

194
56

Mach Bands, Hermann Grid, and Lateral Inhibition in the Retina

by

Nuri Aktuna

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

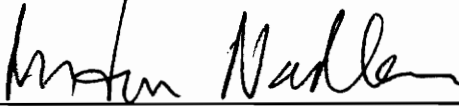
in partial fulfillment of the requirements for the degree of

Master of Science

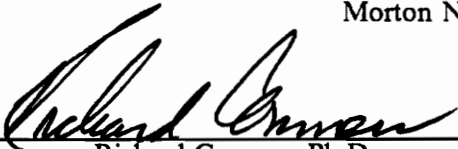
in

Electrical Engineering

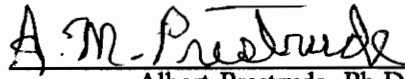
APPROVED:



Morton Nadler, Ph.D., Chairman



Richard Conners, Ph.D.



Albert Prestrude, Ph.D.

August, 1990

Blacksburg, Virginia

LD

5655

V855

1990

A368

C.2

Mach Bands, Hermann Grid, and Lateral Inhibition in the Retina

by

Nuri Aktuna

Morton Nadler, Ph.D., Chairman

Electrical Engineering

(ABSTRACT)

Literature about two illusions that are traditionally linked to lateral inhibitory processes in the retina is surveyed. Despite much research and a long history, the Mach bands illusion has not yet been very well-understood. Quantitative data is lacking, while qualitative data is controversial. The Hermann Grid illusion, equally old, has not yet been fully explored either. A critical and a comparative review of the experiments and the conclusions is carried out, noting common results as well as contradictory ones. New experiments and a different method of measurement are then proposed. Some preliminary experiments and their results are included in the appendix together with a review of recent knowledge about the primate retina and its functioning.

Acknowledgements

I would foremost like to acknowledge the support of my committee chairman Dr. Morton Nadler, and my committee members, Drs. Richard Conners, and Albert Prestrude, without whom this study would not be possible. I would also like to extend my thanks to the Turkish Education Foundation which financed my education in United States. I am further grateful to Nikos Asimopoulos of Image Processing Technologies for providing the equipment for the preliminary experiments, and for his generous support. Special thanks go to my roommates Marwan Bikdash and Cihan Tuncbilek, as well as always helpful friend, Emin Demirbag.

Table of Contents

- 1. Introduction 1
- 2. Mach Bands 3
 - 2.1 Historical Overview 3
 - 2.2 Properties of Mach Bands 6
 - 2.2.1 Width 6
 - 2.2.2 Location 8
 - 2.2.3 Magnitude 8
 - 2.2.4 Thresholds 9
 - 2.2.5 Flicker and Short Exposure Durations 10
 - 2.2.6 Edges 11
 - 2.2.7 Nonlinearity 12
 - 2.3 Critique 12
 - 2.3.1 Achromatic Mach Band Research 13
 - 2.3.2 Chromatic Mach Band Research 15
 - 2.4 Edge Detection Theories and Mach Bands 17

3. Hermann Grid Illusion	20
3.1 Physiological locus and global factors	20
3.2 Chromatic effects	22
3.3 Perceptive field sizes	27
4. Proposal	29
4.1 Raster Display Systems	29
4.2 Research Program	32
4.3 Compensation	35
4.4 Achromatic Mach Bands	39
4.5 Chromatic Mach Bands	43
4.6 Incremental Detection Thresholds	45
4.7 Hermann Grid	46
5. Conclusion	48
6. Figures	50
7. Bibliography	89
Appendix 1. Physiological Summary	98
A.1.1 The structure of the primate retina	98
A.1.2 Functional physiology	103
A.1.3 Implications for retinal processing	104
Appendix 2. Programs	107
A.2.1 Program Mach-l1	107
A.2.2 Program Mach-e1	117

A.2.3 Program Mach-l2 127

A.2.4 Program Mach-e2 137

A.2.5 Program Herl 147

Appendix 3. Preliminary Experiments and Results 156

A.3.1 Apparatus 156

A.3.2 Programs 157

A.3.3 Limitations 160

A.3.4 Experiments 162

A.3.5 Projection 164

Vita 165

List of Illustrations

Figure 1. Fourier series expansions	51
Figure 2. Even and odd-symmetric receptive fields	52
Figure 3. Mach Band generating retinal light distributions	53
Figure 4. The discs used by von Bekesy [1972]	54
Figure 5. Stimuli used by Fiorentini and Radici [1957]	55
Figure 6. The five contrast patterns used by Ware and Cowan [1987]	56
Figure 7. The Hermann grid. Small grey dots appear at the intersections.	57
Figure 8. Stimulation of ganglion cells by a chromatic Hermann grid	58
Figure 9. Stimulation of double opponent color cells by a Hermann grid	59
Figure 10. The modified Hermann Grid used by Levine, et al. [1980]	60
Figure 11. Intersection in a hollow Hermann Grid used by Troscianko [1982b]	61
Figure 12. A schematic of proposed experimental compensation stimuli	62
Figure 13. The organization and layers of the primate retina [Rodieck 1988]	63
Figure 14. The cone pedicle and its synaptic connections	64
Figure 15. The four pathways from a cone	65
Figure 16. Receptive field structures of midget ganglion cells	66
Figure 17. Receptive field structures of diffuse ganglion cells	67
Figure 18. On and off-center pathways from a cone	68
Figure 19. Looking at Mach Band formation	69
Figure 20. Double opponent cells in the visual cortex	70
Figure 21. Recursive and non-recursive lateral inhibition	71

Figure 22. Types of lateral inhibition in the eye 72

Figure 23. Slope in the linear and exponential gradients 73

Figure 24. Mach Band compensation using program Mach-11 74

Figure 25. Mach Band compensation using program Mach-12 75

Figure 26. Hermann Grid compensation using program Her1 76

Figure 27. Intensity and compensation profiles - 1 77

Figure 28. Intensity and compensation profiles - 2 78

Figure 29. Intensity and compensation profiles - 3 79

Figure 30. Intensity and compensation profiles - 4 80

Figure 31. Intensity and compensation profiles - 5 81

Figure 32. Intensity and compensation profiles - 6 82

Figure 33. Intensity and compensation profiles - 7 83

Figure 34. Intensity and compensation profiles - 8 84

Figure 35. Intensity and compensation profiles - 9 85

Figure 36. Intensity and compensation profiles - 10 86

Figure 37. Intensity and compensation profiles - 11 87

Figure 38. Intensity and compensation profiles - 12 88

1. Introduction

The complexity and the capabilities of the human visual system have been baffling scientists for a long time. It has been nearly a century now that noted Spanish neurobiologist Ramon y Cajal first took a clear glimpse into the retina and wonderfully depicted it. In spite of important strides in between, however, the retina still remains for the neuroscientist a place of mysteries, this time with perhaps more familiar landmarks. For instance, we have neither a universally accepted classification of ganglion cells, nor an adequate understanding of amacrine cells and their function. In an analogous fashion, although the great debate between the color theories of Young and Helmholtz is now resolved, psychophysical research too has discovered the unending layers of complexity surrounding biological sensory mechanisms. Color constancy is a classic example.

Quite early in vision research, in an unusual instance, psychophysics and neurobiological investigation came together when one psychophysical phenomenon, the Mach Bands illusion, was linked to lateral inhibition in the retinal network, and was later proved in the lateral eye of the horseshoe crab *Limulus*. A number of models, and a series of psychophysical experiments have followed. The results, however, have not been satisfactory or conclusive. This study attempts to critically review these experiments, point out the shortcomings and the contradictions, and, finally, to propose a new set of experiments with currently available technology. In addition, a related but relatively unexplored psychophysical phenomenon, the Hermann Grid illusion, is also being investigated. Al-

though the emphasis of this work is psychophysical, a physiological summary of the retina is included to provide a sound biological perspective. Finally, a series of preliminary experiments designed to test the feasibility of some of the proposed ideas and the results of these experiments are discussed.

It is the conviction of the author that Mach Bands and related phenomena are important windows into the functioning of the human visual system and are not merely enjoyable illusions. We hope that this study will act as a reference point for future research into the subject.

2. Mach Bands

2.1 Historical Overview

Since its discovery by the famous German physicist Ernst Mach in 1865 [Mach 1865], the Mach bands illusion has drawn the interest of many scientists partly because it is a very powerful and enduring illusion, and partly because it has been seen as a gateway for understanding the processing of visual information. There are also immediate practical applications for printing, Roentgen photography, and microscopy where the illusion commonly occurs. Mach bands illusion has also been demonstrated in other senses [von Bekesy 1967].

During the first half of this century there were relatively few and isolated reports on the psychophysics of Mach bands [Fry 1948; Koffka & Harrower 1931; McDougall 1903; Thouless 1922]; this changed with a series of papers published between 1955-1972 [Ercoles-Guzzoni & Fiorentini 1958; Ercoles & Fiorentini 1959; Fiorentini 1957,1972; Fiorentini et al. 1955; Fiorentini & Radici 1957, 1958; Lowry & de Palma 1961; Matthews 1966,1967; McCollough 1955; Thomas 1965,1966; von Bekesy 1967,1968a,1968b,1968c,1972; Watrasiewicz 1966; Whiteside & Davidson 1971]. A large number of these papers signified a concerted effort, first by a group of Italian sci-

entists and then by von Bekesy, to determine the properties of the Mach bands and to obtain quantitative data on their behaviour. The papers published during this period still form the bulk of our psychophysical data about achromatic Mach bands.

At about the same time, theoretical interest in Mach bands culminated in several quantitative models, elegantly described by Ratliff [1965] in his comprehensive book. Most of these models are very similar and are based on lateral inhibition in the retina. With the advent of single-cell microelectrode neurophysiology, Hartline and Ratliff were able to confirm and measure not only the extent of lateral inhibition, but also the empirical Mach bands illusion in the composite eye of the horseshoe crab, *Limulus* [Hartline & Ratliff 1954,1957; Hartline, et al. 1956,1961; Ratliff & Hartline 1959]. Their results, therefore, have given rise to a more concrete model that has found greater universal acceptance, and have established the concept of lateral inhibition in early stages as the dominant explanation of Mach bands.

The engineering community, in its attempt to understand the human visual system and to exploit its properties for image processing and transmission, has also been interested in the Mach bands illusion [Granrath 1981; Hall & Hall 1977; Schreiber 1978; Stockham 1972]. Citing Hartline's equations, Stockham [1972] has seen the illusion as a manifestation of linear filtering and has included a dramatic example of the usefulness of the underlying mechanism in everyday vision, while wisely pointing out the inaccuracy of earlier experiments and suggesting an adjustment method for newer measurements.

Late fifties and sixties witnessed the application of linear systems theory and Fourier Analysis, as a tool if not necessarily as a model, in understanding the processing of spatial visual information. Over the years, this approach influenced much of visual psychophysics and led Morrone, et al. [1986] to publish an alternative explanation of the Mach bands illusion. Morrone, et al. asserted that Mach bands were phase-dependent and that the out-of-phase Fourier series components in linear illumination gradients were the cause of the illusion (Fig. 1). In light of relevant work in the area [Bennett & Banks 1987; Field & Nachmias 1984], it is clear that lateral inhibition is necessary

to form the odd and even-symmetric receptive fields that are essential for phase detection (Fig. 2). Mach bands, however, are not thought to be a direct and immediate result of lateral inhibition in the retina, as Hartline and Ratliff's models have proposed, with minimal processing in the brain; instead, they are the result of a complicated cortical mechanism that extracts and processes phase information.

Another explanation for Mach bands was offered recently by Watt and Morgan [1983] in their article exploring the spatial primitives in vision. They modelled the response of the visual system at very early stages (most probably at the retina) by a second derivative operation (Laplacian), which resulted in a high-pass filter. The response to a linear illumination gradient revealed a 0-valued inflection point-neighbourhood of substantial extent which was singled out as the cause of the Mach bands illusion (Fig. 3). Although lateral inhibition is certainly the high-pass filter mentioned, the detection mechanism in this model must be in higher centers since it depends on the extraction of a spatial primitive: a 0-valued inflection point region. If the region is below some critical length, the illusion is not seen.

After the achromatic domain, it was natural for scientists to search for Mach bands in the chromatic domain, but the evidence so far is inconclusive. An important reason for this lack of agreement is the complexity of the perception of color: The percepts of hue, saturation and brightness can only roughly be correlated with wavelength, purity and intensity. It is therefore very difficult to control one sensation without simultaneously affecting the others. Another is semantic: "Color" itself is an ambiguous term: in the vision literature it is used at times synonymously with hue and at other times to refer to the whole perceptual experience, including brightness, hue and saturation" [Pease 1978]. Chromatic Mach bands were recently investigated only by Ware and Cowan [1987], who found evidence for their existence.

It is clear from the preceding overview that, in spite of the numerous contributions to the expanding knowledge base, our knowledge about Mach bands is still very much incomplete. In fact, as we shall

see in many cases, it is contradictory. Nor are we sure of the underlying mechanisms. Below, a summary of the properties of the Mach bands and a critical review of past research are presented.

2.2 Properties of Mach Bands

The behaviour of the bright and dark Mach bands will be considered separately under each property heading; this dichotomy is pursued because, in many instances, there are differences in the behaviour of the bands. Furthermore, many articles have been devoted to the investigation of only one Mach band, and many authors have stated that each band is of a different nature, refraining, most of the time, however, from specifying or speculating exactly where the difference lies.

Since even the existence of chromatic Mach bands is controversial, and research so far is ambiguous, chromatic Mach bands are not discussed here.

In the following summary, the gradient or the gradient field refers to the area of transition between the dark and bright uniform fields (this is not necessarily a ramp) and slope refers to the slope of the gradient field.

2.2.1 Width

The width of Mach bands varies from 1 minute to 10-12 minutes of arc [Fiorentini 1972].

The bright band is perceived to be narrower than the dark band [Fiorentini & Radici 1958; Whiteside & Davidson 1971].

Most studies have shown that as the slope of the gradient increases, the width of the bright band decreases [Lowry & de Palma 1961; McCollough 1955; von Bekeſy 1972]. Von Bekeſy's and Lowry and de Palma's observations are qualitative while McCollough uses, as the measure of slope, the logarithm of maximum luminance of the nonlinear intensity distribution used. This, admittedly, is very difficult to interpret without reference to the article. (In the experiment, different slopes are created by increasing the intensity at each point proportional to the already existing intensity)

McCollough [1955] has found that the width of the bright band also decreases with increasing luminance, but Watrasiewicz [1966] has reported the opposite: Up to 1000 trolands, the width of the bright band increases with increasing luminance, and there is no indication of a consistent decrease afterwards.

The decrease in width with increasing slope holds for the dark band, too [Lowry & de Palma; von Bekeſy 1972]. Watrasiewicz [1966] has indicated that with increasing log of retinal illuminance, the width of the dark band decreases linearly.

Finally, we would like to mention an interesting observation by von Bekeſy [1972]: the width of the Mach bands increases linearly as the viewing distance is increased. To explain the phenomenon, von Bekeſy notes that the angle of the ramp pattern (angle x in fig. 4) stays the same in all viewing distances and forms the basis for the perception of Mach bands. Thus, at a larger distance, although the size of the stimulus is reduced, the same angle is perceived and the same Mach bands are created, but this time over a larger part of the shrunk stimulus. Convinced that Mach bands are primarily produced on the retina, von Bekeſy likens this to Emmert's law and predicts an Emmert's law for the Mach bands themselves. (Emmert's law states that the size of an image depends on the focusing distance; it is blown up if the eyes are focused to a distance greater than the original focusing distance, and is shrunk if the eyes are focused to a closer distance)

Von Bekeſy's justification, however, is vague. What the retina perceives when the disc is revolved is an intensity gradient, and the only clue to the angle is the slope, which depends on the extent of

the gradient in minutes of arc. At a greater distance, this extent is reduced, and the slope is increased. According to previous experiments thinner Mach bands should be produced. Due to the shrinkage of the stimulus, however, even these thinner Mach bands may be perceived to occupy a larger area of the stimulus pattern.

2.2.2 Location

The bright band is perceived to be nearer to the gradient than the dark band [Fiorentini & Radici 1958] and for some observers, shifts toward the gradient field as slope or luminance increases [McCollough 1955].

The greater portion of the dark band, on the other hand, lies in the dark uniform field [Fiorentini & Radici 1958].

2.2.3 Magnitude

It is now generally agreed that the bright band is superior to the dark band in magnitude [Fiorentini & Radici 1957; von Békésy 1968a]. Also, in studies with exponential gradients, that were designed to cancel out the nonlinear transduction of light intensity in the retina to create the most symmetric bands, perfect symmetry was never reported [Whiteside & Davidson 1971].

It is not certain whether an increase in slope increases the brightness of the bright band [Fiorentini & Radici 1958; Lowry & de Palma 1961], but there is some evidence, though not worded quite clearly, that it reduces it [Thomas 1966].

The dark bands present us with the same uncertainty. While some researchers argue that as the slope of the gradient is increased, the apparent brightness of the dark band decreases (the dark band becomes darker) [Fiorentini & Radici 1958; Thomas 1966], results of other researchers show that

there is almost no difference [Lowry & de Palma 1961], and still some others describe an opposite effect [Watrasiwicz 1966].

2.2.4 Thresholds

Thresholds are very dependent on how they are measured and researchers of Mach bands have employed different criteria. There are two kinds of thresholds: detection thresholds and discrimination (incremental) thresholds.

With respect to the behaviour of the detection threshold of the bright band with luminance, Ercoles and Fiorentini [1959] have reached the conclusion that, as luminance is increased, the detection threshold, which is the slope/luminance necessary to see the bright Mach band, decreases and later reaches a more or less constant value. Thus a shallower slope is sufficient to detect the bright band with increasing luminosity. The same research has found that, when both bands can be detected, the detection threshold for the bright band is lower than that of the dark band, in fact, by a ratio of 1:4. At low luminance values, however, only the dark band is seen; the bright band begins to be seen at higher minimum values. (The results of this experiment should be approached with some reservation. The detection threshold is in units of slope/luminance. Slope is defined as the luminance difference between the uniform regions/maximum width of the gradient where the bright band is first visible. The luminance used in the threshold unit, however, varies. It is the luminance of the bright uniform area in case of the bright Mach band, and the luminance of the dark uniform area in case of the dark Mach band. This is difficult to interpret and since they are not tested against a common variable, it is doubtful whether the results of the bright and dark Mach bands can be compared.)

Thomas [1965], using luminance difference/slope as the dependent variable (threshold criteria), has reported, as the result of one experiment, that the detection threshold for the bright band seems to be a function of both luminance and slope, but obtained ambiguous results in another experiment.

In the same paper he indicated that data suggested the independence of the dark and bright band thresholds.

Von Bekesy [1968a] measured the discrimination thresholds across the whole subjective distribution, plotting a graph of ΔI vs. location. Interestingly, there is an overshoot at the location of the bright band. His results were supported by Matthews [1966] who plotted a graph of $\log \Delta I$ vs. location.

Ercoles and Fiorentini's results [1959] about the behaviour of the detection threshold of the bright Mach band with luminance, summarized above, also hold for the dark band.

Thomas [1963] found that dark bands had lower detection thresholds than bright bands, thus contradicting Ercoles and Fiorentini's results.

Von Bekesy [1968a], on the other hand, noted that although the band itself is clearly visible, no increase in incremental thresholds in the location of the dark Mach band is seen.

2.2.5 Flicker and Short Exposure Durations

Flicker elicits different responses from the bright and the dark Mach bands, suggesting different kinds of inhibition [von Bekesy 1968b]. Mach bands are enhanced with large grain random noise; this is linked to the flicker-like quality of large-grain random noise [von Bekesy 1968b].

2 millisecond exposures are sufficient for the detection of the bright band, ruling out the suggestion that eye movements are essential for the detection of Mach bands [Matthews 1966]. At brief exposures, however, the incremental threshold, which rises at the point where the bright Mach band is located, does not exhibit this behaviour, although the bright band itself is visible [Matthews 1966].

2.2.6 Edges

Since the mechanism causing the Mach band illusion is very probably involved in extracting edge information, the results of some Mach band experiments that also explored the detection of the edges, are reviewed below. This is a tricky area; as will be seen, there is no consensus on what to measure and what it means.

Watrasiewicz [1966], defining the visual edge "as the borderline between the bright and dark field", has indicated that as retinal illumination increases, the visual edge shifts toward the bright side of the geometric edge, but shows the opposite behaviour as the wavelength of the monochromatic illumination gradient is increased.

Whiteside and Davidson [1971] have asked observers to relate the brightness of the area between the Mach bands to either the dark uniform region or the bright uniform region. When the gradient was changing from linear to exponential, "at or slightly below" the ideal exponential stimulus (which compensates the logarithmic transduction in the photoreceptors), the brightness of the intermediate area - we might just as well call it the subjective edge - changed from light to dark. When, instead of the exponential, power series stimuli were presented, unless the power series function resembled the exponential very closely, the area between the bands was always judged bright.

An important finding comes from Watt and Morgan [1983] who argue that the edges are subjectively perceived to be on the dark side of the geometric edge.

These experiments suggest that the subjective edge is quite different from the geometric edge.

2.2.7 Nonlinearity

It is known that the transduction of light intensity in the human eye is highly nonlinear. In an experiment designed to measure the nonlinearity, von Békésy [1968a] found that when luminance was increased (with the ratio of the luminance step kept constant), the magnitude of the bright band increased more slowly than the overall increase in luminance, implying that the nonlinearity is before the lateral inhibition.

In their experiment to determine the nature of the nonlinearity, Whiteside and Davidson [1971] concluded that since the bands in patterns that were closest to exponential gradients were most symmetrical, there was a logarithmic nonlinearity before lateral inhibition.

2.3 Critique

Like most psychophysical phenomena, Mach bands have proved to be elusive and uncooperative partners in experimentation. From a glimpse at pioneering articles (for example [von Békésy 1968a]), one cannot but feel admiration for the innovative techniques and the meticulous work that have characterized the measurements of the various properties of the illusion. Like all scientific work, however, previous experiments are not impervious to close scrutiny and criticism. Below, some common deficiencies in Mach band research are summarized and deviations from the ideal are noted.

2.3.1 Achromatic Mach Band Research

1) In many experiments, the designated independent variables (such as luminance or slope, which both influence the bands) are not controlled independently, but manipulated together, making it impossible to isolate the underlying cause in the behaviour of the dependent variable (such as brightness or width) and to precisely materialize the relationship [McCollough 1955; Thomas 1965,1966].

2) Ratliff et al.'s work [1983] on the susceptibility of the illusion to the disruptive effects of adjacent stimuli suggest that the comparison stimuli in brightness matching experiments (and even the indicator lines in width determination experiments) have a disruptive effect on the test pattern and thus on the Mach bands themselves. This is admitted by some of the researchers. [Lowry & de Palma 1961; McCollough 1955; Thomas 1965,1966]. In fact, this is sometimes so severe that comparison stimuli have to be relocated [Thomas 1966].

3) Placing the comparison stimuli to one side of the test stimulus to prevent the disruption of the Mach bands [Thomas 1966], has its own problems. Basically the observer has to remember the brightness of the bands and the comparison stimulus while his eyes are racing back and forth, making the method inappropriate for fine measurements.

4) Some of the experimenters have used binocular matching to alleviate the above-mentioned problem but not without some drawbacks of its own [Fiorentini & Radici 1957]. In their words: "...binocular match requires fixation of the eye and extrafoveal vision of the field. One might expect that if the measurements were performed by a method allowing central vision and scanning eye movements, the result would be slightly different."

5) In some brightness matching experiments [Fiorentini & Radici 1957; Thomas 1966], the brightness of the dark band cannot be duplicated without a bright inducing field around the comparison stimulus, which, through the mechanism of simultaneous brightness contrast, lowers the brightness

of the comparison stimulus (Fig. 5). As admitted by Fiorentini and Radici, this lends indirectness and arbitrariness to the measurement of the dark field.

6) As a comparison stimulus, von Bekesy [1968a] has used a thin marker that flickered with the portion of the Mach band whose brightness was measured. While other parts of the field were flickering, "the luminance of the marker was adjusted so that no flicker could be observed between the Mach band and the red slit." Equal brightness was thus obtained.

Although the method and the technique is ingenious and no doubt was implemented with painstaking care and precision, the use of a different colored comparison slit lends some arbitrariness to the results. We now know that different colors have different brightness percepts, therefore, the equivalence of the brightness of two patches of red and green does not necessarily indicate that they have the same luminance, making comparable objective measurements impossible. Furthermore, successive color contrast effects are another possibility.

7) In Thomas' experiments [1965], the stimuli were presented for periods as brief as 1 second. As pointed out by Watrasiewicz [1966], the transient response of the Mach bands illusion has not yet been explored in depth and some results on incremental thresholds [Matthews 1966] suggest that 1 second may not be sufficient for very fine measurements. In fact, Thomas himself remarks that "present information about the effects of brief exposure is not entirely consistent."

8) Results from many experiments have not been entirely satisfactory. Low spatial sampling, which hinders an accurate depiction of the bands, has sometimes been a problem [Fiorentini & Radici 1957], and sometimes the amount or significance of the data presented is not commensurate with the difficulty and the creativity of the method [von Bekesy 1968a,1972].

A major difficulty arises in the extent of interpersonal variations: All Mach band data are characterized by extremely large interpersonal variations; this lends ambiguity to the results and makes it impossible to draw concrete conclusions. Yet almost all research on Mach bands has employed

a handful of observers. Although this might be a characteristic of much of psychophysical research, in this case where the extent of interpersonal variations is unusually large, the statistical significance of the results is seriously weakened. Quantitative data suffer most; certainly that is one of the reasons why no single model of lateral inhibition in the human retina has emerged as a clear winner, and in popular models such as von Békésy's [Ratliff 1965], why no parameters have been agreed upon.

At times the presentation of data has not been very informative. An example is McCollough's paper [1955]. Although, ultimately, relevant data can be extracted, it is not readily available from the graphs or from the text, which defeats the purpose of the publication as a tool of communication.

9) Some topics in Mach band research are noticeably controversial. For example, the existence of Mach bands in sudden changes of illumination (square gradients) has ardently been debated for some time. The scientific community has been split: On one side are supporters of the existence of the illusion in sharp edges [Matthews 1966; von Békésy 1968a; Watrasiewicz 1966] and on the other side are critics [Fiorentini & Radici 1957; Ratliff 1984; Thomas 1965]. Independent research from both sides makes a comparative review and criticism extremely difficult. For instance, to what extent are von Békésy's results (who is an adherent to the view that the bands exist in sharp edges) comparable to those of Thomas or Fiorentini (who carried out their research on ramp gradients) ?

2.3.2 Chromatic Mach Band Research

We have not attempted a critical review of research on chromatic Mach bands in this proposal because there already exists an exceptional one [Pease 1978]. In his review of the field, Pease not only summarizes the previous experiments, providing us with their stimulus configurations, but also points out the ambiguity in the definition of chromaticity. He then roughly divides the color Mach band research into categories of brightness Mach bands, saturation Mach bands, and hue Mach bands. Stating that there is (as of then) no evidence of hue Mach bands, he proposes subjective

definitions for hue, brightness and saturation which could equip the researchers with a guideline "for describing the Mach effect in color domain."

After a period of dormancy following Pease's review, Ware and Cowan [1987] have very recently investigated chromatic Mach bands. They used five contrast patterns (including a ramp - the classical Mach band stimulus) in both chromatic and achromatic versions, and concluded that all patterns showed contrast effects, albeit with great differences in strength (Fig. 6). While they reported that there was "no evidence that the Mach band gradients were qualitatively different", they also noted that "classical Mach band gradient gave the smallest effect for both chromatic and achromatic versions."

Ware and Cowan performed two other experiments, one to alleviate the achromatic channel in measuring chromatic contrast and thus to ascertain that the chromatic Mach band effect was indeed due to chromatic components, and the other to alleviate the suggestion that small eye movements might cause the chromatic Mach bands.

In the first experiment the chromatic Mach bands remained even when all achromatic components were nulled. In the second experiment eye movement effects were prevented by intermittent viewing but the chromatic Mach bands did not disappear.

Ware and Cowan later suggested that these Mach bands could be due to lateral inhibition and based their model on the concentric color-sensitive double-opponent cells in the monkey striate cortex.

In achromatic vision simultaneous brightness contrast uniformly extends for large areas and is differentiated from the Mach bands illusion which is much more local. Ware and Cowan, however, used simultaneous brightness and color contrast stimuli together with Mach band stimuli, and although they did not attribute simultaneous color contrast to the Mach bands illusion, they considered and interpreted the results as a whole. Their assertion was that the same kind of mechanism mediated all the contrast effects (except the ambiguous Chevreul illusion pattern), and that the

mechanism was incorporated separately in the red-green, yellow-blue and achromatic channels. It is not certain, if not unlikely, that the same lateral inhibition mechanism accounts for all the contrast phenomena. In fact, Chevreul illusion, which remained inexplicable in this experiment, provides some evidence for the existence of different types of lateral inhibition. Von Békésy [1968b], in a paper devoted to Mach bands and Chevreul illusion, has found substantial differences between the type of inhibition causing the Mach bands and the one causing the Chevreul illusion (which he christened as Hering-type lateral inhibition).

Ware and Cowan's reopening of chromatic Mach band research is definitely encouraging. There remains, however, much to be done; above all, both qualitative confirmation and quantitative documentation are necessary.

2.4 Edge Detection Theories and Mach Bands

It is believed by some researchers that the purpose of early visual processing is to extract primitive features (or spatial primitives) that can be used to construct a compact, yet a very informative representation of the image upon which higher level mechanisms can operate [Marr and Hildreth 1980; Watt and Morgan 1983]. Marr, who contributed prolifically to the field of computational neuroscience [Marr 1976,1982; Marr and Hildreth 1980; Marr and Poggio 1979; Marr, et al. 1979], based his theories of early visual information processing on the detection of zero-crossings in the second derivative of the intensity distribution and their amplitudes (defined as the slope of the directional derivative taken perpendicular to the segment of zero-crossings). According to the theory of edge detection proposed by Marr and Hildreth [1980], the image is initially filtered by different-sized optimally smoothing Gaussian filters, and at each fineness scale thus established, the Laplacian operator is applied to obtain the zero-crossings information which is then relayed to higher centers. Watt and Morgan [1983] broaden the category of spatial primitives to include the

maxima and minima of the processed distribution and remark that the length of the zero-crossings, being the crucial element for the detection of Mach bands in their theory, is also noticed by the visual system. Not all researchers share this point of view about early visual processing. Srinivasan, et al. [1982] argue that the major duty of early visual processing is not to extract some features or reconstruct the scene but to encode the signals as efficiently as possible so as to protect them from intrinsic neural noise. They point out, however, that their model, predictive encoding, and edge detection as modeled by Marr and Hildreth, are not contradictory or mutually exclusive descriptions of the function of lateral inhibition, but complementary ones.

Whatever the purpose and organization of early visual processing in a computational hierarchy are, the detection of edges, as suggested by much evidence, is crucial for visual perception and is accomplished efficiently. While there is no agreed-upon theory of edge detection, it is possible to note two major guidelines in the past, which will also shape future research. One is lateral inhibition and its biology. No computational theory can simply ignore the biological knowledge base accumulated over the last hundred years, and should, at its best, try to account for such information as the inhibitory circuits in the retina, the scatter of receptive fields of ganglion cells, or the X and Y channels. The second guideline is psychophysical research. For theories of human visual perception this is even more important as *in vivo* studies cannot be performed in humans. Marr and Hildreth's theory is a good example. It was through psychophysical research that the existence of parallel channels in vision, which forms the backbone of the theory, was demonstrated [Wilson and Bergen 1979; Wilson and Giese 1977]. And it is precisely such research about Mach Bands that must be accounted for or used by computational theories.

There is, however, one very important missing guideline. This can be described as the "goodness" of the resulting representation. By "goodness" we mean that the edge-picture obtained should be as accurate and as complete a line drawing of the image as our previous experience dictates; the position, thickness, and orientation of the edges should be such that we should be comfortable with the reduction of the image to the edge-picture. This, admittedly, is a subjective criterion, but rightly so; after all, we are trying to explain a subjective phenomenon. It should also be noted, however,

that the “goodness” of the edge-picture is not a high-level criterion; what is perceived is not a caricature or an artists’ depiction of the scene, which would necessarily involve complicated cognitive processes, but a simple line drawing. Almost all previous edge detection research has seriously underestimated the abilities of human visual perception by choosing images where fine detail or subtle nuances are either irrelevant to the understanding of the scene (such as office furniture) or impossible to appreciate (such as vegetation). We are neither hard-wired nor trained for these purposes. To a naturalist, the legs of a gazelle, the horns of an oryx, or the spots of a jaguar may not seem “right” after they are rendered in a line drawing based on an edge-picture, while to us ignorant many they would seem perfect. There is, however, one area where all of us are experts: human appearance, more specifically, human facial appearance and expression. Here, we know that the most subtle variations are noticed. Very few edge detection techniques (and none described above) have been used to depict human faces. One exception is Letellier, et al.’s paper [Letellier, et al. 1985], where a dramatic comparison of major computational models is included. Letellier, et al. argue and convincingly demonstrate that the edge-picture formed by the zero-crossings of the Laplacian-processed image does not give us a “good” drawing of a human face, but the edge-picture formed by the maxima does. According to these results, the edges are “optimally” perceived to be in the dark side. That is precisely what Watt and Morgan’s findings [1983] on the subjective edge indicate.

At this point, it is clear that psychophysical Mach Band research is not only helpful, but also necessary. In fact, it is our only clue as to how the visual system perceives very simple edges. Through such research, not only can we answer some unresolved qualitative questions, but also we can produce quantitative models of the response of the visual system to edges at early stages.

3. Hermann Grid Illusion

The Hermann Grid illusion is probably the clearest manifestation of lateral inhibitory processes in the human visual system, but has received much less attention than it deserves and has been only inadequately investigated by vision scientists. Much of the effort belongs to the last decade. As in the Mach Bands illusion, research has produced contradictory results, challenged old explanations, and posed new questions, leaving much of the terrain unexplored. Since this will be a primary focus of investigation in our research programme, we find it necessary to review some of the major findings, controversies, and unresolved issues in Hermann Grid research. The number of reports are few; therefore, we take them individually.

3.1 Physiological locus and global factors

Lavin and Costall [1978] prepared a dichoptic presentation of the illusion and measured detection thresholds for both monocular and dichoptic viewing. They reported that the illusion was present in dichoptic presentation but that dichoptic detection thresholds were much higher than monocular detection thresholds. In spite of the difficulty and temporal instability of fusion due to severe

binocular rivalry, they concluded (from the dichoptic existence of the illusion) that "the central (cortical) site is a major if not the major site for its generation." This premature conclusion or emphasis was criticised by Troscianko [1982a] who pointed out the following additional inadequacies in Lavin and Costall's methodology:

a) The strength of the illusion was not measured independently

b) To reach the dichoptic detection threshold very large (and hence suboptimal) stimuli were used, whereas "when stimulus subtense is kept constant, a dichoptic presentation of the grid results in an abolition of the illusion."

Troscianko reinvestigated the problem using a more cleverly built stereoscopic presentation apparatus capable of giving each eye a degraded (or modified) stimulus, which, however, would give the classical Hermann Grid illusion in depth when fused in the cortex. The fused image did not give a stronger illusion than the degraded stimulus in each eye which led Troscianko to conclude that the peripheral site was the major cause of the illusion and that "the central mechanism only provide(d) an illusion for much larger stimulus sizes."

There are inherent difficulties in separating peripheral and cortical mechanisms. If an illusion occurs only stereoscopically, then we can confidently point out to the cortex as the generator of the illusion. (A good example is Julesz's random dot patterns which give a structured image in depth only when fused - the question here, however, was not the site of fusion, as this was already known, but whether structure contributed to stereopsis) The converse is not true, i.e., if an illusion occurs in both monocular and stereoscopic viewing, and if we manage, as Troscianko does, to show that stereoscopic viewing does not enhance the monocularly viewed images, the only conclusion we can draw is that stereoscopic input is not necessary to produce the illusion. We cannot rule out cortical mechanisms as any perception of an illusion requires the involvement of the visual cortex and since stereoscopic information is not a prerequisite for the formation of an image in the cortex.

The investigation of global factors in the Hermann Grid [Wolfe 1984] indicates the importance of and warns us about the subtle effects of cortical involvement. Using a magnitude estimation technique, Wolfe varied the number of intersections of the grid and found that as the number of intersections increased so did the magnitude of the illusion. To determine if the number of intersections and not the orderliness of the grid was the important factor, he compared an irregular collection of crosses with a regular grid which had the same number of intersections and found that the irregular cluster gave weaker percepts. The illusion was modified in oblique viewing, suggesting orientation specific processes. (Wolfe further reported that when the street width was held constant, the size of the squares was of moderate significance only for white grid-on-black background illusions, which Wolfe hypothesized to be the result of possible differences between on- and off-center ganglion cells). Wolfe reasoned that such global effects could not be explained locally in terms of concentric on- and off-center receptive fields, and since no other substitute cortical mechanism was yet obvious, proposed more experimentation to find out the effects of brief exposure times and eye movements as well as to investigate the illusion itself in the fovea.

Wolfe points out an inherent inadequacy in the magnitude estimation method: "Of course, little can be made of the precise shape of the (magnitude vs. no. of intersections) function given the use of a rating scale. One cannot base conclusions on the assumption that the difference between ratings 4 & 5 is the same as between 6 & 7." At this point, the usefulness of a compensation method becomes apparent; in fact, a compensation method has been successfully used by some researchers for the investigation of the Hermann Grid [Troscianko 1982a, 1982b].

3.2 Chromatic effects

The first systematic research into the chromatic properties of the Hermann Grid belongs to McCarter [1979]. Her basic stimulus was a Hermann Grid formed by 9 squares. There were three

independent variables, all of which were systematically varied: lightness contrast between squares and the grid, saturation of the squares, and hue of the squares. From these, first the sharpness and the clarity of the dots were rated by a magnitude estimation technique, and then Munsell chips were matched to the saturation and lightness of the squares (The hue of the chips was the same as the hue of the squares). McCarter reported that nobody complained about the hue of the matching chips indicating that the hue of the spots was the same as or similar to the hue of the squares. Statistical analysis of the data revealed that the sharpness and clarity of illusory dots were most affected by lightness contrast between the squares and the grid; lightness contrast, however, was not necessary for the perception of the illusory dots, a very important result indeed. Statistically, the lightness value of the matched Munsell chips was not affected to a significant degree by any of the factors. The illusory dots were lighter than the squares, and only in cases of high contrast (very light grid) were they darker than the grid. Saturation of the chips, however, was influenced by both the hue and the saturation of the squares. As the saturation of the squares was increased, so did the saturation of the illusory spots. Furthermore, longer wavelengths (red and yellow) resulted in more saturated matches than the shorter wavelengths (blue and green).

Showing that the opponent color cells (such as green on-center, red off-surround or green off-center, red on-surround or vice versa) would produce illusory spots of complementary colors against the findings of the experiment, McCarter attributed the results to monochromatic antagonistic center-surround cells (such as red on-center, red off-surround, or vice versa) or to double opponent cells in the striate cortex (such as red on-, green off- center, red off-, green on-surround and vice versa). (Figs. 8 and 9) She also noted that the activation of the achromatic mechanisms were possibly responsible for the sharpening and the darkening of the spots with increasing lightness contrast while the stability of the saturation of the spots under different lightness contrasts showed the independence of chromatic and achromatic mechanisms. (The effect of saturation of the squares on the saturation of the illusory spots could, for instance, be easily modeled by monochromatic antagonistic center-surround cells)

McCarter, nevertheless, did not hesitate to put a word of caution for some of the experimental results. She remarked that the lightness contrast polarity of the stimulus was not reversed in the experiment (the grid was always lighter), which, if it were carried out, could possibly lend "more powerful support to the existence of the chromatic induction effect in the absence of lightness contrast." Furthermore, she noted that the step sizes were not adequately small. (We would once again like to point out the arbitrariness of the magnitude estimation technique)

The saturation effects in the chromatic Hermann Grid illusion were investigated in detail by Levine, Spillmann, and Wolf [1980] who used a modified version of the grid (Fig. 10a). Munsell papers of identical hue and lightness value but of different saturation (5 levels) were used to create all possible saturation combinations of upper stripe, lower stripe and the background (squares). This was repeated for three different hues. In the first part of the experiment, subjects simply observed whether the illusory spots were more or less saturated than the upper stripe. "Alternatively, the subjects used the terms brighter or darker instead." In the second part, twelve representative intersections (four in each hue) were matched to six Munsell chips, each one one Munsell step along either of the three dimensions from the designation of the upper stripe.

The authors found that for a strong illusion, the saturation difference between the upper stripe and the squares should be null or very small, while at the same time, the saturation difference between the upper stripe (or the squares for that matter) and the lower stripe should be large. For maximal illusion this means a "gap" in the lower stripe (Fig. 10b). When the upper stripe and the lower stripe were of the same saturation, the illusion was moderate, and when the lower stripe and the squares had similar saturation levels, the illusion was minimal. This lends support to, but also clearly modifies McCarter's conclusion which linked the saturation level of the illusory spots to the saturation level of the squares in the classical grid.

The matching experiments showed that the saturation changes (of the illusory spot with respect to the upper stripe where it lies) was in the direction of the saturation level of the lower stripe, which is in accordance with McCarter's results. Levine, et al. also reported that the red-purple grids

(longer wavelength) resulted in more responses than the green-yellow grids (shorter wavelength), an observation which they linked to different receptive field sizes of different color sensitive cells. This result is similar to McCarter's. Like McCarter, Levine, et al. accounted for their data by monochromatic antagonistic center- surround cells or double opponent cells. An important auxiliary finding they mentioned, however, was the disappearance of weaker spots when the grid was rotated by 45 degrees, which again suggests the role of orientation selective cells in the cortex.

It is important at this point to note that the results obtained with the modified Hermann Grid point to a nonlinearity in the perception of the illusion. This is most apparent when the classical Hermann Grid is compared with the modified grid which gives the maximal illusion strength, i.e. the one where the upper stripe and the squares have the same saturation that is very different from the saturation of the lower stripe. Needless to say, a detailed analysis and documentation of this nonlinearity would be very helpful in understanding the functioning of chromatic mechanisms in human vision.

One discomforting element in Levine, et al.'s experiment was their willingness to substitute the terms "darker" and "brighter" instead of exclusively using "more saturated" and "less saturated." We would also like to point out that magnitude estimation was done on a limited scale and with a limited number of available steps.

Hue effects in the Hermann Grid illusion were investigated by Oehler and Spillmann [1981]. The structure of this experiment is identical to the experiment described above which was concerned with the effects of saturation. Modified Hermann Grids were made of Munsell papers of identical lightness and saturation but different hue. Ten different hues and grey were used. In the first part, the subjects were simply required to report whether they saw an illusion at an intersection or not. The strongest illusions were the ones most noticed. In the second part, subjects used some selected representative intersections to match the illusory dots to 31 Munsell chips varying in one or two steps along the direction of lightness value, saturation, or hue. One of the auxiliary experiments carried out was to find the effect of orientation.

The results are similar to the ones obtained from the previous experiments about saturation. Instead of different levels of saturation, here we have three major hue groups: The red group, the blue-green group and the neutral group. The red and the blue-green group are antagonistic in their behaviour. For a powerful illusion, the upper stripe and the squares should belong to the same hue group while the lower stripe should belong to the opposite hue group. The strongest illusion occurs when the upper stripe and the squares have the same hue, leading to a "gap" in the lower stripe. Unlike saturation, when the upper and lower stripes are of the same hue group but opposite to the hue of the squares (classical Hermann Grid), no illusion occurs. This result is very significant because it contradicts not only McCarter's results but also the explanation of the illusion in terms of double-opponent cells. In the matching experiment, the illusory spots were equated with Munsell chips that were shifted in either hue or saturation or both, but not in lightness. Generally the saturation of the illusory spots was greater than that of the upper stripe whereas the results for hue were too varied to draw general conclusions. The involvement of saturation in the illusory spots on grids differing only in hue is another example of the complexity of color perception. It is also a signal for the necessity for experiments whereby saturation and hue can be independently controlled.

In the auxiliary experiment, the strength of the illusion was again reduced when the grid was viewed at an angle of 45 degrees.

Oehler and Spillmann remarked that the colors in a given hue group all lay on the same tritanopic confusion line, indicating that blue cones did not participate in the hue illusion. Thus they modified their basic conclusion as such: "For a strong illusion to occur, the tritanopic purity difference (difference in other than blue) of the lower stripe and the background (squares) must be greater than the difference between the upper stripe and the background."

Oehler and Spillmann concluded their paper by citing the unresolved issues most of which are remarked above. One that is not, and which is particularly interesting is the asymmetry of the illusion in protanopes. Although strong illusory spots are seen by protanopes on grids with blue-green

lower stripes and red squares, the converse is not true. No illusion is seen by protanopes on grids with red lower stripes and blue-green backgrounds whereas color-normal observers see a strong illusion in both configurations.

3.3 Perceptive field sizes

The idea to use the Hermann Grid to investigate the distribution of perceptive field sizes was pursued by Troscianko [1982b]. Troscianko used a hollow variant of the Hermann Grid (Fig. 11) which would create the illusion by activating perceptive fields of only certain sizes and characteristics. These characteristics, however, are not rigid, but approximate. The important element in Troscianko's method is the cancellation (compensation) technique whereby a spot of light is shone to the intersection to cancel the illusion and to determine the increment thresholds. The apparatus limitations prevented the author from examining foveal perceptive fields, and he limited his dependent variable, the width of the streets, to seven values.

The results revealed that at photopic levels, as the eccentricity increased, the illusion increased in comparative strength (and also in overall strength) at larger street sizes. Expectedly, in mesopic adaptation levels, there was a marked increase in the perceptive field size, and at scotopic levels, the illusion ceased to exist except for the largest stimuli. Troscianko asserted that his results compared favorably with that of previous researchers and added that "relative diameters of the centers and surround may not be typical; it is these possibly atypical ones which seem to be responsible for this illusion." This, according to Troscianko, also explains why foveal Mach Bands extend to about one mm. although foveal receptive fields are known to be much smaller: "On the present theory ... Mach Bands would extend out to the extent of the largest receptive field present at appropriate quantity at that eccentricity." Troscianko found that orientation did not change the strength of the

illusion (although it modified the grid somewhat), which is a result that contradicts Levine, et al. and Oehler and Spillmann's findings.

4. Proposal

4.1 Raster Display Systems

Previous sections have pointed out to a major problem in Mach band research: the difficulty in measuring the various properties of the bands without altering the bands themselves. Most of the techniques used have been invasive, i.e., they have disrupted the original appearance of the bands; this has not escaped the attention of some experimenters. Perhaps the ideal method to measure the brightness of the bands (if not the width) is a compensation method as suggested by Stockham [1972]. Von Békésy [1972] is the only scientist that we know of who has used a compensation method. He compensated the Mach bands in rotating discs by altering the white/black ratio in some areas which were not directly influential in the production of Mach bands but were superimposed on the subjective bands when the disc rotated. The results he obtained are mostly qualitative; furthermore, as von Békésy himself notes, due to the nonlinearity, they may not be comparable to the brightness change produced by the bands themselves.

Much more sophisticated devices are available now. A very high resolution CRT, coupled to a computer, can provide the precision and speed so much sought for in earlier experiments. Using

compensation techniques, it is possible to make reliable measurements of the brightness and width of the bands under different conditions with little interference. Moreover, because the subjects will need only a minimum amount of training, the study may be extended to a larger population, giving statistically more significant and credible results.

A limiting factor in earlier experiments is the experimental set-up. Once a specific property or behaviour of the Mach bands was targeted, and the method decided upon, very little could be changed. A computerized raster display system, on the other hand, offers tremendous flexibility when compared with the time-honoured Maxwellian methods used in psychophysical vision research. Many set-ups can be stored in computer's memory, and each subject can do a variety of experiments or measurements in one sitting. As Cowan [1983] points out "... new stimulus configurations and modulation methods can be produced quickly, without the need for construction of mechanical and optical components."

Needless to say, a computer can also generate much more complicated shapes, simulate neural network models or neural computation algorithms, and, thus, actively test various theories about Mach bands. A high definition raster display system connected to a computer is particularly helpful with edge detection theories and algorithms. For instance, the edges generated by an algorithm can be traced on top of the digitized picture, or displayed beside, and the "goodness" of the theory behind the algorithm can immediately be assessed. Or the image may be in motion. It is not difficult to imagine the immense variety and availability of stimuli and the speed with which processing, assessment, and comparison can be carried out.

Perhaps the broadest and single most important area that computer controlled raster display systems will revolutionize is color vision research. Color vision research is difficult; this difficulty lies in the production and control of stimuli that will excite the different cone mechanisms in a predictable manner. Classical methods are hard to implement and offer very little flexibility not only for changes in spatial configuration or movement pattern of the stimulus, a characteristic also shared by achromatic vision research, but also in hue. It is thus beset with problems much less

straightforward, such as broadband transfer characteristics of some monochromatic filters [Matthews 1967], or the disruptive effects of a background that does not cease to exist in locations where color stimuli are superimposed [Rodieck 1983].

The most visible contribution of current color display systems is palette size [Taylor & Murch 1986]. Literally millions of shades of color in almost any stimulus configuration are possible. Furthermore, each color is exactly specified and controlled by the intensities of the emissions of three different phosphors. Since the CIE standards are also defined by tristimulus values and the systems are linear, transformations are straightforward. Based on these principles, Rodieck [1983] has shown how to use a color monitor to differentially excite only one cone system, up to now an elusive and much sought-after goal.

Spatial manipulation of color stimuli has already begun to produce very interesting observations; for example, spatial information is "... curiously degraded when objects are portrayed using purely chromatic contrast without any luminance variation; apparently the chromatic information cannot be used in the construction of three-dimensional shape from shading." [MacLeod 1986] Using moving colored stimuli has produced equally remarkable results: "... purely chromatic stimuli have very little capacity ... to excite motion-sensing mechanisms." [MacLeod 1986]

Some areas of research in color vision have readily found a suitable environment in color displays. Simultaneous color constancy (analogous to simultaneous brightness constancy), traditionally explored with Munsell papers in Mondrian-type displays, has successfully been investigated with color monitors [Arend & Reeves 1986]. The most dramatic use of raster displays, however, has been in heterochromatic brightness matching, where areas of different color are matched in brightness until the border between becomes minimally distinct [Taylor & Murch 1986].

The use of raster display systems in vision research is very new. There is only a handful of articles, while most of the work is not published [MacLeod 1986]. This hesitation, at least during later years, can be linked to a lack of information about the limitations of the display systems. These problems

are discussed in recent articles by Rodieck [1983], Cowan [1983], Cowan and Rowell [1986], and MacLeod [1986]. For the display devices, nonuniformity of the screen, nonlinearity of light emission from the phosphors, the lack of complete independence of the guns, temporal instability, and calibration are among the major problems [Cowan & Rowell 1986; MacLeod 1986]. Control devices, on the other hand, have quantization problems both in spatial and in temporal domains [Cowan 1983]. The pace of development both in the quality of CRTs and the power of information-handling devices, however, is very rapid, and for much of stimuli in vision research, raster display systems are ideal. As the problems and limitations of these systems, like any new technology, are investigated and dealt with, much of vision research, especially color vision research will rely on them in the near future. As Boynton [1988] points out in his review of color vision research, "It seems probable that raster displays will be used for most of the studies that will be cited in the next *Annual Review of Psychology* chapter on color vision in 1994."

Computerized graphic display systems have been extensively used in recent visual psychophysical research. Here, we will cite a few examples from 1989 and 1988, which have, if not exactly the same, a similar direction and flavor: (Bradley, et al. 1988), (Gorea and Papathomas, 1989), (Lindsay and Teller, 1989), (Bijl, et al., 1989), (Paradiso, et al. 1989).

4.2 Research Program

On the premise that a high resolution color raster display system be available, we propose the following research program to investigate the Mach bands and the Hermann Grid illusion and to solve the ambiguities in the current knowledge base:

- 1) It is evident from the preceding critique that psychophysical research on Mach bands has not produced satisfactory results, and that many unresolved issues and contradictory findings remain.

Therefore, we propose a series of experiments to investigate, using recent technology, the major psychophysical properties of Mach bands. Since it is not certain that the Mach bands illusion exists in discontinuous gradients, a range of gradients will be used with emphasis on limiting conditions. The following is a concise list of controversial topics and properties which can conveniently be explored with a raster display system.

1. The effect of the luminance level on the width of the bright band.
2. The effect of the slope of the gradient on the brightness (magnitude) of both bands.
3. The independent variables that affect the detection threshold.
4. The behaviour of incremental thresholds in dark bands.
5. The existence of Mach Bands in steep intensity gradients, and if they do disappear, the existence of a threshold.
6. The existence of chromatic Mach Bands and their behaviour.

2) The Hermann Grid illusion, unlike the Mach Bands illusion, has not been widely investigated. Still, as noted in the critique above, contradictions exist in the few published reports. The effect of the saturation and the hue of the squares on the illusory spots, for instance, is controversial. Psychophysical measurements of the Hermann grid may not only bring clarity and precision to the subject, but also help determine the parallels and differences with the Mach bands illusion.

3) It is controversial whether further processing in the striate cortex is necessary to perceive the Mach bands. Binocular fusion experiments, using random dot stereograms, can provide a convincing answer. Admittedly, the precision that is possible with straightforward psychophysical experiments is difficult to achieve with random dot stereograms. Since the eyes should not receive too disparate an input, the luminance difference between the two patterns should not be large. The

experiment aims to present each eye a random dotted gradient pattern, which, by itself should not exhibit the Mach bands illusion (or exhibit it minimally), but when binocularly fused, might do so (or exhibit it maximally). The feasibility of this experiment can not be known beforehand; it should be determined during the course of the investigation.

4) Heterochromatic brightness matching methods described earlier can settle most of the ambiguity in chromatic Mach band research. Much of the early research has relied on equiluminant areas of different hue. The 1931 CIE luminous efficiency function, which was used in measuring subjective luminance in color vision experiments, is outdated now [Howett 1986]. Furthermore, additivity of luminances of different hues was assumed. This assumption is false [Howett 1986]. The minimally distinct border criterion, on the other hand, gives a much more accurate description of the subjective sensation of brightness, and fulfills Pease's recommendation [1978] for subjective criteria. Hence, each subject participating in chromatic Mach band experiments will first perform a heterochromatic brightness match between the two areas of different hue and then will proceed with the main body of the experiment. If the existence of the bands is established, a chromatic compensation procedure will be used for fine measurements.

5) There are various models of lateral inhibition, though not necessarily in the human retina, most of which are summarized in Ratliff's book [1965]. Recently, there has been renewed interest in the area [Srinivasan, et al. 1982]. These models can be tested, using, again, compensation techniques. Furthermore, when combined with the results of the psychophysical experiments, the data base should be sufficient for improvements on already existing models or for developing new models. The Mach bands illusion is the result of a mechanism that serves an important purpose in vision, and Mach band research should be directed and interpreted in this context.

With the following research program, and with raster display systems, a number of pitfalls that have appeared in earlier experiments will also be avoided. Many independent variables, for instance, can be manipulated singly because of increased flexibility. Similarly, low spatial sampling will not be a problem. Very probably, the experiments will involve a larger number of people to assure statis-

tical reliability. Finally, the use of the compensation method will eliminate the problems of binocular matching and comparison stimuli.

Before elaborating on the procedure of psychophysical experiments based on compensation, it is necessary to look at the compensation method more closely, and recognize its limitations.

4.3 *Compensation*

The basic component of the nonlinearity in the transduction of light intensity to spike frequency in the human eye is believed to take place in the photoreceptors themselves, and can be approximated with a logarithmic or a power series function. Introducing a complementary nonlinearity to the stimulus as a whole can compensate for this initial nonlinearity in the eye. In this discussion and the proposed experiments it will be assumed that such a transformation is being applied to the stimulus. Consequently, the stimulus is more or less as the retina will receive it after its detection and transduction to electric potential at the outer segments of the photoreceptors. For simplicity, a linear gradient as shown in fig. 12 will be used to create Mach Bands.

In the following discussion, we are considering a one dimensional intensity gradient (along, for instance, the x-axis) in a two dimensional rectangular stimulus pattern. Therefore, all transformations act on this gradient.

If we denote the intensity gradient by G , the transformation in the visual system which gives rise to Mach bands is $T(G)$. If M denotes the Mach Bands, then

$$M = T(G) - G \tag{1}$$

The compensated stimulus is $G + C$. (We will not yet assume that C is $-M$) The transformed, compensated stimulus is $T(G + C)$. Therefore, if the compensation is successful, i.e., the Mach Band percept is prevented, then

$$T(G + C) = G \quad (2)$$

This equation, however, assumes that the same transformation applies to $G + C$ as to G alone. While this might seem logical, it is not necessarily true. Given that retina has ganglion cells of different receptive field sizes, if C is a narrow band, for instance, it is quite possible that a slightly different transformation might apply.

Substituting from (1),

$$T(G + C) = T(G) - M \quad (3)$$

We have two cases:

a) If the transformation is linear (or it exhibits linearity for the range of photopic intensities we are interested in), and its inverse exists, we have

$$\begin{aligned} T(G) + T(C) &= T(G) - M \\ T(C) &= -M \\ C &= T^{-1}(-M) \end{aligned} \quad (4)$$

If we try to make the compensating stimulus the negative of Mach Bands, the transformation reduces to the identity transformation:

$$C = -M \quad (5)$$

If so, however, when we look at (1),

$$M = G - G = 0 \quad (6)$$

and the whole problem is trivialized.

b) If the transformation is not linear, and its inverse exists, we have from (3)

$$\begin{aligned} T(G + C) &= T(G) - M \\ G + C &= T^{-1}(T(G) - M) \\ C &= T^{-1}(T(G) - M) - G \end{aligned} \tag{7}$$

There is nothing we can do at this point. We can attempt to solve the equation only after the transformation is known. But it is exactly for this reason, to understand the transformation, that we are using a compensation stimulus.

It is possible, as mentioned above, that different transformations may affect the stimulus patterns.

In this instance, we have two cases:

c) The transformation on the compensated stimulus is different from that on the original one. Denote it T' . Then we have from (3)

$$T'(G + C) = T(G) - M \tag{8}$$

If T' is linear, and its inverse exists,

$$\begin{aligned} T'(G) + T'(C) &= T(G) - M \\ T'(C) &= T(G) - M - T'(G) \\ C &= T'^{-1}(T(G) - M - T'(G)) \\ C &= T'^{-1}(T(G) - M) - G \end{aligned} \tag{9}$$

If T' is not linear, and assuming its inverse exists,

$$C = T'^{-1}(T(G) - M) - G \tag{10}$$

d) The transformation acting on the compensated stimulus can be separated into two distinct transformations; one belonging to the original stimulus, i.e., T ; the other to the compensation, i.e., T' . Then from (3) we have:

$$\begin{aligned} T(G) + T'(C) &= T(G) - M \\ T'(C) &= -M \end{aligned} \tag{11}$$

and if the inverse exists

$$C = T'^{-1}(-M) \tag{12}$$

If, as in a), we make T' the identity transformation such that

$$C = -M$$

Unlike case a, equation (1) is not trivialized.

$$M = T(G) - G \tag{14}$$

The discussion above is simplified since it doesn't take 2-D spatial extent and effects into account. Certainly it is valid as long as the compensation stimulus is added to the original (gradient stimulus) at all locations. If the compensated area, however, is small and located so as not to disturb the critical Mach Band producing areas of the stimulus (namely the gradient), as shown in fig. 12, then its distortion of the original stimulus may be tolerable.

The center-surround configuration of ganglion and bipolar cell receptive fields which gives rise to Mach bands in the classical explanation of the illusion allow such a separation of areas whereby the area where the compensation stimulus is placed is neither actively involved in the creation of the bands, nor greatly disturbs the areas essential for the production of the bands.

According to the classical explanation, the bright Mach Band in fig. 12 is formed when the lateral inhibitory surround of the ganglion cell receives less illumination than that of its counterparts in the bright equiluminant side. The portion of the surround which does not receive as much illumination is located in the gradient, while the center is located on the border between the gradient and the bright side, at the location of the Mach Bands. If the compensatory stimulus is small, and its location coincides with the center (or the Mach Band), it will not interfere much with the surround. If, furthermore, the lateral inhibition in the human eye is not recursive (see the physiological summary), this interference will be minimal.

Since we do not know the nature of the transformations taking place in the retina (or elsewhere in the visual system), it is very difficult to determine the functional shape of the compensation stimulus. If the compensated area is small, however, we may assume that different transformations act on the final composite stimulus. This is case d. It may also be assumed that the compensation is not transformed, or its transformation is very close to the identity transformation. It is then possible to assign $-M$ for the compensation stimulus. The following experiments are mostly based on these assumptions.

4.4 Achromatic Mach Bands

The function to be used to compensate Mach Bands can be selected based on the shape and size of the compensated area and on the knowledge of the approximate shapes of Mach Bands. The shapes of the bands are depicted in several articles. If a small circular spot is used as the compensation area, probably a uniform intensity distribution is a more suitable choice. If a rectangular band is being used as the compensation area, then several easily manipulated and versatile functions can be used. A sine wave, a Gaussian and a decaying exponential (reflected along the ordinate) can serve as approximate models of the Mach Band peaks. In the discussion below, an exponential is

used. A series of such exponentials can be used to approximate any function effectively. The experimental methodology, however, can be applied to any function; the main idea is to carry out a number of adjustment experiments initially, and then to fine-tune the results with forced-choice experiments. Furthermore, as dictated by time constraints, the following procedure can be shortened in a number of ways.

The damped exponential that will be used initially in the compensation of the bands is of the form

$$i = ae^{(-bx)}$$

where i is the intensity, and x is the value of the x -coordinate. There are two constants: a modifies the overall amplitude; b is the time constant of the exponential specifying its rate of decay. The maximum of this function occurs at $x = 1/b$; therefore, when we change the time constant, we also change the location of the maximum.

When we are using this function to compensate, we can change the maximum intensity by varying a , the "spread" of the function and the location of its maximum by varying b , and the location of the whole function on the screen by varying the origin (horizontal translation, say l).

Initially, a series of adjustment experiments will be carried out. Each subject will find the compensation exponential for a specific intensity distribution using an adjustment technique whereby one or more of the changes described can be made. Varying the horizontal displacement of the function and the time constant at the same time, however, can be difficult, time-consuming, and indecisive, mostly because both variations cause the location of the maximum to change and create an infinite number of alternatives. Therefore, this section will be preceeded by a preliminary investigation where the experimenters will determine, with an adjustment technique, freely varying all parameters, an approximate time constant for the exponential for each intensity distribution. A preliminary set of exponentials with specified time constants and each corresponding to a predetermined intensity profile will thus be obtained. The subjects will first be shown the original Mach Band stimulus and will be asked to indicate with a marker the maximum or the minimum of the

intensity distribution on the screen. These locations will be tentative. The subjects then will be shown the stimuli with the corresponding exponentials superimposed, and will be asked to vary the horizontal displacement only. The criterion will be the greatest possible neutralization of the minimum or the maximum of the intensity distribution under the markers. At this point, complete neutralization of the maximum or minimum or the neutralization of the entire extent of the bands are not sought. The mean of the responses from each subject will be chosen as the location of the minimum or the maximum of the band for the given intensity distribution. We believe that this will give a better measurement of the location of the bands than estimating from the Mach Band stimulus alone. The next step of this initial stage will be the adjustment of the amplitude a so that the maximum or the minimum will be completely neutralized. As with the previous set, repeated measurements from each subject will be taken, and a mean will be found for each intensity distribution. This will be followed by a set of experiments where the time constant b will be the independent variable to be changed by the subjects. Again, for each exponential, the mean location and the mean amplitude found above will be used. Here, we are more concerned with the neutralization of the whole extent of the bands. Repeated measurements will yield a mean time constant for each subject for each intensity distribution.

In the second part of the experiment, a three alternative forced-choice technique will be used for greater accuracy. A collection of exponentials will be generated for each subject based on their "mean" replies. All three variables (amplitude, time constant, and horizontal location) will be varied in discrete amounts in all combinations for this collection. In order not to cause a combinatorial explosion in the number of stimuli in each collection, these discrete variations will be limited to three along each independent variable. Therefore, for each Mach Band stimulus we have twenty seven compensation exponentials all based on the exponential with the mean values determined in the previous section. The subjects will be presented this collection superimposed on that particular Mach Band pattern three at a time (not to overcrowd the display area) with two of the independent variables held constant and the third one varied. They will be asked to select the one with the most neutralized (or prominent) Mach Bands. The selected ones in each group will be presented again

(or discarded), in the same fashion inasmuch as possible, and the last set of three winners will be presented for a final decision. There is, however, an alternative to this method of successive elimination. Although, there are a limited number of compensation exponentials, each one may be shown several times, in each case, as a member of a different threesome. For instance, we may have a set where the location and the time constant is kept constant, and the amplitude is varied, say set $a_1b_1l_1$, $a_2b_1l_1$, $a_3b_1l_1$. One of the members of this set, let's say the first, will be in another set where the location and amplitude is kept constant and the time constant is varied, for instance the set $a_1b_1l_1$, $a_1b_2l_1$, $a_1b_3l_1$. If each exponential is displayed several times among different partners, superimposed of course on the Mach band stimulus, then we might use a more statistical criterion to choose the best compensating exponentials. This will be carried out for each subject for each intensity distribution examined.

Because of inherent nonlinearities in the human visual system, a residual Mach Band may not remain once the greater portion is neutralized. If it remains though, a similar series of exponential compensation will be undertaken. This time, there will be no preliminary stage by the experimenters; the subjects will select an initial "coarse" time constant themselves. The forced-choice technique will be carried out for two collections of stimuli this time: one will be the residual intensity distribution compensated with the new exponential, with the exponential varying as described above; the other one will be the original intensity distribution compensated by the sum of the initial exponential (the last one remaining in the first experiment) and the new exponential, again varying across the important parameters.

For each intensity distribution, the data from different subjects will be pooled and statistically analyzed to understand the variation in the Mach Bands with the change of slope and luminance level. If, however, interpersonal differences are statistically significant, this analysis will be carried out for each subject separately.

4.5 *Chromatic Mach Bands*

The investigation of the chromatic Mach Bands will be based on the same concepts. Chromatic Mach Bands, however, pose some problems. One difficulty is quantifying a "chromaticity" or hue difference as precisely as possible so as to obtain an objective exponential function. The major difficulty, however, is to keep the luminance constant throughout the stimulus. We can resort to two different methods. In the first one, we can use the standard CIE system to find equiluminant stimuli of opposite hues, and then can transform the tristimulus values of the CIE primaries to the tristimulus values of the phosphor primaries. Since both are linear systems, this is quite straightforward. CIE system, however, assumes additivity of luminances from each primary, which is not exactly true. Furthermore, computational work is extensive, and the concepts are difficult to relate to. More recently, MacLeod and Boynton [1978] came up with another chromaticity diagram based on stimuli of equal luminance, which makes the computations easier and is more directly related to cone excitation space. The simplification comes from the findings that blue cones do not contribute to brightness sensation. This system, we believe, is a better choice for choosing equiluminant stimuli.

The other method approaches the problem more empirically and relies on the purely subjective measure of a minimally distinct border between two different hues. In minimally distinct border experiments, the stimulus is made up of two neighboring areas of different but uniform hue. The subject manipulates the luminance of each area until the border between becomes minimally distinct. Since contrast sensitivity depends heavily on luminance differences, the two areas are now thought to be equiluminant. This, we would like to point out, is a subjective measure, not universally accepted. It is, however, no worse, in our opinion than the CIE system; in fact, it may be an excellent subjective criterion, the lack of which, according to Pease [1978], has cast doubt on earlier chromatic Mach Band research.

In this method, we will use the emission spectra of the phosphors (which are narrow band) and the quantal absorption spectra of the human photopigments to determine the quantal catch of each cone for each hue that will be tested in the uniform area. Therefore, for each Mach Band generating pattern, there will be two quantal catch profiles. Since the brightness sensation is not necessarily proportional to these numbers and since we only want to investigate the Mach Bands due to chromaticity, the luminance of each area will initially be equated after a subject specific minimally distinct border experiment. The quantal catches will be determined after this. The gradient will be established by dividing the difference of the quantal catches into equal increments for each cone system. It is desirable to have hues with a quantal difference restricted to two cone systems only (for instance, long and middle wavelength cones and not short wavelength ones). Thus, i in the intensity equation for the achromatic Mach Bands will be the difference between the excitation of two cone systems (the quantal catch difference of the two cone systems over and above the values for the plateau) for the chromatic Mach Bands. The computer will use the value of i (the parameters of which are adjusted by the subject through a potentiometer) to calculate the intensities of the electron guns at each location.

At this point, the contribution from luminance is minimized but we are not certain whether the bands receive contribution from saturation difference of the two areas. Saturation, the amount of white light in the perception of colour, is closely tied to the other and relatively more independent percepts of hue and luminance. We are not aware of any quantitative relationship established between wavelength and intensity distribution, which we can objectively measure, and the percept of saturation. As long as both areas have opposite hues, and very close saturation levels (to the observer), however, we do not think that saturation will be a major interference. On the other hand, a Mach Band experiment devoted to saturation is the subject of a different and more involved experiment whereby saturation, hue and intensity are linked to the strengths of the electron guns and the emission spectra of the phosphors. Difficulties similar to that of hue exist in testing for saturation Mach Bands exclusively. While the quantal catch difference in both uniform fields and the gradient can be kept constant, the addition of a uniform white light to one of the fields and gradu-

ally to the gradient without changing the intensity, is an unsolved question and will probably require further experimentation. Furthermore, the relationship between saturation and lightness value has long been noted by subjective classification systems such as Munsell's. High lightness values, and high chroma values, for instance, are mutually exclusive. Therefore, at this point, we restrict ourselves to hue Mach Bands.

The initial step in determining the properties of the chromatic Mach Bands, however, is to find out whether they exist. The individual results may widely vary in spite of the quality of the experimental apparatus. Thus, before a compensation experiment, a series of matching experiments will be carried out. The experimenters will set markers on specific locations of the test field which will be on the upper half of the screen. On the bottom half, there will be a comparison stimulus: a band located with the markers will be embedded in a grey equiluminant field in alignment with the markers. The subjects will be asked to note whether they see any difference. Alternatively, this can be presented as a matching experiment whereby the subjects will be asked to choose from three comparison "bands" which vary slightly in hue and one of which is identical to the marked area in the test field. This, however, is to determine whether any Mach Bands are perceived; we do not believe, as stated elsewhere in this proposal, that a matching method will give as accurate a measurement of the profile of the bands as the compensation method. The results from the subjects will be used to choose a subset with the higher scores in the perception of the bands. These selected subjects will proceed to the compensation experiment. Naturally, if a statistically significant proportion of subjects do not see the bands, the controversy will not have been cleared.

4.6 Incremental Detection Thresholds

Incremental detection threshold experiments can give us a good understanding of the magnitude of the illusion. With a computerized raster system, they can be done easily and effectively.

The stimulus to be used in detection threshold experiments will be a small circular spot of uniform intensity or chromaticity. As with compensation experiments, the procedure can be divided into two parts. In the first one, a series of experiments carried out using the method of adjustments gives us an approximate but still fairly accurate estimate of the power of the illusion. Since the stimulus is uniform in intensity, the number of independent variables is much smaller than that in the compensation stimulus. With achromatic Mach Bands, this is only intensity and location; with chromatic Mach Bands the intensities of each electron gun should be considered.

The locations where stimulus will be presented will be predetermined, and at each location subjects will make two sets of adjustments; one using an ascending series, and the other using the descending series. The stopping rule is the detection of the spot. From each series, the mean results will be obtained, and a set of intensities between those two will be selected as the stimuli for the next part of the experiment.

In the second part, a two-alternative forced-choice technique will again be used for better accuracy. For each location, the determined set of intensities will be divided into two groups, equally distributed above and below their mean. Two stimuli, one from each group, will be presented at each of the predetermined locations, one on top of the other. The subject will be asked to identify the least discernible one. After successive eliminations, the intensity or the chromaticity of the least discernible spot is the incremental threshold at that location.

4.7 Hermann Grid

It is clear that research on the Hermann Grid illusion has revealed many interesting properties of human vision, and has given a glimpse of how this illusion can be used to explore the functioning of the visual system. Yet, as pointed out in some of the papers, questions and undecided issues are

more numerous than answers. Based on the availability of a high resolution computer graphic sytem, we are proposing the following experiments:

1. An exhaustive investigation on the magnitude of the illusory spots: For the achromatic Hermann Grids, this means the lightness value; for the chromatic grids, it means hue. We propose a compensation method which will eliminate the arbitrariness of rating scales. This will be very similar to the previous Mach Band experiments. A method of adjustments will be used initially to determine the approximate settings of different parameters; for final measurements a fixed choice procedure relying on both ascending and descending series will be used. The stopping rule is the full compensation of the illusory spots. Since the inhibitory areas (squares) are not compensated, the interference of the compensating stimulus will be small, especially if the inhibition is not recursive. During the adjustment section, for achromatic grids, the subjects will be able to change the lightness value of the compensating spot, while for chromatic grids, they will be able to control the wavelength distribution and the intensity. In chromatic grids, as with chromatic Mach Band experiments, the equivalent luminances (lightness values) for selected grid colors will be determined using either MacLeod-Boynton system or the minimally distinct border criterion in a series of preliminary experiments.
2. Since extremely high resolution is possible, foveal Hermann Grids will be examined, probably for the first time.
3. Perceptive field distribution for the chromatic Hermann Grid illusion will be examined with a procedure similar to Troscianko's. Since the stimulus configuration will be highly flexible, a grid changing in size with eccentricity and keeping the illusion constant throughout the visual field, will also be used.
4. The nonlinearity in the modified Hermann Grid will be investigated using stimulus configurations which will offer flexibility in size as well as chromaticity or lightness coordinates. The compensation method will again be the measurement tool.

5. Conclusion

The preceding chapters have shown that previous Mach Band and Hermann Grid research has not been conclusive. It is our opinion that this fact alone warrants a reinvestigation of the subject. The tools of modern technology can be employed to give the experimenters greater freedom and flexibility, and to produce more accurate and consistent results. The preliminary trials suggest that such experiments are feasible. These experiments, however, will also be a starting point in understanding a basic component of the human visual system. The ultimate goal is a mathematical model.

We should, however, be cautious. Although lateral inhibition in some organisms like the horseshoe crab is well documented and understood, lateral inhibition in the human retina is not. Psychophysical experiments suggest that there are different kinds of inhibitory processes in the human visual system. There is no doubt, however, that the human retina and the possible inhibitory processes are more complicated and perhaps fundamentally different than those of the horseshoe crab. Indeed, since even the retinas of the vertebrates are very different in wiring, what is known about species other than the old world monkeys cannot be of much help. On the other hand, since single-cell recordings from the monkey retina have not been possible (except in a few instances), research has concentrated in two isolated areas: Psychophysics and cytology. (In this respect cortex research has been luckier; at least single-cell recordings are widespread) Recent advances in cytology have made it possible to understand the types of cells in the human retina and their functions more

clearly, although much remains unknown. It seems that a synthesis is necessary; this should be a computational approach based on physiological findings and supported by psychophysical experiments. It is difficult to isolate the retina as the major cause of a psychophysical finding, but then again, this kind of ambiguity is inherent in all psychophysical research. Furthermore, if a computational retinal model predicts some psychophysical results, there is nothing wrong with attributing the results, at least tentatively, to the retina. From the viewpoint of physiology, there are already some models of wiring in the human retina [Gouras & Evers 1985]. Supported by concurrent research on the visual cortex, this synthesis can lead to a better understanding of the nature and purpose of visual processing. Computational vision is a new area, but one destined to be crucially important.

6. Figures

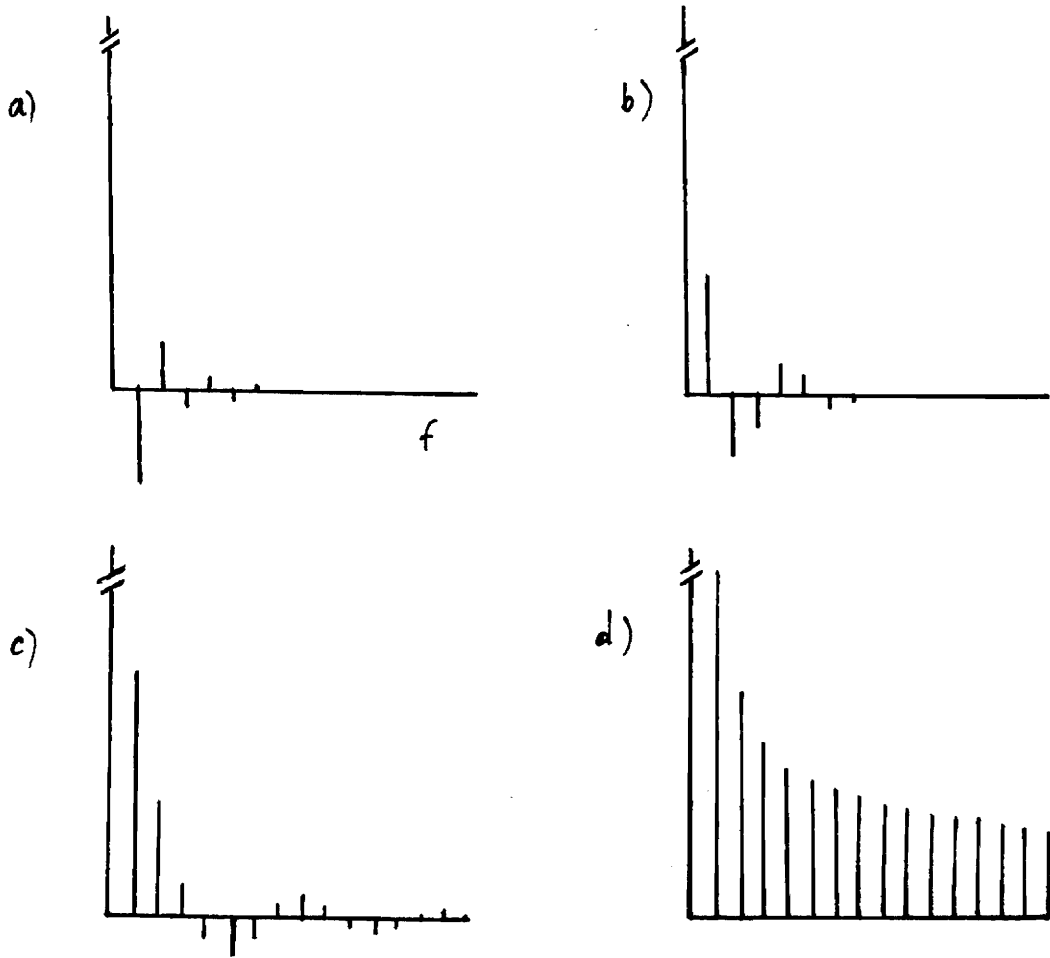


Figure 1. Fourier series expansions: (a) triangular wave, (b) and (c) trapezoidal waves, (d) square wave, [from Morrone, et al. 1986]. According to Morrone, et al., Mach bands are not perceived in square waves (discontinuous gradients) because they do not have any out-of-phase (here, negative) components. It should be noted, however, that the existence of Mach bands in discontinuous gradients is controversial.

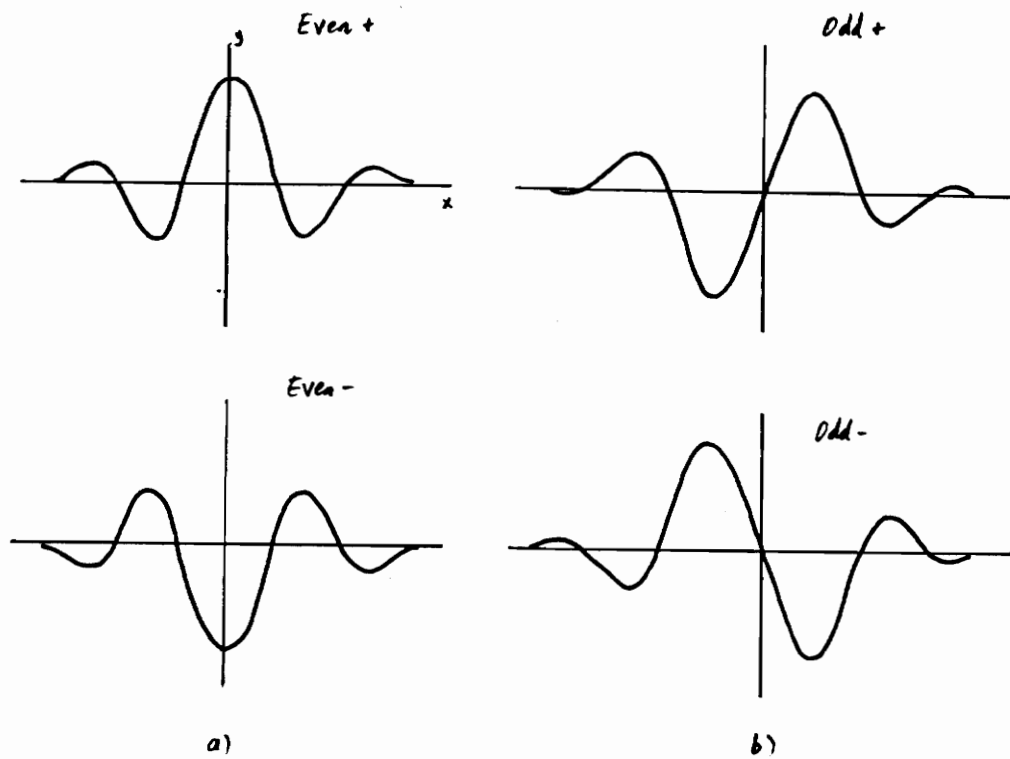


Figure 2. Even and odd-symmetric receptive fields: (a) Even, and (b) Odd-symmetric receptive fields. The one dimensional section of the field is along the x-axis, the response is along the y-axis. The negative areas are due to lateral inhibition. (from Bennett & Banks [1987])

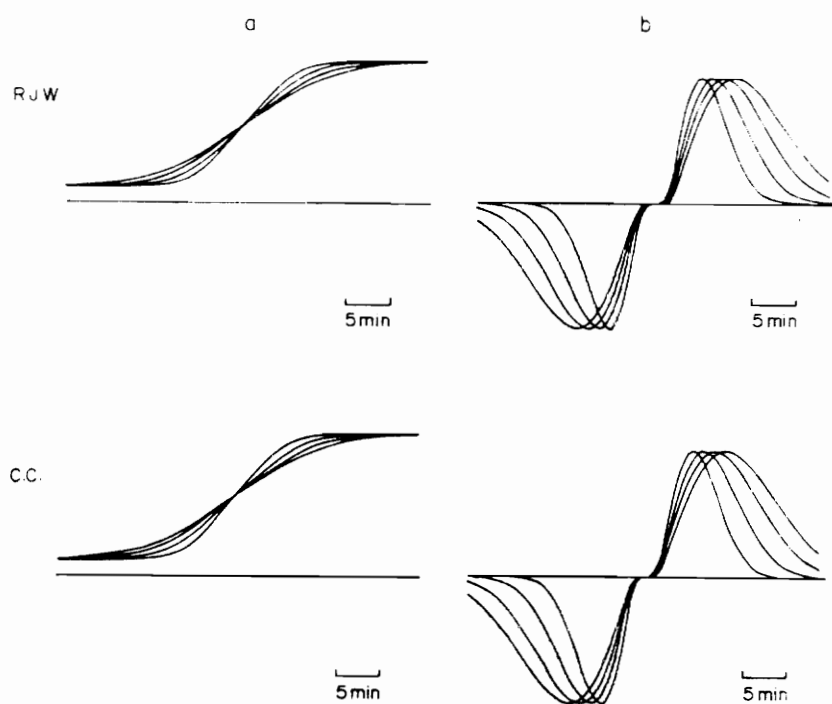


Figure 3. Mach Band generating retinal light distributions: For two subjects: (a) Retinal light distributions of an edge blurred with four different composite - gaussian + rectangular - blurring functions. According to Watt and Morgan [1983], these distributions give rise to the perception of Mach bands. (b) All normalized second derivatives show an inflection point of substantial width. [from Watt & Morgan 1983]

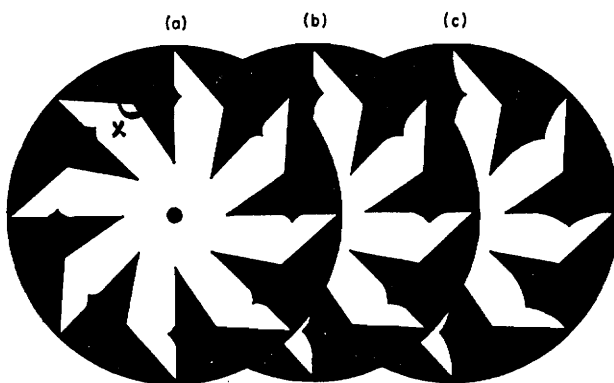


Figure 4. The discs used by von Bekesy [1972]

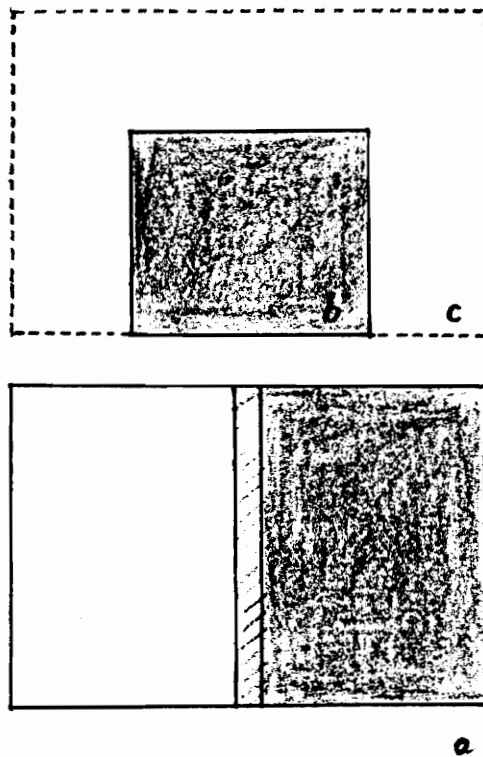


Figure 5. Stimuli used by Fiorentini and Radici [1957]: The test field (a) is presented to the observer's right eye; the comparison field (b) is presented to the left eye. For dark band measurements, an inducing field (c) is used; the configuration shown above is for dark band measurements. When binocularly fused, the test field appears below the comparison field, as seen above.

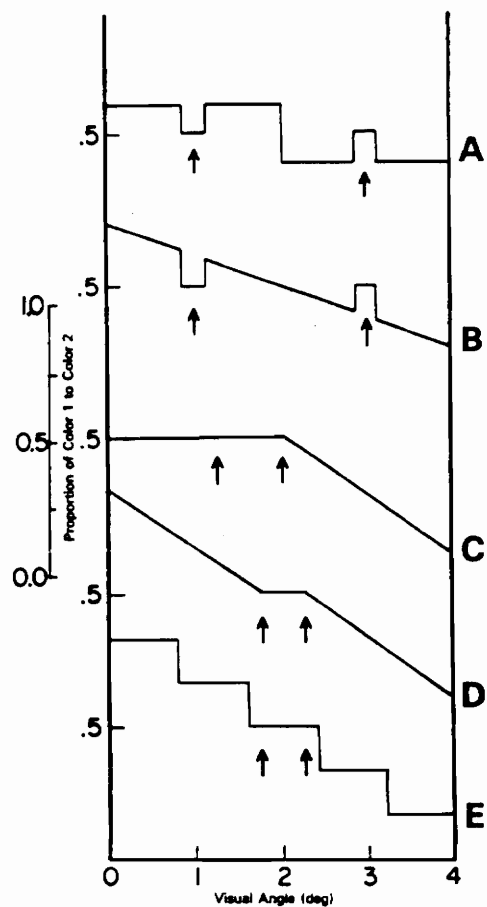


Figure 6. The five contrast patterns used by Ware and Cowan [1987]: A and B give rise to simultaneous brightness and color contrast phenomena. C is the classical Mach band pattern without the dark or opposite color uniform field; D is a variation. E creates the Chevreul illusion in the achromatic domain. For chromatic measurements, the ordinate shows the proportion of one of the two component colors out of a possible total of 1.

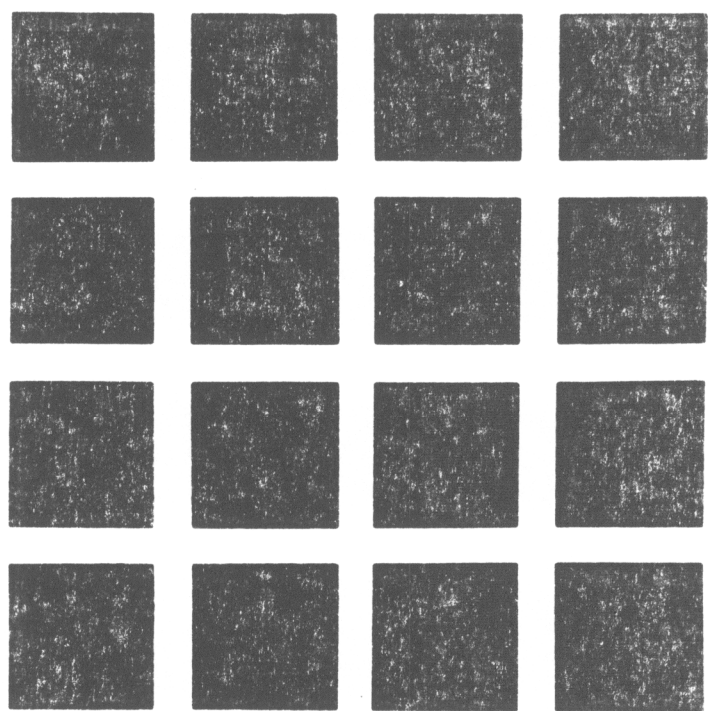


Figure 7. The Hermann grid. Small grey dots appear at the intersections.

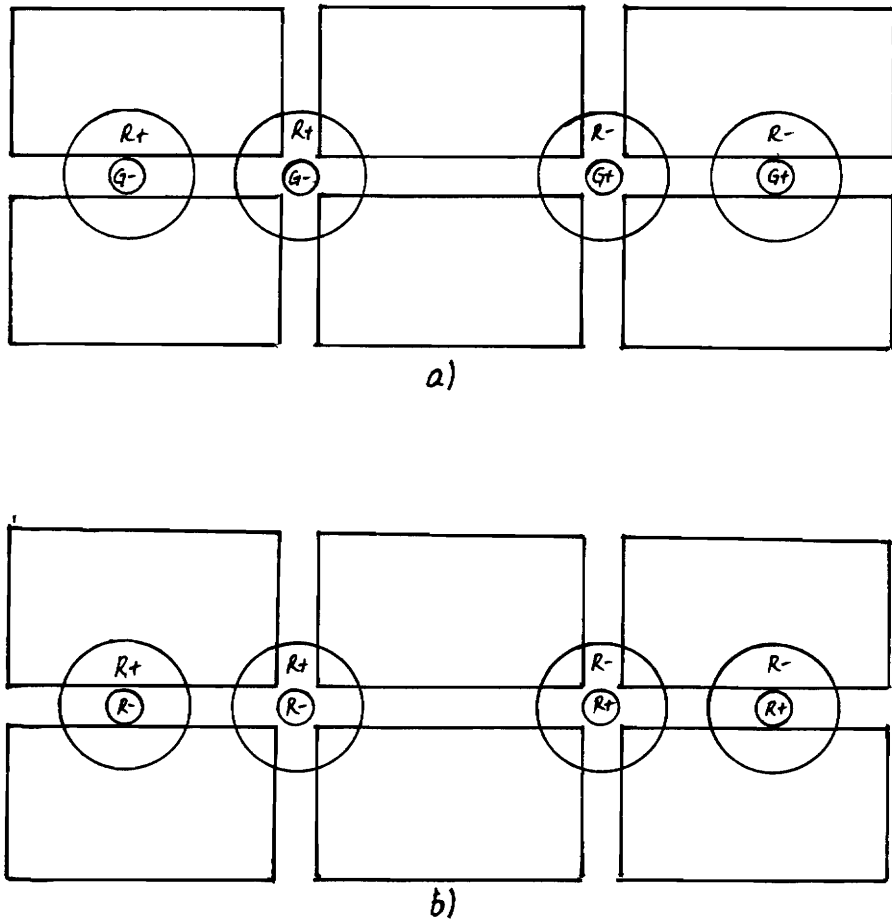


Figure 8. Stimulation of ganglion cells by a chromatic Hermann grid: The grid consists of red squares and gray intersecting bars. In a) the ganglion cells are both spectrally and spatially opponent with off-center cells on the left, and on-center cells on the right. In both cases, an enhancement of green at the intersections is predicted. In b) the ganglion cells are spatially opponent but spectrally non-opponent with off-center cells on the left, and on-center cells on the right. In both cases, an enhancement of red at the intersections is predicted. (from McCarter [1979])

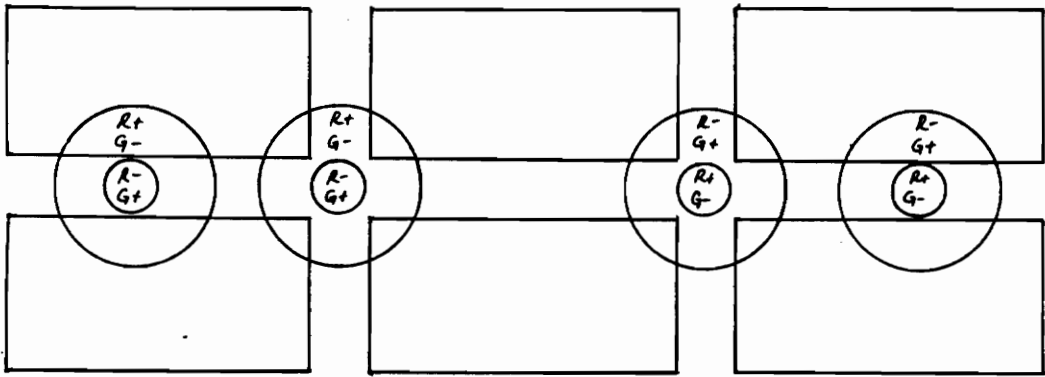


Figure 9. Stimulation of double opponent color cells by a Hermann grid: The grid consists of red squares and grey intersecting bars. In both cases, an enhancement of red at the intersections is predicted. (from McCarter [1979])

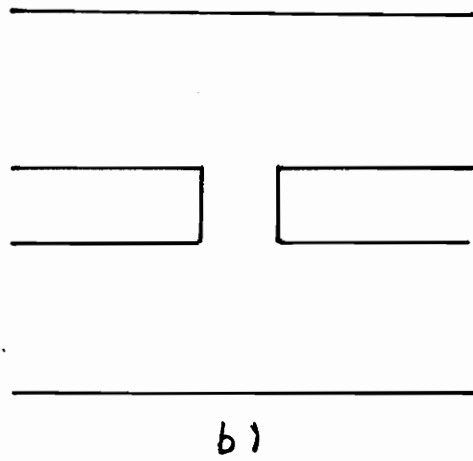
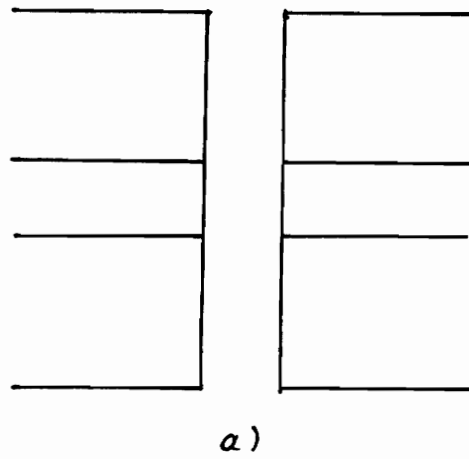


Figure 10. The modified Hermann Grid used by Levine, et al. [1980]: a) The intersection in the modified grid, b) For the creation of maximal illusion, the upper stripe and the squares should have the same saturation which should also be quite different from the saturation of the lower stripe.

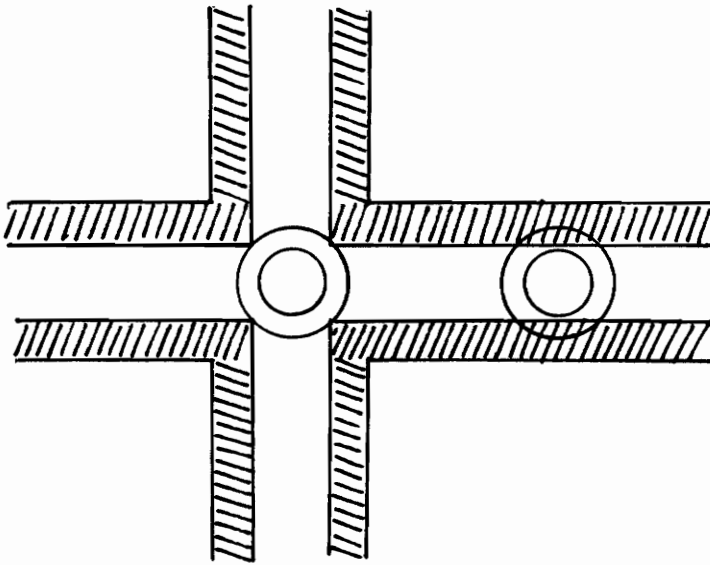


Figure 11. Intersection in a hollow Hermann Grid used by Troscianko [1982b]

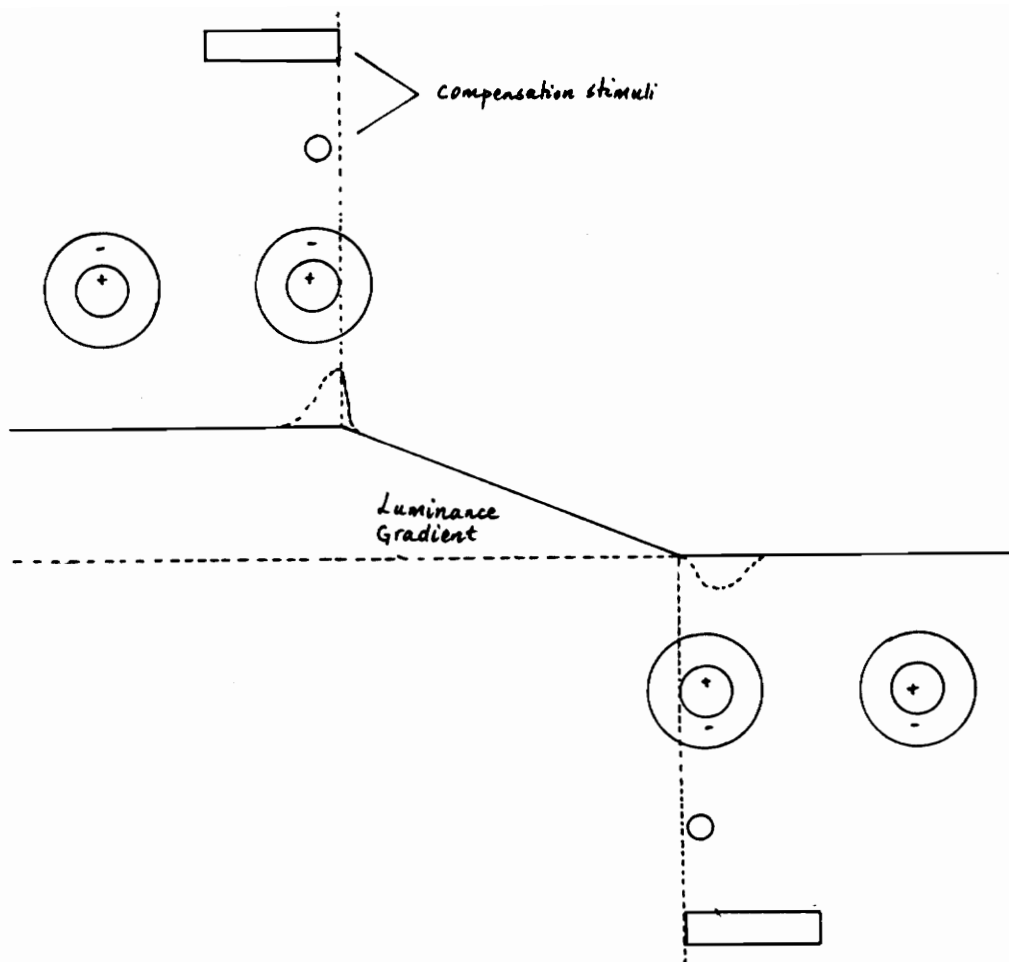


Figure 12. A schematic of proposed experimental compensation stimuli: The Mach Band generating gradient and two compensation stimuli, which, according to the classical explanation of the phenomena, would have minimal interference in the perception of the bands.

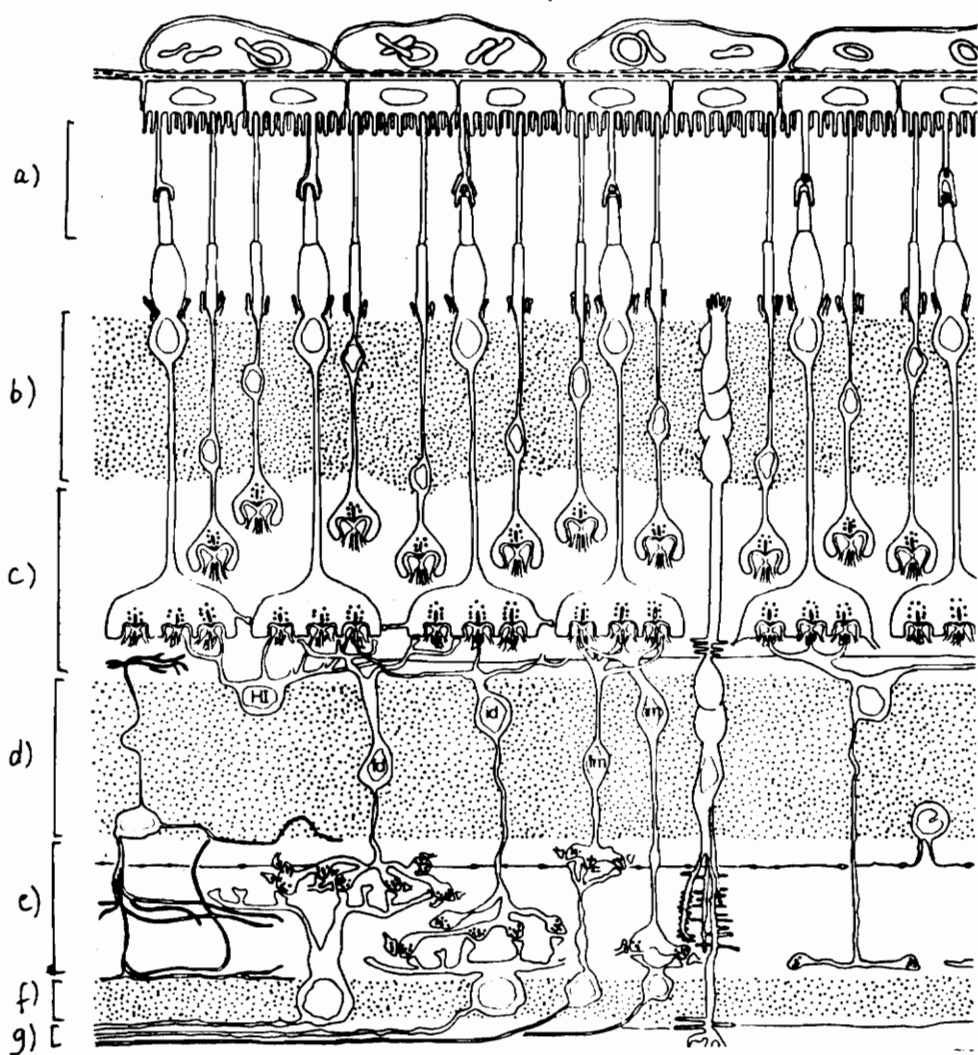


Figure 13. The organization and layers of the primate retina [Rodieck 1988]: a) Outer segment layer, b) Outer nuclear layer, c) Outer plexiform layer, d) Inner nuclear layer, e) Inner plexiform layer, f) Ganglion cell layer, g) Optical fiber layer.

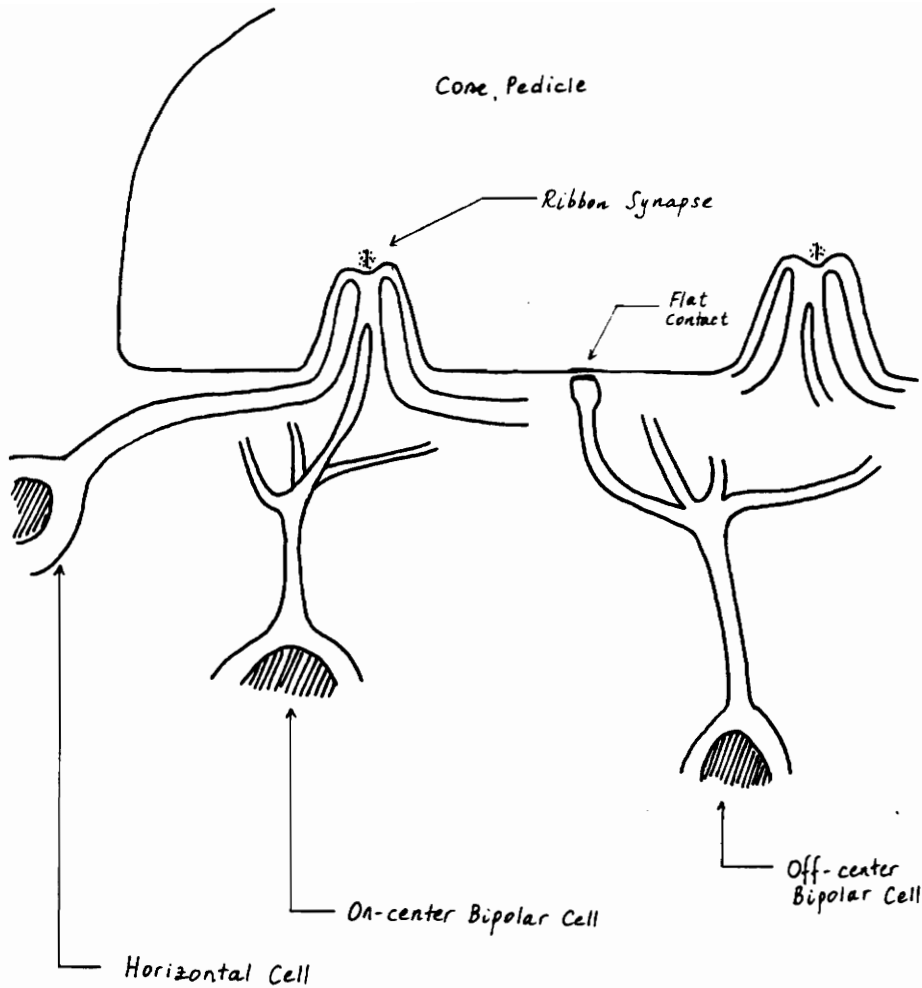


Figure 14. The cone pedicle and its synaptic connections: Invaginating on-center bipolar dendrites, flanking horizontal cell dendrites, and the flat contact of an off-center (diffuse) bipolar are shown. At the top of the invaginations there are ribbon synapses.

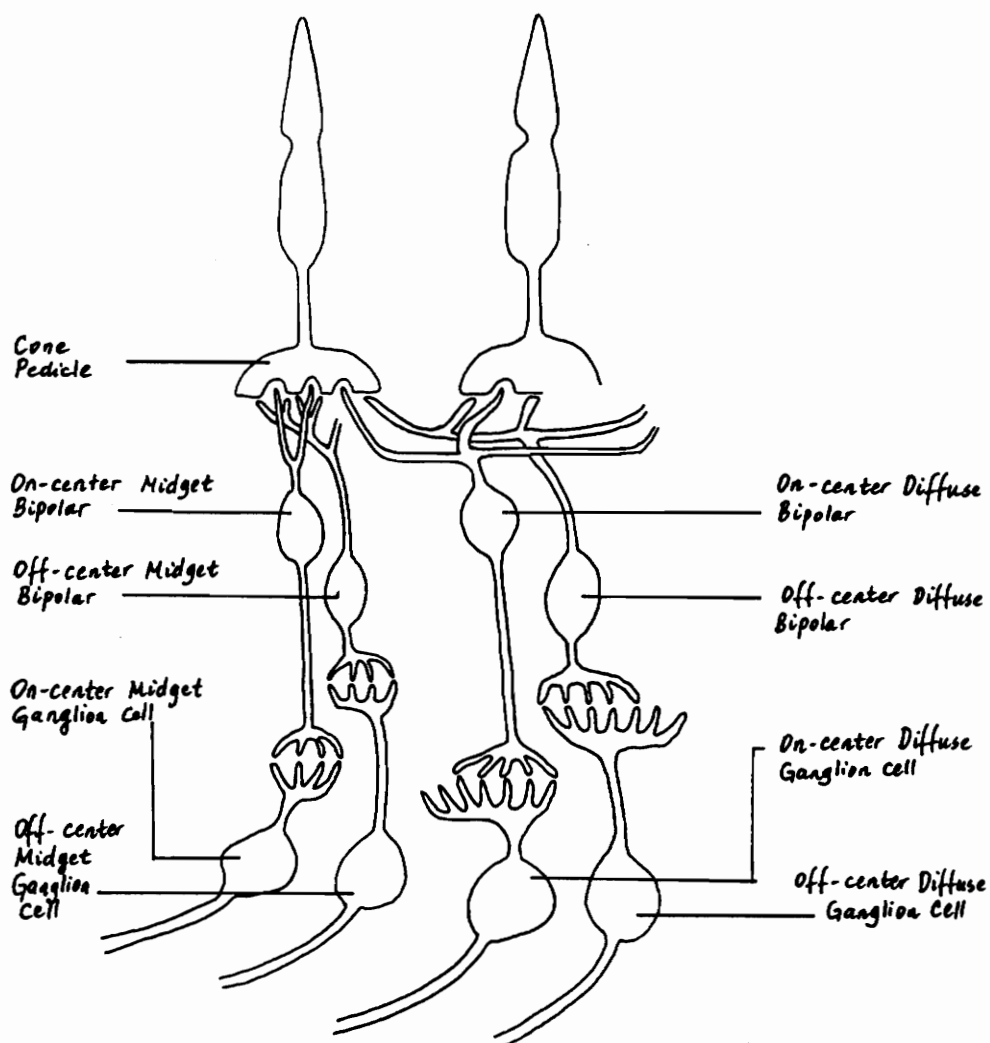


Figure 15. The four pathways from a cone: In this figure, the cone on the left is served by four bipolar, and subsequently, four ganglion cells. The two pathways on the left include two midget bipolar and two midget ganglion cells, which have small receptive fields and thus high-acuity; while the two pathways on the right include two diffuse bipolar and two diffuse ganglion cells which have much greater receptive fields, and therefore less acuity. For simplicity, horizontal and amacrine cells are not shown.

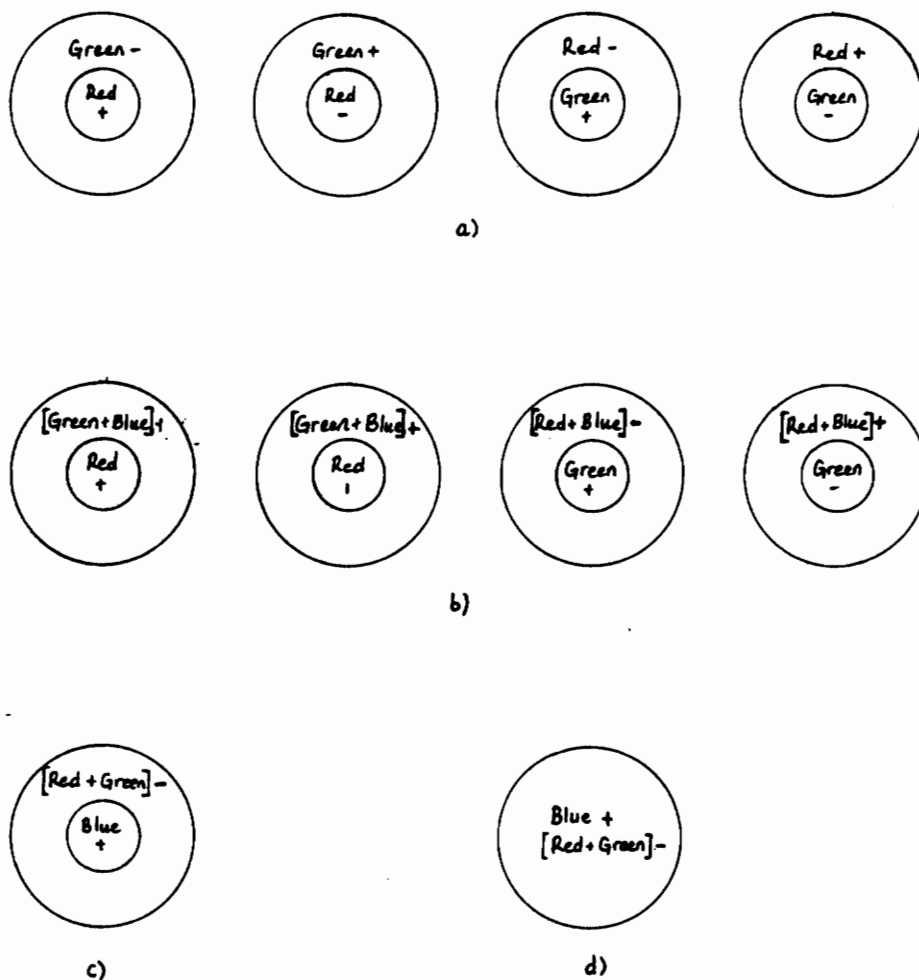


Figure 16. Receptive field structures of midget ganglion cells: a) Receptive fields of red-green midget ganglion cells, b) Receptive fields of a subtype of red-green midget ganglion cells whereby the surround also receives input from blue cones, c) Receptive fields of blue-yellow midget ganglion cells, d) Receptive fields of nonconcentric blue-yellow midget ganglion cells.

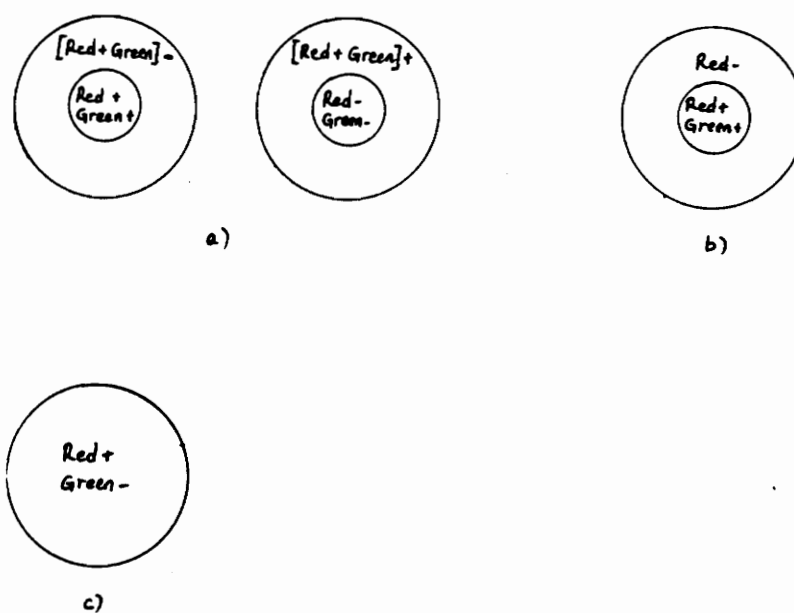


Figure 17. Receptive field structures of diffuse ganglion cells: a) Receptive fields of spectrally non-opponent, but spatially opponent diffuse ganglion cells, b) Receptive fields of spectrally and spatially opponent diffuse ganglion cells, c) Receptive fields of on-off diffuse ganglion cells.

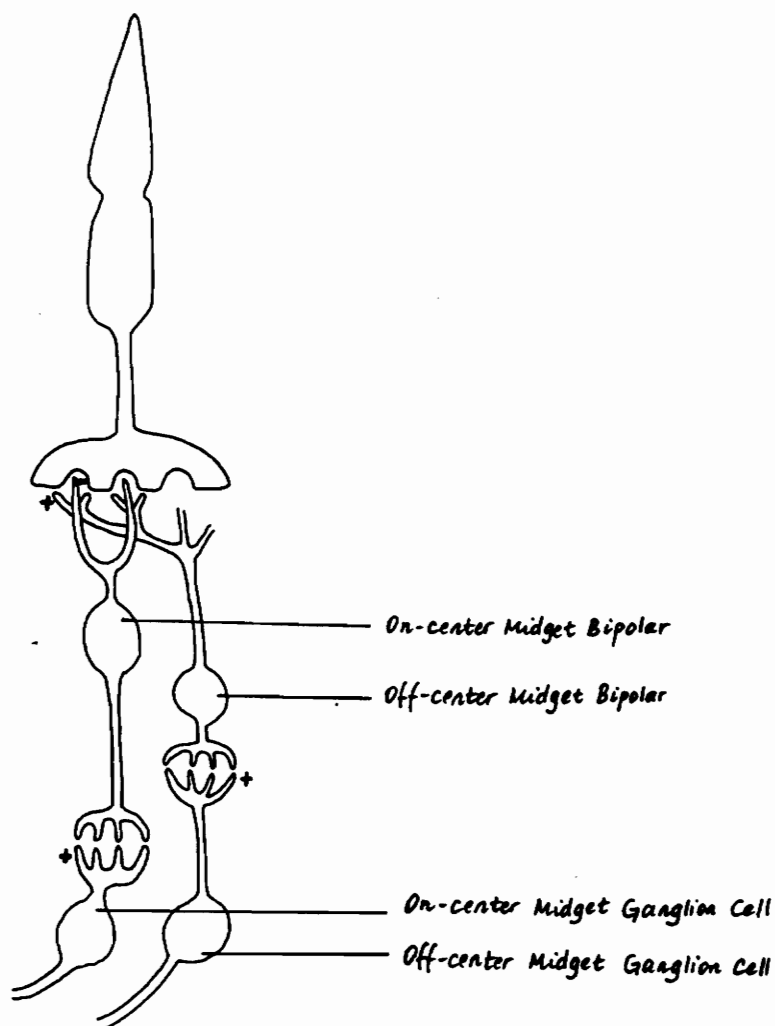


Figure 18. On and off-center pathways from a cone: Only midget bipolar and ganglion cells are shown. Sign-preserving synapses are marked with a (+) sign at the postsynaptic terminal; sign-inverting synapses with a (-) sign.

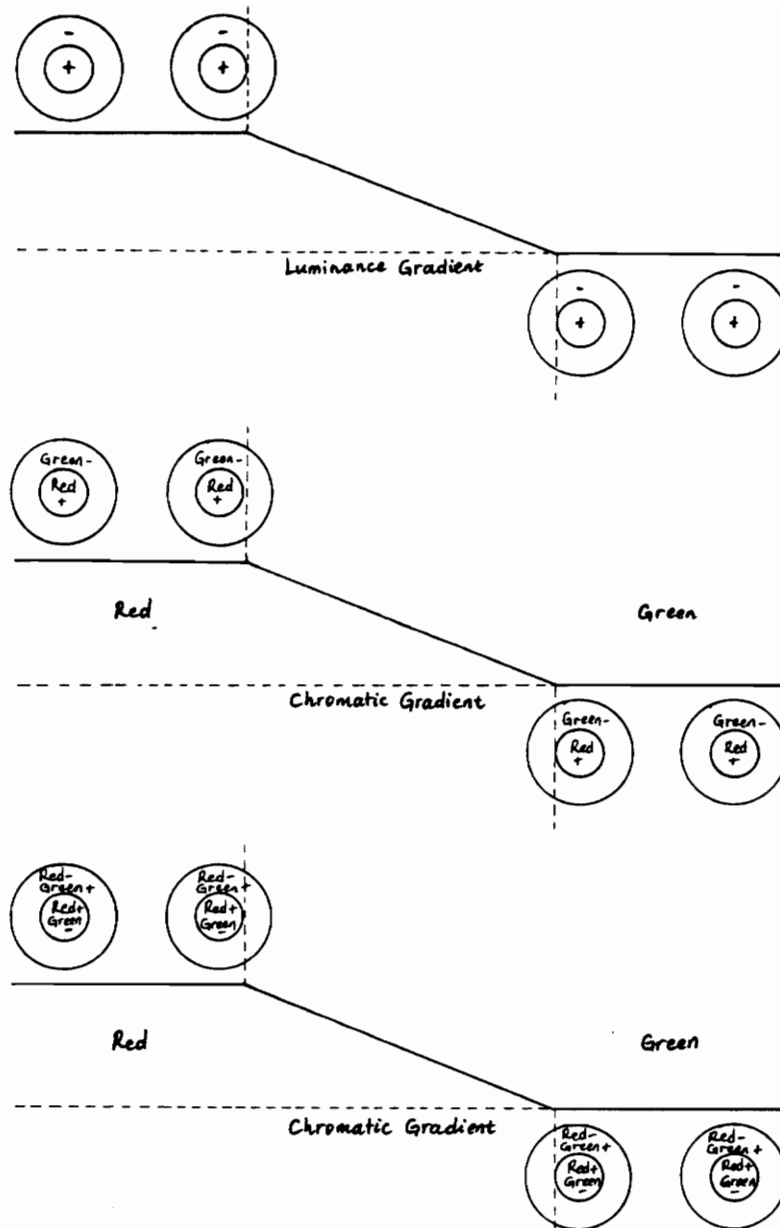


Figure 19. Looking at Mach Band formation: A lateral inhibition point of view a) With a luminance gradient and spatially opponent ganglion cells, b) With a chromatic (red-green) gradient and red-green midget ganglion cells (fig. 16a) (This one is unsuccessful), c) With a chromatic (red- green) gradient and double opponent cells in the cortex (fig. 20).

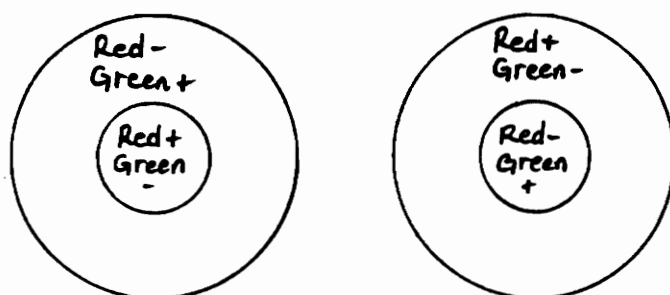


Figure 20. Double opponent cells in the visual cortex

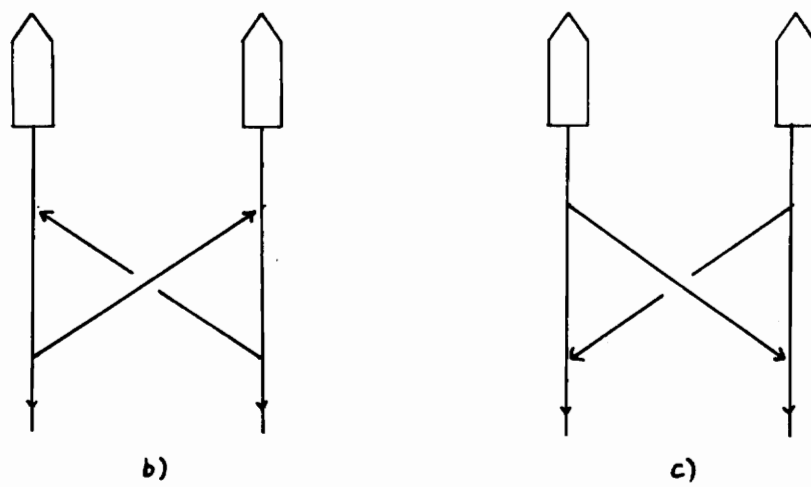


Figure 21. Recursive and non-recursive lateral inhibition: a) Recursive lateral inhibition, b) Non-recursive lateral inhibition. Both a) and b) are from Ratliff [1965].

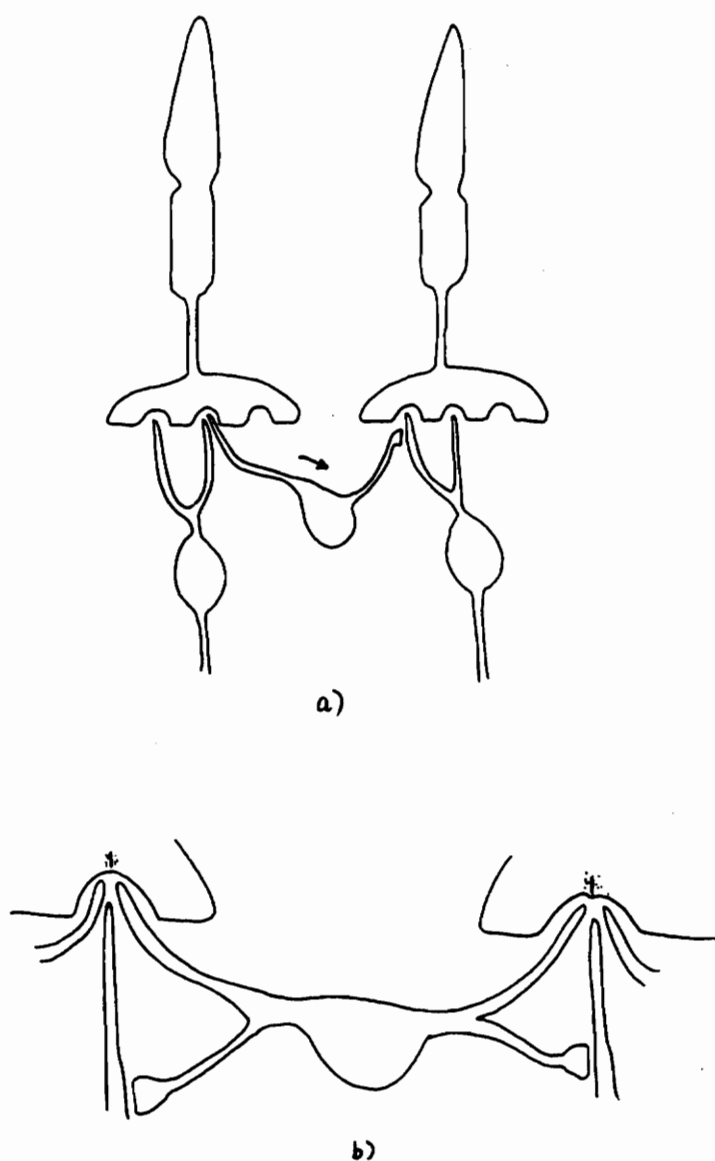


Figure 22. Types of lateral inhibition in the eye: a) Non-recursive lateral inhibition by a cone type-specific one-way horizontal cell in the model proposed by Gouras and Evers [1985], b) Lateral inhibition by a two-way horizontal cell where recursiveness, although seemingly quite unlikely, is not impossible.

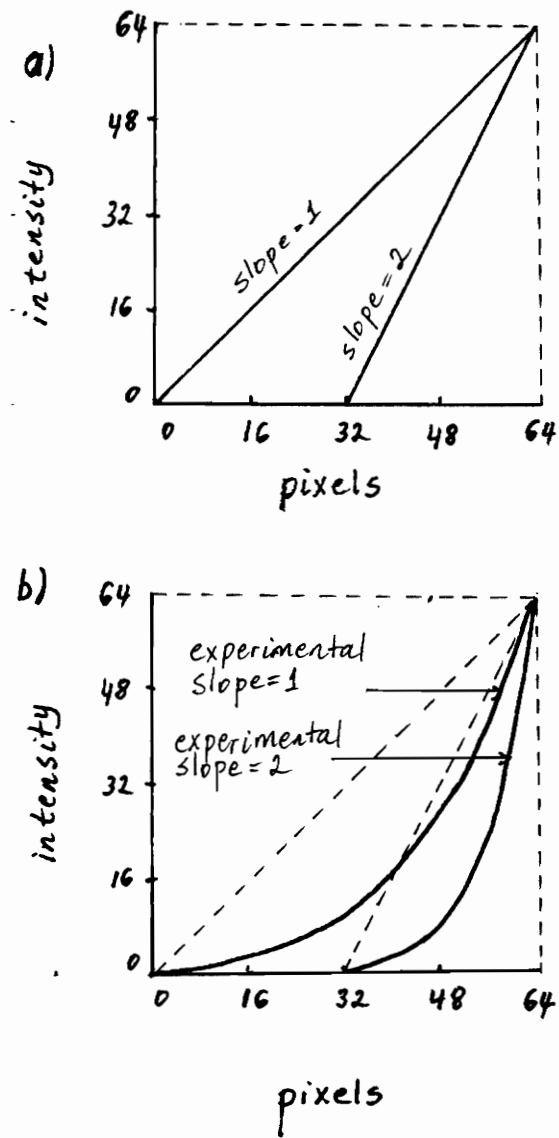


Figure 23. Slope in the linear and exponential gradients: a) Linear gradient, b) exponential gradient.

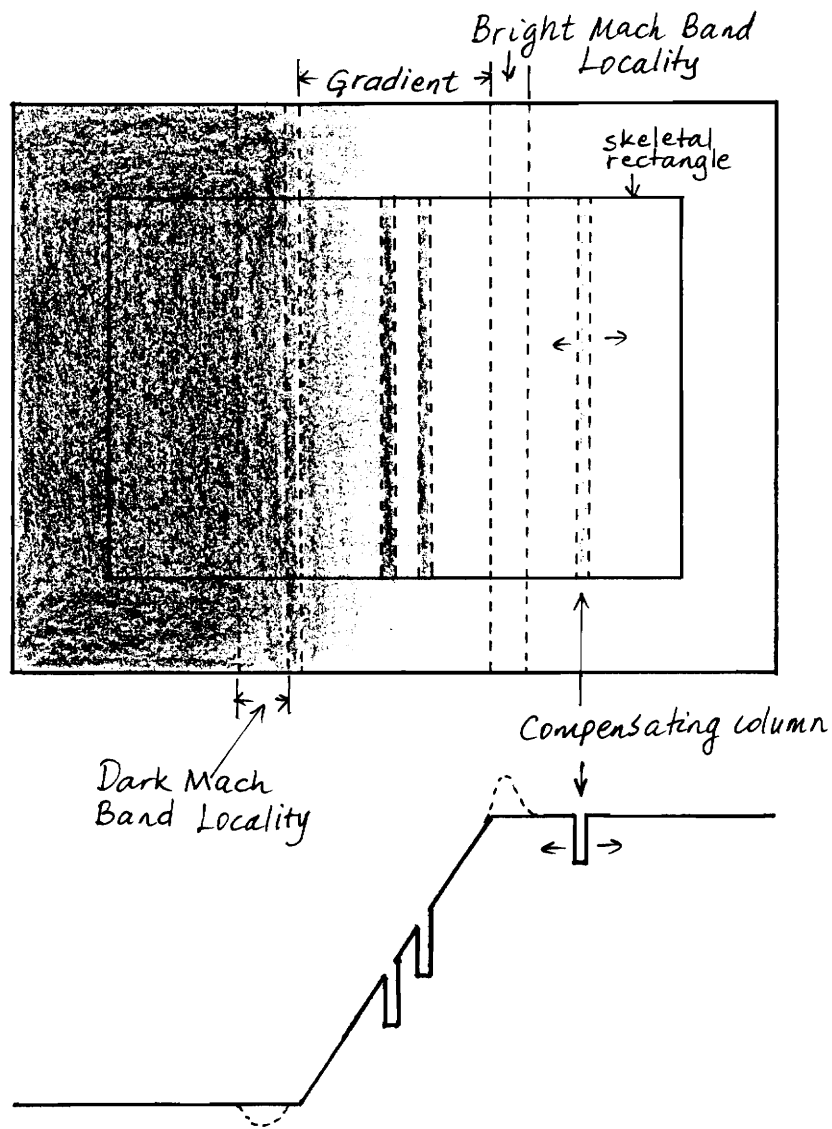


Figure 24. Mach Band compensation using program Mach-II: In the top part of this graph, the monitor screen is depicted. In the bottom section, a graph of the intensity distribution is given. (Both are not to scale) The user compensates with a thin, one pixel-wide column and the changes are permanent.

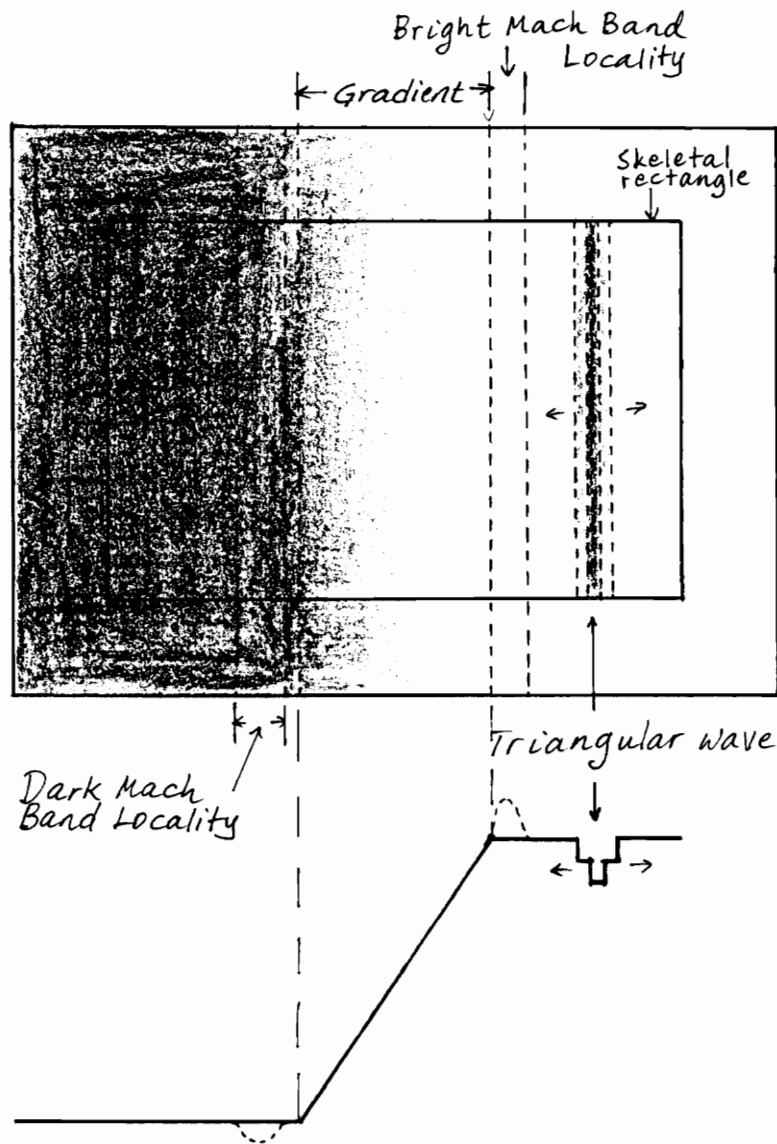


Figure 25. Mach Band compensation using program Mach-l2: In the top part of this graph, the monitor screen is depicted. In the bottom section, a graph of the intensity distribution is given. (Both are not to scale) The user compensates with a column or rectangle whose width is variable and whose intensity distribution is a jagged triangular wave. Although the rectangle can move, the changes are not permanent, and only the area currently within the rectangle is compensated.

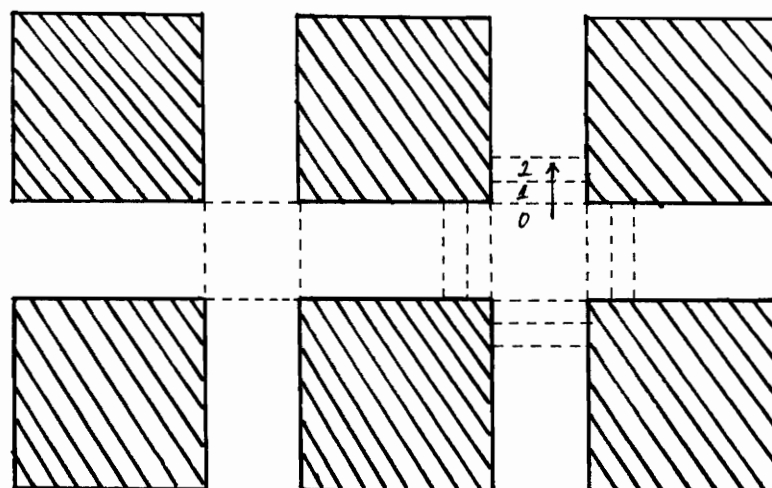


Figure 26. Hermann Grid compensation using program Her1: In the intersection at the left, the compensating intensity is directed towards the intersection alone. This is the 0-smear condition. In the intersection at the right, the compensating intensity is directed both to the intersection itself, and to a decreasing extent (in the direction of the arrow) to areas 1 and 2. Therefore, the compensation intensity is "smeared" towards the streets, and in this case, the smear is 2.

INTENSITY AND COMPENSATION PROFILES WHEN SLOPE IS 1

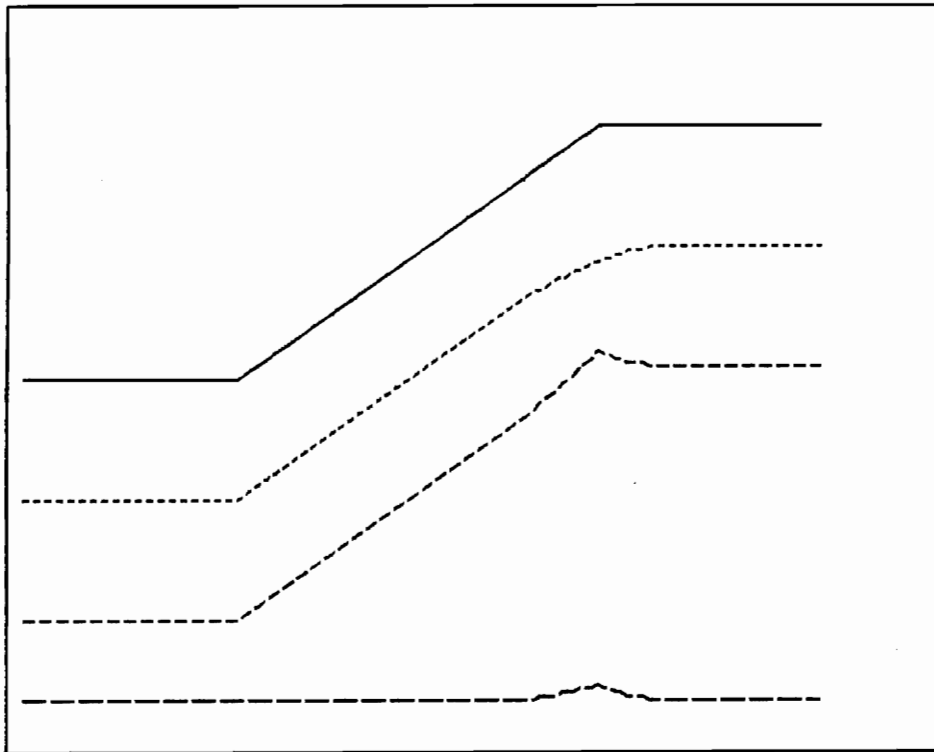


Figure 27. Intensity and compensation profiles - 1: This figure gives the intensity and compensation profiles for subject N. A. for a run of program mach-11. The slope was 1. From top to bottom, the depicted graphs belong to : (a) the original intensity distribution, (b) the compensated intensity distribution, (c) the intensity distribution resulting from the addition of the compensation to the original profile, (d) the compensation, i.e. the difference between the compensated and the original intensity distributions.

INTENSITY AND COMPENSATION PROFILES WHEN SLOPE IS 4

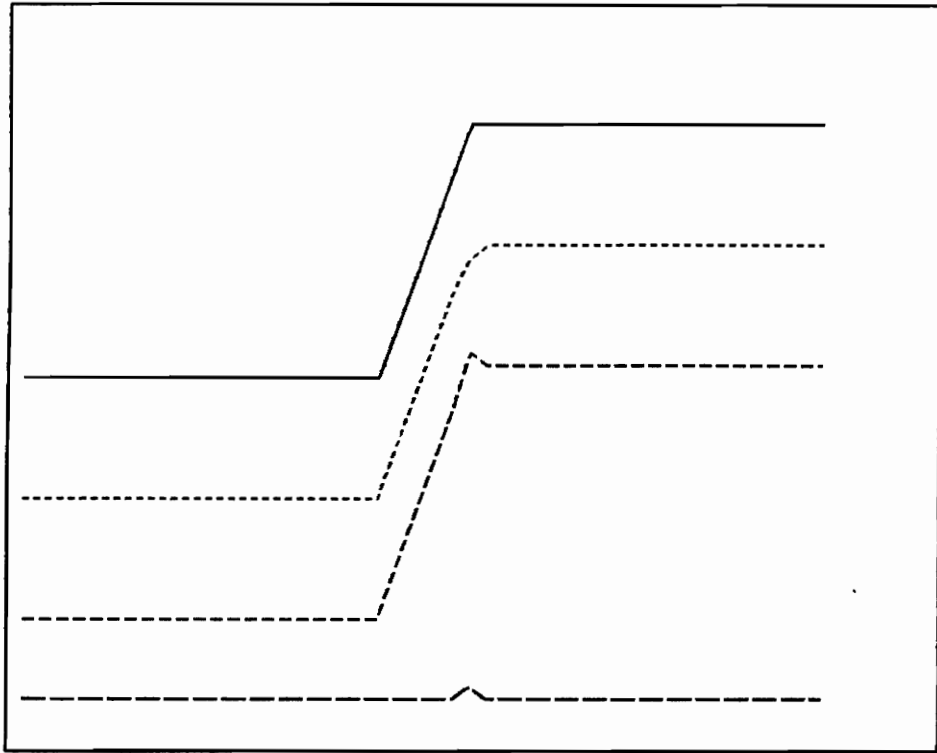


Figure 28. Intensity and compensation profiles - 2: This figure gives the intensity and compensation profiles for subject N. A. for a run of program mach-11. The slope was 4. From top to bottom, the depicted graphs belong to : (a) the original intensity distribution, (b) the compensated intensity distribution, (c) the intensity distribution resulting from the addition of the compensation to the original profile, (d) the compensation, i.e. the difference between the compensated and the original intensity distributions.

INTENSITY AND COMPENSATION PROFILES WHEN SLOPE IS 1

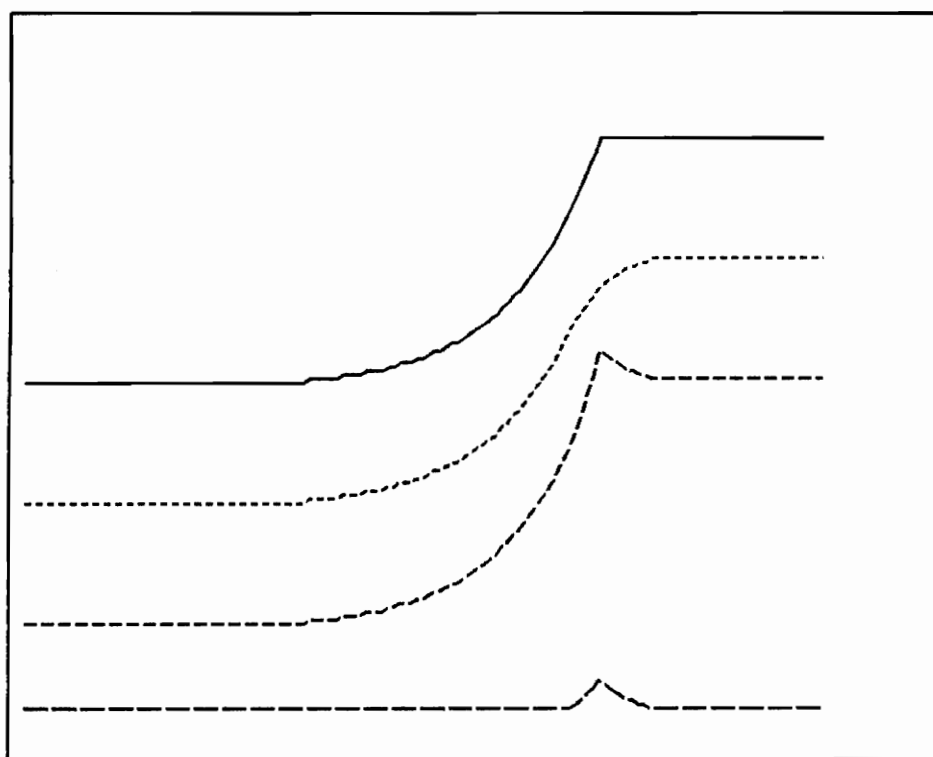


Figure 29. Intensity and compensation profiles - 3: This figure gives the intensity and compensation profiles for subject N. A. for a run of program mach-e1. The slope was 1. From top to bottom, the depicted graphs belong to : (a) the original intensity distribution, (b) the compensated intensity distribution, (c) the intensity distribution resulting from the addition of the compensation to the original profile, (d) the compensation, i.e. the difference between the compensated and the original intensity distributions.

INTENSITY AND COMPENSATION PROFILES WHEN SLOPE IS 4

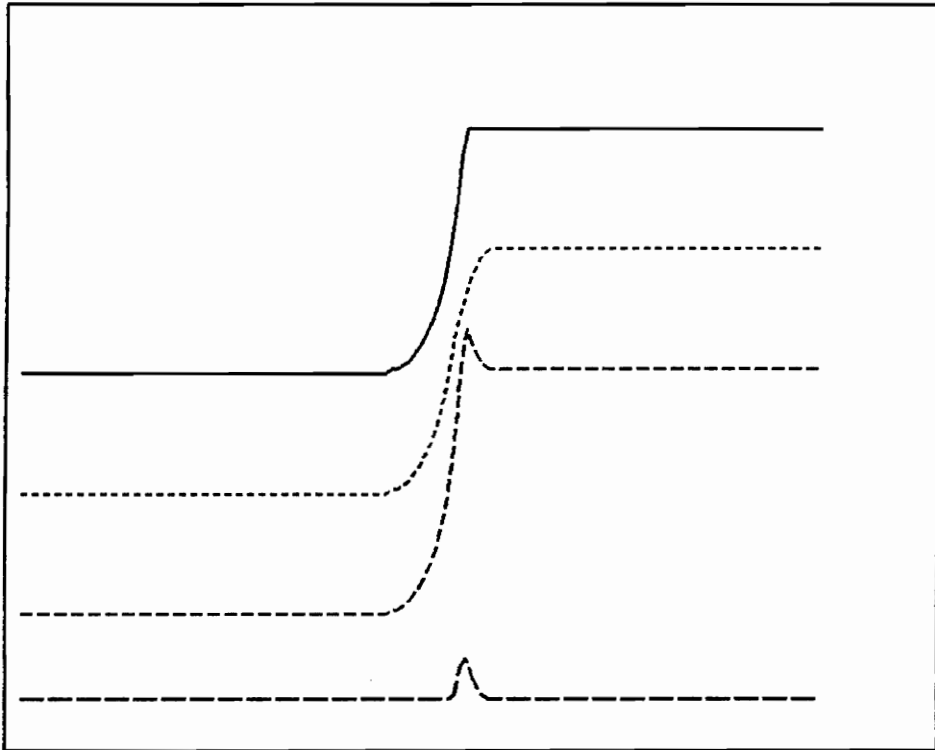


Figure 30. Intensity and compensation profiles - 4: This figure gives the intensity and compensation profiles for subject N. A. for a run of program mach-e1. The slope was 4. From top to bottom, the depicted graphs belong to : (a) the original intensity distribution, (b) the compensated intensity distribution, (c) the intensity distribution resulting from the addition of the compensation to the original profile, (d) the compensation, i.e. the difference between the compensated and the original intensity distributions.

INTENSITY AND COMPENSATION PROFILES WHEN SLOPE IS 1

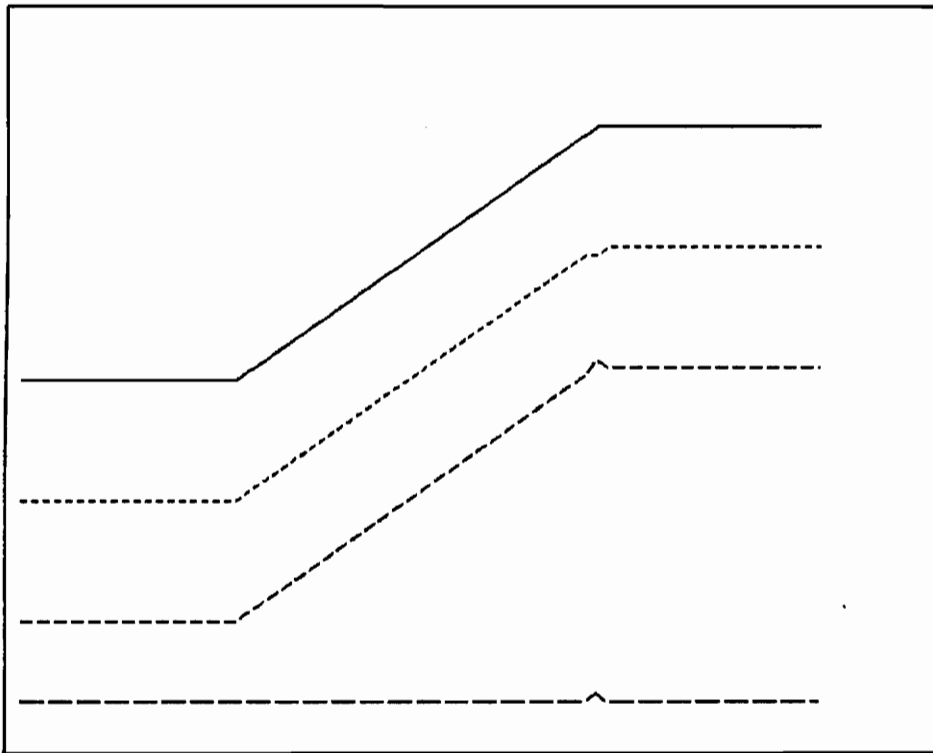


Figure 31. Intensity and compensation profiles - 5: This figure gives the intensity and compensation profiles for subject N. A. for a run of program mach-12 (which uses a triangular wave). The slope was 1. From top to bottom, the depicted graphs belong to : (a) the original intensity distribution, (b) the compensated intensity distribution, (c) the intensity distribution resulting from the addition of the compensation to the original profile, (d) the compensation, i.e. the difference between the compensated and the original intensity distributions.

INTENSITY AND COMPENSATION PROFILES WHEN SLOPE IS 4

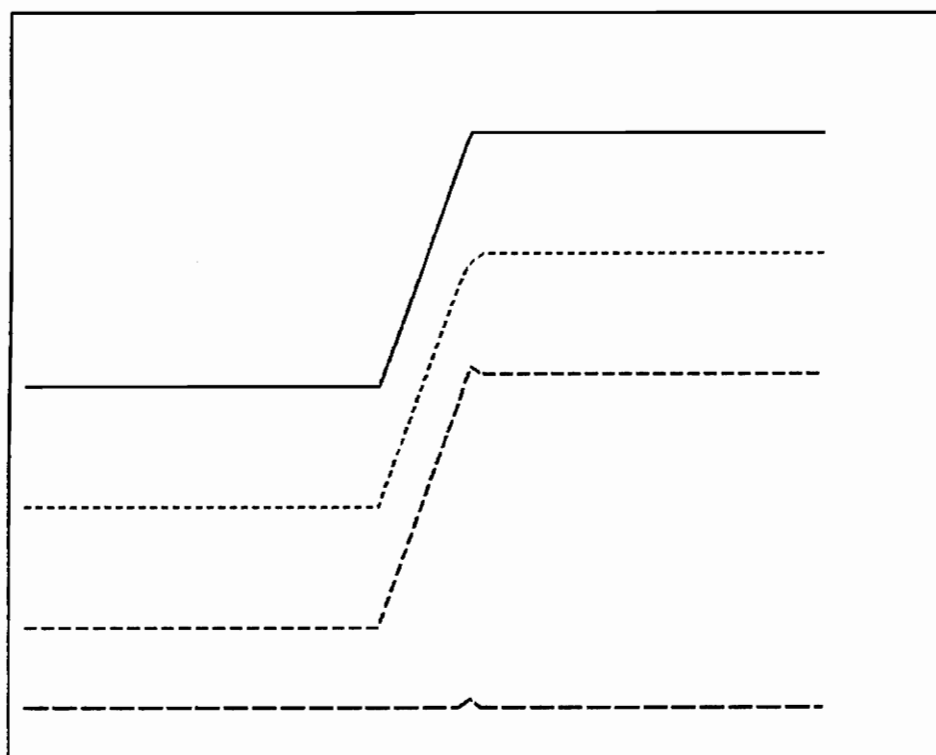


Figure 32. Intensity and compensation profiles - 6: This figure gives the intensity and compensation profiles for subject N. A. for a run of program mach-12 (which uses a triangular wave). The slope was 4. From top to bottom, the depicted graphs belong to : (a) the original intensity distribution, (b) the compensated intensity distribution, (c) the intensity distribution resulting from the addition of the compensation to the original profile, (d) the compensation, i.e. the difference between the compensated and the original intensity distributions.

INTENSITY AND COMPENSATION PROFILES WHEN SLOPE IS 1

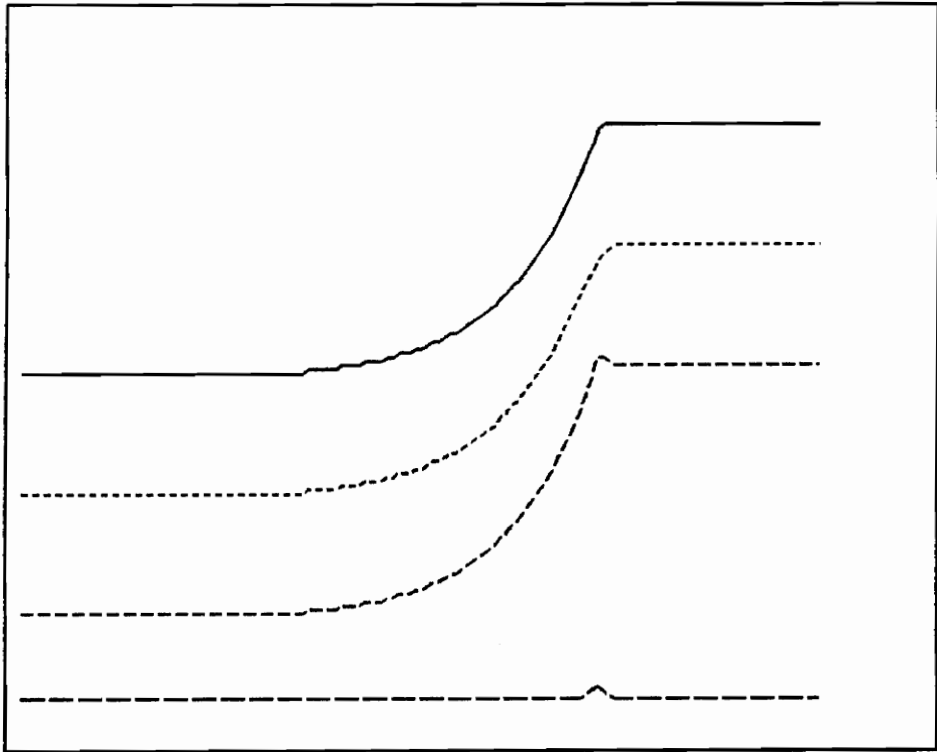


Figure 33. Intensity and compensation profiles - 7: This figure gives the intensity and compensation profiles for subject N. A. for a run of program mach-e2 (which uses a triangular wave). The slope was 1. From top to bottom, the depicted graphs belong to : (a) the original intensity distribution, (b) the compensated intensity distribution, (c) the intensity distribution resulting from the addition of the compensation to the original profile, (d) the compensation, i.e. the difference between the compensated and the original intensity distributions.

INTENSITY AND COMPENSATION PROFILES WHEN SLOPE IS 4

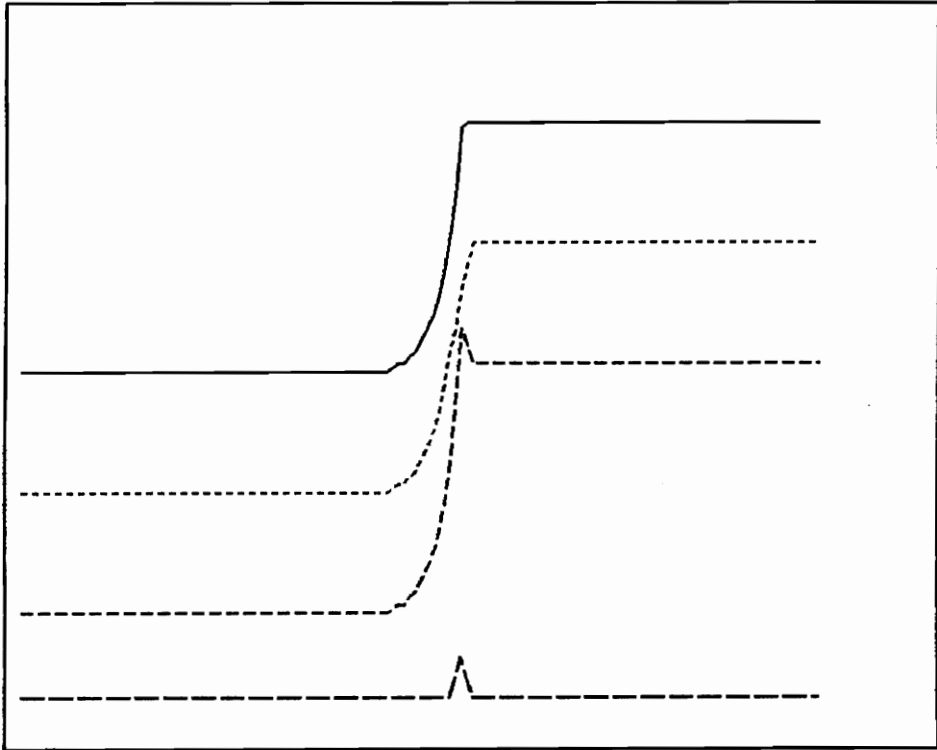


Figure 34. Intensity and compensation profiles - 8: This figure gives the intensity and compensation profiles for subject N. A. for a run of program mach-e2 (which uses a triangular wave). The slope was 4. From top to bottom, the depicted graphs belong to : (a) the original intensity distribution, (b) the compensated intensity distribution, (c) the intensity distribution resulting from the addition of the compensation to the original profile, (d) the compensation, i.e. the difference between the compensated and the original intensity distributions.

INTENSITY AND COMPENSATION PROFILES WHEN SLOPE IS 1

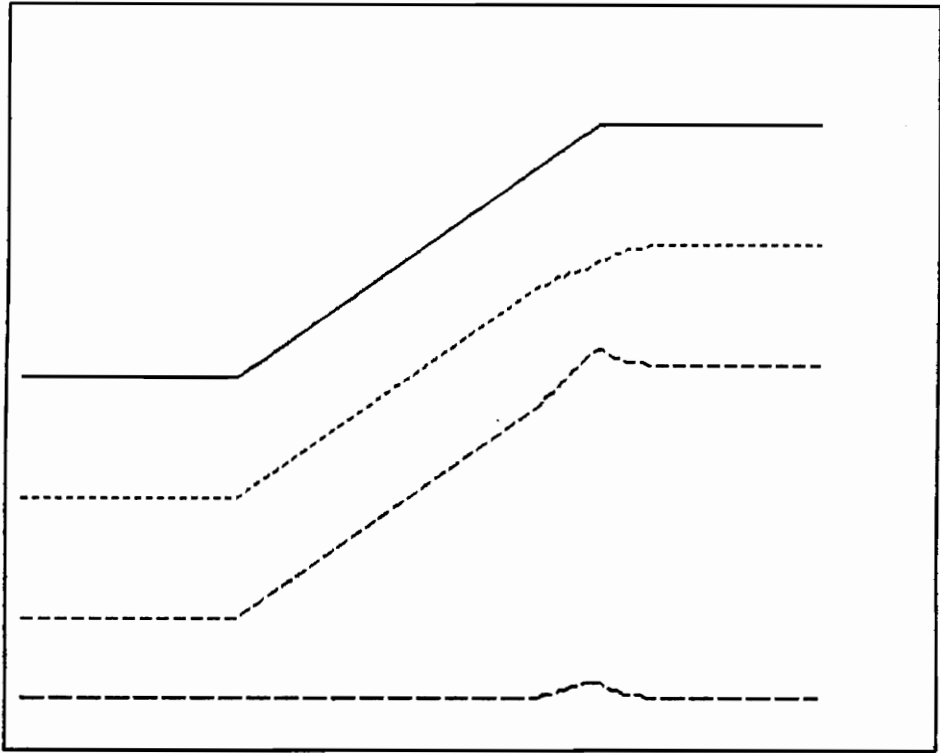


Figure 35. Intensity and compensation profiles - 9: This figure gives the intensity and compensation profiles for subject C. H. T. for a run of program mach-11. The slope was 1. From top to bottom, the depicted graphs belong to : (a) the original intensity distribution, (b) the compensated intensity distribution, (c) the intensity distribution resulting from the addition of the compensation to the original profile, (d) the compensation, i.e. the difference between the compensated and the original intensity distributions.

INTENSITY AND COMPENSATION PROFILES WHEN SLOPE IS 1

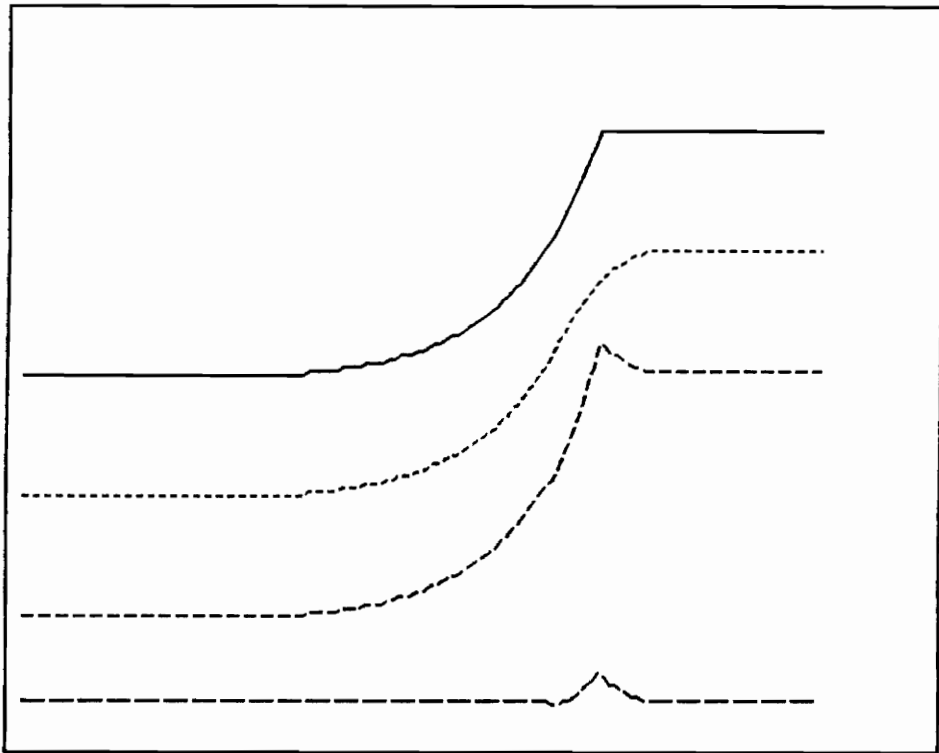


Figure 36. Intensity and compensation profiles - 10: This figure gives the intensity and compensation profiles for subject C. H. T. for a run of program mach-e1. The slope was 1. From top to bottom, the depicted graphs belong to : (a) the original intensity distribution, (b) the compensated intensity distribution, (c) the intensity distribution resulting from the addition of the compensation to the original profile, (d) the compensation, i.e. the difference between the compensated and the original intensity distributions.

INTENSITY AND COMPENSATION PROFILES WHEN SLOPE IS 1

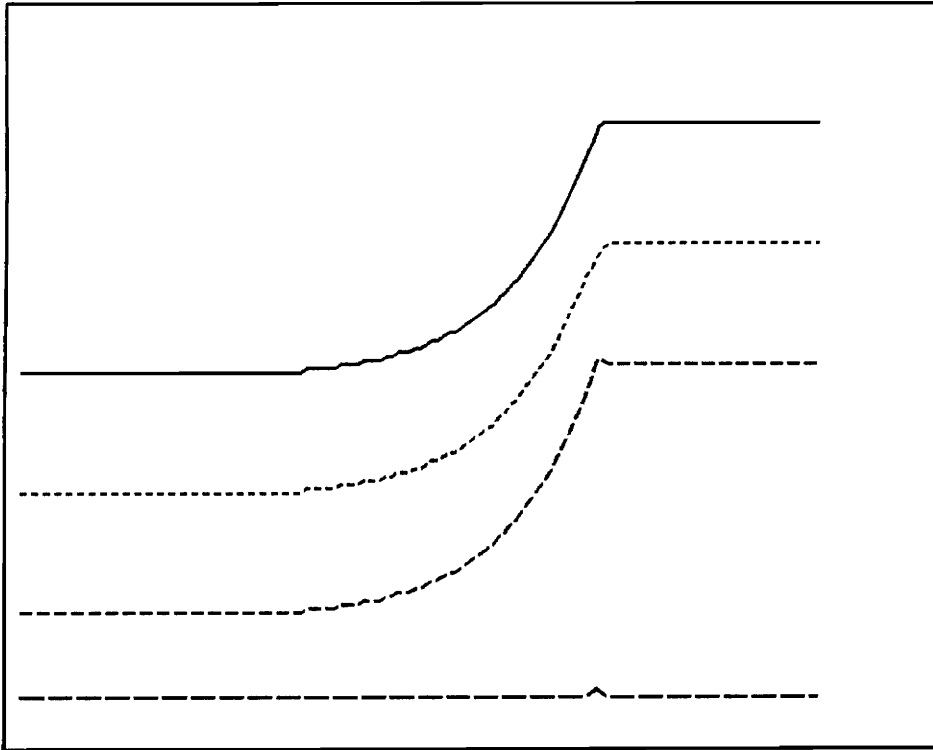


Figure 37. Intensity and compensation profiles - 11: This figure gives the intensity and compensation profiles for subject C. H. T. for a run of program mach-e2 (which uses a triangular wave). The slope was 1. From top to bottom, the depicted graphs belong to : (a) the original intensity distribution, (b) the compensated intensity distribution, (c) the intensity distribution resulting from the addition of the compensation to the original profile, (d) the compensation, i.e. the difference between the compensated and the original intensity distributions.

INTENSITY AND COMPENSATION PROFILES WHEN SLOPE IS 4

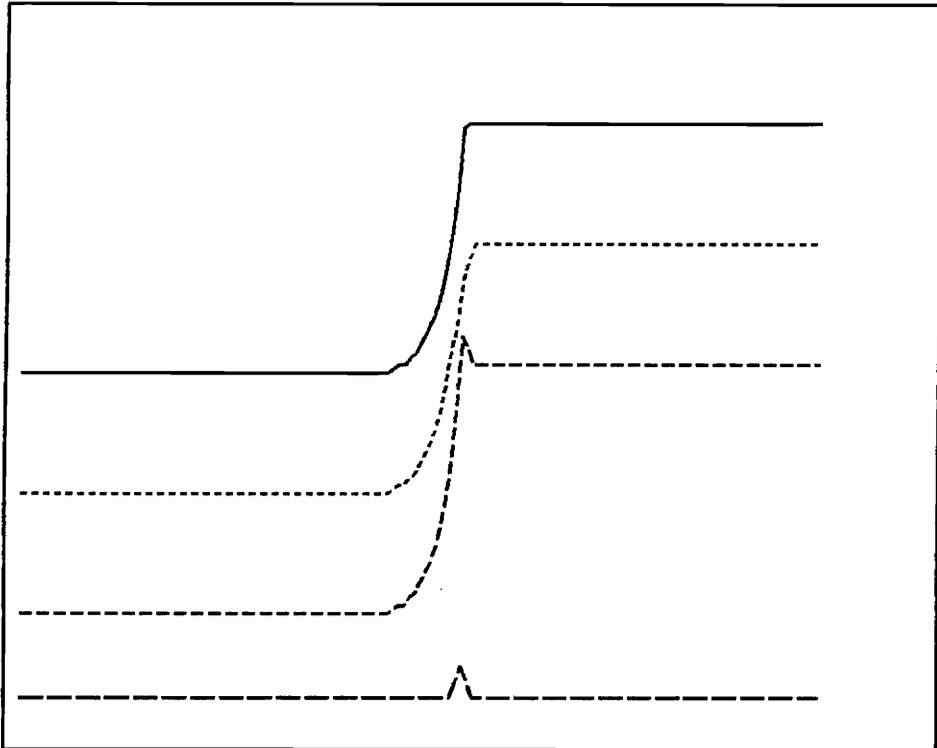


Figure 38. Intensity and compensation profiles - 12: This figure gives the intensity and compensation profiles for subject C. H. T. for a run of program mach-e2 (which uses a triangular wave). The slope was 4. From top to bottom, the depicted graphs belong to : (a) the original intensity distribution, (b) the compensated intensity distribution, (c) the intensity distribution resulting from the addition of the compensation to the original profile, (d) the compensation, i.e. the difference between the compensated and the original intensity distributions.

7. Bibliography

1. Arend L. and Reeves A. [1986], "Simultaneous color constancy," *Journal of the Optical Society of America A*, vol. 3, no. 10, pp. 1743-1751.
2. Bennett P. J. and Banks M. S. [1987], "Sensitivity loss in odd-symmetric mechanisms and phase anomalies in peripheral vision," *Nature*, vol. 326, pp. 873-876.
3. Bijl P., Koenderink J. J., and Toet A. [1989], "Visibility of blobs with a Gaussian luminance profile," *Vision Research*, vol. 29, pp. 447-456.
4. Boynton R. M. [1988], "Color Vision," *Annual Review of Psychology*, vol. 39, pp. 69-100.
5. Bradley A., Switkes E., and De Valois K. [1988], "Orientation and spatial frequency selectivity of adaptation to color and luminance gratings," *Vision Research*, vol. 28, pp. 841-856.
6. Cowan W. B. [1983], "Discreteness artifacts in raster display systems," in *Colour Vision*, ed. J. D. Mollon and L. T. Sharpe, London: Academic Press, pp. 145-153.
7. Cowan W. B. and Rowell N. [1986], "On the gun independence and phosphor constancy of colour video monitors," *Color Research and Application*, vol. 11, supplement, pp. s34-s38.

8. Dowling J. E. and Dubin M. W., "The vertebrate retina," in *Handbook of Physiology*, American Physiological Society, pp. 317-339.
9. Ercoles-Guzzoni A. M. and Fiorentini A. [1958], "Simultaneous contrast effect produced by non-uniform coloured fields," *Atti Fond. Giorgio Ronchi*, vol. 13, pp. 135-144.
10. Ercoles A. M. and Fiorentini A. [1959], "Visibility of the Mach bands as a function of field luminance," *Atti Fond. Giorgio Ronchi*, vol. 14, pp. 230-235.
11. Field D. J. and Nachmias J. [1984], "Phase reversal discrimination," *Vision Research*, vol. 24, no. 4, pp. 333-340.
12. Fiorentini A. [1957], "Foveal and extrafoveal contrast threshold at a point of a nonuniform field," *Atti Fond. Giorgio Ronchi*, vol. 12, pp. 180-186.
13. Fiorentini A. [1972], "Mach band phenomena," in *Handbook of Sensory Physiology*, vol. VII/4, *Visual Psychophysics*, ed. D. Jameson and L. M. Hurvich, Berlin: Springer-Verlag, pp. 188-201.
14. Fiorentini A., Jeanne M., and Toraldo di Francia G. [1955], "Measurements of differential threshold in the presence of spatial illumination gradient," *Atti Fond. Giorgio Ronchi*, vol. 10, pp. 371-379.
15. Fiorentini A. and Radici T. [1957], "Binocular measurements of brightness on a field presenting a luminance gradient," *Atti Fond. Giorgio Ronchi*, vol. 12, pp. 453-461.
16. Fiorentini A. and Radici T. [1958], "Brightness, width and position of Mach bands as a function of the rate of variation of the luminance gradient," *Atti Fond. Giorgio Ronchi*, vol. 13, pp. 145-155.

17. Fry G. A. [1948], "Mechanisms subserving simultaneous brightness contrast," *American Journal of Optometry and Archives of the American Academy of Optometry*, vol. 25, pp. 162-178.
18. Gorea A. and Papathomas T. V. [1989], "Motion processing by chromatic and achromatic pathways," *Journal of the Optical Society of America A*, vol. 6, p. 590.
19. Gouras P. and Evers H. U. [1985], "The neurocircuitry of primate retina," in *Neurocircuitry of the Retina - A Cajal Memorial*, ed. A. Gallego and P. Gouras, New York: Elsevier, pp. 233-244.
20. Granrath D. J. [1981], "The role of human visual models in image processing," *Proceedings of the IEEE*, vol. 69, no. 5, pp. 552-561.
21. Hall C. F. and Hall E. L. [1977], "A nonlinear model for the spatial characteristics of the human visual system," *IEEE Transactions in Systems, Man, and Cybernetics*, vol. SMC-7, no. 3, pp. 161-170.
22. Hartline H. K. and Ratliff F. [1954], "Spatial summation of inhibitory influences in the eye of *Limulus*," *Science*, vol. 120, p. 781.
23. Hartline H. K. and Ratliff F. [1957], "Inhibitory interaction of receptor units in the eye of *Limulus*," *Journal of General Physiology*, vol. 40, pp. 357-376.
24. Hartline H. K., Wagner H. G., and Ratliff F. [1956], "Inhibition in the eye of *Limulus*," *Journal of General Physiology*, vol. 39, pp. 651-673.
25. Hartline H. K., Ratliff F., and Miller W. H. [1961], "Inhibitory interaction in the retina and its significance in vision," in *Nervous Inhibition*, ed. E. Florey, New York: Pergamon Press, pp. 241-284.

26. Howett G. L. [1986], "The coming redefinition of photometry," *Journal of the Illuminating Engineering Society*, vol. 15, pp. 5-18.
27. Koffka K. and Harrower M. R. [1931], "Color and organization, Part I," *Psychologische Forschung*, vol. 15, pp. 146-192.
28. Lavin E. and Costall A. [1978], "Detection thresholds of the Hermann Grid illusion," *Vision Research*, vol. 18, pp. 1061-1062.
29. Letellier P., Nadler M., and Abramatic J. [1985], "The Telesign project," *Proceedings of the IEEE*, vol. 73, no. 4, 813-827.
30. Levine J., Spillmann L., and Wolf E. [1980], "Saturation enhancement in colored Hermann Grids varying only in chroma," *Vision Research*, vol. 20, pp. 307-313.
31. Lindsay D. T. and Teller D. Y. [1989], "Influence of variations in edge blur on minimally distinct border judgements: a theoretical and empirical investigation," *Journal of the Optical Society of America A*, vol. 6, p. 446.
32. Lowry E. M. and de Palma J. J. [1961], "Sine wave response of the visual system, 1: The Mach phenomenon," *Journal of the Optical Society of America*, vol. 51, pp. 740-746.
33. Mach E. [1865], "Über die Wirkung der raumlichen Vertheilung des Lichtreizes auf die Netzhaut, I," *Sitzungsberichte der mathematisch-naturwissenschaftlichen Classe der kaiserlichen Akademie der Wissenschaften*, vol. 52, pp. 303-322.
34. MacLeod D. I. A. [1986], "Computer controlled color displays in vision research: Possibilities and problems - Extended abstract," *Color Research and Application*, vol. 11, supplement, pp. s45-s46.

35. MacLeod D. I. A. and Boynton R. M. [1979], "Chromaticity diagram showing cone excitation by stimuli of equal luminance," *Journal of the Optical Society of America*, vol. 69, no. 8, pp. 1183-1186.
36. Marr D. [1976], "Early processing of visual information," *Philosophical Transactions of the Royal Society of London B*, vol. 275, pp. 483-524.
37. Marr D. [1982], *Vision: A computational investigation into the human representation and processing of visual information*, San Francisco: W. H. Freeman and company.
38. Marr D. and Hildreth E. [1980], "Theory of edge detection," *Proceedings of the Royal Society of London B*, vol. 207, pp. 187-217.
39. Marr D. and Poggio T. [1979], "A computational theory of human stereo vision," *Proceedings of the Royal Society of London B*, vol. 204, pp. 301-328.
40. Marr D., Poggio T., and Ullman S. [1979], "Bandpass channels, zero-crossings, and early visual information processing," *Journal of the Optical Society of America*, vol. 69, pp. 914-916.
41. Matthews M. L. [1966], "Appearance of Mach bands for short durations and at sharply focused contours," *Journal of the Optical Society of America*, vol. 56, pp. 1401-1402.
42. Matthews M. L. [1967], "Mach band increment thresholds and the mechanisms of color vision," *Journal of the Optical Society of America*, vol. 57, no. 8, pp. 1033-1036.
43. McCarter A. [1979], "Chromatic induction effects in the Hermann Grid illusion," *Perception*, vol. 8, pp. 105-114.
44. McCollough C. [1955], "The variation in width and position of Mach bands as a function of luminance," *Journal of Experimental Psychology*, vol. 49, pp. 141-152.

45. McDougall W. [1903], "Intensification of visual sensation by smoothly gradual contrast," *Proceedings of the Physiological Society*, vol. 1, pp. 19-21.
46. Michael C. R. [1978], "Color vision mechanisms in monkey striate cortex: Dual-opponent cells with concentric receptive fields," *Journal of Neurophysiology*, vol. 41, pp. 572-578.
47. Morrone M. C., Ross J., Burr D. C., and Owens R. [1986], "Mach bands are phase dependent," *Nature*, vol. 324, pp. 250-253.
48. Oehler R. and Spillmann L. [1981], "Illusory color changes in Hermann Grids varying only in hue," *Vision Research*, vol. 21, pp. 527-541.
49. Paradiso M. A., Carney T., and Freeman R. D. [1989], "Cortical Processing of hyperacuity tasks," *Vision Research*, vol. 29, pp. 247-254.
50. Pease P. L. [1978], "On color mach bands," *Vision Research*, vol. 18, pp. 751-755.
51. Ratliff F. [1965], *Mach Bands: Quantitative studies on neural networks in the retina*, San Francisco: Holden-Day.
52. Ratliff F. [1984], "Why Mach bands are not seen at the edges of a step," *Vision Research*, vol. 24, no. 2, pp. 163-165.
53. Ratliff F. and Hartline H. K. [1959], "The response of *Limulus* optic nerve fibers to patterns of illumination on the receptor mosaic," *Journal of General Physiology*, vol. 42, pp. 1241-1255.
54. Ratliff F., Milkman N., and Bennert N. [1983], "Attenuation of Mach bands by adjacent stimuli," *Proceedings of the National Academy of Sciences USA*, vol. 80, pp. 4554-4558.
55. Rodieck R. W. [1988], "The primate retina," in *Comparative Primate Biology*, vol. 4: *Neurosciences*, eds. H. D. Steklis and J. Erwin, Alan R. Liss, Inc. 203-278.

56. Rodieck R. W. [1983], "Raster based colour simulators," in *Colour Vision*, ed. J. D. Mollon and L. T. Sharpe, London: Academic Press, pp. 131-146.
57. Schreiber W. F. [1978], "Image processing for quality improvement," *Proceedings of the IEEE*, vol. 66, no. 12, pp. 1640-1651.
58. Srinivasan M. V., Laughlin S. B., and Dubs A. [1982], "Predictive coding: A fresh view of inhibition in the retina," *Proceedings of the Royal Society of London B*, vol. 216, pp. 427-459.
59. Stockham T. G., Jr. [1972], "Image processing in the context of a visual model," *Proceedings of the IEEE*, vol. 60, no. 7, pp. 828-842.
60. Taylor J. M. and Murch G. M. [1986], "The effective use of color in visual displays: Text and graphics applications," *Color Research and Application*, vol. 11, supplement, pp. s3-s9.
61. Thomas J. P. [1965], "Threshold measurement of Mach bands," *Journal of the Optical Society of America*, vol. 55, no. 5, pp. 521-524.
62. Thomas J. P. [1966], "Brightness variations in stimuli with ramp-like contours," *Journal of the Optical Society of America*, vol. 56, no. 2, pp. 232-242.
63. Thouless R. M. [1922], "Some observations on contrast effects in graded discs," *British Journal of Psychology*, vol. 13, pp. 301-307.
64. Troscianko T. [1982a], "A stereoscopic presentation of the Hermann Grid," *Vision Research*, vol. 22, pp. 485-489.
65. Troscianko T. [1982b], "A given visual field location has a wide range of perceptive field sizes," *Vision Research*, vol. 22, pp. 1363-1369.

66. von Bekesy G. [1967], "Mach band type lateral inhibition in different sense organs." *The Journal of General Physiology*, vol. 50, pp. 519-532.
67. von Bekesy G. [1968a], "Brightness distribution across the Mach bands measured with flicker photometry and the linearity of sensory nervous interaction," *Journal of the Optical society of America*, vol. 58, no. 1, pp. 1-8.
68. von Bekesy G. [1968b], "Mach- and Hering-type lateral inhibition in vision," *Vision Research*, vol. 8, pp. 1483-1499.
69. von Bekesy G. [1968c], "Location of maxima and minima in sensation patterns influenced by lateral inhibition," *Journal of Applied Physiology*, vol. 25, no. 2, pp. 200-206.
70. von Bekesy G. [1972], "Mach bands measured by a compensation method," *Vision Research*, vol. 12, pp. 1485-1497.
71. Ware C. and Cowan W. B. [1987], "Chromatic Mach bands: Behavioral evidence for lateral inhibition in human color vision," *Perception and Psychophysics*, vol. 41, pp. 173-178.
72. Watrasiewicz B. M. [1966], "Some factors affecting the appearance of the Mach bands," *Journal of the Optical Society of America*, vol. 56, no. 4, pp. 499-503.
73. Watt R. J. and Morgan M. J. [1983], "The recognition and representation of edge blur: Evidence for spatial primitives in human vision," *Vision Research*, vol. 23, no. 12, pp. 1465-1477.
74. Whiteside J. A. and Davidson M. L. [1971], "Symmetrical appearance of the bright and dark Mach bands from an exponential illumination gradient," *Journal of the Optical Society of America*, vol. 61, no. 7, pp. 958-961.

75. Wilson H. R. and Bergen J. R. [1979], "A four mechanism model for spatial vision," *Vision Research*, vol. 19, pp. 19-32.
76. Wilson H. R. and Giese S. C. [1977], "Threshold visibility of frequency gradient patterns," *Vision Research*, vol. 17, pp. 1177-1190.
77. Wolfe J. M. [1984], "Global factors in the Hermann Grid illusion," *Perception*, vol.13, pp. 33-40.

Appendix 1. Physiological Summary

A.1.1 The structure of the primate retina

The mammalian retina is a highly organized structure where morphologically (and functionally) different layers have been identified even during early years of cytological research. Without going into great detail, this layerization (Fig. 13) can be summarized as follows:

Most distally located (with respect to the center of the eye) are the photoreceptors; rods and cones, their segments containing photosensitive membranes forming the more distal outer segment layer, and their nuclei forming the outer nuclear layer. The photoreceptors contact other cells in the outer plexiform layer (OPL) where they form basal structures called pedicles in cones and spherules in rods. These structures serve as a host to a complex neural network which also includes the dendrites of horizontal and bipolar cells. Horizontal cell functioning is confined to this region and has been identified to be inhibitory. Bipolar cells, on the other hand, form a conducting pathway to ganglion cells which further transmit the information to the lateral geniculate nucleus in the thalamus and the superior colliculus. The nuclei of both horizontal and bipolar cells lie in the inner nuclear layer. The second neural network in the retina is the inner plexiform layer (IPL) where bi-

polar axons meet ganglion cell dendrites. This layer is the least understood part of the retina in large part due to many types of amacrine cells which contact both bipolar dendrites and ganglion cells. The first transient responses in retina emerge from this layer. The inner plexiform layer can be morphologically divided into 2 distinct layers: some bipolar-ganglion interactions occur in the inner part of the IPL, and some in the outer part. As will be explained later, this also has functional significance. The next layer in the retina contains the nuclei of the ganglion cells; this is followed by the optical fiber layer made up of axons of ganglion cells.

Since here we are concerned with photopic vision, those aspects of retinal architecture pertaining to scotopic (rod) vision are skipped.

Each cone pedicle in the OPL has two major sites of contact with other cells. One is through invaginations: Foveal cones have about 12, parafoveal cones have about 24 invaginations.

Each invagination in the OPL is penetrated by 3 dendritic processes called a "triad" (Fig. 14). The central process is the dendrite of the cone bipolar; it is accompanied on two flanks by horizontal cell dendrites. At the top of the invagination the cone pedicle has a ribbon synapse. All three dendritic processes are postsynaptic to the cone. Horizontal cell dendrites, however, also make conventional synapses with bipolar cells where they are presynaptic (and inhibitory). A single horizontal cell contacts at least two and sometimes all three cone types.

The second site of contact is at the base of the pedicles called basal or flat contact, and is made to other bipolar cell dendrites. Although no synaptic vesicles have been seen, it is proved to be synaptic.

The bipolars differ according to the site of contact. Those which contact the cone pedicle at the invaginations are on-center, while the ones which make the basal contact with the cones are off-center. There are, moreover, two types of both on-center and off-center bipolars. Those which

contact a few cones (in the fovea only one) are called midget bipolars, and those which contact 6-7 cones are called diffuse (parasol) bipolar cells.

The flat contacts that off-center midget and diffuse bipolar cells make to photoreceptors are somewhat different: Off-center midget bipolar dendrites seem to surround the invaginating dendrites of on-center midget bipolars but make flat contacts with the pedicle, while the flat contacts made by off-center diffuse bipolar cells are at the base of the pedicle, their dendrites widely separated from the invaginating dendrites of on-center diffuse bipolar cells (Fig. 15).

On-center bipolar axons make synapses with on-center ganglion cells in the inner layer of the IPL, whereas off-center bipolar axons meet ganglion cell dendrites in the outer layer.

Ganglion cells preserve the same functional categorization as bipolars. Therefore, there are:

1. On-center midget ganglion cells
2. Off-center midget ganglion cells
3. On-center diffuse ganglion cells
4. Off-center diffuse ganglion cells

Thus there are four pathways from each cone: A pair of high acuity (in most cases private) pathways through midget cells, and a pair of low-acuity shared pathways through diffuse cells (Fig. 15).

The midget ganglion cells have small receptive fields, especially near or at fovea, and make up 80 % of the optical nerve fibers. All project to the parvocellular laminae of the LGN. They have a high level of maintained activity, and are therefore called tonic by some researchers.

The diffuse ganglion cells have larger dendritic fields and make up only 10 % of optical nerve fibers. These project to the magnocellular laminae of the LGN and the superior colliculus. They have a low level of maintained activity, show strong transiencies, and have been labeled as phasic.

Shorter wavelengths undergo the greatest chromatic aberration in the eye, and not surprisingly blue cones are not used in high acuity vision. They are conspicuously absent in the fovea, and elsewhere they are rare and spaced in an equidistant fashion. A distinct bipolar, the blue bipolar, serves the pathway from the blue cones. The blue bipolar is strictly on-center (it makes invaginating contacts with blue cones) and contacts with more than one blue cone. Other bipolars, however, may also contact blue cones.

If spectral opponency as well as spatial opponency is also added, the picture gets quite complicated. The classification of ganglion cells is not yet final; at a finer scale 26 categories have been listed. In this summary, only the most important ganglion cell categories will be described.

Most of the midget ganglion cells (about 70 % of all cells) are spatially opponent in an on-center or off-center fashion. About 80 % of these are organized with red-center, green-surround or green-center, red-surround. Sometimes the surround takes input also from blue cones, resulting in a magenta (red + blue) or cyan (green + blue) surround. These eight types, depicted in Fig. 16a and 16b, form the red-green pathway. The remaining 20 % is organized in a blue vs. yellow (red + green) fashion and form the blue-yellow pathway, depicted in Fig. 16c. Almost all the cells in this group are on-center.

A fraction of midget ganglion cells which are infrequent in the fovea do not have spatial opponency, but preserve spectral opponency. They get input from all cones in a blue vs. yellow (red + green) fashion, blue cones providing the excitatory, while red and green cones providing the inhibitory input. All this input, however, is sampled from the same area (Fig. 16d).

Diffuse ganglion cells can also be divided into spatially opponent and spatially non-opponent categories.

Far more common is the spatially opponent type, comprising 20-30 % of all ganglion cells in the retina. These have large receptive fields over which the summation is nonlinear. In contrast to midget cells, they have a low level of spontaneous activity and short conduction latencies. Some of these are spectrally non-opponent, getting input from red and green cones both to the center and the surround (Fig. 17a), while others get red + green input to the center and red input to the surround (Fig. 17b). The latter are exclusively on-center.

The less common spatially non-opponent type gives transient responses to on-off stimuli. Some of them show no antagonistic surround, and get their on responses from red, off-responses from green cones (Fig. 17c). Others have a high spontaneous activity, not usual in this group, and are inhibited by motion.

From this data a complicated picture emerges. From the viewpoint of modeling retinal architecture, some researchers favor to start with a more simplified (and in this case also somewhat different) version: A good example is that given by Gouras and Eggers [1985]. According to their basic classification, there are two midget cell (tonic) categories, and one diffuse (phasic) category.

In the first midget category, the ganglion cells take their input only from red and green cones. The center receives input from one kind, and the surround from the other (Fig. 16a). (This is the red-green pathway) The horizontal cells provide the inhibitory channel and are believed to be cone-specific, i.e., the horizontal cells that get their input from red cones inhibit "green" bipolars, and vice versa.

In the second midget category, the ganglion cells receive input from all cones in a spatially non-opponent fashion. The ganglion cell receives an excitatory input from blue cones, and an inhibitory input from both red and green cones from the whole receptive field (fig. 16d). This fits into one

of the rarer categories mentioned earlier, but not into the dominant one where the concentric opponent organization of the receptive field is preserved (Fig 16c).

The diffuse system in Gouras and Eggers' modeling is spatially opponent, but spectrally non-opponent. It gets all its input from red and green cones both at the center and at the surround. This model, which is the same as the dominant diffuse category (Fig. 17a) at the earlier classification, requires amacrine cell intervention in the IPL to acquire transiency.

Since there is so little known about amacrine cells, and since probably they do not play a major role in the tonic systems, they are not considered here.

A.1.2 Functional physiology

Before an analysis of this data can be carried out, the excitation mechanism and the inhibitory interactions in the OPL should be reviewed.

Unlike other cells in the nervous system, photoreceptors, horizontal cells, and bipolar cells do not produce impulses, but show sustained, graded potential changes to stimuli. It is at the ganglion cell layer that spikes are introduced.

Normally, nerve cells depolarize for impulse production to commence. The distal neurons of the retina, not only do not generate impulses, but also do not necessarily depolarize to excitation. Photoreceptors hyperpolarize to the stimulation of light. Horizontal cells which get their input from photoreceptors may respond to excitation by hyper- or depolarization. The luminance sensitive horizontal cells in some animals are kept in a state of continual depolarization by a transmitter from photoreceptors. When light hits the eye and the photoreceptor hyperpolarizes, this transmitter is

cut off, and the horizontal cell also hyperpolarizes. The behaviour of chromaticity sensitive horizontal cells is not yet known. In humans it seems likely that hyperpolarization takes place.

The cone to on-center bipolar synapse is sign-inverting and the bipolar depolarizes upon the excitation of the photoreceptor. The cone to off-center bipolar synapse is sign-preserving, and the bipolar hyperpolarizes upon the excitation of the photoreceptor (Fig. 18).

The remarkable thing is that the amount of voltage change in the polarization status of the photoreceptors required for neurotransmitter release is in the order of a few millivolts when compared with 20-30 millivolts for other neurons.

The transmitter release versus depolarization is an S-shaped function. By keeping the photoreceptor cells in depolarization, the visual system may be taking advantage of the "knee" of this function, thus boosting sensitivity.

Local adaptation is offered as the reason for the center-surround organization of bipolar cells. The photoreceptor sensitivity range is 3-4 log units of illumination whereas that of the eye is about 10 log units. It is thought that horizontal cells provide a lateral shift of the dynamic operating region. In fact, "... for a brighter surround illumination, a brighter center illumination is necessary for the (on-center) bipolar cell to even start responding." [Dowling & Dubin] This explanation does not preclude the possible role of horizontal cells in deblurring and edge-extraction.

A.1.3 Implications for retinal processing

Probably the most interesting feature of the retina to emerge from this brief summary is the existence of off-center ganglion cells. For a dark object on a light background, these cells increase firing while on-center cells decrease firing. The presence of an off-center pathway for both midrange and

diffuse ganglion cells clearly indicates that "... except at threshold, visual information is reliably coded only by an increase in the firing rate. Thus if objects darker than the background are to be distinguished as clearly as objects lighter than the background, then two pathways are required." [Rodieck 1988]

Previous models of lateral inhibition in the retina which sought to answer the formation of Mach Bands dependent only upon on-center ganglion cells [Ratliff 1965]. In many, this created symmetrical bright and dark Mach Bands which is contrary to many observations about their magnitude. Without including off-center cells, this asymmetry can possibly be explained by an inherent non-linearity in the firing rate: The surround of the ganglion cell at the edge of the bright side is not as brightly illuminated on the gradient side, and does not have much inhibitory power - thus the bright band is created. The surround of the ganglion cell at the edge of the dark side is more strongly illuminated on the side of the gradient, and has more inhibitory power - hence the dark band (Fig. 19a). The increase of the firing rate for an decreased illumination of the surround might be much different from the decrease of the firing rate for an increased illumination even if the change in the illumination is the same in both cases. Some researchers, however, believe that there are other qualitative differences between bright and dark Mach Bands, such as incremental threshold. If the visual system does not rely on a decrease of the firing rate, as the presence of off-center cells indicate, then it is quite possible that dark Mach Bands are signaled by off-center ganglion cells. This might give the dark Band its different characteristics. (It is also possible, though unlikely, that the information from both on-center and off-center systems are combined. This would require complicated circuitry and could be redundant)

The spectral opponency of midget ganglion cells, whether it is red-center, green-surround (or vice versa), or blue-center, yellow(red + green) surround, immediately indicates that chromatic Mach bands, if they exist, are not retinal. Imagine a red-center, green-surround cell coming across an equiluminant chromatic gradient that changes from red to green (Fig. 19b). At location A we get not a reddish but a more greenish (or grayish) band.

Michael has reported cells in the visual cortex which have the capability to account for the chromatic Mach Bands; these are called double-opponent cells. In Fig. 20, the center-surround organization can be seen. These cells should be able to detect specs of one color in the background of another color. When confronted with the gradient in Fig. 19c, such cells can produce a reddish band.

Another subject of interest in the human retina is the nature of lateral inhibition. Models of lateral inhibition based on the lateral eye of the *Limulus* are recursive, i.e., the site of inhibition on a cell is before the response from that cell is diverted to inhibit other cells (Fig. 21a). According to Gouras and Eggers' model, a horizontal cell that gets input from one type of cone inhibits the bipolar of a different kind of cone. The reverse does not occur within the same cell, i.e., a horizontal cell cannot be both post and pre-synaptic in the same cone. This is a one-way street, and characterizes a non-recursive lateral inhibition as seen in Figs. 21b & 22a. On the other hand, some horizontal cells are shown to make synapses to the bipolars receiving input from the same cone as the horizontal cell [Dowling & Dubin] (Fig. 22b). Furthermore they also make serial alternate synapses with other cells (but what types of cells?). Recursiveness, therefore, remains as a possibility. At this point we do not have enough information.

As can be seen from this summary, very detailed (albeit conflicting) wiring models of the retina are emerging. Combined with increased knowledge about the electrical conduction and transmitter release properties of the cells, it will be possible to make good, approximate information processing models of the retina soon.

Appendix 2. Programs

A.2.1 Program Mach-11

```
/* Program Mach-11 */
```

```
/* This program creates the setting for a psychophysical experiment  
whereby a compensation method is used in getting a measure of the  
strength and width of Mach Bands. It uses VGA graphics in displaying  
a gradient pattern which gives rise to dark and bright Mach Bands.  
The slope, as well as the intensities of both the bright and the dark  
uniform areas are user defined. On top of this pattern, the skeleton of  
a rectangle whose dimensions are in the define file "dims.h" is drawn.  
A thin column extending from the top to the bottom of the rectangle  
and of width of one pixel is then generated. This column can be moved  
in both directions and its intensity can be changed by the user;  
therefore, it acts as the compensation tool. The compensated intensity  
distribution in the rectangle as well as the original intensity  
distribution are stored in two integer data files. This data can also be  
viewed during the execution of the program. Looping constructs provide  
several capabilities of redoing the compensation.
```

```
*/
```

```
#include <stdio.h>  
#include <conio.h>  
#include <graph.h>  
#include <memory.h>  
#include <dos.h>  
#include <stdlib.h>  
#include "dims.h"
```

```
/* Below, the variables used in this program are explained:
```

```
"dark_intensity" is the user-defined intensity of the dark uniform area.  
"intensity" is the user-defined intensity of the bright uniform area.  
"scale" is the slope - not necessarily the mathematical slope, but one  
normalized to the smallest possible value without compromising  
resolution (for more explanation see text).  
"steps" is the number of distinct intensity levels at the given scale.  
"step_size" is the intensity difference between each step.
```

"ls" is the starting point of the gradient, calculated so as to justify the gradient.

"left" and "right" are the amounts the skeletal rectangle (enclosing the area where compensation is to be carried out) extends to the left and right respectively from the ends of the gradient area.

"top_border" and "bottom_border" give the upper and lower_bounds of this skeletal rectangle.

"intensity_array" is a one dimensional array containing the x-coordinate values for the original Mach band generating intensity profile.

"intensity_array_comp" is an array of the same type but this time containing intensity values modified by compensation.

"column_no" indicates the x-coordinate of the thin compensating rectangle.

"column_no_store" stores the above value for any repeat loop.

"i", "j", and "jj" are loop counters.

"correct_flag" is a flag testing the "correctness" of a value entered by the user in choosing what step should be taken next. In function Compensation it is also used as the test flag for the loop.

"repeat_flag", as the name suggests, is used to signal to the program that the whole display is going to be recreated after a look at the graphs of intensity_array and intensity_array_comp arrays.

"con_flag" indicates that the original profile within the skeletal rectangle is to be (re)created.

"times" is another flag signalling whether the program has been run once (1) or more (0).

"im_buffer1" is an image buffer that stores the original intensity distribution within the skeletal rectangle.

"im_buffer2" does it after compensation.

"im_buffer3" stores the original intensity of the places where the small pink dot indicating the top of the thin compensation rectangle will go

"seekresult1" and "seekresult2" are return values for the fseek library function.

"written1" and "written2" are return values for the fwrite library function.

"infile" is a binary file storing the values of intensity_array.

"infile_comp" does it for the values of intensity_array_comp.

"length" is the size (in bytes) of each item of data written to the data files.

"offset" is the offset number of bytes the pointer in the data files is located with respect to the last parameter in fseek function call (in this case, where SEEK_SET was used, with respect to the beginning)

*/

```
int i, intensity, length,
    column_no, con_flag, times, intensity_array[325],
    correct_flag, column_no_store,
    intensity_array_comp[325], j, jj,
    repeat_flag, scale,
    ls, steps, step_size, dark_intensity,
    written1, written2, seekresult1,
    seekresult2, f1;

long offset;
char dummy[10], target1[5], parameter, buffer1[40],
    buffer2[40], buffer3[40], target2[5], filename1[15],
```

```

        filename2[15];

        unsigned char far *im_buffer1;
        unsigned char far *im_buffer2;
        unsigned char far *im_buffer3;

/* ***** */

void build_palette()
{
    short pix;
    long color,value;

    for (pix=0; pix <= 63; pix++ ) {
        value = (long)(pix);
        color= value | (value < 8) | (value < 16);
        _remappalette(pix, color);
    }
} /* end of build_palette */

/* ***** */

/* The following function displays a Mach Band generating intensity
distribution on the screen. The gradient is linear. The display is
created in three major steps: First, the bright uniform area on the left
is created using the _rectangle graphics function. Then the gradient is
displayed as a succession of thin rectangles each corresponding to a
different intensity level. Finally the dark uniform area is created
as a single rectangle. After each step the intensity values at the
corresponding x-coordinates are written into the intensity_array. The
variable dots is a flag defined in the "dims.h" include file. If
activated, it displays two pink dots indicating the beginning and the end
of the gradient.
*/

void build_display()
{
    steps = (intensity-dark_intensity)/scale;
    step_size = (intensity-dark_intensity)/steps;
    ls = (320-steps)/2;

    /* first step - displaying the dark uniform area */

    _setcolor(dark_intensity);
    _rectangle(_G_FILLINTERIOR,0,0,ls,200);

    for(j=0; j <= ls; ++j)    /*putting into intensity array*/
    {
        intensity_array[j] = dark_intensity;
    }

    if(dots)
    {
        _setcolor(65);
        _setpixel(ls,top_border-10);
    }
}

```

```

    }

/* second step - displaying the gradient */

for (i= 1; i <= steps; i + +)
{
    _setcolor(dark_intensity + i*step_size);
    _rectangle(_G_FILLINTERIOR,ls + i,0,ls + (i + 1),200);
    intensity_array[ls + i] = dark_intensity + i*step_size;

    /*putting into int. array*/
}

/* third step - displaying the bright uniform area */

_setcolor(intensity);
_rectangle(_G_FILLINTERIOR,ls + 1 + steps,0,320,200);

if(dots)
{
    _setcolor(65);
    _setpixel(ls + 1 + steps,top_border-10);
}

for(j= ls + 1 + steps; j <= 319; + + j) /*putting into int. array*/
{
    intensity_array[j] = intensity;
}
} /* end of build_display */

/* ***** */

/* The following function reads inputs sent by the user */

void reading_inputs()
{
    while(correct_flag == 0)
    {
        parameter = (char) getch();
        if ((parameter == '1') || (parameter == '2') || (parameter == '3')
            || (parameter == '4') || (parameter == '5'))
            correct_flag = 1;
        else
            correct_flag = 0;
    } /* while */
} /* end of reading_inputs */

/* ***** */

/* The following function carries out the compensation procedure.
The compensated area is a thin (1 pixel) rectangle that extends vertically
in the skeletal rectangle. Outside the rectangle, a pink dot
points in the direction of this thin column to inform the user of its
location. When "4" is entered as the parameter, the rectangle moves right,
when "3" is entered, it moves left. When "2" is entered as the parameter,
the rectangle darkens (by the smallest possible amount each time "2" is

```

pressed), when "1" is entered it brightens in the same way. All the final changes are retained when exited from the loop by pressing "c".
 */

void compensation()

```

{
    while(correct_flag == 0)
    {
        parameter = (char)getch();
        if (parameter == '4')
        {
            ++column_no;
            if (column_no > ls-1 + steps + right)
                column_no = ls-1 + steps + right;
            _putimage(column_no-1, top_border-2, im_buffer3, _GPSET);
            _getimage(column_no, top_border-2, 1 + column_no,
                top_border-1, im_buffer3);
            _setcolor(65);
            _setpixel(column_no, top_border-2);
        }
        else if (parameter == '3')
        {
            --column_no;
            if (column_no < ls + 1-left)
                column_no = ls + 1-left;
            _putimage(column_no + 1, top_border-2, im_buffer3, _GPSET);
            if (column_no == ls + 1-left)
                _getimage(column_no + 1, top_border-2, 2 + column_no,
                    top_border-1, im_buffer3);
            else
                _getimage(column_no, top_border-2, 1 + column_no,
                    top_border-1, im_buffer3);
            _setcolor(65);
            _setpixel(column_no, top_border-2);
        }
        else if (parameter == '1')
            ++intensity_array_comp[column_no];
        else if (parameter == '2')
            --intensity_array_comp[column_no];
        else if (parameter == 'c')
            correct_flag = 1;
        else
            correct_flag = 0;
        _setcolor(intensity_array_comp[column_no]);
        _rectangle(_GFillINTERIOR, column_no, top_border + 1,
            column_no, bottom_border-1);
    } /* while */
} /* end of compensation */

/* ***** */

/* The following function displays a graph of the two arrays,
intensity_array and intensity_array_comp, and their element-by-element
difference.
*/
```

```

void display_graph()
{
    for(j = 76; j <= 225; ++j)
    {
        jj = jj + 2;
        _setcolor(63);
        _setpixel(jj, 120 - intensity_array_comp[j]);
        _setpixel(jj, 150 - intensity_array[j]);
        _setpixel(jj, 175 - (intensity_array[j] -
                                intensity_array_comp[j]));
    } /* for */
} /* end of display_graph */

/* *****
***** */

/* The main program first requests the names of the data files, and
then enters a main loop (Loop0) where it firsts asks the parameters
of the intensity distribution and then by calling the function
Build_display displays the Mach Band generating intensity distribution.
If repeat_flag is up, then these questions are not asked but the
display is carried out as before. If repeat_flag is down (the program
is run afresh) the arrays, intensity_array and intensity_array_comp, are
initialized and the values of intensity_array are also copied into
intensity_array_comp. Again, if repeat_flag is down the skeletal
rectangle is drawn, the image buffers are prepared, and the original
intensity profile in the rectangle is stored in im_buffer1. Otherwise,
these steps, which have been done previously, are skipped, and the
stored contents of the skeletal rectangle (where compensation
has been carried out during the earlier run) are retrieved from
im_buffer2 and displayed.

The program then enters the second loop (Loop1) controlled by con_flag.
The first step is the compensation.
Once the desired compensation is carried out, the user has 2 choices:
Either he (or she) can start anew (parameter = "1-4"),
disregarding the compensation carried out so far (restarting Loop1),
or he can display a graph of the original and compensated intensity
distribution (parameter = "5"). If the graph is displayed, he
is presented with another choice: He can return to the compensation
procedure from where he left (parameter = "6" and thus restarting Loop0 -
goes back to the beginning of the program and redisplay everything
with the contents of the skeletal rectangle retrieved from memory)
or he can start the whole program afresh with new parameters for
dark_intensity, intensity, and slope (parameter = "anything else"
and thus restarting Loop0, but this time with new parameters).
If the return character is entered for slope, the program writes
the contents of the intensity_array and intensity_array_comp into
designated files and stops.
*/

```

```

main() {

```

```

/* preparing the files */

FILE *infile, *fopen();
FILE *infile_comp, *fopen();

printf("\nEnter name of data file1 (profile): ");
gets(filename1);
printf("\nEnter name of data file2 (compensation): ");
gets(filename2);

infile = fopen(filename1, "w + b");
infile_comp = fopen(filename2, "w + b");

offset = 0;
length = sizeof(int);

_setvideomode(_DEFAULTMODE);
repeat_flag = 0;

do {
    /* entering Loop0 */
    if (repeat_flag == 0)
    {
        /* learning parameters */
        printf("\nEnter slope: ");
        gets(dummy);

        if (dummy[0] == '\0')
        {
            /* writing into the files */

            seekresult1 = fseek(infile, offset, SEEK_SET);
            written1 = fwrite((char *)intensity_array, length, 320, infile);
            seekresult2 = fseek(infile_comp, offset, SEEK_SET);
            written2 = fwrite((char *)intensity_array_comp, length,
                             320, infile_comp);

            fclose(infile);
            fclose(infile_comp);
            break;
        } /* if (dummy) */

        scale = atoi(dummy);
        printf("\nEnter dark intensity: ");
        gets(target1);
        dark_intensity = atoi(target1);
        printf("\nEnter target intensity: ");
        gets(target2);
        intensity = atoi(target2);
        parameter = '3';
    }
    /* if (repeat) */

    if (repeat_flag == 0)
    {
        for (jj = 0; jj < 325; ++jj)          /* initializing the array */

```

```

    {
        intensity_array[jj] = 0;
        intensity_array_comp[jj] = 0;
    } /* if */
} /* if */

_setvideomode(_MRES256COLOR);

build_palette();

_settextcolor(65);
_settextposition(0,0);

/* creating the intensity distribution */

build_display();

if (repeat_flag == 0)
{
    for(j = 0; j <= 319; ++j) /* copying to compensation array */
    {
        intensity_array_comp[j] = intensity_array[j];
    } /* */
}

if(repeat_flag == 0)
{
    /* displaying the skeletal rectangle, preparing the image
       buffers, and getting the contents of the original
       profile into memory */

    _setcolor(intensity/2);
    _rectangle(_GBORDER, ls-left,top_border,
               ls + steps + right,bottom_border);
    im_buffer1 = (char far *)malloc((unsigned int)
        _imagesize(ls-left,top_border,
                   ls + steps + right,bottom_border));
    im_buffer2 = (char far *)malloc((unsigned int)
        _imagesize(ls-left,top_border,
                   ls + steps + right,bottom_border));
    _getimage(ls-left,top_border,ls + steps + right,
              bottom_border,im_buffer1);
} /* if */
else

    /* displaying the already modified contents of the skeletal
       rectangle */

    _putimage(ls-left,top_border,im_buffer2,_GPSET);

times = 0;
con_flag = 1;

if (repeat_flag == 0) /* if it is not repeated */
    times = 1;

```

```

/* ----- entering Loop1 ----- */

while(con_flag)
{
/* -----compensation section----- */

for(j = 0; j < ls-left; + + j)
{
intensity_array_comp[j] = intensity_array[j];
}

for (j = ls + steps + right; j < 325; + + j)
{
intensity_array_comp[j] = intensity_array[j];
}

_setcolor(65);
_setpixel(ls-left, top_border-10);
_setpixel(ls + steps + right, top_border-10);

if ((times == 1) && (repeat_flag == 0))
column_no = ls;

/* if it's repeated, start from where you left */

if (repeat_flag == 1)
column_no = column_no_store;

im_buffer3 = (char far *) malloc((unsigned int)
_imagesize(ls-left, top_border-2,
ls + steps + right, top_border-1));

correct_flag = 0;

_settextposition(0,0);
printf(buffer1, "enter parameter:(1,2,3,4) or "
"\n c for continuation");
_outtext(buffer1);

/* compensation */

compensation();

column_no_store = column_no;

/* -----end of compensation section----- */

times = 0; /* after the first run */

correct_flag = 0; i = 0;

/* -----choices----- */

_settextposition(0,0);

```

```

printf(buffer2,"enter 1-4 for clear or      "
      "\n    5 for graphics"); /* getting parameter */
_outtext(buffer2);

reading_inputs();

if (parameter == '5')          /* graphical representation */
{
    _getimage(ls-left,top_border,ls+steps+right,
              bottom_border,im_buffer2);
    _clearscreen(_GCLEARSCREEN);

    jj = 0;

    display_graph();

    printf(buffer3,"enter 6 for repeat,"
          "\n anything else for once again");
    _outtext(buffer3);

    parameter = (char)getch();
    if (parameter == '6') /* repeat */
    {
        /* get out of loop, set rep. flag */
        con_flag = 0;
        repeat_flag = 1; _clearscreen(_GCLEARSCREEN);
    }
    else /* once again from the beginning*/
    {
        /* get out of loop */
        con_flag = 0; repeat_flag = 0;
        _clearscreen(_GCLEARSCREEN);
    }
} /* if */

else /* if parameter != 5 */
{
    /* don't get out of loop */
    repeat_flag = 0;
    _setcolor(intensity_array[column_no_store]);
    _setpixel(column_no_store, top_border-2);
    _putimage(ls-left,top_border,im_buffer1, _GPSET);
    for(i = ls-left; i <= ls-1 + steps + right; ++i)
    {
        intensity_array_comp[i] = intensity_array[i];
    } /* for */

}

/* -----end of choices----- */

}

/* ----- end of while (Loop1) ----- */

    _setvideomode(_DEFAULTMODE);
}while (1); /* end of Loop0 */

} /* end of main */

```

A.2.2 Program Mach-e1

```
/* Program Mach-e1 */

/* This program creates the setting for a psychophysical experiment
whereby a compensation method is used in getting a measure of the
strength and width of Mach Bands. It uses VGA graphics in displaying
a gradient pattern which gives rise to dark and bright Mach Bands.
The slope, as well as the intensities of both the bright and the dark
uniform areas are user defined. On top of this pattern, the skeleton of
a rectangle whose dimensions are in the define file "dims.h" is drawn.
A thin column extending from the top to the bottom of the rectangle
and of width of one pixel is then generated. This column can be moved
in both directions and its intensity can be changed by the user;
therefore, it acts as the compensation tool. The compensated intensity
distribution in the rectangle as well as the original intensity
distribution are stored in two integer data files. This data can also be
viewed during the execution of the program. Looping constructs provide
several capabilities of redoing the compensation.
*/

#include <stdio.h>
#include <conio.h>
#include <graph.h>
#include <memory.h>
#include <dos.h>
#include <stdlib.h>
#include <math.h>
#include "dims.h"

/* Below, the variables used in this program are explained:

"dark_intensity" is the user-defined intensity of the dark uniform area.
"bottom_log_intensity" is the log of dark_intensity.
"intensity" is the user-defined intensity of the bright uniform area.
"top_log_intensity" is the log of intensity.
"scale" is the slope - not necessarily the mathematical slope, but one
used to calculate the number of log intensity levels between
the dark and the bright intensity levels and normalized.
"no_of_steps" is the number of distinct log intensity levels at the
given scale. This real value is then converted to integer "steps".
"step_size" is the log intensity difference between each step.
"power" is the log intensity value at each step.
"f_intensity" is the true real intensity at each step. It is when this
value is converted to integer (and thus truncated) that loss of
resolution occurs, and that unlike the linear gradient, we do not
have maximum resolution at the minimum slope (1).
"ls" is the starting point of the gradient, calculated so as to justify
the gradient.
```

"left" and "right" are the amounts the skeletal rectangle (enclosing the area where compensation is to be carried out) extends to the left and right respectively from the ends of the gradient area.

"top_border" and "bottom_border" give the upper and lower_bounds of this skeletal rectangle.

"intensity_array" is a one dimensional array containing the x-coordinate values for the original Mach band generating intensity profile.

"intensity_array_comp" is an array of the same type but this time containing intensity values modified by compensation.

"column_no" indicates the x-coordinate of the thin compensating rectangle.

"column_no_store" stores the above value for any repeat loop.

"i", "j", and "jj" are loop counters.

"correct_flag" is a flag testing the "correctness" of a value entered by the user in choosing what step should be taken next. In function Compensation it is also used as the test flag for the loop.

"repeat_flag", as the name suggests, is used to signal to the program that the whole display is going to be recreated after a look at the graphs of intensity_array and intensity_array_comp arrays.

"con_flag" indicates that the original profile within the skeletal rectangle is to be (re)created.

"times" is another flag signalling whether the program has been run once (1) or more (0).

"im_buffer1" is an image buffer that stores the original intensity distribution within the skeletal rectangle.

"im_buffer2" does it after compensation.

"im_buffer3" stores the original intensity of the places where the small pink dot indicating the top of the thin compensation rectangle will go

"seekresult1" and "seekresult2" are return values for the fseek library function.

"written1" and "written2" are return values for the fwrite library function.

"infile" is a binary file storing the values of intensity_array.

"infile_comp" does it for the values of intensity_array_comp.

"length" is the size (in bytes) of each item of data written to the data files.

"offset" is the offset number of bytes the pointer in the data files is located with respect to the last parameter in fseek function call (in this case, where SEEK_SET was used, with respect to the beginning).

*/

```
int i, intensity, ls,
    column_no, w, con_flag, x, times, intensity_array[325],
    correct_flag, column_no_store,
    intensity_array_comp[325], j, jj, increment, sum,
    repeat_flag,
    steps, dark_intensity, length,
    seekresult1, seekresult2, written1, written2;

char target1[5], parameter, warning[50], buffer1[40],
    width[2], buffer2[40], buffer3[40], b1[50], filename1[15],
    filename2[15],
    dummy[10], target2[5];

unsigned char far *im_buffer1;
```

```

unsigned char far *im_buffer2;
unsigned char far *im_buffer3;

long offset;

double e,power,step_size,top_log_intensity,f_intensity,no_of_steps,
       scale, bottom_log_intensity, diff;

/* ***** */

void build_palette()
{
    short pix;
    long color,value;

    for (pix=0; pix <= 63; pix++) {
        value = (long)(pix);
        color = value | (value < < 8) | (value < < 16);

        _remappalette(pix, color);
    } /* end of build_palette */

/* ***** */

/* The following function displays a Mach Band generating intensity
distribution on the screen. The gradient is exponential. The display is
created in three major steps: First, the bright uniform area on the left
is created using the _rectangle graphics function. Then the gradient is
displayed as a succession of thin rectangles each corresponding to a
different log, but not necessarily distinct integer intensity level.
Finally the dark uniform area is created
as a single rectangle. After each step the intensity values at the
corresponding x-coordinates are written into the intensity_array. The
variable dots is a flag defined in the "dims.h" include file. If
activated, it displays two pink dots indicating the beginning and the end
of the gradient.
*/

void build_display()
{
    e = 2.718281828;
    top_log_intensity = log((double)intensity);
    bottom_log_intensity = log((double)dark_intensity);
    no_of_steps = (intensity-dark_intensity)/scale;
    steps = (int)no_of_steps;
    if (steps >= 250)
        exit(2);
    diff = top_log_intensity-bottom_log_intensity;
    step_size = (top_log_intensity-bottom_log_intensity)/no_of_steps;
    ls = (320-steps)/2;

    /* first step - displaying the dark uniform area */

```

```

        _setcolor(dark_intensity);
        _rectangle(_G_FILLINTERIOR,0,0,ls,200);

        if (dots)
        {
            _setcolor(65);
            _setpixel(ls,top_border-10);
        }

        for(j=0; j <= ls; ++j)    /*putting into intensity array*/
        {
            intensity_array[j]=dark_intensity;
        }    /* */

    /* second step - displaying the gradient    */

    for (i=1; i <= steps; i++)
    {
        power=bottom_log_intensity+i*step_size;
        f_intensity=pow(e,power);
        _setcolor((int)f_intensity);
        _rectangle(_G_FILLINTERIOR,ls+i,0,ls+(i+1),200);
        intensity_array[ls+i]=(int)f_intensity; /*putting into int. array*/
    }

    /* third step - displaying the bright uniform area    */

    _setcolor((int)pow(e,top_log_intensity));
    _rectangle(_G_FILLINTERIOR,ls+1+steps,0,320,200);

    if(dots)
    {
        _setcolor(65);
        _setpixel(ls+1+steps,top_border-10);
    }

    for(j=ls+1+steps; j <= 319; ++j) /*putting into int. array*/
    {
        intensity_array[j]=(int)pow(e,top_log_intensity);
    }
} /* end of build_display */

/* ***** */

/* The following function carries out the compensation procedure.
The compensated area is a thin (1 pixel) rectangle that extends vertically
in the skeletal rectangle. Outside the rectangle, a pink dot
points in the direction of this thin column to inform the user of its
location. When "4" is entered as the parameter, the rectangle moves right,
when "3" is entered, it moves left. When "2" is entered as the parameter,
the rectangle darkens (by the smallest possible amount each time "2" is
pressed), when "1" is entered it brightens in the same way. All the
final changes are retained when exited from the loop by pressing "c".
*/

void compensation()

```

```

{
    while(correct_flag == 0)
    {
        parameter = (char)getch();
        if (parameter == '4')
        {
            ++column_no;
            if (column_no > ls-1 + steps + right)
                column_no = ls-1 + steps + right;
            _putimage(column_no-1, top_border-2, im_buffer3, _GPSET);
            _getimage(column_no, top_border-2, 1+column_no,
                top_border-1, im_buffer3);
            _setcolor(65);
            _setpixel(column_no, top_border-2);
        }
        else if (parameter == '3')
        {
            --column_no;
            if (column_no < ls + 1-left)
                column_no = ls + 1-left;
            _putimage(column_no + 1, top_border-2, im_buffer3, _GPSET);
            if (column_no == ls + 1-left)
                _getimage(column_no + 1, top_border-2, 2+column_no,
                    top_border-1, im_buffer3);
            else
                _getimage(column_no, top_border-2, 1+column_no,
                    top_border-1, im_buffer3);
            _setcolor(65);
            _setpixel(column_no, top_border-2);
        }
        else if (parameter == '1')
            ++intensity_array_comp[column_no];
        else if (parameter == '2')
            --intensity_array_comp[column_no];
        else if (parameter == 'c')
            correct_flag = 1;
        else
            correct_flag = 0;
        _setcolor(intensity_array_comp[column_no]);
        _rectangle(_GFillINTERIOR, column_no, top_border + 1,
            column_no, bottom_border-1);
    }
} /* while */
} /* end of compensation */

/* ***** */

/* The following function reads inputs sent by the user */

void reading_inputs()
{
    while(correct_flag == 0)
    {
        parameter = (char) getch();
        if ((parameter == '1') || (parameter == '2') || (parameter == '3'))

```

```

        ||(parameter == '4')||(parameter == '5'))
        correct_flag = 1;
    else
        correct_flag = 0;
    } /* while */
} /* end of reading_inputs */

/* ***** */

/* The following function displays a graph of the two arrays,
   intensity_array and intensity_array_comp, and their element-by-element
   difference.
   */

void display_graph()
{
    for(j = 76; j <= 225; ++j)
    {
        jj = jj + 2;
        _setcolor(63);
        _setpixel(jj, 120 - intensity_array_comp[j]);
        _setpixel(jj, 150 - intensity_array[j]);
        _setpixel(jj, 175 - (intensity_array[j] -
                               intensity_array_comp[j]));
    } /* for */
} /* end of display_graph */

/* ***** */

/* The main program first requests the names of the data files, and
   then enters a main loop (Loop0) where it firsts asks the parameters
   of the intensity distribution and then by calling the function
   Build_display displays the Mach Band generating intensity distribution.
   If repeat_flag is up, then these questions are not asked but the
   display is carried out as before. If repeat_flag is down (the program
   is run afresh) the arrays, intensity_array and intensity_array_comp, are
   initialized and the values of intensity_array are also copied into
   intensity_array_comp. Again, if repeat_flag is down the skeletal
   rectangle is drawn, the image buffers are prepared, and the original
   intensity profile in the rectangle is stored in im_buffer1. Otherwise,
   these steps, which have been done previously, are skipped, and the
   stored contents of the skeletal rectangle (where compensation
   has been carried out during the earlier run) are retrieved from
   im_buffer2 and displayed.

   The program then enters the second loop (Loop1) controlled by con_flag.
   The first step is the compensation.
   Once the desired compensation is carried out, the user has 2 choices:
   Either he (or she) can start anew (parameter = "1-4"),
   disregarding the compensation carried out so far (restarting Loop1),
   or he can display a graph of the original and compensated intensity
   distribution (parameter = "5"). If the graph is displayed, he
   is presented with another choice: He can return to the compensation
   procedure from where he left (parameter = "6" and thus restarting Loop0 -

```

goes back to the beginning of the program and redisplay everything with the contents of the skeletal rectangle retrieved from memory) or he can start the whole program afresh with new parameters for dark_intensity, intensity, and slope (parameter = "anything else" and thus restarting Loop0, but this time with new parameters). If the return character is entered for slope the program writes the contents of the intensity_array and intensity_array_comp into designated files and stops.

```
*/
```

```
main() {

    /* preparing the files    */

    FILE *infile, *fopen();
    FILE *infile_comp, *fopen();

    printf("\nEnter the name of data file1 (profile): ");
    gets(filename1);
    printf("\nEnter the name of data file2 (compensation): ");
    gets(filename2);

    infile = fopen(filename1, "w + b");
    infile_comp = fopen(filename2, "w + b");

    _setvideomode(_DEFAULTMODE);
    repeat_flag = 0;

    length = sizeof(int);
    offset = 0;

    do {
        /* entering Loop0 */
        if (repeat_flag == 0)
        {
            printf("\nEnter slope:"); /* learning parameters */
            gets(dummy);
            if(dummy[0] == '\0')
            {
                /* writing to the files */

                seekresult1 = fseek(infile,offset,SEEK_SET);
                written1 = fwrite((char *)intensity_array,length,320,infile);
                seekresult2 = fseek(infile_comp,offset,SEEK_SET);
                written2 = fwrite((char *)intensity_array_comp,length,
                                320,infile_comp);

                fclose(infile);
                fclose(infile_comp);
                break;
            }
            scale = (double) atoi(dummy);
            if (scale == 0.0)
                exit(1);
        }
    }
}
```

```

        printf("\nEnter dark side intensity (greater than 0): ");
        gets(target2);
        dark_intensity = atoi(target2);
        printf("\nEnter bright side intensity: ");
        gets(target1);
        intensity = atoi(target1);
        parameter = '3';
    }
    /* if */

if (repeat_flag == 0)
{
    for (jj = 0; jj < 325; ++jj)          /* initializing the array */
    {
        intensity_array[jj] = 0;
        intensity_array_comp[jj] = 0;
    }
    /* */
} /* if */

_setvideomode(_MRES256COLOR);

build_palette();

_settextcolor(65);
_settextposition(0,0);

/* creating the display */

build_display();

if (repeat_flag == 0)
{
    for(j = 0; j <= 319; ++j)          /* copying to compensation array */
    {
        intensity_array_comp[j] = intensity_array[j];
    }
    /* */
}

if(repeat_flag == 0)
{
    /* displaying the skeletal rectangle, preparing the image
       buffers, and getting the contents of the original
       profile into memory */

    _setcolor(intensity/2);
    _rectangle(_GBORDER, ls-left,top_border,
               ls + steps + right,bottom_border);
    im_buffer1 = (char far *)malloc((unsigned int)
                                     _imagesize(ls-left,top_border,
                                                  ls + steps + right,bottom_border));
    im_buffer2 = (char far *)malloc((unsigned int)
                                     _imagesize(ls-left,top_border,
                                                  ls + steps + right,bottom_border));
}

```

```

        _getimage(ls-left,top_border,ls+steps+right,
                bottom_border,im_buffer1);
    } /* if */
else

    /* displaying the already modified contents of the skeletal
    rectangle */

    _putimage(ls-left,top_border,im_buffer2,_GPSET);

times = 0;
con_flag = 1;

if (repeat_flag == 0) /* if it is not repeated */
    times = 1;

/* ----- entering Loop1 ----- */
while(con_flag)
{
    /*----- compensation section ----- */

    for(j = 0; j < ls-left; ++j)
    {
        intensity_array_comp[j] = intensity_array[j];
    }

    for (j = ls+steps+right; j < 325; ++j)
    {
        intensity_array_comp[j] = intensity_array[j];
    }

    /* creating the rectangle */

    _setcolor(65);
    _setpixel(ls-left, top_border-10);
    _setpixel(ls+steps+right,top_border-10);

    if ((times == 1)&&(repeat_flag == 0))
        column_no = ls;

    /* if it's repeated, start from where you left */

    if (repeat_flag == 1)
        column_no = column_no_store;

    im_buffer3 = (char far *)malloc((unsigned int)
        _imagesize(ls-left,top_border-2,
            ls+steps+right, top_border-1));

    correct_flag = 0;

    _settextposition(0,0);
    sprintf(buffer1,"enter parameter:(1,2,3,4) or "

```

```

        "\n c for continuation");
_outtext(buffer1);

compensation();

column_no_store = column_no;

/* ----- end of compensation section ----- */

times = 0;      /* after the first run */

correct_flag = 0; i = 0;

/* ----- choices ----- */

_settextposition(0,0);
sprintf(buffer2,"enter 1-4 for clear or      "
        "\n      5 for graphics"); /* getting parameter */
_outtext(buffer2);

reading_inputs();

if (parameter == '5')      /* graphical representation */
{
    _getimage(ls-left,top_border,ls + steps + right,
              bottom_border,im_buffer2);
    _clearscreen(_GCLEARSCREEN);

    jj = 0;

    display_graph();

    sprintf(buffer3,"enter 6 for repeat,"
              "\n anything else for once again");
    _outtext(buffer3);

    parameter = (char)getch();
    if (parameter == '6') /* repeat */
    {
        /* get out of loop, set rep. flag */
        con_flag = 0;
        repeat_flag = 1; _clearscreen(_GCLEARSCREEN);
    }
    else /* once again from the beginning */
    {
        /* get out of loop */
        con_flag = 0; repeat_flag = 0;
        _clearscreen(_GCLEARSCREEN);
    }
} /* if */

else /* if parameter != 5 */
{
    /* don't get out of loop */
    repeat_flag = 0;
    _setcolor(intensity_array[column_no_store]);
    _setpixel(column_no_store, top_border-2);
}

```

```

        _putimage(ls-left,top_border,im_buffer1, _GPSET);
        for(i = ls-left; i <= ls-I + steps + right; + + i)
        {
            intensity_array_comp[i] = intensity_array[i];
        } /* for */

    }
    /* ----- end of choices ----- */

}
/* ----- end of Loop1 ----- */

    setvideomode(_DEFAULTMODE);
} while (1); /* end of Loop0 */

} /* end of main */

```

A.2.3 Program Mach-l2

```

/* Program Mach-l2 */

/* This program creates the setting for a psychophysical experiment
whereby a compensation method is used in getting a measure of the
strength and width of Mach Bands. It uses VGA graphics in displaying
a gradient pattern which gives rise to dark and bright Mach Bands.
The slope, as well as the intensities of both the bright and the dark
uniform areas are user defined. On top of this pattern, the skeleton of
a rectangle whose dimensions are in the define file "dims.h" is drawn.
A compensation rectangle extending from the top to the bottom of
the skeletal rectangle and of user-defined width is then generated.
This rectangle, in cross-section, has an intensity distribution
corresponding to a triangular wave. The rectangle can be moved
in both directions and its intensity can be changed by the user;
therefore, it acts as the compensation tool. The compensated intensity
distribution in the skeletal rectangle as well as the original intensity
distribution are stored in two integer data files. This data can also be
viewed during the execution of the program. Looping constructs provide
several capabilities of redoing the compensation.
*/

#include <stdio.h>
#include <conio.h>
#include <graph.h>
#include <memory.h>
#include <dos.h>
#include <stdlib.h>
#include "dims.h"

```

/* Below, the variables used in this program are explained:

"dark_intensity" is the user-defined intensity of the dark uniform area.
 "intensity" is the user-defined intensity of the bright uniform area.
 "scale" is the slope - not necessarily the mathematical slope, but one normalized to the smallest possible value without compromising resolution (for more explanation see text).
 "steps" is the number of distinct intensity levels at the given scale.
 "step_size" is the intensity difference between each step.
 "ls" is the starting point of the gradient, calculated so as to justify the gradient.
 "left" and "right" are the amounts the skeletal rectangle (enclosing the area where compensation is to be carried out) extends to the left and right respectively from the ends of the gradient area.
 "top_border" and "bottom_border" give the upper and lower bounds of this skeletal rectangle.
 "intensity_array" is a one dimensional array containing the x-coordinate values for the original Mach band generating intensity profile.
 "intensity_array_comp" is an array of the same type but this time containing intensity values modified by compensation.
 "changing_intensity" is the array where the triangular wave values are stored.
 "increment" is the intensity difference between successive triangular wave values.
 "column_no" is the array index for changing_intensity.
 "w" is the user-defined width of the compensating rectangle (must be odd).
 "x" is the location of the rectangle.
 "i", "j", and "jj" are loop counters.
 "correct_flag" is a flag testing the "correctness" of a value entered by the user in choosing what step should be taken next. In function Compensation it is also used as the test flag for the loop.
 "repeat_flag", as the name suggests, is used to signal to the program that the whole display is going to be recreated after a look at the graphs of intensity_array and intensity_array_comp arrays.
 "con_flag" indicates that the original profile within the skeletal rectangle is to be (re)created.
 "times" is another flag signalling whether the program has been run once (1) or more (0).
 "im_buffer1" is an image buffer that stores the original intensity distribution within the skeletal rectangle.
 "im_buffer2" does it after compensation.
 "seekresult1" and "seekresult2" are return values for the fseek library function.
 "written1" and "written2" are return values for the fwrite library function.
 "infile" is a binary file storing the values of intensity_array.
 "infile_comp" does it for the values of intensity_array_comp.
 "length" is the size (in bytes) of each item of data written to the data files.
 "offset" is the offset number of bytes the pointer in the data files is located with respect to the last parameter in fseek function call (in this case, where SEEK_SET was used, with respect to the beginning).
 */

```

int i, intensity, changing_intensity[15],ls,
    dark_intensity, steps, step_size, length,

```

```

        column_no, w, con_flag, x, times, intensity_array[325],
        correct_flag, scale,
        intensity_array_comp[325], j, jj, increment, repeat_flag,
        seekresult1, seekresult2, written1, written2;

    char dummy[10], target1[5], parameter, warning[50], buffer1[40],
        buffer2[40], buffer3[40], width[2], target2[5], filename1[15],
        filename2[15];

    unsigned char far *im_buffer1;
    unsigned char far *im_buffer2;

    long offset;

/* ***** */

void build_palette()
{
    short pix;
    long color,value;

    for (pix=0; pix <= 63; pix++) {
        value = (long)(pix);
        color = value | (value < 8) | (value < 16);

        _remappalette(pix, color);
    }
} /* end of build_palette */

/* ***** */

/* The following function displays a Mach Band generating intensity
distribution on the screen. The gradient is linear. The display is
created in three major steps: First, the bright uniform area on the left
is created using the _rectangle graphics function. Then the gradient is
displayed as a succession of thin rectangles each corresponding to a
different intensity level. Finally the dark uniform area is created
as a single rectangle. After each step the intensity values at the
corresponding x-coordinates are written into the intensity_array. The
variable dots is a flag defined in the "dims.h" include file. If
activated, it displays two pink dots indicating the beginning and the end
of the gradient.
*/

void build_display()
{
    steps = (intensity-dark_intensity)/scale;
    step_size = (intensity-dark_intensity)/steps;
    ls = (320-steps)/2;

    _settextcolor(65);

    /* first step - displaying the dark uniform area */

    _setcolor(dark_intensity);

```

```

_rectangle(_G_FILLINTERIOR,0,0,ls,200);

if (dots)
{
    _setcolor(65);
    _setpixel(ls,top_border-10);
}

for(j=0; j <= ls; ++j)    /*putting into intensity array*/
{
    intensity_array[j]=dark_intensity;
}    /* */

/* second step - displaying the gradient */

for (i=1; i <= steps; i++)
{
    _setcolor(dark_intensity+i*step_size);
    _rectangle(_G_FILLINTERIOR,ls+i,0,ls+(i+1),200);
    intensity_array[ls+i]=dark_intensity+i*step_size; /*putting
                                                    into int. array*/
}

/* third step - displaying the bright uniform area */

_setcolor(intensity);
_rectangle(_G_FILLINTERIOR,ls+1+steps,0,320,200);

if (dots)
{
    _setcolor(65);
    _setpixel(ls+1+steps,top_border-10);
}

for(j=ls+1+steps; j <= 320; ++j) /*putting into int. array*/
{
    intensity_array[j]=intensity;
}
} /* end of build_display */

/* ***** */

/* The following function reads inputs sent by the user */

void reading_inputs1()
{
    while (correct_flag==0)
    {
        parameter=(char)getch();
        if ((parameter=='1')||(parameter=='2')||(parameter=='3')
            ||(parameter=='4'))
            correct_flag=1;
        else
            correct_flag=0;
    } /* while */
}

```

```

    } /* end of reading_inputs1 */

/* ***** */

/* This function also reads inputs sent by the user */

void reading_inputs2()
{
    while(correct_flag == 0)
    {
        parameter = (char) getch();
        if ((parameter == '1') || (parameter == '2') || (parameter == '3')
            || (parameter == '4') || (parameter == '5'))
            correct_flag = 1;
        else
            correct_flag = 0;
    } /* while */
} /* end of reading_inputs2 */

/* ***** */

/* The following function carries out the compensation procedure.
The compensated area is a rectangle of user-defined width that
extends vertically in the skeletal rectangle. This rectangle has an
intensity profile of a triangular wave. The intensity changes are
carried out by changing the "increment" between the successive values of
the triangular wave.
When "4" is entered as the parameter, the rectangle moves right,
when "3" is entered, it moves left. When "2" is entered as the parameter,
the rectangle darkens (by the smallest possible amount each time "2" is
pressed), when "1" is entered it brightens in the same way.
*/

void compensation()
{
    /* calculating the intensity of columns in the triangular wave */

    for(column_no = 0; column_no <= (w-1)/2-1; ++column_no)
    {
        changing_intensity[column_no] = intensity_array[ls-left + x + column_no]
            + increment*(column_no + 1);
    }

    for(column_no = w-1; column_no >= (w-1)/2; --column_no)
    {
        changing_intensity[column_no] = intensity_array[ls-left + x + column_no]
            + increment*(w-column_no);
    }

    /* creating the rectangle based on the triangular wave */

    for(column_no = 0; column_no <= w-1; ++column_no)
    {

```

```

        _setcolor(changing_intensity[column_no]);
        _rectangle(_G_FILLINTERIOR,ls-left + x + column_no,top_border + 1,
                    ls-left + x + column_no,bottom_border-1);
        intensity_array_comp[ls-left + x + column_no] = changing_intensity[column_no];
    } /* end of compensation */

/* ***** */

/* The following function displays a graph of the two arrays,
   intensity_array and intensity_array_comp, and their element-by-element
   difference.
   */

void display_graph()
{
    for(j = 76; j <= 225; ++j)
    {
        jj = jj + 2;
        _setcolor(63);
        _setpixel(jj,120-intensity_array_comp[jj]);
        _setpixel(jj,150-intensity_array[jj]);
        _setpixel(jj,175-(intensity_array[jj]-
                                intensity_array_comp[jj]));
    } /* for */
} /* end of display_graph */

/* ***** */

/* The main program first requests the names of the data files, and
   then enters a main loop (Loop0) where it firsts asks the parameters
   of the intensity distribution and then by calling the function
   Build_display displays the Mach Band generating intensity distribution.
   If repeat_flag is up, then these questions are not asked but the
   display is carried out as before. If repeat_flag is down (the program
   is run afresh) the arrays, intensity_array and intensity_array_comp, are
   initialized and the values of intensity_array are also copied into
   intensity_array_comp. Again, if repeat_flag is down the skeletal
   rectangle is drawn, the image buffers are prepared, and the original
   intensity profile in the rectangle is stored in im_buffer1. Otherwise,
   these steps, which have been done previously, are skipped, and the
   stored contents of the skeletal rectangle (where compensation
   has been carried out during the earlier run) are retrieved from
   im_buffer2 and displayed.

   The program then enters the second loop (Loop1) controlled by con_flag.
   The first step is the compensation.
   Once the desired compensation is carried out, the user has 2 choices:
   Either he (or she) can start anew (parameter="1-4"),
   disregarding the compensation carried out so far (restarting Loop1),
   or he can display a graph of the original and compensated intensity
   distribution (parameter="5"). If the graph is displayed, he
   is presented with another choice: He can return to the compensation

```

procedure from where he left (parameter = "6" and thus restarting Loop0 - goes back to the beginning of the program and redisplay everything with the contents of the skeletal rectangle retrieved from memory) or he can start the whole program afresh with new parameters for dark_intensity, intensity, and slope (parameter = "anything else" and thus restarting Loop0, but this time with new parameters). If the return character is entered for slope the program writes the contents of the intensity_array and intensity_array_comp into designated files and stops.

```
*/
```

```
main() {

    /* preparing the files */

    FILE *infile, *fopen();
    FILE *infile_comp, *fopen();

    printf("\nEnter name of data file1 (profile): ");
    gets(filename1);
    printf("\nEnter name of data file2 (compensation): ");
    gets(filename2);

    infile = fopen(filename1, "w + b");
    infile_comp = fopen(filename2, "w + b");

    length = sizeof(int);
    offset = 0;

    _setvideomode(_DEFAULTMODE);
    repeat_flag = 0;

    do {
        /* entering Loop0 */
        if (repeat_flag == 0)
        {
            /* learning parameters */
            printf("\nEnter slope:");
            gets(dummy);
            if (dummy[0] == '\0')
            {
                /* writing into the files */

                seekresult1 = fseek(infile, offset, SEEK_SET);
                written1 = fwrite((char *)intensity_array, length, 320, infile);
                seekresult2 = fseek(infile_comp, offset, SEEK_SET);
                written2 = fwrite((char *)intensity_array_comp, length,
                                320, infile_comp);

                fclose(infile);
                fclose(infile_comp);
                break;
            }
            scale = atoi(dummy);
            printf("\nEnter dark intensity: ");
            gets(target1);
            dark_intensity = atoi(target1);
        }
    } while (repeat_flag == 0);
}
```

```

printf("\nEnter target intensity: ");
gets(target2);
intensity = atoi(target2);
printf("\nEnter width of rectangle (odd) : ");
gets(width);
w = atoi(width);
parameter = '3';
}
/* if */

for (jj = 0; jj < 325; ++jj)      /* initializing the array */
{
    intensity_array[jj] = 0;
    intensity_array_comp[jj] = 0;
}                                  /* */

_setvideomode(_MRES256COLOR);

build_palette();

build_display();

for(j = 0; j <= 320; ++j)      /* copying to compensation array */
{
    intensity_array_comp[j] = intensity_array[j];
}                                /* */

if(repeat_flag == 0)
{
    /* displaying the skeletal rectangle, preparing the image
       buffers, and getting the contents of the original profile
       into memory */

    _setcolor(intensity/2);
    _rectangle(_GBORDER,ls-left, top_border,
               ls + steps + right, bottom_border);
    im_buffer1 = (char far *)malloc((unsigned int)
                                     _imagesize(ls-left,top_border,
                                                  ls + steps + right,bottom_border));
    im_buffer2 = (char far *)malloc((unsigned int)
                                     _imagesize(ls-left,top_border,
                                                  ls + steps + right,bottom_border));
    _getimage(ls-left,top_border,
              ls + steps + right,bottom_border,im_buffer1);
}                                /* if */
else

    /* displaying the already modified contents of the skeletal
       rectangle */

    _putimage(ls-left,top_border,im_buffer2,_GPSET);

times = 0;
con_flag = 1;

```

```

if (repeat_flag == 0) /* if it is not repeated */
{
    x = 10;
    increment = 1;
    times = 1;
}

/* ----- entering Loop1 ----- */

while(con_flag)
{
    if (repeat_flag == 1) /* getting parameter in repeat */
    {
        _settextposition(0,0);
        sprintf(buffer1,"enter parameter(1,2,3,4)");
        _outtext(buffer1);
        correct_flag = 0; i = 0;

        reading_inputs1();

    } /* if */

    if (times == 0)
    {
        if (parameter == '1')
        {
            ++increment;
            if(increment == 0) ++increment;
        }
        else if (parameter == '2')
        {
            --increment;
            if(increment == 0) --increment;
        }
        else if (parameter == '4')
        {
            ++x;
            if(x == ((right + left)-w/2 + steps))
                --x;
        }
        else if (parameter == '3')
        {
            --x;
            if(x == 1)
                ++x;
        }
        else
        {
            _setvideomode(_DEFAULTMODE);
            exit(1);
        }
    } /* if times = 0, we're changing pos or inc */
    /* in all cases after the first run */

    for(j = 0; j < 325; ++j) /* initializing comp array */

```

```

    {
        intensity_array_comp[j] = intensity_array[j];
    }

    /* compensation */
    compensation();

    times = 0;    /* after the first run */

    _settextposition(0,0);
    sprintf(buffer2, "enter parameter(1,2,3,4)"
                " or 5 for graph");
    _outtext(buffer2);
    correct_flag = 0; i = 0;    /* getting parameter */

    reading_inputs2();

    if (parameter == '5')    /* graphical representation */
    {
        _getimage(ls-left, top_border,
                  ls + steps + right, bottom_border, im_buffer2);
        _clearscreen(_GCLEARSCREEN);

        jj = 0;

        display_graph();

        _settextposition(0,0);
        sprintf(buffer3, "enter 6 for repeat"
                    " or anything for once again");
        _outtext(buffer3);
        parameter = (char) getch();
        if (parameter == '6') /* repeat */
        {
            /* get out of loop, set rep. flag */
            con_flag = 0;
            repeat_flag = 1; _clearscreen(_GCLEARSCREEN);
        }
        else /* once again from the beginning */
        {
            /* get out of loop */
            con_flag = 0; repeat_flag = 0;
            _clearscreen(_GCLEARSCREEN);
        }
    } /* if */

    else /* if parameter != 5 */
    {
        /* don't get out of loop */
        repeat_flag = 0;
        _putimage(ls-left, top_border, im_buffer1, _GPSET);
    }
}

/* ----- end of Loop1 ----- */

_setvideomode(_DEFAULTMODE);
}while (1); /* end of Loop0 */

```

```
} /* end of main */
```

A.2.4 Program Mach-e2

```
/* Program Mach-e2 */
```

```
/* This program creates the setting for a psychophysical experiment
whereby a compensation method is used in getting a measure of the
strength and width of Mach Bands. It uses VGA graphics in displaying
a gradient pattern which gives rise to dark and bright Mach Bands.
The slope, as well as the intensities of both the bright and the dark
uniform areas are user defined. On top of this pattern, the skeleton of
a rectangle whose dimensions are in the define file "dims.h" is drawn.
A compensation rectangle extending from the top to the bottom of
the skeletal rectangle and of user-defined width is then generated.
This rectangle, in cross-section, has an intensity distribution
corresponding to a triangular wave. The rectangle can be moved
in both directions and its intensity can be changed by the user;
therefore, it acts as the compensation tool. The compensated intensity
distribution in the skeletal rectangle as well as the original intensity
distribution are stored in two integer data files. This data can also be
viewed during the execution of the program. Looping constructs provide
several capabilities of redoing the compensation.
*/
```

```
#include <stdio.h>
#include <conio.h>
#include <graph.h>
#include <memory.h>
#include <dos.h>
#include <stdlib.h>
#include <math.h>
#include "dims.h"
```

```
/* Below, the variables used in this program are explained:
```

```
"dark_intensity" is the user-defined intensity of the dark uniform area.
"bottom_log_intensity" is the log of dark_intensity.
"intensity" is the user-defined intensity of the bright uniform area.
"top_log_intensity" is the log of intensity.
"scale" is the slope - not necessarily the mathematical slope, but one
used to calculate the number of log intensity levels between
the dark and the bright intensity levels and normalized.
"no_of_steps" is the number of distinct log intensity levels at the
given scale. This real value is then converted to integer "steps".
"step_size" is the log intensity difference between each step.
"power" is the log intensity value at each step.
"f_intensity" is the true real intensity at each step. It is when this
value is converted to integer (and thus truncated) that loss of
resolution occurs, and that unlike the linear gradient, we do not
```

have maximum resolution at the minimum slope (1).

"ls" is the starting point of the gradient, calculated so as to justify the gradient.

"left" and "right" are the amounts the skeletal rectangle (enclosing the area where compensation is to be carried out) extends to the left and right respectively from the ends of the gradient area.

"top_border" and "bottom_border" give the upper and lower_bounds of this skeletal rectangle.

"intensity_array" is a one dimensional array containing the x-coordinate values for the original Mach band generating intensity profile.

"intensity_array_comp" is an array of the same type but this time containing intensity values modified by compensation.

"changing_intensity" is the array where the triangular wave values are stored.

"increment" is the intensity difference between successive triangular wave values.

"column_no" is the array index for changing_intensity.

"w" is the user-defined width of the compensating rectangle (must be odd).

"x" is the location of the rectangle.

"i", "j", and "jj" are loop counters.

"correct_flag" is a flag testing the "correctness" of a value entered by the user in choosing what step should be taken next. In function Compensation it is also used as the test flag for the loop.

"repeat_flag", as the name suggests, is used to signal to the program that the whole display is going to be recreated after a look at the graphs of intensity_array and intensity_array_comp arrays.

"con_flag" indicates that the original profile within the skeletal rectangle is to be (re)created.

"times" is another flag signalling whether the program has been run once (1) or more (0).

"im_buffer1" is an image buffer that stores the original intensity distribution within the skeletal rectangle.

"im_buffer2" does it after compensation.

"seekresult1" and "seekresult2" are return values for the fseek library function.

"written1" and "written2" are return values for the fwrite library function.

"infile" is a binary file storing the values of intensity_array.

"infile_comp" does it for the values of intensity_array_comp.

"length" is the size (in bytes) of each item of data written to the data files.

"offset" is the offset number of bytes the pointer in the data files is located with respect to the last parameter in fseek function call (in this case, where SEEK_SET was used, with respect to the beginning).

*/

```

int i, intensity, changing_intensity[15],ls,
    dark_intensity,steps,length,
    column_no, w, con_flag, x, times, intensity_array[325],
    correct_flag,
    intensity_array_comp[325], j, jj, increment, repeat_flag,
    seekresult1, seekresult2, written1, written2;

char dummy[10], target1[5], parameter, warning[50], buffer1[40],

```

```

        buffer2[40], buffer3[40], width[2], target2[5], filename1[15],
        filename2[15];

    unsigned char far *im_buffer1;
    unsigned char far *im_buffer2;

    long offset;

    double top_log_intensity,bottom_log_intensity,power,f_intensity,
           scale,e, no_of_steps, step_size;

/* ***** */

void build_palette()
{
    short pix;
    long color,value;

    for (pix=0; pix <= 63; pix++ ) {
        value = (long)(pix);
        color = value | (value < 8) | (value < 16);

        _remappalette(pix, color);
    }
} /* end of build_palette */

/* ***** */

/* The following function displays a Mach Band generating intensity
distribution on the screen. The gradient is exponential. The display is
created in three major steps: First, the bright uniform area on the left
is created using the _rectangle graphics function. Then the gradient is
displayed as a succession of thin rectangles each corresponding to a
different log, but not necessarily distinct integer intensity level.
Finally the dark uniform area is created
as a single rectangle. After each step the intensity values at the
corresponding x-coordinates are written into the intensity array. The
variable dots is a flag defined in the "dims.h" include file. If
activated, it displays two pink dots indicating the beginning and the end
of the gradient.
*/

void build_display()
{
    e = 2.718281828;
    top_log_intensity = log((double)intensity);
    bottom_log_intensity = log((double)dark_intensity);
    no_of_steps = (intensity-dark_intensity)/scale;
    steps = (int)no_of_steps;
    if (steps >= 250)
        exit(2);
    step_size = (top_log_intensity-bottom_log_intensity)/steps;
    ls = (320-steps)/2;

    _settextcolor(65);

```

```

/* first step - displaying the dark uniform area */

_setcolor(dark_intensity);
_rectangle(_G_FILLINTERIOR,0,0,ls,200);

if (dots)
{
    _setcolor(65);
    _setpixel(ls,top_border-10);
}

for(j=0; j <= ls; ++j) /*putting into intensity array*/
{
    intensity_array[j]=dark_intensity;
} /* */

/* second step - displaying the gradient */

for (i=1; i <= steps; i++)
{
    power=bottom_log_intensity+i*step_size;
    f_intensity=pow(e,power);
    _setcolor((int) f_intensity);
    _rectangle(_G_FILLINTERIOR,ls+i,0,ls+(i+1),200);
    intensity_array[ls+i]=(int)f_intensity; /*putting
                                         into int. array*/
}

/* third step - displaying the bright uniform area */

_setcolor(intensity);
_rectangle(_G_FILLINTERIOR,ls+1+steps,0,320,200);

if (dots)
{
    _setcolor(65);
    _setpixel(ls+1+steps,top_border-10);
}

for(j=ls+1+steps; j <= 320; ++j) /*putting into int. array*/
{
    intensity_array[j]=intensity;
}

} /* end of build_display */

/* ***** */

void reading_inputs1()
{
    while (correct_flag == 0)
    {
        parameter=(char)getch();
        if ((parameter == '1')||(parameter == '2')||(parameter == '3')
            ||(parameter == '4'))
    }
}

```

```

        correct_flag = 1;
    else
        correct_flag = 0;
    } /* while */
} /* end of reading_inputs1 */

/* ***** */

/* The following function carries out the compensation procedure.
The compensated area is a rectangle of user-defined width that
extends vertically in the skeletal rectangle. This rectangle has an
intensity profile of a triangular wave. The intensity changes are
carried out by changing the "increment" between the successive values of
the triangular wave.
When "4" is entered as the parameter, the rectangle moves right,
when "3" is entered, it moves left. When "2" is entered as the parameter,
the rectangle darkens (by the smallest possible amount each time "2" is
pressed), when "1" is entered it brightens in the same way.
*/

void compensation()
{
    /* calculating the intensity of columns in the triangular wave */
    for(column_no = 0; column_no <= (w-1)/2-1; ++column_no)
    {
        changing_intensity[column_no] = intensity_array[ls-left + x + column_no]
            + increment*(column_no + 1);
    }

    for(column_no = w-1; column_no >= (w-1)/2; --column_no)
    {
        changing_intensity[column_no] = intensity_array[ls-left + x + column_no]
            + increment*(w-column_no);
    }

    /* creating the rectangle based on the triangular wave */
    for(column_no = 0; column_no <= w-1; ++column_no)
    {
        _setcolor(changing_intensity[column_no]);
        _rectangle(_G_FILLINTERIOR, ls-left + x + column_no, top_border + 1,
            ls-left + x + column_no, bottom_border - 1);
        intensity_array_comp[ls-left + x + column_no] = changing_intensity[column_no];
    }
} /* end of compensation */

/* ***** */

void reading_inputs2()
{
    while(correct_flag == 0)
    {
        parameter = (char) getch();
    }
}

```

```

        if ((parameter == '1') || (parameter == '2') || (parameter == '3')
            || (parameter == '4') || (parameter == '5'))
            correct_flag = 1;
        else
            correct_flag = 0;
    } /* while */
} /* end of reading_inputs2 */

/* ***** */

/* The following function displays a graph of the two arrays,
intensity_array, and intensity_array_comp, and their element-by-element
difference.
*/

void display_graph()
{
    for(j = 76; j <= 225; ++j)
    {
        jj = jj + 2;
        _setcolor(63);
        _setpixel(jj, 120 - intensity_array_comp[jj]);
        _setpixel(jj, 150 - intensity_array[jj]);
        _setpixel(jj, 175 - (intensity_array_comp[jj] - intensity_array[jj]));
    } /* for */
} /* end of display_graph */

/* *****
***** */

/* The main program first requests the names of the data files, and
then enters a main loop (Loop0) where it firsts asks the parameters
of the intensity distribution and then by calling the function
Build_display displays the Mach Band generating intensity distribution.
If repeat_flag is up, then these questions are not asked but the
display is carried out as before. If repeat_flag is down (the program
is run afresh) the arrays, intensity_array and intensity_array_comp, are
initialized and the values of intensity_array are also copied into
intensity_array_comp. Again, if repeat_flag is down the skeletal
rectangle is drawn, the image buffers are prepared, and the original
intensity profile in the rectangle is stored in im_buffer1. Otherwise,
these steps, which have been done previously, are skipped, and the
stored contents of the skeletal rectangle (where compensation
has been carried out during the earlier run) are retrieved from
im_buffer2 and displayed.

The program then enters the second loop (Loop1) controlled by con_flag.
The first step is the compensation.
Once the desired compensation is carried out, the user has 2 choices:
Either he (or she) can start anew (parameter = "1-4"),
disregarding the compensation carried out so far (restarting Loop1),
or he can display a graph of the original and compensated intensity
distribution (parameter = "5"). If the graph is displayed, he
is presented with another choice: He can return to the compensation
procedure from where he left (parameter = "6" and thus restarting Loop0 -

```

goes back to the beginning of the program and redisplay everything with the contents of the skeletal rectangle retrieved from memory) or he can start the whole program afresh with new parameters for dark_intensity, intensity, and slope (parameter = "anything else" and thus restarting Loop0, but this time with new parameters). If the return character is entered for slope the program writes the contents of the intensity_array and intensity_array_comp into designated files and stops.

```
*/
```

```
main() {

    /* preparing the files */

    FILE *infile, *fopen();
    FILE *infile_comp, *fopen();

    printf("\nEnter the name of data file1 (profile): ");
    gets(filename1);
    printf("\nEnter the name of data file2 (compensation): ");
    gets(filename2);

    infile = fopen(filename1, "w + b");
    infile_comp = fopen(filename2, "w + b");

    length = sizeof(int);
    offset = 0;

    _setvideomode(_DEFAULTMODE);
    repeat_flag = 0;

    do {
        /* entering Loop0 */
        if (repeat_flag == 0)
        {
            /* learning the parameters */
            printf("\nEnter slope:");
            gets(dummy);
            if (dummy[0] == '\0')
            {
                /* writing into the files */

                seekresult1 = fseek(infile, offset, SEEK_SET);
                written1 = fwrite((char *)intensity_array, length, 320, infile);
                seekresult2 = fseek(infile, offset, SEEK_SET);
                written2 = fwrite((char *)intensity_array_comp, length,
                                320, infile_comp);

                fclose(infile);
                fclose(infile_comp);
                break;
            }
            scale = (double) atoi(dummy);
            printf("\nEnter dark intensity (greater than 0): ");
            gets(target1);
        }
    } while (repeat_flag == 0);
}
```

```

dark_intensity = atoi(target1);
printf("\nEnter target intensity: ");
gets(target2);
intensity = atoi(target2);
printf("\nEnter width of rectangle (odd) : ");
gets(width);
w = atoi(width);
parameter = '3';
}
/* if */

for (jj = 0; jj < 325; ++jj)      /* initializing the array */
{
    intensity_array[jj] = 0;
    intensity_array_comp[jj] = 0;
}                                  /* */

_setvideomode(_MRES256COLOR);

build_palette();

build_display();

for(j = 0; j <= 320; ++j)      /* copying to compensation array */
{
    intensity_array_comp[j] = intensity_array[j];
}                                  /* */

if(repeat_flag == 0)
{
    /* displaying the skeletal rectangle, preparing the image
       buffers, and getting the contents of the original
       profile into memory */

    _setcolor(intensity/2);
    _rectangle(_GBORDER,ls-left, top_border,
               ls + steps + right, bottom_border);
    im_buffer1 = (char far *)malloc((unsigned int)
                                     _imagesize(ls-left,top_border,
                                                  ls + steps + right,bottom_border));
    im_buffer2 = (char far *)malloc((unsigned int)
                                     _imagesize(ls-left,top_border,
                                                  ls + steps + right,bottom_border));
    _getimage(ls-left,top_border,
              ls + steps + right,bottom_border,im_buffer1);
}                                  /* if */
else

    /* displaying the already modified contents of the skeletal
       rectangle */

    _putimage(ls-left,top_border,im_buffer2,_GPSET);

```

```

times = 0;
con_flag = 1;

if (repeat_flag == 0) /* if it is not repeated */
{
    x = 5;
    increment = 1;
    times = 1;
}

/* ----- entering Loop1 ----- */

while(con_flag)
{
    if (repeat_flag == 1) /* getting parameter in repeat */
    {
        _settextposition(0,0);
        sprintf(buffer1,"enter parameter(1,2,3,4)");
        _outtext(buffer1);
        correct_flag = 0; i = 0;

        reading_inputs1();

    } /* if */

    if (times == 0)
    {
        if (parameter == '1')
        {
            + + increment;
            if(increment == 0) + + increment;
        }
        else if (parameter == '2')
        {
            --increment;
            if(increment == 0) --increment;
        }
        else if (parameter == '4')
        {
            + + x;
            if(x == ((right + left)-w/2 + steps))
                --x;
        }
        else if (parameter == '3')
        {
            --x;
            if(x == 1)
                + + x;
        }
        else
        {
            _setvideomode(_DEFAULTMODE);
            exit(1);
        }
    } /* if times = 0, we're changing pos or inc */
}

```

```

        /* in all cases after the first run */

for(j=0; j<325; ++j)    /* initializing comp array */
{
    intensity_array_comp[j]=intensity_array[j];
}

/* compensation */

compensation();

times=0;    /* after the first run */

_settextposition(0,0);
sprintf(buffer2,"enter parameter(1,2,3,4)"
           "or 5 for graph");
_outtext(buffer2);
correct_flag=0;i=0;    /* getting parameter */

reading_inputs2();

if (parameter == '5')    /* graphical representation */
{
    _getimage(ls-left,top_border,
              ls+steps+right,bottom_border,im_buffer2);
    _clearscreen(_GCLEARSCREEN);

    jj=0;

    display_graph();

    _settextposition(0,0);
    sprintf(buffer3,"enter 6 for repeat"
               "or anything for once again");
    _outtext(buffer3);
    parameter=(char)getch();
    if (parameter == '6') /* repeat */
    {
        /* get out of loop, set rep. flag */
        con_flag=0;
        repeat_flag=1; _clearscreen(_GCLEARSCREEN);
    }
    else /* once again from the beginning */
    {
        /* get out of loop */
        con_flag=0; repeat_flag=0;
        _clearscreen(_GCLEARSCREEN);
    }
} /* if */

else /* if parameter !=5 */
{
    /* don't get out of loop */
    repeat_flag=0;
    _putimage(ls-left,top_border,im_buffer1, _GPSET);
}

}

/* ----- end of Loop1 ----- */

```

```

        _setvideomode(_DEFAULTMODE);
    }while (1);    /* end of Loop0 */

} /* end of main */

```

A.2.5 Program Her1

```

/* Program Her1.c */

/* The following program creates a Hermann Grid pattern on the screen
and allows compensation measures. The size of the squares,
the width of the streets, as well as the intensity of both
are controlled by the user. This is also true of the compensation
intensity. The compensation does not have to stay at the intersection
of the streets, but can extend to the streets at the user's discretion.
This is called smear in this program. There are several looping
constructs that allow for uninterrupted or more basic interaction.
*/

#include <stdio.h>
#include <conio.h>
#include <graph.h>
#include <memory.h>
#include <dos.h>
#include <stdlib.h>

/* Below is an explanation of variables used in the program

"no_of_squares_in_y" is the number of squares in the vertical direction.
"no_of_squares_in_x" is the number of squares in the horizontal
direction. These are calculated by the program.
"square_width" is the user-defined width of the squares in pixels.
"street_width" is the user_defined width of the streets in pixels.

"vert_unused" is the total vertical space not used by the grid pattern.
"top_unused" is the unused vertical space at the top.
"bottom_unused" is the same for the bottom.
"hor_unused" is the total horizontal space not used by the grid pattern.
"left_unused" is unused horizontal space at left.
"right_unused" is the same for the right.

"y_counter" and "x_counter" are loop counters.

"start_pos_x" is the x-coordinate where a square starts.
"end_pos_x" is the x-coordinate where a square ends.
"start_pos_y" is the y-coordinate where a square starts.
"end_pos_y" is the y-coordiante where a square ends.
"start_pos_x_comp", "end_pos_x_comp", "start_pos_y_comp", and
"end_pos_y_comp" are the corresponding terms for the compensation
squares.

```

"bk_intensity" is the user-defined intensity of the background upon which squares are laid.
 "sq_intensity" is the user-defined intensity of the squares.
 "comp_intensity" is the user-defined maximum intensity of the compensation squares.
 "extent" is the user-defined "smear", i.e., by how many pixels the compensation squares will extend beyond the square area (where streets intersect) into the streets.
 "comp_intensity_increment" is the intensity increment at each smear step in the compensation area.

"x_coord" and "y_coord" are the x and y coordinates of the fixation point.
 "prev_x_coord" and "prev_y_coord" are storage variables for the future x and y coordinates of the fixation point. They are used to capture the intensity at that point (to be outputted after the point has moved away).

"correct_flag" is a flag testing whether an input entered by the user is one of the correct choices; it is also used as a loop flag.
 "finish_flag" is a flag raised by the user to indicate that he is satisfied by the position of the fixation point and that the program can continue.
 "interactive_flag" is a flag which, when raised, creates a loop where the user can quickly change the smear and the comp_intensity_increment without losing the pattern for a long time.
 "interactive_times" is the counter of this loop.
 "repeat_flag" signals that the program should be repeated with new values of comp_intensity, bk_intensity, and smear.

*/

```
int no_of_squares_in_y, no_of_squares_in_x, street_width,
    square_width, vert_unused, top_unused, bottom_unused, hor_unused,
    left_unused, right_unused, x_counter, y_counter, start_pos_x,
    end_pos_x, start_pos_y, end_pos_y, start_pos_x_comp,
    end_pos_x_comp, start_pos_y_comp, end_pos_y_comp, comp_intensity,
    no_of_streets, comp_intensity_increment,
    bk_intensity, sq_intensity, no_of_comp_squares, extent,
    i, column_no, times, repeat_flag, correct_flag,
    finish_flag, x_coord, y_coord, prev_x_coord, prev_y_coord,
    interactive_flag, interactive_times;
```

```
char dummy[10], dummy0[5], parameter, warning[50], buffer1[100],
    dummy1[5], dummy2[5], dummy3[5], dummy4[5], buffer2[100],
    dummy5, filename[15], buffer3[100];
```

```
char far *im_buffer;
```

```
/* ***** */
```

```
void build_palette()
{
    short pix;
    long color,value;
```

```

    for (pix=0; pix <= 63; pix + + ) {
        value = (long)(pix);
        color = value | (value < < 8) | (value < < 16);

        _remappalette(pix, color);
    }
} /* end of build_palette */

/* ***** */

/* This function builds the Hermann Grid on the screen as well as
the compensation intensity and smear specified by the user.

Since the user specifies the pixel width of the squares and the
streets, the function finds out the number of squares in the
vertical and the horizontal directions, determines the unused space,
and justifies the grid. Since the graphics software used overwrites
any new creation on top of existing ones, the compensated areas
are created first. The smear, i.e., the extension of the compensated
area from the intersection of the streets into the streets
themselves, is simply done by writing overlapping squares of
increasing compensation. The superfluous sections of these squares
(those that do not belong to the streets or the intersection) are
later overwritten by the squares of the grid themselves. Although
it may look complicated, the creation of the grid is quite
straightforward.
*/

void build_grid()
{
    vert_unused = (200 - street_width) % (square_width + street_width);
    if (vert_unused/2)
        top_unused = vert_unused/2 + 1;
    else
        top_unused = vert_unused/2;
    bottom_unused = vert_unused - top_unused;

    no_of_squares_in_y = (200 - vert_unused - street_width) /
        (square_width + street_width);

    hor_unused = (320 - street_width) % (square_width + street_width);
    if (hor_unused/2)
        left_unused = hor_unused/2 + 1;
    else
        left_unused = hor_unused/2;
    right_unused = hor_unused - left_unused;

    no_of_squares_in_x = (320 - hor_unused - street_width) /
        (square_width + street_width);

    no_of_comp_squares = 1 + extent;

```

```

comp_intensity_increment = (comp_intensity - bk_intensity)/
                           (no_of_comp_squares);

_setcolor(bk_intensity);
_rectangle(_G_FILLINTERIOR,0,0,320,200);

if(left_unused > 0 && top_unused > 0)
{
--left_unused; --top_unused;
}

/* ----- writing the compensation and the squares ----- */

for(y_counter=0; y_counter <= no_of_squares_in_y-1; ++y_counter)
{
for(x_counter=0; x_counter <= no_of_squares_in_x-1; ++x_counter)
{
start_pos_x=left_unused+street_width+(x_counter*(square_width+
street_width))+1;
end_pos_x=start_pos_x+square_width-1;
start_pos_y=top_unused+street_width+(y_counter*(square_width+
street_width))+1;
end_pos_y=start_pos_y+square_width-1;

if ((y_counter <= (no_of_squares_in_y - 2))&&
(x_counter <= (no_of_squares_in_x - 2)))
{
start_pos_x_comp=end_pos_x+1;
end_pos_x_comp=start_pos_x_comp+street_width-1;
start_pos_y_comp=end_pos_y+1;
end_pos_y_comp=start_pos_y_comp+street_width-1;

for(i=0; i <= no_of_comp_squares-1; ++i)
{
_setcolor(bk_intensity+comp_intensity_increment*(i+1));
_rectangle(_G_FILLINTERIOR, start_pos_x_comp+extent+i,
start_pos_y_comp+extent+i,
end_pos_x_comp+extent-i, end_pos_y_comp+extent-i);
}
}

_setcolor(sq_intensity);
_rectangle(_G_FILLINTERIOR, start_pos_x, start_pos_y, end_pos_x,
end_pos_y);

} /* end of loop in X-direction */
} /* end of loop in Y-direction */
} /* end of build_grid */

/* ***** */

/* The following function sets the fixation point. The user
signals that he is satisfied with the location of the point
by setting the finish_flag.

```

```

    */
void set_fixation_point()
{
    x_coord = 160;
    y_coord = 100;

    correct_flag = 0;
    finish_flag = 0;

    im_buffer = (char far *)malloc((unsigned int)
                                   _imagesize(1,1,1,1));

    _getimage(x_coord, y_coord, x_coord, y_coord, im_buffer);

    _setcolor(65);
    _setpixel(x_coord, y_coord);

    _settextcolor(65);
    _settextposition(0,0);
    sprintf(buffer1,"enter 1, 2, 3, 4 to move dot"
            "\n f to continue");
    _outtext(buffer1);

    while(correct_flag!= 1)
    {
        prev_x_coord = x_coord;
        prev_y_coord = y_coord;
        parameter = (char)getch();
        if(parameter == '1')
            ++y_coord;
        else if (parameter == '2')
            --y_coord;
        else if (parameter == '4')
            ++x_coord;
        else if (parameter == '3')
            --x_coord;
        else if (parameter == 'f')
        {
            finish_flag = 1;
            correct_flag = 1;
        }
        else
            correct_flag = 0;

        if (finish_flag == 0)
        {
            _putimage(prev_x_coord,prev_y_coord,im_buffer,_GPSET);
            _getimage(x_coord, y_coord, x_coord, y_coord,im_buffer);
            _setpixel(x_coord, y_coord);
        } /* if */
    } /* while */
} /* end of set_fixation_point */

```

```

/* *****
***** */

/* The main program first prompts the user for the specifications
of the grid and the compensation. After displaying the grid on the
screen, the pink fixation point is set. Two nested loops follow.
In the inner so-called interactive loop (Loop2), the user can
change the compensation and the smear slowly without losing the
grid except for a second. In the outer so-called repeat loop
(Loop2), the user can change the compensation and the smear
by inputting the new values outside the grid. Of course, there is
also the option of starting anew by exiting to the outermost loop
(Loop0).
The parameters are outputted to the screen and to an output file
for reference when exiting from the program.
*/

main() {
    /* preparing the files */

    FILE *outfile, *fopen();

    printf("Enter the name of the output file: \n");
    gets(filename);

    outfile = fopen(filename, "r + w");

    repeat_flag = 0;

    /* ----- */
    /* start of main loop (Loop0) */

    do {
        printf("\nEnter the pixel width of the squares: ");
        gets(dummy);
        if (dummy[0] == '\0')
        {
            fprintf(outfile, "square width is %d , street width is %d , \n"
                "background intensity is %d , intensity of the squares is \n"
                "%d , compensation intensity is %d , smear is %d",
                square_width, street_width, bk_intensity, sq_intensity,
                comp_intensity, extent);

            printf("square width is %d , street width is %d , \n"
                "background intensity is %d , intensity of the squares is \n"
                "%d , compensation intensity is %d , smear is %d",
                square_width, street_width, bk_intensity, sq_intensity,
                comp_intensity, extent);

            fclose(outfile);

            break;
        } /* if */
    }

```

```

square_width = atoi(dummy);
printf("\nEnter the pixel width of the streets: ");
gets(dummy0);
street_width = atoi(dummy0);
printf("\nEnter intensity of background: ");
gets(dummy1);
bk_intensity = atoi(dummy1);
printf("\nEnter intensity of squares: ");
gets(dummy2);
sq_intensity = atoi(dummy2);

times = 1;

/* ----- */

/* start of repeat loop (Loop1) */

while((repeat_flag == 1)|| (times == 1))
{
    if(repeat_flag == 1)
        _setvideomode(_DEFAULTMODE);
    printf("\nEnter compensation intensity: ");
    gets(dummy3);
    comp_intensity = atoi(dummy3);
    printf("\nEnter extent of smear in pixels: ");
    gets(dummy4);
    extent = atoi(dummy4);

    _setvideomode(_MRES256COLOR);
    build_palette();

    interactive_flag = 1;
    dummy5 = '0';
    interactive_times = 0;

/* ----- */

/* start of interactive loop (Loop2) */

while(interactive_flag == 1)
{
    if(dummy5 == '1')
        ++comp_intensity;
    if(dummy5 == '2')
        --comp_intensity;
    if(dummy5 == '3')
        ++extent;
    if(dummy5 == '4')
        --extent;

    build_grid();

```

```

if (interactive_times == 0)
    set_fixation_point();

/* interactive loop (Loop2) questions */

correct_flag = 0;
_settextposition(0,0);
sprintf(buffer2, "enter 1/2 for compensation, "
              "\n3/4 for smear, c for cont");
_outtext(buffer2);

while(correct_flag == 0)
{
    dummy5 = (char)getch();
    if((dummy5 == '1') || (dummy5 == '2') || (dummy5 == '3') || (dummy5 == '4'))
    {
        sprintf(buffer2, "\nobtained");
        _outtext(buffer2);
        correct_flag = 1;
        interactive_flag = 1;
        interactive_times = 1;
    }
    else if(dummy5 == 'c')
    {
        correct_flag = 1;
        interactive_flag = 0;
    }
    else
        correct_flag = 0;
} /* correct_flag while */

} /* end of Loop2 */

/* ----- */

/* setting repeat criteria (Loop1) */

correct_flag = 0;

_settextposition(0,0);
sprintf(buffer2, "enter r for repeat, "
              "\n s for start ");
_outtext(buffer2);

while(correct_flag != 1)
{
    parameter = (char)getch();
    if(parameter == 'r')
    {
        repeat_flag = 1;
        correct_flag = 1;
    }
    else if (parameter == 's')

```

```

        {
            repeat_flag=0;
            correct_flag=1;
        }
        else
            correct_flag=0;
    } /* while */

    times=0;

} /* end of Loop1 */

/* ----- */

    _setvideomode(_DEFAULTMODE);
}while (1);    /* end of Loop0 */
}    /* end of main */

```

Appendix 3. Preliminary Experiments and Results

To test the feasibility of an experimental design based on a a computer-controlled high resolution graphics terminal, and incorporating the compensation method, a number of programs were developed, and a series of simple preliminary experiments were carried out. They are described below in more detail.

A.3.1 Apparatus

The computer used was a Zenith Data Systems PC which controlled a NEC Multisync 3D color graphics terminal. The graphics capabilities were provided by Microsoft QuickC Compiler.

A.3.2 Programs

Programs mach-l1, mach-e1, mach-l2 and mach-e2 display a Mach Band-generating intensity distribution on the screen. The parameters of this distribution, i.e., the intensities of the uniform dark and bright areas as well as the slope, can be varied by the user each time the program is run. Programs mach-l1 and mach-l2 employ a linear intensity distribution (a ramp) whereas programs mach-e1 and mach-e2 use an exponential one. In the following discussion, intensity will be taken as one of the 64 levels available with the graphics software, not as lumens/sq.m or candelas/sq.m. This simplifies looking at intensity, and the results can easily be converted to basic dimensions once the luminance output at each intensity level and at that specific setting of the brightness dial is photometrically measured. Since terminals will show variations, it also makes the argument more transferable. The mathematical slope thus will be defined to have dimensions of (difference of intensity level between the two uniform areas) / (spatial extent of the gradient in pixels). The user-specified slope in the experiments is not necessarily the mathematical slope. In programs using the exponential gradient, for instance, it is important to note that since there is no constant slope, the user-specified slope is used as the mathematical slope of a "hypothetical" linear gradient between the two intensity levels.

Since the graphics software was not using the maximum spatial resolution possible (the smallest dot being 4 physical pixels on each side), the pixel references below refer to the maximum resolution possible with the software.

When the gradient is linear, the user-specified slope is the same as the mathematical slope. Therefore, a ramp ascending from lowest intensity (0) to highest intensity (63) and displaying every intensity level (with each level extending only one pixel) has a slope of 1 (64 intensity levels/64 pixels). A ramp which extends between the same intensity levels but at half the spatial extent and thus

displays not every but everyother intensity level in neighboring pixels has a slope of 2 (64 intensity levels/32 pixels). (Fig. 23a)

In the case of exponential intensity distribution, where the slope in the mathematical sense changes continuously across the gradient, the user-specified "experimental" slope is the mathematical slope of a hypothetical linear gradient between the two uniform areas. In this sense, the user-specified slope is used to change the spatial extent of the exponential gradient. For instance, if the gradient ascends from the lowest intensity (0) to the highest intensity (63), the spatial extent be will be 64 pixels for a slope of 1, and 32 pixels for a slope of 2 and so forth (Fig. 23b) The gradient follows the formula:

$$i = e^{(bx)}$$

where i is intensity, and b is the user-specified slope. Due to truncation to integer intensity values as required by the graphics software, the calculated intensity at neighboring pixels of the gradient will not, unlike the linear gradient, necessarily be distinct this time. If we assume a perfectly natural logarithmic transduction in the eye, following the photoreceptors, the exponential gradient will collapse into a linear gradient having the user-specified slope.

In all of the programs, a skeletal rectangle, i.e., a rectangle with only borders is drawn. All compensation is carried inside this rectangle, whose dimensions are in an "include" file.

In programs mach-11 and mach-e1, a thin, tall rectangle, or more accurately, a column spanning the height of the skeletal rectangle, and of the smallest possible width (1 pixel), forms the immediate area compensated by the user. A marker is placed outside the skeletal rectangle to show the location of this column. The user is free to manipulate the intensity of the column as he/she sees fit (Fig. 24). Successful compensation of successive thin areas requires some experience but after a while can be managed quite satisfactorily and to a high degree of accuracy.

In programs mach-l2 and mach-e2 , on the other hand, a different compensation scheme is employed. The compensated area, this time, is a movable rectangle of variable (but odd-numbered) width (in pixels), and having the intensity distribution of a triangular wave with the apex (or zenith) at the center pixel. This triangular wave "rides" on the original intensity distribution. It is important to note that, unlike programs mach-l1 and mach-e1, where a column whose intensity is modified stays put as the marker moves on, when the triangular wave moves, the the area it leaves behind is restored to its original intensity distribution (Fig. 25). The user can increase or decrease the intensity difference between neighboring pixels in the triangular wave (a difference of 0 is not allowed), but since the difference is constant throughout the neighboring pixels in the wave, this action amounts to a vertical expansion or a constriction of the wave. The freedom of movement in these programs is, therefore, much more restricted and the choices are simple.

Program her1 generates a Hermann Grid pattern on the screen and incorporates a compensation procedure. In this program, the widths of the squares and the streets, as well as their intensities, are specified by the user. This is also true for the "compensation" intensity, which is beamed at (or subtracted from) the intersections to alleviate the illusion. A movable pink marker is available for the fixation of the eye. During the development of the program, it was found that if the compensated area was restricted to the intersections of the streets, undesirable edge effects developed at intensities necessary for compensation. For this reason, the compensated area extends from the intersection into the streets by an amount called "smear" and specified by the user. The compensation intensity decreases (or increases) as one goes from the intersection (where it is highest (or lowest)) to the streets along the direction of the smear, creating an intensity ladder of the shape of "+ " sign (Fig. 26).

Two other programs were written for possible Mach Band experiments, one employing again a linear, the other an exponential gradient. In these programs, instead of a skeletal rectangle, a full-bodied rectangle of half the maximum intensity (31) was displayed. The method was similar to programs mach-l1 and mach-e1 in that a thin column rather than a triangular wave was used. Instead of compensation, however, this time, the intensity of the column was matched to the intensity

of the Mach Band (naturally, the intensity of the bright uniform area was kept below the maximum intensity level so that the Mach Band percept was not above this level - of this we could not be certain until many results are evaluated). Preliminary trials with these programs revealed that it was impossible to match the intensity accurately. Large changes in the intensity of the matching column could go unnoticed, due in part to the small border between the column and the location in the intensity profile of the Mach Band. Unlike compensation experiments, however, even when a large number of successive thin columns had been "coarsely" matched, it was generally impossible to see any direction for further improvement or whether improvement was being done. Therefore, experiments with these programs were not carried out. It should also be noted that these trials suggest how tricky it is to match two neighboring intensities, less intensities at different locations which was how many matching experiments were carried out.

A.3.3 Limitations

In the Mach Band experiments there were 2 major limitations, and both were related to the capability of the graphics software. First, the maximum intensity range was limited to 64 steps. With this limited range, it is difficult to create subtle intensity differences. Although this is evident when one thin column is considered, it is especially true when the increment between the successive pixels in the triangular wave is considered. In the latter case, this lack of intensity resolution severely limits acceptable compensation to a few increment values. Furthermore, this narrow range degrades the smoothness of transition between the dark and bright uniform areas in the Mach Band gradient. In the linear case, the individual rectangles making up the gradient (see program for construction) are visible up close, but barely at distances where the experiments were carried out. In the exponential case, however, because of the conversion and subsequent truncation of double values to integers in the calculation, these rectangles do not all have distinct intensity levels, thus extend over greater distances and are clearly visible. The second limitation was the lack of optimum spatial re-

solution. As explained previously, four physical pixels make up one "experimental" pixel. This creates a problem when steep slopes are used. In the programs where compensation was achieved with tall, thin columns, it would be more difficult and time-consuming to compensate with a rectangle (or column) thinner than what was available, but the programs where the triangular wave was used could be designed with a much better and smoother envelope such as a Gaussian.

The luminance output at different points on the screen was measured at different intensity levels and with different manual brightness dial settings by a photometer borrowed from the Human Factors Laboratory with the permission of Dr. Bob Beaton. The difference in luminance output at different locations on the screen is considerable, but the transition is smooth, and since Mach Bands seem to be a much more localized phenomena [von Békésy 1968b], it may not affect their perception to the same degree. Indeed, the difference of brightness at different locations is hardly visible, certainly not at viewing ranges used in the experiments. Nevertheless, this should be corrected in any comprehensive study by obtaining an intensity signature of the screen and compensating for it. Due to the limitations of the graphics software used in these programs, this was not possible. In a system where each pixel is individually controlled, this can be achieved, but the computer must be a fast one, preferably a workstation.

Another important point regarding screen luminance is about the manual contrast and brightness dial settings. In the experiment, the contrast setting was understandably used at the maximum level to get the greatest dynamic range. The behaviour of the brightness dial, however, is more complex and relevant, and thus was examined by taking measurements with the photometer. Increasing the brightness setting from the lowest to the highest level increased the luminance output at all 64 intensity levels (0-63), but the lower intensities showed a more pronounced increase. Therefore, the brightness dial is a nonlinear device changing the luminance output at different intensities unequally. Since the brightness dial is unmarked, to avoid arbitrariness between different experimental runs, two sets of experiments were carried out, one when the brightness dial was at the highest, the other when it was in the lowest setting. The results reported below belong to the highest setting.

Unexpected glare problems occurred with the Hermann Grid, especially after prolonged exposure to the stimulus. Therefore, experiments with the Hermann Grid have been postponed until the acquisition of a fine mesh glare screen. Informal investigation by the experimenters, however, was encouraging.

A.3.4 Experiments

Due to glare problems with the Hermann Grid, as stated above, only Mach Band experiments were carried out. Although the dark band was also visible, measurements were taken only for the bright band as it was much more prominent. The viewing distance was 30 in., and free viewing was used. Two subjects, both graduate students in their twenties and both with corrected vision, participated in the experiments. Both did a complete set of experiments at both the highest and the lowest brightness dial settings. Maximum intensity difference between the dark and bright uniform areas (64) was used, since at significantly lower intensity levels (32 for instance), the experiments were difficult and the results were ambiguous. The following experiments were carried out:

1. Column-by-column compensation experiments on a Mach Band-generating linear intensity gradient (using program mach-l1) at slopes of 1, 2, 4, and 8. The slope when the Mach Band disappeared was also reported.
2. Column-by-column compensation experiments on a Mach Band-generating exponential intensity gradient (using program mach-e1) at slopes of 1, 2, 4, and 8. The slope when the Mach Band disappeared was also reported.
3. Triangular wave compensation experiment on a linear intensity gradient (using program mach-l2) at slopes of 1 and 4.

4. Triangular wave compensation experiment on an exponential intensity gradient (using program mach-e2) at slopes of 1 and 4. In both of these experiments (3 and 4), both subjects used a rectangle of three columns as this gave the best results. (The results at slope 4 were especially good) At steeper slope values, the smallest multipixel wave (3, which was used) became too large, and the experiments at these slope values were not carried out.

The results of these experiments are depicted in graphs 27-38. The first 8 belong to subject N. A.; the remaining 4 belong to subject C. H. T. It is important to note that the word "compensation" used in the graph descriptions is the negative of the "compensation" used in section 4.3 of the thesis. Here, it is the amount to be subtracted from the original profile to obtain a gradient free of Mach Band percepts. (This minor modification was adopted to make the descriptions more legible) The third profile from the top is the approximation of an actual Mach Band percept by adding the compensation (in the new sense) to the original profile.

Since this experiment was preliminary in nature, and since little data was collected, a quantitative analysis of the data will not be pursued here to avoid premature conclusions. Qualitatively, and assuming that we have good approximations of the Mach Band percepts, the following observations can be made:

- For both linear and exponential gradients, the width of the Mach Band, decreased sharply as the slope was increased. (Compare Fig. 27 with Fig. 28, and Fig. 29 with Fig. 30 - all for N. A.) This agrees with the earlier findings.
- Such a clear conclusion cannot be reached about the magnitude of the bright band. Although, in one clear instance, both subjects have used higher compensation when the slope of the exponential gradient was 4, than when the slope was 1 (compare Fig. 33 with Fig. 34 for N. A., and Fig. 37 with Fig. 38 for C. H. T.), this can be attributed more to the fact that a triangular wave was used for compensation, and thus that the number of choices was limited. (Compare, for instance, Fig. 27 with Fig. 31 or Fig. 28 with Fig. 32 - all for N. A.) Indeed,

other results do not seem to support such a conclusion. Interpersonal differences are noticable, with N. A. using, in general, larger compensation than C. H. T. especially for the exponential gradient (not shown). These results are again consistent with earlier research.

- For both subjects, the exponential gradient seems to create more powerful Mach Band percepts. (Compare Fig. 27 with Fig. 29 or Fig. 28 with Fig. 30 for N. A., and Fig. 35 with Fig. 36 for C. H. T.)
- The Mach Band became very ambiguous or dissapeared at slope 16 for N. A. and at slope 20 for C. H. T. when the linear gradient was used. For both observers, this occured at slope 20 with the exponential gradient.

A.3.5 Projection

Since all the parameters of the intensity distribution are specified by the user, these programs offer a flexibility which is not readily available with Maxwellian systems. Furthermore, separation of independent variables such as slope and absolute luminance is also possible with these programs. This has not yet been explored in depth in these preliminary experiments; but future experiments can easily be planned. Compensation experiments can be carried out more precisely with a higher resolution system or software, which would, among other advantages, also allow the construction of a better compensation envelope. It is our belief that compensation gives us an important if indirect measure of Mach Bands, and that these programs and preliminary experiments can be extended and improved to give us a better understanding of this phenomenon. This is also true for the Hermann Grid.

Vita

Nuri Aktuna was born in Istanbul, Turkey in 1962. He completed his B.S. in Electrical Engineering at Bogazici (Bosphorus) University in Istanbul in 1986. He is expected to obtain the M.S. degree in Electrical Engineering from Virginia Polytechnic Institute and State University in August 1990.