INFORMATION TO USERS

While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. For example:

- Manuscript pages may have indistinct print. In such cases, the best available copy has been filmed.
- Manuscripts may not always be complete. In such cases, a note will indicate that it is not possible to obtain missing pages.
- Copyrighted material may have been removed from the manuscript. In such cases, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or as a 17”x 23” black and white photographic print.

Most photographs reproduce acceptably on positive microfilm or microfiche but lack the clarity on xerographic copies made from the microfilm. For an additional charge, 35mm slides of 6”x 9” black and white photographic prints are available for any photographs or illustrations that cannot be reproduced satisfactorily by xerography.
Hue and luminance multiplexing in Type I r-g cells

Billock, Vincent Alan, Ph.D.
The Ohio State University, 1987

Copyright ©1987 by Billock, Vincent Alan. All rights reserved.
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark √.

1. Glossy photographs or pages
2. Colored illustrations, paper or print
3. Photographs with dark background
4. Illustrations are poor copy
5. Pages with black marks, not original copy
6. Print shows through as there is text on both sides of page
7. Indistinct, broken or small print on several pages
8. Print exceeds margin requirements
9. Tightly bound copy with print lost in spine
10. Computer printout pages with indistinct print
11. Page(s) lacking when material received, and not available from school or author.
12. Page(s) seem to be missing in numbering only as text follows.
13. Two pages numbered. Text follows.
14. Curling and wrinkled pages
15. Dissertation contains pages with print at a slant, filmed as received
16. Other

University Microfilms International
HUE AND LUMINANCE MULTIPLEXING IN TYPE I r-g CELLS

DISertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of the Ohio State University

By

Vincent Alan Billock, B.S., M.S.

* * * * *

The Ohio State University

1987

Dissertation Committee:
C.R. Ingling Jr.
P.E. King-Smith
S.W. Smith

Approved by

Graduate Program in Biophysics
While St. Francis may have relied on the kindness of strangers, doctoral candidates depend on the support of friends and colleagues. My thanks to you all:

To Scott Grigsby, for the use of "Grigsbian". Thanks for unpublished material also go to Dr. William Swanson, Dr. Leo Lipetz and Dr. Marion Wienrich.

To Dr. Stanley Smith and Dr. Leo Carney, for serving on my dissertation committee.

To Dr. Ewen King-Smith, for his great patience, exemplary teaching, and enthusiasm.

To Dr. Carl R. Ingling Jr., whose insight, genius, and humanity set a standard to admire and emulate.

To John Letscher and Wesley Gray, for friendship and support, for flashes of wit and flares of brilliance, for serving as a reminder that degrees are a means and life is an end.

To my parents, for their constant support and unwavering love.

Most of all, to Bambi, my wife, for the sacrifice of time that can never be recovered, in the pursuit of a goal she never would have sought for herself. She is a great blessing.
VITA

July 12, 1959 ...................... Born - Youngstown, Ohio.

1981 .............................. B.S., Physics, Miami University.

1983 .............................. M.S., Physics, Miami University.

PUBLICATIONS


FIELDS OF STUDY

Major Field: Sensory Biophysics - Dr. Carl R. Ingling Jr.

Studies in Electrophysiology - Dr. P. Ewen King-Smith.

- Dr. Karl Kornacker.

Studies in Psychophysics - Dr. Carl R. Ingling Jr.

- Dr. P. Ewen King-Smith.

- Dr. Lester Krueger.
TABLE OF CONTENTS

ACKNOWLEDGMENTS ................................................................. ii
VITA ........................................................................................... iii
LIST OF TABLES ................................................................. v
LIST OF FIGURES .......................................................... vi
PREFACE ................................................................................. 1

CHAPTER PAGE

I. MODELLING r-g X-CELL RESPONSE TO SPATIOTEMPORAL CHROMATIC STIMULI ......................... 5

II. SPECTRAL SENSITIVITY OF THE r-g CHANNEL FOR AN ACUITY CRITERION ........................................... 30

III. RECOVERY OF THE HUE AND LUMINANCE SIGNALS OF THE r-g CHANNEL ............................................. 38

LIST OF REFERENCES ........................................................... 59
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The V function compared to probability summation</td>
<td>36</td>
</tr>
<tr>
<td>2. Munsell values, reflectances and chromaticity coordinates for the Mondrian</td>
<td>55</td>
</tr>
<tr>
<td>3. Original and recovered R-G and R+G Mondrian values</td>
<td>58</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The tuning of the r-g channel - Space time domain</td>
<td>24</td>
</tr>
<tr>
<td>2. The tuning of the r-g channel - Frequency domain</td>
<td>25</td>
</tr>
<tr>
<td>3. Wienrich's data for slow moving edge</td>
<td>26</td>
</tr>
<tr>
<td>4. Wienrich's data for faster moving edge</td>
<td>27</td>
</tr>
<tr>
<td>5. Predictions of the model for data in figure 3</td>
<td>28</td>
</tr>
<tr>
<td>6. Predictions of the model for some data in figure 4</td>
<td>29</td>
</tr>
<tr>
<td>7. Vₐ and probability summation</td>
<td>35</td>
</tr>
<tr>
<td>8. Russell's Criterion</td>
<td>54</td>
</tr>
<tr>
<td>9. Spatial configuration of the Mondrian</td>
<td>56</td>
</tr>
<tr>
<td>10. Chromaticity diagram of the Mondrian</td>
<td>57</td>
</tr>
</tbody>
</table>
Type I cells (spatially and spectrally opponent) make up over 90 percent of foveal ganglion cells (Lennie, 1980). By virtue of form and number, this must be the cell that constitutes the r-g opponent channel. By virtue of size and position, it must also be the cell of choice for acuity. It is therefore not entirely surprising that Kelly has shown that the Type I cell can account for almost all of the chromatic and achromatic spatiotemporal contrast sensitivity surfaces (Kelly, 1983). Kelly comments "...that one well behaved type of retinal receptive field can account for so many spatial, temporal and chromatic properties of vision is startling to say the least...clearly other types of pathways must perform other functions, but (except at very low temporal frequencies...) we apparently do not need other pathways to account for sensitivity to stabilized, moving gratings, chromatic or achromatic." The Type I cell must be considered the predominant cell of the primate fovea, the cell that distinguishes primate vision from the vision of magnocellularly dominated animals.

Because of its enormous importance, one would expect that this cell would have been subjected to exhaustive study. Unhappily, the Type I cell has been more a subject of confusion than of penetrating analysis. This is due in part to the difficulty of working with such
small cells, to the preference for the magnocellularly dominated cat as an experimental animal, and to confusion over the relative roles of hue and luminance signals, mixed in the overall output of the cell. Until compelling electrophysiological evidence to the contrary had been produced, it was easier to assume the existence of separate channels to carry each signal. Until recently, the parvocellular channel has received little attention, while great efforts have been expended on the magnocellular X cells (see Enroth-Cugell and Robson (1983), as a survey of the exemplary work that has been done on this cell). While interesting, these results cannot easily be extended to the parvocellular system, because of differences in the organization of these cell's receptive fields and the resultant intermixture of spectral and spatial variables (De Valois, 1975). One illustration of these differences is the bimodal surround of the r-g cells (De Monasterio, 1978). Such a surround gives an amplification of intermediate spatial frequencies that might not be expected from the overall "mexican hat" structure of the receptive field - superficially similar to the "DOG" (difference of gaussians) structure of the magnocellular unit (Ratliff, 1969).

Nineteen years after the discovery of the Type 1 cell (Hubel and Wiesel, 1966), Ingling and Martinez-Uriegas presented a formal linear systems analysis of this unit (Ingling and Martinez-Uriegas, 1985). For the first time, the roles that hue and luminance play in the overall output of the cell are clear. With the removal of this blinder, it is incumbent on us to convert the formal linear systems
analysis into a quantitative analysis. The ability to specify exactly what the cortex will receive from the fovea (and much of the rest of the retina), will open up new vistas for those who study information processing in the visual system. Yet, all of this transmitted information is for naught, if it cannot be decoded. As discussed previously, not only are two independent variables jumbled up in this channel, but the multiplexing of these variables overlaps as well. All schemes that base cortical reconstruction of scenes on zero-crossings (Marr, 1980) or other second derivative features shall come to ruin, unless an algorithm can be concocted to reconstitute the pure signals. Of the reconstruction techniques that might be tried, one has been selected for further study - the Russell Filter (Russell, 1979; Ingling and Martinez-Uriegas, 1983a).

The significance of the Russell Filter is that it is derived not from some optimal mathematical transformation, but rather from psychophysical evidence, and suffers from the same defects as at least one portion of the human visual system. Lu and Fender (1972) found that there is no stereopsis for random dot stereograms when the random dots are equal luminance red and green pixels and that the amount of achromatic contrast required for stereopsis is a function of the two wavelengths. This result can be explained if a certain achromatic contrast is required to raise the output of the bandpass spatial filter until it appears as a criterion "Mach Band" type enhancement on the chromatic lowpass signal which masks it. It follows that the achromatic signal can be inferred from the magnitude of the
enhancement and the chromatic signal can be inferred from the background on which the enhancement sits. This scheme is by no means optimal since its efficiency depends on wavelength; however, it is simple, realistic and firmly grounded on the known properties of the transmission channel. If such a simple algorithm successfully reconstitutes a two variable scene, then we have gained an insight into the calculations of the cortex. Furthermore, the bandpass signal information processing techniques - developed so laboriously by the computational vision experts - need not go to waste. Other applications of this type of analysis might include the explanation of Benham's top and other optical illusions that confuse luminance with color. Finally, the demultiplexing process points up new ways to construct cortical color contrast cells. The confusion that reigns over this area is reminiscent of the early misunderstanding of the role of the r-g cell. Mayhap it will yield to a similar form of analysis.
In 1966, Wiesel and Hubel codified the spatial and spectral properties of the Type I ganglion cell. Their description of a spectrally and spatially opponent cell should have been sufficient. In retinas, as in architecture, structure follows function. Unfortunately, in the Type I cell, function is obscured by structure and the role of the cell remained unknown. Nineteen years after the discovery of the Type I cell, Ingling and Martinez-Uriegas presented a formal linear systems analysis of this unit (Ingling and Martinez-Uriegas, 1985), which demonstrates that the Type I cell response is equivalent to the responses of four achromatic and chromatic, spatially and temporally tuned subchannels. The formal model for an r-g Type I cell is:

$$\text{Response} = \frac{1}{4}(R+G)(\text{Lowpass Spatial})(\text{Bandpass Temporal})$$
$$+ \frac{1}{4}(R-G)(\text{Bandpass Spatial})(\text{Lowpass Temporal})$$
$$+ \frac{1}{4}(R-G)(\text{Bandpass Spatial})(\text{Bandpass Temporal})$$
Where R and G are cone spectral sensitivities, and the filters (e.g; lowpass spatial) are temporal impulse or spatial weighting functions.

For the first time, the roles that hue and luminance play in the overall output of the cell are clear. With the removal of this blinder, it is incumbent on us to convert the formal linear systems analysis into a quantitative analysis. Such an analysis requires specific functions to replace the lowpass and bandpass filters above. Such studies have been performed very successfully for magnocellular X cells and somewhat successfully for Y cells. Most studies of the magnocellular system use flickering or drifting spatial stimuli. Such stimuli are not modulated in color because the magnocellular system is known to be color blind. Until recently, due perhaps to the influence of the magnocellular studies, physiological studies of the parvocellular system suffered from a curious stimulus deficiency. Either the parvocellular system was treated like its magnocellular kin and its color processing capabilities were ignored or color was studied to the exclusion of either spatial or temporal factors. (see DeValois and Pease, 1971; DeValois et. al., 1977) However, in 1982, M. Wienrich completed a study of color opponent cells whose stimuli were a set of colored edges of various hues and luminances, drifting at various velocities. The subjects of the study are Macaca Fascicularis, the primate with color vision most similar to humans (virtually identical). At last, data that can test a full theory of the Type I cells is at hand.
BACKGROUND

To predict the output of a Type I cell for an arbitrary stimulus we need to substitute physiologically inspired spectral response, spatial weighting and temporal impulse response functions. The functions for spatial vision are approximations to the weighting functions of individual cones and to the effects of spatial summation. In the simplest case, the center of a midget ganglion cell is contributed by a bipolar cell whose sole input is a single cone. Invaginating bipolar cells form "on centers" and have dendritic trees that branch low in the inner plexiform layer, while flat bipolar cells form "off centers" and have dendritic trees that branch higher in the inner plexiform layer (Famiglietti and Kolb, 1976; Nelson et al., 1978). The shape of such a center is roughly a "gaussian with a skirt". This shape is the result of the optical properties of both eye and cone. (Richter and Ullman, 1980; Baylor and Hodgkin, 1973). A gaussian approximation to a single cone in the human retina has a space constant of about 25 seconds of arc. (Richter and Ullman, 1980). Surrounds are formed by horizontal cells, each of which can contact six to nine cones in the fovea (Kolb, 1970). Horizontal cells sum the cones that form the surround and inhibit either the bipolar cells (a linear subtraction), or the cones that serve it (a nonlinear shunting process that can be approximated by division). Given that psychophysically the r-g channel is modelled by a linear subtraction (Guth, 1980) and the perfect subtraction of surround from center shown by other X cells (Enroth-Cugell and Pinto, 1970), it is likely that
the horizontal cell inhibits the bipolar cell directly.

Because the cones in the retina are hexagonally packed, a perfectly circular surround is unlikely. As the extent of the surround grows so should the circularity of the surround. Given the absence of impulse conduction in horizontal cells, a cone's contribution to a surround should be an exponentially declining function of distance. In practice, this restriction has little effect on foveal surrounds, because the lack of horizontal interconnections implies that surround extent is restricted to those horizontal cells whose small dendritic fields overlap the midget bipolar.

Outside the fovea, centers and surrounds of most cells scale upwardly in size as the dendritic fields of bipolar and horizontal cells contact more cones. This may be less true of r-g cells than others since data in Hubel and Wiesel (1960) show that a graph of receptive field size as a function of retinal eccentricity is flatter for Type I cells than for other retinal cells. De Monasterio (1980) goes even farther and states that center size of r-g cells "shows little variation with the distance from the foveola and it is typically smaller than 0.02 degrees when measurements are corrected for aberration effects". Compared to the rest of the literature, this statement seems severe. The discrepancy probably lies in the corrections De Monasterio makes for aberrations (optical spread) - aberrations that probably belong in most modeling, if not in electrophysiology. This controversy aside, the relatively flat
receptive field size function bodes well for a channel whose primary
duty is the transmission of high spatial frequencies.

A REVIEW OF FUNCTIONS FOR MODELLING RECEPTIVE FIELDS

There are many functions that can be used to approximate the
spatial response of the centers and surrounds of X cells, the only
requirements being rough symmetry and lowpass response. Comparison of
these functions can be found in Budrikis (1973), in Marr and Hildreth
(1980), and many other sources.

Probably the best data on the spatiotemporal response of the r-g
channel is that of Kelly (1983). Kelly found, as Ingling and
Martinez-Urelias (1983b) had suggested, that the chromatic surface
could be modelled by summing center and surround mechanisms, and the
achromatic surface could be modelled by a difference of the same
mechanisms. The center and surround mechanisms employed are in the
form of spatial and temporal excitatory and inhibitory contrast
sensitivity functions (Burbeck and Kelly, 1980) and follow a template
previously found by Kelly to be useful:

\[ S(f) = K + Af|\log(bf)| \exp(-bf) \]  

Unfortunately, the inverse transform of this template function is
difficult. Although point spread and impulse response functions could
be calculated numerically, such an approach is cumbersome, if one wishes to make small changes in the model. Probably the best course would be to replace Kelly and Burbeck's functions with a more mathematically tractable form, even if the fit to their CSF data is slightly degraded.

Spatial Functions

The most widely used lowpass function for spatial vision is the gaussian. The popularity of the gaussian lies in its familiarity, in its long historical tenure and in its mathematical properties. The gaussian is unique among functions, in that its Fourier Transform is also a gaussian. This property led Marr to declare the gaussian to be uniquely localized in space and frequency, a property which had value in Marr's engineering based analysis of an ideal perceiver.

\[
F(r) = \frac{1}{2\pi a^2} e^{-\frac{r^2}{2a^2}} \quad \text{Equation 3}
\]

\[
G(f) = e^{-2\pi a f^2} \quad \text{Equation 4}
\]

A useful property of the gaussian is separability. The radially symmetric gaussian can be separated into the product of two functions, each of a single variable. This simplifies convolution, since a 2D
convolution can be broken into one dimensional convolutions in each direction.

Another useful function is a simple square or "delta" weighting function often used in machine vision. It has the advantage of simplicity and its parameters are easily estimated from receptive field maps. Although unlikely to be representative of large cells, whose receptive fields are formed by spatial summation weighted by distance, the square weighting function is probably not a bad representation of a midget ganglion cell receptive field element. Unfortunately, the square weighting function has side lobes in its frequency transform and could distort the perception of complex images. This can be taken care of by preconvolving the image with a gaussian to mimic the effect of the eye's optics and then convolving the image with a receptive field made up of square elements. This is the origin of Marr's Laplacian of a Gaussian and is a versatile approach since one can preconvolve either the target or the receptive field with the gaussian.

Another function of some interest is the Cauchy point spread function, a function which shares some of the best attributes of the square and gaussian functions. This function is somewhat different in appearance than the gaussian, being rather squared off. It may therefore be closer to the cone point spread function than the gaussian. Unlike the square function, it remains limited in the frequency domain, with no side bands. Its transform is an exponential
function of frequency. Interestingly, Budrikis (1973) in his comparison of filter types found that the Cauchy function gave a best fit to contrast sensitivity data when employed in a simple center-surround model.

\[ F(r) = \frac{a}{2 \pi (a + r)^{2 + 3/2}} \]  
\[ G(f) = e^{-2af^2} \]

Equation 5

Equation 6

These three functions, when used to represent centers and surrounds, are good models for receptive fields. There are two approaches in the literature for the combination of centers and surrounds. Both use gaussians. The most popular approach for modelling X-cells is the difference-of-gaussians (DOG). This method is more appropriate for the magnocellular system, where it has been employed with great success in electrophysiology (Enroth-Cugell and Robson, 1983; Rodieck, 1965) and in psychophysics (Wilson et. al., 1983). There is at least one attempt to apply the model to the parvocellular system (Richter and Ullman, 1980). As in all DOG models, the center is a narrow gaussian and the surround is a wider gaussian, extending through, and peaking at the center. This is not an appropriate model for the Type I cell, because a color opponent cell must keep the R-cone from contaminating the G-cone center, and vice-versa. Computationally, it is important to get the form of the surround
correct because a bimodal surround acts in synergy with the center to produce a gain for intermediate frequencies, while a cell with a unimodal surround can only reduce gain to nonpreferred frequencies (Ratliff, 1969).

Ingling and Martinez's square wave approximations are therefore closer to the filter characteristics we expect for the cell, but suffer from not being bandlimited. The sharpness of the corners of these functions introduces side lobes into the frequency spectrum and shows up as a distortion in the zero-crossings of the convolution product of the filter and the stimulus (Marr, Poggio, and Ullman, 1979). A better approximation would be a cauchy function or a gaussian for each cone. These mimic more closely the effect of convolving the cone photon collector surface with the optical point spread function and the transforms of these model functions are not afflicted with frequency side lobes. These functions may be set apart at the proper distance for cones in the primate retina or given the width found in physiological or psychophysical experiments. Still, it will be of interest to see if the simple square wave approximation will suffice, and this possibility will be explored.

At least one researcher has explored the use of the difference of offset gaussians (DOOG) (Young, 1986). Pioneered by Hallet (1971) for rod receptive fields, the DOOG is an elegant device. Young fit a DOOG to a primate r-g point spread function and calculated a line spread function. Young reports an excellent fit to the human line spread
function – better than that obtainable from DOG's or Gabor functions.

Temporal Functions

In cell modelling the gaussian does double duty, since it can represent the temporal impulse response function. To use the gaussian, shift the entire function to sufficiently positive values of time such that

\[ F(t) = \frac{1}{(2\pi T)^{1/2}} e^{-\frac{1}{2}((t-To)/2T)^2} \]

Equation 7

\[ G(f) = e^{-(2\pi T f - j2\pi f To)} \]

Equation 8

Budkris (1973) suggests that for To/T>3 one may assume that V(t) is the function for all t, including t<0.

A function which has been used more extensively is the simple exponential response.

\[ F(t) = \frac{1}{T} e^{-t/T} \]

Equation 9

\[ G(f) = \frac{1}{1 + j2\pi f T} \]

Equation 10
This function is popular for two reasons. First, it is the response of an RC filter, often used to model neural membrane response (Jack et al., 1970). Second, there exists an approximation to convolution with an exponential. Define $R(t)$ as the convolution of the stimulus $S(t)$ and the exponential filter (equation 9). It can be shown that the recursive equation for exponential convolution is

$$R(t) = e^{-T dt} R(t-dt) + (1-e^{-T dt})S(t)$$

Equation 11

See Fleet et al. (1984) for a derivation of this relation. The recursive relationship defined above will hold true for $dt$ small enough to represent the frequencies present in the signal (Shannon's Sampling Theorem). Since use of the exponential recursion relation saves a great deal of storage space and computing time, it is little wonder that it has been employed so frequently.

One troublesome point about the exponential is its immediate response. Unlike psychophysical impulse response functions, the exponential function begins with its maximum response and then decays over time. Ratliff et al. (1969) have offered a similar function that avoids this criticism and offers a variety of options.
Here $K$ is a constant equal to the desired integrated response, $T$ is the relaxation constant, $T_0$ is the onset before response, and $n$ affects the nature of the onset: $n=0$ abrupt, $n=1$ linear, $n=2$ parabolic. For limulus, Ratliff found the best fit to data with $n=0$ abrupt.

A variety of papers have suggested that diffusion is the culprit in high frequency flicker falloff (Kelly, 1971a; 1971b; Kelly and Wilson, 1978).

\[
F(t) = \frac{K}{n!} \frac{(t-T_0)}{T^n} e^{-(t-T_0)/T} \tag{Equation 12}
\]

Kelly and coworkers have achieved notable success in the use of one or more stages of diffusion in modelling psychophysical results. One stage probably originates in the inner segment of the cones, the other is in the outer segment. Budrikis' study showed that for a simple Excitatory - Inhibitory model, the best fit to the contrast
sensitivity surface was obtained using diffusion for the temporal response.

Another model which has had wide applicability is the n-stage linear filter, basically an n-stage RC filter. The model suggests that the temporal signal is filtered through n RC filters of similar time constants.

\[
F(t) = \frac{(n-1) \cdot (-t/T)}{(t/T) \cdot e} \cdot \frac{1}{T(n-1)!} \quad \text{Equation 15}
\]

\[
G(f) = (2\pi f)^2 \cdot (c + f) \quad \text{Equation 16}
\]

The system has a corner frequency \( c \) of \( 1/(2\pi T) \) - below this frequency the amplitude asymptotes at \( T^n \), above this frequency it approaches an asymptote of \( (2\pi w)^{-n} \). In log-log coordinates the first portion of the curve is a flat line and the second portion has a slope of \(-n\).

**SPECTRAL SENSITIVITY AND ADAPTATION**

Cone spectral sensitivity functions are hard to obtain directly and are best inferred from linear transformations of the psychophysical color matching functions of normal and color vision deficient observers. Once obtained, one can weight the radiances of the stimulus at each wavelength with the appropriate cone
sensitivities to obtain cone outputs. When the radiance of the target is large, the output of the cone is affected by pigment bleaching and adaptation. Both phenomena arise from similar physical processes and have the same functional form - the hyperbolic tangent function. Let \( I_0 \) be the half bleach constant of 20000 Trolands, \( S \) is the cone's sensitivity, \( I \) is the light intensity, \( K \) is a half saturation constant of 10000 Trolands and \( n=0.7 \). The overall equation for the cone output is

\[
\text{Response} = \frac{\left(\frac{ISI_0}{I+I_0}\right)^n}{\left(\frac{ISI_0}{I+I_0}\right)^n + K}
\]

Equation 17

Finally, the output of the cone must be weighted by a numerical factor which accounts for its participation in the channel. For example, in an r-g cell, for cone sensitivity functions with a maximum height of one, the ratio of \( R \) and \( G \) weighting functions must be roughly 2/3 so that the function r-g will have a zero crossing at 580 nm.

**MODELLING WIENRICH'S DATA**

Spatiotemporal convolution can then be implemented by treating the moving stimulus as a function of time and weighting it with the spatial function as the stimulus passes. The resulting spatial
convolution as a function of time, can then be convolved with the
temporal impulse response function. These convolutions must be carried
out for center and surround or for each subchannel in the Ingling and
Martinez-Uriegas expansion. Each method yields its own insights. The
total response is added to a constant firing rate and may be compared
to actual data, such as Wienrich's.

From Wienrich's extensive study, I have selected for analysis the
poststimulus firing histograms of an eight degree parafoveal r-g cell
to a couple of moving colored edges. The stimulus consists of adjacent
red (631nm) and violet (451nm) fields. The cell starts out covered by
the 451 nm field, but the field drifts until the cell is well within
the confines of the 631 nm stimulus, then the direction of drift
reverses and the 451 stimulus again passes over the cell. This
alternation elicits a characteristic response from the cell, which
varies with edge contrast and edge velocity.

To model this cell's response I selected a square wave line
spread function for the spatial response. The width of each square was
set at .1 degree, a value consistent with the literature on this cell.
(Hubei and Wiesel, 1960, Zrenner 1982, Young 1985) Although other
functions might have been employed, it would be difficult to tell the
difference between a simple edge convolved with a square wave line
spread function and the same edge convolved with a gaussian or cauchy
function. The temporal response function chosen was the venerable
linear filter. This choice was dictated by availability of plausible
temporal data. Actual data on the temporal response of the r-g channel is sparse, actual fits to filter models are almost unknown. Swanson has fit data on the sensitivity of the r-g channel to combinations of luminance and chromatic flicker to the multiple stage linear filter model (Swanson, 1987). A particularly good fit is found for a five stage excitatory center with a time constant of .00655 seconds and a seven stage inhibitory surround with a time constant of .00765 seconds. The choice of spectral sensitivity functions is slightly more complicated and is based on the following considerations.

Figure 3 illustrates Wienrich’s results for one edge (451 nm, 5.8 log units of quanta/sec/square micrometer; 631 nm, 5.7 log units of quanta/sec/square micrometer) and Figure 4 shows results for another (6.0 log units of quanta/sec/square micrometer for both wavelengths). As the two figures show, there is a luminance enhancement at the edge when the edge is moving at moderate velocities. Unfortunately, the model predicts that for the given stimuli and human cone sensitivities, no enhancement occurs. The stimulus violates Russell’s criteria for enhancement, i.e., the human G-cone is more sensitive to the 451 nm stimulus than the 631 nm stimulus (see Chapter 3, also Russell, 1979). The cure is to not use human cone sensitivities. The similarity between human and macaque color vision is slightly misleading. The wavelength locations of human and macaque maximum cone sensitivities vary by as much as nine nm. At moderate wavelengths this has little effect but at the spectral extremes the effects can be quite large. Cone sensitivities are difficult to shift because the
shape of the spectral sensitivity function changes with the location of the maxima, while preretinal absorptions do not. Using an algorithm by Lipetz (1987), I have shifted the G-cones from 531 to 532-537 and the R-cones from 558 to 565-570 nm, using a rhodopsin template. Just a small shift in the G-cones is sufficient to create the desired enhancement (Figures 2 and 4). The best values were found to be around 537 nm for G-cones and 567 nm for R-cones. This value was initially troublesome for two reasons. First of all, of the range of maximum values for this cone, 537 nm is at the extreme long wavelength side. Second, the general trend in the literature for G-cones is toward shorter wavelengths. Typical of these is the particularly meticulous work of Baylor (in press) who reports a maxima of 531 nm. However, Baylor notes that when the G-cone function is fit by shifting iodopsin templates, the best maximum is 538 nm, a value remarkably similar to our own, obtained by the method most appropriate to our own.

Swanson's temporal model included a parameter which set the strength of the surround relative to the center. In the case of the Smith and Pokorny sensitivities that Swanson employed, this worked out to a crosspoint of about 600 nm for the best temporal filter parameters. This is slightly at variance with the well known 580 nm channel crosspoint for the r-g system. Interestingly, a 600 nm crosspoint gives a good fit to our data. This should cause little consternation, since individual r-g cells have crosspoints that vary considerably. Swanson's fit is to an entire channel and the deviation from the channel crosspoint is curious. It may arise from Swanson's
model cell (G on center, R off surround). It may be that a more sophisticated version of Swanson's model which used both R and G centers, could redress this discrepancy.

Figures 5 and 6 illustrate the performance of the model. As noted above all of the parameters of the model were chosen from the literature of physiology and psychophysics and were allowed to vary only within published ranges. The only truly free parameters were resting firing rates of the cell and the gain of the responding cell relative to its resting background. These parameters were useful for scaling the output of the model to the magnitude of Wienrich's recordings but have no real effect on the convolution profile. Given these restrictions, the fit of theory to data is quite reasonable. Several points about this fit deserve notice.

The increase in enhancement with velocity is associated with an increase in the $(R+G)\text{(Lowpass Spatial)}(\text{Bandpass Temporal})$ component to the total response, from less than 1% at 3 deg/sec to about 4% at 17 deg/sec.

As pointed out in the discussion, all of the parameters involved are physiologically or psychophysically inspired. It is therefore not surprising that the fit between the model and experimental evidence is not perfect. The model is meant to reflect the performance of the r-g channel, not the r-g cell. However it is apparent that the fit can easily be improved. With some deviation from Swanson's parameters we
could get a better fit to the width of the overshoots. Changes in filter type, or small changes in size are of less help, as the data lacks the resolution to aid our selection. Other problems with the fit are more difficult to fix. For example, in Figure 4 the decrement in response to the end of the red field is much greater than the predicted decrement.

The ability to specify exactly what the cortex will receive from the fovea, (and much of the rest of the retina) will open up new vistas for those who study information processing in the visual system.
The r-g spatiotemporal chromatic receptive field is factored into its center and surround components, and rewritten as a determinant. This determinant is then expanded into a set of terms with psychophysically meaningful spectral sensitivities (R+G and R-G). The analysis shows that the r-g cell behaves as if it were composed of four simple receptive fields: achromatic bandpass spatial, lowpass temporal; achromatic lowpass spatial, bandpass temporal; chromatic lowpass spatial, lowpass temporal; and chromatic bandpass spatial, lowpass temporal. (Reproduced from Ingling and Martinez-Uriegas (1985) with the permission of C. R. Ingling Jr.)
Fourier transformation of the impulse response and spatial weighting functions shown in Figure 1 to the frequency domain. (Reproduced from Ingling and Martinez-Urriegas (1985), by permission of C. R. Ingling Jr.)
Response patterns for an R on-center, G off-surround cell (Macaca Fascicularis, eight degrees perifoveal). The cell is stimulated by a moving edge. One side is 451 nm, 5.8 log units quanta/sec/square micrometer; the other side is 631 nm, 5.7 log units quanta/sec/square micrometer. The edge moves back and forth across the receptive field, from 15 degrees into the 451 nm side to 15 degrees into the 631 nm side to 15 degrees into the 451 nm side, and so on. Edge velocities are 3 and 17 degrees/second. Note the edge enhancement for the high velocity case. (Reproduced from Wienrich and Zrenner, (private communication) by permission of M. Wienrich.)
Response patterns for an R on-center, G off-surround cell (Macaca Fascicularis, eight degrees perifoveal). The cell is stimulated by a moving edge. One side is 451 nm, 6.0 log units quanta/sec/square micrometer; the other side is 631 nm, 6.0 log units quanta/sec/square micrometer. The edge moves back and forth across the receptive field, from 15 degrees into the 451 nm side to 15 degrees into the 631 nm side to 15 degrees into the 451 nm side. Edge velocities range from 34 to 71 degrees/second. (Reproduced from Wienrich and Zrenner, (private communication) by permission of M. Wienrich.)
Comparison of the model to Wienrich's results in Figure 3. See text for details on the model.
Comparison of the model to Wienrich's results in Figure 4. See text for details on the model.
It is a curious fact that luminance measured by an acuity criterion tapping the parvocellular system, strongly resembles luminance as measured by tapping the magnocellular channel. This is curious for two reasons.

First, the parvocellular system is subtractive, not additive. This contradiction was addressed by Ingling and associates in a series of papers (Ingling, 1982; Ingling and Martinez, 1983a; Ingling and Martinez-Uriegas, 1983b; and others) which made clear the presence of an additive signal \((R+G)\) signal in the \(r-g\) cell. Considering only spatial and chromatic interactions, the total signal is

\[
(R-G)(\text{LOWPASS SPATIAL FILTER})/2 + (R+G)(\text{BANDPASS SPATIAL FILTER})/2
\]

Equation 18

Thus, any target that contains a luminance change over space produces a signal in the \(r-g\) cell that is additive. Therefore, attempts to measure luminance with such targets produce additive luminance.
QED? Not quite. The CIE $V_\lambda$ function, which represents luminosity for the standard human observer, implies a weighting of R-cones to G-cones of about 5R:3G (13R:8G to be more exact), while the r-g cell has a weighting of 2R:3G, as determined by the fact that r-g=0 at 580nm. (These weightings assume Smith and Pokorny cone sensitivity functions with a maximum sensitivity of 1.) There are various ways to view these weightings and for the r-g cell the 2R:3G weighting probably reflects a difference in potency between centers and surrounds. Such differences represent the number of cones in both and the difference in gain between center and surround mechanisms. This explanation is unlikely to explain the 5:3 ratio for Y cells because large receptive field size suggests that both centers and surrounds should have the same ratio of cones. This 5:3 ratio suggests an alternate hypothesis: that R-cones are about twice as numerous as G-cones - a notion with much support in the literature (Walraven, 1974; Cicerone and Nerger, 1985; and others). In other words; flicker counts cones while acuity should not. The problem is that acuity produces a luminosity function indistinguishable from $V_\lambda$. (Pokorny et. al., 1968).

Ingling and Tsou (1987) have developed a model which explains this finding by exploiting the differences between center and surround spatial sensitivities. Centers, being smaller than surrounds, have much higher spatial frequency cutoffs. In the case of the simple DOG
function described in Chapter 1, with a surround 3 times wider than the center, the spatial frequency cutoff for centers is four times greater than the cutoff for surrounds (Enroth-Cugell and Robson, 1984). Hence acuity targets are resolved by centers alone, and if the ratio of R-cone centers to G-cone centers is the same as the ratio of R-cones to G-cones, then acuity counts cones, as does flicker.

There is a tacit assumption here - that somehow the R-cone responses of r-g cells and the G-cone responses of g-r cells are added. A plausible model for this addition is probability summation. The cortex receives the R-cone signals and the G-cone signals. It has a certain probability of detecting an R-cone signal and a certain probability of detecting a G-cone signal amidst the noisy neural background. The cortex has an even better probability of detecting something if both signals are present, a sensitivity enhancement known as probability summation. Probability summation is an inherently nonlinear process, and given 5R:3G it should not produce a 5R+3G luminosity function. The question addressed by this study is: how different from $V_A$ is a probability summation model?

Vision researchers use an approximation to probability summation that has great utility. The approximation, shown by Quick (1974) to be appropriate is

$$\text{Response} = \frac{1}{n} (a + b + c + \ldots \ldots)$$

Equation 19
Where $N$ is the slope of a frequency of seeing curve for the stimulus, and $a, b, c \ldots$ are the signals of independent channels. A great variety of these studies have been made in vision; some typical values from studies involving luminance and chromatic mechanisms are $n=2$ (Guth, 1980; Ingling and Tsou, 1977), $N=2.5$ to $3.5$ under various conditions (Cowan et. al., 1984) and $n=4$ (Kranda and King-Smith, 1979). In the absence of frequency of seeing data, researchers have resorted to exploring the range of 2 to arbitrary values (Lindsey, et. al., 1986); our study explores the range of $n$ from 2 to 10. The equation used is

$$n \quad n \left(\frac{1}{n}\right)$$

$$\text{Response} = \left(\left(aR\right) + \left(bG\right)\right)$$

Equation 20

The results (Table 1) are encouraging. For example, for $n$ equal to 2, $V_\lambda$ and probability summation agree fairly closely. The acuity spectral sensitivity function is difficult to measure and hence probability summation is a viable model for acuity spectral sensitivity over the usual range of $n$.

It would be interesting however, to ask the question - what should the weighting of R-cones to G-cones be to produce $V$? A useful statistic is cumulative percent error, where

$$\% \text{ ERROR} = \frac{\text{ABS}(V_\lambda - \text{Prob Summation})}{V_\lambda} \cdot 100$$

Equation 21
for each wavelength. In this way, the low sensitivity spectral extremes are given the same weight as other regions of the spectrum. The results are interesting. Although 13R:8G (a 490 nm cone crosspoint) produces good results for n=2, a 2R:3G (a 580 nm crosspoint) produces an even better approximation to \( V_A \) (after being exposed to probability summation) than the 490 crosspoint universally associated with \( V_A \). To be more exact, when n=2, a 560 nm crosspoint is optimal. For n=3, a 570 nm crosspoint is best. For n=4 to 10, a 580 nm crosspoint is desirable. Is the proximity of these crosspoints to the 580 nm crosspoint of the r-g channel a coincidence? Probably. To take the result seriously implies that red centers and green centers have gains which preserve the weighting of red to green cones, rather than the difference being in the number of cones present in the surround. This may be true but there is no supporting evidence for this contention.

In conclusion, the probability summation model for receptive field centers supports the contention that r-g cells produce \( V_A \)-like luminosity functions for an acuity criterion. The model also demonstrates the need for more information on receptive field formation in the parvocellular system. The entire problem, taken in context, raises an interesting point. Four separate criteria - critical flicker fusion, heterochromatic flicker photometry, minimum border and acuity - from two different systems - magno and parvocellular - all produce the same spectral sensitivity. Nature has gone to great lengths to produce a system of consistent spectral
sensitivity. For such a system to have evolved implies a survival advantage in luminance consistency, a notion that deserves further consideration.
| \( \lambda \) (nm) | \( V(\lambda) \) (log10) | 490 | 560 | 570 | 580 | 590 | 600 | 610 | 620 | 630 | 640 | 650 | 660 | 670 | 680 | 690 | 700 |
|-----------------|-----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 400             | 0.655           | 0.648| 0.666| 0.674| 0.685|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 410             | 0.971           | 0.962| 0.984| 0.994| 1.008|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 420             | 1.245           | 1.234| 1.263| 1.277| 1.294|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 430             | 1.438           | 1.424| 1.468| 1.489| 1.514|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 440             | 1.581           | 1.564| 1.627| 1.658| 1.690|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 450             | 1.672           | 1.655| 1.736| 1.776| 1.813|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 460             | 1.780           | 1.766| 1.858| 1.904| 1.943|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 470             | 1.961           | 1.947| 2.039| 2.085| 2.124|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 480             | 2.145           | 2.129| 2.209| 2.249| 2.286|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 490             | 2.320           | 2.303| 2.371| 2.404| 2.437|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 500             | 2.511           | 2.495| 2.552| 2.580| 2.611|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 510             | 2.704           | 2.688| 2.737| 2.761| 2.788|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 520             | 2.853           | 2.839| 2.880| 2.899| 2.923|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 530             | 2.938           | 2.926| 2.958| 2.973| 2.992|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 540             | 2.982           | 2.972| 2.995| 3.005| 3.020|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 550             | 2.999           | 2.995| 3.007| 3.012| 3.020|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 560             | 3.000           | 3.000| 3.000| 3.000| 3.000|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 570             | 2.981           | 2.989| 2.974| 2.968| 2.958|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 580             | 2.942           | 2.961| 2.929| 2.920| 2.897|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 590             | 2.881           | 2.916| 2.866| 2.857| 2.825|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 600             | 2.802           | 2.856| 2.789| 2.786| 2.750|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 610             | 2.704           | 2.776| 2.699| 2.702| 2.667|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 620             | 2.583           | 2.671| 2.587| 2.596| 2.563|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 630             | 2.425           | 2.525| 2.438| 2.450| 2.417|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 640             | 2.245           | 2.353| 2.265| 2.279| 2.246|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 650             | 2.032           | 2.145| 2.056| 2.071| 2.039|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 660             | 1.788           | 1.904| 1.815| 1.830| 1.798|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 670             | 1.507           | 1.626| 1.536| 1.552| 1.520|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 680             | 1.233           | 1.354| 1.264| 1.280| 1.248|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 690             | 0.915           | 1.037| 0.947| 0.963| 0.931|     |     |     |     |     |     |     |     |     |     |     |     |     |
| 700             | 0.615           | 0.741| 0.651| 0.667| 0.635|     |     |     |     |     |     |     |     |     |     |     |     |     |
Comparison of the CIE V function (solid line) to the predictions of probability summation for n=2 and a 490 nm crosspoint (broken line).
There are several advantages to the design of the Type I cell. Because of multiplexing, the parvocellular system enjoys a fifty percent reduction in bandwidth. Spatial vision benefits from not having to sacrifice any nerves to color vision, and both variables are kept in perfect registration. Nevertheless, a transmitter that multiplexes hue and luminance would be of little use to the visual system if there were no way to decode the signal at the receiver. As we will see, a lack of understanding on the composition and decomposition of the r-g cell signal has led to at least two serious misunderstandings on the roles of retinal and cortical cells.

There are several ways to go about decomposing the r-g signal. This chapter explores three such approaches in various detail. One should be cautious in adopting any of these competing schemes. It may help to keep the following principles in mind when contemplating any reconstruction scheme.
First, the system should be generalized to almost any stimulus and not require a lot of advance knowledge about the stimulus.

Second, the system should not outperform real humans. Human subjects do not perfectly reconstruct complex stimuli, as the Lu and Fender experiment (1972) illustrates. We should expect illusions and reconstruction failures to occur mathematically as well as psychophysically. Their occurrence is proof that we have not strayed too far from the psychophysiological rock upon which all such models should be built.

Three, algorithms should be robust—in the sense that they obey Huggins and Licklider's (1951) "principle of sloppy workmanship":

"The principle of sloppy workmanship states that it is dangerous to postulate a neural structure that is precisely arranged in detail. The nervous system is the product of a superb architect and a sloppy workman, and in his plans the architect took into account the fact that the workman would not get all the terminal boutons where they belonged. One of the basic facts of neurophysiology is that the nervous system works despite a considerable amount of misarrangement of detail. That fact should be taken into account in constructing theory. Taking it into account is often incompatible with simplicity of representation. One neuron in the diagram must represent a thousand in the nervous system. Nevertheless, it is important to keep in mind that a statistical interpretation of details is required."
There are probably a limitless supply of clever reconstruction algorithms that can be created, most of which will violate the axioms above. For an example of an algorithm that violates all of these rules, consider the following - created by this author in 1983.

Given an abrupt edge and a standard r-g receptive field, with a total width three times the width of the center, and the integrated area of the surround equal to the center; convolve the edge with the cell, using the spatio-chromatic identity (See Chapter 2). When completely on either side of the edge, (R+G) Bandpass will have zero signal, and (R-G) Lowpass will have a constant signal. Therefore, the only points of interest from the convolution will be the signals from two center widths before the edge to two center widths after the edge. Assume the R-G signal is greater for the first point than the last. Then the edge can be reconstructed by calculating the lowpass drop and differencing this from the overall signal of the cell, yielding the bandpass signal. For example, at the point of which half of the surround has entered the dim side of the field, the lowpass signal should drop to seventy five percent of its previous value. After the receptive field has moved another 1/3 of its width the lowpass signal is down to a quarter of its original value. If we denote the (R-G) lowpass signal on the bright side of the field as (R-G)1 then
\[(R+G)2 \text{ Bandpass} = \text{Total signal} - 0.75((R-G)1 - (R-G)4)\]

Equation 22

\[(R+G)3 \text{ Bandpass} = \text{Total signal} - 0.25((R-G)1 - (R-G)4)\]

The algorithm performs to perfection, a definite tipoff that it is wrong. More importantly, it requires special knowledge on the location and type of edge involved and it would require an unreasonable neural network to implement.

A second method for demultiplexing the r-g signal was created by Martinez-Uriegas (1985). This elegant algorithm postulates the existence of spectrally conjugated X-cell pairs, i.e.; for every r-g cell, there is g-r cell in the same location. Martinez-Uriegas denotes these two cells as X and Xc. Each of these two cells has similar spatial and temporal properties. At the cortex, X and Xc cells serve as excitatory and inhibitory inputs to other cells. Given

\[X = RSeTe - GSiTi = (R+G)(SeTe - SiTi)/2 + (R-G)(SeTe + SiTi)/2\]

and

Equation 23

\[Xc= GSeTe - RSiTi = (G+R)(SeTe - SiTi)/2 + (G-R)(SeTe + SiTi)/2\]

Where Se, Si, Te and Ti are the spatial and temporal excitatory and inhibitory functions, then simple addition and subtraction yield
\[ X + X_c = (R+G)(SeTe - SiTi) \]

and

\[ X - X_c = (R-G)(SeTe + SiTi) \]

Of course, one can also postulate \( X_c - X \), in which case one computes a G-R signal.

In many respects this is an excellent model. It reproduces the stimulus precisely, in a simple, easy to implement method, consistent with known physiology (ordinary excitation and inhibition). However, there are several objections that can be raised to this model. First, one can object to the definition of location. The model suggests that a R center and a G center share the same space. Clearly the best that these two opponent color cells can do, is to have the surround of one cell contain the center of the other.

A more serious objection is related to bandwidth considerations. As mentioned previously, one suspected motivation for using r-g cells, despite the ambiguity of the signal, is that given an adequate receiver, only half as many transmitters are required to signal hue and luminance. This is an important consideration in a system like the fovea, that requires the highest possible acuity, and color information as well. In essence, we are trying to get by with one equation and two unknowns, because one equation is all we can afford. Martinez-Urriegas attempts to solve the problem by supplying a second equation, at the cost of the benefit we desired. Why have the r-g cell
at all? Martinez-Uriegas speculates: "to keep in spatiotemporal register the information about chromatic and achromatic changes that occur simultaneously and in the same place" (Martinez-Uriegas, 1985). I do not believe that this is a sufficient reason for nature to have gone to all this trouble.

A third problem is that the recovery is just too good to be true. The available evidence suggests that recovery is dependent on a criterion enhancement (Russell, 1979). Martinez is aware of this and suggests that the brain may use both systems, exploiting in some cases a possible speed difference between the two (private communication, 1987). If examples of near perfect recovery could be found, or if the Lu and Fender data (see below) could be explained in a different way, Martinez's theory would gain more credibility.

It may be possible to meet some of these objections with the following refinement. Let each cone give rise to a receptive field center (true at least in the fovea). Therefore, each surround gives rise to several receptive field centers, arrayed radially about the original center and of opposite spectral response. For example, an r-g cell is surrounded by g-r cells, each of whose center is contributed by cones that also are contributing to the r-g cell surround. Likewise, the g-r cells are surrounded by r-g cells constructed in the same way. Then, all we need postulate is that nearest neighbors add and subtract in exactly the way suggested by Martinez-Uriegas, with a weighting factor that makes a cell's response equal to the sum of its
nearest neighbors. Obviously, real retinas do not contain equal numbers of R-G and G-R cells, arrayed in perfectly ordered two dimensional lattices. Such a model would only be of use to machine vision. It would be of interest to construct a model retina of r-g, g-r, -r+g and -g+r cells, scattered about in some imperfectly ordered way, and apply this algorithm to it. Certainly, reconstruction would not be perfect, but it would be interesting to note what sort of imperfections arise.

There is a third model for reconstruction available, which is based on psychophysical data and is at least physiologically plausible. Lu and Fender (1972) found that there is no stereopsis in Julez random dot stereograms when the stereograms are composed of equal luminance colored pixels of various hues. The luminance difference required for stereopsis was not constant and turned out to be a rather odd function of wavelength. Russell (1979) found that the result could be explained by an early model of Ingling's (Ingling and Drum, 1973). Ingling and Drum showed that the response of the r-g cell to a chromatic edge is monotonic, while the response to a luminance edge is enhanced. The combination of the two will yield a Mach band type signal if a set of inequalities is obeyed. These inequalities relate center and surround response. For example, if the response of the R-cone center to the first side of an edge is greater than the response of the center to the second side, then for enhancement to occur, the response of the G-cone surround to the the first side of the edge must also be greater than the response of the surround to the
second side, i.e., if $R_1 > R_2$ then $G_1 > G_2$ for enhancement to occur. In essence, this requirement means that when the cell is convolved with the edge, as a portion of the surround moves into the dimmer side of the edge, it inhibits the center less, producing an enhancement on the overall response of the cell. Likewise, when the center moves into the dimmer side it is still being inhibited by that portion of the surround responding to the stronger side of the edge and an undershoot is produced on the cell's overall response. These overshoots and undershoots bracket the position of the edge. Russell postulated that their existence and detection was essential for stereopsis to proceed.

Another, but equivalent view, based on their subchannel analysis, was presented by Ingling and Martinez (1983) and is displayed in Fig. 8, reproduced from Ingling and Martinez-Uriegas (1983). Enhancement occurs when the increase in the $(R+G)$ bandpass signal across an edge is greater than the decrease in the $(R-G)$ lowpass signal. This occurs when

$$I_1(R_1+G_1) - I_2(R_2+G_2) > I_1(R_1-G_1) - I_2(R_2-G_2)$$

Equation 25

Ingling and Martinez (1983) postulate that the output of the cell is filtered at the cortex to remove the DC response. A simple Laplacian filter would be adequate for the task. The DC response of the cell provides an estimate of hue. The height of the enhancements provide an estimate of luminance. The position of the enhancements mark the edges. Since hue between the edges is known, a two variable
specification of a scene can be obtained by interpolating luminance from the heights of the enhancements. Ingling and Martinez (1983) call this procedure "Russell Filtering".

At first blush, this is an algorithm that seems to satisfy the principles set at the beginning of this chapter. It requires no special knowledge of the scene. It succeeds and fails as a human observer would, as displayed in the data of Lu and Fender. The model is robust, dependent only on the detection of an enhancement known to occur, and is physiologically sound, dependent on cortical filtering and interpolation, well established mechanisms. Not being perfect, like the previous two models described, mandates a test of this model's capabilities. One good test would be to encode and reconstruct a Mondrian.

RECONSTRUCTION OF THE MONDRIAN BY RUSSELL FILTERING

Mondrians are stimuli created by juxtaposing polygons of various sizes, hues and luminances. Originally created by the artist Paul Mondrian, Mondrians are usually constructed from patches of papers, but are sometimes simulated on display terminals as well. They are generally used to test theories of brightness and color constancy and hence are natural stimuli with which to test hue and luminance reconstruction. Grigsby (private communication) has created a Mondrian that is well suited for this purpose. The individual patches are assigned random Munsell chip parameters. Hue, Value and Chroma were
selected at random from the Munsell range. Munsell chips were used to
guarantee that the Mondrian would be physically realizable.
Reflectance and chromaticity coordinates for these Munsell stimuli
were taken from Table 6.9 of Wyszecki and Stiles (1967). For the
purpose of this study, chromaticity coordinates were converted to cone
outputs by assuming a tritanopic retina with a 580 nm crosspoint,
utilizing equation 3.3.6 in Wyszecki and Stiles (1967). The Mondrian
is displayed in Fig. 9 and Table 2, and a chromaticity chart of the
Mondrian is displayed in Fig. 10.

Cone outputs were used to construct R-G and R+G. Each of the
sixty-six transitions were convolved with the lowpass and bandpass
filters of Fig. 8. The output of the (R+G)(Bandpass) convolution and
the (R-G)(Lowpass) convolution were added together. By Russell's
Criterion fifty-nine of the sixty-six edges were enhanced. The Russell
filter was applied and the estimated R-G and R+G signals on each side
of the edge were recovered. As expected the R-G signal was easy to
recover, since the convolution profile does not contain an R+G signal
until the edge is crossed. R+G is more difficult. The signal can be
estimated by throwing away the DC level of the convolution profile
leaving only the enhancements. In practice this can be done by
convolving the output of the r-g cell with a laplacian, identical to
the bandpass filter of the (R+G)(Bandpass) component, but opposite in
sign. The output of the negative laplacian is an estimate of the R+G
signal across the edge. If it were proportional to the difference of
R+G across the edge, the edge could be reconstructed perfectly.
Unfortunately, this is not the case. This is due to a portion of the R+G bandpass signal being hidden by an inverse change in the R-G lowpass signal. Like an iceberg, much of the R+G signal is hidden. However, assuming that enhancement is proportional to luminance difference does lead to a fairly accurate recovery of the edge, with errors that range from less than one percent error to an occasional fifty percent error. Armed with these estimates one can reconstruct the Mondrian by going from block to block and adding or subtracting the luminance difference across the edge to the R+G value of the previous block. The results, displayed in Table 3, are a fairly good reconstruction of the original Mondrian. While luminance is not perfectly recovered, the correlation coefficient between objective and recovered luminance is in excess of 0.95. To improve the reconstruction one would need to use information about the lowpass drop, as we did in the first algorithm of this chapter.

APPLICATION TO MARR'S THEORY

As alluded to previously, there are a number of problems whose solution may lie in a successful demultiplexing scheme. One such problem is the confusion in the computational vision community over the role of the r-g cell. This problem is treated in detail by Ingling, Billock and Grigsby (1987). The MIT Artificial Intelligence Laboratory, and others, have constructed a web of ingenious algorithms for accomplishing a variety of human visual processing tasks. Almost
all of these processes use the zero crossings from the output of a receptive field that is shaped like the Laplacian of a Gaussian. In Marr's (1982) landmark study, Vision, the following statements can be found:

"According to physiological descriptions, some opponent-color cells in the retina of the monkey have receptive fields with rather mixed properties, like a red center and a green surround. There seems to be no internal reasons for doubting these reports; nevertheless, I find such cells extremely difficult to understand in general and impossible to fit into the $\nabla G$ framework...a cell with such a receptive field... signals a complex mixture of spatial and chromatic information...it is not like a second derivative and its zero crossings are meaningless..."

Marr goes on to detail how hue and luminance ought to be encoded:

"...lightness and brightness should be separated from color. Luminance boundaries correspond effectively to change in the contributions of the red and green channels, which we can write (R+G). To detect these boundaries requires a $\nabla G$ operator running on this sum...to detect changes in color, on the other hand...can be done by a $\nabla G$ operation on (R-G)..."

In other words, Marr requires the output of double opponent cells to do color processing and Type III cells to do luminance processing. Type I cells are a problem for the MIT paradigm, but could be ignored as a sort of visual appendix if they are rare and double opponent and Type III cells are plentiful. This situation does not obtain. As we have seen, Type I cells dominate central vision. Type III cells are more important outside the central 40 degrees, and double opponent cells are absent in the primate retina (De Monasterio, 1980).
Fortunately, the very information Marr needs is available from the Type I $r-g$ cell if a demultiplexing scheme is available. The $(R+G)$ bandpass component is very similar to the Type III receptive field (the only difference being that the surrounds of Type III cells extend through the centers), and the $(R-G)$ lowpass component provides a color signal, which can be used to construct double opponent cells. With a demultiplexing algorithm, Marr's algorithms can be confidently applied to human central vision. Without a demultiplexing algorithm, such applications are primarily of theoretical interest.

APPLICATION TO DOUBLE-OPPONENT CELLS

Another area where a good demultiplexing theory might explain much is the origin of some cortical receptive fields. Michael (1978a, 1978b, 1978c) reports a class of cortical cells that are double-opponent and better equipped to analyze simultaneous color contrast than the ordinary single opponent cells we have been analyzing. Michael makes a case for the formation of these cells from Type II and Type III LGN fibers, although, as we shall see, he is probably in error.

A typical concentric double-opponent receptive field has a center that is $R-G$ and a surround that is $-(R-G)$. Michael also finds a variety of double opponent simple and complex cells, built up from concentric double opponent cells. This is interesting, because regular simple and complex cells are built out of LGN cells. If, as Michael
suggests, double opponent cells are built from LGN cells, why aren't
double opponent simple cells built directly from their LGN precursors?
As Michael writes, "it is not clear why an additional step should be
necessary".

The second problem with the current concentric double opponent
cell model concerns their receptive field structure. Michael finds
that the receptive fields of concentric double opponent cells are
composed of subunits that are either Type II or Type III cells with
single cone spectral sensitivities. Now this is a very curious result.
Michael finds that about 80 percent of the LGN fibers projecting to
this region of striate cortex are Type I, but not one of these color
transmitting cells are used to produce cells used to detect
simultaneous color contrast. Instead, we find Type II inputs, despite
the rarity of these cells and a version of the Type III cell that has
never been reported in retina or LGN. Real Type III cells have R+G
spectral sensitivities. In every case Michael's proposed subunit, has
inputs only from R or only from G. The Type II report is odd too,
because most Type II cells are B-Y. In fairness to Michael, he
believes that something is amiss here, but does not offer an
explanation.

All of these problems can be resolved if Michael's cells are built
out of Type I r-g cells. The need to filter these cells to find their
constituents explains why the cells in the double opponent pathway
require an extra step not found in the usual geniculate, simple cell,
complex cell pathway. The output of any filtering operation that isolates the (R-G)(Lowpass Spatial) channel will look just like the output of a Type II cell. Even the discrepancy that Michael noted from Type II behavior is readily explainable. The cells contributing to the double opponent cell have no reaction to diffuse white fields, difficult for Michael to explain because Type II cells are known to be sensitive to such fields. The (R-G) lowpass spatial component of the r-g cell will have no sensitivity to diffuse white fields because its parent r-g cell has none. The R and G sensitive Type III inputs that Michael finds are harder to explain but the following argument can be made.

Recall the argument made in Chapter 2 on the spectral sensitivity of acuity. If the stimulus is too small for the surround to resolve, then the spectral sensitivity of the entire receptive field will be R or G, depending on which occupies the center. If one were searching for such a signal with a matched filter, would a lowpass or a bandpass filter be better? Recall that the cutoff for the surround is four times smaller than the center. Therefore, a signal that exceeds the surround cutoff is at an intermediate to high frequency. Recall too, that a bandpass filter whose surround does not extend through the center, has a gain for intermediate frequencies. It is clear that there will exist a range of frequencies for which the spectral sensitivity transmitted will be the sensitivity of the center and the filter characteristic for which will appear to be bandpass. It remains to be seen what in Michael's field could have triggered this response.
Perhaps we are seeing the higher harmonics at the edge of the two fields.

In summary, it would seem that the visual system has two different ways to wire a double opponent receptive field, each method arising out of a different component of the r-g cell response.
Response of the r-g cell for three different edges (edges expressed as R-cone and G-cone signals). Edge a has equal luminances on each side, therefore the R+G signal is zero. Edge b is edge a with achromatic contrast; the right side is decreased by 15%. This produces an R+G signal, but this signal is masked by the R-G signal and does not appear in the overall response of the cell. Edge c has its right side decreased by 30% - sufficient achromatic contrast for the overall signal of the cell to show enhancement. Of the three edges only edge c satisfies Russell's criterion, and only this edge can extracted by a Russell Filter. (Reproduced from Ingling and Martinez (1983a) with the permission of the C. R. Ingling Jr.)
<table>
<thead>
<tr>
<th>Block</th>
<th>Hue</th>
<th>V/C</th>
<th>Reflect</th>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.5R</td>
<td>5/8</td>
<td>.1977</td>
<td>.4252</td>
<td>.3101</td>
</tr>
<tr>
<td></td>
<td>10.0Y</td>
<td>3/2</td>
<td>.0656</td>
<td>.3513</td>
<td>.3789</td>
</tr>
<tr>
<td></td>
<td>5.0BG</td>
<td>7/2</td>
<td>.4306</td>
<td>.2898</td>
<td>.3225</td>
</tr>
<tr>
<td></td>
<td>7.5YR</td>
<td>6/8</td>
<td>.3005</td>
<td>.4596</td>
<td>.4064</td>
</tr>
<tr>
<td></td>
<td>7.5P</td>
<td>3/4</td>
<td>.0656</td>
<td>.3072</td>
<td>.2448</td>
</tr>
<tr>
<td></td>
<td>7.5R</td>
<td>3/10</td>
<td>.0656</td>
<td>.5730</td>
<td>.3240</td>
</tr>
<tr>
<td></td>
<td>2.5G</td>
<td>7/8</td>
<td>.4306</td>
<td>.2861</td>
<td>.4129</td>
</tr>
<tr>
<td></td>
<td>5.0BG</td>
<td>5/6</td>
<td>.1977</td>
<td>.2360</td>
<td>.2370</td>
</tr>
<tr>
<td></td>
<td>7.5R</td>
<td>5/2</td>
<td>.1977</td>
<td>.3425</td>
<td>.3229</td>
</tr>
<tr>
<td></td>
<td>5.0RP</td>
<td>8/2</td>
<td>.5910</td>
<td>.3180</td>
<td>.3120</td>
</tr>
<tr>
<td></td>
<td>5.0G</td>
<td>5/2</td>
<td>.1977</td>
<td>.2978</td>
<td>.3392</td>
</tr>
<tr>
<td></td>
<td>10.0YR</td>
<td>6/4</td>
<td>.3005</td>
<td>.3861</td>
<td>.3767</td>
</tr>
<tr>
<td></td>
<td>7.5PB</td>
<td>7/6</td>
<td>.4306</td>
<td>.2687</td>
<td>.2612</td>
</tr>
<tr>
<td></td>
<td>7.5GY</td>
<td>8/10</td>
<td>.5910</td>
<td>.3463</td>
<td>.4791</td>
</tr>
<tr>
<td></td>
<td>2.5RP</td>
<td>7/6</td>
<td>.4306</td>
<td>.3338</td>
<td>.2854</td>
</tr>
<tr>
<td></td>
<td>2.5PB</td>
<td>2/4</td>
<td>.0313</td>
<td>.2175</td>
<td>.2245</td>
</tr>
<tr>
<td></td>
<td>7.5G</td>
<td>7/2</td>
<td>.4306</td>
<td>.2972</td>
<td>.3333</td>
</tr>
<tr>
<td></td>
<td>2.5RP</td>
<td>3/8</td>
<td>.0656</td>
<td>.3598</td>
<td>.2233</td>
</tr>
<tr>
<td></td>
<td>2.5R</td>
<td>6/2</td>
<td>.3005</td>
<td>.3318</td>
<td>.3166</td>
</tr>
<tr>
<td></td>
<td>10.0RP</td>
<td>2/6</td>
<td>.0313</td>
<td>.4139</td>
<td>.2608</td>
</tr>
<tr>
<td></td>
<td>10.0B</td>
<td>4/4</td>
<td>.1200</td>
<td>.2429</td>
<td>.2648</td>
</tr>
<tr>
<td></td>
<td>7.5p</td>
<td>7/4</td>
<td>.4306</td>
<td>.3111</td>
<td>.2880</td>
</tr>
<tr>
<td></td>
<td>7.5R</td>
<td>2/2</td>
<td>.0313</td>
<td>.3751</td>
<td>.3181</td>
</tr>
<tr>
<td></td>
<td>2.5YR</td>
<td>5/4</td>
<td>.1977</td>
<td>.3925</td>
<td>.3494</td>
</tr>
<tr>
<td></td>
<td>5.0YR</td>
<td>7/6</td>
<td>.4306</td>
<td>.4091</td>
<td>.3701</td>
</tr>
<tr>
<td></td>
<td>7.5P</td>
<td>7/2</td>
<td>.4306</td>
<td>.3109</td>
<td>.3037</td>
</tr>
<tr>
<td></td>
<td>10.0Y</td>
<td>2/2</td>
<td>.0313</td>
<td>.3556</td>
<td>.3848</td>
</tr>
<tr>
<td></td>
<td>10.0PB</td>
<td>4/6</td>
<td>.1200</td>
<td>.2618</td>
<td>.2263</td>
</tr>
<tr>
<td></td>
<td>2.5GY</td>
<td>6/4</td>
<td>.3005</td>
<td>.3572</td>
<td>.4038</td>
</tr>
</tbody>
</table>
Spatial arrangement of the Mondrian used in this chapter. This Mondrian contains 66 edges (including corners) of which 59 are extractable by Russel Filtering. (Diagram created by S.S. Grigsby (private communication) and is reproduced by permission of the author.)
Chromaticity diagram shows the Chromaticities of the Mondrian studied in this chapter. This Mondrian contains 66 edges (Diagram created by S.S. Grigsby (private communication) and is reproduced by permission of the author).
<table>
<thead>
<tr>
<th>Block #</th>
<th>(R-G)</th>
<th>(R+G)</th>
<th>(R+G) Real</th>
<th>(R+G) Recovered</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.0103</td>
<td>.1977</td>
<td>.1977</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.0066</td>
<td>.0656</td>
<td>.0924</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-.0620</td>
<td>.4306</td>
<td>.4767</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.0013</td>
<td>.3005</td>
<td>.2999</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-.0044</td>
<td>.0656</td>
<td>.0904</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.0176</td>
<td>.0656</td>
<td>.0904</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-.0768</td>
<td>.4306</td>
<td>.4903</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-.0415</td>
<td>.1977</td>
<td>.2451</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-.0151</td>
<td>.1977</td>
<td>.2209</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-.0602</td>
<td>.5910</td>
<td>.6218</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>-.0279</td>
<td>.1977</td>
<td>.2326</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>-.0177</td>
<td>.3005</td>
<td>.3173</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>-.0612</td>
<td>.4306</td>
<td>.4760</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>-.0848</td>
<td>.5910</td>
<td>.6442</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>-.0258</td>
<td>.4306</td>
<td>.4436</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>-.0065</td>
<td>.0313</td>
<td>.0609</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>-.0602</td>
<td>.4306</td>
<td>.4751</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>.0042</td>
<td>.0656</td>
<td>.0825</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>-.0259</td>
<td>.3005</td>
<td>.3248</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>.0033</td>
<td>.0313</td>
<td>.0520</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>-.0220</td>
<td>.1200</td>
<td>.1562</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>-.0414</td>
<td>.4306</td>
<td>.4580</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>-.0009</td>
<td>.0313</td>
<td>.0558</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>-.0061</td>
<td>.1977</td>
<td>.2127</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>-.0118</td>
<td>.4306</td>
<td>.4579</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>-.0458</td>
<td>.4306</td>
<td>.4619</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>-.0031</td>
<td>.0313</td>
<td>.0578</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>-.0158</td>
<td>.1200</td>
<td>.1505</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>-.0317</td>
<td>.3005</td>
<td>.3301</td>
<td></td>
</tr>
</tbody>
</table>


