of the Human Ocular Modulation Transfer Function. *Ophthalmic and Physiological Optics*, *11*(1), 91–93. , doi:10.1111/j.1475-1313.1991.tb00200.x

Denes, G., Maruszczyk, K., Ash, G., & Mantiuk, R. K. (2019). Temporal Resolution Multiplexing: Exploiting the Limitations of Spatio-Temporal Vision for More Efficient VR Rendering. *IEEE transactions on visualization and computer graphics*, *25*(5), 2072–2082. , doi:10.1109/TVCG.2019.2898741

Derrington, A. M., Krauskopf, J., & Lennie, P. (1984). Chromatic Mechanisms in Lateral Geniculate Nucleus of Macaque. *The Journal of physiology*, *357*(1), 241–265. , doi:10.1113/jphysiol.1984.sp015499

De Vries, HL. (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. , doi:10.1016/S0031-8914(43)90575-0

Dobkins, K. R., Gunther, K. L., & Peterzell, D. H. (2000). What Covariance Mechanisms Underlie Green/Red Equiluminance, Luminance Contrast Sensitivity and Chromatic (Green/Red) Contrast Sensitivity? *Vision Research*, *40*(6), 613–628. , doi:10.1016/S0042-6989(99)00211-4

Donner, K., & 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. , doi:10.1017/S0952523800007215

Duinkharjav, B., Chen, K., Tyagi, A., He, J., Zhu, Y., & Sun, Q. (2022). Color-Perception-Guided Display Power Reduction for Virtual Reality. *ACM Transactions on Graphics (TOG)*, *41*(6), 1–16.

Eskew, R., McLellan, J. S., & Giulianini, F. (1999). Chromatic detection and discrimination. In K. Gegenfurtner & L. Sharpe (Eds.), *Color vision: From molecular genetics to perception* (pp. 343–366). Cambridge, UK: Cambridge University Press.

Ferry, E. S. (1892). Persistence of Vision. *American Journal of Science (1880-1910)*, *44*(261), 192–207. , doi:10.2475/ajs.s3-44.261.192

García-Pérez, M. A., & Peli, E. (1997, September). The Transition from DeVries-Rose to Weber's Laws: Comments on Rovamo, Mustonen and Näsänen (1995). *Vision Research*, *37*(18), 2573–2576. , doi:10.1016/S0042-6989(97)00071-0

Gegenfurtner, K. R., & Hawken, M. J. (1996). Interaction of Motion and Color in the Visual Pathways. *Trends in neurosciences*, *19*(9), 394–401. , doi:10.1016/S0166-2236(96)10036-9

Hansen, T., Pracejus, L., & Gegenfurtner, K. R. (2009). Color Perception in the Intermediate Periphery of the Visual Field. *Journal of vision*, *9*(4), 26–26. , doi:10.1167/9.4.26

Haun, A. M., & Peli, E. (2013). Is Image Quality a Function of Contrast Perception? In *Human vision and electronic imaging XVIII* (Vol. 8651, pp. 55–59). SPIE.

Hecht, S., & Shlaer, S. (1936). Intermittent Stimulation by Light: V. The Relation between Intensity and Critical Frequency for Different Parts of the Spectrum. *The Journal of general physiology*, *19*(6), 965–977.

Hecht, S., & Smith, E. L. (1936). Intermittent Stimulation by Light: VI. Area and the Relation between Critical Frequency and Intensity. *The Journal of general physiology*, *19*(6), 979–989.

Hecht, S., & Verrijp, C. D. (1933). Intermittent Stimulation by Light: III. The Relation between Intensity and Critical Fusion Frequency for Different Retinal Locations. *The Journal of general physiology*, *17*(2), 251–268.

Hernández, C., Doménech, B., Seguí, M. M., & Illueca, C. (1996, July). The Effect of Pupil and Observation Distance on the Contrast Sensitivity Function. *Ophthalmic and Physiological Optics*, *16*(4), 336–341. , doi:10.1046/j.1475-1313.1996.95001166.x

Hess, R., & Snowden, R. (1992). Temporal Properties of Human Visual Filters: Number, Shapes and Spatial Covariation. *Vision Research*, *32*(1), 47–59. , doi:10.1016/0042-6989(92)90112-V

Hood, D. C., & Birch, D. G. (1990). A Quantitative Measure of the Electrical Activity of Human Rod Photoreceptors Using Electroretinography. *Visual neuroscience*, *5*(4), 379–387. , doi:10.1017/S0952523800000468

Hood, D. C., & Birch, D. G. (1993). Human Cone Receptor Activity: The Leading Edge of the a–Wave and Models of Receptor Activity. *Visual neuroscience*, *10*(5), 857–871. ,

doi:10.1017/S0952523800006076

Horiguchi, H., Winawer, J., Dougherty, R. F., & Wandell, B. A. (2013). Human Trichromacy Revisited. *Proceedings of the National Academy of Sciences*, *110*(3), E260–E269. , doi:10.1073/pnas.1214240110

Hubel, D. H., & Wiesel, T. N. (1962). Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex. *The Journal of physiology*, *160*(1), 106.

ITU. (2017). *Image Parameter Values for High Dynamic Range Television for Use in Production and International Programme Exchange* (Tech. Rep.). ITU-R BT.2100-1.

Jindal, A., Wolski, K., Myszkowski, K., & Mantiuk, R. K. (2021). Perceptual Model for Adaptive Local Shading and Refresh Rate. *ACM Transactions on Graphics*, *40*(6). , doi:10.1145/3478513.3480514

Johnson, C. A., Keltner, J. L., & Balestrery, F. (1978). Effects of Target Size and Eccentricity on Visual Detection and Resolution. *Vision research*, *18*(9), 1217–1222. , doi:10.1016/0042-6989(78)90106-2

Kaplanyan, A. S., Sochenov, A., Leimkühler, T., Okunev, M., Goodall, T., & Rufo, G. (2019). DeepFovea: Neural Reconstruction for Foveated Rendering and Video Compression Using Learned Statistics of Natural Videos. *ACM Transactions on Graphics (TOG)*, *38*(6), 1–13. , doi:10.1145/3355089.3356557

Kelly, D. H. (1979). Motion and Vision. I. Stabilized Images of Stationary Gratings. *Journal of the Optical Society of America*, *69*(9), 1266–1274. , doi:10.1364/JOSA.69.001266

Kelly, D. H. (1983). Spatiotemporal Variation of Chromatic and Achromatic Contrast Thresholds. *Journal of the Optical Society of America*, *73*(6), 742–750. , doi:10.1364/JOSA.73.000742

Kim, K. J., Mantiuk, R. K., & Lee, K. H. (2013). Measurements of Achromatic and Chromatic Contrast Sensitivity Functions for an Extended Range of Adaptation Luminance. In *Human vision and electronic imaging XVIII* (Vol. 8651, pp. 319–332). SPIE.

Kim, Y. J., Reynaud, A., Hess, R. F., & Mullen, K. T. (2017). A Normative Data Set for the Clinical Assessment of Achromatic and Chromatic Contrast Sensitivity Using a qCSF Approach. *Investigative ophthalmology & visual science*, *58*(9), 3628–3636. , doi:10.1167/iovs.17-21645

King-Smith, P. E., & Kulikowski, J. J. (1975). Pattern and Flicker Detection Analysed by Subthreshold Summation. *The Journal of physiology*, *249*(3), 519–548. , doi:10.1113/jphysiol.1975.sp011028

Kong, X., Pérez, M. R. B., Vogels, I. M., Sekulovski, D., & Heynderickx, I. (2018). Modelling Contrast Sensitivity for Chromatic Temporal Modulations. In *Color and imaging conference* (Vol. 2018, pp. 324–329). Society for Imaging Science and Technology.

Krajancich, B., Kellnhofer, P., & Wetzstein, G. (2023). Towards Attention-Aware Rendering for Virtual and Augmented Reality. *arXiv preprint arXiv:2302.01368*.

Kulikowski, J. J. (1976). Effective Contrast Constancy and Linearity of Contrast Sensation. *Vision Research*, *16*(12), 1419–1431. , doi:10.1016/0042-6989(76)90161-9

Kulikowski, J. J., & Robson, AG. (1999). Spatial, Temporal and Chromatic Channels: Electrophysiological Foundations. *Journal of Optical Technology*, *66*(9), 797. , doi:10.1364/JOT.66.000797

Laird, J., Rosen, M., Pelz, J., Montag, E., & Daly, S. J. (2006). Spatio-Velocity CSF as a Function of Retinal Velocity Using Unstabilized Stimuli. In *SPIE 6057, human vision and electronic imaging XI.*

Legge, G. E. (1984). Binocular contrast summation — II. quadratic summation. *Vision Research*, *24*(4), 385–394. , doi:10.1016/0042-6989(84)90064-6

Levi, D. M., & Harwerth, R. S. (1982). Psychophysical Mechanisms in Humans with Amblyopia. *American journal of optometry and physiological optics*, *59*(12), 936–951. , doi:10.1097/00006324-198212000-00002

Levi, D. M., Klein, S. A., & Aitsebaomo, AP. (1985). Vernier Acuity, Crowding and Cortical Magnification. *Vision research*, *25*(7), 963–977.

Li, Q., Gomez-Villa, A., Bertalmío, M., & Malo, J. (2022). Contrast Sensitivity Functions in Autoencoders. *Journal of Vision*, *22*(6), 8–8.

Lucassen, M., Lambooij, M., Sekulovski, D., & Vogels, I. (2018). Spatio-Chromatic Sensitivity Explained by Post-Receptoral Contrast. *Journal of Vision*, *18*(5), 13–13. , doi:10.1167/18.5.13

Luebke, D., & Hallen, B. (2001). Perceptually Driven Simplification for Interactive Rendering. In *Rendering techniques 2001: Proceedings of the eurographics workshop* (p. 223).

Mandler, M. B., & Makous, W. (1984). A Three Channel Model of Temporal Frequency Perception. *Vision research*, *24*(12), 1881–1887. , doi:10.1016/0042-6989(84)90021-X

Mantiuk, R. K., Ashraf, M., & Chapiro, A. (2022). StelaCSF: A Unified Model of Contrast Sensitivity as the Function of Spatio-Temporal Frequency, Eccentricity, Luminance and Area. *ACM Trans. Graph.*, *41*(4). , doi:10.1145/3528223.3530115

Mantiuk, R. K., Daly, S. J., & Kerofsky, L. (2008). Display Adaptive Tone Mapping. *ACM Transactions on Graphics (Proc. of SIGGRAPH)*, *27*(3), 68. , doi:10.1145/1360612.1360667

Mantiuk, R. K., Kim, K. J., Rempel, A. G., & Heidrich, W. (2011). HDR-VDP-2: A Calibrated Visual Metric for Visibility and Quality Predictions in All Luminance Conditions. *ACM Transactions on Graphics*, *30*(4), 1–14. , doi:10.1145/2010324.1964935

Mantiuk, R. K., Kim, M., Ashraf, M., Xu, Q., Luo, M. R., Martinovic, J., et al. (2020). Practical Color Contrast Sensitivity Functions for Luminance Levels up to 10000 Cd/M2. In *Color and imaging conference* (Vol. 2020, pp. 1–6). Society for Imaging Science and Technology.

Mantiuk, R. K., Krawczyk, G., Myszkowski, K., & Seidel, H.-P. (2004). Perception-Motivated High Dynamic Range Video Encoding. *ACM Transactions on Graphics (Proc. of SIGGRAPH)*, *23*(3), 733. , doi:10.1145/1015706.1015794

Metha, A. B., & Mullen, K. T. (1996). Temporal Mechanisms Underlying Flicker Detection and Identification for Red–Green and Achromatic Stimuli. *JOSA A*, *13*(10), 1969–1980. , doi:10.1364/JOSAA.13.001969

Miller, S., Nezamabadi, M., & Daly, S. J. (2013). Perceptual Signal Coding for More Efficient Usage of Bit Codes. *SMPTE Motion Imaging Journal*, *122*(4), 52–59. , doi:10.5594/j18290

Moreland, JD., & Cruz, A. (1959). Colour Perception with the Peripheral Retina. *Optica Acta: International Journal of Optics*, *6*(2), 117–151. , doi:10.1080/713826278

Mullen, K. T. (1985). The Contrast Sensitivity of Human Colour Vision to Red-Green and Blue-Yellow Chromatic Gratings. *The Journal of physiology*, *359*, 381–400. , doi:10.1113/jphysiol.1985.sp015591

Mullen, K. T., et al. (2002). Differential distributions of red–green and blue–yellow cone opponency across the

visual field. *Visual neuroscience*, *19*(1), 109–118. , doi:10.1017/S0952523802191103

Mullen, K. T., Sakurai, M., & Chu, W. (2005). Does L/M Cone Opponency Disappear in Human Periphery? *Perception*, *34*(8), 951–959. , doi:10.1068/p5374

Newton, J. R., & Eskew, R. T. (2003). Chromatic Detection and Discrimination in the Periphery: A Postreceptoral Loss of Color Sensitivity. *Visual neuroscience*, *20*(5), 511–521. , doi:10.1017/S0952523803205058

Noorlander, C., Koenderink, J. J., Den Olden, R. J., & Edens, B. W. (1983). Sensitivity to Spatiotemporal Colour Contrast in the Peripheral Visual Field. *Vision research*, *23*(1), 1–11. , doi:10.1016/0042-6989(83)90035-4

Norcia, A. M., Tyler, C. W., & Hamer, R. D. (1990). Development of Contrast Sensitivity in the Human Infant. *Vision research*, *30*(10), 1475–1486. , doi:10.1016/0042-6989(90)90028-J

Ortiz, C., Jiménez, J. R., Pérez-Ocón, F., Castro, J. J., & González-Anera, R. (2010). Retinal-Image Quality and Contrast-Sensitivity Function in Age-Related Macular Degeneration. *Current eye research*, *35*(8), 757–761. , doi:10.3109/02713683.2010.486521

Owsley, C., & Sloane, M. E. (1987). Contrast Sensitivity, Acuity, and the Perception of 'real-World' Targets. *British Journal of Ophthalmology*, *71*(10), 791–796. , doi:10.1136/bjo.71.10.791

Peli, E. (1995). Suprathreshold Contrast Perception across Differences in Mean Luminance: Effects of Stimulus Size, Dichoptic Presentation, and Length of Adaptation. *Journal of the Optical Society of America A*, *12*(5), 817. , doi:10.1364/JOSAA.12.000817

Perez-Ortiz, M., Mikhailiuk, A., Zerman, E., Hulusic, V., Valenzise, G., & Mantiuk, R. K. (2019). From Pairwise Comparisons and Rating to a Unified Quality Scale. *IEEE Transactions on Image Processing*, *29*, 1139–1151. , doi:10.1109/TIP.2019.2936103

Pointer, JS., & Hess, RF. (1989). The Contrast Sensitivity Gradient across the Human Visual Field: With Emphasis on the Low Spatial Frequency Range. *Vision research*, *29*(9), 1133–1151. , doi:10.1016/0042-6989(89)90061-8

Pokorny, J., Smith, V. C., & Lutze, M. (1987). Aging of the Human Lens. *Applied Optics*, *26*(8), 1437–1440. , doi:10.1364/ao.26.001437

Porter, T. C. (1902). Contributions to the Study of Flicker. Paper II. *Proceedings of the Royal Society of London*, *70*(459-466), 313–329. , doi:10.1098/rspl.1902.0032

Riccó, Annibale. (1877). Relazione fra il minimo angolo visuale e l'intensitá luminosa. *Memorie della Societa Degli Spettroscopisti Italiani, vol. 6, pp. B29-B58*, *6*, B29–B58.

Robson, J. G. (1966). Spatial and Temporal Contrast-Sensitivity Functions of the Visual System. *Journal of Optical Society of America*, *56*(8), 1141–1142. , doi:10.1364/JOSA.56.001141

Robson, J. G., & Graham, N. (1981). Probability Summation and Regional Variation in Contrast Sensitivity across the Visual Field. *Vision research*, *21*(3), 409–418. , doi:10.1016/0042-6989(81)90169-3

Rohatgi, A. (2022). *Webplotdigitizer: Version 4.6.*

Rose, A. (1948). The Sensitivity Performance of the Human Eye on an Absolute Scale. *JOSA*, *38*(2), 196–208. , doi:10.1364/JOSA.38.000196

Rovamo, J., Kankaanpää, M. I., & Kukkonen, H. (1999). Modelling Spatial Contrast Sensitivity Functions for Chromatic and Luminance-Modulated Gratings. *Vision Research*, *39*(14), 2387–2398. , doi:10.1016/S0042-6989(98)00273-9

Rovamo, J., Luntinen, O., & Näsänen, R. (1993). Modelling the Dependence of Contrast Sensitivity on Grating Area and Spatial Frequency. *Vision Research*, *33*(18), 2773–2788. , doi:10.1016/0042-6989(93)90235-O

Rovamo, J., Mustonen, J., & Näsänen, R. (1995). Neural Modulation Transfer Function of the Human Visual System at Various Eccentricities. *Vision Research*, *35*(6), 767–774. , doi:10.1016/0042-6989(94)00171-H

Rovamo, J., Näsänen, R., & Mustonen, J. (1997, September). Transition from DeVries-Rose to Weber's Law: Reply to García-Pérez and Peli (1997). *Vision Research*, *37*(18), 2576–2578. , doi:10.1016/S0042-6989(97)00072-2

Sakurai, M., & Mullen, K. T. (2006). Cone Weights for the Two Cone-Opponent Systems in Peripheral Vision and Asymmetries of Cone Contrast Sensitivity. *Vision research*, *46*(26), 4346–4354. , doi:10.1016/j.visres.2006.08.016

Sharpe, L. T., Stockman, A., Jagla, W., & Jägle, H. (2005). A Luminous Efficiency Function, V*(λ), for Daylight Adaptation. *Journal of Vision*, *5*(11), 3–3. , doi:10.1167/5.11.3

SID. (2022). *Information Display Measurement Standard* (1.1a

ed.; I. C. f. D. M. ICDM, Ed.). Society for Information Display.

Snowden, R. J., Hess, R. F., & Waugh, S. J. (1995). The Processing of Temporal Modulation at Different Levels of Retinal Illuminance. *Vision Research*, *35*(6), 775–789. , doi:10.1016/0042-6989(94)00158-I

Stabell, B., & Stabell, ULF. (1996). Peripheral Colour Vision: Effects of Rod Intrusion at Different Eccentricities. *Vision Research*, *36*(21), 3407–3414. , doi:10.1016/0042-6989(96)00079-X

Swanson, W. H., Ueno, T., Smith, V. C., & Pokorny, J. (1987). Temporal Modulation Sensitivity and Pulse-Detection Thresholds for Chromatic and Luminance Perturbations. *JOSA A*, *4*(10), 1992–2005. , doi:10.1364/JOSAA.4.001992

Switkes, E., & Crognale, M. A. (1999). Comparison of Color and Luminance Contrast: Apples versus Oranges? *Vision research*, *39*(10), 1823–1831. , doi:10.1016/S0042-6989(98)00219-3

Tariq, T., Matsuda, N., Penner, E., Jia, J., Lanman, D., Ninan, A., et al. (2023). Perceptually Adaptive Real-Time Tone Mapping. In *SIGGRAPH asia 2023 conference papers.*

Thibos, L. N., Still, D. L., & Bradley, A. (1996). Characterization of Spatial Aliasing and Contrast Sensitivity in Peripheral Vision. *Vision research*, *36*(2), 249–258. , doi:10.1016/0042-6989(95)00109-D

Tolhurst, DJ. (1973). Separate Channels for the Analysis of the Shape and the Movement of a Moving Visual Stimulus. *The Journal of Physiology*, *231*(3), 385. , doi:10.1113/jphysiol.1973.sp010239

Tursun, O. T., Arabadzhiyska-Koleva, E., Wernikowski, M., Mantiuk, R., Seidel, H.-P., Myszkowski, K., et al. (2019). Luminance-Contrast-Aware Foveated Rendering. *ACM Transactions on Graphics (TOG)*, *38*(4), 1–14. , doi:10.1145/3306346.3322985

Van der Horst, G. J., & Bouman, M. A. (1969). Spatiotemporal Chromaticity Discrimination. *JOSA*, *59*(11), 1482–1488. , doi:10.1364/JOSA.59.001482

Vangorp, P., Myszkowski, K., Graf, E. W., & Mantiuk, R. K. (2015, October). A Model of Local Adaptation. *ACM Transactions on Graphics*, *34*(6), 1–13. , doi:10.1145/2816795.2818086

Van Hateren, J. H. (1993). Spatiotemporal Contrast Sensitivity of Early Vision. *Vision research*, *33*(2), 257–267. ,

doi:10.1016/0042-6989(93)90163-Q

Van Meeteren, A. (1974). Calculations on the Optical Modulation Transfer Function of the Human Eye for White Light. *Optica Acta: International Journal of Optics*, *21*(5), 395–412. , doi:10.1080/713818902

Verbaken, JH., & Johnston, AW. (1986). Population Norms for Edge Contrast Sensitivity. *American Journal of Optometry and Physiological Optics*, *63*(9), 724–732. , doi:10.1097/00006324-198609000-00005

Virsu, V., & Rovamo, J. (1979). Visual Resolution, Contrast Sensitivity, and the Cortical Magnification Factor. *Experimental Brain Research*, *37*(3), 475–494. , doi:10.1007/BF00236818

Virsu, V., Rovamo, J., & Laurinen, P. (1982). Temporal Contrast Sensitivity and Cortical Magnification. *Vision Research*, *22*, 1211–1217. , doi:10.1016/0042-6989(82)90087-6

Volbrecht, V. J., Shrago, E. E., Schefrin, B. E., & Werner, J. S. (2000). Spatial Summation in Human Cone Mechanisms from 0 to 20 in the Superior Retina. *JOSA A*, *17*(3), 641–650. , doi:10.1364/josaa.17.000641

Wagner, G., & Boynton, R. M. (1972, December). Comparison of Four Methods of Heterochromatic Photometry. *Journal of the Optical Society of America*, *62*(12), 1508. , doi:10.1364/JOSA.62.001508

Ward-Larson, G., Rushmeier, H., & Piatko, C. (1997). A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes. *IEEE Transactions on Visualization and Computer Graphics*, *3*(4), 291–306. , doi:10.1109/2945.646233

Watson, A. B. (2000). Visual Detection of Spatial Contrast Patterns: Evaluation of Five Simple Models. *Optics Express*, *6*(1), 12–33. , doi:10.1364/OE.6.000012

Watson, A. B. (2013). A Formula for the Mean Human Optical Modulation Transfer Function as a Function of Pupil Size. *Journal of Vision*, *13*(6), 18–18. , doi:10.1167/13.6.18

Watson, A. B. (2018). The Field of View, the Field of Resolution, and the Field of Contrast Sensitivity. *Journal of Perceptual Imaging*, *1*(1), 1–11. , doi:10.2352/J.Percept.Imaging.2018.1.1.010505

Watson, A. B. (2021). *Stanford Talks: The Chromatic Pyramid of Visibility.*

Watson, A. B., & Ahumada, A. J. (2016). The Pyramid of Visibil-

ity. *Human Vision and Electronic Imaging 2016, HVEI 2016*, 37–42. , doi:10.2352/ISSN.2470-1173.2016.16HVEI-102

Watson, A. B., & Yellott, J. I. (2012). A Unified Formula for Light-Adapted Pupil Size. *Journal of vision*, *12*(10), 12. , doi:10.1167/12.10.12

Weale, RA. (1953). Spectral Sensitivity and Wave-Length Discrimination of the Peripheral Retina. *The Journal of Physiology*, *119*(2-3), 170. , doi:10.1113/jphysiol.1953.sp004837

Weber, E. H. (1831). *De Pulsu, Resorptione, Auditu et Tactu: Annotationes Anatomicae et Physiologicae*. Koehler.

Wilson, ME. (1970). Invariant Features of Spatial Summation with Changing Locus in the Visual Field. *The Journal of physiology*, *207*(3), 611–622. , doi:10.1113/jphysiol.1970.sp009083

Wright, M. J., & Johnston, A. (1983). Spatiotemporal Contrast Sensitivity and Visual Field Locus. *Vision Research*, *23*(10), 983–989. , doi:10.1016/0042-6989(83)90008-1

Wuerger, S., Ashraf, M., Kim, M., Martinovic, J., Pérez-Ortiz, M., & Mantiuk, R. K. (2020). Spatio-Chromatic Contrast Sensitivity under Mesopic and Photopic Light Levels. *Journal of Vision*, *20*(4), 23. , doi:10.1167/jov.20.4.23

Wuerger, S., Watson, A. B., & Ahumada Jr, A. J. (2002). Towards a Spatio-Chromatic Standard Observer for Detection. In *Human vision and electronic imaging VII* (Vol. 4662, pp. 159–172). SPIE.

Xin, D., & Bloomfield, S. A. (1999). Dark-and Light-Induced Changes in Coupling between Horizontal Cells in Mammalian Retina. *Journal of Comparative Neurology*, *405*(1), 75–87. , doi:10.1002/(SICI)1096-9861(19990301)405:1¡75::AID-CNE6¿3.0.CO;2-D

Xu, Q., Luo, M. R., & Sekulovski, D. (2020). Investigation of Spatial Chromatic Contrast around 5 Colour Centres. In *London imaging meeting* (Vol. 2020, pp. 1–4). Society for Imaging Science and Technology.

Xu, Q., Ye, Q., Mantiuk, R. K., & Luo, M. R. (2022). A Study of Spatial Chromatic Contrast Sensitivity Based on Different Colour Background. In *Color and imaging conference* (Vol. 30, pp. 236–240). Society for Imaging Science and Technology.

Yi, S., Jeon, D. S., Serrano, A., Jeong, S.-Y., Kim, H.-Y., Gutierrez, D., et al. (2022, January). Modelling Surround-aware Contrast Sensitivity for HDR Displays. *Computer Graphics Forum*. , doi:10.1111/cgf.14439

Zeng, W., Daly, S. J., & Lei, S. (2002). An Overview of the Visual Optimization Tools in JPEG 2000. *Signal Processing: Image Communication*, *17*(1), 85–104. , doi:10.1016/S0923-5965(01)00029-7

Zhang, X., & Wandell, B. A. (1997). A Spatial Extension of CIELAB for Digital Color-Image Reproduction. *Journal of the Society for Information Display*, *5*(1), 61. , doi:10.1889/1.1985127

Appendix: fitted parameters

Tables 5–7 contain the fitted parameters of castleCSF.

Table 5: The fitted parameters for S_{Ach}.

Part	Parameters
Size and spatial freq.	$k_{a,\mathcal{S}}^{Ach} = 0.1002$, $k_{b,\mathcal{S}}^{Ach} = 0.000213$, $a_{0,\mathcal{S}}^{Ach} = 157.1$, $\rho_{0,\mathcal{S}}^{Ach} = 0.7023$, $k_{a,\mathcal{T}}^{Ach} = 0.0002412$, $k_{b,\mathcal{T}}^{Ach} = 2.676$, $a_{0,\mathcal{T}}^{Ach} = 3.816$, $\rho_{0,\mathcal{T}}^{Ach} = 3.014$,
Temporal channels	$\beta_{\mathcal{S}}^{Ach} = 1.331$, $\sigma_{\mathcal{S}}^{Ach} = 10.58$, $\beta_{\mathcal{T}}^{Ach} = 0.1898$, $\sigma_{\mathcal{T}}^{Ach} = 0.08448$, $m_{\omega}^{Ach} = 2.415$, $c_{\omega}^{Ach} = 4.704$,
Luminance	$k_{s1,\mathcal{S}}^{Ach} = 56.49$, $k_{s2,\mathcal{S}}^{Ach} = 7.547$, $k_{s3,\mathcal{S}}^{Ach} = 0.1445$, $k_{s4,\mathcal{S}}^{Ach} = 5.583e-7$, $k_{s5,\mathcal{S}}^{Ach} = 9.669e9$, $k_{\rho1,\mathcal{S}}^{Ach} = 1.781$, $k_{\rho2,\mathcal{S}}^{Ach} = 91.57$, $k_{\rho3,\mathcal{S}}^{Ach} = 0.2567$, $k_{s1,\mathcal{T}}^{Ach} = 0.1934$, $k_{s2,\mathcal{T}}^{Ach} = 2748$, $k_{\rho,\mathcal{T}}^{Ach} = 0.0003167$,
Eccentricity	$k_{e1}^{Ach} = 0.0189$, $k_{e2}^{Ach} = 0.02399$, $k_{e1,nasal}^{Ach} = 0.008136$, $k_{e2,nasal}^{Ach} = 0.04007$

Table 6: The fitted parameters for S_{RG}.

Part	Parameters
Size and spatial freq.	$k_{b,\mathcal{S}}^{RG} = 2.421$, $a_{0,\mathcal{S}}^{RG} = 2.816442e3$, $\rho_{0,\mathcal{S}}^{RG} = 0.0711058$,
Temporal channels	$\beta_{\mathcal{S}}^{RG} = 1.156$, $\sigma_{\mathcal{S}}^{RG} = 16.43$,
Luminance	$k_{s1,\mathcal{S}}^{RG} = 681.4$, $k_{s2,\mathcal{S}}^{RG} = 38$, $k_{s3,\mathcal{S}}^{RG} = 0.4804$, $k_{\rho,\mathcal{S}}^{RG} = 0.01784$,
Eccentricity	$k_{e1}^{RG} = 2.05e-69$, $k_{e2}^{RG} = 0.05914$, $k_{e1,nasal}^{RG} = 0.1811$, $k_{e2,nasal}^{RG} = 2.896e-5$

Table 7: The fitted parameters for S_{YV}.

Part	Parameters
Size and spatial freq.	$k_{b,\mathcal{S}}^{YV} = 2.682$, $a_{0,\mathcal{S}}^{YV} = 2.827890e7$, $\rho_{0,\mathcal{S}}^{YV} = 0.000635093$,
Temporal channels	$\beta_{\mathcal{S}}^{YV} = 0.9691$, $\sigma_{\mathcal{S}}^{YV} = 7.15$,
Luminance	$k_{s1,\mathcal{S}}^{YV} = 166.7$, $k_{s2,\mathcal{S}}^{YV} = 62.9$, $k_{s3,\mathcal{S}}^{YV} = 0.4119$, $k_{\rho,\mathcal{S}}^{YV} = 0.004258$,
Eccentricity	$k_{e1}^{YV} = 0.008066$, $k_{e2}^{YV} = 0.003569$, $k_{e1,\text{nasal}}^{YV} = 0.01107$, $k_{e2,\text{nasal}}^{YV} = 5.858e - 141$