A New Representation of the Contrast Sensitivity Function for Human Vision

Bert W. Rust
Mathematical and Computational Sciences Division
National Institute of Standards and Technology
Gaithersburg, MD 20899

Holly E. Rushmeier
IBM T.J. Watson Research Center
Hawthorne, NY 10532

Abstract *This paper presents a new mathematical representation of the contrast sensitivity function for the human visual system. It is based on optimal smoothing spline fits to vertical and oblique contrast sensitivity measurements. Unlike most of the approximations in current use, which assume isotropy, this new representation incorporates the dependence on orientation revealed by the measurements. One of the uses of the representation is demonstrated by applying it to an example luminance field.*

Keywords: human vision, contrast sensitivity, luminance fields

1 Introduction

The psychophysical response of the human visual system to spatial variations in a perceived luminance field is generally believed to be characterized by a *contrast sensitivity function* (CSF) which measures the threshold ability to detect such variations as a function of the spatial frequencies in the horizontal and vertical directions. The CSF is measured by observing sinusoidal gratings with varying contrasts and recording the contrast at which the grating becomes detectable [1]. The contrast of a grating is defined by

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

where $L_{\max}$ and $L_{\min}$ are the maximum and minimum luminances of the grating. For a grating with a given spatial frequency, the reciprocal of the threshold contrast needed for detection is the contrast sensitivity at that frequency.

Figure 1 gives plots of vertical and oblique CSFs measured by Campbell et al. [2]. The usual convention is to plot the CSF only on a positive x-axis since the values at negative frequencies are identical to those at the corresponding positive frequencies. Figure 1 does not follow this convention for reasons that will be explained in the following. In both plots, the sensitivity has a maximum value between the frequencies 5 and 10 cycles/degree, and drops off precipitously at lower frequencies and exponentially at higher frequencies, and the whole structure is reflected about the origin for the negative frequencies. Of course a CSF is actually a function of two variables, spatial frequency in the x-direction and in the y-direction. Assuming isotropy allows this bivariate function to be constructed from a single univariate function, like one of the two shown in the figure, by rotating the whole graph around the vertical axis, which would become the z-axis in the resulting 3-dimensional structure. If both positive and negative frequencies are included in the univariate plot, then the bivariate function is completed by a rotation of 180°, but if only the positive frequencies are

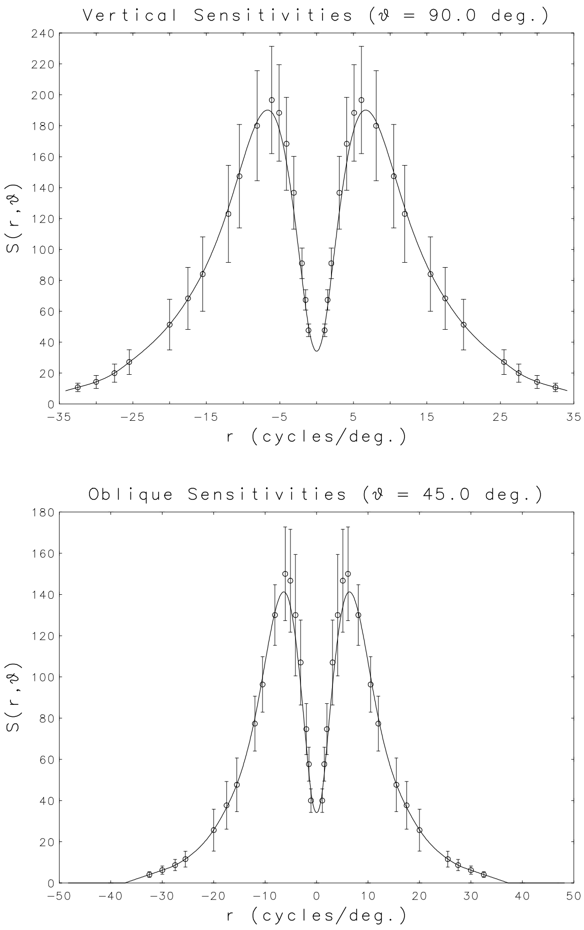

Figure 1: The measured vertical and oblique sensitivities

included then a rotation of 360° is required. In either case the resulting bivariate function has circular symmetry about the z-axis. Viewed from above, the level surfaces would all be circles, with the highest contour corresponding to the maximum in the univariate plot.

2 The Measured Data

The isotropy assumption would be justified only if the contrast sensitivity of the visual system were independent of the orientation of the sinusoidal gratings used to make the measurements. Campbell, Kulikowski and Levinson [2] measured the CSFs for three observers using two different orientations for the gratings – vertical and oblique (inclined at an angle of 45° from the vertical). Their results are summarized in Figure 1 where each of the discrete data points is the average of three separate measurements and the attached error bar gives the ± 1 standard deviation for those measurements. Although we have plotted the measurements at both positive and negative spatial frequencies, the points at negative frequencies do not represent independent measurements. They are simply the reflections about the vertical axis of the measurements plotted at positive frequencies. This procedure is valid because the CSF is the Fourier transform of a real function of the spatial coordinates. It is clear by inspection that for any given spatial frequency, the oblique sensitivities are significantly smaller than the vertical ones. Thus the isotropy assumption is untenable.

A great deal of work has been devoted to finding continuous function representations for univariate CSFs, e.g., [3, 4, 5]. Some investigators have tried to derive functions based on physical models for the eye's response, but most have simply tried to find formulas which could be fit to measured data. Our approach is to let the measured data specify the function, but to use a cubic smoothing spline fit rather than asumming some functional form for the relation. A cubic spline is a piecewise polynomial function composed of cubic polynomial segements joined smoothly at the nodes. The nodes are taken at the same abcissas as the measurements, and the polynomials in neighboring segments are chosen to guarantee continuity at the nodes of not only the composite function but also its first and second derivatives. If the spline is forced to pass through all the data points, it is called an interpolating spline, but it is possible to relax that requirement by introducing a nonnegative parameter, called the smoothing constant, which can be used to produce smoothed representations of the data, with the amount of smoothing depending on the choice of the constant. Choosing zero produces an interpolating spline, and choosing infinity (or just a very large value) produces a least squares straight line fit to the data.

The curves shown in Figure 1 are *optimal* *smoothing spline* fits obtained from the algorithm CUBGCV [6] which chooses the smoothing constant to minimize the generalized cross validation statistic [7]. For a given value of the smoothing constant, that statistic is computed as the sum of squared errors obtained by successively dropping one data point from the set, fitting the corresponding smoothing spline to the remaining points and calculating the error in predicting the point that was dropped. The value chosen for the smoothing constant is the one which minimizes the sum of squares of those prediction errors. Each prediction error was also weighted inversely with the uncertainty (standard deviation) in the dropped point so that points with smaller uncertainties were given more influence in the choice of the smoothing constant. The advantage of this procedure is that it allows the data and their uncertainties to make an objective choice of the smoothing constant rather than relying on the subjective judgement of the user.

3 The Baseline and Amplitude Functions

Since the two univariate CSFs given in Figure 1 differ significantly, a realistic bivariate CSF cannot be formed by simply rotating one, or the other, or even the average of the two about the vertical axis. Ginsburg [1] and Gervais et al. [8] have presented bivariate CFSs which take these differences into account, but they did not give the recipe for generating their representations. Following their lead, we have assumed that the vertical sensitivities represent the CSF for the orientation angles $\vartheta = 0°$, $90°$, $180°$, $270°$ and the oblique sensitivites represent the angles $\vartheta = 45°$, $135°$, $225°$, $315°$. There were measurements for 17 spatial frequencies r in both the vertical and oblique cases, so we had meaured values of the bivariate CSF at a total of $17 \times 4 \times 2 = 136$ discrete points (r, ϑ) in the 2-dimension spatial frequency space, but the coverage was far from uniform. Therefore attempts to fit 2-dimensional spline surfaces to these data produced large spurious folds and ripples at intermediate r and ϑ values. To avoid this problem, we sought a representaion of the form

$$S(r, \vartheta) = b(r) + A(r) cos(4\vartheta) \; , \qquad (2)$$

where $S(r, \vartheta)$ is the bivariate CSF, and $b(r)$ and $A(r)$ are a base-line function and an amplitude function which are both independent of the orientation. If r is fixed and ϑ increases from $0°$ to $360°$, then $S(r, \vartheta)$ passes through the maximum value $[b(r) + A(r)]$ at angles $\vartheta = 0°$, $90°$, $180°$, $270°$ and the minimum value $[b(r) - A(r)]$ at angles $\vartheta = 45°$, $135°$, $225°$, $315°$.

The vertical and oblique measurements were made at the same 17 spatial frequencies r_i, so it follows that for $i = 1, 2, \ldots, 17$,

$$b(r_i) = \frac{1}{2} \left[S_i(90°) + S_i(45°) \right] \; , \qquad (3)$$

and

$$A(r_i) = \frac{1}{2} \left[S_i(90°) - S_i(45°) \right] \; , \qquad (4)$$

where the $S_i(90°)$ and $S_i(45°)$ are the measured sensitivities plotted in Figure 1. The same values are also obtained by computing separate sets of $b(r_i)$ and $A(r_i)$ values for each of the 3 test subjects and averaging the 3 values obtained for each measurement. These derived sensitivity measurements are plotted as discrete points in Figure 2 with error bars which are the ± 1 standard deviations for the separate averages over the three test subjects. Note that the derived uncertainties are the same for the two curves, so the relative uncertainties are larger for the $A(r_i)$ than for the $b(r_i)$.

The curves in Figure 2 are weighted, optimal smoothing spline fits like the ones described in the preceding section. Each point was weighted inversely with the uncertainty represented by the corresponding error bar. In order to avoid an ambiguity at the origin (i.e., to assure that $S(0, 0)$ is uniquely defined) it was also necessary to require that the $A(r)$ curve be zero at $r = 0$. This was accomplished

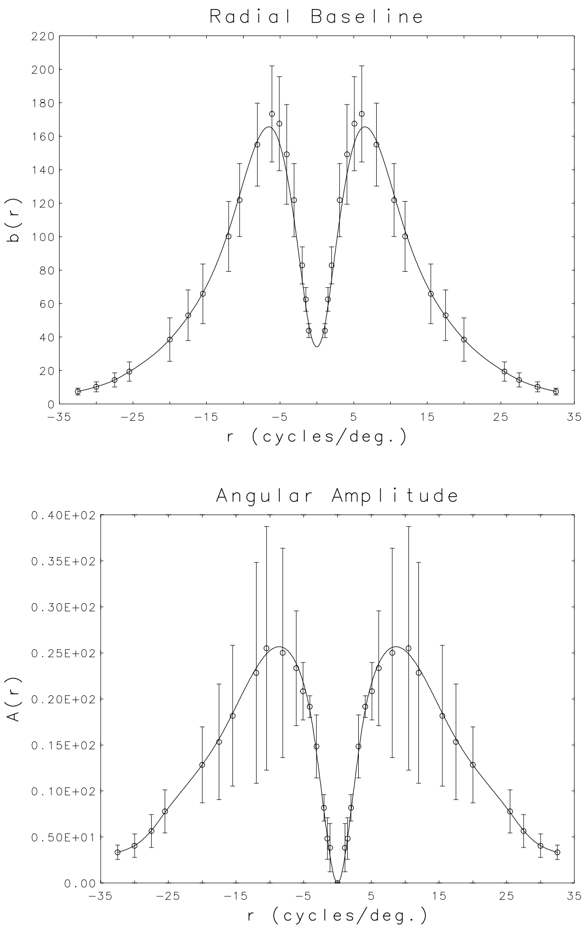

Figure 2: The radial baseline and angular amplitude functions

by adding a zero measurement, with an extremely small uncertainty, at $r = 0$. The lower plot in the figure shows that this procedure gives a reasonable extrapolation of the measured data and Eq. (4) shows that it amounts to assuming that the vertical and oblique sensitivities are equal at $r = 0$.

The question of the value of the $b(r)$ function at $r = 0$ is a more difficult matter. We have not been able to find any reference that specifies the contrast sensitivity at spatial frequency zero, but we have found formulas fitted to measured data that give negative values for the CSF at $r = 0$. Presumably their originators did not notice this fact because they followed the universal practice of truncating their plots well short of frequency zero. Some workers have cleverly skirted this sticky issue by plotting their CSFs on logarithmic scales. Of course, it is quite difficult to measure the CSF for very low frequencies and presumably impossible to measure it at $r = 0$, since a zero-frequency grating would have an infinite period. But when the CSF is used as a filter for observed luminance fields, the zero frequency value multiplies the amplitude of the zero frequency cycle in the observed data. Since the latter quantity is just the average luminance of the observed field, and since the visual system is surely sensitive to this quantity, it does not seem reasonable to simply leave the CSF undefined at $r = 0$. Negative values are certainly unreasonable, and it can be seen by inverting Eq. (1) that a zero value would be obtained only if either $L_{\mathrm{max}} + L_{\mathrm{min}} = 0$ or $L_{\mathrm{max}} - L_{\mathrm{min}} = \infty$. The first case corresponds to total darkness and the second could only be approximated at the center of a star or a nuclear bomb blast.

Lacking any clear prescription for $b(0)$, we decided to let the measured data specify the value. This was accomplished by performing the spline fit on the negative as well as the positive frequencies (see the upper plot in Figure 2). The value at $r = 0$ is thus the minimum value of the cubic polynomial segment between the measurement at the lowest positive frequency and its mirror image at the corresponding negative frequency. Of course there is no guarantee that a cubic polynomial is a good approximation of the desired function in this region, but it is at least well-behaved. Not only is it continuous, but also its first and second derivatives are continuous, and the first derivative is zero at $r = 0$. The cubic spline also has the important advantage that local errors do not propagate globally, so errors in the polynomial segment through $r = 0$ to not induce errors in the neighboring segments.

4 The Bivariate Contrast Sensitivity Function

It would be tedious to write formulas for the 66 cubic polynomials forming the two splines in

Figure 2. We have written a computer program to evaluate both functions, not only between frequencies -32.5 and $+32.5$ cycles/degree, but also outside that range by using linear extrapolation until the extrapolation lines intersect the frequency axis and beyond that setting $b(r) = 0$ and $A(r) = 0$. Thus, using Equation (2), it is easy to compute $S(r, \vartheta)$ for any (r, ϑ). Figure 3 was obtained by such a calculation on a 1024×1024 equally spaced grid with both the $x-$ and $y-$frequencies ranging between -34.13333 and $+34.06667$ cycles/degree.

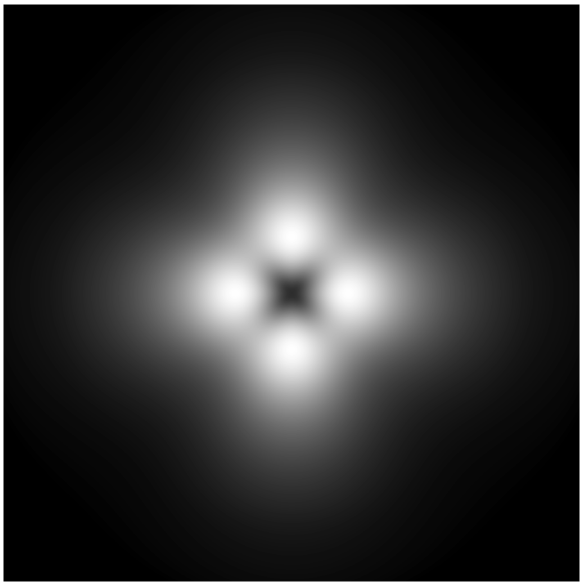

Figure 3: The contrast sensitivity function

With 1023 frequency intervals in each direction, the frequency spacing for the grid is 0.06666667 cycles/degree, and the origin $(r, \vartheta) = (0, 0)$ is one of the grid points. The value of $S(r, \vartheta)$ is plotted on a gray-scale ranging from the value 0 in the corners to a maximum of about 191 at the maxima at angles $\vartheta = 0°$, $90°$, $180°$, $270°$. Any apparent discontinuities in the brightness levels are not real, but rather are artifacts of the plotting and/or printing process. Note the pronounced differences between the vertical and oblique sensitivities.

5 Using the Contrast Sensitivity Function

To illustrate the use of the bivariate CSF in Figure 3, consider the image in Figure 4 which was generated on a 1024×1024 equally spaced grid with discrete luminance levels of 1, 101, 201, 301, and 401. For the purpose of the example, this image should not be considered to be what is actually seen, but rather to be the underlying luminance field to be acted on by the human visual system to get what is actually seen (this distinction will be explained more fully in the following). The bivariate CSF

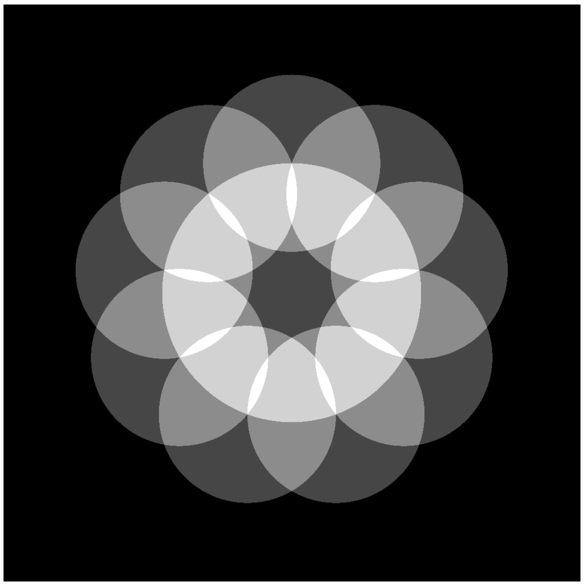

Figure 4: A luminance field test pattern

acts as a filter multiplying the discrete Fourier transform of the image. If the image intensity of the test field at grid point (m, n) is denoted by $t_{m,n}$, $m, n = 0, 1, \ldots, 1023$, then its discrete Fourier transform is the 1024×1024 set of complex coefficients

$$\mathcal{T}_{k,l} = \sum_{m=0}^{1023} \sum_{n=0}^{1023} t_{m,n} \exp\left\{-\frac{j2\pi}{1024}\left[mk + nl\right]\right\} ,$$
$$k,\ l = -512, \ldots, -1, 0, 1, \ldots, 511 , \tag{5}$$

where $j = \sqrt{-1}$. These coefficients correspond to the spatial frequency array

$$\begin{aligned} (\xi_k, \eta_l) &= \left(\tfrac{k}{1024\Delta}, \tfrac{l}{1024\Delta}\right) \ , \\ k,\ l &= -512, \ldots, -1, 0, 1, \ldots, 511 \ , \end{aligned} \tag{6}$$

where ξ_k and η_l are the horizontal and vertical spatial frequencies and Δ is the observed grid spacing. Choosing the value $\Delta = 0.01464844°$ made the above frequency grid exactly the same as the one used in Figure 3. This reverses the procedure that would normally be followed in analyzing a real image, where Δ would be determined by the size of the image, its distance from the observer and the number of grid points along each edge of the image and the CSF would be computed on the same grid.

Having determined the Fourier transform on the frequency mesh (6), Equation (2) is used to compute the CSF on the same mesh, i.e.,

$$\begin{aligned} S_{k,l} &= S\left(r_{k,l}, \vartheta_{k,l}\right) \\ &= b\left(r_{k,l}\right) + A\left(r_{k,l}\right)\cos\left(4\vartheta_{k,l}\right) \ , \end{aligned} \tag{7}$$

where

$$\begin{aligned} r_{k,l} &= \sqrt{\xi_k^2 + \eta_l^2} \ , \\ \vartheta_{k,l} &= \tan^{-1}\left(\frac{\eta_l}{\xi_k}\right) \ . \end{aligned} \tag{8}$$

The effect of the filter is then gotten by forming the products

$$\begin{aligned} \hat{\mathcal{T}}_{k,l} &= S_{k,l}\mathcal{T}_{k,l} \ , \\ k,\ l &= -512, \ldots, -1, 0, 1, \ldots, 511 \ , \end{aligned} \tag{9}$$

where $\hat{\mathcal{T}}_{k,l}$ denotes the filtered transform. Since the Fourier transform is complex, the CSF would multiply both the real and imaginary parts, but the effect of the filtering can be summarized in one image by considering the action of the filter on the amplitude spectrum which is defined by

$$\begin{aligned} \mathcal{A}_{k,l} &= \tfrac{1}{1024\Delta}\left|\mathcal{T}_{k,l}\right| = \tfrac{1}{1024\Delta}\sqrt{\mathcal{T}_{k,l}\mathcal{T}_{k,l}^*} \ , \\ k,\ l &= -512, \ldots, -1, 0, 1, \ldots, 511 \ , \end{aligned} \tag{10}$$

where $\mathcal{T}_{k,l}^*$ is the complex conjugate of $\mathcal{T}_{k,l}$. This amplitude spectrum is shown in Figure

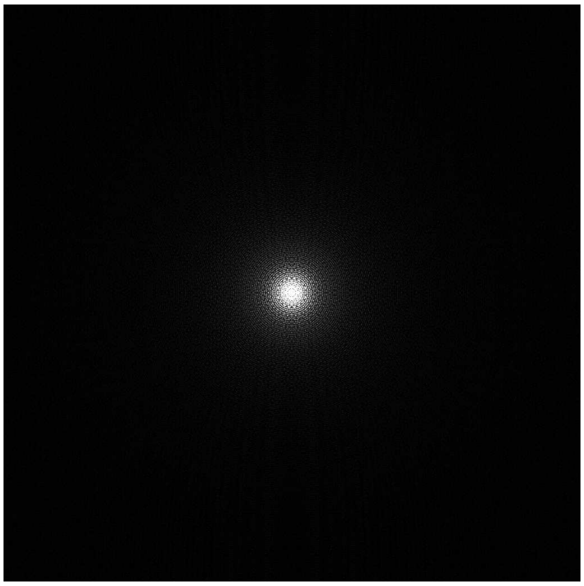

Figure 5: The Fourier amplitude spectrum of the test pattern

5. The minimum and maximum frequencies in both the $x-$ and $y-$directions are the same as those in Figure 3, namely -34.13333 and $+34.06667$ cycles/degree. The intensities are almost negligible (less than 0.1) everywhere except at the very low frequencies. The highest intensity is 83.450, and since it occurs at the origin, it is the average value of the luminance in Figure 4. The filtered amplitude spectrum

$$\hat{\mathcal{A}}_{k,l} = S_{k,l}\mathcal{A}_{k,l} = \frac{1}{1024\Delta}\sqrt{\hat{\mathcal{T}}_{k,l}\hat{\mathcal{T}}_{k,l}^*} \ , \tag{11}$$

is shown in Figure 6 which has the same minimum and maximum frequencies as Figures 3 and 5. Comparing Figures 5 and 6 shows that the intensities at the intermediate frequencies are ehanced relative those at the lower frequencies which dominate Figure 5. Note also the cross-shaped pattern produced by the anisotropy in the CSF.

To see the filtered image, it is necessary to transform the product (9) back into the real space of Figure 4. This back-transformed image is just the 2-dimensional inverse Fourier

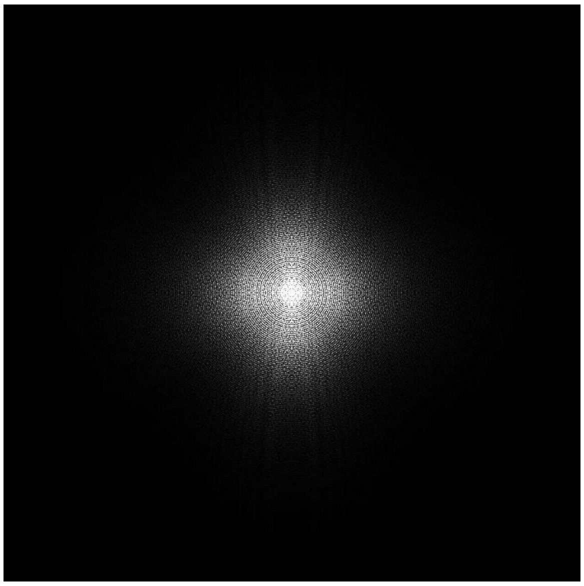

Figure 6: Filtered amplitude spectrum of the test pattern

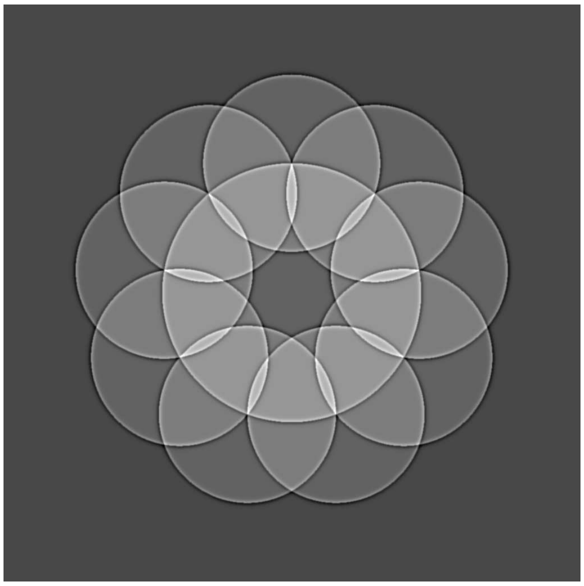

Figure 7: The filtered test pattern

transform of $\hat{T}_{k,l}$ array, i.e.,

$$\hat{t}_{m,n} = \sum_{k=-512}^{511} \sum_{l=-512}^{511} \hat{T}_{k,l} \exp\left\{\frac{j2\pi}{1024}[mk+nl]\right\} , \quad m,\ n = 0,\ 1,\ 2,\ \ldots,\ 1023\ . \tag{12}$$

The result of this back-transforming calculation is shown in Figure 7. It represents the luminance field that would actually be perceived if Figure 4 were the true field. Of course, what one sees when looking at either image has been transformed by the viewer's own personal CSF, but the implied relationship between the two perceived images is the same as the one between the underlying image for Figure 4 and what the viewer perceives when looking at it. To see that this is correct, let $\mathcal{F}$ denote the two dimensional discrete Fourier transform operator and note that the observed image for Figure 4 can be written

$$\hat{t}_{m,n} = \mathcal{F}^{-1}\left[S_{k,l}\mathcal{F}(t_{m,n})\right]\ . \tag{13}$$

This is just a shorthand way of writing Equations (5) and (12). The observed image $\hat{t}_{m,n}$ is also the underlying image for Figure 7, so if $\hat{o}_{m.n}$ is the image that is perceived for that figure, then

$$\hat{o}_{m,n} = \mathcal{F}^{-1}\left[S_{k,l}\mathcal{F}(\hat{t}_{m,n})\right]\ . \tag{14}$$

Comparing Equations (13) and (14) shows that the relatioship between the perceived images $\hat{o}_{m,n}$ and $\hat{t}_{m,n}$ is the same as that between the perceived $\hat{t}_{m,n}$ and the underlying $t_{m,n}$.

Figure 7 differs from Figure 4 in two important ways. The first is that the contrasts between different luminace levels is considerably reduced. The second is the strong Mach bands [9, 10] which form on the boundaries beween adjacent regions with sharply different luminace levels. It is possible to see these bands in Figure 4 by holding the page between 11 and 12 inches from the face, or even closer to really exaggerate the effect.

Acknowledgements

We would like to thank Drs. Philip Sanders, Christine Piatko, and Anastase Nakassis for their advice and encouragement.

References

[1] A.P. Ginsburg, "Spatial filtering and visual form perception," in *Handbook of Per-*

ception and Human Performance: Volume II , Cognitive Processes and Performance, ed. by K.R. Boff, L. Kaufman, and J.P. Thomas, John Wiley and Sons (New York, 1986) Chapter 34.

[2] F.W. Campbell, J.J. Kulikowski, and J. Levinson, "The effect of orientation on the visual resolution of gratings," *Journal of Physiology*, 187 (1966) pp. 427-436.

[3] J.L. Mannos and D.J. Sakrison, "The effects of a visual fidelity criterion on the encoding of images," *IEEE Transactions on Information Theory*, IT-20 (1974) pp. 525-536.

[4] C.F. Hall and E.L. Hall, "A nonlinear model for the spatial characteristics of the human visual system," *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-7 (1977) pp. 161-170.

[5] P.G.J. Barten, "Physical model for the contrast sensitivity of the human eye," *Proc. SPIE, Vol. 1666 (Human Vision, Visual Processing, and Digital Display III)* (1992) pp. 57-72.

[6] M.F. Hutchinson, "Algorithm 642: A fast procedure for calculating minimum cross-validation cubic smoothing splines," *ACM Transactions on Mathematical Software*, 12 (1986) pp. 150-153.

[7] P. Craven and G. Wahba, "Smoothing noisy data with spline functions," *Numerische Mathematik*, 31 (1979) pp. 377-403.

[8] M.J. Gervais, L.O. Harvey, Jr. and J.O. Roberts, "Identification confusions among letters of the alphabet," *Journal of Experimental Psychology: Human Perception and Performance*, 10 (1984) pp. 655-666.

[9] E.M. Lowry and J.J. DePalma, "Sine-wave response of the visual system. I. The Mach phenomenon," *Journal of the Optical Society of America*, 51 (1961) pp.740-746.

[10] T.G. Stockham, Jr., "Image processing in the context of a visual model," *Proceedings of the IEEE*, 60 (1972) pp. 828-842.