A Model of Visual Masking for Computer Graphics

James A. Ferwerda, Cornell University *
Peter Shirley, University of Utah
Sumanta N. Pattanaik, Cornell University
Donald P. Greenberg, Cornell University

Abstract

In this paper we develop a computational model of visual masking based on psychophysical data. The model predicts how the presence of one visual pattern affects the detectability of another. The model allows us to choose texture patterns for computer graphics images that hide the effects of faceting, banding, aliasing, noise and other visual artifacts produced by sources of error in graphics algorithms. We demonstrate the utility of the model by choosing a texture pattern to mask faceting artifacts caused by polygonal tesselation of a flat-shaded curved surface. The model predicts how changes in the contrast, spatial frequency, and orientation of the texture pattern, or changes in the tesselation of the surface will alter the masking effect. The model is general and has uses in geometric modeling, realistic image synthesis, scientific visualization, image compression, and image-based rendering.

CR Categories: I.3.0 [Computer Graphics]: General;

Keywords: visual perception, masking, image quality, error metrics

1 Introduction

In "A Framework for the Analysis of Error in Global Illumination Algorithms", Arvo et. al. [1994] introduce a taxonomy for classifying sources of error in realistic image synthesis. They define three categories of error: errors in the input data caused by limitations in measurement or modeling; discretization errors introduced when analytical functions are replaced by finite-dimensional linear systems that can be computed; and computational errors that occur because the numerical precision of calculations is limited.

These errors can produce visual artifacts in synthetic images. *Faceting* due to tesselation of curved surfaces, *banding* caused by quantization, *aliasing* due to insufficient sampling, and *noise* introduced by stochastic methods are all well known visual consequences of error in computer graphics algorithms. For many years graphics practitioners have observed that visual texture can mask these image artifacts.

*Program of Computer Graphics, 580 Rhodes Hall, Cornell University, Ithaca, NY 14853, USA. http://www.graphics.cornell.edu.

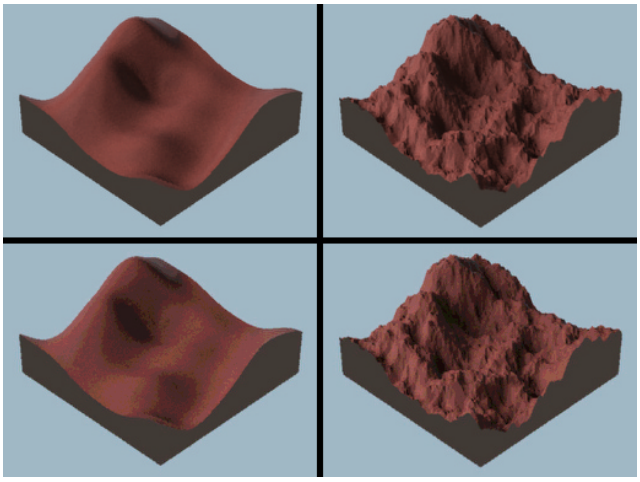

Figure 1: Masking in computer graphics: The upper pair of images are quantized to 8 bits. The lower pair are quantized to 4 bits. Banding is visible in the smooth surface on the lower left but not in the rough surface on the lower right due to masking effects from the visual texture created by the rough surface. From [Bolin95].

A recent example is shown in Figure 1 where banding due to quantization is much more apparent in the smooth surface on the left than in the rough surface on the right. Here the visual texture produced by the rough surface masks the banding artifact. With few exceptions [Bolin95, Mitchell87], the effects of masking have been applied in an ad hoc manner with unpredictable results.

Masking is a robust perceptual phenomenon that has been studied for more than 30 years by physiologists and psychologists. Masking was first observed in auditory perception [Fletcher52] but analogues in the visual domain were soon discovered [Campbell66, Pantle69]. Figure 2 from a classic study by Harmon and Julesz [1973] illustrates the characteristics of visual masking.

A continuous tone photograph of Abraham Lincoln was low-pass filtered to 10 cycles/picture height and then coarsely sampled and quantized to produce the image shown in Figure 2a. Notice how this processing seriously disturbs our ability to recognize the subject. If this blocky image is once again low pass filtered as in Figure 2b, recognition is restored. Thus it first appears that the image discontinuities introduced by high spatial frequencies in the block edges interfere with recognition. However Harmon and Julesz showed that it is not simply high frequencies that disturb recognition, but frequencies adjacent to the picture spectrum. They termed this *critical band masking*. Thus in Figure 2c where spatial frequencies above 40 cycles have been removed, the block edges are softened but recognition is still difficult. However in Figure 2d where frequencies between 12 and 40 cycles have been removed, the block edges are still

apparent, but the subject is identifiable. This shows that masking is caused by interactions within a limited spatial frequency band because removal of a critical band of frequencies directly adjacent to the picture's 10 cycle spectrum limit eliminates the masking effect but removal of higher frequencies does not.

In this paper we develop a computational model of visual masking derived from psychophysical experiments. The model predicts how the presence of one visual pattern affects the detectability of another. The model is general and has uses in geometric modeling, realistic image synthesis, scientific visualization, image compression, and image-based rendering. We demonstrate the utility of the model by predicting when a texture pattern will mask visual artifacts caused by tesselation of a flat-shaded curved surface. The model takes into account how changes in the contrast, spatial frequency, and orientation of the texture pattern, or changes in the tesselation of the curved surface will alter the masking effect.

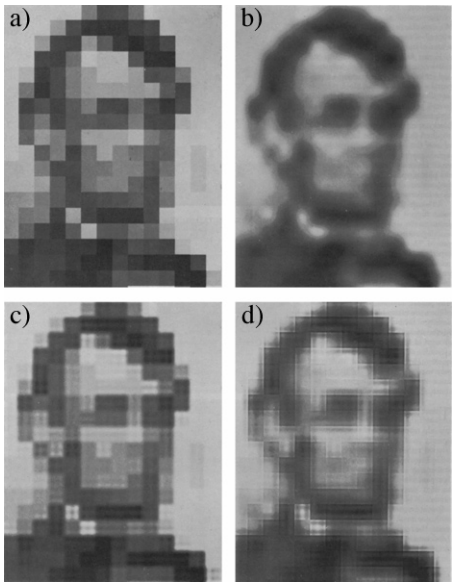

Figure 2: Demonstration of critical band masking. From [Harmon73].

2 Spatial Vision

Spatial vision is the field of psychology that studies of how patterns of light on the retina are interpreted by the visual system. The goal of the research in spatial vision is to understand the visual mechanisms that transform the patterns of light in the retinal image into the colors, sizes, shapes, locations, and motions of the three-dimensional objects we perceive in the world around us. The field has a long tradition which draws on both *physiological* studies of the electrical responses of cells in the visual pathways of primates and lower animals, as well as on *psychophysical* studies of the responses of human observers to simple visual stimuli.

2.1 Physiological foundations of spatial vision

One of the most fundamental physiological findings in the field of spatial vision is that the rod and cone photoreceptors which transduce light into electrical impulses in our nerve

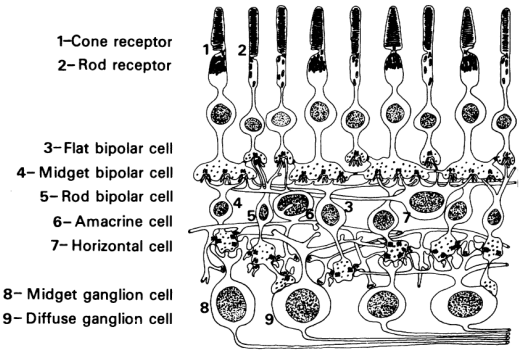

Figure 3: Neural networks in the primate retina: The rod and cone photoreceptors synapse on a variety of cells in the plexiform layers of the retina. These cells form networks which comprise the receptive fields of the retinal ganglion cells whose axons make up the optic nerve. From [Dowling66].

fibers are not independent of one another but interact in various ways. Figure 3 shows a diagram of a cross section through the retina. Amacrine, bipolar, and horizontal cells form neural networks in the plexiform layers of the retina that synapse on ganglion cells whose axons make up the optic nerve.

2.1.1 Receptive fields

To understand the properties of these neural networks, Kuffler [1953] made electrophysiological measurements of the responses of retinal ganglion cells in the cat. He found that each ganglion cell took its input from a spatially localized region of the retina called its *receptive field*.

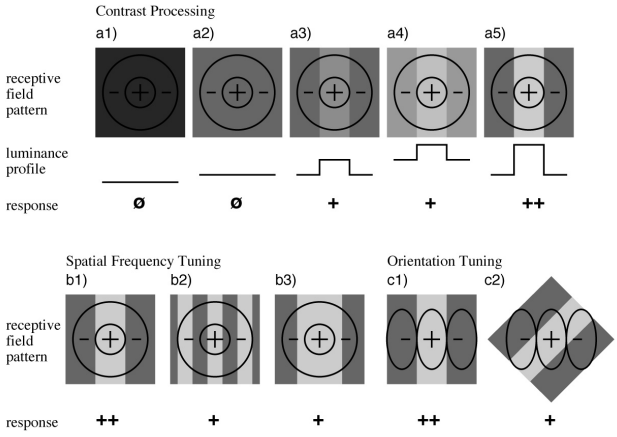

Figure 4: Properties of visual system receptive fields: (a) contrast processing; (b) spatial frequency tuning; (c) orientation tuning.

Kuffler found that these receptive fields had a characteristic center/surround organization with antagonism between the center and surround. Center/surround antagonism in receptive fields results in ganglion cells that respond primarily to contrast rather than to simple light intensity. This is illustrated in Figure 4a which shows the response of an idealized ganglion cell to various types of stimuli. In the dark, (Figure 4a1) the ganglion cell fires spontaneously at its base rate.

If the intensity of light falling on the ganglion cell's receptive field is raised uniformly, (Figure 4a2) the excitatory and inhibitory regions of the field cancel and the cell continues to fire at its base rate. If however, a bar pattern is introduced with contrast between the bar and the background (Figure 4a3), the excitation produced by the center will exceed the inhibition produced by the surround and the cell will increase its firing rate. Figures 4a4 and 4a5 show that the cell's response depends upon the contrast of the pattern rather than it's absolute intensity. In Figure 4a4 the luminance of the bar and background have both increased, but the cell continues to give the same response. However when the contrast between the bar and background is increased, (Figure 4a5) the response goes up as well.

Researchers have also found that different ganglion cells have receptive fields of different sizes and that these receptive fields overlap in the retina so that at any retinal location receptive fields of many sizes can be found.

2.1.2 Spatial tuning in receptive fields

Enroth-Cugell and Robson [1966] measured the response properties of retinal ganglion cells in the cat to sinusoidal grating patterns of different spatial frequencies. They found that the cells responded to limited ranges of spatial frequencies related to the sizes of their receptive fields. This spatial frequency selectivity of ganglion cell receptive fields is illustrated in Figure 4b.

The receptive field of the idealized ganglion cell has an excitatory center and inhibitory surround. If the receptive field is illuminated with the grating pattern shown in Figure 4b1 where the spatial frequency of the grating is such that the bars match the widths of the center and surround, there will be significant excitation from the center and not much inhibition from the surround so the cell will respond near its maximum rate. If however, we raise or lower the grating's spatial frequency as shown in Figures 4b2 and 4b3 there will be both less excitation from the center, and more inhibition from the surround so the cell will respond at a lower rate. The *spatial frequency tuning* of a cell depends upon the size of its receptive field. Cells with small receptive fields will respond to high ranges of spatial frequencies. Cells with larger fields will respond to lower ranges.

Although early studies focused on the response properties of cells in the retina, as more sophisticated electrophysiological techniques became available researchers began to investigate higher levels in the visual system including the visual cortex. These studies found that the receptive field organization first seen in the retina is in evidence throughout the visual system.

2.1.3 Orientation tuning in receptive fields

Hubel and Wiesel [1962,1968] did electrophysiological studies of cells in the visual cortex of the cat and monkey, mapping the properties of cortical receptive fields. They found that many cells responded maximally to patterns at a particular orientation and that response declined rapidly as the pattern was tilted away in either direction. Figure 4c1 shows an idealized receptive field for a cortical cell. The receptive field still has an antagonistic center surround organization, but the field is elongated in a particular direction. This elongation of the field accounts for the cell's orientation selectivity. If a grating pattern of the right spatial frequency and orientation stimulates the cell's receptive field then there will be significant excitation and little inhibition and the cell will respond maximally. However, if the orientation of the grating is changed as in Figure 4c2 then there will be a mix of excitation and inhibition and the response will be reduced. Thus the cell exhibits *orientation tuning*.

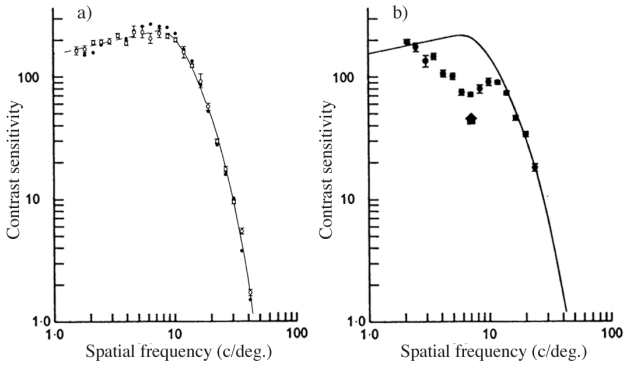

Figure 5: (a) The contrast sensitivity function of subject F.W.C.: Patterns were sine-wave gratings. Mean luminance of the gratings was 100 cd/m^2. Contrast sensitivity is plotted on a logarithmic scale against spatial frequency. Filled and open symbols show two independent measurements on the same subject. (b) Contrast sensitivity function for F.W.C. after adaptation to a sine-wave grating of 7.1 cpd. Note the depression in sensitivity in the spatial frequency band near the adapting frequency. Adapted from [Blakemore69].

2.2 Psychophysics of Spatial Vision

Given the physiological evidence for visual mechanisms in animals selective for contrast, spatial frequency, and orientation, psychophysicists began to test for the existence of similar mechanisms in human vision.

2.2.1 Contrast processing

The psychophysical evidence for contrast processing mechanisms in human vision has a long history going back at least as far as Mach [Ratliff65] who suggested that *lateral inhibition* could account for the bright and dark bands seen at discontinuities in luminance profiles, and Hering who proposed in his *opponent process theory* that antagonism between visual mechanisms was a fundamental principle of color and lightness perception and could explain such visual phenomena as simultaneous contrast and color constancy (see [Hurvich81] for a review). Modern psychophysical evidence for these mechanisms comes from the work of Campbell and Robson [1968] who measured the *contrast sensitivity function* of human vision for sine wave gratings of different spatial frequencies.

Campell and Robson tested contrast thresholds for sine wave gratings over a range of spatial frequencies and plotted the contrast sensitivity function shown in Figure 5a. In the fovea, at the luminance level tested, contrast sensitivity peaks for a pattern of 4-5 cycles/degree where a contrast of 0.5% can be detected. The graph shows that threshold contrast sensitivity declines for both higher and lower spatial frequencies.

2.2.2 Spatial frequency tuning

As was shown in the previous section, the receptive field organization of visual processing in cats and primates leads to

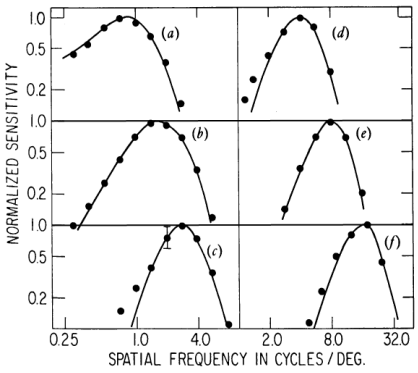

Figure 6: Model of spatial frequency tuned mechanisms in the human visual system: Points show mean data from three subjects on the spatial frequency tuning of six visual mechanisms. The curves show difference-of-Gaussian (DOG) function fits to the data for each mechanism. Mechanisms a-f are arranged in order of increasing peak spatial frequency. Each curve is plotted on a normalized sensitivity scale. Note that the spatial frequency scales in the right and left halves of the figure are different. From [Wilson84].

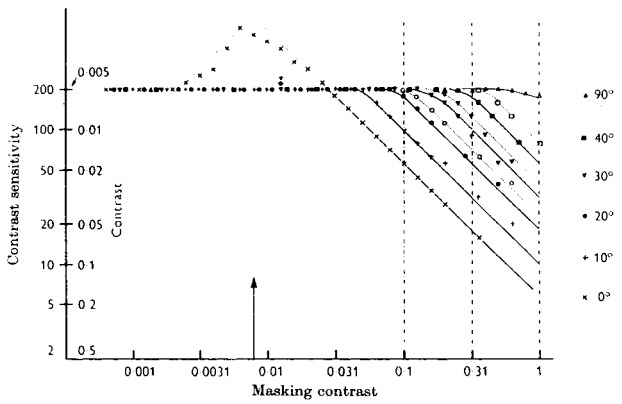

Figure 7: Orientation tuned mechanisms in the human visual system: Curves show contrast sensitivity for a vertical sine-wave test grating as a function of the contrast of a masking grating. Individual curves show how the orientation of the masking grating modulates the masking effect. Note that the magnitude of masking diminishes with the angular difference between the masking and test gratings. The anomalous data seen at low contrasts of the 0° masker is a *facilitation* effect that will be described in a later section. From [Campbell66].

visual mechanisms that are tuned to different ranges of spatial frequencies. Blakemore and Campbell [1969] conducted a series of psychophysical experiments to see if frequency tuned mechanisms exist in human vision.

They used an *adaptation* paradigm in their experiments. Prior to the experiment they measured the subject's contrast sensitivity function. They then had the subject inspect a grating pattern of a particular spatial frequency for one minute, instructing the subject to move their eyes constantly to avoid afterimages. They then re-measured the subject's contrast sensitivity function. Their results are shown in Figure 5b.

Contrast sensitivity is depressed for spatial frequencies close to the adapting frequency. The loss of sensitivity is greatest at the adapting frequency, but sensitivity is also depressed for spatial frequencies within a 2 octave band around the adapting frequency. Frequencies outside of this range are unaffected.

Wilson and Gelb [1984] performed a set of related experiments on spatial frequency discrimination to estimate the spatial frequency tuning of visual mechanisms in the fovea. Their *multiple mechanism* model illustrated in Figure 6 has six spatial frequency tuned mechanisms with different peak frequencies and spatial bandwidths. While there is ongoing debate about the number, peak frequencies, and bandwidths of spatially tuned mechanisms in human vision (see [Wilson91] for a review), the number of mechanisms in this model and their tuning parameters were derived by fitting experimental data and therefore provide a good account of actual visual performance.

2.2.3 Orientation tuning

A similar pattern of results can be found from psychophysical experiments testing the orientation tuning of human visual mechanisms. Campbell and Kulikowski [1969] used a *masking* paradigm to measure contrast sensitivity for a test grating in a vertical orientation, superimposed on a background grating which varied in orientation. Their results are shown in Figure 7.

When the test and background gratings have the same orientation (indicated by the x's in the 0° curve) sensitivity for the test grating drops in direct proportion to the suprathreshold contrast of the background grating. The apparent enhancement in sensitivity at low background contrasts is a *facilitation* effect that will be described in the following section on visual masking.

When the test and background gratings have different orientations, the drop in contrast sensitivity is a function of the angle between the gratings. The greater the angle between the gratings the less effect the background grating has on sensitivity for the test grating. This is indicated by the parallel curves in Figure 7 which show that as the angle between the gratings is increased, higher and higher background contrasts are needed to produce the same reduction in contrast sensitivity.

Phillips and Wilson [1984] performed a related set of experiments to determine the orientation tuning of human visual mechanisms at different spatial frequencies. The test pattern was a spatially localized grating patch superimposed upon a background grating that varied in orientation. Figure 8 shows the orientation tuning half-bandwidths of the visual mechanisms at different spatial frequencies. The results show that the visual system is more tightly tuned to orientation at high spatial frequencies than at low spatial frequencies. At a spatial frequency of 0.5 cycles/degree the orientation bandwidth of the visual system is approximately 60° (half-bandwidth × 2) and at 11 cycles/degree it has narrowed to approximately 30°.

2.3 Visual masking

The visual mechanisms described above are selective for bands of spatial frequencies and orientations. Interactions between image components within these bands result in masking effects like the ones illustrated in Figures 1 and 2 where the visual response to one component depends upon the presence of other components. The parameters of these masking effects were investigated by Legge and Foley [1980].

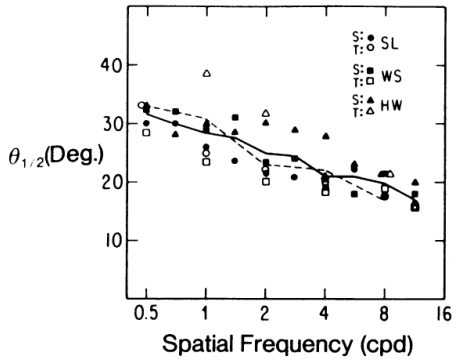

Figure 8: Bandwidth estimates of orientation-tuned mechanisms in the human visual system. The data shows the 50% amplitude, half-bandwidths of orientation-tuned visual mechanisms at different spatial frequencies. Different symbols are used for each of the three subjects. The filled symbols are for sustained presentations. The open symbols are for transient presentations. The solid line runs through the average half-bandwidth value at each spatial frequency. The dashed line compares these results to physiological data from primates [DeValois82]. Note that the orientation bandwidths of the mechanisms become progressively narrower with increasing spatial frequency. From [Phillips84].

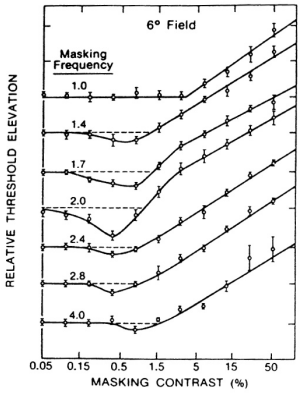

Figure 9: Facilitation and threshold elevation due to masking: Curves show contrast thresholds for a 2.0 cycle/degree sine-wave grating as a function of the masking grating contrast. The individual curves show the results for different spatial frequency masks. Each curve is plotted on its own arbitrary scale. The dotted line through each curve indicates the unmasked threshold for the 2.0 cycle/degree test grating. Note that the curves show a pattern of facilitation or increased sensitivity at low mask contrasts and threshold elevation at higher mask contrasts. From [Legge80].

They performed a series of experiments to determine how the presence of one grating affects the detectability of another. The first grating is called the *mask* and the other is the *test*. Their test grating was a sine-wave grating of 2.0 cycles/degree. The masks were phase-coherent sine-wave gratings that ranged in frequency from 1.0 to 4.0 cycles/degree. They measured the threshold contrast necessary to detect the test grating while varying the contrast and spatial frequency of the mask. Their results are shown in Figure 9.

The individual curves show the results for each mask frequency. Each curve is plotted on its own vertical scale showing in arbitrary units, the relative threshold elevation produced by the mask at different mask contrasts. The general form of the results is that very low mask contrasts have no significant effect on the detectability of the test grating, but as the mask contrast is increased, at first the threshold drops showing increased sensitivity or *facilitation* and then rises again showing a loss in sensitivity or *threshold elevation* for high contrast masks. The shape of the threshold elevation curve is evidence of a contrast nonlinearity in the visual system caused by masking. This contrast nonlinearity is an accelerating function at low mask contrasts and a compressive function at higher mask contrasts.

The curves in Figure 9 also shows the spatial frequency tuning of visual masking. Threshold elevation is greatest when the mask and test gratings have the same spatial frequency. As the spatial frequencies of the mask and test become different greater and greater mask contrasts are necessary to produce the same threshold elevation. The effects of orientation tuning on masking can be understood in a similar way.

2.4 Extensions to color

There is now substantial physiological and psychophysical evidence for *spectrally-tuned opponent mechanisms* in color vision [Hurvich81]. The evidence supports a description of color vision in terms of the responses of an *achromatic channel* and two *chromatic channels*, one tuned to a red/green dimension and the other to a yellow/blue dimension. If we want to predict masking effects in complex color images we need to correctly model masking in both the chromatic and achromatic visual channels.

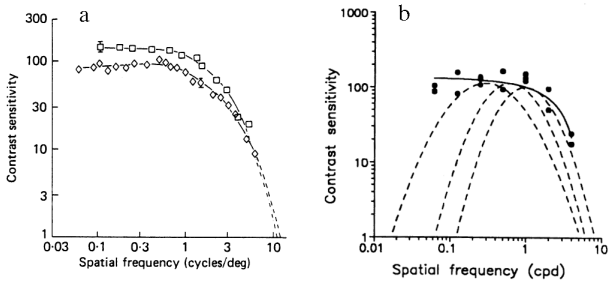

Figure 10: (a) Chromatic contrast sensitivity functions: The curves show contrast sensitivity as a function of spatial frequency for red/green (□ 602, 526nm) and yellow/blue (◇ 470, 577nm) isoluminant sine-wave gratings. From [Mullen85]. (b) Bandpass chromatic mechanisms underlying the chromatic CSF. The solid line shows the average contrast sensitivity data from three subjects for a red/green grating. The dashed curves show Gaussian fits to the average spatial frequency tuning functions measured in the experiments. Peak frequencies are 0.25, 0.5 and 1 cpd. The sensitivities of the mechanisms have been adjusted to fit the shape of the CSF. From [Losada94].

Mullen [1985] measured the contrast sensitivity function of the chromatic channels. The results are shown in Figure 10a. The separate curves show the results for red/green and yellow/blue gratings. Several differences between these *chromatic* CSF's and the *achromatic* CSF shown in Figure 5a should be noted. First, the chromatic CSF's have a low-pass frequency characteristic unlike the bandpass nature of the achromatic CSF. Second, the high frequency cutoff is

approximately 11 cycles/degree rather than more than 30 for the comparable achromatic CSF. This means that the chromatic channels have much lower spatial resolution than the achromatic channel. Finally, the absolute sensitivity of the chromatic channels is lower than that of the achromatic channel except at very low spatial frequencies.

Losada and Mullen [1994] measured the spatial frequency tuning of the red/green chromatic channel. The results are shown in Figure 10b. The three bandpass mechanisms shown provide a good fit to the psychophysically measured chromatic CSF.

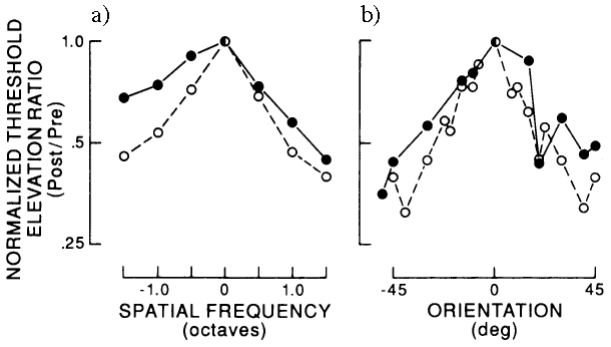

Figure 11: Normalized tuning functions for the achromatic and chromatic channels: (a) Spatial frequency tuning. Tuning is measured in octaves (doubling of spatial frequency). Open circles are for the achromatic channel, filled circles are for the red/green chromatic channel. (b) Orientation tuning. Tuning is measured in degrees. In both cases note the similarity between the tuning characteristics of the achromatic and chromatic channels. From [Bradley88].

Bradley [1988] has shown that the absolute spatial frequency and orientation tuning of the chromatic channels is broader than the achromatic channel, however when this data is normalized for contrast sensitivity differences between the channels, the relative spatial frequency and orientation tuning of the channels is very similar. This is shown in Figure 11. This allows us to use same spatial frequency and orientation tuned mechanisms for both the chromatic and achromatic channels in our masking model. Of course the scaling factors will be different to fit their respective CSF's.

Finally to predict masking effects in color images we must consider interactions between the chromatic and achromatic channels. Switkes [1988] found an asymmetric relationship between luminance masking and chromatic masking for sine-wave grating patterns. While luminance masks can facilitate detection of chromatic test gratings, high contrast chromatic masks raise thresholds for luminance test gratings. However, Gegenfurther [1992] failed to find significant cross-channel effects for sine-wave gratings presented in chromatic and achromatic noise. Given the lack of consistent data on cross-channel effects we have decided to model only in-channel masking effects which should provide good results except for pathological cases.

3 Computational Model of Visual Masking

We will now draw on the physiological and psychophysical evidence outlined in the previous sections to develop a computational model of visual masking. Our goal is to create an algorithm that can predict when a texture will mask visual artifacts in a synthetic image.

Figure 12 illustrates our approach. To determine the visibility of image artifacts we take a *reference* image and a *test* image containing artifacts. The images are processed through the masking model. The model contains four stages. In the first stage the spectral radiances in the images are transformed into responses in an opponent color space to produce a *color representation*. In the next stage this color representation is decomposed into a *pattern representation* that accounts for the sensitivities of the spatial frequency and orientation tuned mechanisms in the visual system. In the third stage an appropriate *masking function* is applied to each visual mechanism to account for masking effects between the image components contained within a mechanism's sensitive band. Finally the responses of the mechanisms are compared in the *detection model* to determine when the artifacts will be visible and when they will be masked.

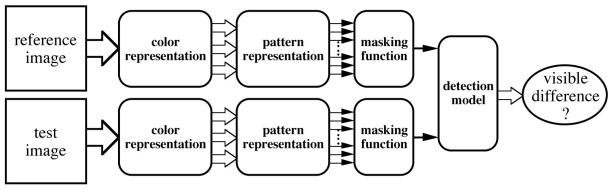

Figure 12: Components of the masking model.

3.1 Color representation

Figure 13 illustrates how the color representation is formed. At the first stage, the spectral irradiance at each point in the retinal image is encoded into the responses of three different classes of cone photoreceptors sensitive to long (L), middle (M) and short (S) wavelength ranges of visible light. The response (R) of each cone is calculated as a nonlinear function (Γ) of the effective light absorption by the cone, where the effective light absorption is given by the light spectrum $I(\lambda)$, times the cone's spectral sensitivity, $S(\lambda)$, and integrated over the wavelength range. Thus we have,

$$\mathrm{R}_j = \Gamma\left(\int_{380nm}^{780nm} I(\lambda)S_j(\lambda)d\lambda\right),$$

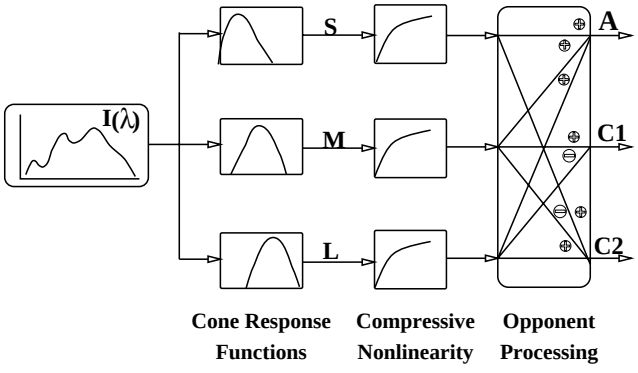

Figure 13: Color representation.

where j = S, M, or L. The function Γ is normally chosen to be a compressive nonlinearity, e.g., a square root, or a logarithm (which yields Weber's law behaviour [Wilson91]).

At the next stage, the cone responses are transformed by opponent mechanisms. This opponency results in a color representation that can be described in terms of the signals in an achromatic channel and two chromatic channels. This model is based on work by Bolin [1995] and Meyer [1988].

3.2 Pattern representation

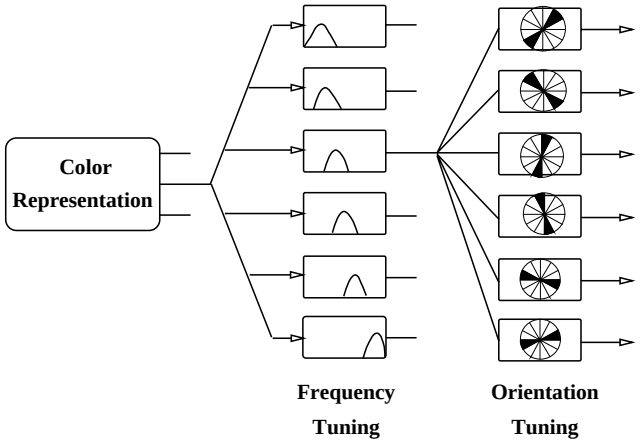

Figure 14: Pattern representation.

As was discussed in detail in earlier sections, the visual system processes spatial patterns with mechanisms tuned to various ranges of frequencies and orientations. The achromatic and chromatic representations of the image undergo transformation by each of these mechanisms and give rise to the pattern representation. Figure 14 illustrates this process.

To implement a model of these mechanisms, we need to choose a particular set of mechanism response functions. A number of mathematical functions such as Gabor functions [Marcelja80], Cortex transforms [Watson87], and differences-of-Gaussians (DOGs) [Marr82] have been proposed to emulate these mechanisms. While all of these functions have been used to model visual processing, and some like the Cortex transform offer greater computational efficiency, we have chosen to use the DOG functions defined by [Wilson84] because they provide an accurate quantitative fit to both physiological and psychophysical data.

Figure 6 shows the normalised response curves for these functions. The expression for the corresponding visual filter functions is:

$$\mathcal{RF}_i = A_i \left[e^{-x^2/\sigma_{x1i}^2} - B_i e^{-x^2/\sigma_{x2i}^2} + C_i e^{-x^2/\sigma_{x3i}^2} \right] e^{-y^2/\sigma_{yi}^2} \quad (1)$$

where x, y are spatial positions in foveal degrees. Different scalings of the gaussians provide filters tuned to different spatial frequencies.

The parameter, A_i, in the equation determines the gain of each mechanism which varies with adaptation and chromatic conditions. Since the psychophysically measured CSF is the envelope of these mechanism gains, by determining the chromatic and achromatic CSF's for our display conditions we can compute the appropriate A_i's. CSF data were generated from the equations given in [Martin92]. We computed the A_i's by least-squares fits to this data.

To account for the orientation tuning of the mechanisms, the filters are used with rotated coordinates. The data given in Figure 8 was used to determine the orientation bandwidths of the mechanisms which range from approximately 60° (for low frequencies) to 30° (for high frequencies). These differences in orientation tuning mean that the number of orientation channels in the pattern representation is different for each mechanism.

Given a visual mechanism $\mathcal{RF}$ with spatial frequency tuning i and orientation tuning θ, the sensitivity of this mechanism to a pattern $I(x', y')$ is given by the expression:

$$S_{i,\theta} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathcal{RF}_{i,\theta}(x - x', y - y') I(x', y') dx' dy'. \quad (2)$$

3.3 Masking function

If the visual system behaved in a linear fashion then it would be possible to predict the appearance of an image directly from the mechanism sensitivities $S_{i,\theta}$ given in equation 2. However, the response (R) of each mechanism is nonlinear due to masking interactions among the spatial frequency components within the mechanism's sensitive band. Wilson [1984] developed the following expression to take these nonlinearities into account in determining the responses of each mechanism in the pattern representation.

$$R_{i,\theta} = R(S_{i,\theta}) = \frac{S_{i,\theta}^2 + K_i S_{i,\theta}^{3-\epsilon_i}}{K_i + S_{i,\theta}^2} \quad (3)$$

In this expression $R_{i,\theta}$ is the response of the (i, θ)-th mechanism to an input pattern. $S_{i,\theta}$ is the mechanism's sensitivity given by equation 2. K_i is given by

$$K_i = 1/H_i(1 - \epsilon_i) \quad (4)$$

and H_i and ϵ_i are empirically derived constants specific to each mechanism that Wilson [1984] determined by fitting the expression to the results of masking experiments performed by [Nachmias74].

3.4 Detection model

Given the mechanism responses we can determine how detectable artifacts in the test image will be. If we neglect response variability for the moment then the responses to two identical images should be the same. Artifacts in one of the images should result in differences in the responses. For each mechanism this difference can be described by:

$$\Delta R_{i,\theta} = \mid R_{i,\theta}(\text{Image}_1) - R_{i,\theta}(\text{Image}_2) \mid \quad (5)$$

where $R_{i,\theta}(\text{Image}_1)$ and $R_{i,\theta}(\text{Image}_2)$ are the responses of the (i, θ)-th mechanism to each of the images.

These differences can be used to determine how visible the artifacts will be by the following formula:

$$\Delta R = \left[\sum_i \sum_\theta \Delta R_{i,\theta}^Q \right]^{\frac{1}{Q}} \quad (6)$$

In this formulation each image is represented as a point in a multidimensional response space and the visibility of artifacts in the test image is related to the distance between the reference and test images in this space. This model is

essentially similar to the classic line-element model for color discrimination developed by Stiles [1978]. However in this case the metric is based on differences in the responses of the spatial frequency and orientation tuned mechanisms in the pattern representation. Choosing a value of 2.0 for Q gives the response space a Euclidean distance metric which has been shown to provide a good fit to experimental data on pattern discrimination [Wilson84].

Finally, there is variability in visual response due to noise within the visual mechanisms as well as uncertainty in the observer's decision making processes. These effects can be modeled by the *psychometric function* [Graham89], which in this case relates the distance between the reference and test images in the response space, to the likelihood that the images will be discriminable. The psychometric function is given by:

$$P(\Delta R) = 1 - 2^{-(1+k\Delta R)^3} \quad (7)$$

The constant k has been given a value of 0.2599, to scale the function so that a ΔR value of 1.0 corresponds to P =0.75, the standard threshold value for discrimination in a two-alternative forced choice (2AFC) task [Wilson91].

3.5 Applying the masking model

To demonstrate the utility of the masking model we applied the model in an algorithm to predict how masking provided by a texture map affects the visibility of shading artifacts caused by polygonal tesselation of a curved surface. The results are shown in Figure 15.

(a1) shows the flat-shaded approximation to the cylinder used in the first three rows of the figure. We examined how three aspects of the texture: its contrast, spatial frequency spectrum, and orientation, affected masking and the visibility of the artifacts. We generated synthetic textures for our study using Perlin's [1989] texture synthesis algorithm.

Row (a) shows how masking is affected by the contrast of the texture map. The texture contrast increases from left to right. Faceting is visible at the lower contrasts shown in (a2) and (a3) but not at the higher contrasts shown in (a4) and (a5).

Row (b) shows how masking is affected by the relationship between the spatial frequency components of the cylinder's luminance profile and the texture map. Since masking only occurs within limited frequency bands, high frequencies will not mask low frequencies and vice-versa. Thus we expect masking to occur only when the texture frequencies and facet frequencies are similar, and this can be observed in the figure. At each end of the row where the spatial frequencies of the texture and the spatial frequencies due to the faceting are very different, the faceting is visible. As we move toward the center of the row where the frequencies become more similar masking increases.

Row (c) shows how masking and the visibility of faceting is affected by the relative orientation of the facets and the texture map. In (c1) where the facets and the dominant orientation of the texture are orthogonal, little masking occurs and the facets can be seen. As the texture is rotated toward the vertical in (c2)-(c5) the masking effect increases. In (c4) some faceting is still detectable, but only with careful scrutiny. In (c5) the faceting has become invisible.

Finally in row (d) we show that the masking effect can be used to find tesselations for curved surface so that faceting artifacts won't be seen. Given a particular texture we want to find low tesselations where the faceting will still be masked by the texture. The tesselation increases from left to right. At the lower levels of detail the given texture fails to mask the facet artifacts, but in (d3) the faceting is just barely detectable and when it is increased in (d4) it is no longer visible. Thus the minimum visually acceptable tesselation for this texture falls between these two levels.

All these effects are predicted by the masking model. The numbers listed at the top of each image give the ΔR values computed by the model for the visibility of faceting in each image. Values smaller than 1.0 indicate that the image is not visibly different from the reference image in a standard forced-choice discrimination task.

4 Conclusions and future work

In this paper we have developed a computational model of visual masking for computer graphics. The model analyzes how the presence of one visual pattern affects the detectability of another. The model allows us to choose texture patterns for synthetic images that hide visual artifacts caused by errors in graphics algorithms. We have demonstrated how the model can be used to determine when a texture will mask the faceting artifacts caused by polygonal tesselation of a curved surface, but the model can be applied to in the same manner to predict masking of other artifacts such as banding, aliasing, and noise. Since the model takes into account how masking is affected by the relationships between the contrast, spatial frequency spectrum, and orientation of the texture and the artifacts it can be used to choose the texture features required to mask artifacts with given characteristics. Conversely if the texture properties are given, the model can predict the precision required in modeling and rendering processes to bring the artifacts to a level that can be masked by the texture. Because the model is based on data from psychophysical experiments on human vision, it is *predictive* and allows us to determine a priori whether a texture pattern will mask a given visual artifact.

There is still much work to be done on this model: structured patterns and unstructured noise masks exhibit different patterns of masking due to learning effects [Watson97]; Chromatic and achromatic cross-channel masking effects are still being investigated in the psychophysics literature [Switkes88, Gegenfurther92]; Foley [1994] has shown that better fits to psychophysical masking experiments may require a model with a divisive inhibitory term that pools activity across mechanisms; and finally, dynamic effects and changes in visual function across the visual field must certainly play a role in masking. All these dimensions should be incorporated into future models and the models should be tested and validated in a wider variety of applications.

In his "Visible Difference Predictor" Daly [1992] incorporated a model of visual masking that is similar in many ways to our own. Daly uses data on threshold elevation from masking studies to predict when an original image and image with artifacts will be visibly different. His method determines visibility differences by calculating the differences in contrast between "images" that represent the activity within spatial frequency and orientation tuned visual mechanisms. We originally tried to apply his approach to our texture masking problem but encountered difficulty in developing a meaningful formulation of the effects of the texture mask in terms of the differences in contrast between our reference and test images. Therefore we based our masking model on Wilson's formulation which has been shown to predict actual psychophysical results, and which offers a more natural interpretation of visible differences in terms of differences in the *responses* of visual mechanisms and the distance between

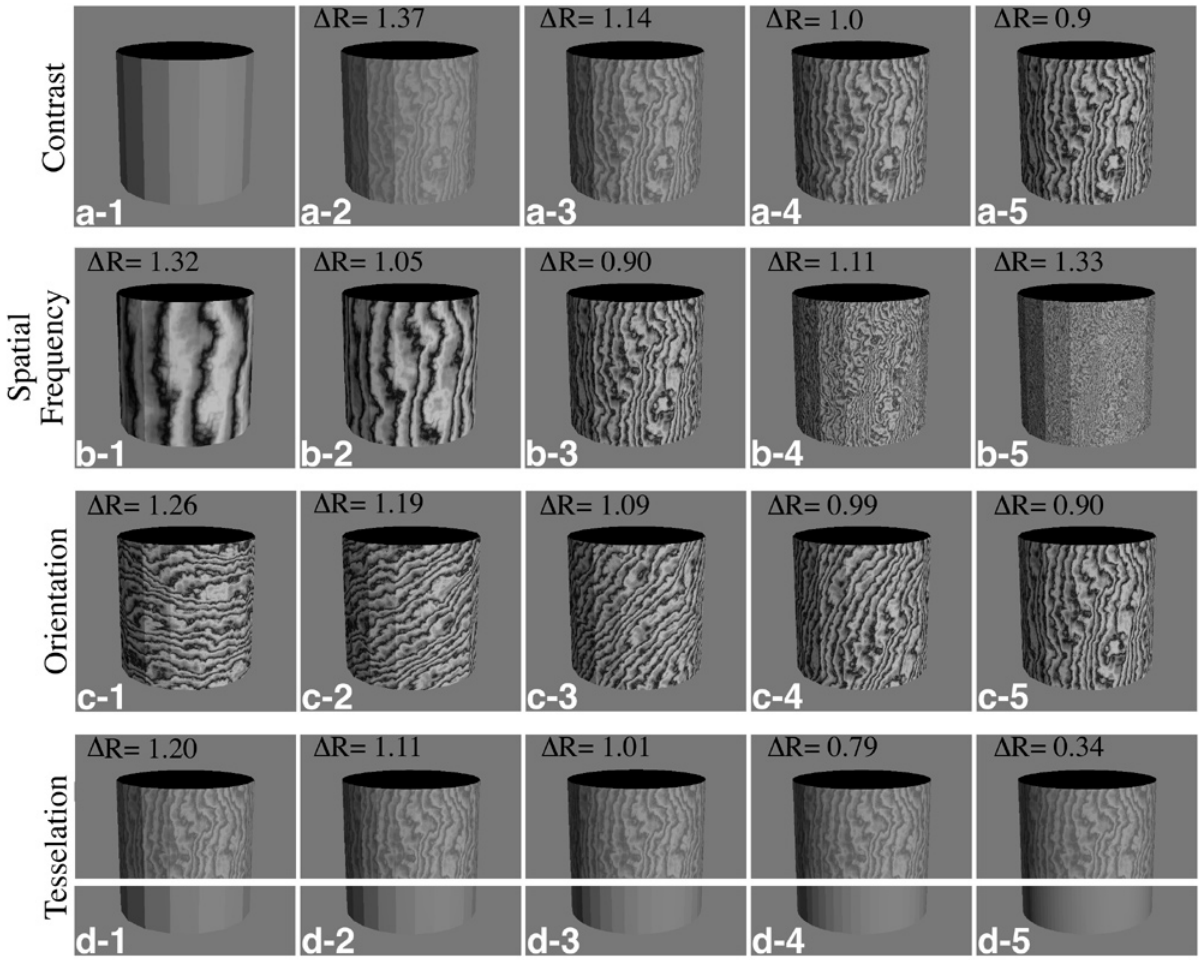

Figure 15: Applying the masking model.

images in a multidimensional response space. Although we have ended up with a somewhat different model than Daly's there is much to commend his approach especially in terms of accounting for learning effects in masking, and in producing visualizations of visible difference maps. We hope that further work will be able to unify the two models.

While we have demonstrated the utility of our masking model in a specific computer graphics application, the model is general and has uses in geometric modeling, realistic image synthesis, scientific visualization, image compression, and image-based rendering. In geometric modeling the masking model should allow us to determine the minimum level-of-detail necessary so a textured object will not exhibit visible shading artifacts. In realistic rendering, Monte Carlo methods often produce noisy images. Here, the model should allow us to determine the thresholds for visible noise in textured scenes to improve the quality and efficiency of these algorithms. In scientific visualization, abstract data is often coded as color or texture patterns superimposed on the surface of a shaded three-dimensional object. At times these patterns can obscure the shape of the underlying geometry leading to misinterpretations of the data. By using the model to predict when masking *won't* occur, we should be able to choose scales for abstract data that don't obscure the object's shape. In compression, methods such as JPEG and MPEG follow the same general visual model we've used, but do not consider the effects of masking. Applying the masking model should allow more aggressive quantization of the DCT coefficients and thereby achieve higher compression factors without losses in visual fidelity (see also [Watson93]). Finally, the model may have applicability in newly developed image-based rendering schemes [Torborg96] where objects are rendered and then their images are warped to efficiently simulate perspective changes. We may be able to use the masking model to determine how much perspective distortion is acceptable for a given relationship between the surface texture and the underlying geometry. This could provide a perceptual basis for the distortion criterion which may allow higher levels of distortion and thereby improve performance.

Acknowledgements

Special thanks to Linda Stephenson for securing permissions for the figures used in this paper.

This work was supported by the NSF/ARPA Science and Technology Center for Computer Graphics and Scientific Visualization (ASC-8920219) and by NSF CCR-9401961 and NSF ASC-9523483 and performed on workstations generously provided by the Hewlett-Packard Corporation.

References

[Arvo94] Arvo, J., Torrance, T., and Smits, B. (1994). A framework for the analysis of error in global illumination algorithms. Proceedings SIGGRAPH 94, 75-84.

[Blakemore69] Blakemore, C. and Campbell, F.W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. J. Physiol., 203, 237-260.

[Bolin95] Bolin, M.R. and Meyer, G.M. (1995). A frequency based ray tracer. Proceedings SIGGRAPH 95, 409-418.

[Bradley88] Bradley, A., Switkes, E., and DeValois, K.K. (1988). Orientation and spatial frequency selectivity of adaptation to color and luminance patterns. Vision Res., 28, 841-856.

[Campbell66] Campbell, F.W. and Kulikowski, J.J. (1966). Orientation selectivity of the human visual system. J. Physiol., 187, 437-445.

[Campbell68] Campbell, F.W. and Robson, J.G. (1968). Application of Fourier analysis to the visibility of gratings. J. Physiol., 197, 551-566.

[Daly92] Daly, S. (1992). The visible difference predictor: an algorithm for the assessment of image fidelity. Human Vision, Visual Processing and Digital Display, SPIE Vol. 1666, 2-15.

[DeValois82] DeValois, R.L., Yund, E.W., and Hepler, N. (1982). The orientation and direction selectivity of cells in macaque visual cortex. Vision Res., 22, 531-544.

[Dowling66] Dowling, J.E. and Boycott, B.B. (1966). Organization of the primate retina: electron microscopy. Proc. Royal Soc. Lond. Ser. B., 166, 80-111.

[Enroth-Cugell66] Enroth-Cugell, C. and Robson, J.G. (1966). The contrast sensitivity of retinal ganglion cells of the cat. J. Physiol., 187, 517-552.

[Fletcher52] Fletcher, H. (1952). *Speech and hearing* (revised ed.). New York: van Nostrand.

[Foley94] Foley, J.M. (1994). Human luminance pattern-vision mechanisms: masking experiments require a new model. J. Opt. Soc. Am. A, 11(6), 1710-1719.

[Gegenfurther92] Gegenfurther, K.R. and Kiper, D.C. (1992). Contrast detection in luminance and chromatic noise. J. Opt. Soc. Am. A, 9(11). 1880-1888.

[Graham89] Graham, N.V. (1989). *Visual Pattern Analyzers*. New York: Oxford University Press.

[Harmon73] Harmon, L.D. and Julesz, B. (1973). Masking in visual recognition: effects of two-dimensional filtered noise. Science, 180, 1194-1197.

[Hubel62] Hubel, D.H. and Wiesel, T.N. (1962). Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. J. Physiol., 160, 106-154.

[Hubel68] Hubel, D.H. and Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. J. Physiol., 195, 215-243.

[Hurvich81] Hurvich, L. (1981). *Color Vision*. Sunderland, MA: Sinauer Assoc.

[Kuffler53] Kuffler, S.W. (1953). Discharge patterns and functional organization of the mammalian retina. J. Neurophysiol., 16, 37-68.

[Legge80] Legge, G.E. and Foley, J.M. (1980). Contrast masking in human vision. J. Opt. Soc. Am., 70, 1458-1470.

[Losada94] Losada, M.A. and Mullen, K.T. (1994). The spatial tuning of chromatic mechanisms identified by simultaneous masking. Vision Res., 34(3), 331-341.

[Marcelja80] Marcelja, S. (1980). Mathematical description of the responses of simple cortical cells. J. Opt. Soc. Am., 70, 1297-1300.

[Marr82] Marr, D. *Vision*. San Francisco: W.H. Freeman.

[Martin92] Martin, R.A., Ahumada, A.J. and Larimer J.A. (1992). Color matrix display simulation based upon luminance and chromatic contrast sensitivity of early vision. Human Vision, Visual Processing and Digital Display, SPIE Vol. 1666, 336-342.

[Meyer88] Meyer, G.M. (1988). Wavelength selection for synthetic image generation. Computer Vision, Graphics, and Image Processing, 41, 57-79.

[Mitchell87] Mitchell, D.P. (1987). Generating antialiased images at low sampling densities. Proceedings SIGGRAPH 87, 65-72.

[Mullen85] Mullen, K.T. (1985). The contrast sensitivity of human color vision to red-green and blue-yellow chromatic gratings. J. Physiol., 359, 381-400.

[Nachmias74] Nachmias, J. and Sansbury, R.V. (1974). Grating contrast: discrimination may be better than detection. Vision Res., 14, 1039-1042.

[Pantle69] Pantle, A. and Sekuler, R.W. (1969). Contrast response of human visual mechanisms sensitive to orientation and direction of motion. Vision Res., 9, 397-406.

[Perlin89] Perlin, K. (1989). Hypertexture. Proceedings SIGGRAPH 89, 253-262.

[Phillips84] Phillips, G.C. and Wilson H.R. (1984). Orientation bandwidths of spatial mechanisms measured by masking. J. Opt. Soc. Am. A, 1, 226-232.

[Ratliff65] Ratliff, F. (1965). *Mach Bands: Quantitative Studies on Neural Networks in the Retina*. San Francisco: Holden-Day.

[Stiles78] Stiles, W.S. (1978). *Mechanisms of Color Vision*. London: Academic Press.

[Switkes88] Switkes, E., Bradley, A. and DeValois, K.K. (1988). Contrast dependence and mechanisms of masking interactions among chromatic and luminance gratings. J. Opt. Soc. Am. A, 5, 1149-1162.

[Torborg96] Torborg, J. and Kajiya, J.T. (1996). Talisman: commodity realtime 3d graphics for the PC. Proceedings SIGGRAPH 96, 353-363.

[Watson87] Watson, A.B. (1987). Efficiency of a model human image code. J. Opt. Soc. Am. A, 4, 2401-2417.

[Watson93] Watson, A.B. (1993) DCT quantization matrices visually optimized for individual images. Human Vision, Visual Processing and Digital Display, SPIE

[Watson97] Watson, A.B., Borthwick, R. and Taylor, M. (1997). Image quality and entropy masking. To appear in: Proceedings SPIE, Vol. 3016.

[Wilson84] Wilson H.R. and Gelb, D.J. (1984). Modified line-element theory for spatial-frequency and width discrimination. J. Opt. Soc. Am. A, 1, 124-131.

[Wilson91] Wilson, H.R. (1991). Psychophysical models of spatial vision and hyperacuity. in D. Regan (Ed.) *Spatial Vision*, Vol. 10, Vision and Visual Dysfunction. Boca Raton, FL, CRC Press.