

A Model of Local Adaptation supplementary information

Peter Vangorp*
Bangor University, UK &
MPI Informatik,
Germany

Karol Myszkowski†
MPI Informatik,
Germany

Erich W. Graf‡
University of Southampton,
United Kingdom

Rafał K. Mantiuk§
Bangor University, UK &
The Computer Laboratory,
University of Cambridge, UK

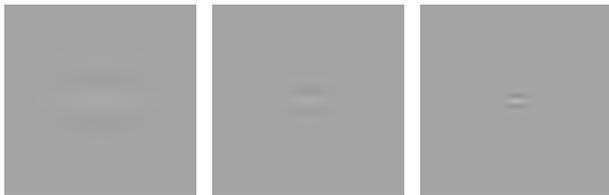


Figure 1: Gabor patches of different spatial frequency, typically used in detection experiments.

Abstract

This document contains additional details on the experiments and methods for the paper *A Model of Local Adaptation* [Vangorp et al. 2015]. In particular, the document contains some background information on detection and discrimination models used to derive the model contrast detection and adaptation pooling. The experiments are described and discussed in more detail than in the main paper.

1 Background: Detection, discrimination and adaptation

Because our local adaptation model is derived from detection data, it is important to understand this phenomenon, how it is measured and modelled. When an observer is asked to detect barely visible patterns on a uniform background, the minimum amplitude or contrast of such a pattern is known as a detection threshold. The patterns are usually designed to contain a single or a narrow band of spatial frequencies. Examples of such patterns are Gabor patches, such as shown in Figure 1, which contain only a single spatial frequency and are limited in their size by modulating their profile with a Gaussian envelope.

The detection performance has been thoroughly measured and modelled. The models of detection of frequency-limited patterns are known as Contrast Sensitivity Functions (CSFs) [Barten 1999; Watson and Ahumada 2005; Mantiuk et al. 2011; Kim et al. 2013]. They model *sensitivity* of the visual system for detecting patterns of a specified spatial frequency, background luminance, orientation, eccentricity (angular distance from the gaze point), stimulus size, etc. In practice, spatial frequency and background luminance have the biggest impact on sensitivity, and therefore those relations are shown in Figure 2.

The sensitivity is defined as the inverse of the detection contrast

$$S = \left(\frac{\Delta L}{L} \right)^{-1} = \frac{L}{\Delta L}, \quad (1)$$

where ΔL is the smallest detectable difference of luminance (or the

amplitude of the sinusoidal pattern in a Gabor patch), and L is the background luminance.

The CSF correctly predicts visual performance only for barely noticeable patterns shown on a uniform background. When the amplitude is much above the detection threshold, the gratings of different spatial frequency appear to have the same magnitude. This phenomenon is known as *contrast constancy* [Georgeson and Sullivan 1975].

Also, when the background is non-uniform, the CSF cannot predict visual performance. For example, when a band-limited noise pattern is shown on top of another noise pattern of similar frequency, it can be very hard to detect, as shown in the 4th column in Figure 3. However, when the spatial frequency of the background pattern differs from the spatial frequency of the detected pattern, the detected pattern is more visible, as shown in the leftmost and rightmost columns in Figure 3. This phenomenon is known as *contrast masking* and an experimental task of detecting a pattern on top of another pattern is known as *discrimination*. A number of contrast masking models can be found in the literature [Wilson 1980; Daly 1993; Watson and Solomon 1997].

Neither detection nor discrimination experiments can measure another important factor affecting visual performance, which is the luminance of adaptation. Different components of the visual system, such as photoreceptors and retinal neural cells, contain a gain-control mechanism, which adapts their sensitivity according to the intensity of the light [Shapley and Enroth-Cugell 1984]. Such gain control is necessary to allow for seeing over an enormous range of lighting conditions, which can vary from bright sunlight to dark moonlight. Some components of the adaptation mechanism are very fast (response in less than 50 ms), some are considerably slower. For example, full dark adaptation is reached after about 20-30 minutes. The detection and discrimination experiments assume that the luminance of adaptation is equal to the luminance of the background. This is achieved by displaying detection targets on a large uniform fields and ensuring that the observer had enough time to adapt. Even masking experiments are constructed in such a way that the average luminance of a masking pattern is predetermined so that the luminance of adaptation is fixed. However, the luminance in natural images is far from uniform and the adaptation luminance is usually unknown.

Because the adaptation mechanism is believed to reside mostly in receptors and post-receptor retinal cells [Dunn et al. 2007], it is fairly localized. One part of the visual field of the eye can be adapted to one luminance level, and another part to another luminance level. However, the spatial extent of such adaptation pools and the way the signal is combined from several receptors to control the adaptation is unknown. In this work we conduct a series of new psychophysical experiments to determine how the signal from photoreceptors is pooled to drive the adaptation mechanism.

2 Experiment 1: Probe-on-Flash experiment

A classical Probe-on-Flash psychophysical paradigm offers a method for measuring visual system performance when the eye is adapted

*e-mail: p.vangorp@bangor.ac.uk

†e-mail: karol@mpi-inf.mpg.de

‡e-mail: e.w.graf@soton.ac.uk

§e-mail: mantiuk@bangor.ac.uk

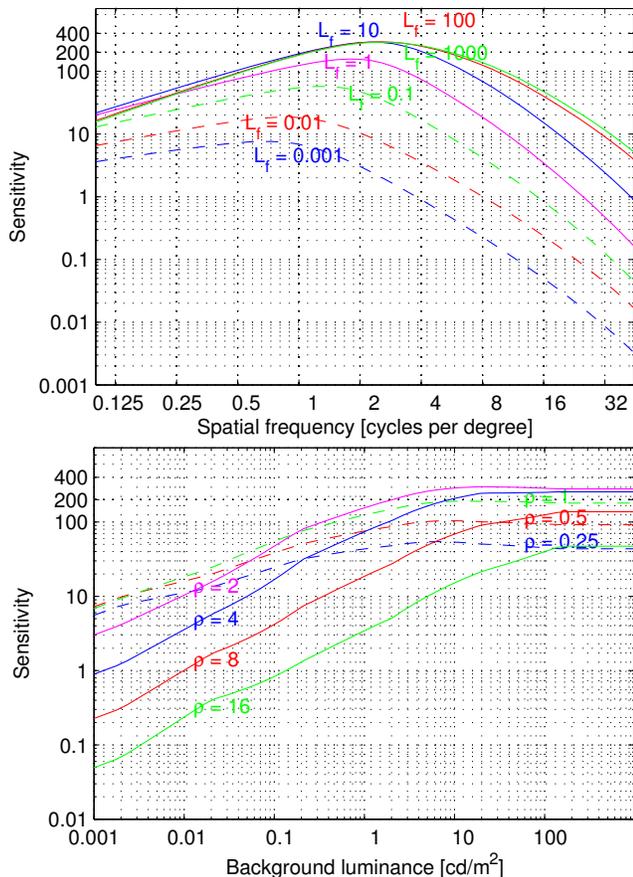


Figure 2: Contrast Sensitivity Function (CSF) plotted as the function of frequency (top) and luminance (bottom). Different line colors denote different background luminance (L_f), or spatial frequency (ρ). The plots are based on the model from [Mantiuk et al. 2011] and [Kim et al. 2013].

to different luminance than the luminance of the background. The data collected in Probe-on-Flash experiments form psychophysical evidence that the response of the photoreceptors can be modelled by an S-shaped curve, the so called Naka–Rushton model [Naka and Rushton 1966; Hood et al. 1979]. But the method is useful to study other aspects of luminance adaptation.

Figure 4 depicts a single trial of a probe-on-flash experiment. An observer is first presented with a uniform luminance field of luminance L_f to get the eye adapted to that level. Then, a test stimulus (a flash) is shown briefly for 200 ms. Usually a small disk (0.2° diameter) of a certain luminance L_p is used as a flash stimulus. Such short presentation time and possibly small size of the disk ensure that the adaptation state is least affected. The test stimulus contains a probe that needs to be detected. In our case, the probe is an edge modulated by a Gaussian envelope and the task is to determine whether the edge is vertical or horizontal. The larger the contrast of the edge, the easier the task. The psychophysical procedure is meant to test the probability of detecting the right orientation of the edge at different edge contrast levels. Each response from the participant was used to establish a detection threshold. Several techniques are available for determining threshold performance; here we used a QUEST staircase procedure [Watson and Pelli 1983], where each stimulus condition required at least 40 trials to establish a detection threshold. Once the data for several contrast levels is collected, the contrast level at which the right orientation of the edge is detected in 75% of

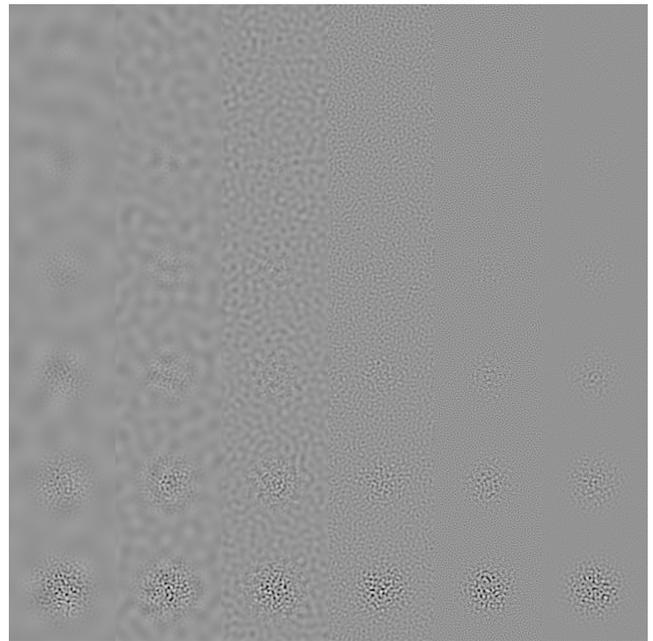


Figure 3: A 4 cpd band-limited noise pattern of different amplitude (rows) is superimposed on a noise pattern of different frequency (columns). The pattern is the least visible when frequency of both patterns is similar (4th column) and the easiest to detect when the frequencies are the most different.

the cases is assumed to be the detection threshold.

The thresholds measured in our experiment are shown in Figure 5. Each of the v-shaped curves in this figure was measured for different luminance of the flash (disk on which the edge was shown). The tip of the “v” represents the condition in which the adaptation luminance was the same as the flash luminance, which means that the probe was displayed on a uniform background (classical detection task). The visual performance worsens (the threshold contrasts increase) as the adaptation luminance starts to vary from the background luminance. This result is evidence of the adaptation process, which requires time and a stimulus of a certain size to reach a state of complete adaptation. The brief and small flashes used in the experiment do not allow the visual system to fully adapt, resulting in elevated detection thresholds.

3 Local adaptation experiments

To find a model of local adaptation we conducted a series of experiments, each measuring a different aspect of the adaptation field. The experimental procedure was similar to Experiment 1, however, the pedestals remained visible the whole time and only the detection target (the same edge or a Gabor patch) was briefly displayed for 200 ms. A single trial of such experiment is depicted in Figure 6.

3.1 Experiment 2: Frequency selectivity

The photoreceptors in the retina are organized in so-called receptive fields, in which a group of photoreceptors contribute to a single signal that is transmitted to the brain. Perhaps the best understood type of receptive fields are center-surround, in which centrally located photoreceptors elicit a positive signal and surrounding photoreceptors a negative signal (on-center, refer to Figure 7). Some neurophysiological measurements of animal retinæ suggest that the positive summation area of the receptive field contributes to gain control,

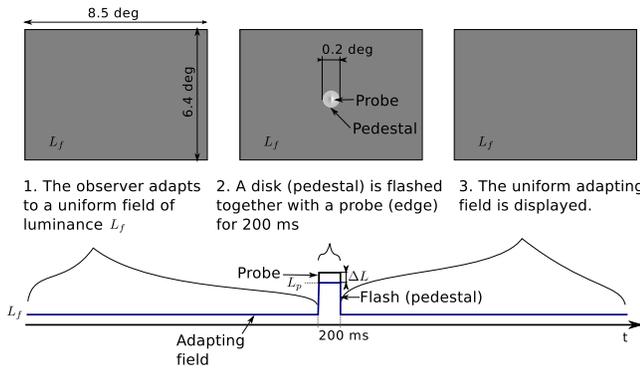


Figure 4: A single trial of a probe-on-flash experiment. The observer task is to detect whether the probe contains a horizontal or vertical edge.

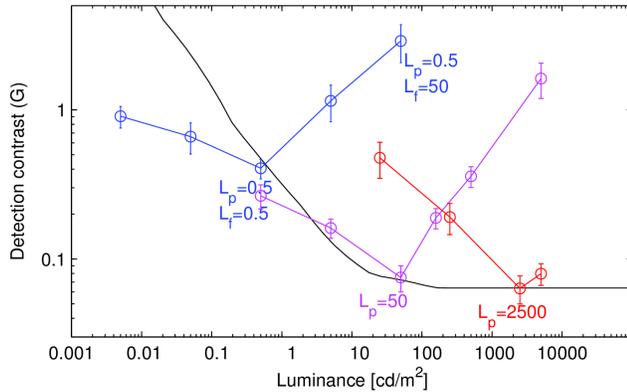


Figure 5: The results of the probe-on-flash experiment. The experimental results (solid color lines) are plotted as a function of varying adaptation luminance $L_a \approx L_f$, and one of three fixed levels of pedestal luminance L_p . Error bars represent the within-observer standard error of the mean (SEM). The black line is the twi function (plotted as logarithmic contrast).

and thus to luminance adaptation [Shapley and Enroth-Cugell 1984, pp. 304–309]. If this is the case, small features, which are detected by small receptive fields, will also have a small adaptation pool.

To test this hypothesis, we design an experiment in which the observers detect Gabor patches of two different sizes (spatial frequencies), which are shown on a Gaussian pedestal of varying size. The patches have a spatial frequency of 2 or 8 cycles per degree and the peak luminance is set to 500 cd/m^2 . The background is a uniform field of 5 cd/m^2 . Example stimuli are shown in Figure 8-top.

Patches of each frequency will be detected by receptive fields of different size. According to the hypothesis, the adaptation pool in both cases should be different, resulting in a different drop of sensitivity as the size of the pedestal background changes. However, as shown in the results of this experiment in Figure 8-bottom, the detection characteristics for Gabor patches of both frequencies are almost identical. The only visible difference can be observed for the largest width of the Gaussian which, however, cannot be caused by different size of the adaptation pool as the size of the Gaussian pedestal is too large. Given only subtle differences in the detection characteristics of both Gabor patches, we have no evidence for spatial selectivity of the adaptation mechanism.

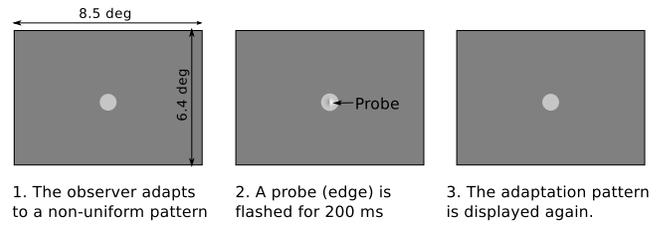


Figure 6: A single trial of a detection experiment. The observer task is to detect whether the probe contains a horizontal or vertical edge.

3.2 Experiment 3: Extent

If the adaptation pool has a certain size, the most obvious stimulus to determine that size, also used before [Hood et al. 1979], is a disk of varying diameter. Figure 9-top shows such disks used in our third experiment. The disk had a luminance of either 50 or 2500 cd/m^2 and the background was set to 5 cd/m^2 . Figure 9-bottom shows that the detection thresholds, and hence the adaptation luminance, levels off around a diameter of 0.5° of visual angle, which is smaller than the 1° extent used in most ad hoc models but larger than the extent of about 0.1° proposed by Wilson [1997] based purely on retinal physiology. The significant drop in the detection thresholds is strong evidence that the adaptation pools have limited size, which is larger than the size of a single photoreceptor but probably smaller than 0.5° .

It is important to note that our experiments are meant to determine the effective size of the area contributing to the adaptation field, but not necessarily the exact size of the biological adaptation mechanism. In particular, the effective size of the adaptation area is larger for our data due to eye movements and optical glare.

3.3 Experiment 4: Long-range effects

The visual field in the direct vicinity of the detected target is likely to have the strongest influence on the state of adaptation. However, adaptation can be also affected by more distant parts of the visual field. Such long-range effects can be much weaker and they cannot be detected in the presence of a strong adaptation field close to the detection target, such as in the case of the disks used in Experiment 3.

To measure such likely, long-range effects, a 0.2° diameter pedestal of 2500 cd/m^2 was surrounded by a concentric ring of the same luminance with varying inner and outer diameters. Different configurations of inner and outer diameters are listed in Table 1. Three different groups of rings were tested: a) rings with a fixed area; b) rings with a fixed outer diameter; and c) rings of which the area increased with the inner diameter to compensate for the weaker effect of more distant regions. The background luminance was set to 5 cd/m^2 .

Figure 10 shows that the long-range effect can be observed up to about 5° . The presence of a ring at that distance from the center lowers the threshold as compared to the case where there is no ring (dashed line). However, the strength of that effect is rather limited. The thresholds drop from 0.3 to about 0.16 , while the close-range adaptation field can reduce the thresholds to 0.05 (refer to Figure 9). The short-range effects are clearly dominant.

3.4 Experiment 5: Non-linear pooling

The two previous experiments measured adaptation pooling as a function of distance from the fixation point. However, they cannot

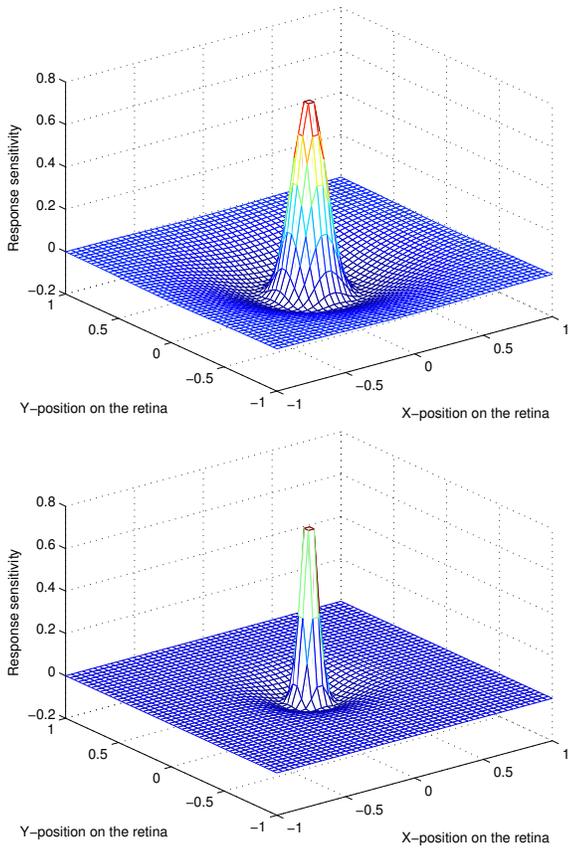


Figure 7: The idealized response sensitivity of the receptive fields in the retina for the two receptive fields tuned to different size and spatial frequency. The high positive values of response sensitivity mean that those receptors at those locations will contribute positive signal, and negative values that those receptors will inhibit positive signal. This way the receptive fields respond to important features, such as edges, rather than to uniform fields. Receptive fields can overlap and vary in size.

explain what kind of non-linearity is involved: pooling might occur in linear (luminance) space, in logarithmic space, or in any other non-linear space. To determine this non-linearity, we flanked the detection target in the center with rings or half rings of different luminance, but fixed outer diameter of 1° . The luminance of the ring varied from 0.5 to 5000 cd/m^2 . The half ring was cut diagonally to reduce any possible interference with the vertical or horizontal detection target. The background was fixed at 0.5 cd/m^2 . Examples of such stimuli surrounded by rings are shown in Figure 11-top.

If the adaptation pooling mechanism operates in the linear domain, the luminance from both halves of the half ring should be summed linearly. Therefore, the detection threshold for a half ring of luminance 99.5 cd/m^2 should be the same as the detection threshold for a full ring of luminance 50 cd/m^2 since the mean of 99.5 and 0.5 is 50 . The results in Figure 11-bottom show that this is clearly not the case; the threshold for the full ring at 50 cd/m^2 is substantially lower than for the half-ring at 99.5 cd/m^2 . If the adaptation pooling operates in the logarithmic domain, the detection threshold for a half ring at 5000 cd/m^2 should be equal to the detection threshold for a full ring at 50 cd/m^2 since the logarithmic mean of 0.5 and 5000 is 50 . Again, the results do not confirm this hypothesis.

One salient feature of the results shown in Figure 11 is that the effect of adaptation is asymmetric for lower and higher luminance of the

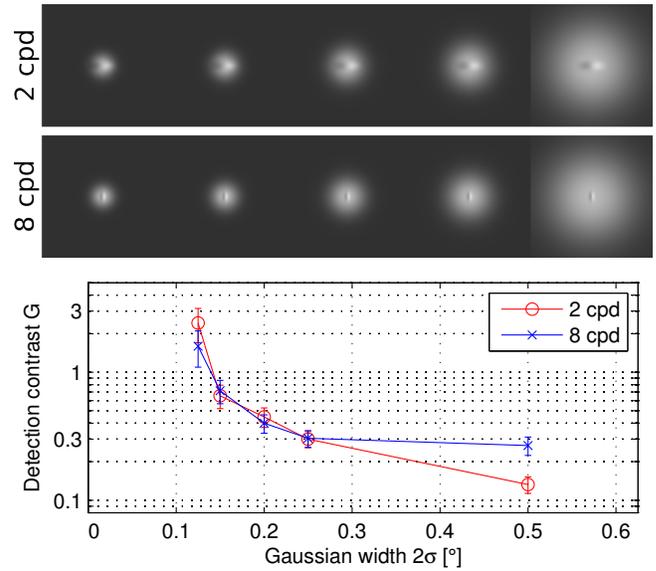


Figure 8: Top: The stimuli used in the Experiment 2: frequency selectivity. Bottom: the results of that experiment.

Table 1: Specifications of the rings in Experiment 4.

Area [$^\circ^2$]	Inner diameter [$^\circ$]	Outer diameter [$^\circ$]	
8.04	0.5	1.68	} fixed area
8.04	1	1.89	
8.04	3	3.4	
8.04	4	4.31	
8.04	6	6.21	} fixed outer diameter
42.60	5	6.21	
92.87	3	6.21	
118.00	1	6.21	
0.19	0.3	0.39	} area increases with diameter
0.27	0.39	0.49	
0.38	0.49	0.60	
0.58	0.60	0.74	
1.03	0.74	0.93	
3.38	0.93	1.40	

half-ring. This is evidence of the strong effect of glare. However, the exact form of the non-linearity is difficult to determine without considering other elements of the adaptation model.

3.5 Experiment 6: Orientation and contrast masking

The sharp contrast edge between the pedestal and background not only changes adaptation luminance, but it also creates a strong masking signal. Contrast masking is usually associated with lower visibility, or higher detection thresholds, for patterns that are shown on top of another pattern. For example, a band-limited noise pattern is more difficult to detect when shown on top of another noise pattern of similar spatial frequency, as shown in Figure 3. Contrast masking, however, also affects patterns that are in the vicinity of a masking signal, especially for lower frequencies. A strong edge between the pedestal disk and background in our experiments will create a masking signal in all frequencies, some extending over the region where the edge to be detected is located. Therefore, masking could explain a higher threshold for detecting our probe if the pedestal edge is close to the detected edge.

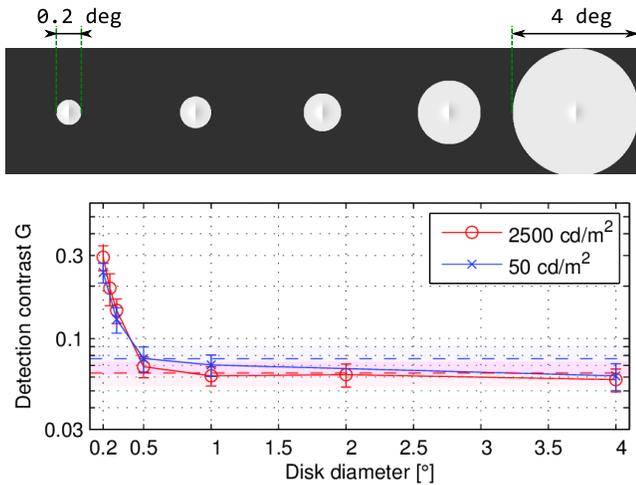


Figure 9: Top: Stimuli used in the experiment 3: extent. The edge in the centre was briefly shown for 200 ms. The observer task was to detect whether the edge was vertical or horizontal. Bottom: Results of that experiment. The detection thresholds for targets on pedestals of different diameters levels off around 0.5° of visual angle. The horizontal dashed lines indicate the detection threshold for a uniform field (or pedestal that covers the entire screen), which was measured in Experiment 1.

Contrast masking is the strongest when the masking signal has the same frequency and orientation as the detected (masked) signal. We use this property to isolate the effect of masking from the effect of local adaptation. We generate two pairs of stimuli, in which the edge between the pedestal and background either has the same orientation as the detected edge in the center (Figure 12-left), or is rotated by 45° relative to the detected edge (Figure 12-right). The area of the squares is identical in both cases. Therefore, if local adaptation is the dominant effect, thresholds for both orientations should be the same. However, if the elevated thresholds are due to masking, the thresholds should be higher when the pedestal edge is aligned with the detected edge. This experiment was also meant to confirm the radially symmetric characteristic of the pooling we assumed in all other experiments.

The bright squares were 2500 cd/m^2 , the dark squares in the checkerboards were 1 cd/m^2 , and the background was 5 cd/m^2 . The squares of the checkerboards had side length 0.2° .

The results in Figure 12 indicate little difference between the two orientations of the pedestal. This shows no evidence to support the hypothesis that the elevated thresholds are caused by contrast masking. Radially symmetric pooling was also observed in the context of lightness perception [Allred et al. 2012].

3.6 Experiment 7: Mondrian and complex images

To enrich the dataset with more real-life adaptation patterns, we also measured detection thresholds for more complex scenarios in which we did not try to isolate any effects.

The first set of images contained a Mondrian-style pattern of square patches of side length 2° with exponentially distributed luminances from 0.25 to 5000 cd/m^2 , roughly corresponding to a uniform distribution of perceived brightness. The detection target was placed at 9 different positions on a central patch of 2500 cd/m^2 , numbered in Figure 13.

The results in Figure 13 show a moderate variation in detection thresholds between 0.06 and 0.1 . The thresholds are the highest

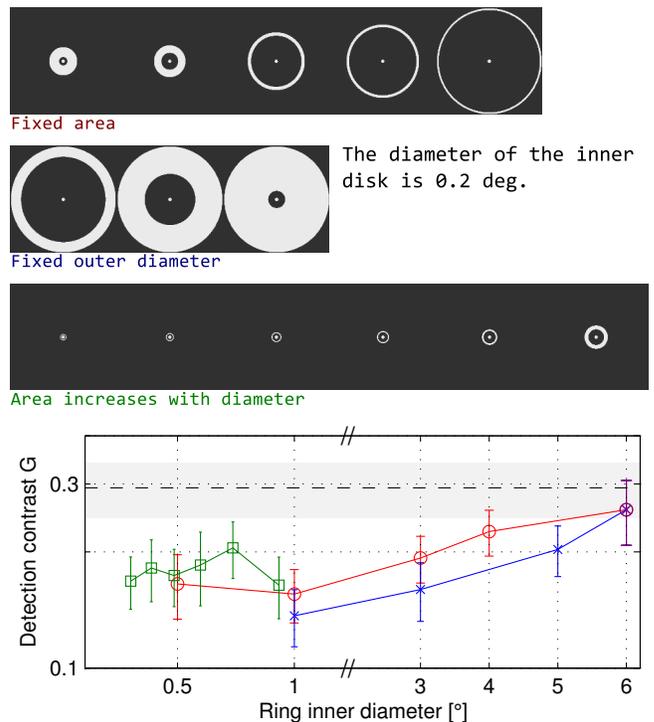


Figure 10: Top: Stimuli used in the experiment 4: long-range effects. Refer to Table 1 for detailed specification of three types of stimuli: fixed area, fixed outer diameter and the area increases with the diameter. Bottom: The detection thresholds for targets surrounded by a ring with different inner diameters, using the same color coding as Table 1. The horizontal dashed black line and gray band indicate the detection threshold and SEM for a pedestal without a ring (from Experiment 3).

for locations 3, 6, and 9 on the right edge of the central square, which borders a much darker square on the right. These results are consistent with our previous observations, in which a much darker area near the detection target elevates detection thresholds.

The second set of stimuli consisted of 4 natural images from the HDR Photographic Survey [Fairchild 2008] in which the detection target was positioned to maximize maladaptation. The images and the positions of detection targets are shown in Figure 14.

As shown in Figure 14, the threshold was the highest for the *Sunrise* image in which the detection target was located on the very bright disk of the sun. This shows that detection thresholds are higher when the pedestal is smaller or when the luminance difference with the surround is larger.

4 OTF model

Our results indicated that the best fit to the local adaptation data is achieved when the optical transfer function (OTF) of Deeley et al. [1991] is used to model optical glare in the eye. Such modeling is performed by multiplying the input luminance map L with the OTF M in the Fourier domain:

$$\mathcal{F}\{L_O\}(u, v) = \mathcal{F}\{L\}(u, v) M(\sqrt{u^2 + v^2}) \quad (2)$$

where \mathcal{F} is the Fourier transform, u and v are horizontal and vertical spatial frequency respectively, and the OTF M is given by:

$$M(\rho) = \exp\left(-\frac{\rho}{(20.9 - 2.1\rho)^{1.3 - 0.07\rho}}\right), \quad (3)$$

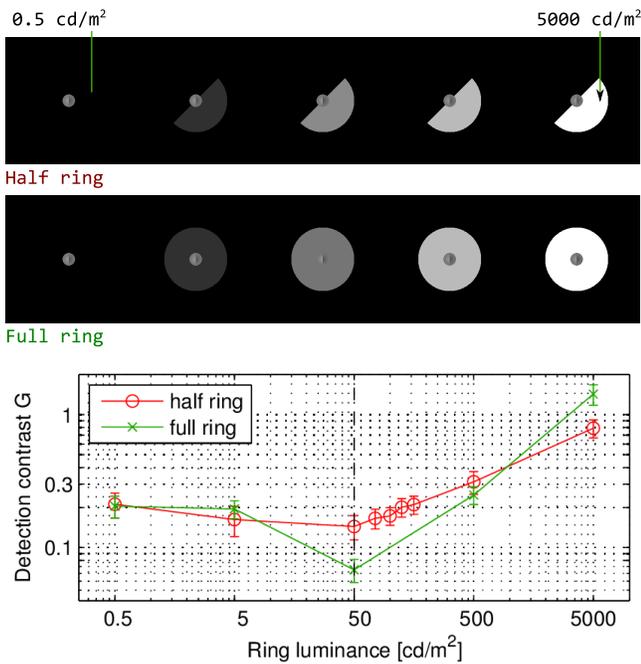


Figure 11: Top: the stimuli used in the Experiment 5: non-linear pooling. Bottom: The detection thresholds for that experiment. The pedestal luminance (the small disk in the center) is indicated with a vertical dashed line.

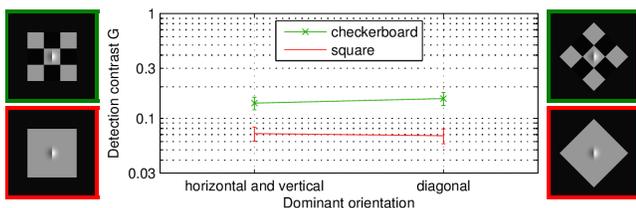


Figure 12: The detection thresholds for targets embedded in squares or checkerboards with different orientations and luminances.

where ρ is the spatial frequency in cycles per degree and p is the pupil diameter (assumed to be 4 mm in our model).

References

- ALLRED, S. R., RADONJIĆ, A., GILCHRIST, A. L., AND BRAINARD, D. H. 2012. Lightness perception in high dynamic range images: Local and remote luminance effects. *Journal of Vision* 12, 2, 7.
- BARTEN, P. G. J. 1999. *Contrast sensitivity of the human eye and its effects on image quality*. SPIE Press.
- DALY, S. 1993. The Visible Differences Predictor: An Algorithm for the Assessment of Image Fidelity. In *Digital Images and Human Vision*, A. B. Watson, Ed. MIT Press, 179–206.
- DEELEY, R. J., DRASDO, N., AND CHARMAN, W. N. 1991. A simple parametric model of the human ocular modulation transfer function. *Ophthalmic and Physiological Optics* 11, 1 (Jan.), 91–93.
- DUNN, F. A., LANKHEET, M. J., AND RIEKE, F. 2007. Light adaptation in cone vision involves switching between receptor and post-receptor sites. *Nature* 449, 7162 (Oct.), 603–6.

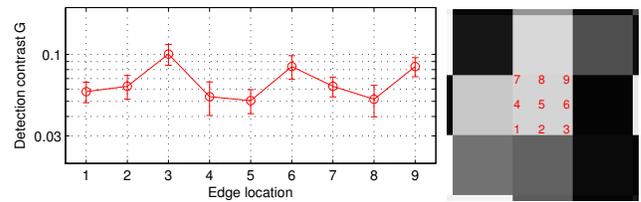


Figure 13: The detection thresholds for targets placed at different positions on the central patch in a Mondrian-style pattern as shown on the right.

- FAIRCHILD, M. D. 2008. *The HDR Photographic Survey*. MDF Publications. <http://rit-mcsl.org/fairchild/HDR.html>.
- GEORGEON, B. Y. M. A., AND SULLIVAN, G. D. 1975. Contrast constancy: deblurring in human vision by spatial frequency channels. *The Journal of Physiology* 252, 3, 627–656.
- HOOD, D. C., FINKELSTEIN, M. A., AND BUCKINGHAM, E. 1979. Psychophysical tests of models of the response function. *Vision Research* 19, 401–406.
- KIM, K. J., MANTIUK, R., AND LEE, K. H. 2013. Measurements of achromatic and chromatic contrast sensitivity functions for an extended range of adaptation luminance. In *Human Vision and Electronic Imaging*, 86511A.
- MANTIUK, R., KIM, K. J., REMPEL, A. G., AND 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 (July), 40:1–40:13.
- NAKA, K. I., AND RUSHTON, W. A. H. 1966. S-potentials from luminosity units in the retina of fish (Cyprinidae). *Journal of Physiology* 185, 587–599.
- SHAPLEY, R., AND ENROTH-CUGELL, C. 1984. Chapter 9 Visual adaptation and retinal gain controls. *Progress in Retinal Research* 3 (Jan.), 263–346.
- VANGORP, P., MYSZKOWSKI, K., GRAF, E. W., AND MANTIUK, R. K. 2015. A model of local adaptation. *ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH Asia 2015)* 34, 6.
- WATSON, A., AND AHUMADA, JR., A. J. 2005. A standard model for foveal detection of spatial contrast. *Journal of Vision* 5, 9, 717–740.
- WATSON, A. B., AND PELLI, D. G. 1983. QUEST: a Bayesian adaptive psychometric method. *Perception & Psychophysics* 33, 2, 113–120.
- WATSON, A., AND SOLOMON, J. 1997. Model of visual contrast gain control and pattern masking. *Journal of the Optical Society of America A* 14, 9, 2379–2391.
- WILSON, H. R. 1980. A transducer function for threshold and suprathreshold human vision. *Biological Cybernetics* 38, 3 (Oct.), 171–178.
- WILSON, H. R. 1997. A neural model of foveal light adaptation and afterimage formation. *Visual Neuroscience* 14, 03 (June), 403–423.

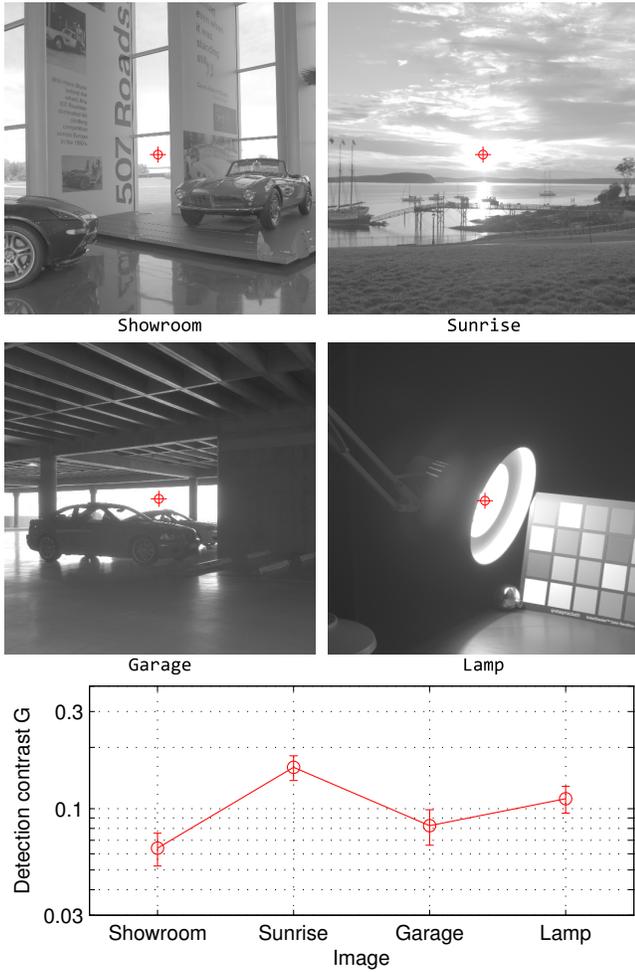


Figure 14: Top: The four images used in an experiment with natural images. The crosses show where the detection target was located in the image. Bottom: The detection thresholds for those detection targets.