Contrast in complex images

Eli Peli

Eye Research Institute, 20 Staniford Street, Boston, Massachusetts 02114

Received November 30, 1989; accepted March 26, 1990

The physical contrast of simple images such as sinusoidal gratings or a single patch of light on a uniform background is well defined and agrees with the perceived contrast, but this is not so for complex images. Most definitions assign a single contrast value to the whole image, but perceived contrast may vary greatly across the image. Human contrast sensitivity is a function of spatial frequency; therefore the spatial frequency content of an image should be considered in the definition of contrast. In this paper a definition of local band-limited contrast in images is proposed that assigns a contrast value to every point in the image as a function of the spatial frequency band. For each frequency band, the contrast is defined as the ratio of the bandpass-filtered image at that frequency to the low-pass image filtered to an octave below the same frequency (local luminance mean). This definition raises important implications regarding the perception of contrast in complex images and is helpful in understanding the effects of image-processing algorithms on the perceived contrast. A pyramidal image-contrast structure based on this definition is useful in simulating nonlinear, threshold characteristics of spatial vision in both normal observers and the visually impaired.

INTRODUCTION

Apparent or perceived contrast is a basic perceptual attribute of an image. Many techniques of contrast manipulation and modification have been developed within the field of digital image processing. The study of contrast sensitivity has dominated visual perception research in the past two decades. However, the measurement and evaluation of contrast and contrast changes in arbitrary images are not uniquely defined in the literature. In this paper I propose a definition of local band-limited contrast in complex images that is closely related to the common definition of contrast in simple pattern tests. The purpose of this new definition is better to link measured physical contrast with visual contrast perception. This definition provides new insights into the perception of suprathreshold contrast in complex images and permits better simulations of the effects of the threshold nonlinear nature of contrast sensitivity on the appearance of images.

Definitions of Contrast in Simple Patterns

Two definitions have been commonly used for measuring the contrast of test targets. The contrast C of a periodic pattern such as a sinusoidal grating is measured with the Michelson formula[1]

$$C = \frac{L_{max} - L_{min}}{L_{max} + L_{min}}, \tag{1}$$

where L_{max} and L_{min} are the maximum and minimum luminance values, respectively, in the gratings. The Weber fraction definition of contrast [Eq. (2) below] is used to measure the local contrast of a single target of uniform luminance seen against a uniform background:

$$C = \frac{\Delta L}{L}, \tag{2}$$

where ΔL is the increment or decrement in the target luminance from the uniform background luminance L. One usually assumes a large background with a small test target, in which case the average luminance will be close to the background luminance. If there are many targets, or if there is a repetitive target as in the case of a grating stimulus, these assumptions do not hold. The processing of images in the visual system is believed to be neither periodic nor local; therefore the representation of contrast in images should be quasi-local as well.

The difference between the two definitions becomes apparent when the Michelson contrast is expressed similarly to the Weber contrast:

$$C = \frac{\Delta L}{L + \Delta L}, \tag{3}$$

where $\Delta L = (L_{max} - L_{min})/2$ and $L = L_{min}$. These two measures of contrast do not coincide or even share a common range of values. The Michelson contrast value ranges from 0 to +1.0, whereas the Weber contrast value ranges from -1.0 to $+\infty$. Other definitions of contrast that share similar problems [for example, $C = 2\Delta L/(2L + \Delta L)$] have been presented by Westheimer.[2] However, all the definitions represent the contrast as a dimensionless ratio of luminance change to mean background luminance.

Previous Definitions of Contrast in Images

Because of the difficulties in defining contrast in images, many definitions of contrast in a complex scene found in the literature are restricted to the assessment of contrast changes in the same image displayed in two different ways. One such definition of contrast change was given by Ginsburg.[3] For an image spanning the full range of displayed gray levels (i.e., 0–255 gray levels), the contrast was defined as 100%, but when the same image was linearly compressed to span only half of the range (i.e., 0–127), the contrast was reduced to 50%. With this definition of contrast change, the mean luminance decreases with contrast and, thus, based on some of the other definitions, the contrast should be left unchanged by compression. More commonly, the contrast change of images was evaluated by using the Michelson

0740-3232/90/102032-09$02.00

definition [Eq. (1)]. Image contrast was changed by linear scaling while the average luminance was held constant.[4] This approach appears to assess properly the relative contrast change between two presentations of the same image (difficulties with this are addressed below).

Absolute measurement of contrast using the Michelson definition is not appropriate because one or two points of extreme brightness or darkness can determine the contrast of the whole image. For example, if a single bright highlight or an especially dark shadow point is added to a fairly low-contrast image, the image Michelson contrast increases dramatically, but the perceived contrast may be decreased. For the same reason, comparison of contrast in two different images, such as two faces, may be affected largely by incidental occurrences, such as reflections from the cornea or from a small, dark birthmark.

In studying the effects of masking[5,6] by using two different images superimposed to create an intensity-mixed image, the relative intensity in percent of each image was used[6] instead of contrast. However, even this measure cannot be used when the two superimposed images are band limited in two different bands of spatial frequencies.[6,7]

A common way to define the contrast in an image so that the contrast of two different images can be compared is to measure the root-mean-square (rms) contrast.[8,9] The rms is defined as

$$\text{rms} = \left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^{1/2}, \tag{4a}$$

where x_i is a normalized gray-level value such that $0 \le x_i \le 1$ and $\bar{x}$ is the mean normalized gray level:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{4b}$$

With this definition, images of different human faces have the same contrast if their rms contrast is equal.[9] The rms contrast does not depend on spatial frequency content of the image or the spatial distribution of contrast in the image.

Loshin and Banton,[10] working with face images, recognized the need to define contrast locally in the images. They defined a local, low-contrast feature by arbitrarily measuring a local mean luminance along the chin line relative to the background and a local high-contrast feature by measuring a mean luminance of the forehead and the dark hair above the forehead.

Band-Limited Contrast

The issue of contrast of complex scenes at different spatial frequencies in the context of image processing and perception was addressed explicitly by Hess *et al.*[11] Contrast was defined in the Fourier domain as

$$C(u, v) = \frac{2A(u, v)}{\text{DC}}, \tag{5}$$

where $A(u, v)$ is the amplitude of the Fourier transform of the image, u and v are the horizontal and vertical spatial frequency coordinates, respectively, and DC is the zero-frequency component. This definition was applied globally to the whole image as well as to one-quarter or one-sixteenth subimages in nonoverlapping windows.

Local Contrast Definitions

The local nature of contrast changes across an image and spatial frequency content are related and should be considered together. This is done implicitly when the contrast of a laser speckle pattern is defined as a local rms contrast.[12] In this approach, the same definition used in Eqs. (4) over the whole image is applied locally to a small subimage of the speckle pattern. Thus for each, possibly overlapping, subimage a local rms contrast is defined, which represents the contrast in the spatial frequency band corresponding to the speckle spatial period.

Watson *et al.*[13] defined a contrast at each point for their test results, which were composed of a sinusoidal grating patch with a two-dimensional Gaussian envelope. A target was described generally as

$$I(x, y) = I_0[1 + C(x, y)], \tag{6}$$

where $C(x, y)$ is the contrast at each point and I_0 is the background luminance. For the targets used, which were band limited, this definition of contrast implicitly addresses the spatial frequency context and explicitly assigns a contrast value to every point in the image. In this scheme, however, the background luminance was constant, and only the peak contrast value for each pattern was used.

Badcock[14] defined measures of local contrast for his complex grating pattern, composed of first and third harmonics. These *ad hoc* measures were based on observers' suggestions and do not apply to any generalization for other types of pattern. Hess and Pointer[15] adapted the same definitions, but they calculated the contrast only around the peaks of the first harmonic and not around the troughs, thus ignoring the effect of the local luminance mean on the contrast of the higher harmonic. This effect is the central issue of the discussion here.

NEW DEFINITION: LOCAL BAND-LIMITED CONTRAST

To avoid many of the problems of other definitions of contrast as reviewed above, the new definition proposed here addresses several issues together. Since human contrast sensitivity is highly dependent on spatial frequency, especially at threshold, contrast for each spatial frequency band is calculated separately. The contrast at each point in the image is calculated separately to address the variation of contrast across the image. Thus we term the calculated contrast local band-limited contrast. This local band-limited contrast corresponds to the quasi-local processing in the visual system. The most important aspect of the local band-limited contrast[16] definition proposed here is that the level of the local luminance mean should be considered in calculating the contrast at every point.

To define local band-limited contrast for a complex image, we will first obtain a band-limited version of the image in the frequency domain $A(u, v)$. This can be done by using a radically symmetric, band-pass filter $G(r)$. The bandpass profile should approximate the Gaussian envelope of the Gabor function in the frequency domain. It is appropriate to select sections of 1-octave bandwidth, because they simulate the bandwidth of cortical simple cells,[17] produce an efficient image code,[18] and contain roughly equal amounts of energy in images of natural scenes.[19] Thus, in the frequency

domain, the band-limited image can be represented in the following way:

$$A(u, v) \equiv A(r, \theta) = F(r, \theta)G(r), \tag{7}$$

where u and v are the respective horizontal and vertical spatial frequency coordinates and r and θ represent the respective polar spatial frequency coordinates: $r = \sqrt{u^2 + v^2}$ and $\theta = \tan^{-1}(u/v)$, and $F(r, \theta)$ is the Fourier transform of the image $f(x, y)$.

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

$$a(x, y) = f(x, y) * g(x, y), \tag{8}$$

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

$$c(x, y) = \frac{a(x, y)}{l(x, y)}, \tag{9}$$

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. See Appendix A for details of implementation of the contrast pyramid.

IMPLICATIONS OF THE CONTRAST DEFINITION

The contrast at a spatial frequency or a band of spatial frequencies is usually considered to be dependent only on the local amplitude at that frequency. The contrast in Eq. (9) depends also on the amplitude at lower spatial frequencies. The effect of this difference can be easily appreciated with a one-dimensional, two-frequency pattern (Fig. 1):

$$f(x, y) = I_0(1 + a_1 \cos wx + a_2 \cos 8wx), \tag{10}$$

where I_0 is the mean luminance and $a_1 I_0$ and $a_2 I_0$ are the amplitude of the first and eighth harmonics, respectively.

Although the amplitude of the eighth harmonic is constant across the image, the apparent contrast is higher near the troughs of the first harmonic than near the peaks, as predicted by Eq. (9). This observation was recently verified psychophysically by Thomas.[20] The contrast of the eighth harmonic c_8 may vary in the range

$$\frac{a_2}{1 + a_1} \le c_8 \le \frac{a_2}{1 - a_1}. \tag{11}$$

For low-contrast patterns (i.e., $a_1 \ll 1$) the contrast variation across the pattern is reduced, and the contrast c_8 may be safely approximated by a_2. Thus the analysis of the results of threshold experiments will not be significantly altered by this definition of contrast in most cases. Only for high-contrast images with contrast levels of more than 0.3 should the analysis consider these local variations and their role in perception. Such contrast levels are commonly encountered in everyday images.

Many investigators have evaluated the contrast of face images and other scenes at various bands by simply bandpass filtering the image and displaying the band images added to an arbitrarily selected DC level, the mean luminance of the image, or the midrange value. However, as can be seen from Fig. 2, this will result in contrast that is substantially different from the one calculated by Eq. (9). In particular, the contrast at high-frequency bands will be much higher over dark areas of the image. In face images this frequently implies that the contrast at high-spatial-frequency bands is higher around the eyes and the mouth than the corresponding amplitudes of the bandpass-filtered image (Fig. 2). Details that are subthreshold and therefore undetected in the bandpass-filtered image and thus assumed to have no relevance to perception[21] may actually be suprathreshold in the image, add to image sharpness, and aid in recognition. The effect of this on the perception of the image may be simulated by adding in superposition the various contrast bands rather than the amplitude bands.

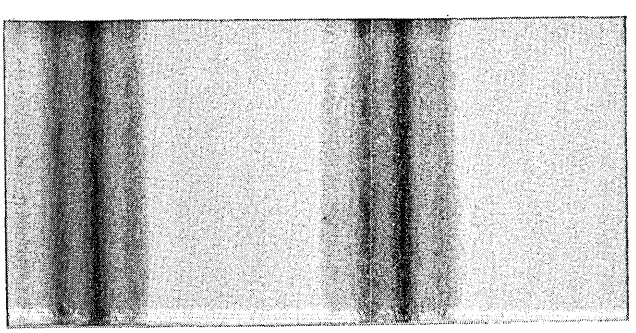

Fig. 1. Compound grating image as described in Eq. (10). The apparent contrast of the high-frequency component changes across the image although the amplitude is fixed.

Fig. 2. Comparison between bandpass amplitude image (left) and local band-limited contrast image (right) for two spatial frequencies, 16 (top) and 32 (bottom) cycles per picture. Note the relative increase of contrast around the eyes and over dark areas in the original image (at left in Fig. 3 below).

The resultant simulated perceived image (Fig. 3) is much sharper, has higher contrast, and enhances those details that occur against darker backgrounds. The details of the filters used in the generation of the images in Fig. 2 and the reconstruction in Fig. 3 are given in Appendix A.

Linear scaling of an image gray scale, a common image-enhancement technique, is frequently used to modify images to study the effect of contrast.[5,9,10] It is usually assumed that linear rescaling will change the contrast of all frequencies in the same way. Indeed, the amplitudes of all frequencies will be modified linearly by the same amount, but the contrast as defined by Eq. (9) will change differently for different spatial frequencies. For example, if in Eq. (10) we multiply a_1 and a_2 by k, the new contrast for the first harmonic will be ka_1, but the contrast of the eighth harmonic will span a new range:

$$\frac{ka_2}{1+ka_1} \le c_8 \le \frac{ka_2}{1-ka_1}. \tag{12}$$

This effect is illustrated in Fig. 4. Each of the two images at the right has two sinusoidal components of the same amplitude with the higher component of equal spatial frequency on both images. The images at the left represent equal linear rescaling of the two images at the right but result in a noticeable difference in the apparent contrast of the higher-frequency components in the two images. Thus linear rescaling of gray levels actually increases the contrast of high spatial frequencies more over dark areas than it does for the same spatial frequency over light areas, and both are changed differently from the amount of change in the contrast of low spatial frequencies.

Reverse scaling or polarity inversion of the display is sometimes used for image enhancement.[22] The fact that such a process results in enhancement of details is usually attributed to the nonlinearity of the display. However, even with a linear display, an improvement in details may be observed with such processing, while at the same time the visibility of other details is reduced. These results are clearly understandable within the framework of the contrast definition proposed here. Reversing the polarity of the display will not change the magnitude of the local high-frequency information. Contrast, on the other hand, will be increased in areas transferred from higher to lower local luminance mean and will be lower for areas transferred from low to high luminance mean.

Changing the polarity of text from black on white to white on black has little effect on normal reading. Legge *et al.*[23] have shown that some low-vision observers read as much as 50% faster with reversed contrast text. These effects, which have been known clinically for many years, are usually attributed to abnormal light scatter in eyes with cloudy media. Part of the effect may be explained by the change in contrast at the critical band of frequencies that occurs with change in polarity. The contrast at a 1-octave-wide band of spatial frequencies, extending upward from the fundamental frequency of the letters, has been shown to contain sufficient information for fast reading. The contrast of details at this band will change substantially with a change of polarity from black-on-white to white-on-black text, according to our definition. Thus a patient's reading performance that declines with a decrease in contrast at high contrast levels will improve with the reversal of text polarity irrespective of the nature of the patient's disability. This indeed appears to be true for the two cases reported by Rubin and Legge.[24] Since for many low-vision patients performance becomes dependent on contrast only at fairly low contrast levels, this effect is apparent only with a small portion of the population. Pelli[25] analyzed similarly the contrast of lines of text in the two polarities on a video display. His patterns, however, span different nonoverlapping luminance ranges and thus had different contrasts globally (defined by Michelson contrast) as well as locally. Only this global difference was considered in his case.

Fig. 3. Simulation of the perceived contrast image. This image was reconstructed by adding the local band-limited contrast images (right) instead of the original bandpass-filtered images (left).

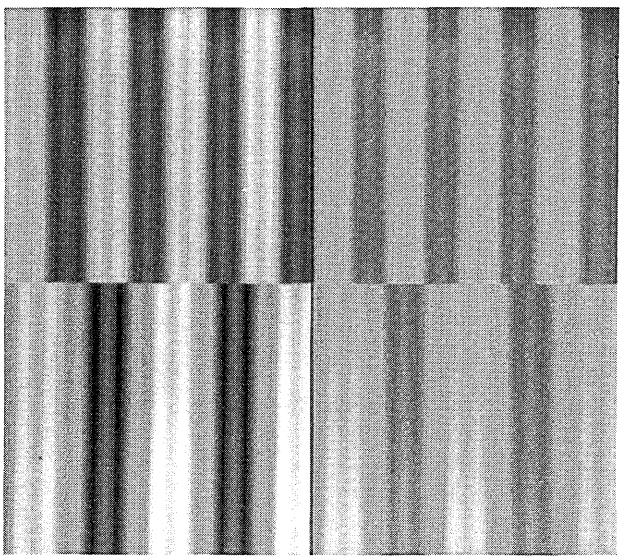

Fig. 4. Illustration of the different effects of linear rescaling on patterns of different spatial frequency composition. The compound gratings at the right were linearly scaled equally (2×), resulting in their respective gratings on the left. The amplitudes of the two sinusoidal components in each image pair are equal, and the high-frequency component is of the same period in all images. Note the relative increase in contrast of this component in the lower-left-hand image compared with the upper-left-hand image.

APPLICATION: SIMULATIONS OF THE APPEARANCE OF IMAGES

In this section two applications of the pyramidal image contrast structure described in Appendix A are illustrated. This type of processing enables us to implement the nonlin-

ear response of the visual system locally. This application was not possible until now.

The Fourier analysis of images in the context of image perception has frequently been interpreted to imply that the contrast sensitivity function measured at various spatial frequencies can be implemented as a modulation transfer function of the system in the Fourier domain for filtration of images.[3,26–28] In most cases, such applications were limited to increasing or decreasing the amplitudes at various spatial frequencies without explicit reference to the possible interactions among amplitudes at different frequencies. When applied to the simulation of appearances of images to observers with normal[3] or abnormal[28] vision, this linear process ignores the highly nonlinear characteristics of the visual system. Despite large differences in contrast sensitivity thresholds for different frequencies at different eccentricities, appearances of superthreshold images are constant or almost constant.[29,30] Hess *et al.*[11] included this nonlinear

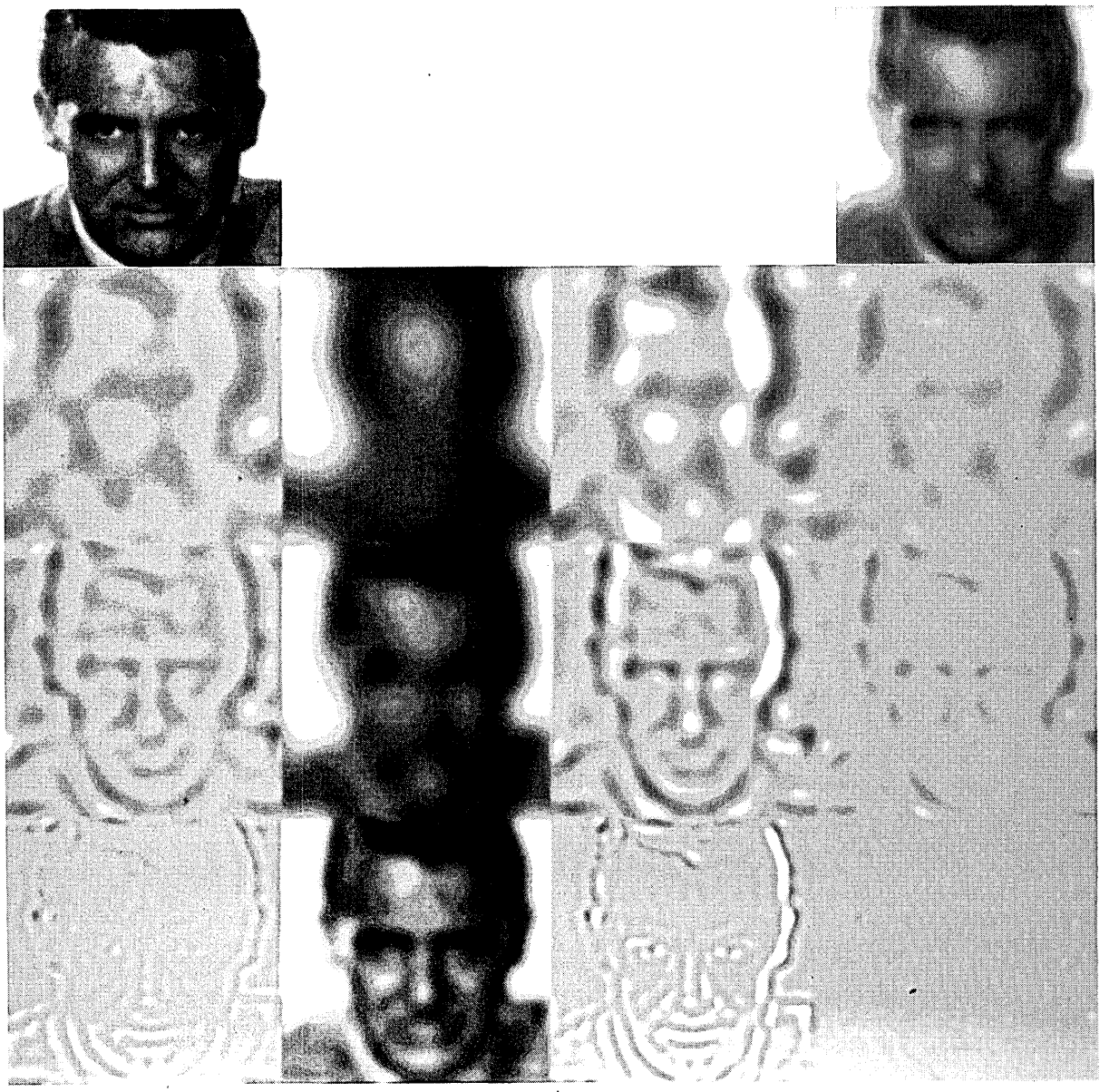

Fig. 5. Simulation of the appearance of a face image (spanning 4 deg of visual angle) to a low-vision patient whose contrast sensitivity function is illustrated in Fig. 6. Top left, the original image; top right, the simulated appearance of the same image to the patient. The three rows of four images represent processing at different spatial frequencies on the pyramid. The far-left-hand image in each row is the bandpass-filtered image obtained from the original image. The second column shows the corresponding low-pass-filtered version for the same scale, i.e., all the energy below the band represented in the first column, or the local luminance mean. The third column represents the contrast images. Contrast arrays are bipolar, and a DC level of 128 has been added arbitrarily to present those arrays as images. Images in the fourth column on the far-right-hand side represent the thresholded, bandpass-filtered images. For each image in the third column, each point was tested against the threshold value illustrated in Fig. 6 for the corresponding spatial frequency. If the contrast of the image at that point is above threshold, the corresponding point from the far-left image is reproduced in the far-right column. If the contrast at a certain point is below threshold, the corresponding point is set to zero (gray) in the far-right image. The simulated appearance image (top right) is generated by summing all the images in the far-right column. Actual processing included two more rows at 2 and 32 cycles per picture (not shown).

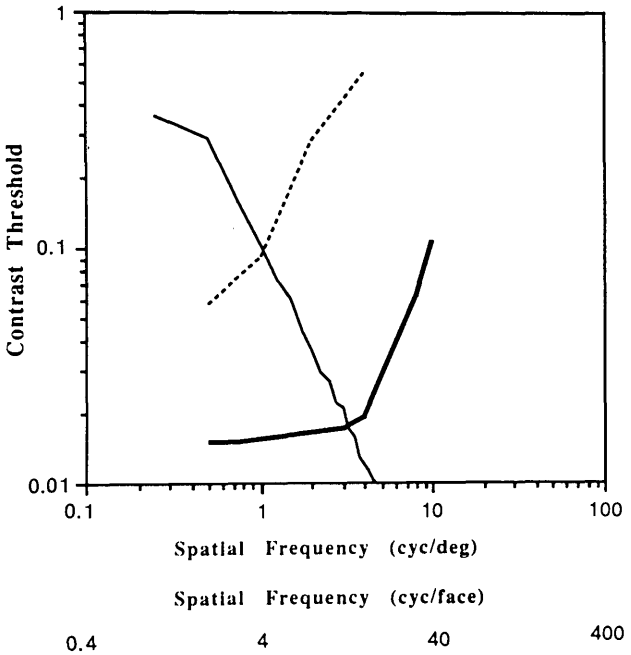

Fig. 6. Contrast detection thresholds (dotted curve) of a low-vision patient with central scotoma owing to age-related maculopathy used in the simulation of Fig. 5. Contrast detection thresholds of 15 normal observers are illustrated by the thick curve. The thin curve represents mean, radially averaged contrast spectra of five different faces.

characteristic in their simulation of vision in amplyopes by applying the threshold in the Fourier domain. Such global processing is insufficient, as it does not address the local variability of contrast across the image. To improve their simulation, they have also applied the same process to subimages.

The pyramidal image contrast structure described in Appendix A enables us to use nonlinear processing to simulate the appearance of images for normal- and low-vision observers point by point and for every spatial frequency in the image. An example of this process is illustrated in Fig. 5. The contrast sensitivity function of a patient with a central scotoma due to macular disease was measured, using 1-octave-bandwidth sinusoidal patches of grating in a two-dimensional Gaussian envelope.[31] The patient's contrast detection thresholds used in the processing of Fig. 5 are illustrated, together with the mean response of 15 normal observers, in Fig. 6. Thus the final image in Fig. 5, top right, represents the appearance of the original image to this patient. The original image was processed with the stipulation that the face span 4 deg of visual angle. On this scale, this patient's visual loss had little effect on information at 4 cycles per picture (top row of four images), a minimal effect on information at 8 cycles per picture (middle row), and a substantial effect on information at 16 cycles per picture (bottom row). Full processing included also the bands of 2 and 32 cycles per picture (both not shown). This simulation differs from previous such simulations[3,28] because suprathreshold contrast features retain their contrast and are not washed out by the processing as with other techniques. Thus the 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 the appearance of images seen through cataracts may be.[26]

The same pyramidal image contrast structure also enables us to simulate the appearance of images with a nonuniform retina. Using data on the contrast threshold at different spatial frequencies at different eccentricities on the retina,[29] we can simulate the appearance of images to the nonhomogeneous visual system by selecting a center of fixation representing the foveal position on the image and then comparing the local threshold at each spatial frequency and each eccen-

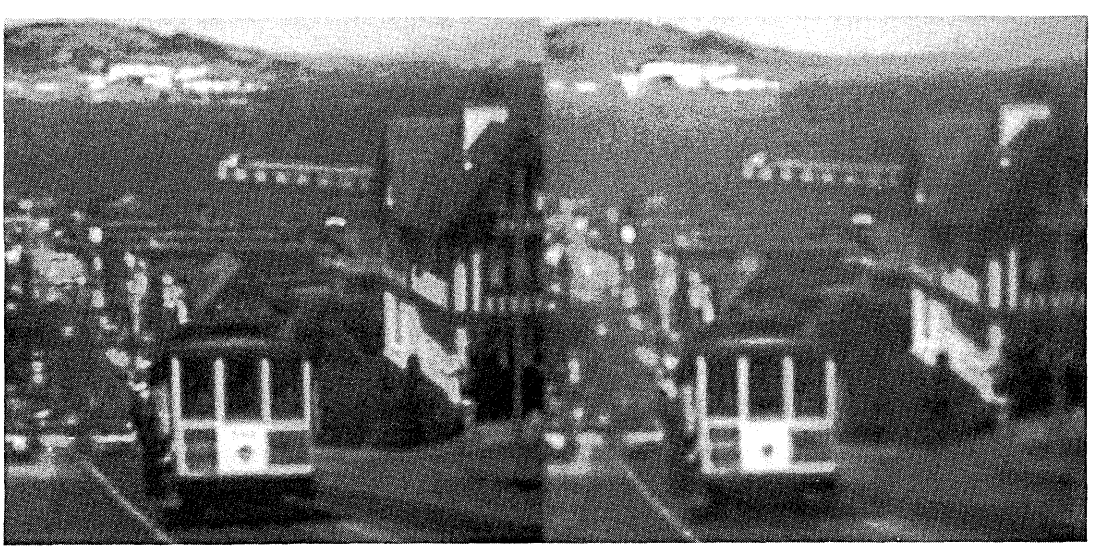

Fig. 7. Simulation of the appearance of an image to a normal observer including the nonuniform characteristic of the visual system. Simulation is carried out with the assumption of fixation at the center of the image. The technique applied is similar to the one used for Fig. 5, except that for every point in the contrast image the distance from the center of fixation in degrees of visual angle was calculated, and the contrast detection threshold corresponding to spatial frequency and retinal eccentricity was used in thresholding the images. The image at the left represents processing when the scene was considered to span 32 deg of visual angle. The image at the right represents the same image processed as if it spanned only 2 deg of visual angle. The most striking effect is the small variability across the visual field in both cases. Note that more heterogeneity is expressed over the image at the right (2 deg of visual angle).

tricity with the measured value. Images shown in Fig. 7 illustrate the appearance of the same image when it spans 2 and 32 deg of visual angle, respectively. The relatively small effect of the nonuniform retina in the appearance of both images is striking. The effect is much smaller than the effect previously simulated by Schwartz and colleagues, using cortical surface data[32,33] or an arbitrarily selected nonuniform function.[34] The same simulation may be expanded to represent the full visual field, as data for the lower spatial frequencies and the higher eccentricities were recently published by Pointer and Hess.[35]

DISCUSSION

The basic assumption of this study was that image contrasts should be expressed as the dimensionless ratio of the local amplitude and the local average luminance similarly to that expressed in the definition of Michelson contrast or Weber fraction. The use of such a ratio implies that the human sensitivity to amplitude of change in luminance varies with the adaptation level associated with the local average luminance.[36] This is known to be the case for threshold contrast sensitivity at all spatial frequencies at high luminance levels. For low frequencies (<4 cycles/deg), the same relation is true for a large portion of the photopic range.[37] 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. Such reduction may actually be necessary to avoid phase inversions in extreme cases. Low-pass-filtered versions of an image may, in extreme but possible images, contain negative values, even if the image is defined as positive only. Such negative values indicate the existence of negative values of the filters. In the visual system such values exist and are referred to as areas of inhibition in the receptive field or the filter's point-spread function. These negative values may result in an inversion of the contrast as defined here, an undesirable result. However, if partial adaptation is applied, it can be adjusted easily to reduce the magnitude of such discontinuities.

The degree or level of local luminance adaptation in suprathreshold contrast sensitivity has, to our knowledge, not been determined. Experiments using a dichoptic presentation found that contrast matching at high-contrast levels indeed approximated contrast as defined by the ratio of amplitude to local luminance mean.[38] Although the methodology used cannot be applied directly to normal viewing of an image, the ability to set the contrast for apparent match under such diverse conditions suggests that similar results may be obtained with monocular viewing of multiple targets over a variable local luminance mean. We are currently attempting to measure directly the level of local luminance adaptation within one image.

Enhancement of images and deblurring in the visual system have been discussed by various authors. Mechanisms such as lateral inhibition or the transfer function calculated from the contrast sensitivity function were used to explain these enhancement effects. Active enhancement using adaptive gain control in different spatial frequency channels was proposed by Georgeson and Sullivan.[38] Enhancement or sharpening of the image reconstructed from contrast rather than amplitude components is proposed here as a local mechanism for enhancement of complex images but cannot explain the experimental results obtained with single sinusoidal targets. Thus it could be postulated that such local enhancement occurs in addition to the reported global sharpening. If such local enhancement does occur, it should be measurable.

The same enhancement that occurs in the visual system may be useful in image-processing algorithms. Indeed, the enhancement capabilities of similar pyramids of contrast-related images have been used in image-processing applications.[39,40] In both, the pyramids were of ratios of low-pass-filtered versions of the image at different scales, and in both cases the visual contrast sensitivity was cited as the motivation. Toet *et al.*[39] used a ratio of 2-octave-spaced low-pass images to merge visual-optical and thermal images. Their contrast ratio was defined as

$$r_i(x, y) = \frac{l_i(x, y)}{l_{i-1}(x, y)} = c_i(x, y) + 1. \tag{13}$$

They argued that the contrast-related bandpass-filtered image version of the optical image is more appropriate to use since it more closely represents visually important features.

The main difference between their definition and the one used here is that in their contrast there is no sign change to distinguish between objects that are darker or brighter than the background. The importance of this sign change in the visual system has been reported by Shapley and Enroth-Cugell.[41]

Schenker *et al.*[40] used a similar ratio of two low-pass filtered images and compressed the output by a logarithmic transformation in an algorithm used to detect image edge structure. In our notation it can be written as

$$I_{0_i}(x, y) = \ln \frac{l_i(x, y)}{l_{i-1}(x, y)}. \tag{14}$$

Logarithmic transformation restores the sign change and also results in relative enhancement of negative contrast or in increased sensitivity to decrements versus increments as found commonly in psychophysical experiments.[42] Hilsenrath and Zeevi[43] implemented a similar process of adaptation in one scale only for designing an adaptive, locally gain-controlled detector. Such adaptation permits imaging over wider dynamic ranges than is possible with standard cameras.

Contrast measured by filtering as suggested here defines only incremental or decremental changes from 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. 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.[44] Incorporation of these additions in a one-dimensional case, using oriented filters, is straightforward.[45] Complete two-dimensional application is difficult owing to the lack of definition of Hilbert's transform for the two-dimensional case.[46]

APPENDIX A: PYRAMIDAL STRUCTURE USED IN ANALYSIS AND SIMULATIONS

A pyramidal image transform was calculated in the frequency domain.[47] For digital processing of images, it is convenient to select center frequencies (in cycles per picture) that are a power of 2 for each segment. Thus the image in the frequency domain may be represented as

$$F(u, v) = F(r, \theta) = L_0(r, \theta) + \sum_{i=1}^{n-1} A_i(r, \theta) + H_n(r, \theta), \quad \text{(A1)}$$

where L_0 and H_n represent the low and high residuals, respectively. They contain the energy in the low and high frequencies after the various bandpass layers, A_i, have been subtracted from the image. The low residual is essential in our application and therefore is maintained. The high residual has little information, and in most applications it may be discarded without any perceptual change in the image.[48]

Although the use of a Gaussian filter is attractive because of the mathematical convenience in transformation from the frequency to the spatial domain, this filter has several shortcomings.[19,48] To obtain an approximation to the shape of a Gaussian and at the same time to satisfy the requirement of symmetrical shape on a log frequency axis, together with the requirement that the image must be able to be reconstructed from the various segments by simple addition,[48] we have used cosine log filters (Fig. 8). A cosine log filter of (1-octave) bandwidth centered at frequency 2^i cycles/picture is expressed as

$$G_i(r) = \tfrac{1}{2}[1 + \cos(\pi \log_2 r - \pi i)]. \quad \text{(A2)}$$

The small difference between these functions and the commonly used Gabor filters or derivatives of Gaussians is of little consequence for the concept described here and its potential applications.

Thus A_i is obtained by multiplying the Fourier transform of the image with a torus-shaped dome filter described in Eq. (A2). The filtered image is transformed back to the space domain, where it can be represented as

$$f(x, y) = l_0(x, y) + \sum_{i=1}^{n-1} a_i(x, y) + h_n(x, y). \quad \text{(A3)}$$

The residual l_0 is calculated simply to maintain the ease of reconstruction with simple addition, but h_n is not used in our model. For every $a_i(x, y)$, the corresponding $l_i(x, y)$ is

$$l_i(x, y) = l_0(x, y) + \sum_{j=1}^{i-1} a_j(x, y), \quad \text{(A4)}$$

and $c_i(x, y)$ is calculated as in Eq. (9):

$$c_i(x, y) = \frac{a_i(x, y)}{l_i(x, y)}. \quad \text{(A5)}$$

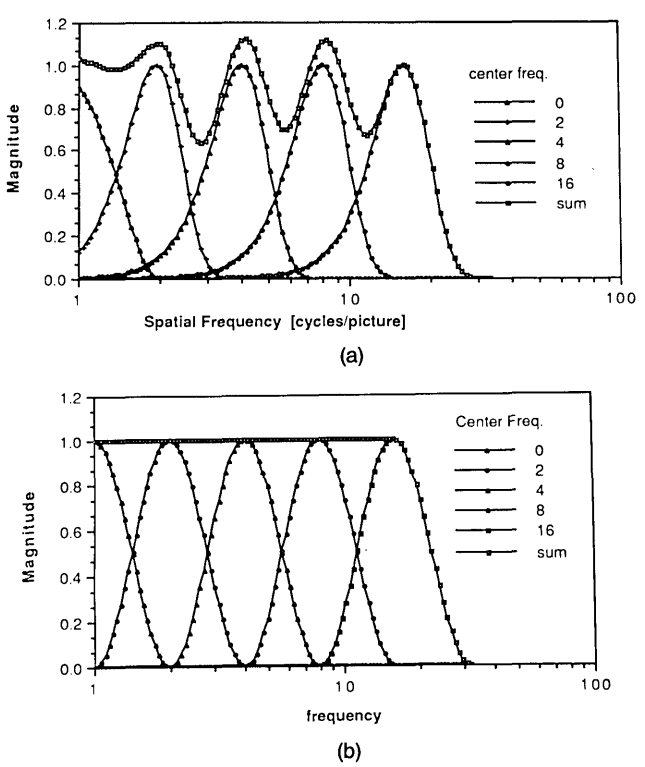

Fig. 8. Comparison of Gaussian (Gabor filters) with the cosine log filters used here. (a) Filter bank of 1-octave-wide Gaussian filters and the sum of all filters. (b) Filter bank of 1-octave-wide cosine log filters. Here the summation of all the filters adds to the unity. Note also the symmetry of the cosine log filters on a logarithmic scale.

ACKNOWLEDGMENTS

This study is supported in part by grant EY05957 from the National Institutes of Health and by a grant from the James H. and Alice Teubert Charitable Trust. I thank Steve Burns for valuable discussions and Robert Goldstein for important programming help.

The author is also with the Department of Ophthalmology, Harvard Medical School.

REFERENCES AND NOTES

1. A. A. Michelson, *Studies in Optics* (U. Chicago Press, Chicago, Ill., 1927).
2. G. Westheimer, "The oscilloscopic view: retinal illuminance and contrast of point and line targets," Vision Res. **25**, 1097–1103 (1985).
3. A. P. Ginsburg, "Visual information processing based on spatial filters constrained by biological data," Ph.D. dissertation, Aerospace Medical Research Laboratory Rep. AMRL-TR-78-129 (University of Cambridge, Cambridge, 1978).
4. C. Owsley, R. Sekuler, and C. Boldt, "Aging and low-contrast vision: face perception," Invest. Ophthalmol. Vis. Sci. **21**, 362–365 (1981).
5. R. Sekuler, C. Owsley, and L. Hutman, "Assessing spatial vision of older people," Am. J. Optom. Physiol. Opt. **59**, 961–968 (1982).
6. M. Hubner, I. Rentschler, and W. Encke, "Hidden-face recognition: comparing foveal and extrafoveal performance," Hum. Neurobiol. **4**, 1–7 (1985).
7. T. R. Riedl and G. Sperling, "Spatial frequency bands in complex visual stimuli: American Sign Language," J. Opt. Soc. Am. A **5**, 606–616 (1988).
8. M. Pavel, G. Sperling, T. Riedl, and A. Vanderbeek, "Limits of visual communication: the effect of signal-to-noise ratio on the intelligibility of American Sign Language," J. Opt. Soc. Am. A **4**, 2355–2365 (1987).

9. G. S. Rubin and K. Siegel, "Recognition of low-pass filtered faces and letters," Invest. Ophthalmol. Vis. Sci. Suppl. **25,** 71 (1984).
10. D. S. Loshin and T. A. Banton, "Local contrast requirements for Facial recognition in patients with central field defects," Invest. Ophthalmol. Vis. Sci. Suppl. **29,** 43 (1988).
11. R. F. Hess, A. Bradley, and L. Piotrowski, "Contrast-coding in amblyopia. I. Differences in the neural basis of human amblyopia," Proc. R. Soc. London Ser. B **217,** 309–330 (1983).
12. J. D. Briers and A. F. Fercher, "Retinal blood-flow visualization by means of laser speckle photography," Invest. Ophthalmol. Vis. Sci. **22,** 255–259 (1982).
13. A. B. Watson, H. B. Barlow, and J. G. Robson, "What does the eye see best?" Nature (London) **302,** 419–422 (1983).
14. D. R. Badcock, "Spatial phase or luminance profile discrimination?" Vision Res. **24,** 613–623 (1984).
15. R. F. Hess and J. S. Pointer, "Evidence for spatially local computations underlying discrimination of periodic patterns in fovea and periphery," Vision Res. **27,** 1343–1360 (1987).
16. E. Peli and R. B. Goldstein, "Contrast in images," in *Visual Communication and Image Processing '88,* T. R. Hsing, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **1001,** 521–528 (1988).
17. R. L. De Valois, D. G. Albrecht, and L. G. Thorell, "Spatial frequency selectivity of cells in macaque visual cortex," Vision Res. **22,** 545–559 (1982).
18. A. B. Watson, "Efficiency of a model human image code," J. Opt. Soc. Am. A **4,** 2401–2417 (1987).
19. D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," J. Opt. Soc. Am. A **4,** 2379–2394 (1987).
20. J. P. Thomas, "Independent processing of suprathreshold spatial gratings as a function of their separation in spatial frequency," J. Opt. Soc. Am. A **6,** 1102–1111 (1989).
21. A. Fiorentini, L. Maffei, and G. Sandini, "The role of high spatial frequencies in face perception," Perception **12,** 195–201 (1983).
22. W. K. Pratt, *Digital Image Processing,* (Wiley, New York, 1978), pp. 307–344.
23. G. E. Legge, G. S. Rubin, D. G. Pelli, and M. M. Schleske, "Psychophysics of reading. II. Low vision," Vision Res. **25,** 253–266 (1985).
24. G. S. Rubin and G. E. Legge, "Psychophysics of reading. VI. The role of contrast in low vision," Vision Res. **29,** 79–91 (1989).
25. D. G. Pelli, "Reading and contrast adaptation," in *Digest of Topical Meeting on Applied Vision* (Optical Society of America, Washington, D.C., 1989), pp. 102–103.
26. E. Peli and T. Peli, "Image enhancement for the visually impaired," Opt. Eng. **23,** 47–51 (1984).
27. T. B. Lawton, "Improved word recognition for observers with age-related maculopathies using compensation filters," Clin. Vision Sci. **3,** 125–135 (1988).
28. B. L. Lundh, G. Derefeldt, S. Nyberg, and G. Lennerstrand, "Picture simulation of contrast sensitivity in organic and functional amblyopia," Acta Ophthalmol. **59,** 774–783 (1981).
29. M. W. Cannon, Jr., "Perceived contrast in the fovea and periphery," J. Opt. Soc. Am. A **2,** 1760–1768 (1985).
30. M. W. Cannon, Jr., and S. C. Fullenkamp, "Perceived contrast and stimulus size: experiment and simulation," Vision Res. **28,** 695–709 (1988).
31. E. Peli, R. Goldstein, G. Young, and L. Arend, "Contrast sensitivity functions for analysis and simulation of visual perception," in *Digest of the Topical Meeting on Noninvasive Assessment of the Visual System* (Optical Society of America, Washington, D.C., 1990).
32. E. L. Schwartz, B. Merker, E. Wolfson, and A. Shaw, "Applications of computer graphics and image processing to 2D and 3D modeling of the functional architecture of visual cortex," in *Digest of Meeting on Computer Graphics and Applications* (Institute of Electrical and Electronics Engineers, New York, 1988), pp. 12–23.
33. E. L. Schwartz and B. Merker, "Computer-aided neuroanatomy differential geometry of cortical surfaces and on optimal flattening algorithm," in *Digest of Meeting on Computer Graphics and Applications* (Institute of Electrical and Electronics Engineers, New York, 1986), pp. 36–49.
34. Y. Y. Zeevi, N. Peterfreund, and E. Shlomot, "Pyramidal image representation in nonuniform systems," in *Visual Communications and Image Processing, '88,* T. R. Hsing, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **1001,** 563–571 (1988).
35. J. S. Pointer and R. F. Hess, "The contrast sensitivity gradient across the human visual field with emphasis on the low spatial frequency range," Vision Res. **29,** 1133–1151 (1989).
36. J. G. Robson, "Linear and non-linear operations in the visual system," Invest. Ophthalmol. Vis. Sci. **29,** 117 (1988).
37. D. H. Kelly, "Visual contrast sensitivity," Opt. Acta **24,** 107–129 (1977).
38. M. A. Georgeson and G. D. Sullivan, "Contrast constancy: deblurring in human vision by spatial frequency channels," J. Physiol. (London) **252,** 627–656 (1975).
39. A. Toet, L. G. van Ruyven, and J. M. Valeton, "Merging thermal and visual images by contrast pyramid," Opt. Eng. **28,** 789–792 (1989).
40. P. S. Schenker, D. R. Urangst, T. F. Knaak, D. T. Huntley, and W. R. Patterson III, "Pyramidal normalization filter: visual model with application to image understanding," in *Real Time Signal Processing V,* J. Trimble, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **341,** 99–108 (1982).
41. R. Shapley and C. Enroth-Cugell, "Visual adaptation and retinal gain controls," in *Progress in Retinal Research,* N. N. Osborne, ed. (Pergamon, Oxford, 1984), Vol. 3, pp. 263–343.
42. R. W. Bowen, J. Pokorny, and V. C. Smith, "Sawtooth contrast sensitivity: decrements have the edge," Vision Res. **298,** 1501–1509 (1989).
43. O. A. Hilsenrath and Y. Y. Zeevi, "Adaptive two-dimensional neighborhood sensitivity control by a one-dimensional process," in *Visual Communication and Image Processing '88,* T. R. Hsing, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **1001,** 717–723 (1988).
44. C. F. Stromeyer III and S. Klein, "Evidence against narrowband spatial frequency channels in human vision: the detectability of frequency modulated gratings," Vision Res. **15,** 899–910 (1975).
45. M. C. Morrone and D. C. Burr, "Feature detection in human vision: a phase-dependent energy model," Proc. R. Soc. London Ser. B **235,** 221–245 (1988).
46. E. Peli, "Hilbert transform pairs mechanisms," Invest. Ophthalmol. Vis. Sci. **30,** 110 (1989).
47. The pyramidal image transform used here is conceptually identical to any of the commonly used pyramids of bandpass-filtered images. Since for our application images of equal size are used at all bands, we avoided the common approach of subsampling the images recursively, filtering, and then unsampling the reduced size images, instead; all filtering was done in the frequency domain. Thus the content of our final images in the pyramid of image scales was identical to the images that would be calculated by upsampling images obtained on a pyramid of image resolution.
48. A. B. Watson, "The cortex transform, rapid computation of simulated neuro images," Comput. Vision Graphics Image Process. **39,** 311–327 (1987).