https://doi.org/10.2352/EI.2024.36.11.HVEI-209

Spatiotemporal contrast sensitivity functions: predictions for the critical flicker frequency

Ali Bozorgian, Maliha Ashraf, Rafał K. Mantiuk

Abstract

Contrast sensitivity functions (CSFs), which provide estimations of detection thresholds, have far-reaching applications in digital media processing, rendering, and transmission. There is a practical interest in obtaining accurate estimations of spatial and temporal resolution limits from a spatiotemporal CSF model. However, current spatiotemporal CSFs are inaccurate when predicting high-frequency limits such as critical flicker frequency (CFF). To address this problem, we modified two spatiotemporal CSFs, namely Barten's CSF and stelaCSF, to better account for the contrast sensitivity at high temporal frequencies, both in the fovea and eccentricity. We trained these models using 15 datasets of spatial and temporal contrast sensitivity measurements from the literature. Our modifications account for two features observed in CFF measurement: the increase of CFF at medium eccentricities (of about 15 deg), and the saturation of CFF at high luminance values. As a result, the prediction errors for CFF obtained from the modified models improved remarkably.

Introduction

It has been demonstrated that spatiotemporal CSF models fall short of providing accurate estimations for the psychophysical measurements of CFF [1]. Our main objective in this study is to improve the accuracy of CFF estimations derived from spatiotemporal CSF models. We propose several modifications to two spatiotemporal contrast sensitivity models — Barten's CSF and stelaCSF — and assess their effectiveness through comparison with experimental measurements of CFF.

The visual system's performance in detecting light intensity variations is restricted by the quantum nature of light and neural noise [2, 3]. Studies in psychophysics have shown that specific characteristics of a stimulus, such as spatial frequency, temporal frequency, luminance, area, and visual eccentricity, can significantly influence the likelihood of detection. A CSF model can be used to predict the minimum level of contrast required for detection, defined, for example, as the contrast at which there is a 75 percent chance of detection in a two-alternative forced-choice experiment.

The measurements of temporal contrast sensitivity function (TCSF) for flickering discs and spatial contrast sensitivity function (SCSF) for static Gabor patterns show a bandpass profile in photopic conditions, implying that the visual system has a suppressed response to slow changes in light but an enhanced response to transient changes. Suppression in lower frequencies is believed to be an evolutionary accommodation to the regularities found in the natural scenes to use the limited dynamic range of visual neurons efficiently [4, 5]. The cutoff frequency of a TCSF curve represents the highest frequency at which a temporally modulated stimulus will cause a flicker sensation. Beyond this point, the flicker fuses to a steady field. This cutoff point is called CFF (see Figure 1). Similarly, the cutoff frequency of an SCSF curve represents the highest frequency at which a spatially modulated stimulus would appear as a pattern. Beyond this limit, the pattern fades into a uniform field. This point is often referred to as acuity and is constrained by the optics of the eye in fovea [6].

There has been a long-standing interest in CFF as an experimental measure of temporal resolution [7, 8, 9]. It has been shown that CFF increases with the logarithm of time-averaged light intensity of the background [10, 11], yielding a straight line in a semi-logarithmic plot. This relationship is called the Ferry-Porter law and holds for a substantial range of light intensities. Additionally, this relationship holds (with varying slopes) for a range of stimulus parameters, including a range of eccentricities, areas, and wavelengths [12, 13]. Depending on the flickering stimuli's spatial configuration, the CFF variation in the peripheral vision may not be monotonic [14, 15, 16, 17, 18]. While spatial acuity tends to decrease with greater eccentricity, CFF measurements reveal an initial rise in higher eccentricities followed by a subsequent decline well below foveal measurements. It has been shown [1] that existing contrast sensitivity models fail to account for this non-monotonic trend, which might have implications for applications such as foveated rendering that may require accurate temporal resolution predictions. Consequently, we argue that special attention to this issue is needed. A mathematical model

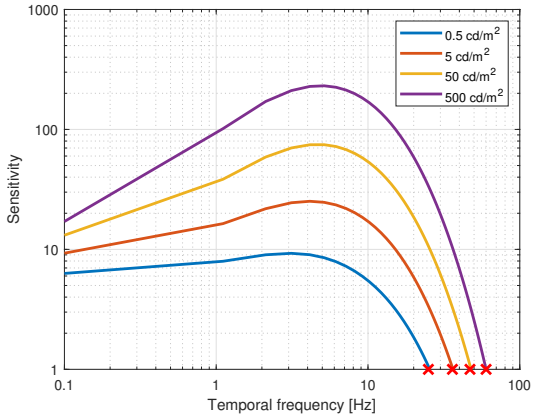


Figure 1: TCSF curves obtained from stelaCSF [1] for a range of luminance values. The red crosses on the curves indicate the temporal cutoff frequency — the point where the TCSF curves intersect with a horizontal line of sensitivity value 1. This value is known as CFF or temporal resolution.

that can predict contrast sensitivity measurements has numerous applications in image/video processing and computer graphics algorithms [19, 20, 21, 22]. There is a practical interest in CFF and acuity predictions, given that they establish a conservative gamut of visible signals and can aid in developing hardware and software technologies tuned toward human perception [23].

Non-monotonic variation of CFF

There are different views on why CFF varies at different retinal locations. Tyler [24] measured CFF as a function of eccentricity for stimuli of varying sizes and observed an increase in CFF values with eccentricity. The stimuli were scaled at each eccentricity so that the number of ganglion cells involved remained fixed (M-scaled, see [25]). Tyler assumed that the temporal parameters of the human visual response are related to anatomical features: the density of ganglion cells and dimensions of the cone outer segment. The former was believed to determine the peak temporal sensitivity, while the latter was related to the response-time constant of the temporal impulse response that characterizes the visual system. The response-time constant was estimated to decrease from 70 milliseconds in the fovea to 35 milliseconds in the mid-periphery.

An alternative view by Rider et al. [13] suggests that CFF measurements do not provide much information about the integration time or speed of temporal processing. Instead, the visual system's gain variation regulates the increase of CFF in mid to elevated light intensities. Gain, in this context, is a scaling value that changes the overall response of the visual system regardless of the temporal frequency. In this view, the variation of CFF as a function of eccentricity might depend on the changes in gain rather than integration time. These interpretations stem from a light adaptation model proposed earlier by the same author [26], comprised of a cascade of leaky integrators with feedforward loops.

Rovamo et al. [18] conducted a study to measure CFF and its variation with eccentricity under different stimulus settings. The study showed that the CFF increased monotonically at higher eccentricities when the stimulus size was M-scaled. Additionally, decreasing the luminance of the stimulus in higher eccentricities, inversely proportional to the area of the center of ganglion cell receptive fields, led to a monotonic decrease in CFF. When both the stimulus size and luminance were scaled, the CFF remained constant from the fovea toward the periphery. Consequently, it was concluded that two counteracting factors determine the non-monotonic change of CFF for fixed-sized and fixed-luminance stimuli: the decrease of ganglion cell density towards the periphery reduces CFF, while the increase of receptive field and summation area enhances CFF. The intuitions behind these measurements came from the observation made by Enroth-Cougell and Shapley that CFF of ganglion cells correlates with flux defined as the product of illuminance and receptive field center area [27, 28].

Spatiotemporal models

Barten proposed a contrast sensitivity model that estimates detection thresholds by considering the contrast energy of stimulus and internal noise [29, 2]. The most well-known version of his model accounts for the effects of spatial frequency, size, and luminance [30]. He also described peripheral and temporal vision extensions by adding eccentricity and flicker frequency parameters to the model [31]. The extended version of the model assumes that a contrast signal undergoes spatial and temporal filtering as it travels through the visual pathway and then causes a subtractive inhibition that faces similar spatial and temporal filtering. To capture the spatio-temporal interaction, Barten [32] suggested the following formulation in the frequency domain for monocular vision:

$$S(\rho,\omega,L,a,e) = \frac{S_0(\rho,L,e)\,T_0(\omega,L)\big(1-S_1(\rho)T_1(\omega,L)\big)}{2k\,m_n(X,Y,T,e,\rho)}, \quad (1)$$

where ρ is the spatial frequency in cycles per visual degree (cpd), ω is the temporal frequency in cycles per second, L is luminance in candela per square meter, e is eccentricity in visual degrees, $S_0(\rho,L,e)$, $S_1(\rho)$, $T_0(\omega,L)$, and $T_1(\omega,L)$ are spatial filter for photoreceptor response, spatial filter for inhibitory response, temporal filter for photoreceptor response, and temporal filter for inhibitory response, respectively. The constant k is signal to noise ratio, and $m_n(X,Y,T,e,\rho)$ is the average modulation of internal noise, which is described by:

$$m_n(X,Y,T,e,\rho) = \sqrt{\frac{\Phi_{ph}(L,e)(1-S_1(\rho))^2+\Phi_0(e)}{XYT}}, \quad (2)$$

where $\Phi_{ph}(L,e)$ is the spectral density of photon noise, $\Phi_0(e)$ is the spectral density of the neural noise, X, Y, and T are the spatial and temporal dimensions of the stimulus covered with noise. The amplitude response of spatial filters is based on Gaussian functions:

$$S_0(\rho,l,e) = \exp\left(-\frac{\rho^2}{\rho_{c_0}(L,e)^2}\right), \quad (3)$$

and:

$$S_1(\rho) = 1-\sqrt{1-\exp\left(-\frac{\rho^2}{\rho_{c_1}^2}\right)}. \quad (4)$$

where $\rho_{c_0}(L,e)$ and ρ_{c_1} are photoreceptor and inhibitory spatial filter cutoff frequencies. The normalized amplitude response of temporal filters is described by:

$$T_j(\omega,l) = \frac{1}{\big(1+(2\pi\omega\tau_j(L,D))^2\big)^{\frac{n_j}{2}}}, \quad j=[0,1], \quad (5)$$

where τ_j is the time constant of the filter, and n is the number of cascaded processing units.

Recently, Mantiuk et al. proposed stelaCSF [1], a unified CSF that accounts for all major dimensions of a stimulus: spatial and temporal frequency, luminance, area, and eccentricity. They combined data from 11 datasets to model the 5-dimensional space of contrast sensitivity. In the stelaCSF, it is assumed that two mechanisms (channels) are responsible for temporal processing: one channel encodes low temporal frequencies (sustained, $R_s(\omega)$), and the other one encodes high temporal frequencies (transient, $R_\tau(\omega)$). For simplicity, instead of cascaded leaky integrators, generalized exponential functions are used to describe the amplitude response of temporal channels:

$$R_s(\omega) = \exp\left(-\frac{\omega^{\beta_s}}{\sigma_s}\right), \quad (6)$$

and:

$$R_\tau(\omega) = \exp\left(-\frac{|\omega^{\beta_\tau} - \omega_0^{\beta_\tau}|^2}{\sigma_\tau}\right), \tag{7}$$

where β_s, β_τ determine the decay slope, and σ_s, σ_τ control the bandwidth of sustained and transient channels, respectively. The spatiotemporal contrast sensitivity in 5-dimensional space is obtained by combining the responses from the channels described above:

$$\begin{aligned} &S(\rho,\omega,L,a,e) = \\ &S_{ecc}(e,\rho)\left(R_s(\omega)S_{A,s}(\rho,a,L) + R_\tau(\omega)S_{A,\tau}(\rho,a,L)\right), \end{aligned} \tag{8}$$

where a is the area of stimulus expressed in visual degrees, $S_{A,s}(\rho,a,L)$, and $S_{A,\tau}(\rho,a,L)$ are 3-dimensional spatial contrast sensitivity functions in sustained and transient channels, respectively. The parameter S_{ecc} models drop of sensitivity in peripheral vision.

Critical flicker frequency predictions

In this section, our goal is to obtain an explicit equation that shows how the CFF values are affected by the various parameters of the two spatiotemporal CSF models we introduced earlier. We take advantage of the separability of spatial and temporal sensitivity in high frequencies [33, 34]. Separability here means that there are no longer any spatiotemporal interactions in higher spatial or temporal frequencies. Hence, changes in the spatial configuration of the stimulus have no impact on temporal sensitivity, except for a vertical shift and vice versa [35]. This behavior in the Barten's CSF would appear in the middle to high temporal frequencies where T_1 is low, so $(1 - S_1(\rho)T_1(\omega,L)) \approx 1$, reducing the Eq. 1 to the following expression:

$$S(\rho,\omega,L,e) \approx \frac{S_0(\rho,L,e)\,T_0(\omega,L)}{2k\,m_n(X,Y,T,e,\rho)}. \tag{9}$$

We assume that CFF is measured for the highest possible contrast of 1, which corresponds to the contrast sensitivity equal to 1. Therefore, CFF can be estimated as the temporal frequency at which CSF crosses the line of sensitivity equal to 1. By introducing Eq. 5 in Eq. 9, setting the sensitivity to 1, solving for temporal frequency, and rearranging one obtains:

$$\omega_{CFF} \approx \frac{\sqrt{\left(\frac{S_0(\rho,L,e)}{2k\,m_n(X,Y,T,e,\rho)}\right)^{\frac{2}{n_0}} - 1}}{2\pi\,\tau_0(L,D)}, \tag{10}$$

the resulting equation exhibits the relationship of CFF with the various parameters of the Barten's CSF.

Similar steps can be taken to demonstrate the separability behavior in stelaCSF. We assume in higher temporal frequencies, the contribution of the sustained channel is negligible; therefore, Eq. 8 reduces to the following expression:

$$S(\rho,\omega,L,a,e) = S_{ecc}(e,\rho)\left(R_\tau(\omega)S_{A,\tau}(\rho,a,L)\right). \tag{11}$$

By introducing Eq. 7 to Eq. 11, setting sensitivity to 1, and solving for temporal frequency one obtains ($\omega > \omega_0$):

$$\omega_{CFF} \approx \left(\sqrt{\sigma_\tau \ln\left(S_{ecc}\,S_{A,\tau}\right)} + \omega_0^{\beta_\tau}\right)^{\frac{1}{\beta_\tau}}. \tag{12}$$

Modifications to Barten's CSF

Barten relied on the variation of on-center parasol cells' density to predict the variation in spatial summation and sampling in the periphery [31]. It has been argued that this assumption may not be accurate and result in overestimation of the fall of acuity and sensitivity with eccentricity [36]. Therefore, to address this issue, we rely on the density of midget cells [37] and local scale [38] to account for the variation of acuity and peak sensitivity, respectively [36]. We assume that the local scale depends on the visual meridian ($k_{temporal}, k_{nasal}$). Aliasing in the periphery is neglected for simplicity.

To account for the variation of CFF with eccentricity in the Barten spatiotemporal CSF, we modify Eq. 5 by making the time constant of the temporal filter, $\tau_0(L,D)$, a function of eccentricity. It is assumed that the variation in the time constant of the normalized amplitude response abstracts the overall influence of gain, integration time, and desensitization on the bandwidth of the temporal CSF curve. Therefore, variation in the time constant of the temporal filter does not necessarily translate to a change in the integration time. From Eq. 10, it is evident that there is an inverse proportionality between CFF and time constant. Hence, a decrease in the time constant of the temporal filter enhances CFF values. According to Barten, the time constant of the photoreceptor filter changes as a function of stimulus size and illuminance based on the following expression:

$$\tau_0(I,D) = \frac{\tau_{10}}{1 + 0.55\ln\left(1 + \frac{I}{3.5}(1+D)^{0.6}\right)}, \tag{13}$$

where τ_{10} is the time constant in the dark, D is the stimulus diameter in visual degrees, and I is the illuminance expressed in Trolands. We model the influence of eccentricity on the time constant by adding an additional term to the denominator:

$$\begin{aligned} &\tau_0(I,D,e) = \\ &\frac{\tau_{10}}{1 + 0.55\ln\left(1 + \frac{I}{k_{\tau_0}}(1+D)^{0.6}\left(\frac{S_{\mathrm{rf}}(e)}{S_{\mathrm{rf}}(0)}\right)\right)}, \end{aligned} \tag{14}$$

where k_{τ_0} is a model parameter, $S_{\mathrm{rf}}(e)$ is the size of the receptive field's center (RF size) as a function of eccentricity, and $S_{\mathrm{rf}}(0)$ is the size of the receptive field at the fovea. The intuition behind this modification comes from the investigations of Rovamo et al. [18] that emphasized the role of luminous flux on the variation of CFF. Therefore, by introducing a normalized RF size term to the equation, we shift the dependency of $\tau_0(I,D,e)$ from illuminance to luminous flux. As mentioned earlier, the luminous flux here is defined as the product of illuminance and RF size expressed in $[(trolands)\,(degree)^2]$. To obtain quantitative estimates of RF size, we rely on a characteristic feature of ganglion cell mosaics: the maintenance of constant dendritic overlap across the retina [39, 40]. This feature implies that the product of ganglion cells' density and their dendritic field area, commonly referred to as the dendritic coverage factor in morphological studies [41], is fixed across the retina:

$$C_{\mathrm{f}} = d_{\mathrm{gf}}(e,k)\,S_{\mathrm{rf}}(e) \tag{15}$$

where $d_{\mathrm{gf}}(e,k)$ is ganglion cells' receptive field density derived from Watson's formula [37], and C_{f} is the dendritic coverage factor. It is assumed here that the dendritic field area represents the

RF size. The ganglion cells' receptive field density is given by the following expression [37]:

$$d_{\mathrm{gf}}(e,k) = d_{\mathrm{gf}}(0)\left(a_{\mathrm{k}}\left(1+\frac{e}{e_{2,\mathrm{k}}}\right)^{-2}+(1-a_{\mathrm{k}})\exp\left(-\frac{e}{e_{\mathrm{r,k}}}\right)\right), \quad (16)$$

where $d_{\mathrm{gf}}(0)$ is ganglion cells' receptive field density in the fovea, $e_{\mathrm{r,k}}$, $e_{2,\mathrm{k}}$, a_{k} are formula's parameters, and the index k denotes visual field meridian. These parameters are optimized with anatomical data and their numerical values are reported in [37]. There is evidence that for certain subsets of human ganglion cells (midget cells), the coverage factor is one and may not vary across the retina [40, 42]. However, several other investigations in primates suggest an increase of coverage factor with eccentricity [43, 44]. Nevertheless, given the limited range of variations observed in primate studies, we assume the coverage factor does not vary with eccentricity or meridian. Note that an increased coverage factor with eccentricity would imply an even larger spatial summation; therefore, our assumption results in a conservative approximation of the RF size. Finally, the RF size is obtained by rearrangement of Eq. 15:

$$S_{\mathrm{rf}}(e) = \frac{C_{\mathrm{f}}}{d_{\mathrm{gf}}(e,k)}, \quad (17)$$

note that due to the normalization of $S_{\mathrm{rf}}(e)$ with its foveal value $S_{\mathrm{rf}}(0)$, we do not need to assume a value for C_{f} in Eq. 14.

Modifications to stelaCSF

Measurements of CFF in bright conditions have revealed a plateau in the Ferry-Porter plots, indicating that CFF's increase slows down and ultimately ceases beyond a certain light intensity [45, 46]. This saturation effect resembles the plateau and subsequent sensitivity decline observed in the SCSF measurements of static (0 Hz) Gabor stimuli [47]. However, the luminance range at which these effects occur in SCSF and CFF varies, offering the involvement of distinct underlying visual mechanisms. Unlike Barten's CSF, stelaCSF operates on two separate temporal channels, allowing for two distinct luminance sensitivity functions based on the temporal frequency regime that stimuli reside in. In the original version of stelaCSF, the variation of peak sensitivity for the transient channel is modeled as a linear increase in a double logarithm plot (see Figure 5 in [1]). To account for the effect of saturation in CFF, we propose the following alternative expression:

$$S_{m,\tau}(L) = k_{\mathrm{s1},\tau}\left(1+\frac{k_{\mathrm{s2},\tau}}{L}\right)^{-k_{\mathrm{s3},\tau}}, \quad (18)$$

where $k_{\mathrm{s1},\tau}$, $k_{\mathrm{s2},\tau}$, $k_{\mathrm{s2},\tau}$ are model parameters , and L is luminance in candela per square meter. From the Eq. 12, a directly proportional relationship between the bandwidth of the transient channel, σ_τ, and CFF is evident. To account for the non-monotonic course of the CFF in peripheral regions, we propose to control the bandwidth of the transient channels (see Eq. 7) as follows:

$$\sigma_\tau(e) = \sigma_\tau\left(1+k_{\sigma,1}\,e^{k_{\sigma,1}}\right), \quad (19)$$

where $k_{\sigma,1}$, $k_{\sigma,2}$ are model parameters, and e is eccentricity in visual degrees. All model parameters are optimized based on the experimental measurements.

Table 1: The fitted parameters for original and modified versions of spatiotemporal CSF models. The HTF name extension stands for "high temporal frequency" and denotes modified models. For a detailed description of the parameters for stelaCSF and Barten's CSF, refer to [1] and [31], respectively.

Models	Parameters
stelaCSF	$k_{\mathrm{a},s}=0.0601$, $k_{\mathrm{b},s}=0.0002197$, $k_{\mathrm{a},\tau}=0.0002726$, $k_{\mathrm{b},\tau}=2.47976$, $\beta_s=1.331$, $\sigma_s=5.3266$, $\beta_\tau=0.1898$, $\sigma_\tau=0.129405$, $k_{\mathrm{s1},s}=57.6634$, $k_{\mathrm{s2},s}=90.5523$, $k_{\mathrm{s3},s}=0.17030$, $k_{\mathrm{s4},s}=8.36556e-7$, $k_{\mathrm{s5},s}=6.9238e9$, $k_{\rho1,s}=1.54334$, $k_{\rho2,s}=11.5343$, $k_{\rho3,s}=0.281654$, $k_{\mathrm{s1},\tau}=0.4378$, $k_{\mathrm{s2},\tau}=216.51$, $k_{\rho,\tau}=0.00199$, $k_{\mathrm{e1}}=0.01552$, $k_{\mathrm{e2}}=0.0196$, $k_{\mathrm{e1,nasal}}=0.00916$, $k_{\mathrm{e2,nasal}}=0.01615$
stelaCSF-HTF	$k_{\mathrm{a},s}=0.0517$, $k_{\mathrm{b},s}=0.00022$, $k_{\mathrm{a},\tau}=0.0002718$, $k_{\mathrm{b},\tau}=2.9898$, $\beta_s=1.331$, $\sigma_s=2.16508$, $\beta_\tau=0.1898$, $\sigma_\tau=0.1176$, $k_{\mathrm{s1},s}=69.20$, $k_{\mathrm{s2},s}=144.74$, $k_{\mathrm{s3},s}=0.1555$, $k_{\mathrm{s4},s}=7.097e-7$, $k_{\mathrm{s5},s}=7.917e9$, $k_{\rho1,s}=1.51$, $k_{\rho2,s}=9.4$, $k_{\rho3,s}=0.2874$, $k_{\mathrm{s1},\tau}=20825.9$, $k_{\mathrm{s2},\tau}=168.77$, $k_{\mathrm{s3},\tau}=0.58755$, $k_{\rho,\tau}=9.6569e-5$, $k_{\mathrm{e1}}=0.02578$, $k_{\mathrm{e2}}=0.01864$, $k_{\mathrm{e1,nasal}}=0.01545$, $k_{\mathrm{e2,nasal}}=0.01563$, $k_{\sigma,1}=0.00001$, $k_{\sigma,2}=94.80$
Barten's CSF	$k=1.44$, $\sigma_0=0.72$, $\rho_{c_1}=14.99$, $\tau_{10}=0.03663$, $\tau_{20}=0.01117$
Barten's CSF-HTF	$k=7.13$, $\sigma_0=0.38$, $\rho_{c_1}=2.45$, $\tau_{10}=0.03666$, $\tau_{20}=0.09171$, $k_{nasal}=0.2$, $k_{temporal}=0.53$, $k_{\tau_0}=1$

Model Fitting and Performance

This section will describe the fitting procedure for the models. To ensure proper training and adequate evaluation of the model, we rely on 15 datasets of contrast sensitivity and CFF measurements. We consider a per-dataset sensitivity adjustment factor to provide a simultaneous fit to all datasets using the same set of model parameters. This factor accounts for variations in experimental methodology and stimulus features not explained by the model's parameters. The fitting procedure and loss function are identical to those used for calibrating stelaCSF [1]. We rely on the "fminunc" function in Matlab software to optimize the free parameters of both models. Table 1 lists the set of optimized parameter values for each model. Note that in the case of the Barten's CSF, we limited both the training and evaluation to the photopic range, as this model is not valid for scotopic vision and results in large prediction errors [31]. The parameters in Barten's CSF that are not listed in Table 1 are not trained on the datasets mentioned here, and the typical values reported in [31] are used for them.

Our analysis includes four CFF datasets, which we refer to by the first author's name for concision: Hecht [45], Delange [55], Krajancich [23], and Chapiro [46]. The stimuli in Hecht, Delange, and Chapiro datasets were discs, while in Krajancich dataset were Gabor patches with spatial frequencies ranging from 0 to 2 cpd. The size of the discs used in these experiments was less than 2 visual degrees, while the standard deviations of the Gaussian envelopes used in the Krajancich study were inversely proportional to the spatial frequency of the Gabor patch. Hecht and Chapiro

Table 2: RMSE for prediction per dataset. The small font numbers in the last row denote the average and standard deviations resulting from 5-fold cross-validation analysis. It is important to note that the stelaCSF and Barten's CSF models were fitted to different portions of the dataset, with Barten's CSF only modeling photopic vision. Therefore, it is not possible to compare the results of both models. For the sake of brevity, datasets are referred to by their first author's names.

	stelaCSF	stelaCSF-HTF	Barten's CSF	Barten's CSF-HTF
Modelfest [48]	4.74	4.41	7.46	3.09
HDR-VDP [49]	3.26	3.29	7.20	3.89
HDR-CSF [47]	3.65	3.55	4.22	3.94
Rovamo [50]	2.70	2.73	5.10	3.97
Robson [33]	3.06	3.15	3.71	2.72
Laird [51]	5.92	6.30	8.14	7.47
Snowden [52]	5.41	5.35	7.47	6.03
Virsu [25]	6.60	5.71	8.44	3.75
Virsu [53]	4.75	4.25	10.45	4.50
Wright [54]	2.79	3.04	6.06	4.50
Anderson [6]	8.15	7.30	11.40	9.61
Krajancich [23]	10.24	7.42	23.14	5.53
Hecht [45]	4.89	3.52	6.17	4.95
Chapiro [46]	4.64	3.61	10.33	4.32
Delange [55]	3.33	2.56	2.50	4.44
Average	4.74	**4.41**	8.05	4.98
Average testing	5.30 ± 0.17	4.87 ± 0.13	9.22 ± 0.23	5.37 ± 0.48

datasets had limited or no surround, while in Delange and Krajancich's experiment, there was an extensive surround with an intensity set to the time-average intensity of the flickering stimuli. Two datasets from Hecht and Chapiro included measurements in high luminance values (above 1000 cd/m^2). In our analysis, we excluded CFF measurements in low luminance conditions (below 1 cd/m^2) for both models due to the possibility of rod intrusion in flicker detection (see appendix in [26]).

A challenge when estimating the contrast sensitivity of discs in a spatiotemporal model is determining an effective value for the spatial frequency parameter. It is not possible to set the spatial frequency of a uniform disc to zero as the spatial configuration of the disc affects temporal sensitivity by determining the strength of inhibitory signals [31]. To estimate the contrast sensitivity of flickering discs using the stelaCSF, we relied on the peak sensitivity of spatial CSF (see [56] for an evaluation of different approaches). On the other hand, when using the Barten's CSF, we follow his approach by setting S_0 to 1 and calculating the spatial frequency of the fundamental wave using $\rho = (\pi D)^{-1}$, where D is the diameter of the flickering disc expressed in visual degrees.

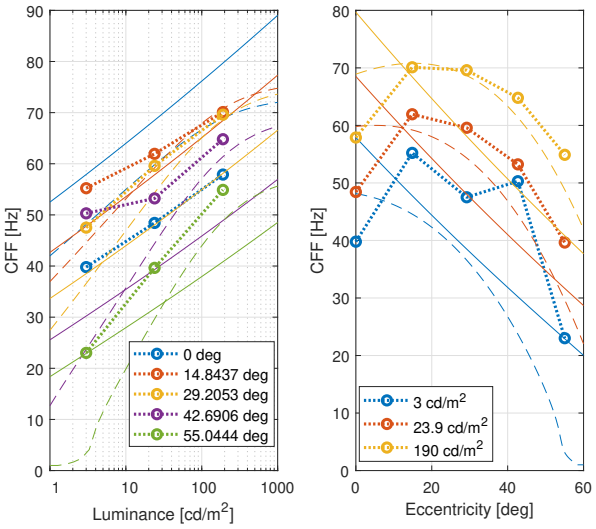


Figure 2: CFF predictions from original (solid lines) and modified (dashed lines) stelaCSF for Krajancich (circles) dataset [23]. The left panel shows the Ferry-Porter lines in different eccentricities. The right panel shows the variation of CFF as a function of eccentricity under different light intensities. In both panels, the spatial frequency of the stimuli is 0.57 cpd.

Psychophysical experiments designed for the measurement of CSF and CFF are usually time-consuming. As a result, most of the available datasets only examine a few sets of stimulus parameters. Due to this restriction, it is not possible to divide the datasets into test/train splits, as test errors for a specific dataset will not be representative of the overall performance of a model [1]. Therefore, we use the entire dataset for training and rely on the fitting error to measure performance. Additionally, we conduct a five-fold cross-validation analysis for each dataset and report the average testing errors along with their standard deviations. The prediction errors are measured in root-mean-squared error (RMSE) and expressed in decibel units.

Results and Discussion

The RMSE prediction errors for the original and modified spatiotemporal models for all datasets are provided in Table 2. The last row reports the average testing errors from a 5-fold cross-validation analysis.

In Table 2, it is evident that the modified models outperform the original models in terms of average prediction errors. The improvement in Barten CSF's performance is especially remarkable after modification. In 14 out of 15 datasets, the modified Barten model consistently produces lower prediction errors. Among all models tested, the modified version of stelaCSF yields the lowest prediction error. One significant advantage of the stelaCSF model over the Barten's CSF model is its capability to handle a broad range of luminance values, even in scotopic conditions.

Out of all the CFF datasets, predictions for the Krajancich dataset [23] show the most significant improvement after modification. This dataset measures CFF values based on spatial frequency, eccentricity, and luminance. When comparing stelaCSF's predictions to the Krajancich dataset (shown in Figure 2), the orig-

inal version resulted in a linear decrease of CFF values, which doesn't match CFF measurements. However, the modified version shows a non-monotonic variation in higher eccentricities. This is due to a modification applied to the transient channel's bandwidth (see Eq. 19), causing an initial increase. Additionally, there are improvements in predicting foveal CFF after the modification.

Figure 3 compares the predictions of Barten's CSF for the Krajancich dataset. The original Barten, like the original stelaCSF, cannot capture the non-monotonic variation of CFF and exhibits a quick drop in eccentricity, resulting in significant prediction errors. However, both foveal and peripheral predictions have improved after adjusting for luminous flux and modifying Equation 14. Interestingly, the modified Barten model delivers the lowest prediction error for the Krajancich dataset. Nevertheless, modified stelaCSF follows the overall data trend more accurately despite larger errors. The initial increase in modified Barten's peripheral predictions appears steeper than the experimental data. Still, it isn't easy to draw a definitive conclusion due to the sparse sampling of the experimental data along the eccentricity dimension. Both models have larger errors for the lowest light intensity, which may be due to rod intrusion [26].

The Ferry-Porter lines from modified stelaCSF tend to have steeper slopes at higher eccentricity values. This trend was previously noted by Tyler and Hamer through a series of peripheral CFF measurements, where they reported a more than twofold increase in the slope of Ferry-Porter lines from 0 to 35 visual degrees [12, 57]. This trend is also evident in the measurements of Krajancich et al. [23], but the sampling along the luminance dimension is quite limited.

Figure 4 shows the predictions made by models for the Delange, Hecht, and Chapiro datasets. All datasets exhibit an almost linear increase of CFF values up to 1000 cd/m^2 as the logarithm

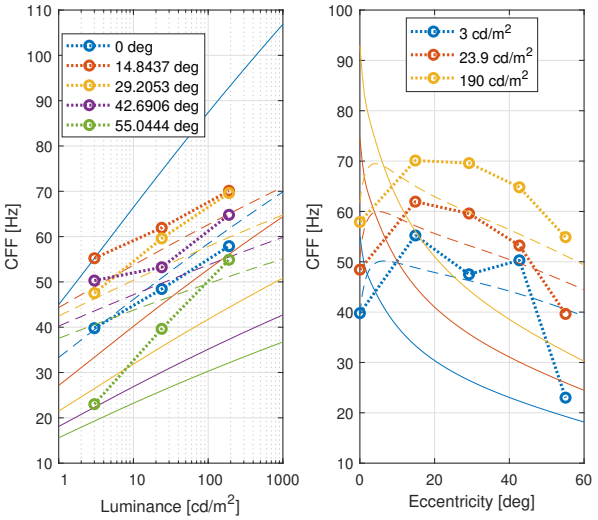


Figure 3: CFF predictions from original (solid lines) and modified (dashed lines) Barten's CSF for Krajancich (circles) dataset [23]. The left panel shows the Ferry-Porter lines in different eccentricities. The right panel shows the variation of CFF as a function of eccentricity under different light intensities. In both panels, the spatial frequency of the stimuli is 0.57 cpd.

of luminance increases, which follows the well-known Ferry-Porter law previously discussed. The modified stelaCSF model indicates a slower increase in CFF at the luminance value above 1000 cd/m^2. The prediction errors of the stelaCSF decreased for all three datasets after modification to the peak sensitivity function in the transient channel (refer to Eq. 18).

A limitation in our analysis and model fitting is the absence of explicit consideration for the impact of the surrounding field. The majority of CFF measurements are conducted in a dark setting, but some, like Delange, Krajancich, and Hartmann [17], have conducted them in a surrounding field with an intensity matching the time-averaged intensity of the flicker. Barten has suggested modeling the impact of the surrounding field by adjusting the effective field size of the disk to the surrounding area [31].

Conclusions

We introduced several modifications to Barten's CSF and stelaCSF to enhance the accuracy of CFF predictions. Our comparison to psychophysical measurements from 15 datasets revealed qualitative improvements in following the characteristic trends in data and quantitative enhancements in prediction errors. Specifically, the modified version of stelaCSF resulted in the least average RMSE values across all datasets. The stelaCSF's dual-channel mechanism was particularly useful in explaining the saturation of CFF measurements at high luminance values.

References

[1] Rafał K. Mantiuk, Maliha Ashraf, and Alexandre Chapiro. stelaCSF - A Unified Model of Contrast Sensitivity as the Function of Spatio-Temporal Frequency, Eccentricity, Luminance and Area. *ACM Transactions on Graphics*, 41(4):16, 7 2022.

[2] Andrew B. Watson, H. B. Barlow, and John G. Robson. What does the eye see best? *Nature*, 302(5907):419–422, 1983.

[3] Albert Rose. *Vision: human and electronic*. Springer Science & Business Media, 2013.

[4] Rosario M. Balboa and Norberto M. Grzywacz. The role of early retinal lateral inhibition: More than maximizing luminance information. *Visual Neuroscience*, 17(1):77–89, 2000.

[5] Floyd Ratliff. Mach bands: quantitative studies on neural networks. *Retina. San Francisco, CA: Holden-Day*, 1965.

[6] S. J. Anderson, K. T. Mullen, and R. F. Hess. Human peripheral spatial resolution for achromatic and chromatic stimuli: limits imposed by optical and retinal factors. *The Journal of Physiology*, 442(1):47–64, 1991.

[7] E. SIMONSON and J. BROZEK. Flicker fusion frequency; background and applications. *Physiological reviews*, 32(3):349–378, 7 1952.

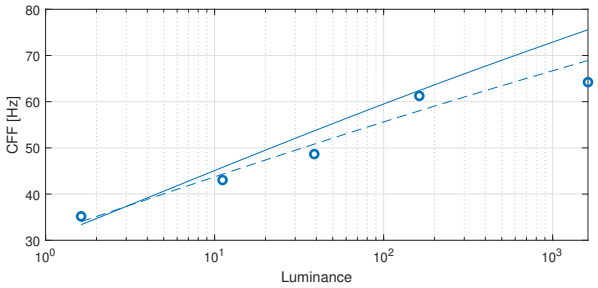
[8] C. LANDIS. Determinants of the Critical Flicker-Fusion Threshold. *https://doi.org/10.1152/physrev.1954.34.2.259*, 34(2):259–286, 4 1954.

[9] Kristian Donner. Temporal vision: Measures, mechanisms and meaning, 7 2021.

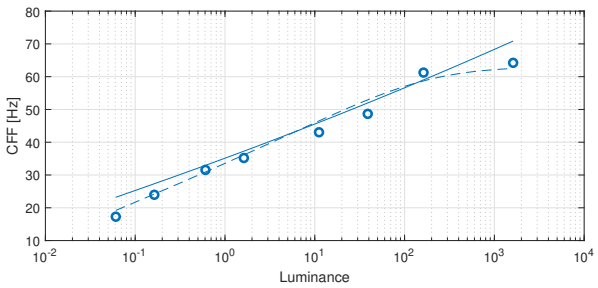
[10] Ervin S Ferry. ART. XXVI.–Persistence of Vision. *American Journal of Science (1880-1910)*, 44(261):192, 1892.

[11] Thomas Conrad Porter. Contributions to the study of flicker. Paper II. *Proceedings of the Royal Society of London*, 70(459-466):313–329, 1902.

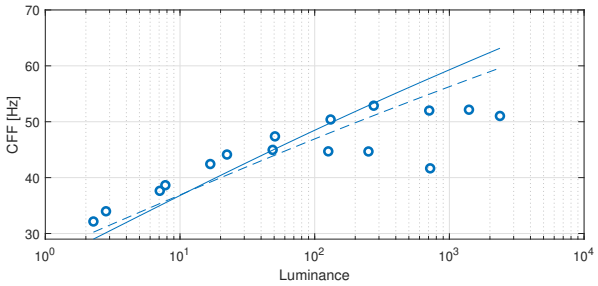
[12] Christopher W. Tyler and Russell D. Hamer. Analysis of visual mod-



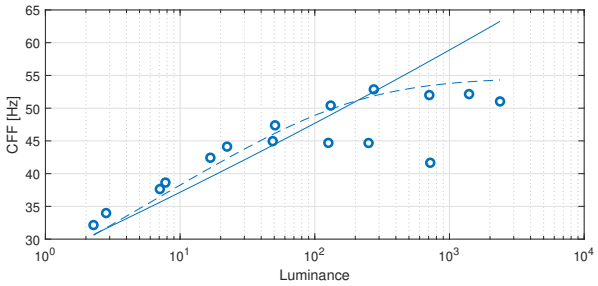
(a) Barten's CSF and Barten's CSF-HTF for Delange



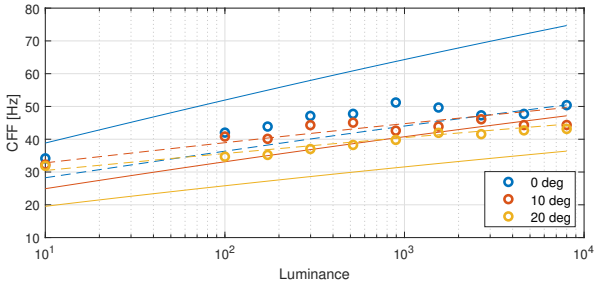
(b) stelaCSF and stelaCSF-HTF for Delange



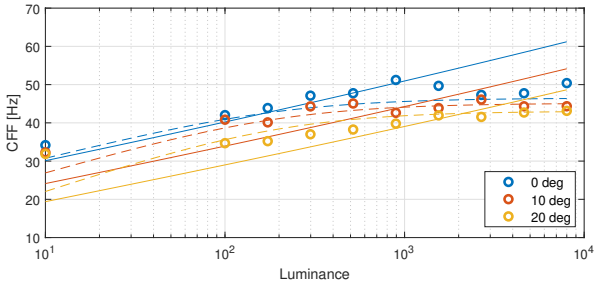
(c) Barten's CSF and Barten's CSF-HTF for Hecht



(d) stelaCSF and stelaCSF-HTF for Hecht



(e) Barten's CSF and Barten's CSF-HTF for Chapiro



(f) stelaCSF and stelaCSF-HTF for Chapiro

Figure 4: The predictions from original (solid lines) and modified (dashed lines) versions of spatiotemporal models for Delange, Hecht, and Chapiro datasets (circles). The modified stelaCSF captures the saturation at high light intensities.

ulation sensitivity IV Validity of the Ferry–Porter law. *Journal of the Optical Society of America A*, 7(4):743, 4 1990.

[13] Andrew T. Rider, G. Bruce Henning, and Andrew Stockman. A reinterpretation of critical flicker-frequency (CFF) data reveals key details about light adaptation and normal and abnormal visual processing. *Progress in Retinal and Eye Research*, 87:101001, 3 2022.

[14] B. S. Hylkema. EXAMINATION OF THE VISUAL FIELD BY DETERMINING THE FUSION FREQUENCY. *Acta Ophthalmologica*, 20(2):181–193, 9 1942.

[15] William De Wiveleslie Abney. VII. The sensitiveness of the retina to light and colour. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, (190):155–195, 1897.

[16] Ragnar Granit and Phyllis Harper. COMPARATIVE STUDIES ON THE PERIPHERAL AND CENTRAL RETINA. *https://doi.org/10.1152/ajplegacy.1930.95.1.211*, 95(1):211–228, 10 1930.

[17] E. Hartmann, B. Lachenmayr, and H. Brettel. The peripheral critical flicker frequency. *Vision Research*, 19(9):1019–1023, 1 1979.

[18] Jyrki Rovamo and Antti Raninen. Critical flicker frequency and M-scaling of stimulus size and retinal illuminance. *Vision Research*, 24(10):1127–1131, 1 1984.

[19] Scott J. Daly. Visible differences predictor: an algorithm for the assessment of image fidelity. In Bernice E. Rogowitz, editor, *Human Vision, Visual Processing, and Digital Display III*, volume 1666, page 2. SPIE, 8 1992.

[20] Rafał K. Mantiuk, Gyorgy Denes, Alexandre Chapiro, Anton Kaplanyan, Gizem Rufo, Romain Bachy, Trisha Lian, and Anjul Patney. FovVideoVDP: A visible difference predictor for wide field-of-view video. *ACM Transactions on Graphics*, 40(4), 7 2021.

[21] Albert J. Ahumada, Jr. and Heidi A. Peterson. Luminance-model-based DCT quantization for color image compression. *Human Vision, Visual Processing, and Digital Display III*, 1666(27):365, 8 1992.

[22] Okan Tarhan Tursun, Elena Arabadzhiyska-Koleva, Marek Wernikowski, Radosław Mantiuk, Hans Peter Seidel, Karol Myszkowski, and Piotr Didyk. Luminance-contrast-aware foveated rendering. *ACM Transactions on Graphics*, 38(4):1–14, 7 2019.

[23] Brooke Krajancich, Petr Kellnhofer, Gordon 2021 Wetzstein, Stanford University, Raxium Gordon, and Gordon 2021 Wetzstein. A

perceptual model for eccentricity-dependent spatio-temporal flicker fusion and its applications to foveated graphics. *ACM Transactions on Graphics*, 40(4):11, 4 2021.

[24] Christopher W. Tyler. Analysis of visual modulation sensitivity II Peripheral retina and the role of photoreceptor dimensions. *Journal of the Optical Society of America A*, 2(3):393, 3 1985.

[25] V. Virsu and J. Rovamo. Visual resolution, contrast sensitivity, and the cortical magnification factor. Technical Report 3, 1979.

[26] Andrew T. Rider, G. Bruce Henning, and Andrew Stockman. Light adaptation controls visual sensitivity by adjusting the speed and gain of the response to light. *PLoS ONE*, 14(8):e0220358, 8 2019.

[27] Christina Enroth-Cugell and R. M. Shapley. Flux, not retinal illumination, is what cat retinal ganglion cells really care about. *The Journal of Physiology*, 233(2):311, 9 1973.

[28] Christina Enroth-Cugell and R. M. Shapley. Adaptation and dynamics of cat retinal ganglion cells. *The Journal of Physiology*, 233(2):271–309, 9 1973.

[29] Albert Ahumada, Jihyun Yeonan-Kim, and Andrew B. Watson. A dual channel spatial-temporal detection model. In *IS and T International Symposium on Electronic Imaging Science and Technology*, volume 30, pages 1–4. Society for Imaging Science and Technology, 1 2018.

[30] Peter G. J. Barten. Formula for the contrast sensitivity of the human eye. *Image Quality and System Performance*, 5294(December 2003):231–238, 2003.

[31] Barten and Peter G. *Contrast Sensivity of the Human Eye and its Effects on Image Quality*. SPIE press, 1999.

[32] Peter G. J. Barten. Spatiotemporal model for the contrast sensitivity of the human eye and its temporal aspects. In *Human Vision, Visual Processing, and Digital Display IV*, volume 1913, pages 2–14. SPIE, 9 1993.

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

[34] Andrew B. Watson. Temporal Sensitivity. *Handbook of Perception and Human Performance*, pages 1–6, 1986.

[35] Andrew B. Watson. Temporal Sensitivity. pages 1–6, 1986.

[36] Aliakbar Bozorgian, Marius Pedersen, and Jean-Baptiste Thomas. Modification and evaluation of the peripheral contrast sensitivity function models. *Journal of the Optical Society of America A*, 39(9):1650, 9 2022.

[37] Andrew B. Watson. A formula for human retinal ganglion cell receptive field density as a function of visual field location. *Journal of Vision*, 14(7):1–17, 2014.

[38] Andrew B. Watson. Estimation of local spatial scale. *Journal of the Optical Society of America A*, 4(8):1579, 8 1987.

[39] H. Wassle and B. B. Boycott. Functional architecture of the mammalian retina. *https://doi.org/10.1152/physrev.1991.71.2.447*, 71(2):447–480, 1991.

[40] Dennis M. Dacey and Michael R. Petersen. Dendritic field size and morphology of midget and parasol ganglion cells of the human retina. *Proceedings of the National Academy of Sciences of the United States of America*, 89(20):9666–9670, 10 1992.

[41] V. H. Perry, R. Oehler, and A. Cowey. Retinal ganglion cells that project to the dorsal lateral geniculate nucleus in the macaque monkey. *Neuroscience*, 12(4):1101–1123, 8 1984.

[42] D. M. Dacey. The mosaic of midget ganglion cells in the human retina. *Journal of Neuroscience*, 13(12):5334–5355, 1993.

[43] F. L. Gomes, L. C.L. Silveira, C. A. Saito, and E. S. Yamada. Density, proportion, and dendritic coverage of retinal ganglion cells of the common marmoset (Callithrix jacchus jacchus). *Brazilian Journal of Medical and Biological Research*, 38(6):915–924, 2005.

[44] Elizabeth S. Yamada, Luiz Carlos L. Silveira, and V. Hugh Perry. Morphology, dendritic field size, somal size, density, and coverage of M and P retinal ganglion cells of dichromatic Cebus monkeys. *Visual Neuroscience*, 13(6):1011–1029, 1996.

[45] Seug Hecht and Cornelis D. Verrijp. Intermittent stimulation by light III. the relation between intensity and critical fusion frequency for different retinal locations. *Journal of General Physiology*, 17(2):251–268, 11 1933.

[46] Alexandre Chapiro, Nathan Matsuda, Maliha Ashraf, and Rafal Mantiuk. Critical flicker frequency (CFF) at high luminance levels. *Electronic Imaging*, 35(10):223–1, 1 2023.

[47] Sophie Wuerger, Maliha Ashraf, Minjung Kim, Jasna Martinovic, María Pérez-Ortiz, and Rafał K. Mantiuk. Spatio-chromatic contrast sensitivity under mesopic and photopic light levels. *Journal of Vision*, 20(4):23–23, 4 2020.

[48] Andrew B. Watson and Albert J. Ahumada. A standard model for foveal detection of spatial contrast. *Journal of Vision*, 5(9):717–740, 10 2005.

[49] Rafał Mantiuk, Kil Joong Kim, Allan G. Rempel, and Wolfgang Heidrich. HDR-VDP-2. *ACM Transactions on Graphics (TOG)*, 30(4):1–14, 7 2011.

[50] Jyrki Rovamo, Olavi Luntinen, and Risto Näsänen. Modelling the dependence of contrast sensitivity on grating area and spatial frequency. *Vision Research*, 33(18):2773–2788, 12 1993.

[51] Justin Laird, Mitchell Rosen, Jeff Pelz, Ethan Montag, and Scott Daly. Spatio-velocity CSF as a function of retinal velocity using unstabilized stimuli. In *Human Vision and Electronic Imaging XI*, volume 6057, page 605705. SPIE, 2 2006.

[52] Robert J. Snowden, Robert F. Hess, and Sarah J. Waugh. The processing of temporal modulation at different levels of retinal illuminance. *Vision Research*, 35(6):775–789, 3 1995.

[53] Veijo Virsu, Jyrki Rovamo, Pentti Laurinen, and Risto Näsänen. Temporal contrast sensitivity and cortical magnification. *Vision Research*, 22(9):1211–1217, 1 1982.

[54] M. J. Wright and A. Johnston. Spatiotemporal contrast sensitivity and visual field locus. *Vision Research*, 23(10):983–989, 1 1983.

[55] H. DE LANGE DZN. Research into the dynamic nature of the human fovea-cortex systems with intermittent and modulated light. I. Attenuation characteristics with white and colored light. *Journal of the Optical Society of America*, 48(11):777–784, 11 1958.

[56] Maliha Ashraf, Rafal Mantiuk, and Alexandre Chapiro. Modelling contrast sensitivity of discs. *Electronic Imaging*, 35(10):246–1, 1 2023.

[57] Christopher W. Tyler and Russell D. Hamer. Eccentricity and the Ferry–Porter law. *Journal of the Optical Society of America A*, 10(9):2084, 9 1993.