SIMULATING NORMAL AND LOW VISION

ELI PELI
The Schepens Eye Research Institute, Harvard Medical School
and The New England Eye Center, Tufts University School of Medicine
Boston, MA 02114 , USA

Simulation may provide insights into the factors that control the appearance of images to normal and visually impaired observers. The linear spectral methods that have been useful in interpreting low contrast, threshold phenomena with periodic stimuli are inappropriate for analyzing the local contrast in images. Linear simulations commonly use the normalized contrast sensitivity function (CSF) as a modulation transfer function (MTF) applied to the image amplitude rather than to contrast. Simulations within the linear model are limited by a number of problems discussed here. To achieve a valid simulation, a measure of local band-limited contrast in images is needed. Such a measure was developed and was used to simulate the threshold nonlinear characteristics of the visual system and suprathreshold contrast constancy. This approach was used to simulate both central vision and vision with peripheral retina. In the latter, the simulations were based on CSF measurements across the retina. Methods to test the validity of these simulations were implemented. The results of preliminary testing confirm our ability to evaluate important parameters of the model and point to previously unnoted image dependent characteristics of visual perception.

1. Introduction

Simulations of various environments are frequently created to provide cost effective or less dangerous methods of training or evaluating operational capabilities of machines and their operators. I will discuss here the simulation of the appearance of a scene in the environment to an observer when viewed on a display. Such pictorial representations have been attempted by many over the years in an effort to illustrate the effects on observers of visual disability[5,15,30], changes in observation distance[10], and the use of peripheral vision[34]. More recently, the simulation of human observers has been integrated with the physical simulation of new display systems[16]. This combination should enable visual effects to be considered as part of the display.

Why do we need simulations? After all, the concepts of linear filters and convolution have become common enough in visual perception research to provide most workers with intuitions about the effects of simple filters on various signals. I will try to show that these basic linear system concepts may not be sufficient or appropriate for the analysis of the appearance of complex images. The simulations that I propose overcome some

In Peli E, editor (1995) Visual Models for Target Detection and Recognition (in memory of Arthur Menendez).World Scientific Publishing Co. Singapore, New Jersey, London, Hong Kong.

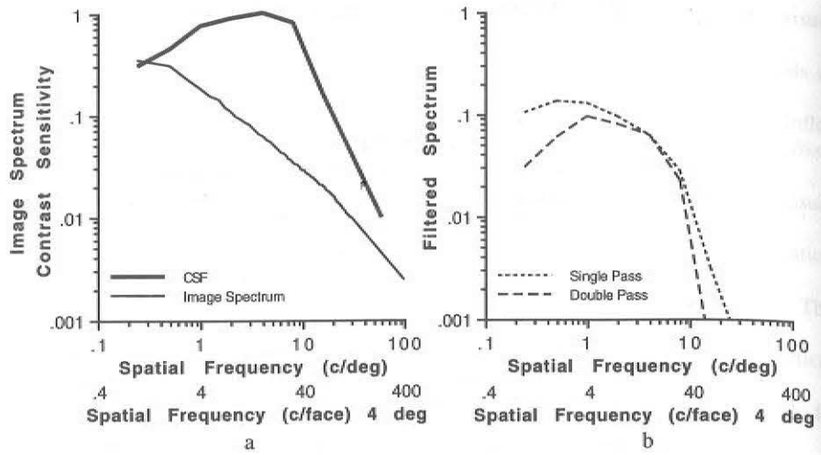


Figure 1. Linear simulation of the appearance of images. a) The normalized contrast sensitivity function (CSF, thick line) used to multiply the Fourier transform of the image. The Fourier transform is illustrated by the radially averaged amplitude spectrum of a face image (thin line). b) The result of the multiplication is illustrated by a dotted line. When the observer examines the simulation it is processed once more by the same contrast sensitivity function (dashed line) leading to the *double-pass* problem.

of these difficulties and thus can provide the required insight. Further, I will demonstrate that the simulated images may be useful for testing the visual system models underlying them and may lead to better, more refined models.

When simulations are intended to illustrate differences among subjects' visual sensitivities, one must be particularly careful about interpreting the resulting images. There are two problems. First, the images must be computed and displayed to the observer in such a way that the observer's spatial sensitivity characteristics do not interact with the phenomena being portrayed. Ginsburg's[10] linear simulations of the appearance of images to normal subjects have been criticized for the *double-pass* effect, resulting when the simulated image is processed again by the reader's visual system[36]. However, by careful design it is possible to present images in which the important details are relatively unaffected by the characteristics of the reader's visual system. For example, the linear systems approach, in which the ratio of the abnormal to the normal contrast sensitivity function (CSF) represents the modulation transfer function (MTF), was used to simulate low vision and represent the appearance of images to amblyopes[15]. When simulating low vision, the double-pass problem is insignificant. The detail loss suffered by the patient and portrayed in the simulation is at a spatial scale and of such a large magnitude that the effects of the normal reader's visual system can be ignored.

The second problem, which arises from treating the CSF as an MTF, is more serious and applies to simulation of both normal and low vision. The linear systems approach to simulation[10,11] ignores many aspects of the nonlinearities of the visual system, at least three of which are important for simulations. First, information that is below threshold is not merely highly attenuated, it is lost and cannot be recovered (the threshold nonlinearity). Second, curves of constant suprathreshold apparent contrast are not multiples of the CSF. Above threshold, apparent contrast is relatively independent of retinal eccentricity and spatial frequency (contrast constancy), while the contrast threshold changes significantly[2]. Third, thresholding should be applied to image contrast, which is a nonlinear function of the local amplitudes in the image (contrast nonlinearity). To be valid, simulations must take these nonlinearities into account.

In the simulations presented here, all three nonlinearities were incorporated. The simulations explicitly applied the threshold nonlinearity, and the processing was done in the nonlinear contrast domain rather than the amplitude domain. This approach has the additional effect of reducing the double-pass problem. When one wishes to portray the loss of details that occurs in viewing an object from a specified distance, the simulations

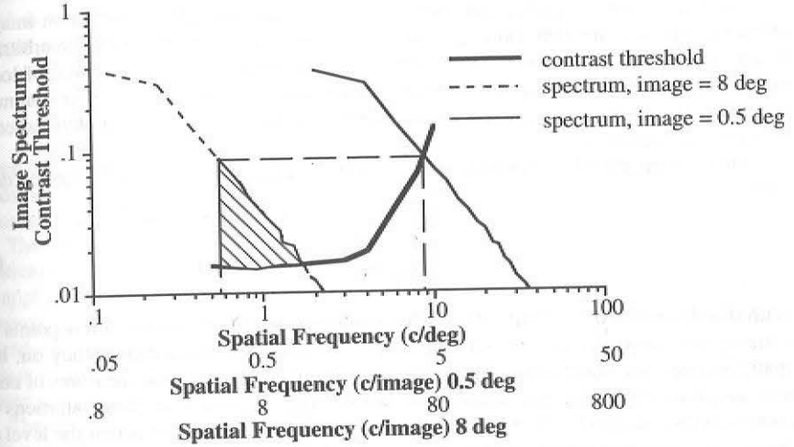


Figure 2. Reducing the double-pass problem within the context of a nonlinear threshold model of the visual system. Assume that the amplitude spectrum (thin solid line) represents a "contrast" spectrum of an image spanning 0.5 deg. Image content beyond the intersection of the observer's contrast detection threshold (thick line) and the spectrum will not be visible to the observer. This could be simulated by removing all image content to the right of the intersection. If the simulated image is then magnified or inspected from a short distance, where it spans 8 deg (dashed line), most of the effects of simulation (shaded area) will be suprathreshold and these effects can be detected by the observer when the simulation is compared to the unprocessed original image.

should remove the details that drop below threshold at that distance with a process that mimics the effect of the visual system[22]. The image then should be magnified (or equivalently, examined from a shorter distance) so that the surviving fine details are large enough to be seen without being affected appreciably by the reader's visual system (Fig. 2). Since the subthreshold information was removed completely, it cannot be altered by the reader's visual system. To use this approach, a method of measuring local contrast in images is required.

In this chapter I will:

1) Define a metric to represent local band-limited contrast in complex images;
2) Discuss the effects of visual nonlinearities on the appearance of images in foveal vision;
3) Illustrate how the simulations can be used to evaluate the contrast thresholds obtained with different psychophysical stimuli;
4) Present a formulation for simulating the appearance of images processed with an inhomogeneous retina.

2. Local Band-limited Contrast

Contrast is a basic perceptual attribute, based on the amplitudes within an image. However, the measurement and evaluation of contrast and contrast variations in arbitrary images are not uniquely defined in the literature. I have proposed a definition of local band-limited contrast in complex images[22] that is more closely aligned to the common definition of contrast in simple test patterns and can better relate measured physical contrast to visual perception.

Most commonly, the contrast of images has been evaluated using the Michelson definition

$$c = \frac{L_{max} - L_{min}}{L_{max} + L_{min}}. \tag{1}$$

With this definition the contrast of the whole image is dependent on only a few points of extreme brightness or darkness. Thus, there is no reference to, or dependency on, the spatial frequency content of the image. In addition, all commonly used measures of contrast assign a single value of contrast to the whole image, ignoring the local variations of contrast within images[18]. The basis of the local band-limited contrast is that the level of the local mean luminance should be considered in defining the contrast at every point and at every frequency band[22].

To define local band-limited contrast for a complex image, a band-limited version of the image in the frequency domain, $A(u,v)$, must first be obtained using a radially symmetric bandpass filter, $G(\rho)$.

$$A(u,v) \equiv A(\rho,\theta) = F(\rho,\theta) \cdot G(\rho) \tag{2}$$

where u and v are the horizontal and vertical spatial frequency coordinates, and ρ and θ represent the polar spatial frequency coordinates:

$$\rho = \sqrt{u^2 + v^2} \tag{3}$$

$$\theta = \tan^{-1}\left(\frac{v}{u}\right); \tag{4}$$

and where $F(\rho,\theta)$ is the Fourier transform of the image $f(x,y)$. We call the spatial frequency support of this bandpass filter the band.

In the space domain, the filtered image $a(x,y)$ can be represented similarly:

$$a(x,y) = f(x,y) * g(x,y) \tag{5}$$

where $(*)$ represents the convolution operator and $g(x,y)$ is the inverse Fourier transform of the bandpass filter $G(\rho)$. We also can define for every $a(x,y)$ the corresponding $l(x,y)$, which is a lowpass filtered image containing all energy below the band of interest. The contrast at the band can be represented as a two-dimensional array $c(x,y)$:

$$c(x,y) = \frac{a(x,y)}{l(x,y)} \tag{6}$$

where $l(x,y) > 0$. This definition provides a local contrast measure for every band that depends not only on the local energy at that band, but also on the local background luminance as it varies from place to place in the image.

The computation of local band-limited contrast as described above is conceptually identical to any of the commonly used pyramids of bandpass-filtered images[38]. Since for the application of the simulation images of equal size are used at all bands, we avoid the common approach of sub sampling the image recursively, filtering, and then upsampling the reduced size images. Instead, all filtering is done in the frequency domain. Thus the content of the final pyramid of image scales is identical to the images that would be calculated by upsampling images obtained on a pyramid of image resolution.

The details of the implementation of this pyramidal image analysis structure in the discrete, digital case have been described elsewhere[22]. Various implications of this definition of contrast, and their relation to the perceived contrast in complex images, were also discussed. The most important implications for the purposes of simulation are the facts that contrast and amplitude differ and cannot simply be interchanged, and that "linear rescaling" of images, which linearly scales amplitudes, results in a nonlinear space-varying change of contrast. Other investigators have recently adopted similar approaches to the calculations of local contrast and the generation of simulations[3,4]. The following section illustrates pictorially the steps involved in calculating the local band-limited contrast for an image and demonstrates the way it is used to generate simulations of image

appearance. The same basic method is applied to the simulation of vision with the central retina and the simulation of image appearance with peripheral retina. Preliminary results of psychophysical testing of the validity of these simulations are described as well.

3. Appearance of Images Using Central Retina

The pyramidal image-contrast architecture described above enables the use of non-linear processing to simulate the appearance of centrally fixated images point-by-point

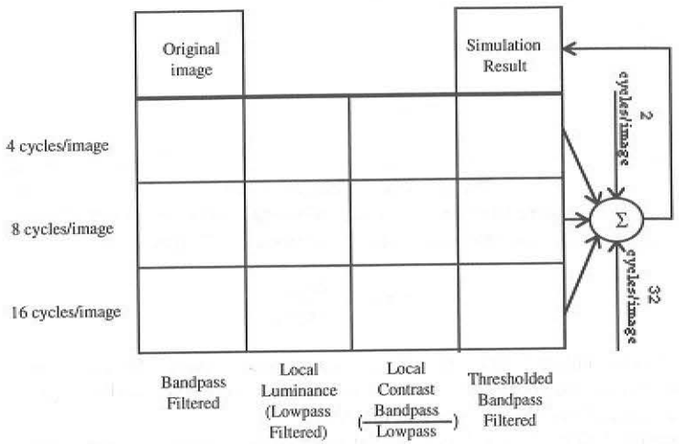


Figure 3a. Schematic illustration of the process of generating the simulations depicted in Fig. 3b. Upper left, the original image; top right, the final simulated appearance of the same image to the observer. The three rows represent processing at different spatial frequencies on the pyramid. The left column in each row represents the bandpass filtered image obtained from the original image. The second column in each row represents the corresponding lowpass filtered version for the same scale, i.e. all the energy below the band represented in the first column. The third column represents the contrast images. The bandpass filtered arrays are bipolar and a DC level of 128 has been added arbitrarily in order to present those arrays as images. Images in the right column represent the thresholded bandpass filtered images. For each image in the third column, each point was tested against the threshold value for the corresponding spatial frequency. If the contrast of the image at that point is above threshold, the corresponding point from the left image is maintained and reproduced in the right column. If the contrast at a point is below threshold, the corresponding point is set to zero (gray) in the right image. The simulated image, top right, is generated by summing all of the images in the right column. Actual processing included two additional rows, one at 2 cycles per image and one at 32 cycles per image, not shown.

Figure 3b (facing page). The process of simulating the appearance of an image (spanning 4 deg of visual angle) to a low vision patient. The rows and columns are explained in the schematic of Fig. 3a.

and for every spatial frequency band in the image. This process is illustrated in Fig. 3 using the CSF of a patient with a central scotoma due to macular disease. (A patient's CSF was used to obtain more noticeable effects than those obtained with a normal observer's CSF.) The images were sectioned in the frequency domain into 1-octave bands.

Contrast at each spatial position was calculated by dividing, for each pixel location, the bandpass filtered pixel value by the lowpass filtered pixel value at the same point. At each pyramid (scale) level, every point was compared with the appropriate observer's contrast threshold for the corresponding spatial frequency. If the contrast at that point was higher than the threshold, the amplitude of that point was not affected. If the contrast at the point was below threshold, the amplitude was set to zero. Thus, the final image in Fig. 3, top right, represents the appearance of the original image to a patient using the

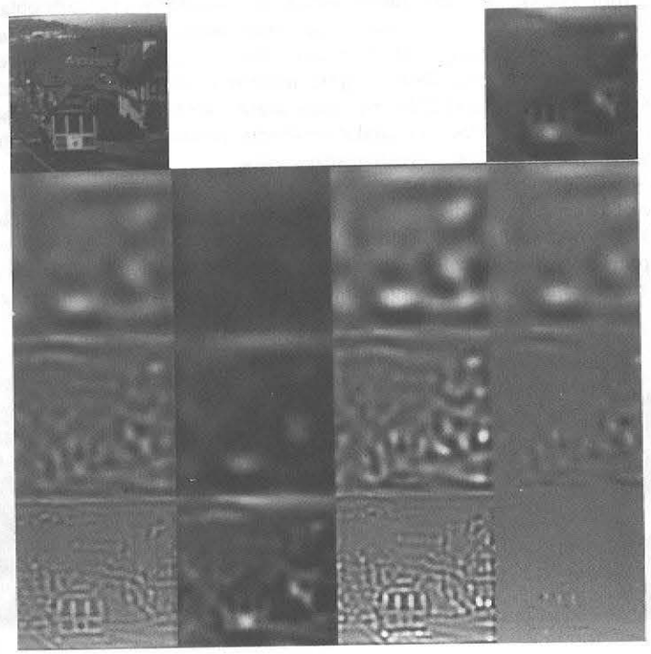
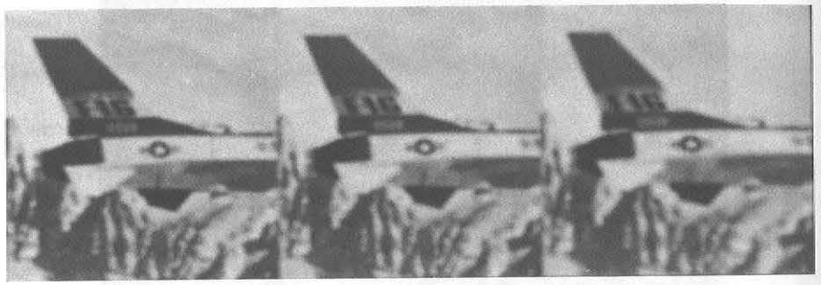


Fig. 3b

same retinal area used to obtain the CSF data. This image was processed with the stipulation that the image span 4 deg of visual angle. On this scale, this patient's visual loss had little effect on information at 4 cycles per image (corresponding to 1 c/deg, top row). A small effect may be noted at 8 cycles per image (middle row), and a substantial effect at 16 cycles per image (corresponding to 4 c/deg, bottom row). The complete processing also included the band at 2 cycles per image and 32 cycles per image, neither of which is shown. The simulated image maintains the full-contrast appearance reported by patients with central visual loss and clear media and is not faded or washed out, as may be the appearance of images simulated with linear filtering.

To simulate the appearance to a normal observer, the mean CSF from 14 normal observers[23] was used. The CSF data were obtained from a task requiring discrimination of horizontal from vertical gratings. Simulations were calculated using CSFs obtained with both 1-octave Gabor patches and fixed-aperture extended gratings. The appearance of an image to a subject with a given CSF depends on the viewing geometry assumed for the modeled observation situation. For example, in the simulations shown in Fig. 4, we assumed that the scene spanned 2 deg of visual angle. To evaluate the simulations and avoid the double-pass problem, the reader should first examine the original image (Fig. 4a) from a distance of 1 meter, at which it spans 2 deg.

This appearance should then be compared with the simulated images examined from a distance of 25 cm. At this distance, the simulations calculated using the patch-CSF (1-octave) (Fig. 4c) should resemble the appearance of the original when viewed from 1 meter. Specifically, the loss of fine details is visible. On the other hand, the simulations based on the fixed-aperture CSF (Fig. 4b) viewed at a distance of 25 cm appear to be almost identical to the original image. Thus, the appearance of the patch simulation at 25 cm more closely represents normal perception of the original (2 deg) image, suggesting that the 1-octave bandwidth Gabor patch may be a better stimulus if one wishes to measure a CSF that characterizes the normal appearance of complex images[25].



a b c

Figure 4. Simulations of the appearance of an image under several viewing conditions. a) Original image. b) Image processed to simulate the original's appearance when it spans 2 deg of visual angle based on the CSF measured with fixed-aperture gratings. c) Same as b, but processed using the CSF measured with the 1-octave patch stimuli. The printed images should span 2 deg at a distance of 1 meter (about 29 times the image width).

A second way to view these images is to examine all three of them side by side from different distances. The original image should appear indistinguishable from the simulated image when both are viewed from a distance larger than that assumed in computing the simulation. However, as the two images are moved closer to the observer, the simulation should be easy to discriminate from the original. The reader may verify that these relations appear to hold when the patch simulation is compared to the original image, but the fixed-aperture simulation is not easily distinguished from the original even at short distances. These effects may be difficult to discern in the printed photographs due to problems in obtaining accurate reproduction. However, both effects are visible when observed on a calibrated display. A formal application of this method to test the simulations is described below.

Figure 5 depicts the relationship between the measured contrast thresholds and the amplitude spectra of the image at various viewing distances. The radially averaged amplitude spectrum of the image gives the approximate contrast at each frequency expressed in cycles per face. (It is only approximate because in the simulations we were working with local contrast, not amplitude.)

The same logic applies to both the patch-CSF simulation (Fig. 5, bottom) and the fixed-aperture CSF simulation (Fig. 5, top). The dashed line in each panel represents the radially averaged amplitude spectra of the real (thin line) and simulated (thick dashed line) images when they subtend 2 deg on the retina. The solid curves represent the spectra of the same images when they subtend 4 deg. Any information in the image that falls below the observer's threshold (i.e., to the right of the point at which the contrast threshold curve intersects the image spectrum curve) is not visible to the observer. To illustrate this, the simulation should (as is shown) remove all that information (thick curve). If the original and simulated images are viewed from the simulated distance or farther (subtending 2 deg or less), they should be indistinguishable because the reader's CSF removes the same information from the original that was removed in the simulation. However, if the original and simulation are viewed from a closer distance, the difference in content (shaded area) between the original and the patch simulation should be visible.

Figure 5 is useful only to illustrate the logic of the simulations. The analysis it represents does not provide the information obtainable from the simulations. The effects of contrast threshold on apparent contrast in the image are local, not global. The effective contrast is not represented accurately by the radially averaged amplitude spectrum, and the simulation algorithm is not represented accurately by the filtering.

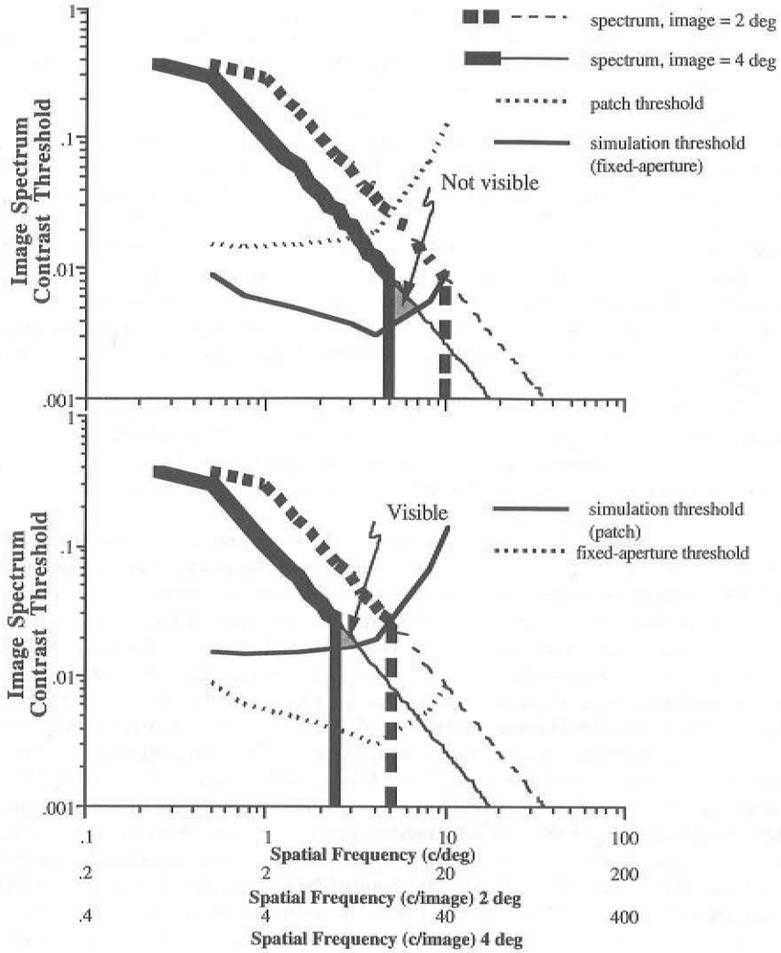


Figure 5. See caption opposite page.

This analysis, used in conjunction with the results of the observations of Fig. 4, described above, lead to the following conclusion: The loss of detail depicted by the shaded area is visible in the case of the patch-CSF simulation (Fig. 5, bottom and Fig. 4c), but not for the fixed-aperture CSF simulation (Fig. 5, top and Fig. 4b), indicating that the simulation-viewer's actual threshold lies above the shaded area in Fig. 5, top and below the shaded area in Fig. 5, bottom. Therefore, the patch-CSF simulation more closely represents the visibility of detail to a normal observer. One can test the validity of the simulations and the CSF data used by formally measuring the observations described above. Preliminary results from such testing are described next.

3.1. Testing the Simulations of Central Vision

The simulations can be tested by presenting the original image side by side with the simulation. If the simulations are valid, the simulated image and the original will be indistinguishable from a distance equal to or larger than the distance assumed in the simulation. The two images should be progressively easier to distinguish at distances shorter than the simulated distance.

Observers viewed image pairs from various distances and were asked to make a forced-choice distinction between the simulated and the original image. In each presentation one of the images was the original and the other was one of the simulations. We used four different images in this experiment. From each image, we calculated three simulated views representing views from three different distances. For the three distances (40, 80, and 160 in) the images spanned visual angles of 4, 2, and 1 deg, respectively. The subjects observed the image pairs from six distances, including shorter (20 in) than the shortest simulated distance and longer (300 in) than the longest simulated distance. From each observation distance the percent correct identification of the processed/simulated image was tabulated. The data for each simulated distance were fitted with a Weibull psychometric function to determine threshold at a 75% correct level of performance (Fig. 6). The CSF data used in the simulations were obtained for each sub-

Figure 5. Relationships among spatial frequency spectra of images and contrast thresholds. Spatial frequency is expressed in c/deg and cycles/image for different image sizes. Thin solid-line spectrum: 2 deg image; thin dotted-line spectrum: 4 deg image. Top: Simulation using fixed-aperture CSF. Bottom: Simulation using patch CSF. The medium thickness solid line represents the CSF used for simulation. Contrast below that curve is below the simulated subject's contrast threshold. Therefore, we remove image components to the right of the point where the threshold curve intersects the 2 deg image spectrum (thick dotted line). At the 2 deg distance, the removed components are below threshold and thus the original image and simulation should appear identical. When the original and simulation are moved to the 4 deg distance, a portion of the removed components (shaded area) will be above threshold and visible if the CSF used for the simulation is an accurate description of the viewer's visual system. Viewers can see the difference at 4 deg between the patch simulation and the original (Fig. 4a and c) but cannot see the difference between the fixed-aperture simulation and the original (Fig. 4a and b), indicating that the viewer's threshold curve lies below the shaded area of Fig. 5, bottom and above the shaded area of Fig. 5, top.

ject individually. In the first set of experiments, simulated images were produced with CSFs obtained using 1-octave Gabor patches and the observer's task was to discriminate gratings of horizontal orientation from those of vertical orientation[23]. Three observers participated in this experiment and their results were similar.

Data from one subject (Fig. 6) illustrates that the simulations generated using the CSF obtained with a discrimination of orientation task of 1-octave stimuli did not support our hypothesis. The hypothesis tested is that if the simulations are veridical the fitted curves should cross the 75% correct level at the simulated distance, marked by diamonds at the bottom of the graph. It appears that the observer could detect the changes at greater distances from the screen than the distance assumed in the simulation. The next set of experiments involved simulated images produced with the CSF obtained for the same type of stimuli but using a simple detection task. When we used CSF data obtained using a detection task for the same 1-octave stimuli, the results were much closer to the predictions. The results from one subject are illustrated in Fig. 7. These results illustrate that,

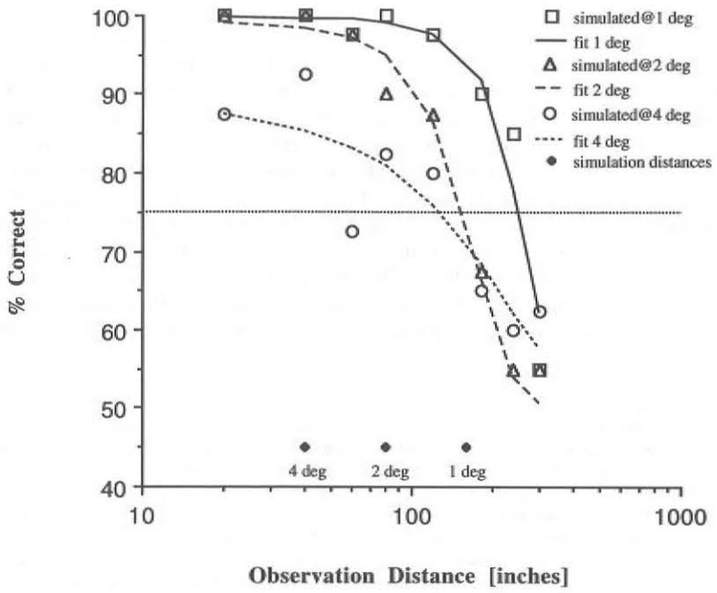


Figure 6. Results of testing central vision simulation using the CSF based on discrimination of orientation of 1- octave Gabor patches. The data and the psychometric function fits indicate that the subject could distinguish the simulation from the original at distances larger than the distances assumed in the simulations (see diamond inserts at bottom of graph).

using this methodology, we can reject the simulations generated using the orientation discrimination threshold but not the simulations using the contrast detection threshold. The differences between the CSFs used in these two modes of testing are illustrated in Fig. 8.

4. Vision with Central Scotoma

Vision with central visual field loss (scotoma) frequently is simulated using a black spot covering a central part of an image. Image details outside the simulated area of the scotoma remain unaltered and thus appear sharp and clear. This simplistic representation is static and does not represent the patient's ability to see details of interest by moving his or her eyes and thus the scotoma. In addition, these simulations, which have been applied to both still print images[5] and video motion picture simulations[30], fail to represent the reduced visual capabilities of the patient's functioning peripheral retina. Patients rarely report their visual experience as that of losing a part of the visual field. Instead they commonly complain of blurred or foggy vision. Experiencing the field loss requires directed

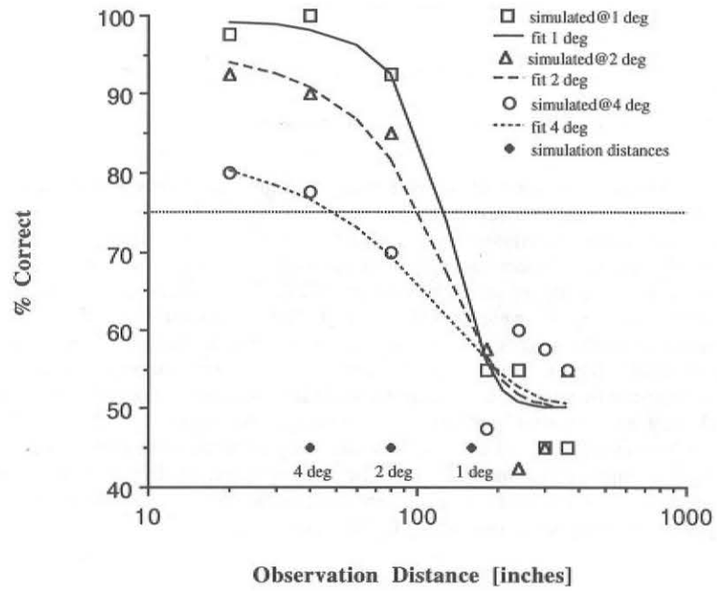


Figure 7. Results of simulation testing using the CSF based on detection of 1-octave Gabor patches. Here the subject could distinguish the simulation from the original approximately at the distance assumed in the simulations (see diamond inserts at bottom of graph).

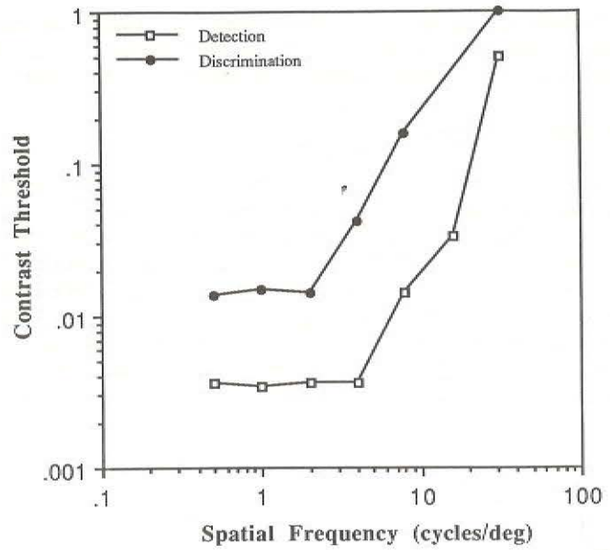


Figure 8. The contrast threshold (1/CSF) used in the simulation testing shown in Figs. 6 and 7.

examination. Our simulations attempt to illustrate what the patient sees with residual vision using the functioning area next to the scotomatous retina[26,28].

In one mode of such simulation, a patient's measured CSF is used to represent the vision outside the scotoma. Processing equal to the process described above for the simulation of normal vision is applied using the measured CSF. The result of this type of simulation is illustrated using one patient's CSF in Fig. 9. For comparison, a linear simulation of the appearance of the images to a cataract patient with the same CSF is illustrated as well (Fig. 9, right column). The nonlinearly simulated image maintains the full-contrast appearance reported by patients with central visual loss and clear media and is not faded or washed away, as is the simulated appearance of images seen through cataracts[27]. This type of simulation was used by Peli et al.[26] to tune the parameters of an enhancement algorithm used to improve face recognition by the visually impaired (Fig. 9, bottom row). For this simulation it is assumed that the patient is using the same functioning retinal area for both examining images and responding to CSF test patterns.

4.1. Vision with Nonuniform Retina

Another way to simulate and understand vision with central scotoma is to include in the simulation the variability in sensitivity across the retina of a normal observer and then examine the consequences of losing the central area to the way images will be perceived in the periphery.

Spatial inhomogeneities are an important feature of the visual system's organization. Nonuniform processing starts with the spatially variable sampling rate of the photoreceptors, featuring a high density centrally and gradually decreasing density toward the retinal periphery. The nonuniform organization continues throughout the system up to the retinotopic mapping of the visual field onto the surface of the striate cortex[32]. This nonuniform structure, a visually unfavorable characteristic of the visual system (commonly considered biologically necessary), is assumed to be a response to limitations of the nonuniform imaging by the eye's optics or a method to reduce data rate in response to limited processing capabilities[40].

In contrast, we showed that the changes in contrast sensitivity across the retina may play an important role in maintaining size (distance) invariance[28]. The spatial frequency spectra of images frequently are defined in terms of cycles per picture-width rather than cycles per degree[6,11]. This is done under the assertion that "form perception is largely independent of distance" (ref. 6, p. 196). Such distance invariance has been reported for identification of bandpass filtered letters embedded in Gaussian noise[20], for recognition

Figure 9. Simulations of the appearance of original images (top row) and enhanced face images (bottom row) for patients with central scotoma and cataract. Both patients were assumed to have the same contrast sensitivity function, and the images were assumed to span 4 deg of visual angle. The left column represents the original and enhanced images. The middle column simulates the appearance of both images to a patient with central scotoma (nonlinear processing). The right column illustrates the appearance of the same images to a patient with cataract (linear filtering). Note the improvement in visibility of detail for both simulated patients when the images are enhanced.

of bandpass filtered face images[13], and for identification of toy tanks' images[19]. Invariant perception also requires that important image features remain visible (do not drop below threshold) with changes in retinal spatial frequencies associated with the change in distance. This property of the visual system that causes its *detection* of image contrast to be nearly invariant to changes in size due to distance changes should be included in a proper simulation of vision. We have shown that the threshold distance invariance away from the fovea is as good as the invariance at the fovea[28]. For spatial frequencies straddling the peak of the CSF, the deviation from optimal invariance is small. The invariance model contains the foveal CSF in addition to the fundamental eccentricity constant (FEC), representing the drop in contrast sensitivity as a function of retinal eccentricity. Specifically, the FEC is the slope of the line representing contrast threshold as a function of eccentricity for a stimulus of 1 c/deg on a log-log graph. The CSF across the whole retina for all spatial frequencies can be described as a function of the foveal CSF and the FEC.

This enables the simulation of the appearance of wide angle images that are affected by the nonuniformity of the visual system. By using the same pyramidal image-contrast structure discussed above, together with the variation across the visual field, we can simulate the appearance of images with a nonuniform retina. As above, the image is sectioned into a series of bandpass filtered versions of 1-octave bandwidth. For each section, we calculated the corresponding local band-limited contrast for each point in the image. We then assigned a fixation locus, and for each point in the image the distance from this fixation point was determined. Based on that distance, in degrees, and the spatial frequency associated with the bandpass filtered version, each point can be tested against the appropriate threshold at that location to determine if it will be visible. A suprathreshold point is left unchanged while a subthreshold point is set to zero contrast. The appearance of the same image from two different distances is obtained by applying the same method of simulation while assigning a different angular span to the image depending on the observation distance. The simulations presented in Fig. 10 were obtained using the fit to the data of Cannon[2].

The simulated appearance of a scene to a normal observer from a distance where it spans 32 deg of visual angle is compared in Fig. 10 (top) to the appearance of the same scene spanning only 2 deg. In both cases, fixation is assumed to be at the center of the image, and there is little apparent change in image quality at different eccentricities. Furthermore, the two images appear similar despite the large difference in observation distances. The only noticeable effect is a slight blurring of the fine details of the "smaller" image. Such effects are associated with the suboptimality of the invariance and are included in our simulation. This constancy breaks down for an observer with a central visual loss of 5 deg radius as illustrated in Fig. 10 (bottom). Here we assumed that the observer placed the image of the scene on the retina adjacent to the scotomatous area, i.e. on the most centrally available functional retina[35]. At 32 deg, the wide field simulated image appears almost the same to the low vision observer as to the normal observer. However, when the scene spans only 2 deg of visual angle, the low vision observer suffers from a substantial loss of detail, and the invariant appearance of the image cannot be

Figure 10. Changes in the appearance of a scene with a large change in observation distance are simulated for a normal observer (top) and for a patient with a 10 deg-diameter central scotoma (bottom). Images on the left represent the appearance of the scene to the observers when it spans 32 deg of visual angle. Images on the right represent the appearance of the same scene to the observers when it spans only 2 deg of visual angle. The normal observer is assumed to fixate at the center of the image in both cases. The patient is assumed to place the edge of the scotoma at the left edge of the image in both cases. The changes in appearance for the normal observer are small, limited to the very high spatial frequencies and compatible with the filtering of the image by the eye's optical media. The effect of change in distance is much more detrimental for the patient with a central scotoma. Note that at close range, the appearance of the scene to the patient is almost identical to its appearance to the normal observer. The effect in both cases is somewhat under represented in the simulation because the 512×512 pixels image at 32 deg contains information only up to 8 c/deg. The natural scene contains information at higher frequencies.

maintained. Thus, the same mechanism that serves to maintain the appearance of the image for normals across this 16-to-1 change in distance results in a deterioration of the appearance of the image for a patient with central scotoma.

4.2. Testing the Simulations of Nonuniform Retina

In testing the simulations of the appearance of wide images we look for the ability to distinguish the simulation from the unprocessed original using peripheral retina. The hypothesis in this case is that the two images become indistinguishable when the simulation closely represents the losses of image detail that occur with eccentricity. Changing the distance, however, should have little effect if our model of image invariance with change in distance[28] is of any value. In this case we also wish to determine which CSF stimuli are better estimaters of observers' perception and whether our estimate of the FEC is valid.

The testing stimuli were generated from two high resolution (1024×1024) images processed to simulate their appearance when fixated at the center and spanning 64 deg of visual angle[24]. Half of an original (unprocessed) image was displayed side by side with the simulation applied to the mirror image of the same image (Fig. 11) using two different CSF data sets. These data sets were obtained using 1-octave Gabor patches[28], and Cannon's[2] fixed aperture 2 deg extended gratings. We implemented these two types of simulations with the FEC values found for two tasks, 0.035 and 0.055 for contrast detection and discrimination of a vertical grating from a horizontal grating, respectively[28], and with two additional arbitrary levels on each side. The images were displayed on a large screen projection TV spanning 64 deg of visual angle. Image pairs were presented abruptly for 167 msec to prevent eye movements. Subjects were required to identify the processed (slightly blurred) half of the screen while maintaining fixation at the center of the screen. The central 6.4 deg of the field was masked with a gray circular patch of mean image luminance. Note that although these images contained no information above 8 c/deg the size of the central mask assures that for most of the image, the sensitivity of the observer's peripheral retina that had to be used was so poor that even with 100% contrast, higher frequencies could not be detected.

The most important result of this experiment is that the data actually fell along the active slope of the psychometric function (Fig. 12). Our prediction was that if the simulations were veridical, the threshold FEC (arbitrarily set to 75% correct response) should be close to either the 0.035 or the 0.055 values, where it is indeed found. This initial finding is encouraging. It illustrates that the simulations are not grossly incorrect. It would be easy, with simulations based on basic visual function measurements, to have images which are either completely distinguishable or completely indistinguishable.

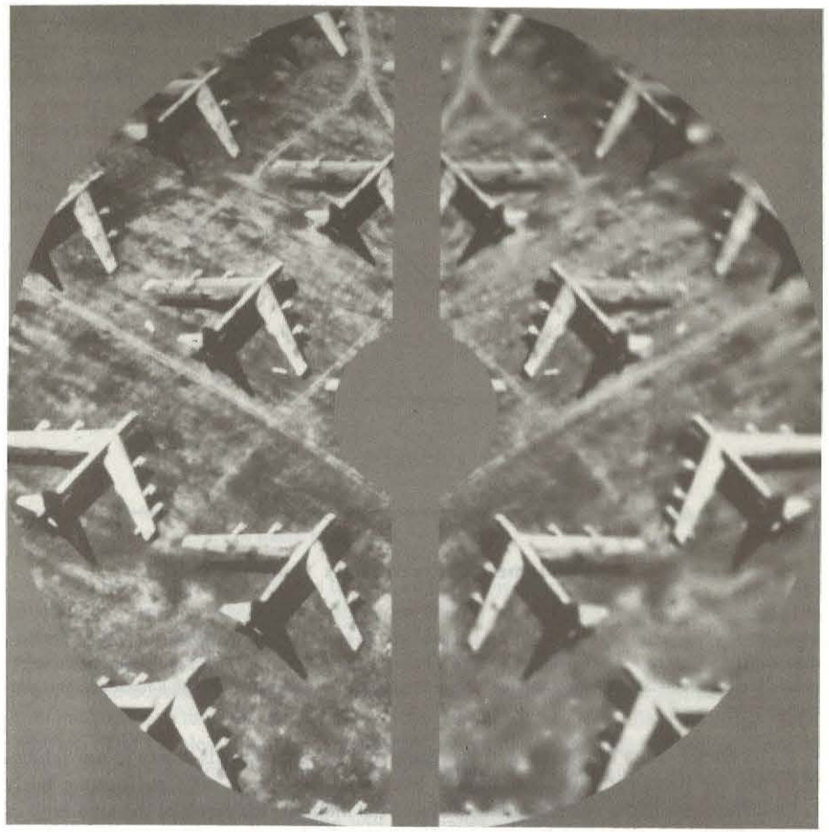


Figure 11. An example of the stimuli used in testing the simulation of wide images. The original unprocessed image on the left was compared with the simulation using the CSF obtained with Gabor patches of 1-octave bandwidth and an FEC of 0.1 on the right. This high value of the FEC results in clearly visible blurring at the edge of the image. The whole image spanned 64 deg at the observer's eye, and the central 6.4 deg were masked as shown. (This work was done in collaboration with G. Geri.)

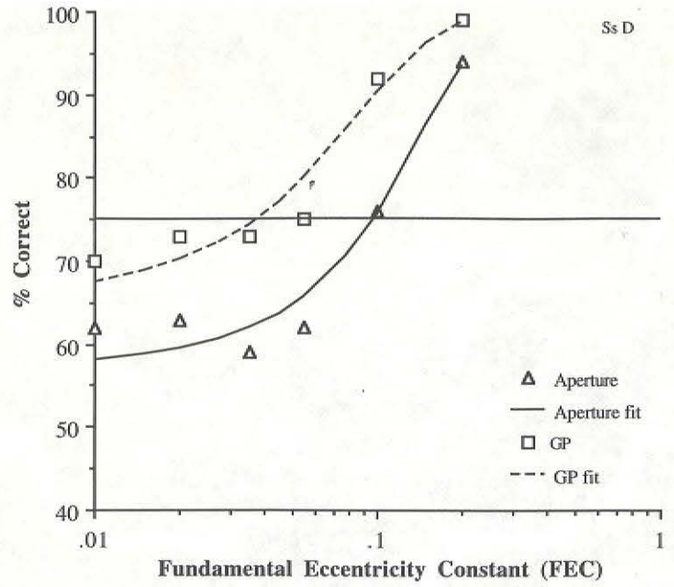


Figure 12. Results of testing the wide field simulations. Percent correct discrimination of the original from the simulation is plotted as a function of the FEC. Data are shown for two sets of simulations, one using 1-octave Gabor patch-CSF data (GP) and the other using Cannon's[2] data obtained with a fixed-aperture grating stimulus.

The Gabor patch based simulations resulted in FEC thresholds closer to our prediction than the fixed-aperture simulations. But this difference is not large despite fairly large differences in the corresponding CSFs. This might be the result of the similarity of the two CSF data sets at low frequencies (Fig. 13). The low frequencies may have played a more important role in the peripheral vision simulated and tested here than the higher frequencies where the CSFs, and therefore the simulations, differ.

A surprising result was that even for very low FEC values the detection of the simulation was much better than 50% (Fig. 12). Such a result is possible since even for an FEC of 0, the images are processed (using the effect of the foveal CSF), and therefore differ from the original. However, it is important to note that for the low values used, the simulations were processed so little that they were practically indistinguishable when examined carefully side by side on the screen with foveal vision for unlimited time. We believe that the high level of discrimination in the periphery is a result of the abrupt presentation[23]. However, extending the time of presentation to 500 msec did not change the results.

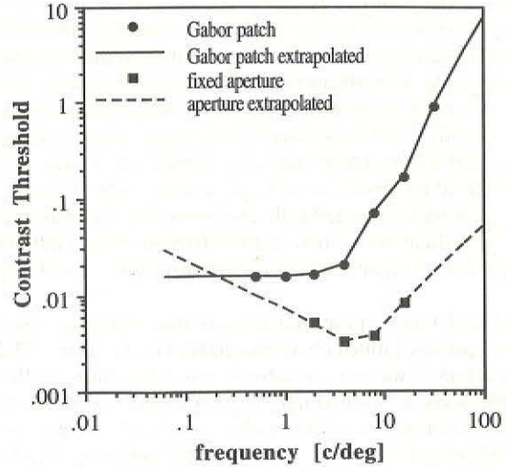


Figure 13. The foveal contrast threshold data sets used in testing the wide field simulations. Note the convergence of the two sets at low spatial frequencies.

Another consistent result was that for all our subjects, one of the images was easier to distinguish than the other. This difference should be explored further to determine the nature of the image content that led to this result. The issue of image dependence has been largely ignored. Watson[39] has recently shown the benefits that can be obtained from image dependent compression coding. The consideration of image dependent factors in simulation may also be of value in the analysis of image quality and image quality metrics. The simplest assumption is that the image dependent effect found is due to the relative spatial frequency content of different images. It is plausible that if an image has little high frequency information the original and the simulation will be harder to discriminate, since the effect of simulation will be smaller. The proper metrics for such spatial frequency content have not yet been determined.

5. Discussion

The basic assumption of this work is that local image contrast should be expressed as the dimensionless ratio of the local amplitude and the local average luminance in a way similar to that expressed in the definition of Michelson contrast or Weber fraction. The use of such a ratio implies that the human's sensitivity to the amplitude of change in luminance varies with the adaptation level associated with the local average luminance[31]. This is known to be the case for threshold contrast sensitivity at all spatial frequencies at high luminance levels. For low frequencies (less than 4 c/deg), the same relation is true

for a large portion of the photopic range[14]. For the rest of the spatial frequencies and luminance ranges, the DeVries-Rose law applies, representing only partial adaptation.

Partial adaptation may be included in the present definition of contrast simply by reducing the effect of the local luminance mean on the high-frequency contrast to some degree. The degree or level of adaptation in suprathreshold contrast sensitivity has not been determined until recently. One experiment using dichoptic presentation found that contrast matching at a high-contrast level indeed approximated contrast as defined by the ratio of amplitude to local luminance mean[9] over a wide range of luminances. In a recent study, we found that perceived suprathreshold contrast does decrease with low luminance levels when testing is done under more natural free viewing conditions[29]. The effect, however, is of significant magnitude only for the very dark parts of displays viewed in a dark room.

García-Pérez[8] and Fleck[7] proposed models incorporating visual inhomogeneity based on multiple, spatially limited channels centered in the fovea. The spatial extent of various frequency channels was measured for sinusoidal gratings. In these models, which were used for simulations, for each spatial frequency there is only one threshold, limiting all features of that frequency to a fixed radius around the fovea. In our simulations, higher-contrast features will be visible farther into the periphery than lower-contrast features of the same spatial frequency. Their models were suggested as useful tools to analyze various visual perceptual phenomena such as the Gestalt frame of reference effect[8]. However, the role of the nonuniform visual system in aiding size-distance invariance was not considered.

This model[22] provides a new, more powerful tool to analyze the visibility of displays[3], generate equal visibility displays, or generate displays of pre-designed variable visibility. The model generalized the idea underlying Anstis'[1] equal visibility acuity chart. This generalization is achieved by adding the dimension of contrast and by providing a computationally efficient algorithm for its application to any image.

5.1. Future Work

Contrast measured by filtering, as suggested here, defines only incremental or decremental changes from the local background. This is analogous to the symmetric (cosine phase) responses of mechanisms or cells in the visual system. Another type of contrast may be defined as a transition from low to high luminance, or vice versa, in a band-limited signal[12]. The latter may be viewed as the response of the antisymmetric (sine phase) mechanisms. A complete description of contrast in a complex image should include both of these contrast representations in a way similar to the analysis of Stromeyer and Klein[33]. Incorporation of both symmetric and antisymmetric responses in a one-dimensional case using oriented filters is straightforward[17]. Complete two-dimensional application is difficult due to the lack of definition of Hilbert's transform for the two-dimensional case[21]. Without such a definition the contrast measure cannot be expanded to include the representation of antisymmetric mechanisms.

This work so far has addressed only the effects of contrast on perception at threshold. All of our work assumed contrast constancy to hold for suprathreshold contrast levels. The changes in suprathreshold contrast perception, and their effects on the perception of images, have to be examined and incorporated into an expanded model of contrast perception used in simulations. To that end, the proper metrics for the measurement of contrast of local features of various types in an image should be determined. We have already shown that a number of candidate metrics do not hold even for relatively simple patterns[12]. The effects of luminance on suprathreshold contrast perception were tested only for one spatial frequency. However, since the effect is spatial frequency dependent at threshold[37], the role of spatial frequency should be investigated at suprathreshold levels as well.

Acknowledgments

Supported in part by NIH grant EY05957 and by grants from the Ford Motor Company Fund and DigiVision, Inc. I thank R. Goldstein for valuable programming help, G. Young for help in data collection, and G. Geri for enabling the testing of the wide field simulations. Elisabeth Fine provided valuable editing help, and Miguel García-Pérez wrote an outstanding, detailed review that greatly improved the quality of this chapter.

References

1. S. N. Anstis, "A chart demonstrating variations in acuity with retinal position," *Vision Res.* **14** (1974) 589-592.
2. M. W. Cannon Jr., "Perceived contrast in the fovea and periphery," *J Opt. Soc. Am. [A]* **2** (1985) 1760-1768.
3. S. Daly, "The visual differences predictor: An algorithm for the assessment of image fidelity," *Proc. of the SPIE Vol. 1666 Human Vision, Visual Processing, and Digital Display III* (1992) 2-15.
4. M. Duval-Destin, "A spatio-temporal complete description of contrast," *Digest of Technical Papers Society for Information Display* (1991) 615-618.
5. E. Faye, "Functional classification of eye disease," in *Clinical Low Vision.* ed. E. E. Faye (Little Brown & Co., Boston, 1976), p. 218.
6. A. L. Fiorentini, L. Maffei, and G. Sandini, "The role of high spatial frequencies in face perception," *Perception* **12** (1983) 195-201.
7. H. J. Fleck, "Measurement and modeling of peripheral detection and discrimination thresholds," *Biol. Cybern.* **61** (1989) 437-446.
8. M. A. García-Pérez, "Visual inhomogeneity and reference frames," *Bulletin of Psychonomic Society* **27** (1989) 21-24.
9. M. A. Georgeson and G. D. Sullivan, "Contrast constancy: Deblurring in human vision by spatial frequency channels," *J Physiol.* **252** (1975) 627-656.

10. A. P. Ginsburg, "Is the illusory triangle physical or imaginary?," *Nature* **257** (1975) 219-220.
11. A. P. Ginsburg, "Visual information processing based on spatial filters constrained by biological data," *AMRL-TR-78-129 Vol. I and II* (Wright-Patterson Air Force Base, OH, 1978).
12. R. Goldstein, E. Peli, and G. Young, "Matching the contrast of luminance increments, decrements, and transitions," *ARVO Abstracts Invest. Ophthalmol. Vis. Sci.* **31(4, suppl)** (1991) 1271.
13. T. Hayes, M. C. Morrone, and D. C. Burr, "Recognition of positive and negative bandpass-filtered images," *Perception* **15** (1986) 595-602.
14. D. H. Kelly, "Visual contrast sensitivity," *Optica Acta* **24** (1977) 107-129.
15. B. L. Lundh, G. Derefeldt, S. Nyberg, and G. Lennerstrand, "Picture simulation of contrast sensitivity in organic and functional amblyopia," *Acta Ophthalmol. (Kbh)* **59** (1981) 774-783.
16. J. Larimer, "Designing tomorrow's displays," *NASA Tech Briefs* **17(4)** (1993) 14-16.
17. M. C. Morrone and D. C. Burr, "Feature detection in human vision: A phase-dependent energy model," *Proceedings of the Royal Society of London* **B235** (1988) 221-245.
18. B. Moulden, F. Kingdom, and L. F. Gatley, "The standard deviation of luminance as a metric for contrast in random-dot images," *Perception* **19** (1990) 79-101.
19. J. Norman and S. Ehrlich, "Spatial frequency filtering and target identification," *Vision Res.* **27** (1987) 87-96.
20. D. H. Parish and G. Sperling, "Object spatial frequencies, retinal spatial frequencies, noise and the efficiency of letter discrimination," *Vision Res.* **31** (1991) 1399-1415.
21. E. Peli, "Hilbert transform pairs mechanisms," *ARVO Abstracts Invest. Ophthalmol. Vis. Sci.* **30(4, suppl)** (1989) 110.
22. E. Peli, "Contrast in complex images," *J Opt. Soc. Am. [A]* **7** (1990) 2030-2040.
23. E. Peli, L. Arend, G. Young, and R. Goldstein, "Contrast sensitivity to patch stimuli: Effects of spatial bandwidth and temporal presentation," *Spatial Vis.* **7** (1993) 1-14.
24. E. Peli and G. Geri, "Putting simulations of peripheral vision to the test," *ARVO Abstracts Invest. Ophthalmol. Vis. Sci.* **34 (4, suppl)** (1993) 820.
25. E. Peli, R. B. Goldstein, G. M. Young, and L. E. Arend, "Contrast sensitivity functions for analysis and simulation of visual perception," *Noninvasive Assessment of the Visual System* **3** (1990) 126-129.
26. E. Peli, R. B. Goldstein, G. M. Young, C. L. Trempe, and S. M. Buzney, "Image enhancement for the visually impaired: Simulations and experimental results," *Invest. Ophthalmol. Vis. Sci.* **32** (1991) 2337-2350.
27. E. Peli and T. Peli, "Image enhancement for the visually impaired," *Opt. Eng.* **23** (1984) 47-51.
28. E. Peli, J. Yang, and R. Goldstein, "Image invariance with changes in size: The role of peripheral contrast thresholds," *J Opt. Soc. Am. [A]* **8** (1991) 1762-1774.
29. E. Peli, J. Yang, R. Goldstein, and A. Reeves, "Effect of luminance on suprathreshold contrast perception," *J .Opt. Soc. Am. [A]* **8** (1991) 1352-1359.

30. D. Pelli, *What is Low Vision?*, video tape presentation (Institute for Sensory Research, Syracuse University, Syracuse, NY, 1990).
31. J. G. Robson, "Linear and non-linear operations in the visual system," *ARVO Abstracts Invest. Ophthalmol. Vis. Sci.* **29(4, suppl)** (1988) 117.
32. E. L. Schwartz, "Spatial mapping in the primate sensory projection: Analytic structure and relevance to perception," *Biol. Cybern.* **25** (1977) 181-194.
33. C. F. Stromeyer III and S. Klein, "Evidence against narrow-band spatial frequency channels in human vision: The detectability of frequency modulated gratings," *Vision Res.* **15** (1975) 899-910.
34. L. N. Thibos and A. Bradley, "The limits of performance in central and peripheral vision," *Digest of Technical Papers Society for Information Display* **22** (1991) 301-303.
35. G. T. Timberlake, E. Peli, E. A. Essock, and R. A. Augliere, "Reading with macular scotoma. II: Retinal locus for scanning text," *Invest. Ophthalmol. Vis. Sci.* **28** (1987) 1268-1274.
36. C. W. Tyler, "Is the illusory triangle physical or imaginary?," *Perception* **6** (1977) 603-604.
37. F. L. Van Nes and M. A. Bouman, "Spatial modulation transfer in the human eye," *J. Opt. Soc. Am.* **57** (1967) 401-406.
38. A. B. Watson, "The cortex transform: Rapid computation of simulated neural images," *Computer Vision, Graphics, and Image Processing* **39** (1987) 311-327.
39. A. B. Watson, "DCTune: A technique for the visual optimization of DCT quantization matrices for individual images," *Society for Information Display '93 Digest* (1993) 946-949.
40. Y. Y. Zeevi, N. Peterfreund, and E. Shlomot, "Pyramidal image representation in nonuniform systems," *Proceedings of the SPIE* **1001** (1988) 563-571.