

# Spatial long-range modulation of contrast discrimination

Chien-Chung Chen<sup>a\*</sup> & Christopher W. Tyler<sup>b</sup>

<sup>a</sup> Ophthalmology Department, University of British Columbia, Vancouver, British Columbia, Canada

<sup>b</sup> Smith-Kettlewell Eye Research Institute, San Francisco, California, USA

## ABSTRACT

Contrast discrimination is an important type of information for establishing image quality metrics based on human vision. We used a dual-masking paradigm to study how contrast discrimination can be influenced by the presence of adjacent stimuli. In a dual masking paradigm, the observer's task is to detect a target superimposed on a pedestal in the presence of flankers. The flankers (1) reduce the target threshold at zero pedestal contrast; (2) reduce the size of pedestal facilitation at low pedestal contrasts; and (3) shift the TvC (Target threshold vs. pedestal contrast) function horizontally to the left on a log-log plot at high pedestal contrasts. The horizontal shift at high pedestal contrasts suggests that the flanker effect is a multiplicative factor that cannot be explained by previous models of contrast discrimination. We extended a divisive inhibition model of contrast discrimination by implementing the flanker effect as a multiplicative sensitivity modulation factor that account for the data well.

**Keywords:** image quality assessment, human vision, dual masking, divisive inhibition

## 1. Introduction

### 1.1 Contrast discrimination and the divisive inhibition models

A good image quality metrics should assess the quality of an image in a way that is consistent with human visual experience. Thus, much effort has been expended in incorporating human visual psychophysics data or models based on the human visual system into image quality metrics<sup>5,19,22,22,24</sup>. The most relevant human psychophysical data for image quality assessment are contrast discrimination thresholds. In a typical contrast discrimination experiment, the task of an observer is to detect a periodic pattern (target) superimposed on another periodic pattern (pedestal) with the same spatiotemporal properties except contrast. The contrast discrimination threshold, or target threshold, is defined as the target contrast that allows the observer to tell the difference between pedestal-alone and pedestal-plus-target with certain percentage of correctness.

A typical result of contrast discrimination experiments is the "dipper" shaped *target contrast vs. pedestal contrast* (TvC) function<sup>6,7,10,16,25</sup>. That is, as pedestal contrast increases, the target contrast threshold first decreases (facilitation) and then increases (masking) as shown in figure 1a. The TvC function reflects the contrast response characteristics in the visual system. As shown in Figure 1, in order to be detected, the target has to have enough contrast to increase the response to the pedestal alone by a certain amount, defined as one unit. Suppose that the response function is accelerating near a pedestal contrast (e.g.,  $C_1$  in Figure 1), it would require less target contrast ( $\Delta C_1$ ) than the unmasked threshold ( $C_0$ , the threshold measured without the pedestal) to increase the response by one unit. On the other hand, when the response function is decelerating near a pedestal contrast (e.g.,  $C_2$  in Figure 1), it would take greater target contrast ( $\Delta C_1$ ) to increase the response by the same amount. Thus, the target threshold at a pedestal contrast is inversely proportional to the slope of the response function at that pedestal contrast.

Currently, the most popular model of contrast response functions is the divisive inhibition model, also called the contrast normalization model<sup>1,3,4,6,8,16,17,19,23</sup>. Different variations of this model have been used in several image quality metrics. All the variations of the divisive inhibition models share the following common features:

---

\* Correspondence: E-mail [chen@ski.org](mailto:chen@ski.org); Telephone: 1(604)6097225

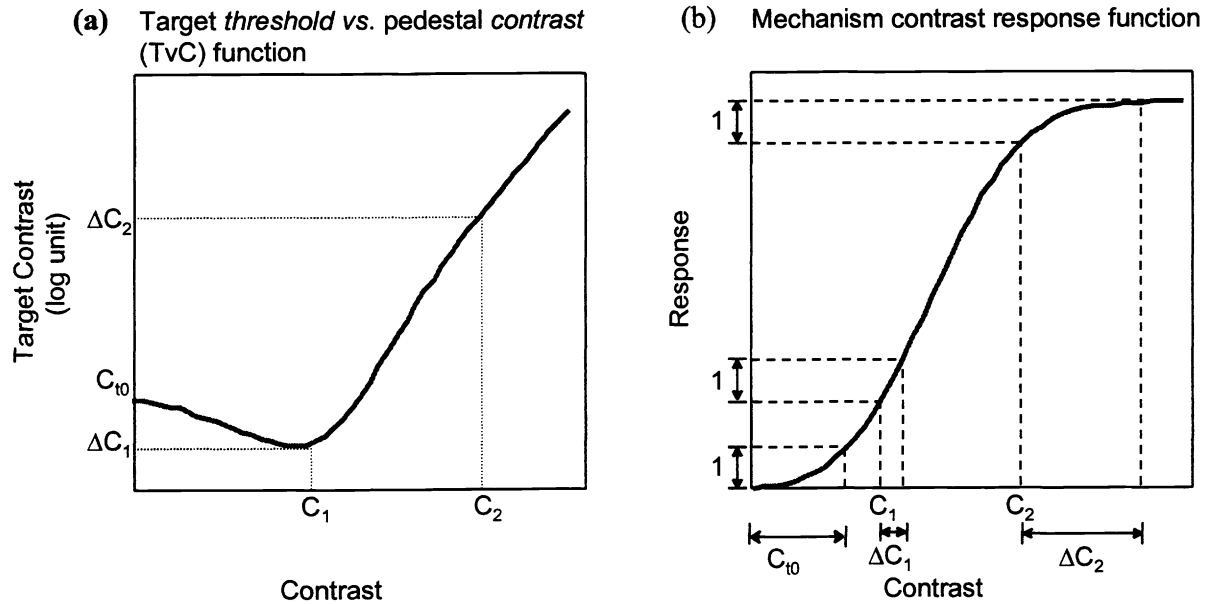


Figure 1. The relationship between the target threshold vs. pedestal contrast (TvC) function (left panel) and the contrast response function (right panel).  $C_{t0}$  is the target absolute (unmasked) threshold.  $\Delta C_1$  and  $\Delta C_2$  are target thresholds measured with the presence of pedestal contrast  $C_1$  and  $C_2$  respectively. At threshold, the target increases the response to pedestal alone by one unit. As a result, the target threshold is inversely proportion to the slope of the response function.

- (1) Linear filters. The input images are processed by a band of linear filters defined by wavelet functions such as Gabor or difference-of-Gaussian. Those linear filters have a limited spatial extent and are localized in both space and the Fourier domain. The output of each linear filter can be simplified as the contrast of the Fourier component to which it is tuned weighted by a constant<sup>4</sup>.
- (2) Divisive inhibition. Each linear filter is followed by a nonlinear response operator. Each nonlinear operator takes two inputs: i) the excitatory input fed directly from the corresponding linear filters and ii) the divisive inhibitory input which is pooled from the outputs of all relevant linear filters (normalization pools). The response of each nonlinear operator is the excitatory input raised by a power and then divided by the divisive inhibitory input plus a constant.
- (3) Decision. In a two-alternative forced-choice contrast discrimination experiment, the observer compares the nonlinear operator responses to the pedestal-alone and to the pedestal-plus-target. The decision is based on the operator that shows the greatest difference between the responses. The target is at threshold when the difference reaches a criterion.

## 1.2 Flanker effect

Notice that the divisive inhibition models are based on localized mechanisms. These models were designed to account for the contrast discrimination experiments where only the targets and the pedestals were presented and they did not take the spatial context of the stimuli into account. However, recent studies have shown that spatial context does affect human visual behavior. For instance, Polat & Sagi<sup>14,15</sup> measured detection thresholds for a target Gabor pattern at the fovea flanked by two other high contrast Gabor patterns (flankers). The target threshold decreased up to about 50% of the absolute threshold (facilitation) when a pair of collinear flankers (where the flanker had the same orientation as the target) presented at about three units of target wavelength away. Conversely, flankers with an orientation that was orthogonal to the target had no effect on target detection. This control result establishes that the effects of the flankers are not generic attention or uncertainty effects but are local or long-range interactions specific to the receptive field and orientation selectivity.

Current divisive inhibition models can deal with the flanker effect in two ways: (1) Increase the size of the linear filters such that the target and the flankers can be covered by the same filter. Thus, the presence of flankers can produce a response in the target mechanism and in turn affect the target threshold<sup>18</sup>. (2) Increase the extent of the contrast normalization pools<sup>17</sup>. Thus, the presence of flankers just adds an extra term in the contrast normalization signal. In either way, the effect of the flanker is an additive term to the effect of the target as proposed both by Snowden & Hammet<sup>17</sup> and Solomon et al.<sup>18</sup>. However, as we will show in this study, the flanker effect is not additive but multiplicative. Our experiment shows that the current divisive inhibition models cannot account for contrast discrimination thresholds in the presence of flankers.

### 1.3 Dual Masking Paradigm

Consider a dual masking experiment in which the task of the observer is to detect a target pattern superimposed on a pedestal (first masker) and in the presence of flankers (second masker). When the pedestal and the target have the same spatial temporal properties except contrast, this dual making experiment is equivalent to a contrast discrimination experiment conducted in the presence of flankers. Thus, if we systematically measure the target contrast on various pedestal contrasts, we can obtain a contrast discrimination TvC function under the influence of flankers.

Figure 2 shows the prediction of current divisive inhibition model to the flanker effect. The solid curve is the contrast discrimination TvC function measured without the presence of flankers. When the flankers are presented, according to the divisive inhibition models, their effects are added to that of the pedestal. Suppose that the flanker contrast is a constant through the experiment. The flanker effect should also be a constant. On the other hand, the effect of pedestal increases with its contrast. Thus, comparing to the pedestal effect, the flanker effect should be less and less salient as the pedestal contrast increases. Therefore, the divisive inhibition model should predict that the TvC functions with the presence of the flankers should converge to the TvC functions without the flankers as pedestal contrast increases. A detailed discussion of the divisive inhibition model predictions can found in Snowden & Hammet<sup>17</sup> in a slightly different context. The dashed curve in Figure 2 illustrates such prediction. The experiment report below will show this prediction fails to capture the characteristics of the data.

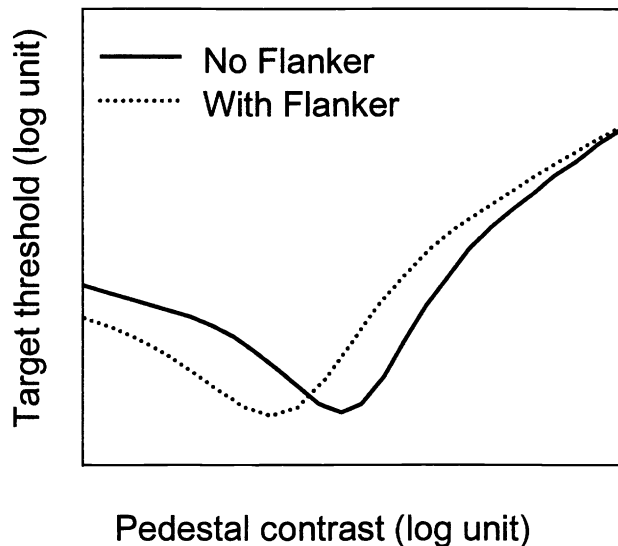


Figure 2. The divisive inhibition model prediction on flanker effects. The solid curve is the TvC function measured without the presence of flankers and dotted curve is that measured with the presence of flankers. The divisive inhibition model predicts that the two TvC functions are getting closer to each other as pedestal contrast increases.

The dashed curve in Figure 2 illustrates such prediction. The experiment report below will show this prediction fails to capture the characteristics of the data.

## 2. Methods

### 2.1 Apparatus

The stimuli were presented on two SONY CPD-1425 monitors each driven by a Radius PrecisionColor graphic board. A Macintosh Quadra Pro computer controlled the graphic boards. The resolution of the monitor was 640 horizontal by 480 vertical pixels. At the viewing distance we used (128 cm), there were 60 pixels per degree. The viewing field was then 10.7° (H) by 8° (V). The refresh rate of the monitor was 60 Hz. We used the LightMouse photometer (Tyler, 1997) to measure the input-output intensity function of the monitor. This information allowed us to compute linear lookup table settings. The mean luminance of the monitor was set at 35 cd/m<sup>2</sup>.

### 2.2 Stimuli

The target, the pedestal and the flankers were all vertical Gabor patches defined by the equation

$$G(x,y) = BG + BG * C * \cos(2\pi kx) * \exp(-x^2/2\sigma^2) * \exp(-(y-u_y)^2/2\sigma^2)$$

where BG was the mean luminance, C, ranged from 0 to 1, was the contrast of the pattern, k was the spatial frequency,  $\sigma$  was the scale parameter (standard deviation) of the Gaussian envelope and  $u_y$  was the displacement of the pattern. All patterns had a spatial frequency (k) of 4 cycles per degree and a scale parameter ( $\sigma$ )  $0.1768^\circ$ . The target and the pedestal were centered at the fixation point, therefore the displacement  $u_y$  was zero. The two flankers were placed at the top and below the target. The displacement was  $0.75^\circ$ . The stimuli were presented concurrently. The temporal waveform of the stimulus was a pulse with duration of 100 msec. The contrast of the patterns is presented in decibels (dB), which is 20 times log base 10 of the linear contrast.

### 2.3 Procedures

We used a temporal two-alternative forced-choice paradigm to measure the target threshold. In each trial, the pedestal and the flankers were presented at both intervals. The target was randomly presented in either one of the intervals. The task of the observer was to determine which interval contains the target. We used the Psi adaptive threshold seeking algorithm<sup>9</sup> to measure the threshold. The experimental control software was written in MATLAB<sup>12</sup> using the Psychophysics Toolbox<sup>2</sup>, which provides high level access to the C-language VideoToolbox<sup>13</sup>.

The target contrast threshold was measured upon several pedestal contrasts ranged from -34dB (2%) to -6 dB (50%). On each trial, the two flankers always had the same contrast. The contrast of the flanker was either 50% (-6dB) or 0%. Each target threshold was measured at least four times for each observer. The thresholds reported here are average of those repeated measurements.

Two observers served in this study. CCC (male, early 30s) is an author of this paper. MDL (female, late 20) is a paid observer naïve to the purpose of the experiment. MDL has a normal and CCC has a corrected-to-normal visual acuity (20/20).

### 3. Results

Figure 3 shows the result from one of the observer. The smooth curves are fits of the sensitivity modulation models discussed below. The closed circles and solid curves show the TvC function measured without the presence of the flankers. When the flanker is absent, the TvC function shows the typical "dipper" shape commonly seen in the spatial contrast discrimination literature<sup>6,7,10,16,25</sup>. That is, the target threshold first decreases and then increases with pedestal contrast. The greatest threshold decrement occurs when the pedestal contrast is at about its own detection threshold. A particularly robust facilitation effect of -9 db is seen in this example.

The open circles and dashed curve show the TvC function measured in the presence of -6dB (50%) flankers. The flankers have three major effects on TvC functions. First, when there was no pedestal (denoted as  $-\infty$  dB contrast pedestal in Figure 3), the flankers reduced the target threshold.

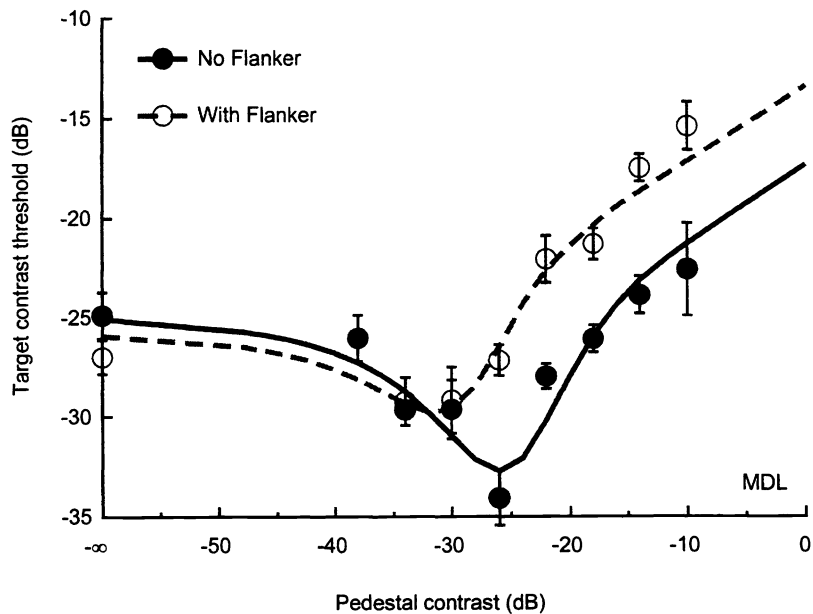


Figure 3. Flanker effect on TvC functions for observer MDL. The closed curves and solid curve denote the TvC functions measured with no flanker presented. The open circles and the dotted curve denote the TvC function measured with the presence of -6 dB or 50% contrast flankers.

This is the lateral masking effect reported by Polat & Sagi<sup>14,15</sup>. Second, as the pedestal contrast increased, the target threshold did not show as much decrement as in the case of no flankers. There was little, if any, low pedestal contrast dip when the flankers were presented. Third, at high pedestal contrasts, the flankers increased the threshold at every pedestal contrast evenly. This effect can be viewed as shifting the TvC function horizontally to left. The two TvC functions are parallel to each other up to the highest pedestal contrast available from our apparatus. There is no sign of merging between the TvC functions measured with and without flankers presenting. The data of different observers are consistent with this result. Thus, the data reject the divisive inhibition models, which predict the two TvC functions should merge together at high pedestal contrast (see section 1.3 and Figure 2). For data shown in Figure 3, the divisive inhibition models can underestimate the target by about 6dB or 2-fold linear contrast unit.

## 4. Modeling

### 4.1 The sensitivity modulation model

In Figure 3, the data are plotted in log-log coordinates. The flankers shift the TvC functions horizontally on these coordinate, suggesting that the flanker effect is multiplicative. Based on this idea, we propose a sensitivity modulation model to account for the flanker effect.

Figure 4 shows a diagram of this model. Our model does share the same linear filter assumption as most models of contrast discrimination. That is, the visual system contains a band of localized linear filters each responds to a Fourier component of input images. Each filter has a limit extent in the space domain and bandwidth in Fourier domain. The output of each linear filter (called excitation, denoted as  $E$  in Figure 4) is the contrast of the input image weighted by a number. The weight is called the sensitivity of the filter to the image and is determined by the cross-correlation of the spatial profile of the filter and the image. A nonlinear response operator follows each linear filter. For a reason that will be obvious shortly, this direct link between the linear filter and nonlinear operator is called the excitatory input to the nonlinear operator. In addition to the excitatory input, the response of the nonlinear operator is also influenced by the other linear filters with a receptive field covering the same spatial location but tuned to different Fourier components of the input image (inside the dotted box in Figure 4). A pooling process combines the excitations of all relevant linear filters together to form the divisive inhibition signal, denoted as  $I$  in Figure 4. Mathematically, the pooling is done by summing the excitations of relevant filters raised by a power  $q$ . When the flanker is not presented, the response of the nonlinear operator is simply the excitation from the excitatory input ( $E$ ) raised by a power  $p$  and then divided by the divisive

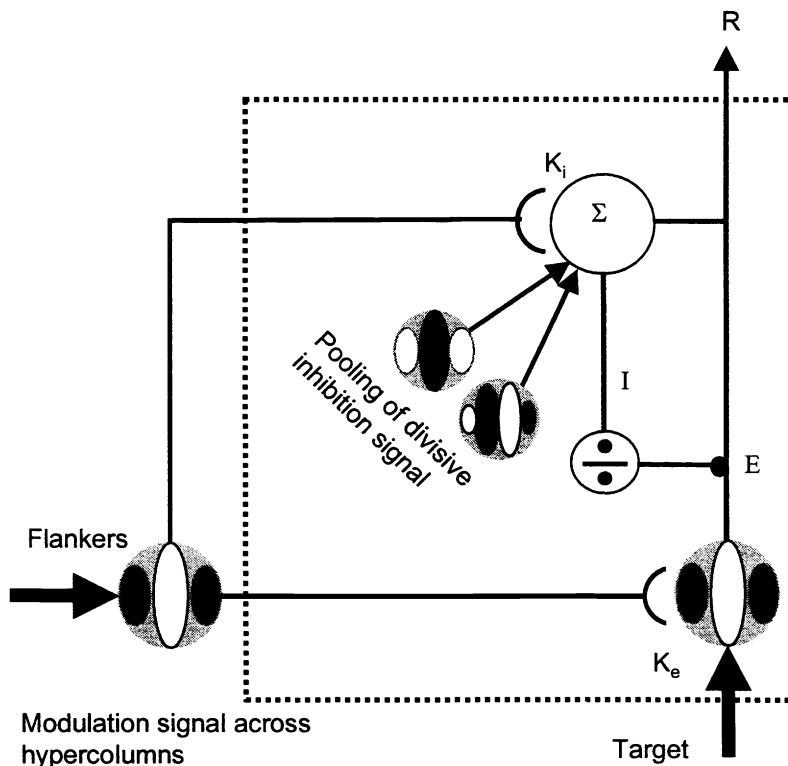


Figure 4. A diagram of the sensitivity modulation model. Inside the dotted box, all linear filters respond to image components presented at the same location. Their behavior is described by the divisive inhibition models. The initial excitation ( $E$ ) of a linear filter is the contrast of the target pattern weighted by the filter's sensitivity to that pattern. The initial excitations of all relevant filters are pooled together to form the divisive inhibitory signal ( $I$ ). The final response is the initial excitation raised by a power and then divided by the normalization signal plus a constant. The flanking filters send signals that change the sensitivities of the contacted filters. See text for further details.

inhibitory input ( $I$ ) plus an additive constant. That is,

$$R = \frac{E^p}{I + \sigma} \quad (1)$$

where  $\sigma$  is an additive constant. This same-location operation is the same as that proposed by the divisive inhibition model.

In contrast to the divisive inhibition model, the flanking linear filters, however, do not contribute anything to the pooling of the divisive inhibitory signal. Instead, their effect is to change the sensitivities of the linear filters located in the same space as the target as well as the mechanism that pools the divisive inhibitory signals. That is, when the flanker is presented, the excitation of the same-location linear filters is the no-flanker excitation multiplied by a factor. After an algebraic operation, the response of the target mechanism can be written as

$$R = \frac{k_e * E^p}{k_i * I + \sigma} \quad (2)$$

where  $k_e$  and  $k_i$  are multiplicative constants dependent on the flanker contrast. When the flanker is not presented,  $k_e = k_i = 1$ . In the presence of (-6dB) 50% contrast flanker,  $k_e$  and  $k_i$  are empirically determined as 1.52~2.63 and 1.92~4.09 respectively. The fits of the model are shown as smooth curves in Figure 3. The model captures all aspects of the flanker effects. The RMSE of the model is between 0.98~1.11, on a par with the standard error of measurement (0.92~1.06).

#### 4.2 Modulation factors

How can the two factors  $k_e$  and  $k_i$  explain the flanker effects? When the pedestal contrast is low, the linear filter excitation ( $E$ ) is dominated by the target contrast. At threshold, the divisive inhibitory input to the nonlinear operator ( $I$ ) is negligible comparing to the additive constant  $\sigma$ . The response function with flankers presented can be simplified as  $R = k_e * E^p / \sigma$ . Thus, a  $k_e$  larger than 1 will boost the response and make the target easier to be detected. This explains the lateral masking effect found by Polat & Sagi and the initial facilitation at lower end of the TvC functions.

As pedestal contrast increases, the divisive inhibitory inputs ( $I$ ) begin to catch up. Since  $k_i$  is larger than  $k_e$ , the flankers have a greater effect in the denominator of the response function than in the numerator. Therefore, the facilitation effect observed at low contrasts should decrease with the pedestal contrast. At medium contrasts, where the TvC function measured without the flanker presented shows a dip, the flankers produce less threshold reduction than at lower contrasts. Compared to the initial facilitation, the presence of the flankers has the effect of reducing, if not eliminating, the dip at medium contrasts. As pedestal contrast further increases, the flanker effect on the denominator of the response function eventually outweighs its effect on the numerator. The presence of the flanker is then increasing the target threshold rather than decreasing it. Finally, when the pedestal contrast is sufficiently high, the additive constant  $\sigma$  is negligible comparing to the inhibitory input  $I$ . Thus, we can simplify equation (1) as  $(E^p/I)$  and equation (2) as  $(k_e/k_i) * (E^p/I)$ . That is, the response function with the flankers presented is a constant times the response function without flankers. Translating the responses to thresholds, it gives the horizontal shift of TvC functions we observed on a log-log coordinate.

### 5. Conclusion

We found that the presence of flankers has the following effects on TvC functions. First, when there is no pedestal, the flankers reduce the target threshold. Second, the magnitude of the dip at low to medium pedestal contrast is greatly reduced. Third, at high pedestal contrasts, the flankers shift the TvC function horizontally to the left on a log-log coordinate. The two TvC functions are parallel to each other up to the highest pedestal contrasts available. Current divisive inhibition models cannot account for these effects. At some points, the discrepancy between the model and the data can be as large as 6dB or 2-fold. This discrepancy may affect the accuracy of an image quality metric. We refined the divisive inhibition models and propose a lateral sensitivity modulation model. This model takes the flanker effect as a multiplicative sensitivity modulator and captures all aspects of the flanker effect observed in our experiment.

## Acknowledgements

This study was supported by NIH grants EY7890 to CWT and a Rachel C. Atkinson Fellowship from the Smith-Kettlewell Eye Research Institute to CCC.

## References

1. Albrecht, D. G. & Geisler, W. S. (1991). Motion selectivity and the contrast response function of simple cells in the visual cortex. *Visual Neuroscience*, **7**, 531-546.
2. Brainard, D.H. (1997). The psychophysics toolbox. *Spatial Vision*, **10**, 433-436.
3. Carandini, M. & Heeger, D. J. (1994). Summation and division by neurons in primary visual cortex. *Science*, **264**, 1333-1336.
4. Chen, C. C., Foley, J. M. & Brainard, D. H. (2000). Detection of chromoluminance patterns on chromoluminance pedestals II: model. *Vision Research*, **40**, 789-803.
5. Daley, S. (1993). The visible differences predictor: an algorithm for the assessment of image fidelity quality. In A. B. Watson (ed.). *Digital images and human vision*. MIT Press, Cambridge, MA.
6. Foley, J. M. (1994). Human luminance pattern-vision mechanisms: Masking experiments require a new model. *Journal of the Optical Society of America A*, **11**, 1710-1719.
7. Foley, J. M. & Chen, C. C. (1997). Analysis of the effect of pattern adaptation on pattern pedestal effects: A two-process model. *Vision Research*, **37**, 2779-2788.
8. Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, **9**, 181-197.
9. Kontsevich, L. L. & Tyler, C. W. (1999) Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, **39**, 2729-2737.
10. Kontsevich, L. L. & Tyler, C. W. (1999). Nonlinearity of near-threshold contrast transduction. *Vision Research*, **39**, 1869-1880.
11. Legge, G.E. & Foley, J.M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America*, **70**, 1458-1470.
12. MathWorks (1993). *Matlab*. Natick: The MathWorks Inc.
13. Pelli, D.G. (1997). The Video Toolbox software for visual psychophysics: Transform numbers into movies. *Spatial Vision*, **10**, 437-442.
14. Polat, U. & Sagi, D. (1993). Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments. *Vision Research*, **33**, 993-999.
15. Polat, U. & Sagi, D. (1994). The architecture of perceptual spatial interactions. *Vision Research*, **34**, 73-78.
16. Ross, J. & Speed, H. D. (1991). Contrast adaptation and contrast masking in human vision. *Proceedings of Royal Society London, Ser. B*, **246**, 61-69.
17. Snowden, R. J. & Hammett, S. T. (1998). The effects of surround contrast on contrast thresholds, perceived contrast and contrast discrimination. *Vision Research*, **38**, 1935-1945.
18. Solomon, J. A., Watson, A. B. & Morgan, M. J. (1999). Transducer model produces facilitation from opposite-sign flanks. *Vision Research*, **39**, 987-992.
19. Teo, P. C. & Heeger, D. J. (1994). Perceptual image distortion. *SPIE proceedings*, **2179**, 127-141.
20. Tyler, C. W. & McBride, B. (1997). The Morphonome image psychophysics software and a calibrator for Macintosh systems. *Spatial Vision*, **10**, 479-484
21. Watson, A. B. (1987). The cortex transform: Rapid computation of simulated neural images. *Computer Vision, Graphics, and Image Processing*, **39**, 311-327.
22. Watson, A. B. (1993). *Digital images and human vision*. Cambridge MA: MIT Press.
23. Watson, A. B. & Solomon, J. A. (1997). A model of visual contrast gain control and pattern masking. *Journal of the Optical Society A*, **14**, 2378 - 2390.
24. Watson, A. B., Taylor, M. & Borthwick, R. (1997). Image quality and entropy masking. *SPIE Proceedings*, **3016**, 2-12.
25. Wilson, H.R., McFarlane, D.K. & Philips, G.C. (1983). Spatial frequency tuning of orientation selectivity units estimated by oblique masking. *Vision Research*, **23**, 873-882.