INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106

73-24,551

DiGIUSEPPE, Jack Leonard, 1939-
PICTURE BANDWIDTH COMPRESSION AND THE CHARACTERISTICS OF HUMAN VISION.

The University of Michigan, Ph.D., 1973
Engineering, electrical

University Microfilms, A XEROX Company, Ann Arbor, Michigan

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED.

PICTURE BANDWIDTH COMPRESSION AND THE CHARACTERISTICS OF HUMAN VISION

by

Jack Leonard DiGiuseppe

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering)
in The University of Michigan
1973

Doctoral Committee:

Professor Howard Diamond, Co-Chairman
Professor Edgar H. Sibley, Co-Chairman
Professor Richard K. Brown
Professor Benjamin F. Cheydleur, Oakland University
Professor Norman R. Scott
Professor Franklin H. Westervelt, Wayne State University

ACKNOWLEDGEMENTS

The author wishes to express his gratitude to the following individuals and organizations for the part that each played in the successful conclusion of this investigation:

Prof. Edgar Sibley, present chairman of the committee, for initial motivation, and constant encouragement and guidance throughout the investigation, and for his invaluable help in reading and editing the first draft;

Prof. Howard Diamond, present co-chairman of the committee, whose encouragement and support through the facilities of the Microelectronics Laboratory enabled me to continue my research;

Prof. Norman Scott, original chairman and present member of the committee, for this author' introduction into the exciting world of digital computation, for this author's first research assistantship, and for his continuous sincere interest;

Professors Richard K. Brown, Benjamin F. Cheydleur, and Frank H. Westervelt, members of the doctoral Committee, for advice, encouragement, and helpful criticism;

Prof. R.C.F. Bartels and the staff of the Computing Center, for understanding assistance and support;

The Air Force Office of Scientific Research (Contract AFOSR-72-2219) for financial support toward obtaining the digitized pictures and preparing the manuscript,

Kenneth Wacks, of the Massachusetts Institute of Technology, for unselfish cooperation and time spent providing this author with digitized pictures and reconstructions of the processed pictures;

Lothar Rossol, of the General Motor Technical Center for generously providing this author with the first set of digitized pictures;

Donna Gabaccia, aided by Paula Harbison and Geraldine Cox, for a fantastic job of typing this manuscript and replicating the figures;

Jean H. DiGiuseppe, my wife, for patient and cheerful transcription of the rough draft and her unfailing confidence which was so essential to the completion of this dissertation.

TABLE OF CONTENTS

page

ACKNOWLEDGEMENTS ii
LIST OF TABLES vi
LIST OF FIGURES vii
LIST OF APPENDICES ix

Chapter

I INTRODUCTION 1

II REDUNDANCY REDUCTION 7

2.0 Introduction 7
2.1 One-dimensional redundancy reduction techniques 10
2.2 Two-dimensional techniques 22
REFERENCES. 35

III CHARACTERISTICS OF HUMAN VISION 38

3.0 Introduction 38
3.1 The structure of the human eye 38
3.2 Characteristics of the optical system . . 44
3.3 Characteristics of the retinal photo-receptors. 63
3.4 Characteristics of the retina and the composite visual pathway 80
3.5 Summary 116
REFERENCES 119

IV PSYCHOPHYSICAL MODELS 122

4.0 Introduction 122
4.1 A model for the transformation of a luminance distribution 122
4.2 Spatial frequency channels in human vision 167
4.3 A model for the transformation of a spatial luminance distribution into a perceived brightness distribution 190
REFERENCES 194

V APPLICATION OF THE PERCEPTION MODEL TO PICTURE BANDWIDTH COMPRESSION 197

5.0 Introduction 197
5.1 The picture viewing environment 239
5.2 A simplified model of the eye 239
REFERENCES 293

TABLE OF CONTENTS, continued

Chapter page

VI SUMMARY AND COMMENTS 294

APPENDIX A 299

APPENDIX B 303

APPENDIX C 308

APPENDIX D 319

APPENDIX E 329

APPENDIX F 342

BIBLIOGRAPHY 346

LIST OF TABLES

2-1 Contour coding data for figures 16, 17, and 18 . . 27
2-2 Bandwidth compression data 34
3-1 Summary of contrast sensitivity measurements and the associated experimental parameters . . .109
5-1 The relative photopic luminous efficiency of radiation 201
5-2 Typical values of luminance and illumination . . .205
5-3 Reflection factors of common paper and inks . . .207
5-4 The specular and diffuse luminance (foot-Lamberts) of various reflecting surfaces 209
5-5 Effect of viewing distance on the limit of resolution 234
5-6 One-dimensional simulation results for "face". . .270
5-7 One-dimensional simulation results for "cameraman" 270
5-8 One-dimensional simulation results for "crowd" 271
5-9 Segment size, number of segments and total number of non-redundant Fourier Coefficients. 282
5-10 Two-dimensional simulation results for "face". . .283
5-11 Two-dimensional simulation results for "cameraman". 284
5-12 Two-dimensional simulation results for "crowd" 285
5-13 Number of non-redundant Fourier Coefficients retained for each picture and each simulated distance 286
5-14 Fourier coefficients retained in the one-dimensional simulation 292

LIST OF FIGURES

1-1 Schematic diagram of a general communications system 5
2-1 Stockham's model of the eye and his experiment to determine H. 21
2-2 12 element statistical encoding pattern. 23
3-1 The visual pathway 40
3-2 The Gullstrand simplified schematic eye 48
3-3 The Gullstrand simplified schematic eye 49
3-4 The relation between pupillary diameter and field luminance 51
3-5 Modulation transfer functions 60
3-6 Calculated contrast transfer functions 62
3-7 Rod receptor cell density 65
3-8 The relative rod spectral sensitivity curve 67
3-9 Densities of rods and cones across the retina 73
3-10 Spectral sensitivities of dark-adapted peripheral rods, foveal cones, and peripheral cones 75
3-11 Absorption spectra for three classes of cones found in the human retina 77
3-12 The variation of visual acuity with luminance. . 81
3-13 The relation between threshold contrast and stimulus area; 6 sec. exposure time 85
3-14 The relation between threshold contrast and stimulus area; 15 sec. exposure time 87
3-15 The effect on the perceived luminance L_m of a test disc L_t due to its surround L_b. 91
3-16 The modulation transfer functions for the visual system at four different mean retinal illuminances 104
3-17 The contrast sensitivity of the composite visual pathway 114
4-1 Modulation transfer functions for the eyes of three subjects 135
4-2 Normalized modulation transfer functions, averaged among three subjects 136
4-3 Diagram of the synaptic connections of the primate retina 139
4-4 Visual system transfer functions 149
4-5 Machbands and the influence of slope of gradient on the appearance of Mach bands . . .152
4-6 Sensory unit response due to spatial interaction 154
4-7 A luminance gradient and the associated response of the sensory unit given in 4-6. . .154
4-8 Graphical and mathematical representations of six models of inhibitory networks 156

4-9 Sine wave response of the complete visual system 160
4-10 The Fourier transform 163
4-11 An idealized composite transfer function . . . 165
4-12 Single-channel model 169
4-13 Multiple-channel model 169
4-14 Contrast thresholds for complex grating patterns 174
4-15 Block diagram of the single channel model. . . 182
4-16 Block diagram of the multiple channel model 182
4-17 Psychometric functions for the single channel model 184
4-18 Psychometric functions for the multiple channel model 184
4-19 Simple and complex psychometric functions . . 186
4-20 A model for the prediction of human brightness perception 192
5-1 Relative brightness versus pupil diameter. . . 202
5-2 Magnification and visual angle of a sine wave gratings 211
5-3 Periodic extension of a two-dimensional picture luminance function 215
5-4 Slope and period of two dimensional sinusoids 220
5-5 A periodic pulse of duration a and period 2L 224
5-6 c_n and the envelope of c_n 224
5-7 A model for the determination of imperceptible spatial frequency components 240
5-8 Computer simulation procedures 243
5-9 Composite sine wave contrast sensitivity . . . 249
C-1 Line # 100 of "CROWD", one-dimensional program, viewing distance 8 times height. 309
C-2 Line # 100 of "CROWD", one-dimensional program, viewing distance 16 times height 310
C-3 Line # 100 of "CROWD", one-dimensional program, viewing distance 32 times height 311
C-4 Line # 100 of "CROWD", one-dimensional program, viewing distance 64 times height 312
C-5 Line # 200 of "CROWD", one-dimensional program, viewing distance 8 times height 313
C-6 Line # 200 of "CROWD", one-dimensional program, viewing distance 16 times height 314
C-7 Line # 200 of "CROWD", one-dimensional program, viewing distance 32 times height 315
C-8 Line # 200 of "CROWD", one-dimensional program, viewing distance 64 times height 316

C-9 Fourier domain for "CROWD", one-dimensional program, viewing distance 32 times height 317
E-1 Line # 100 of "CROWD", two-dimensional program viewing distance 4 times height. 330
E-2 Line # 100 of "CROWD", two-dimensional program viewing distance 8 times height. 331
E-3 Line # 100 of "CROWD", two-dimensional program viewing distance 16 times height 332
E-4 Line # 100 of "CROWD", two-dimensional program viewing distance 32 times height 333
E-5 Line # 100 of "CROWD", two-dimensional program viewing distance 64 times height 334
E-6 Line # 200 of "CROWD", two-dimensional program viewing distance 4 times height. 335
E-7 Line # 200 of "CROWD", two-dimensional program viewing distance 8 times height. 336
E-8 Line # 200 of "CROWD" two-dimensional program viewing distance 16 times height 337
E-9 Line # 200 of "CROWD", two-dimensional program viewing distance 32 times height 338
E-10 Line # 200 of "CROWD", two-dimensional program viewing distance 64 times height 339
E-11 Fourier domain for "CROWD", one-dimensional program, viewing distance 32-times height 340

LIST OF APPENDICES

		page
APPENDIX A:	256 X 256 X 256 Pictures	299
APPENDIX B:	FORTRAN program for Zero-Order Interpolator	303
APPENDIX C:	One-Dimensional simulation Results for "CROWD".	308
APPENDIX D	Two-dimensional Simulation Program	319
APPENDIX E	Two-dimensional Simulation Results for "CROWD".	329
APPENDIX F:	Reconstructed Results	342

CHAPTER I--INTRODUCTION

The compression of pictorial data is one aspect of a more general problem--the efficient and accurate transmission of information. It is the problem of designing a communications system which meets the information requirements of the receiver and, at the same time, minimizes the required resources. One desirable resource to be minimized is channel bandwidth. This is true because generally increased bandwidth implies increased "cost." For example, if the "channel" is a wideband leased communications line, the cost in dollars per mile increases with channel capacity (reference 1-1). If the "channel" is a segment of electromagnetic frequency space, then the "cost" is a decreased supply of a scarce resource, electromagnetic frequency space.

A diagram of a general communication system is given in Figure 1-1. This communications system includes two methods for minimizing the required channel capacity or bandwidth: channel encoding and source-receiver encoding (reference 1-2). A particular class of communication systems will be considered here, namely those which communicate visible information. This means that the ultimate receiver is the human visual sensory system.

The source information for a visual data communication system is generally the image of real objects in motion. This image is a consequence of the reflectivity properties of the objects and the spectral and distributional characteristics of the sources of illumination. The communications system

may process the image optically in which case the source is characterized by intensity and phase information, both of which are functions of wavelength and time. For other forms of processing, the image is treated as a two-dimensional intensity distribution. This study will confine itself to two-dimensional intensity distributions which do not vary with time, and the intensity of each point in the image is independent of wavelength monochrome still pictures.

The two-dimensional intensity distribution serves as input to the source-receiver encoding section of the visual data communications system. In a formal sense, the purpose of this section of the system is to minimize the channel capacity required to transmit the image according to a specified fidelity criterion. This problem has been formally described by Limb (reference 1-2). A source-receiver encoding technique accomplishes bandwidth reduction at the expense of image fidelity, but it is not the mathematical accuracy of the image that is important, but the qualitative evaluation of the received image by the observer. Therefore, there exists some complex and yet unknown relationship between the quality of images, the bandwidth requirements, and the parameters of the encoding technique. So, source-receiver encoding not only depends upon the characteristics of image intensity distribtuions, but also depends upon the characteristics of the receiver, in this case the human visual system.

It is the purpose of this research to perform a comprehensive study of the physiological and psychophysical

characteristics of human vision in order to better understand source-receiver encoding suitable for visual information. One of the goals of this research was to obtain a tractable model of human vision which could be used to predict imperceptible components of an image. Such a model could then be used to prefilter the image before the application of conventional compression techniques, the purpose being to enhance the results obtained from these conventional compression techniques.

Chapter II is a summary of feature extraction and redundancy reduction techniques. This summary establishes a basis for comparative efficiency of bandwidth compression techniques and makes possible the selection of a suitable compression algorithm for evaluation of the perceptual model. In Chapter III, some physiological and psychophysical characteristics of human vision are summarized. Performance measurements of the human eye are described, as well as the experimental conditions under which the measurements were obtained. Chapter IV is concerned with models that have been proposed to describe different components of the visual system, and a model is selected for investigation. In Chapter V the parameters of a typical picture viewing environment are defined in order to establish the operating range of the eye. Suitable parameters are then selected for the model, and computer simulation is used to test the effects of the model on digitized pictures.

The perceptual model, which is based on the contrast sensitivity function for human vision, is found to be useful for an investigation of the effects of distance on human vision and for encoding in the Fourier domain.

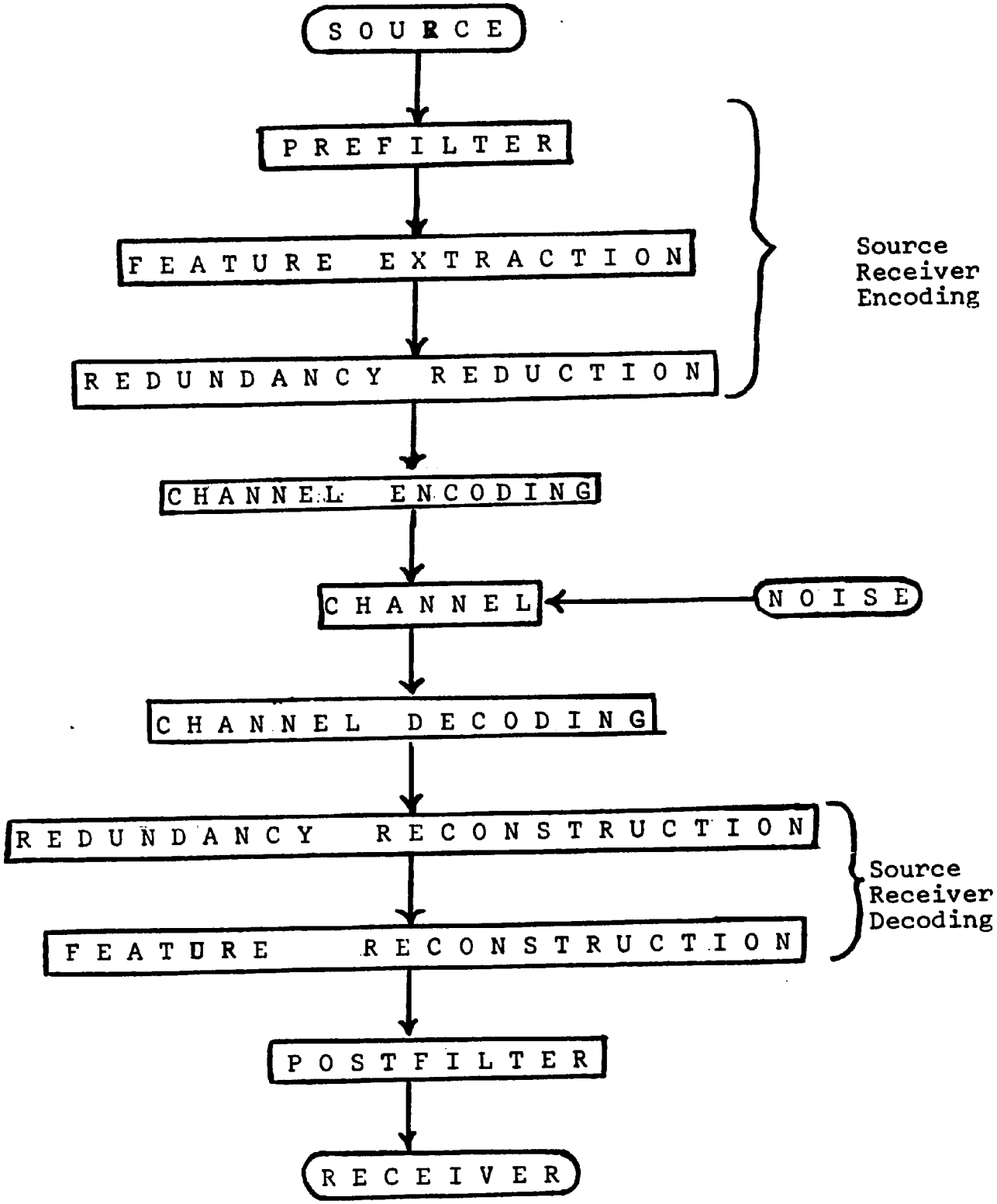


Figure 1-1. Schematic Diagram of a General Communications System.

REFERENCES

1-1 M. Edwards, "Data Transmission Links - How to Get the Most from Today's Lines," The Data Communications User, Aug-Sept 1971, pp. 515-518.

1-2 J.O. Limb, "Source-Receiver Encoding of Television Signals," Proc. IEEE, vol. 55, no. 3, March 1967, pp. 364-379.

CHAPTER II
REDUNDANCY REDUCTION

2.0 Introduction

Consider the ensemble of source messages to be transmitted by a communications system. If the requirements are such that every message be received without error, then no redundancy reduction is possible. If, on the other hand, the specifications include a fidelity criterion which translates into a maximum allowable difference between each source message and each received message, then redundancy reduction techniques can be applied. A fidelity criterion is, in a sense, a mapping whose domain is the ensemble of source messages. Each source message is mapped into a set of messages that satisfy the fidelity criterion. Now consider the messages in a set corresponding to a given source message. These messages require differing amounts of channel bandwidth for their transmission. Some will require less bandwidth than the original source message while others will require more bandwidth. The redundancy reduction problem becomes one of choosing that message in the set which satisfies the fidelity criterion and requires the least channel bandwith. When this problem is solved for all possible source messages, then a new ensemble of messages is obtained which satisfy the fidelity criterion, and which require channel bandwidth less than or equal to that required by the original source ensemble.

A semantic distinction is being made here between channel encoding and redundancy reduction. If a statistical description of the source messages is availalbe, then channel encoding can be applied directly to the source messages or, where redundancy reduction techniques have been used, channel encoding can be applied to the message ensemble derived from the fidelity criterion.

Some redundancy reduction techniques are applied to a transformation of the source data. These transformations have been called parameter extraction techniques (reference 2-1). Fourier transformation is an example of a transformation that can be used in conjunction with a redundancy reduction technique.

Redundancy reduction can be applied to both continuous and sampled data. In the case of sampled data, it is important to distinguish between the loss of fidelity due to sampling and the loss of fidelity due to redundancy reduction. If the continuous source data is bandlimited, then a sampling rate exists which will not result in a loss of fidelity. For a properly sampled bandlimited function, it is possible to determine a reconstruction filter which will exactly reproduce the original function from its samples. Petersen and Middleton (reference 2-2) treat the problem of selecting efficient sampling lattices and reconstruction filters for n dimensional bandlimited functions. However, it is not always possible to sample at the theoretically prescribed rate. Sampling causes the spectrum

of the function being sampled to be periodically extended in frequency space. Let the interval between samples for a one-dimensional function be T. Then the interval between spectra in frequency space is given by 1/T. Therefore, if T is less than or equal to 1/2B where B is the bandwidth of the function being sampled, then the spectra are separated and do not overlap. In such a case, a low pass filter can be used to extract the original spectrum and reconstruct the function.

When T is greater than 1/2B, the spectra overlap, and perfect reconstruction of the function is not possible. The errors introduced by reconstruction of an undersampled function are of two types: errors of omission and errors of commission. These errors, which depend upon the frequency characteristics of the function and reconstruction filter, have been analyzed in detail by Gardenhire (reference 2-3). Errors of omission are due to a loss of the portion of the original spectrum which is outside the passband of the reconstruction filter. Errors of commission are due to the portions of the repeated spectra that lie within the passband of the reconstruction filter. This latter type of error is also called aliasing because it causes some frequency components in the original function to assume a new value of frequency in the reconstructed function.

For finite sampling intervals, the spectrum of a function may vary from one sampling interval to the next. If the sampling rate is fixed, then for certain time intervals, the

function may be oversampled. Therefore, for that sampling interval the resulting sampled function is redundant in the sense that it is equivalent to another sampled function containing fewer samples. Finding the equivalent minimum sampled function that satisfies the fidelity criterion is another instance of the redundancy reduction problem described above. Adaptive sampling is one possible solution to this problem.

2.1 One-dimensional redundancy reduction techniques

Redundancy reduction techniques have been classified as being one-dimensional or two-dimensional on the basis of the functional form of the source data (reference 2-4). Visual data, which is inherently two-dimensional, can be reduced to a one-dimensional form by sampling in one or both dimensions. For example, the picture may be dissected into a series of horizontal line segments. The line segments may then be retained in continuous form or may themselves be dissected into a series of samples.

Redundancy reduction techniques can be considered in their canonical form, the form in which they are applicable to many different types of data. They can also be considered from the standpoint of a particular application. Redundancy reduction, as applied to visual data, has been called psychophysical coding (reference 2-5).

2.1.1 General techniques

In 1937, the theory of pulse coded modulation (PCM) was introduced. This communications technique is advantageous for long distance transmission because it eliminates the problems of accumulated noise and distortion associated with analog systems (reference 2-6). Development of this technique was delayed by the need for economical, high speed switching devices. In 1962, the T1 Carrier PCM System was introduced by the Bell System (reference 2-7). Since its practical introduction, PCM has become the preferred form of communication for voice, video, and data.

However, PCM has the disadvantage that it requires more bandwidth than an equivalent analog system. As a result, there has been a great deal of interest in the development of redundancy reduction techniques applicable to PCM systems. PCM is a digital communications technique. Therefore the data must be in discrete form. Each discrete datum is coded in terms of a train of equally spaced pulses. The absence or presence of a pulse in a given position constitutes one bit in a binary code word. In the Bell T1 Carrier System, the pulses are put into a "bipolar format" before being gated onto the coaxial cable. Bipolar format improves the power spectral density of the pulse train.

A simple PCM technique consists of sampling the one-dimensionalized data at regular intervals and converting the amplitude of the samples to a specified number of discrete levels. There are several parameters of this process that

can be optimized in order to minimize the number of levels required to quantize the sample amplitudes. The optimum choice of these parameters which meets the receiver fidelity criterion is a redundancy reduction problem. The parameters that can be optimized include the selection of presampling and reconstruction filters, and classification of the range of amplitudes into quantization levels. PCM using six-bit linear quantization is often used as a standard against which the efficiency of other digital redundancy reduction techniques are compared.

Delta modulation and differential pulse coded modulation (DPCM) are digital communications techniques derived from PCM. In DPCM, at each sample time, the difference between the present sample and the previous sample is classified into one of 2N + 1 quantization levels lying in the range $-n, \ldots, 0, \ldots, +N$. In Delta modulation, N=1. For these techniques, feedback must be used at the transmitter in order to avoid cumulative errors. DPCM is seriously degraded by transmission errors since all reconstructed samples following the error will themselves be in error.

Interpolators and predictors are digital redundancy reduction techniques. They produce binary code words which can be transmitted using PCM. The binary code words produced by interpolators and predictors do not represent a single sample of the source data, but a sequence of samples. Interpolators and predictors are classified according to their order. The order corresponds to the order of the polynomial being used to interpolate intervening samples from received

samples (interpolator) or to predict future samples on the basis of transmitted samples (predictor). For both techniques a polynomial is selected based on the characteristics of the data. All samples that can be interpolated or predicted from the polynomial, within a specified error tolerance, are redundant. The corresponding polynomial parameters are transmitted for each sequence of redundant samples. Predictors are of limited utility because their effectiveness is dependent upon relatively constant data characteristics. Because complexity increases with increasing order, zero-order and first-order interpolators are most often used. A first-order interpolator can have several degrees of complexity depending upon whether optimization is confined to a single interpolation line segment or several line segments. The memory requirements in the transmitter increase accordingly. (reference 2-1)

If the amplitude samples of the source data are discrete then run-length encoding can be used. This is analogous to a zero-order interpolator with a zero error tolerance.

2.1.2 Psychophysical techniques

When the general techniques described above are applied to visual data, a suitable choice of system parameters can minimize the number of bits per sample and still satisfy the fidelity requirements of the human observer. Huang et al (reference 2-8) summarize results obtained in the course of investigating the optimum choice of PCM parameters for the transmission of low-resolution monochrome still pictures.

They define low resolution as being 256 X 256 sample points per picture frame or fewer. The viewing distance used by human observers to judge fidelity was in the range four to six times picture height. The illumination at the surface of the pictures was about 30 foot-candles. The PCM parameters investigated for optimization included the pre- and post-filters and sample amplitude quantization.

They state that for a very low resolution (64 X 64 samples per frame) system and a square sampling array, ideal low pass pre- and postfilters give reconstructed pictures with the best subjective quality. However, a postfilter with high frequency accentuation improves picture quality for higher resolution (256 X 256 samples per frame) systems. With regard to amplitude quantization, if uniform linear quantization is used, five to seven bits per sample (32-128 brightness levels) are required. Otherwise, artificial contours (quantization noise) are sometimes perceptible. If logarithmic quantization is used, the data requirements can be reduced by one bit per sample. Optimum pre- and post-filters can be determined for three bits per sample quantization resulting in slightly lower quality pictures free of perceptible quantization noise. (reference 2-9)

Roberts (reference 2-10) used an interesting method to decrease the number of bits per sample and at the same time decrease the perceptibility of quantization noise. He added pseudo-random noise to the picture before it was quantized and subtracted the same noise at the receiver.

The effect was to transform correlated quantization noise into random noise with the same rms value. Using four bits per sample, he was able to obtain pictures of good quality.

Limb (reference 2-9) investigated the optimum choice of DPCM quantization levels. His visual data sample space consisted of monochrome still television pictures. The criterion used to choose optimum quantization representative and decision levels was minimum weighted mean square error. Let x_i be a decision level and let y_i and y_{i-1} be representative levels. Then the following equations can be used to determine the representative and decision levels that yield a minimum weighted mean squared error (reference 2-17).

$$x_i=(y_i+y_{i-1})/2 \qquad (2\text{-}1)$$

$$\int_{x_i}^{x_{i+1}} P(x)W(x)(y_i-x)dx=0 \qquad (2\text{-}2)$$

P(x) is the probability density function for the source data amplitude given by x. W(x) is the weighting function. Limb chose a weighting function based on threshold measurements he obtained from a simultaneous contrast experiment. The experiment measured the increment threshold $\Delta B(x)$ of a one element wide line on a uniform background positioned one element from an abrupt discontinuity in luminance. W(x) was set equal to $1/\Delta B(x)$. In addition, P(x) was measured for some set of television signals. For subjective evaluation of his system, he chose three pictures ranging from high to low in the amount of detail. For standard viewing conditions and nine representative levels, the measured entropy of the DPCM signal was between 1.98 and 0.81 bits per sample.

Gardenhire (reference 2-3) describes results obtained from the application of interpolation techniques to monochrome still pictures. On the basis of experience, he selected two pictures which represent the extreme ends of the compressibility scale. The step (zero-order interpolator) and fan (first-order interpolator) methods were applied to these pictures for different values of error tolerance. A facsimilie scanner with a resolution of 200 lines per inch was used to digitize the pictures. The sampled picture contained 1,428 X 1,317 six bit samples. For the complex picture representing the minimum compressibility end of the scale, the difference between the results obtained for the fan method and the step method was very small. A good picture was obtained for an error tolerance of 4.69 per cent (3 quantization levels). The corresponding sample and bit compression ratios were 4.09 and 2.45 resulting in 2.45 bits per sample on the average. For the picture representing the maximum compressibility end of the scale using the same error tolerance, the corresponding figures were 11.05 and 6.63 giving 0.905 bits per sample on the average. The difference between the step and fan methods was somewhat larger for the more compressible picture. The results given above are for the step method and linear quantization. When "companding" was used on the quantization levels, based on equal-perceptibility nonlinear quantization, nearly the same bit compression ratios were obtained. However, the resulting pictures were of even better quality. Results

for other pictures using interpolation would be expected to lie between the values given above.

Gardenhire's results are representative of those that can be obtained using interpolators. These results indicate that for practical error tolerances a zero order interpolator is at least as effective as a first order interpolator and is simpler in terms of equipment and processing time. Therefore, in the computer simulation described in Chapter V a zero order interpolator is employed.

In the preceding paragraphs, a description was given of the application of general redundancy reduction techniques to visual data. Now some techniques will be described that were specifically designed for visual data. These techniques are based on a variation of visual sensitivity with spatial frequency. Portions of a picture containing mostly low spatial frequencies must generally be reconstructed to a greater degree of accuracy than regions containing abrupt brightness "edges." The emphasis in these techniques is on edge detection and "dual mode" methods where the picture is separated into low and high frequency components using filtering techniques. Redundancy reduction is then applied to the two components separately.

Schreiber and Knapp (reference 2-11) experimented with a dual mode system which transmitted the low frequency component using ordinary PCM. The high frequency component was run-length encoded using 3 bits for brightness quantization and 5 bits for run-length. The bit compression ratio achieved was estimated to be 4:1. Newell and Geddes

(reference 2-12) tested a similar system using Gouriet's (reference 2-13) analog technique for run-length transmission and found that a bandwidth reduction of 3:1 was practical. They also tested the "synthetic highs technique of Schreiber et al (reference 2-14). This method is based on the detection of brightness "edges" in a scan line. The "lows" signal is separated using a low-pass filter and is transmitted separately using PCM or reduced bandwidth analog transmission. The "highs" signal is derived by differentiating the original signal and, by means of very few quantization levels, replacing the derived signal by pulses whose height and duration are determined by the portion of the signal within a given quantization level. These pulses are then transmitted in analog fashion or are run-length encoded by transmitting the distance between pulses and the discrete pulse height. At the receiving end the pulses are passed through a specially designed transversal filter which synthesizes a "highs" signal which is recombined with the "lows" to give the reproduced picture. The "synthetic highs" technique gave qualitatively better pictures than the quantized highs method with generally the same noise and bandwidth requirements.

Julesz (reference 2-15) has investigated a technique which combines a type of run-length encoding at the transmitter with linear interpolation at the receiver. The beginning of a run was based on the detection of an edge or abrupt change in brightness in a scan line. These edges

were detected by comparing the differences in brightness of adjacent pairs of unquantized brightness samples. When an edge was detected, a new run was begun and an old run was terminated, quantized, and was transmitted. Sixteen sampling intervals was the maximum allowed run-length, so 4 bits were required. Ten bits were used for brightness quantization (6 bits would have been sufficient) except when the change was within $\pm$ 2 sampling intervals of the last change, then coarser 4-bit quantization was used. At the receiver, linear interpolation was used to reconstruct the signal across a run-length between two edges. The minimum average ratio of edge points to total raster points was 31.5%, with 42.4% being the maximum and 25.3% being the minimum for one scene. Fifty-six per cent, 72% and 44% were the average, maximum, and minimum ratios, respectively, of coarsely quantized samples to selected samples. Statistics were also computed for run lengths between edges. For one of the scenes, less than 10% of the run lengths exceeded 16 Nyquist intervals and, therefore, required additional samples. These statistics were used to compute the entropy in bits per sample in order to estimate the results of statistical encoding. The average, maximum, and minimum entropy was 2.35, 2.80, and 2.01, respectively, giving an average compression ratio of 2.9 compared to 7-bit PCM. The statistical encoding in this case takes into account only the run-length statistics and not the brightness-level statistics.

Julesz found that if any of the scenes were coded with a statistical code tailored for any other scene the results were only slightly different from the results when a scene was coded with its own code.

Gabor and Hill (reference 2-16) describe a method which involves detecting edges of brightness on alternate scan lines and, using linear interpolation, derives the edges for the scan line in between.

Near the completion of this work, I became aware of the content of a particularly significant paper by Stockham (reference 2-18). The paper describes a nonlinear model for human vision which consists of a logarithmic transformation followed by a linear filter. The logarithmic transformation is based on the longstanding assumption that the eye responds logarithmically to intensity. The linear filter was derived experimentally. In the experiment, a luminance distribution was selected which would normally produce simultaneous contrast effects and Mach bands. Before being presented to the eye, this pattern was passed through three transformations in cascade. The first was logarithmic, the second was a linear filter, and the third was exponential. The complete system, including the model of the eye, is shown in Figure 2-1. A value for H' was determined which would cause the simultaneous contrast effects and Mach bands to be cancelled. It was theorized that this would occur when H' was equal to the inverse of H. The exponential and logarithmic transformations would cancel with the net effect being one of a

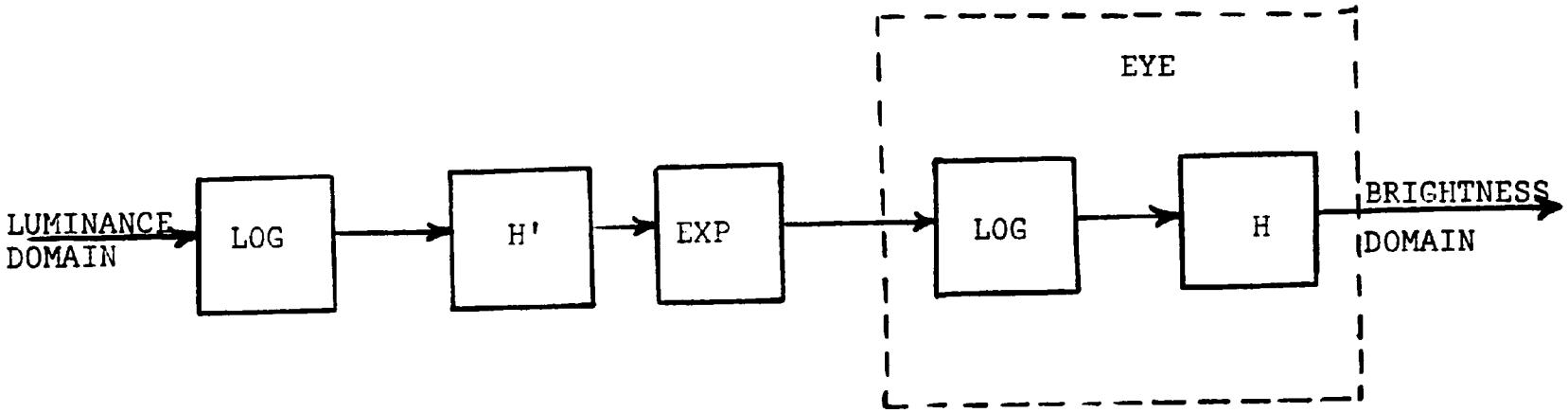


Figure 2-1 Stockham's Model of the eye and his experiment to determine H.

logarithmic transformation. Having determined a suitable value for H, a one-dimensional spatial frequency filter, Stockham investigated the effects of applying redundancy reduction techniques in the brightness domain. Roberts' method using a one bit quantizer was applied to a picture in three ways:

1) Directly in the luminance domain

2) Within a logarithmic preprocessor and an exponential post processor.

3) In the brightness domain with the model of the eye as the preprocessor and its inverse as the post-processor.

The third method was judged to give the best results.

2.2 Two-dimensional techniques

With the exception of optical transmission techniques, where visual data could be transmitted serially by frame and parallel by picture point, the data must eventually be reduced to one dimension and transmitted serially. However, redundancy reduction techniques which code areas, contours or two-dimensional transformations, operate on the data in two-dimensional form. The results are then coded for serial transmission.

2.2.1 General techniques

An area encoding method described by Cunningham (reference 2-19) involves replacing every n X n area of the picture by the average brightness in that area. The intermediate values are determined at the receiver by means

of interpolation. In addition, at the transmitter, the interpolated values are compared to the true values, and a pair of criteria are used to determine whether special correction symbols need to be sent. Averaging over a 3X3 area gave good quality pictures and a bit compression ratio of 6:1; 5X5 averaging gave poor quality pictures and a bit compression ratio of around 9:1. A modified correction scheme applied to 5X5 averaging brought the quality up and the ratio down to that of 3X3 averaging.

An encoding technique based on area-dependent statistics was investigated by Wholey (reference 2-20). He selected the 12 sample pattern shown in Figure 2-2 and and analyzed several pictures of a given class (weather maps) for the statistical relationship between the value of P and the various combinations of values for the 12 x's.

$x_1 x_2 x_3 x_4 x_5$

$x_6 x_7 x_8 x_9 x_{10}$

$x_{11} x_{12} P$

Figure 2-2 12-sample statistical encoding pattern.

Only black and white pictures were considered. These statistics were used to predict the most probable value for P, given value assignments for the 12 x's. A particular picture is processed by scanning the pattern across it and making a prediction for P at every sample point. The picture is assumed to have a white border two elements wide. For each sample point at which a wrong prediction is made, a 1 is placed in the corresponding position in an error matrix. The resulting error matrix is a complete representation of the picture, which is then transmitted using statistical run-length encoding. For ten 7000-sample weather maps the average error matrix had 5.5% 1's and the average compression ratio was 2.6:1. The entropy of the average error matrix was

$$H=[0.055 \log_2 0.055 + 0.945 \log_2 0.945] = 0.31,$$

which implies a compression ratio of 3.2:1.

2.2.2 Psychophysical techniques

Consider the following assumptions: (1) the human eye responds most to edges of brightness and, in fact, emphasizes them; (2) sharp edges of brightness are relatively few in an average picture and they lie along connected contours; (3) less accurate reproduction of the low-frequency low-contrast portions of a picture will not be noticeable. These assumptions were used by Pan (reference 2-21) and Graham (reference 2-22) as the basis for a dual mode picture transmission technique which is a two-dimensional

extension of the edge-detection method of Schreiber, Knapp, and Kay (reference 2-14). A two-dimensional low-pass filter is used to form an image resembling an out-of-focus version of the original picture which is then transmitted using ordinary 6-bit PCM and a much coarser sampling rate. In order to detect the edges of brightness the original picture is operated upon using a function which will emphasize abrupt changes in two dimensions. Either the gradient or the Laplacian operator can be used for this purpose. A nonlinear thresholding operation is performed on the resulting image to detect the presence of contours. Once a potential contour point is detected, neighboring points are examined in order to trace out the entire contour. Extremely short contours are rejected and are assumed to be the result of noise in the original picture. Each contour found is encoded and transmitted to the receiver. To reconstruct the original picture at the receiver, a two-dimensional transversal filter is applied to the image containing the thin line contours, synthesizing a "highs" signal, which, when combined with the "lows" signal, gives the desired result. Graham performed a computer simulation of this technique using a 256X256 sampled picture and 8-bit brightness quantization. A two-dimensional low-pass filter was used to determine the "lows" image, which was transmitted using 6-bit PCM and a 32X32 sampling matrix giving 3000-6000 bits. Both the gradient and Laplacian methods were used for edge detection, and transversal filters were derived for both cases. For one of the test pictures the

the gradient yielded 11,886 edge points while the Laplacian give 26,119; the same threshold was used in both cases. Results for three other test pictures using the gradient are summarized in Table 2-1. Sixteen bits were used to designate a contour starting location, 3 bits designate the direction of the gradient, 2 bits are for the gradient magnitude, and 3 bits designate the contour direction, making 24 bits total. A variable-length Huffman code is used to designate the following elements of contour continuation: contour vector direction change (VDC) which is integral multiples of 45°, $-3 \leq VDC \leq 3$; gradient direction change (GDC) which has the same values as VDC, and gradient magnitude change (GMC) which has values $0 \leq GMC \leq 4$. The end-of-contour symbol is an eighth value for VDC. The four test pictures required, on the average, 5.8 bits per contour-continuation point. A fixed-length non-statistical code would have required 9 bits. The compression ratios achieved by contour encoding (see Table 2-1) were generally higher than those which would be expected if previously described methods were applied to the same pictures. If the 256X256 sampling matrix were replaced by a 512X512 matrix, the number of samples would increase by a factor of 4, but the number of edge points would be expected only to double; therefore the compression ratios would be expected to at least double. There seem to be two disadvantages associated with this scheme. One is the amount of processing time required for each picture. The other is

TABLE 2-1

Contour Coding Data for Figures 16, 17, and 18
(Ref. 2-22)

	Fig. 16	Fig. 17	Fig. 18
Number of edge points	8900	13,980	2301
Number of contours	263	485	55
Number of bits for:			
VDC	20,300	38,400	5,900
GDC	17,400	30,600	4,260
GMC	11,200	18,900	2,880
start point	6,312	11,640	1,320
lows	3,000	3,000	3,000
Total number of bits	58,212	102,540	17,360
Bit compression ratio	6.75	3.8	22.7

sensitivity to transmission errors. Since a contour continuation message contains information only about changes in contour properties, an error in one contour continuation symbol would cause the remaining portion of that contour to be reconstructed improperly.

Cheydleur (reference 2-23) has proposed a scheme for the coding of contours and lines of varying width. It is based on the use of compactly coded syllables which specify such things as contour location and contour alteration. A contour location syllable specifies the beginning of a new contour. It is followed by sub-syllables which specify its horizontal displacement relative to another contour, the slope of the left contour boundary, the rate of change of the slope, the width of the contour, and the rate of change of the width. There is also a sub-syllable which specifies the vertical displacement to the position where there is a change in any of the previously mentioned sub-syllables. The values of the changes are specified by the sub-syllables of a contour alteration syllable. Pictures with grayness levels are coded by replacing the width sub-syllable by one that specifies grayness value to the left of the contour.

Previous methods which have been described generally involve operations which attempt to extract and encode the essential information by operating on the original picture. It would be desirable to find a suitable transformation which, when applied to the original picture, yields an image containing the essential information in a more convenient form.

A suitable transformation would preserve the information content in the original picture and would have an inverse. Andrews (reference 2-24) showed that the discrete two-dimensional Fourier transformation is an information preserving mapping. The entropy of a function in the space domain is equal to its entropy in the Fourier domain. Andrews and Pratt have investigated the effects of transformation techniques on images (references 2-25, 2-26, 2-27, and 2-28).

Andrews (reference 2-27) investigated the effects of the Fourier transform on images. He found that good quality reproduced images could be obtained if the Fourier transform and its inverse were performed on a digital computer. Sources of error, such as truncation, a limited number of Fourier coefficients, and a square, binary observation window had an insignificant effect on the results. The number of sample points in the Fourier domain is equal to the number of sample points in the spatial domain. However, each Fourier coefficient has a real and an imaginary part. Andrews showed that for a non-negative real function the Fourier domain is conjugate symmetric. Therefore, the number of samples that need be transmitted in the Fourier domain is equal to the number that need be transmitted in the spatial domain. But, it was found that if the Fourier coefficients were linearly quantized, using six bits, transmitted and the inverse transform taken at the receiver, the results were unsatisfactory. In order to quantize the Fourier

coefficients so that the quantization error is minimized in a mean square sense, or so that the quantization levels are equally probable, the probability density function of each Fourier coefficient must be known. In general, it would be expected that the probability density function would be dependent on the position of the Fourier coefficient in the Fourier domain. Andrews investigated the relationship between a stochastic process in the spatial domain and its Fourier domain counterpart. He did not obtain an exact solution to the problem, but instead assumed that the real and imaginary part have a normal density function. Now, the statistical analysis indicated that the variance of the Fourier coefficients is a sampled periodic function of the power spectral density of the continuous spatial stocastic process. Experiments on a test scene indicated that a two-dimensional Gaussian function was a suitable approximation for the power spectral density. Therefore, it was assumed that each Fourier coefficient has a normal density function and a variance which is a two-dimensional Gaussian function of spatial frequency. For these assumptions, it can be concluded that the magnitude of a Fourier coefficient is Rayleigh distributed and its phase is uniformly distributed. Using these probability functions, and six bit quantization, levels were determined for each Fourier coefficient such that all quantization levels were equally probable. The resulting transformed, quantized, and inverse transformed pictures were of good quality for both Gaussian real and

imaginary quantization, and Rayleigh magnitude and linear phase quantization. It was found that the quality of the results was relatively immune to the functional form of the variance distribution.

So far it has been shown that it is possible to obtain good quality images by transmitting the quantized Fourier coefficients using the same bandwidth as would be required for six bit PCM transmission of the sampled spatial domain function. This same problem was studied by Pratt and Andrews (reference 2-29) for other transforms, including the Hadamard transformation. They also investigated the use of redundancy reduction techniques in the Fourier domain. They found that satisfactory results could not be obtained if spatial domain techniques were applied in the Fourier domain. Other techniques which were also found to be unsatisfactory were checkerboard sampling, random sampling, random sampling of high frequencies, and zonal low-pass filtering. However, if moderate resolution loss is acceptable, then a low pass zonal filter which transmits 99.99% of the image energy can be used to obtain a bit compression ratio of 4:1. This is true only for the Fourier transform. The Hadamard transform is more susceptible to image degradation when zonal filtering is used. The most successful technique was one in which all Fourier coefficients whose magnitude exceeded a constant threshold were transmitted. However, such a scheme requires that additional data be transmitted which describes the positions of the Fourier coefficients

which are above threshold. For a Fourier domain threshold which resulted in satisfactory pictures, a sample compression ratio of 6.2 was possible. Run-length statistics for this case indicated that four bits would be satisfactory for a fixed length position code. This would result in an approximate bit compression ratio of 3.72, assuming ten bits per Fourier component.

Anderson and Huang (reference 2-29) used a more sophisticated threshold technique. Their method consisted of dividing the picture into square subsections. The dimension of a subsection was based on the period of the spatial frequency at which the contrast sensitivity of the eye, for sinusoidal gratings, is a maximum. Each subsection was logarithmically transformed, and then Fourier transformed. and the number of Fourier components above threshold was defined to be 1.85 times the standard deviation for that subsection. For 16 X 16 subsections no more than 150 Fourier coefficients were ever transmitted. The magnitude and phase of the coefficients above threshold were linearly quantized using 2, 3, 4, or 5 bits depending upon the value of standard deviation for that subsection. The maximum Fourier coefficient magnitude in a given subsection determined that largest quantization level. Ten bits were used for its transmission. The zero frequency term in each subsection was transmitted using seven bits. They found that the run-length statistics for coefficient position were different for different rows of the Fourier transform.

Therefore, the code words were selected accordingly. The results they obtained are summarized in Table 2-2. Since the pictures they used are very similar to the ones used in this research, it will be interesting to compare the results.

Other techniques not mentioned here, but studied in the preparation of this chapter, are referenced in the Bibliography, if they are not specifically referenced below.

TABLE 2-2

Bandwidth Compression Data
(Ref. 2-29)

Picture	Total Number of Fourier Coefficients	Run-Length Data (bits)	Total Bandwidth (bits)	Bits per Picture Point
1	15 240	16 298	82 899	1.265
2	10 880	13 003	60 839	0.928
3	21 718	20 016	103 770	1.583

REFERENCES

2-1 C.M. Kortman, "Redundancy Reduction -- A Practical Method of Data Compression," Proc. IEEE, vol. 55, no. 3, Mar. 1967, pp. 253-263.

2-2 D.P. Peterson and D. Middleton, "Sampling and Reconstruction of Wave-Number-Limited Functions in n-Dimensional Euclidean Spaces," Information and Control, vol. 5, 1962, pp. 279-323.

2-3 L.W. Gardenhire, Sampling and Source Encoding, Radiation Incorporated Systems Division, Melbourne, Florida, April, 1970.

2-4 L.C. Wilkins, P.A. Wintz, "Bibliography on Data Compression, Picture Properties, and Picture Coding,: IEEE Transactions on Information Theory, vol. IT17, no. 2, pp. 180-197.

2-5 W.F. Schreiber, "Picture Coding," Proc. IEEE, vol. 55, no. 3, March 1967, pp. 320-339.

2-6 E.M. Deloraine, A.H. Reeves, "The 25th Anniversary of Pulse Code Modulation," IEEE Spectrum, May 1965, vol. 2, no. 5, pp. 56-63.

2-7 V. Krueger, "The T-Carrier PCM System," Telecommunications, vol. 6, no. 12, December 1972, pp. 31-35.

2-8 T.S. Huang, O.J. Tretiak, B. Prasada, Y. Yamaguchi, "Design Considerations in PCM Transmission of Low Resolution Monochrome Still Pictures," Proc. IEEE, vol. 55, no. 3, March 1967, pp. 331-335.

2-9 J.O. Limb, "Source-Receiver Encoding of Television Signals," Proc. of the IEEE, vol. 55, no. 3, March 1967, pp. 145-154.

2-10 L.G. Roberts,"Picture Coding Using Pseudo-Random Noise," IRE Trans. on Information Theory, vol. IT-8, Feb, 1962, pp. 145-154.

2-11 W.F. Schreiber, C.F. Knapp, "TV Bandwidth Reduction by Digital Coding," IRE Conv. Record, 1958, pp. 88-99.

2-12 G.G. Newell and W.K.E. Geddes, "Tests of Three Systems of Bandwidth Compression of Television Signals," Proc. IEE (London), vol. 109, pt. B, pp. 311-323.

2-13 G.G. Gouriet, "Bandwidth Compression of a Television Signal," Proc. IEE, Paper No. 2357R, May 1957, 104B, p. 265.

2-14 W.F. Schreiber, C.F. Knapp, and N.D. Kay, "Synthetic Highs, an Experimental TV Bandwidth Reduction System," J. Soc. Motion Picture and TV Engrs., vol 68, August 1959, pp. 525-537.

2-15 B. Julesz, "A Method of Coding Television Signals Based On Edge Detection," Bell Sys. Tech. J., vol. 38, July 1959.

2-16 D. Gabor and P.C.J. Hill, "Television Band Compression by Contour Interpolation," Proc. IEE (London), vol. 108, pt. B, no. 39, 1961, p. 303.

2-17 J. Max, "Quantizing for Minimum Distortion," IRE Trans. on Information Theory, vol. IT-6, March 1960, pp. 7-12.

2-18 T.G. Stockham, Jr., "Natural Image Information Compression with a Quantitative Error Model," Conference on Pertinent Concepts in Computer Graphics, Urbana, Illinois, April, 1969.

2-19 J.E. Cunningham, "Picture Processing," Quarterly Progress Report No. 54, MIT Research Lab. for Electronics, 1959, p. 138.

2-20 J.S. Wholey, "The Coding of Pictorial Data," IRE Trans. on Information Theory, vol. IT-7, April 1961, pp. 99-104.

2-21 J. Pan, "Reduction of Information Redundancy in Pictures," ScD. Dissertation, Dept. of Elec. Engrg., MIT, Cambridge, Mass., September 1962.

2-22 D.N. Graham, "Image Transmission by Two-Dimensional Contour Coding," Proc. IEEE, vol. 55, no. 3, March 1967, pp. 336-346.

2-23 B.F. Cheydleur, "Compressed Code," Private Correspondence.

2-24 H.C. Andrews, "Entropy Considerations in the Frequency Domain," Proceedings of the IEEE, vol. 56, no.1, January 1968, pp. 113-114.

2-25 H.C. Andrews, W.K. Pratt, "Television Bandwidth Reduction by Fourier Image Coding," Preprint No. 103-91, 103rd. Technical Conference, Society of Motion Picture and TV Engineers, May 1968.

2-26 H. Andrews, "A High-Speed Algorithm for the Computer Generation of Fourier Transforms," IEEE Trans. on Computers, vol. C-17, no. 4, April 1968, pp. 373-375.

2-27 H.C. Andrews, Fourier Coding of Images, University of Southern California, Electronic Sciences Laboratory, June 1968, USCEE Report 271.

2-28 W.K. Pratt, H.C. Andrews, Transform Processing and Coding of Images, University of Southern California, Electronic Sciences Laboratory, USCEE Report 341, March 1969.

2-29 G.B. Anderson, T.S. Huang, "Piecewise Fourier Transformation for Picture Bandwidth Compression," IEEE Trans. on Comm. Tech., vol. COM-19, vo. 2, April 1971, pp. 133-140.

CHAPTER III

CHARACTERISTICS OF HUMAN VISION

3.0 Introduction

The structure of the human eye will be examined in detail in order to determine the significance of each component of the eye in the picture-viewing process. Particular attention is given to the optical components, the photoreceptors, and the composite visual pathway from retina to brain. Characteristics of these components based on experimental measurements as well as theoretical models are given in the following sections.

3.1 The structure of the human eye

The vertebrate eye is one of the many types of "eyes" that are found in nature. Eye-like mechanisms range in complexity from single photo-sensitive surface cells, or eye spots, through the compound eye of the insect to our own eyes with their ability to focus, adapt, and move.

The fundamental, functional architecture of the human eye consists of the following components (reference 3-1):

1) the eyeball which contains the optics for image formation and sensory mechanisms for image detection and enhancement,

2) the two principal refracting elements of the optics in the eyeball are the cornea and the lens. These act together to produce an inverted, reduced image on the retina. The power and the aperture (pupil) of the lens are variable,

3) the retina is the sensory mechanism of the eye. It consists of several functional layers parallel to its surface. The photoreceptors form the layer nearest the outside diameter of the eyeball. Moving toward the inside of the eyeball, several interconnected layers of nerve cells are encountered. The innermost layer consists of the individual fibers of the optic nerve which converge into a single bundle and exit at the rear of the eyeball,

4) the two optic nerve bundles from each eyeball combine at the optic chiasma and emerge as two optic tracts connected to the two sides of the brain. The optic tract connected to a particular side of the brain contains the optic nerve fibers for the visual half-field on that same side of the two eyeballs, as shown in figure 3-1.

Functionally the visual system can be assumed to consist of two subsystems:

The optical system, which receives and focuses the light incident on the eye, and

The visual information processing system, which converts the light energy into the form of neural signals and transmits these signals to the brain.

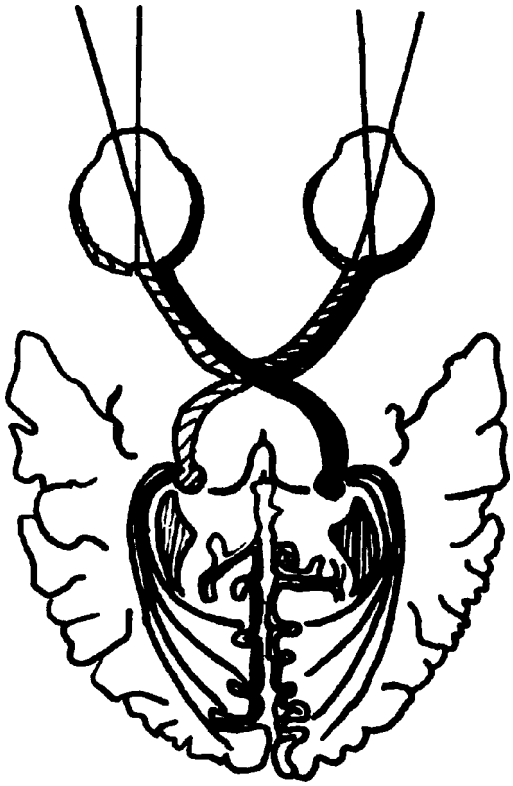


Figure 3-1. The visual pathway. (Reference 3-1)

3.1.1 Optical system components

Light entering the eye passes through an optical system consisting of the cornea, the aqueous humor, the lens, and the vitreous humor. The cornea is the major component in the optical system of the eye. Of the total dioptric power of the human eyeball nearly three-quarters is contributed by the interface between the cornea and the air. (Reference 3-2) The cornea is divided into layers lying parallel to its surface. A layer called the stroma makes up about nine-tenths of the thickness of the cornea. Aqueous humour is an intra-ocular fluid which circulates in the chamber between the lens and the cornea. (Reference 3-3) The lens is an important part of the refractive media of the eye, because it can alter its shape and therefore change the optical properties of the system. This characteristic is known as accommodation. (References 3-4, 3-7) The iris is the F-stop of this optical system. This diaphragm controls the aperture of the refracting system of the eye. (Reference 3-5) The vitreous body is the clear jelly-like structure which fills the space between retina and lens. It is described as being "well-adapted to the purpose of retaining the retina in position and preventing free diffusion of large molecules or cells, while allowing small nutrient molecules to diffuse within it." (Reference 3-6)

3.1.2 Components of the retina and visual pathway

Light entering the eye is focused by the optical system on the retina. The retina is organized into several major layers running parallel to the surface of the retina. The outermost layer, the one nearest the outside of the eye, is occupied by the photoreceptors or rods and cones. Therefore, light entering the eye passes through the optics into the vitreous body, across the eyeball toward the inner layers of the retina, and across the retinal layers into the layer of photoreceptors. The innermost layer contains the fibers of the optic nerve. The layers in between consist of specialized types of cells: horizontal cells which transmit the impulses from the receptors to the surrounding rods and cones; bipolar cells which transmit the impulses from the photoreceptors to the ganglion cells; bipolar cells which receive the impulses from ganglion cells, and also from the brain, and transmit them back to the photoreceptors; amacrine cells which intercept a part of the bipolar impulses and spread them over the surrounding territory; and ganglion cells which receive impulses from the bipolars and transmit them to the brain along the optic nerve fibers (reference 3-8). The optic nerve fibers converge and leave the eye forming the optic nerve. The arrangement of the fibers within the optic nerve is a very orderly one, retaining essentially the retinal image pattern. The two optic nerves combine at the optic chiasma. The arrangement of the optic nerve fibers changes profoundly

within the optic chiasma. The two optic nerve fiber bundles leaving the optic chiasma are called the optic tracts. The left optic tract contains the optic nerve fibers from the temporal portion of the left retina and the optic nerve fibers from the nasal portion of the right retina and vice versa for the right optic tract. From there on, each visual pathway carries only one-half of the retinal image pattern. In this new pattern, however, the two corresponding half-images are combined into one and carried in one compact fiber-tract. The fibers of the optic tract terminate in two distinctly separate areas of the brain. One set of fibers terminates in the pretectal nucleus and is assumed to control the light reflex of the pupil. The other set of fibers is considered to be visual fibers and terminates in the lateral-geniculate body. Of the six layers in the human lateral-geniculate body, layers one, four, and six accept the crossed fibers from the nasal retinal half of the opposite eye while layers two, three, and five serve the uncrossed fibers from the temporal retinal half of the ipsilateral eye. Even at this end of the optic tract, a point-to-point projection, that is the retinal image pattern, is maintained to a very large extent. The optic fibers then leave the lateral-geniculate body as a solid tract and spread out into a big fan or lamella which terminates in the visual cortical area (reference 3-1).

3.2 Characteristics of the optical system

The characteristics of the optical system of the human eye will now be investigated in detail. This investigation will serve several purposes: one purpose will be to set a theoretical bound on the spatial frequency bandwidth of the eye, another is to determine the equivalent magnification power of the optical system. Also this investigation will yield a mathematical model for the optical system which will be useful in incorporating the eye into the picture bandwidth compression problem.

3.2.1 A simplified schematic model of the optical system

Light entering the eye passes first from the air through the anterior refractive surface of the cornea, through the cornea and then through the posterior refractive surface of the cornea into the aqueous humour. The light then encounters the anterior refractive surface of the lens, goes through the lens and passes through the posterior refractive surface of the lens into the vitreous humour where it is then focused on the retina.

In order to understand how the lens system of the human eye might be analyzed, it is necessary to introduce some concepts from optics. Consider a single spherical refracting surface. The surface separates a medium with index of refraction n_1 from a medium whose index of refraction is n_2 . From geometrical optics, the object point O is related to the image point I by equation

$$n_2 / l_2 - n_1 / l_1 = (n_2 - n_1) / R \quad (3\text{-}1)$$

where R is the radius of curvature of the spherical surface and l_1 and l_2 are distances from O and I to the refracting surface. The right side of the equation is called the surface power and has the symbol F. The two terms on the left by definition are called the reduced vergence and are given the symbols L_1 and L_2. Therefore the equation becomes $L_2 = L_1 + F$. Vergence is a measure of the convergence (positive) or divergence (negative) of light rays. For example, light rays traveling parallel to one another have a vergence of zero. When light rays pass through a refracting surface, their degree of vergence is changed by an amount equal to the surface power. The equation $L_2 = L_1 + F$ also holds for a thin spherical lens. In this case, the surface power is the sum of the surface powers of the two refracting surfaces. Vergence is a function of the distance from a point of focus. When light rays pass through a refracting surface, the point of focus changes, becoming either virtual or real. Now as the light rays continue toward the next refracting surface, the vergence changes because the distance to the point of focus changes. In the case of a thick lens, for example, this change of vergence must be taken into account. The equation that relates the object point O to the image point I in the case of two refracting surfaces separated by a distance t is $L_3 = L_1 + F_1 + L_x + F_2$ where F_1 and F_2 are the surface powers of the two refracting surfaces

and L_x is the change in vergence between the two refracting surfaces. L_x is also known as effectivity and is given by equation 3-2

$$L_x = L_2 \; / \; (1 - (t/n_2)\, L_2) \qquad (3\text{-}2)$$

where $L_2 = L_1 + F_1$. $F_1 + L_x + F_2$ can be combined into F to give the equivalent power of the lens system, giving an equation similar to the one for a single refracting surface.

In a similar way, any number of refracting surfaces can be analyzed in terms of a single equivalent refracting surface or thin lens. One important property of a lens system is its principal foci. These are the two points of focus for parallel rays of light entering the lens system from the right and from the left. For a single refracting surface and parallel rays entering from the left, $L_1 = 0$, therefore $L_2 = F$ so the distance to one of the principal foci is given by $l_2 = n_2R \; / \; (n_2 - n_1)$. l_2 is the distance from the refracting surface to the second principal focus. The first principal focus can be found in a similar way. For a complex lens system, represented in terms of an equivalent thin lens, principal foci can also be determined. In this case the principal foci determine the position where the equivalent lens should be located.

In general the position of the equivalent thin lens is different for parallel rays entering from the right and from the left. These two position points are known as the principal points of the equivalent lens system. For a lens

system in which the first and last refractive indices are the same, the principal points have the additional property of being the axial points at which an object and its image subtend equal angles. However if these refractive indices differ, as they do in the case of the eye, a different pair of points have this property. They are known as the nodal points. One last property of a lens system needs to be defined - magnification. This is usually given by the ratio of image height to object height. However magnification is given by L_1 / L_2 . Naturally, all the above results only apply to objects placed at or near the axis of the lens system.

Bennett and Frances (reference 3-9) describe a model proposed by Gullstrand which has been used to analyze the optics of the eye. The model consists of three refracting surfaces; a single surface of curvature chosen to represent the effect of the whole cornea, and the two surfaces of a lens assumed to be homogeneous and with suitable curvatures so that its equivalent power is close to that of the average crystalline lens. The radii of curvature of the cornea and the lens surfaces have been measured, and the refractive indices of the cornea, aqueous, lens, and vitreous are known.

By incorporating these values into the model, it is possible to determine the surface powers of the various refracting surfaces and the equivalent power of the entire optical system. The principal points and the nodal points of the system can also be determined. Figure 3-2 and 3-3

below show equivalent systems for the unaccommodated and accommodated eye. For the unaccommodated eye, this model gives an equivalent optical system power of 59.6 dioptres. Using these values, it is possible to compute the size and position of images.

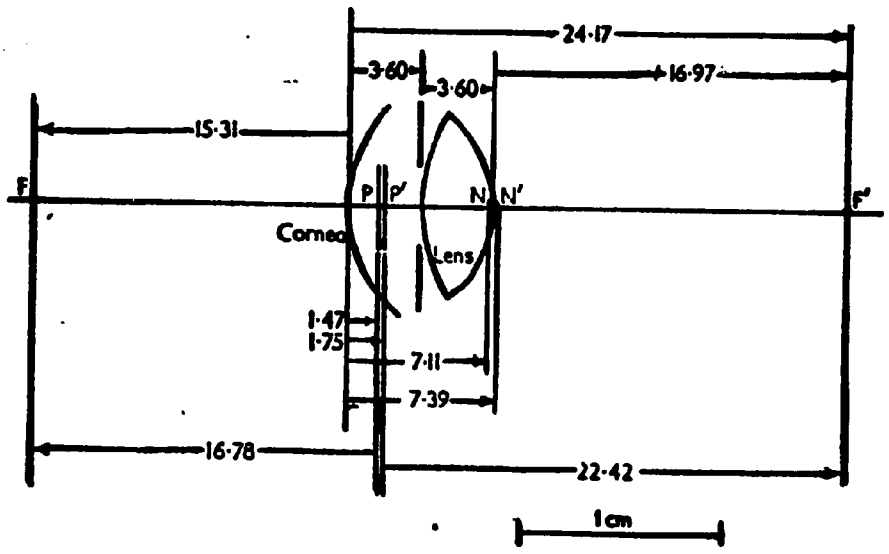


Figure 3-2. The Gullstrand simplified schematic eye: unaccommodated. F=anterior principal focus. F'=posterior focus. P and P'=first and second principal points. N and N'=first and second nodal points. (All distances in mm). (reference 3-9)

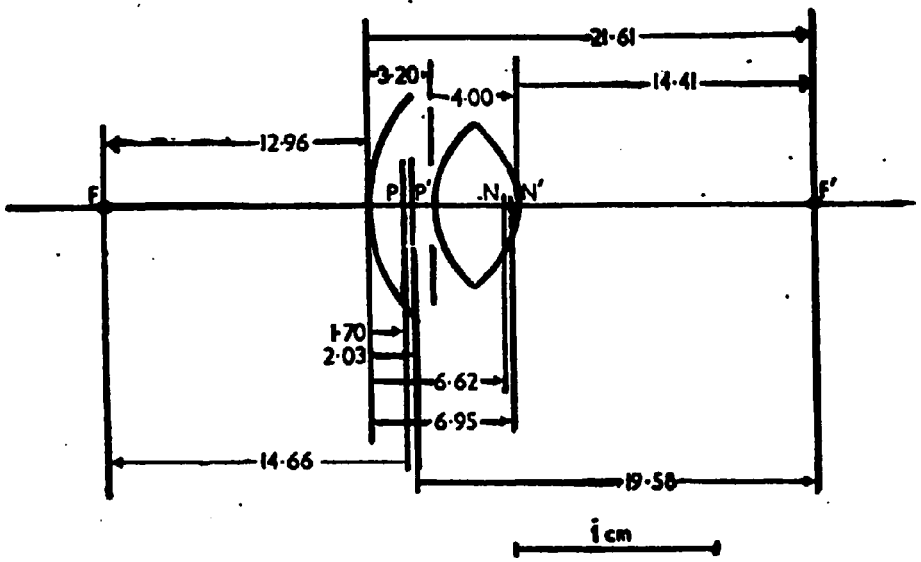


Figure 3-3. The Gullstrand simplified schematic eye: accommodated 8.62 D. (reference 3-9)

3.2.2 Accommodation of the lens

When the eye fixates on an object at a certain distance, the lens of the eye accommodates to a shape appropriate to focusing the image of the object on the retina. Accommodation involves a change in shape of the refractive surfaces of the lens. When the eye accommodates to focus a near object, the lens power increases, the change being accomplished mainly by an increase in curvature of the front surface, a small increase in curvature

of the back surface, and a small forward movement of the front vertex due to the increase in thickness. An example given in reference 3-9 shows that when the eye is focused for a point 11.6 centimeters from the first principal point, the anterior lens surface becomes more steeply curved, its radius decreasing from ten millimeters to five millimeters. This change is accompanied by an increase in axial thickness from 3.6 to 4.0 millimeters and a small increase in the curvature of the posterior surface. Thus the lens alters its equivalent power from 20.28 Dioptres to 30.13 Dioptres. The result is to change the equivalent power of the eye as a whole from 59.6 Dioptres to 68.22 Dioptres.

3.2.3 Effects of illumination on the iris

Within the optical system of the eye, the pupil serves three main functions. First, it regulates the amount of light entering the eye. Second, it increases the depth of focus of the eye by decreasing the aperture of the optical system. Third, it reduces chromatic and spherical aberrations, especially in bright light when the pupil is small (reference 3-5). In response to changing conditions, the pupillary diameter may vary from about 1.3 to almost ten millimeters. The graph in figure 3-4, taken from Walsh (reference 3-10), shows the variation in pupil diameter as a function of field luminance.

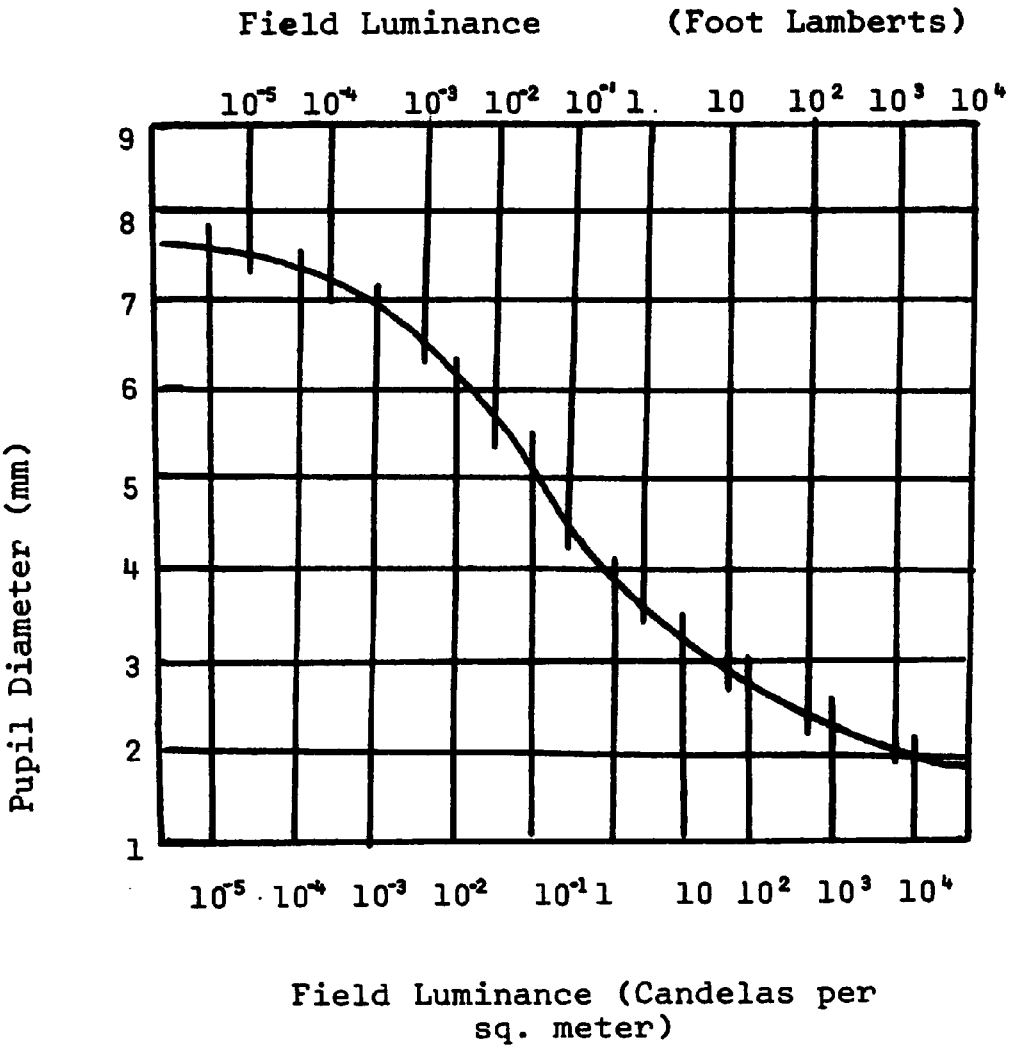


Figure 3-4. The Relation Between Pupillary Diameter and Field Luminance. (reference 3-10)

3.2.4 Pupil size and cut-off frequency

An estimate of the spatial frequency characteristics of the optical system of the eye can be derived by means of some results from linear optics. The system transfer function of a coherently illuminated system is equal to P (-u,-v) where P (-u,-v) has the value 1.0 in a region given by the smallest aperture of the system projected onto the exit pupil, and is zero elsewhere. The transfer function for an incoherently illuminated system is not given by this same two-valued function, but will always have twice the cut-off frequency of a coherently illuminated system (reference 3-11). In the optical system of the eye, the smallest aperture of the system is the pupil. By means of the model described previously, it is possible to project the pupil onto the exit pupil of the system. Therefore, for a given pupil size, the cut-off frequency of the optical system can be determined. Since the pupil diameter varies as a function of illumination, the cut-off frequency will also be dependent on illumination. For a given pupil size and varying object distance, the refractive power of the lens will vary and therefore the image of the pupil on the exit pupil will vary in size. This, in turn, will vary the cut-off frequency of the system.

Bennett and Francis (reference 3-9) calculated the position and magnification of the exit pupil and found it to be 0.08 mm behind the real pupil and 1.03 times as

large for the unaccommodated eye. If the exit pupil to retina distance l = 20.49 mm and the exit pupil diameter d = 3 mm (good interior lighting) and the illumination is incoherent, of wavelength $\lambda = 0.5 \times 10^{-3}$mm, then the diffraction limited, abberration free cut-off frequency is given by l / 2d = 293 cycles / mm or 85.2 cycles/degree.

3.2.5 Optical quality of the human eye

Westheimer and Campbell used an experimental procedure to measure the optical quality of the living human eye. The subject's natural pupil was chemically dilated: and a thin filament of light was focused through the eye's optics onto the retina. The filament then reflected off the retina and back through the eye's optics, where it was separated from the incoming beam and scanned by a photocell. An artificial pupil was used to limit the aperture. An auxillary lens was also used to achieve a best focus of the image on the retina. If it is assumed that the optical system of the eye is reversible and that the reflection at the retina's surface is diffuse, then the distribution of illumination measured by the photocell is also an indication of the linespread function of the system. If the system is assumed to be linear, then the Fourier transform of the light distribution at the photocell is equal to the square of the Fourier transform of the system because of the double-pass through the eye's

optics. The inverse transform of the system transform is the true linespread function of the system; that is, the light distribution of the image of the filament focused on the retina.

Westheimer and Campbell (reference 3-12) showed data for both a three millimeter and a six millimeter artificial pupil. They found that the equation ($f(x) = \varepsilon^{-0.7|x|}$) gave a good approximation of the linespread function for the three millimeter pupil, where f(x) is the illuminance at a distance x, in minutes of arc, from the center of the image. This equation could be used to compute the one-dimensional Fourier transform for the system which in turn is a measure of the response of the system to one dimensional spatial sine waves of varying frequencies.

Krauskopf (reference 3-13) performed some experiments similar to those of Westheimer and Campbell. However, in his apparatus, he took greater care to eliminate the effects of stray light and electronic noise. Drugs were used to paralyze accommodation and dilate the pupil. The optics of the eye were again assumed to form a reversible optical system. The target used was a bright vertical bar 1.6 minutes of arc wide and 2 degrees, 35 minutes high. Krauskopf used Fourier analysis to find the Fourier coefficients of the target. These coefficients were then divided by the coefficients of the final image for corresponding sine and cosine terms, the result being the Fourier coefficients of the effective optical system.

Since the final image is the result of a double passage through the optical system of the eye, the square root of the derived Fourier coefficients gives the desired result. The Fourier coefficients of the system are a measure of the relative response of the system to spatial sine waves for different frequencies. For pupil diameters of three, four, five, six, seven, and eight millimeters, Krauskopf computed the system response function.

The system bandwidth was the greatest for three, four, and five millimeter pupil diameters and gave a measured cut-off frequency of 0.50 lines per minute of arc or 30 cycles per degree. For each pupil diameter the retinal image of the bright vertical line was computed by multiplying the corresponding Fourier coefficients of the target and the system. These images are also a very good approximation of the linespread function of the system. Pupil diameters of three, four, and five millimeters gave the narrowest linespread functions.

A different technique has also been used to measure the optical quality of the eye. This psychophysical method is based on a comparison of the threshold contrast sensitivity of an interference grating produced directly on the retina versus the threshold contrast sensitivity of a grating degraded by and passing through the optics of the eye. A coherent source of light, a laser, is used to produce an interference pattern directly on the retina. For a given interference line spacing, the contrast of

the interference pattern is adjusted until that level of contrast is reached where the subject's perception of the interference pattern is at threshold. Then the interference line spacing is adjusted to a new value and the threshold contrast is determined for that spacing and so on.

In the next phase of the experiment, a grating pattern is produced on the face of a cathode ray tube. With the observer viewing the cathode ray tube, and for a fixed grating line spacing, the contrast is adjusted until threshold is determined. This procedure is repeated for several different grating line spacings. The resulting contrast threshold curves as a function of grating line spacing, or spatial frequency, are a measure of the relative response of the system to spatial sinusoids for different frequencies, with the optics of the eye both included and not included in the system. If the overall-system and retina-brain-only measurements were made at identical wave lengths and luminances, then the ratio of the two threshold contrast sensitivity functions can be interpreted as being due to defects in the refractive system of the eye's optics.

The psychophysical technique was used by Campbell and Green (reference 3-14). They used an oscilloscope with a green phosphor and a spectral luminance peaking at 530 millimicrons. The beam half-width was $\pm$ 30 millimicrons. The laser used emitted at 632.8 millimicrons. Therefore, identical wave-lengths were not used. Additional experiments

indicated that the use of widely differing wave lengths did not affect the results. The retinal illumination produced by the laser was 500 trolands. The luminance of the oscilloscope was 100 candelas per square meter. For a 2mm pupil, the retinal illumination was 314 trolands, a significant difference. The contrast transfer functions obtained for different pupil sizes were compared with the results which would be expected if the eye were an ideal diffraction limited optical system. The results indicate that for small pupil diameters the diffraction limit is approached.

Krauskopf's results from the double pass method for pupil diameters of three, four, five, and six millimeters, indicate that at the 50% relative response point, the corresponding spatial frequency is about six cycles per degree. However, Campbell and Green found, using the psychophysical method, that for pupil diameters of 2, 2.8, 3.8, and 5.8 millimeters the spatial frequencies corresponding to the 50% contrast ratio point were approximately sixteen, thirteen, eight, and five cycles per degree respectively.

Without considering the differences in wavelength and luminance levels used, the results given by the psychophysical method seem to indicate that the optical system is of higher quality than indicated by the double pass method. It is claimed by Campbell and Green that

> "the Ophthalamoscopic technique requires considerable skill in the alignment of the optical apparatus, in the adjustment of the electronics, and in the selection and training of the subjects. Furthermore, the calculations required to obtain the single pass optics from the double pass data depend on a number of untested assumptions about the nature of the reflecting (or diffusing) surface at the fundus. These difficulties and uncertainties could lead to the conclusion that the quality of the image is worse than it is in reality."

Campbell and Gubisch (reference 3-15) were concerned about the apparent difference in results obtained using the double pass and the psychophysical methods:

> "Several factors can possibly account for the different results between these physical and psychophysical methods of estimating optical quality. First, the fundus may act optically as a poor screen and thus degrade the line's image before casting it by reflection out of the eye. However, the image available to the receptors may not have been appreciably blurred in this manner. Secondly, the fundus may behave as a smooth mirror. If this is so, Fourier analysis, which in the objective experiment depends for its validity upon the presence of a perfect diffuser rather than a mirror would yield an estimate of poorer optical quality than that which actually exists. Thirdly, light scatter in the optical media must affect all physical measurements of image degradation but will not show up in the psychophysical technique because scatter occurs in both coherent and incoherent beams and when the difference between two corresponding psychophysical measurements is taken the effect of scatter tends to be cancelled."

They investigated the nature of the fundal (retinal) reflection and concluded that "the human fundus in white light may be regarded as a nearly perfectly diffuse reflector."

Campbell and Gubisch also used the double pass method to compute optical quality. The optical apparatus they used was simple and could be more easily aligned accurately. They used a special purpose computer to average out noise in the image-measuring process. The results they obtained were closer to those obtained using the psychophysical method. They tried to account for the remaining difference between the results obtained using the two methods. Chromatic aberration, which is not present in the psychophysical method due to the use of monochromatic light sources,seemed to account for about one-half the difference between the two methods. This was shown by treating the image of the polychromatic target as the superposition of several monochromatic images. Due to spherical aberration, not all of these monochromatic images are in focus. Therefore, the effective system transfer function for a particular monochromatic image can be approximated by the ideal modulation transfer function defocused by an amount particular to the given wave-length. The equivalent modulation transfer function is then computed as a weighted sum of the defocused modulation transfer functions for each monochromatic image.

In another paper, Gubisch (reference 3-16) summarizes the results that he and Campbell obtained using the double pass method. These modulation transfer functions were obtained using white light and are the average of three subjects for artificial pupil diameters of 2.4, 3.8, and 6.6 millimeters. These curves are shown in figure 3-5.

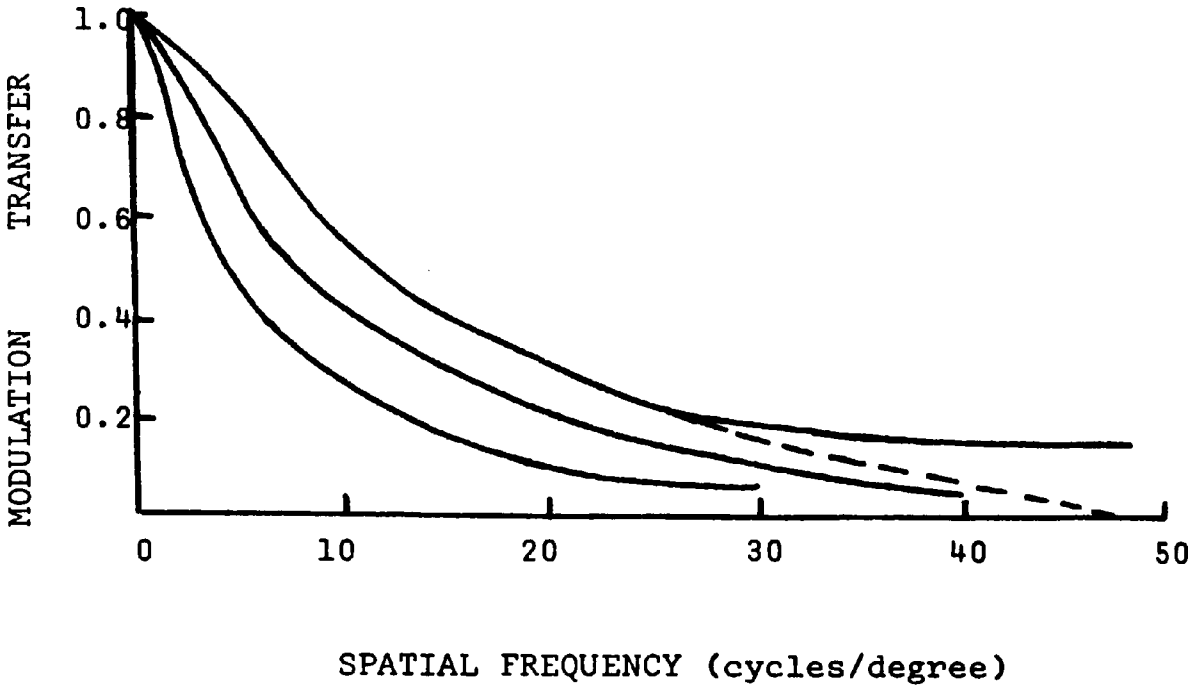


Figure 3-5. Modulation transfer functions of the optics of the human eye in white light, with 2.4 mm, 3.8 mm, and 6.6 mm artificial pupils. The curves lie in order of pupil diameter, with the highest curve corresponding to the smallest pupil. Dashed line gives MTF for 1.5 mm pupil. (reference 3-16).

Results obtained by Campbell and Green using the psychophysical method are shown in figure 3-6.

At this point it must be decided which of the two methods of measuring optical quality yield the most accurate results relative to the picture-viewing environment. This can best be judged by considering the differences between the two methods.

Results obtained using the double pass method are dependent on the nature of the fundal reflection, the effects of scatter within the non-optical elements of the eye such as the aqueous and the vitreous and the corneal surface, and on the effects of chromatic and spherical aberrations.

The psychophysical method on the other hand is only dependent on spherical aberrations. Therefore it would seem that the psychophysical method gives the most accurate measure of optical quality of the eye. However, because white light is a more typical source of illumination for a picture-viewing environment chromatic aberration as well as spherical aberration will be present in the image degrading process. Since the fundal reflection has been shown to be nearly perfectly diffuse, the only remaining source of error in the double pass method would be the effects of light scatter.

The results given by Gubisch were obtained using white light and therefore include the effects of chromatic aberration. Chromatic aberration was shown to account for most

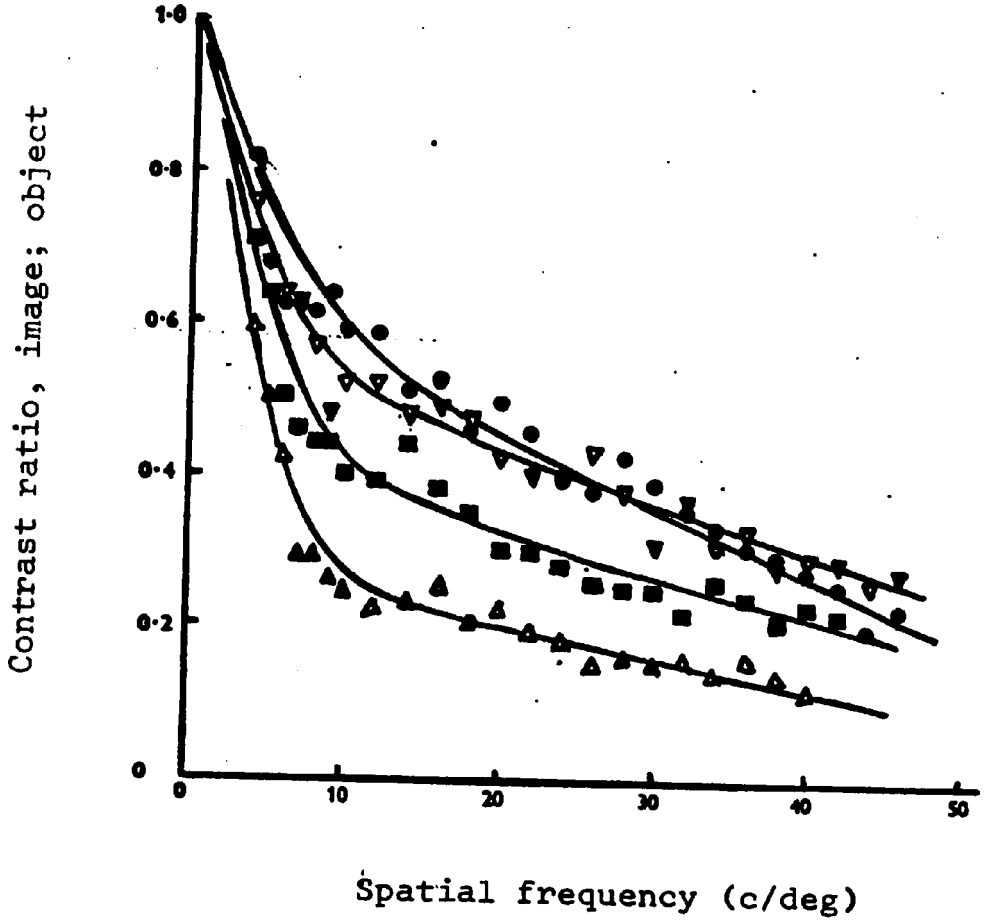


Figure 3-6. Calculated contrast transfer functions for the in-focus optics of the eye at four pupil diameters, ●, 2mm; ▼, 2.8 mm; ■, 3.8 mm; ▲, 5.8 mm (reference 3-14)

of the difference between the double pass and the psychophysical methods. In addition, the results obtained by Gubisch appear to be more accurate than those obtained by his predecessors using the double pass method. Therefore, the modulation transfer functions given by Gubisch are the most suitable measure of the optical performance of the human eye for use in the picture-viewing environment.

3.3 Characteristics of the retinal photo-receptors

It is well known that the detection of light by the retina is performed by two basic types of photoreceptors, the rods and the cones. This concept is known as the duplicity theory of vision. The human retina contains about 6.5 million cones and between 110 and 125 million rods. It should be pointed out that the number of optic nerve fibers is on the order of only one million. (reference 3-15) Therefore, the signal from each individual rod and cone is not transmitted to the brain, but is combined in some way with the signals from other rods and cones before being transmitted to the brain. (reference 3-17) Rods and cones can be distinguished on the basis of their physical shape and size, rods having a thin cylindrical inner and outer segment, the cones having a wider cylindrical inner segment and a thin cone-like outer segment. Rods and cones also differ in other respects, such as receptor density as a function of retinal position, the ability to perceive fine detail, or acuity, and absolute light detection threshold.

3.3.1 Characteristics of the rods

The rods are generally associated with low acuity, colorless, night vision.

3.3.1.1 Density of the rods

The density of rod receptors as a function of distance from the center of the fovea is shown in figure 3-7, taken from Cornsweet (reference 3-18). An area 260 microns in diameter centered at the fovea contains no rods at all. This area contains only cones. Using 4.85 microns as the distance on the retina corresponding to a visual angle of one minute of arc, the rod-free area subtends a visual angle of 54 minutes of arc. (reference 3-19)

3.3.1.2 Spectral sensitivity of the rods

The magnitude of the response evoked in a photoreceptor or group of photoreceptors by incident radiation is dependent upon the wavelength of that radiation. The measure, at every wavelength, of the relative energy required to evoke a fixed visual response is a measure of the luminous efficiency of radiation. The range of wavelengths effective in human vision is relatively narrow, lying roughly between 400 and 700 millimicrons. For rods, the peak luminous efficiency is at about 507 millimicrons. M.H. Pirenne (reference 3-20) describes a method that was used to measure the luminous efficiency of the rods. The method is based on photometric matching at scotopic levels of illumination, 10^{-6} to 10^{-3} candelas per square meter. The field used was a circular area 20 degrees of visual angle in diameter.

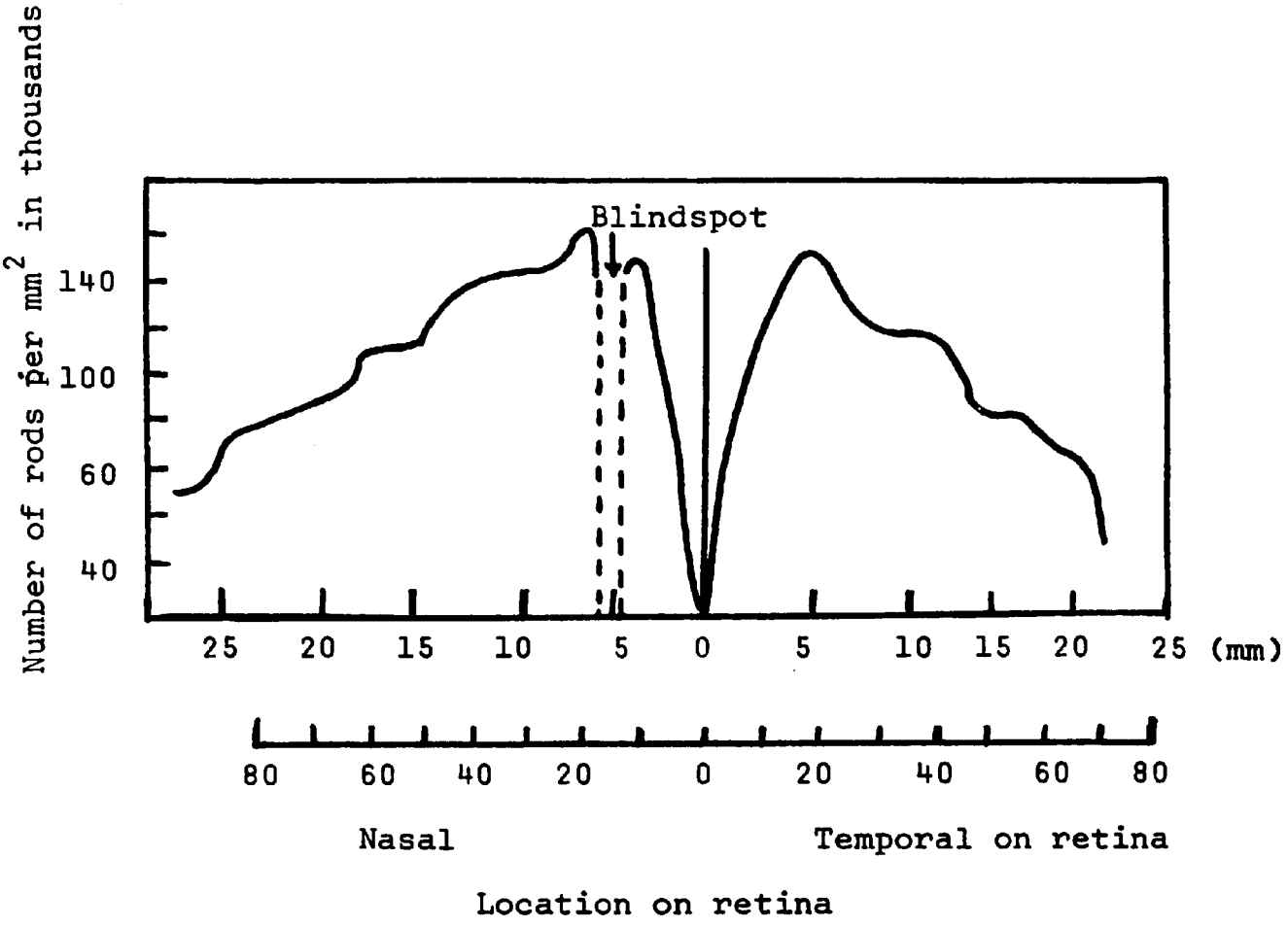


Figure 3-7. Rod receptor cell density as a function of the horizontal location across the retina of the right eye. (reference 3-18)

The circle is divided into two half-circular fields. The circle is viewed so that its center lies ten degrees off the central fovea and therefore in a region predominantly containing rods. One semi-circular field was uniformly lit with white light of luminance 3.2×10^{-5} candelas per square meter. The other semicircle was illuminated with narrow spectral bands centered at different wavelengths. For each narrow-band matching field, the energy was adjusted until it perfectly matched the field containing white light. The matching energy was recorded for narrow bands traversing the entire range of visual wavelengths. The resulting "scotopic" spectral sensitivity curve is shown in figure 3-8.

3.3.1.3 Absolute light detection threshold of rods

There are three types of absolute thresholds commonly measured. They are distinguished on the basis of the size, duration and viewing conditions of the test stimulus.

A large test stimulus of several seconds duration is used to measure the minimum retinal illumination necessary for vision. The photometric unit for illumination is light flux per unit area. Reference 3-23 gives a value for rod illumination threshold in radiometric units assuming a test light whose wavelength is 507 millimicrons. This is the wavelength at which the rods are most sensitive. The value given is 0.45×10^{-14} watts per square centimeter. A constant K_m exists which relates radiometric energy flux to photometric light flux. K_m equal 1745.0 scotopic lumens per watt. (reference 3-27) Therefore, at this wavelength, the threshold illumination is equal to 0.785×10^{-11} scotopic

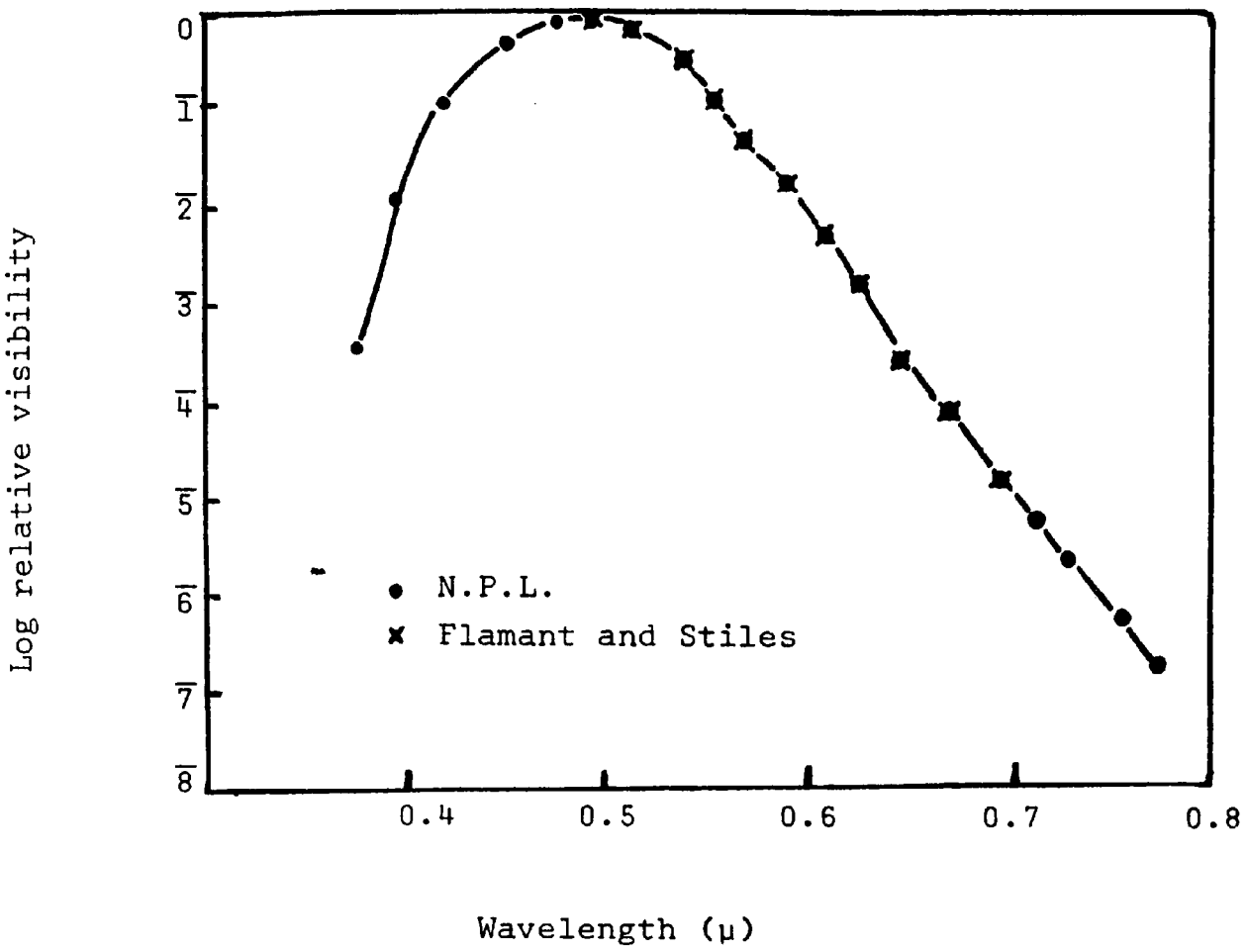


Figure 3-8. The relative rod spectral sensitivity curve. (reference 3-20)

lumens per square centimeter, or 0.0785 microlux. In terms of rate of absorption of quanta by the rods, Pirenne gives 2,300 per second per square centimeter. Another way of expressing threshold illumination is in terms of the luminance of a perfectly diffusing surface illuminated by the test source. For a white light large test field and rod vision using the natural pupil Pirenne (reference 3-23) gives the figure of 0.75×10^{-6} candelas per square meter for absolute threshold luminance. For comparison, Pirenne states that the luminance of the darkest night sky measured at the National Physical Laboratory in England was 14×10^{-6} candelas per square meter. The threshold luminance is noticeably below this value.

A point test stimulus of extremely short duration measures the minimum amount of light energy required for seeing. The rods are truly impressive in this regard. Current thought indicates that a single quantum of light energy is sufficient to stimulate a single rod. (reference 3-21) However, it is necessary for several rods to pool their responses before the flash is actually seen (reference 3-21). Because of the low light levels used at threshold, the few quanta present are randomly distributed in both space and time. Therefore the detection of brief flashes at these low levels of illumination is a random event. Thus a typical definition of absolute threshold is that light level at which the flash is seen fifty or sixty percent of the time.

Hecht, Schlaer and Pirenne (reference 3-22) measured the absolute energy threshold for one millisecond flashes of light of wave length 510 millimicrons. The flash field subtended a visual angle of ten minutes of arc in diameter and was located twenty degrees from the central fovea. They found that the flash was seen sixty percent of the time if it contained from 54 to 148 quanta measured at the cornea. Due to the transmissivity of the eye at that wave length, it was estimated that fifty percent of the quanta entering the cornea actually reach the retina, and of those only twenty percent were absorbed by the rhodopsin in the rods. Therefore it was concluded that even in the case of the brightest flash, containing 148 quanta, only fourteen quanta were effective in stimulating the rods (reference 3-21). The energy associated with fourteen quanta at that wave length is equal to 8×10^{-9} ergs. This is the amount of energy that would be intercepted by the eye from an ideal 100 watt bulb 2,600 miles away flashed for one millisecond.

Measurements of absolute energy threshold indicate some interesting dependencies on flash size, location and duration. For fixed flash location and duration the number of quanta required at threshold is a constant with increasing flash diameter up to a point. Then the number of quanta required begins to increase. This implies that, for each location in the retina, there is a fixed-size area over which rod signals are summed to evoke a response. Similarly if flash duration is the variable, the number of quanta required at threshold is a constant with increasing flash

duration until a duration is reached at which the number of quanta required begins to increase. Thus there is evidence of temporal summation as well. The visual angle subtended by the summation area twenty degrees away from the fovea is ten minutes or smaller in diameter. For a flash two minutes in diameter and fifteen degrees away from the central fovea, the summation time is 100 milliseconds (reference 3-18).

3.3.1.4 Visual acuity of the rods

Visual acuity is defined by Pirenne (reference 3-24) "as the reciprocal of the angle in minutes subtended by the smallest detail which can be seen under given conditions." The "given conditions" will involve the state of the eye's optics, the location of the image on the retina, the spectral characteristics of the illumination used, and the intensity of the illumination used. Acuity is dependent on the eye's optics because the image produced on the retina has been degraded by the optics and the amount of degradation is dependent on the pupil size and the degree of spherical and chromatic aberration. The location of the image on the retina determines the type of photoreceptors that are active and their density in that region. The wavelength properties of the light used will determine the degree of chromatic aberration and the degree of detectability of blur caused by defraction and refraction, this blur tending to obscure details of the image.

Gubisch (reference 3-16) computed the retinal image of test patterns using the modulation transfer function

for the optics of the eye. In general, details in the image with dimensions comparable to the interreceptor spacing become very broad relative to their proper size.

The theoretical maximum acuity of the rods can be approximated using the maximum density of rods, which occurs about twenty degrees away from the fovea. At this point the rod density is about 160,000 rods per square millimeter or 400 rods per millimeter. Therefore the inter-rod spacing is approximately 2.5×10^{-3} millimeters or 2.5 microns. If detection of two point sources requires two stimulated rods separated by one unstimulated rod then the minimum detectable point separation is about 2.5 microns. Since a visual angle of one minute of arc corresponds to 4.85 microns on the retina, (reference 3-23) then the minimum point separation corresponds to 0.52 minutes of visual angle and an approximate acuity of 2.0.

A standard target used to measure visual acuity is the black landolt C. Pirenne (reference 3-24) summarizes several measurements of acuity using this target. The results for white light in the peripheral retina, presumably only for rods, indicate a maximum visual acuity of between 0.05 and 0.1, considerably lower than the theoretical maximum. The maximum light intensity used was 1 millilambert, corresponding to poor interior lighting. (reference 3-25) Therefore, points separated by a visual angle of at least ten to twenty minutes of arc could be distinguished using rod vision, or in terms of high contrast spatial sine wave gratings, this implies a cutoff frequency of 3-6 cycles per degree.

3.3.2 Characteristics of the cones

The cone receptors are associated with high acuity color vision under normal conditions of illumination. Under typical viewing conditions the eye will position itself relative to an object so that the image of that object is in focus and centered on the fovea of the retina. Since this region of the retina, which subtends a visual angle of approximately one degree, contains only cones, it is the cone characteristics that play a predominant role in the viewing of objects such as pictures.

3.3.2.1 Density of the cones

Figure 3-9 describes the density of the cones as a function of retinal position (reference 3-18). As the figure indicates the density is very high in the foveal area and is relatively low in those regions of the retina which also contain rods.

3.3.2.2 Spectral sensitivity of the cones

Measurement of the spectral sensitivity of the cones is complicated by the dual nature of the cone sensation: intensity and color. It is also complicated by the belief that color sensation is due to the existence of perhaps three cone mechanisms. The number of cone mechanisms in the human retina is not yet a closed question. However, best available evidence indicates that there are three mechanisms. As a result, measurements of the spectral

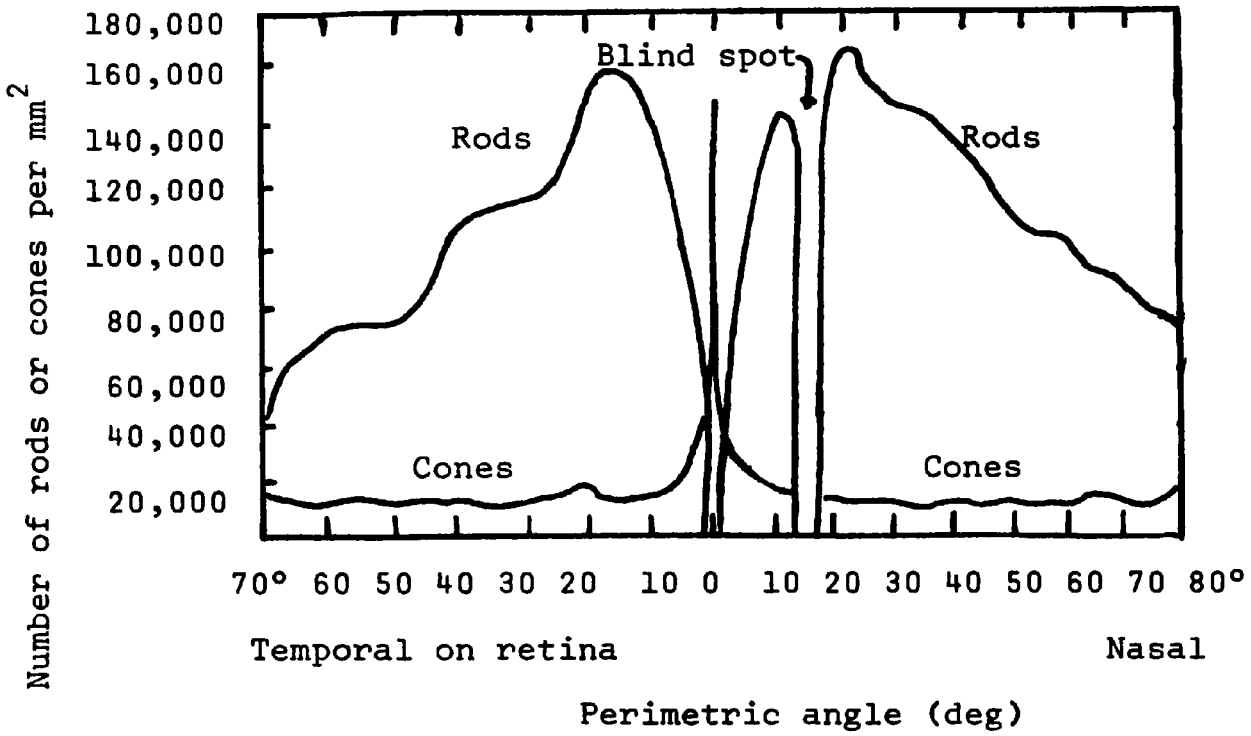


Figure 3-9. Densities of rods and cones across the retina. (reference 3-18)

sensitivity of cones generally result in the composite spectral sensitivity of the three cone mechanisms. Two methods are commonly used to measure spectral sensitivity: the threshold method and the matching method.

In the threshold method, the energy in a flash of light of a certain spot size and duration is adjusted until the flash is just visible, say fifty percent of the time. The relative amount of energy at threshold for each wavelength is a measure of the spectral efficiency. Figure 3-10 compares the spectral efficiency of cones with that of the rods using the threshold method (reference 3-20). While the peak sensitivity of rods occurs near 507 millimicrons, the peak for the cones is at about 560 millimicrons. In addition over a wide range of visible wave lengths, the rods are more sensitive than the cones.

The matching method, while easily applied to rods, is difficult to use on cones because of the difficulty of matching in intensity two semicircular fields of different colors. However, if the standard and matching fields are near to one another in wave length as in the step-by-step method, or if the two fields are alternately flashed, satisfactory results can be obtained (references 3-10, 3-20).

Several methods have been used to attempt to isolate the various cone mechanisms and to measure the spectral sensitivity of the individual cone mechanisms (references 3-10, 3-21, 3-26). Only two of the methods will be mentioned here: the bleaching method and the microspectrophotometry method.

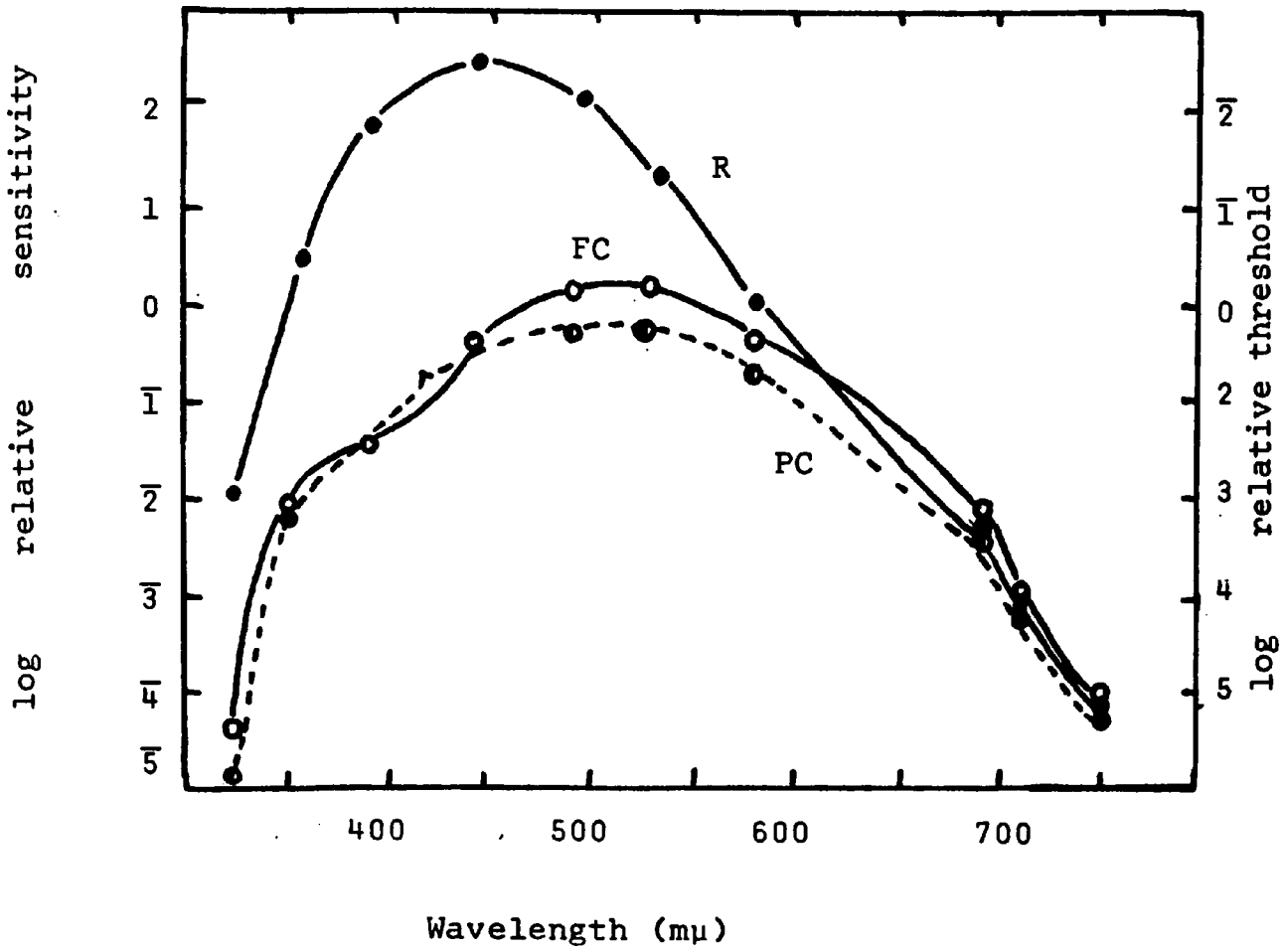


Figure 3-10. Spectral sensitivies (1/threshold) of dark-adapted peripheral rods (full line, R), foveal cones (full line FC) and peripheral cones (broken line, PC). All sensitivities are expressed relative to the maximum sensitivity of the fovea. The relative positions of these functions on the ordinates are therefore those observed in the eye. The test field, 1° in diameter and exposed in 0.04 second flashes was placed at an eccentricity of 8° below the fixation point for the peripheral measurements. (reference 3-20)

This method is based on the idea that if a cone mechanism is exposed to a strong light to which it is uniquely sensitive, then that mechanism will be unable to absorb additional energy. The spectral sensitivity is measured before and after a strong bleach. The "difference spectra" is then a measure of the spectral sensitivity of the bleached cone mechanism.

A more direct method, known as microspectrophotometry has recently been perfected (reference 3-18). In this method, excised human retinas are placed under a specially designed microscope. Fine pencils of light are positioned so that they pass through single photoreceptors and the spectral sensitivity is measured accordingly. Using this method, three classes of cones have been identified. The relative spectral sensitivity of the three cone mechanisms is shown in Figure 3-11.

It has still not been determined whether the three mechanisms are due to different pigments, receptors or structures.

3.3.2.3 Absolute light detection threshold of cones

For cones, measurements of absolute threshold are complicated by the existence of several cone mechanisms.

Graham, Brown, and Mott measured the foveal (cone) threshold illumination using white light and stimuli of radii 0.93 to 30.0 minutes. (reference 3-28) The resulting thresholds expressed in terms of equivalent source intensity were 0.183 candelas per square meter for the smallest radius

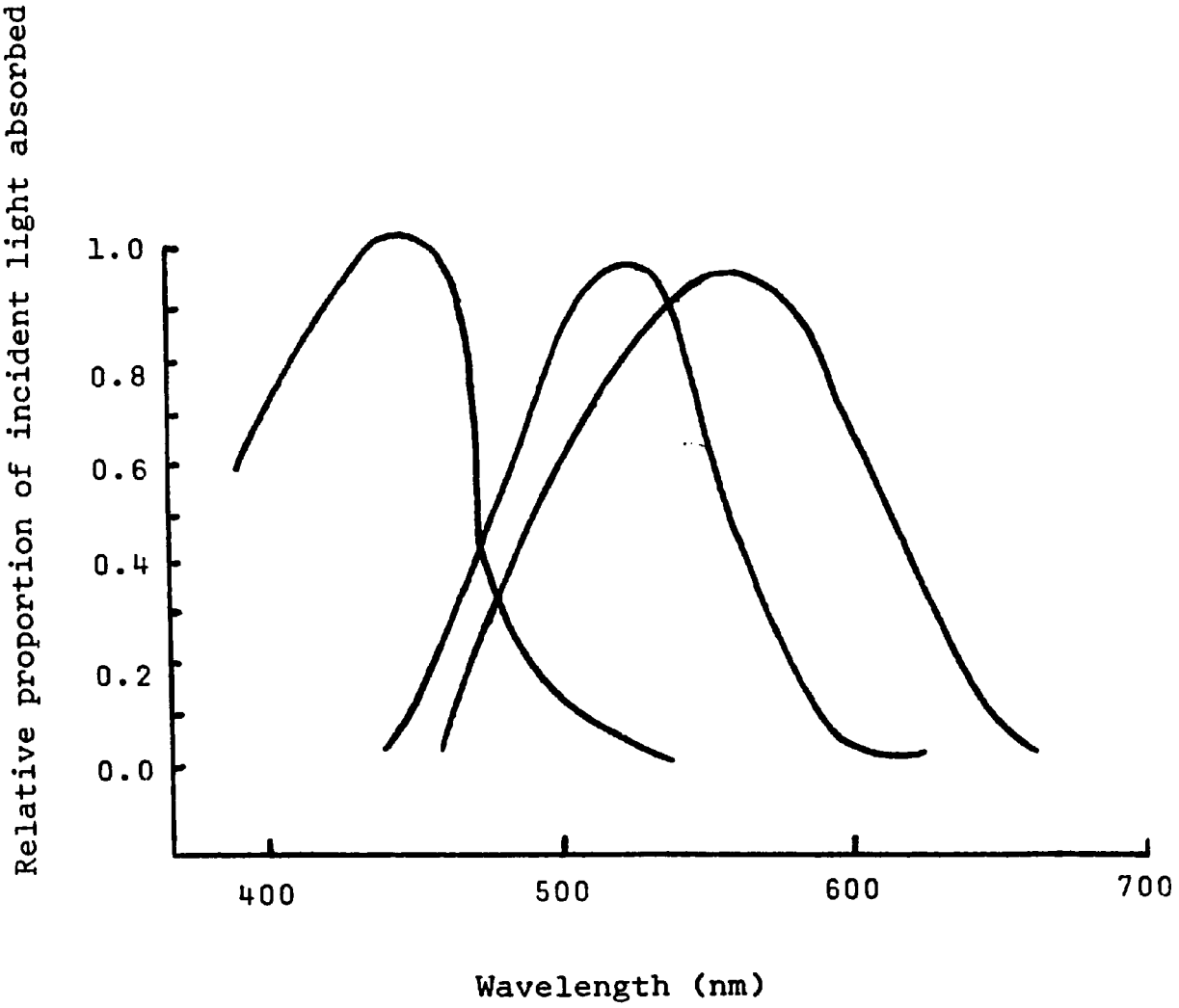


Figure 3-11. Absorption spectra for three classes of cones found in the human retina. The heights have been adjusted arbitarily to have maxima at 1.0 (reference 3-18)

and 0.00664 candelas per square meter for the largest radius. These results are to be compared with their measurements for the peripheral retina (31.5 degrees, horizontally displaced from the fovea). For radii of 0.93 and 30.0 minutes, the rod thresholds were 0.0439 and 0.877×10^{-6} candelas per square meter, respectively.

For a point source of white light the foveal threshold illumination is about 10^{-7} photopic lumens per square meter, or 0.1 microlux (references 3-23, 3-10). When a point test source of long duration is used, then threshold light flux or energy flux is being measured. For rods, using an effective point source of wavelength 507 millimicrons and a long exposure time, the threshold flux is about 100 to 150 quanta per second. Threshold illumination at the pupil is about 4×10^{-10} lumens per square meter, or 0.0004 microlux. For a white light point source, the value is on the order of 0.004 microlux. (references 3-23, 3-10) For cones and a white light point source, the comparable values are 10,000 quanta per second at 555 millimicrons and an equivalent threshold illumination of 0.1 microlux.

When brief flashes confined to fixed regions of the retina are used to measure threshold, the threshold energy is being measured. For rods, using a one millisecond flash (510 millimicrons) located 20 degrees extra-foveally, the threshold was 54 to 148 quanta using a sixty percent criterion. Measurements for brief flashes located in the

fovea indicate an equivalent of 625 quanta for light of 550 millimicrons, the wave length at which the cones are most efficient. Another experiment gave 2,440 quanta at 550 millimicrons.

Therefore, in general the rods have much lower thresholds than the cones. For brief flashes, the rods are an order of magnitude more sensitive than the cones. However, for test fields of long duration, three orders of magnitude separate the absolute threshold of the rods and cones. A typical picture-viewing environment would involve values of illumination in the range 10 to 100 lux. This is to be compared with the cone threshold in white light of 0.1 microlux. Therefore, typical levels of illumination would be eight orders of magnitude above the threshold.

3.3.2.4 Visual acuity of the fovea

The central fovea, subtending a visual angle of one degree, is populated only by cones. The maximum cone density occurs at the center of the fovea and is about 150,000 cones per square millimeter, or 388 cones per millimeter. This implies an intercone spacing of 2.58×10^{-3} millimeters or 2.58 microns. Therefore, theoretically, points separated by a visual angle of 0.53 minutes should be distinguishable. The corresponding visual acuity is 1.88. Theoretical estimates of acuity in the fovea are complicated by the existence of several cone mechanisms and the absence of data on the density of the individual mechanisms in the fovea. Pirenne (reference 3-19) gives the results of

several measurements of acuity. In all cases, the black landolt C was used as the test object. In general, a two millimeter artificial pupil was used. The results indicate the dependence of acuity on the intensity of illumination for white light and an illumination range of one troland (weak interior lighting) to 10,000 trolands (extremely intensive illumination). The visual acuity increases from 0.316 to 2.5. Therefore, the maximum foveal acuity exceeds the theoretical maximum. Walsh (reference 3-10) gives the results for the natural pupil. These results are summarized in figure 3-12.

In terms of high contrast spatial sine wave gratings, an acuity of 2.0 implies a cutoff frequency of 60 cycles per degree.

3.4 Characteristics of the retina and the composite visual pathway

In the preceding sections, characteristics of individual elements of the visual mechanism were discussed. In this section, experimental evidence will be described which cannot be explained in terms of the characteristics of individual elements of the visual mechanism. In general the theories which have been offered to explain the following phenomena are based on interaction among the various elements of the visual mechanism.

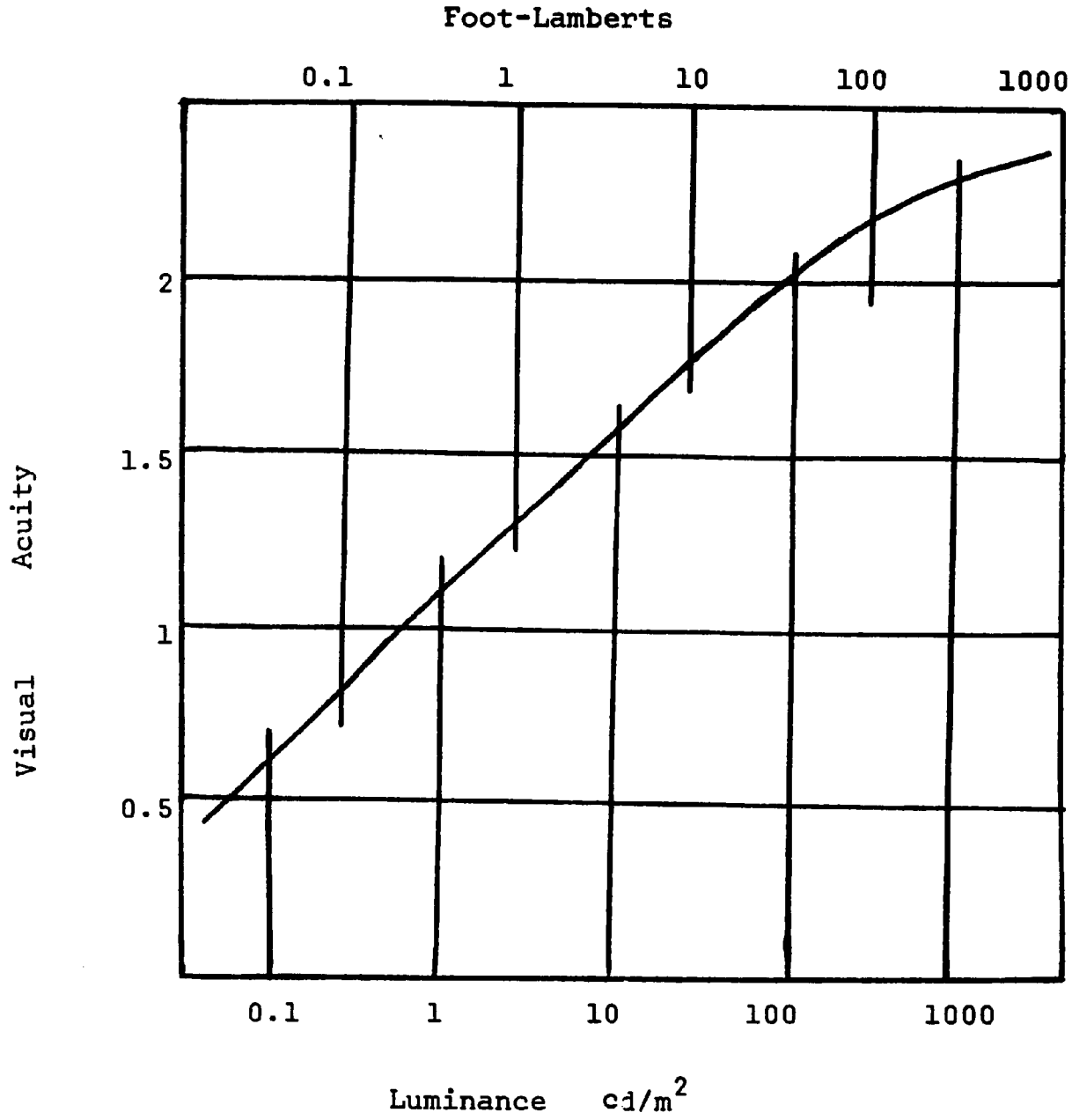


Figure 3-12. The Variation of Visual Acuity with Luminance (reference 3-10)

3.4.1 Spatially varying brightness stimuli

Since any stimulus used in a psychophysical experiment must at some point in time be turned on and then later shut off, the response is always a function of time. Therefore, it is not clear how an experiment could be performed so that the outcome is only a function of spatial variables. However, if the stimulus is presented to the subject a sufficient length of time before the subject is asked to respond, then the response should only be a function of the spatial characteristics of the stimulus.

3.4.1.1 Incremental brightness sensitivity

A common way to measure incremental brightness sensitivity, also called brightness discrimination and contrast threshold, is to use a large, circular background subtending a visual angle of several degrees, with a circular test area at its center. The subject is first adapted to the luminous intensity of the background; then an incremental test flash is superimposed on the background at its center. The subject is asked to indicate whether or not he detected the test flash. The variables of the experiment are background intensity, test flash intensity, test flash duration, test flash size, background wavelength, test flash wavelength, and region of the retina being tested.

If the background intensity is zero, then absolute threshold is being measured as a function of test wavelength and area as well as retinal location.

If different wavelengths are used for the background and test, then because of the different spectral sensitivity characteristics of the various photoreceptor mechanisms, it is possible to study in isolation the individaul photoreceptor mechanisms. The background is selected so as to decrease the incremental brightness sensitivity at the wavelength of the test for the unwanted mechanisms.. (reference 3-26)

Let the intensity of the background be designated by I and the intensity of the test flash be indicated by ΔI. Then, in general, the ratio $\frac{\Delta I}{I}$ decreases as the background intensity I is increased. For large values of I, $\frac{\Delta I}{I}$ approaches a constant. This behavior is known as Weber's law. If $\frac{\Delta I}{I}$ as function of background intensity is being measured for extrafoveal regions of the retina, then as I is increased, the curve will develop two branches, one for the rods and another at higher intensities for the cones. Increasing the test flash area or duration will decrease the contrast threshold. The magnitude of these effects depends upon the region of the retina being tested. If I is held constant and ΔI is measured as a function of flash duration or flash area, it is found that complete summation of the energy in the flash occurs for a range of durations and areas on the low end of the scale. That is, threshold test flash intensity decreases in proportion to increases in test flash duration or area (references 3-29, 3-30).

The data on contrast thresholds most relevant to this work was obtained by Blackwell (reference 3-21). He measured contrast thresholds using white light and natural vision. No fixation or artificial pupil was used. As a result, for each experimental condition, the subjects used the region of the retina that gave best detectability of the stimulus.

The experimental parameters of Blackwell's experiment were background intensity, test field intensity, and test field size. Test field duration was not a parameter. Since the purpose of the experiment was to measure maximum contrast sensitivity, long duration test fields were used. Blackwell used a specially constructed observation room 63 feet long, ten feet high, and ten feet wide. The observation screen at the front of the room subtended a visual angle of ten degrees. The viewing screen and the immediately adjacent walls formed the background. Figure 3-13 below gives the results for a six second exposure of the test field. In this case the stimulus was randomly positioned in one of eight equally spaced points on the circumference of a circle three degrees in radius generated about the center of the observation screen. Additional experiments with longer exposures indicated that exposures as long as sixty seconds were not sufficient to determine minimum thresholds. However, when the test stimulus was always presented in a fixed position, fifteen second exposures were found to be adequate for minimum threshold.

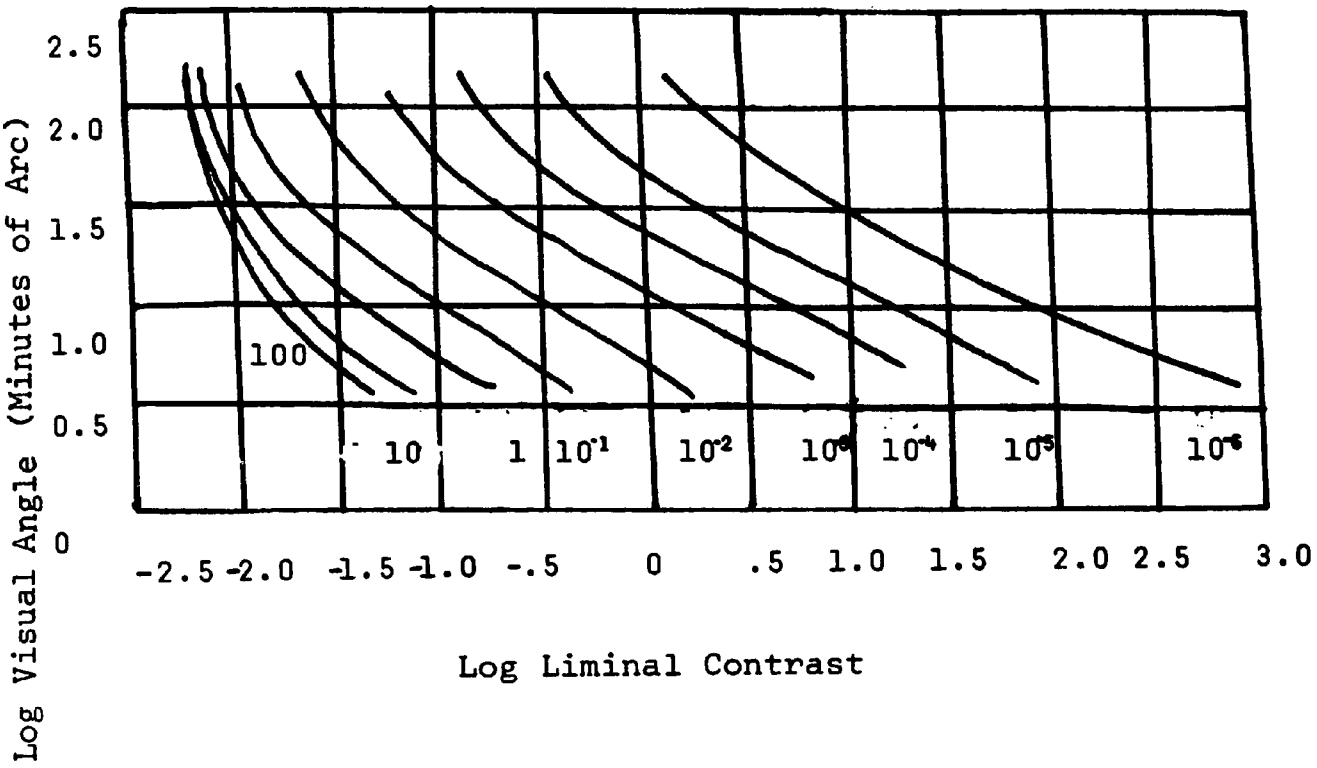


Figure 3-13. The relation between threshold contrast and stimulus area for a given adaptation brightness (foot-lamberts) for 6 sec exposure time and 8 stimulus positions (reference 3-31).

The results for this case are reproduced in Figure 3-14. From this figure it can be seen that for a six-degree stimulus, and a background brightness of 100 foot-lamberts (good interior lighting) the logarithm of the threshold contrast is -2.5. Therefore $\frac{\Delta I}{I}$ equals 0.00316 or approximately a 300 to 1 ratio.

Pirenne (reference 3-23) summarizes the basic theories that have been put forth to explain intensity discrimination. One of these theories postulated that an incremental stimulus ΔI will be detected when it causes a constant increment of photochemical decomposition of the pigment. For example, equations exist which describe the photochemical decomposition of rhodopsin, the rod pigment, as a function of time and intensity. There are also equations that describe the regeneration of rhodopsin as a function of time. For constant intensity, a steady state is reached at which the rate of decomposition and regeneration are equal and the proportion of bleached to unbleached pigment is a constant.

An incremental change in the intensity increases the rate of decomposition. This in turn increases the density of decomposition products causing an increase in the rate of regeneration. However, at the saturation level of the rod mechanism, where the contrast threshold becomes infinite, it was shown that only ten percent of the rod pigment is bleached, yet no increment of photochemical decomposition could be large enough to be detected, even though sufficient unbleached pigment is present.

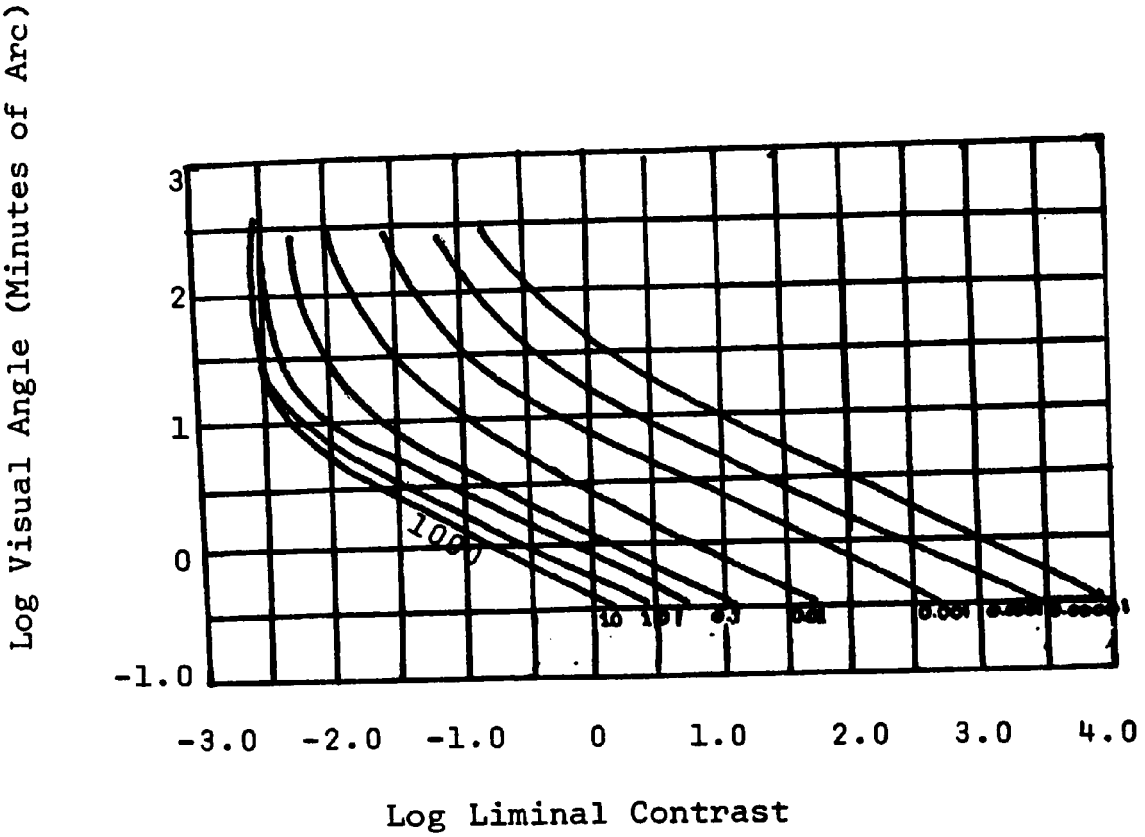


Figure 3-14. The relation between threshold contrast and stimulus area for a given adaptation brightness (foot-lamberts) and for 15 sec. exposure times and single stimulus position (reference 3-31).

Rushton (reference 3-32) describes a model for incremental brightness sensitivity which incorporates a feedback mechanism which causes a non-linear decrease in incremental sensitivity as intensity is increased. The feedback is a function of background intensity and the level of photochemical decomposition of the pigment. The model he describes is primarily for rods. However, he proposes that a separate, independent feedback mechanism exists for each of the different cone mechanisms.

Another explanation given for contrast threshold is the "noise" theory. It is best described by Barlow (reference 3-30).

> "In counting the quanta absorbed from the stimulus a certain number of quanta coming from the background will inevitably be counted with them, and according to the hypothesis to be tested the threshold corresponds to the limit set by the fluctuations in this number. The limit will be lower if this number is reduced, but at the same time the maximum possible number of quanta from the stimulus must be collected. Using stimuli which are sharply delimited in space and time, the best performance will clearly be obtained by 'sampling'the numbers of quanta absorbed over areas and durations matched to the area A and duration T of the stimulus. When A and T are increased, the 'sampling' area and time must also be increased, and the average number of quanta counted from the background will then be increased in direct proportion to the area and duration of the stimulus. The actual numbers absorbed on particular occasions will be distributed according to the Poisson distribution, to which an approximation is given by a normal distribution of standard deviation equal to the square root of the average number of quanta absorbed. Hence if the threshold response is of constant reliability, ΔN is proportional to $A^{1/2}T^{1/2}I^{1/2}$
>
> where ∇N is the incremental quantity of light; or ΔI is proportional to $A^{-1/2}T^{-1/2}I^{-1/2}$
>
> where ΔI is incremental intensity."

However, Barlow found that the square root law only held for an unreliable range of areas and a short range of times and goes on to hypothesize some reasons for the discrepancies.

Spatial interaction among the photoreceptors in the retina has also been used to explain contrast thresholds (references 3-13, 3-33). One type of spatial interaction is described by the following model: consider a particular photoreceptor. The output signal from that photoreceptor is not only a function of the light incident on that photoreceptor, but is also a function of the output signals of the surrounding photoreceptors. A particular surrounding photoreceptor will tend to inhibit the output signal of the central photoreceptor as a function of its distance from that photoreceptor. On the basis of such a model, it is possible to explain how a surrounding background would tend to decrease the response to an incremental central stimulus. This concept will be discussed later in more detail.

3.4.1.2 Simultaneous Contrast

In the preceding section, results were given which describe the ability of the eye to detect brightness differences within the stimulus area. This section is concerned with the relationship between subjective perceived brightness and physical intensity. It is not unreasonable to expect that perceived brightness would be a monotonically increasing function of physical intensity.

However, irrefutable evidence exists which shows that this is not the case. The perceived brightness of an object is highly dependent upon the luminance of its surroundings. By holding the luminance of an object fixed, and varying the luminance of its surroundings over a wide range, the object can be made to appear any shade from black to white.

Cornsweet (reference 3-18) describes some experiments that were used to measure the effect of a surround on the brightness of an object. The stimulus configuration used in the experiment consists of a large, black background upon which are superimposed two disks 28 minutes in diameter. One of the disks is surrounded by an annulus 55 minutes in diameter. The experiment is arranged so that the observer sees one of the disks with one eye and the other disk with the other eye. Let the luminance of the surrounding annulus be L_b and the luminance of the disk that it surrounds be L_t. L_m will designate the luminance of the other disk. In the experiment, a fixed value is chosen for L_t. Then the value of the surrounding luminance L_b is varied from 0 or black to values slightly larger than L_t. For each value of L_b, the observer adjusts L_m, the luminance of the matching disk, until L_m and L_t are perceived to be identical. Figure 3-15 shows the results that are obtained from this experiment for different values of L_t. As would be expected, for $L_b = 0$, the observer choses a value for L_m which is numerically identical to L_t. However, as L_b is increased, L_t "appears" to first increase slightly and then decrease to the point

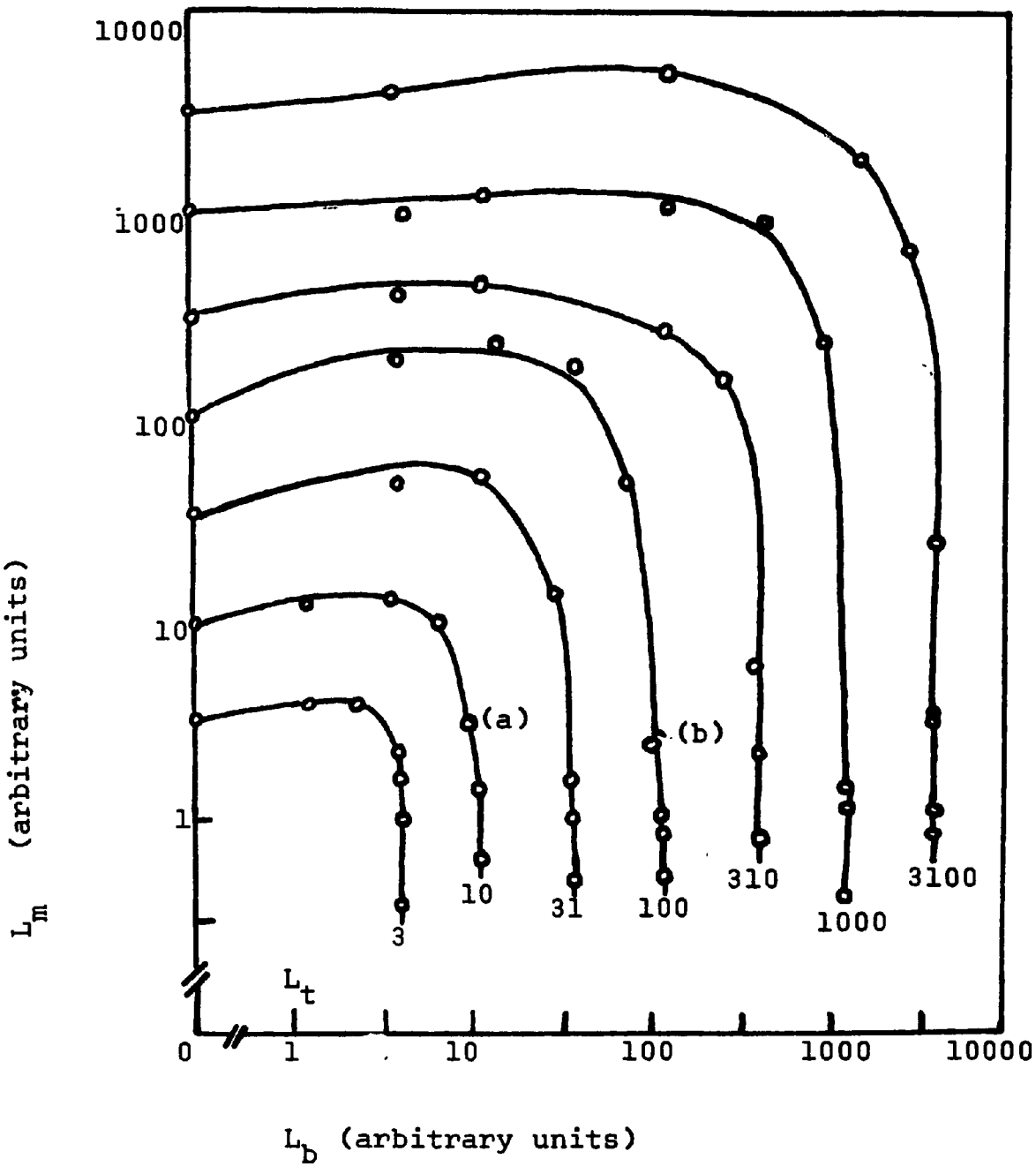


Figure 3-15. The effect on the perceived luminance L_m of a test disc L_t due to its surround L_b. (reference 3-18)

where L_b equals L_t, beyond which L_t "appears" to approach zero very rapidly. Therefore, even though the physical intensity of L_t was held constant, the effect of varying the luminance of its surround was to change its perceived brightness over a wide range.

The preceding experiment shows how perceived brightness can be independent of physical luminance. Now consider an object whose physical luminance measured at the eye is a constant for every point on the object. Will the object appear to be of uniform brightness? Not necessarily. In fact, it is also possible for an object of non-uniform luminance to appear to be of uniform brightness. Cornsweet (reference 3-18) gives photographic evidence for the complex relationship between physical luminance and perceived brightness on pages 272-279 of his book.

One phenomenon which demostrates the complex relationship between physical intensity and perceived brightness has been studied in great detail. This phenomenon has been given the name Mach bands. For a detailed study of Mach bands, as well as an account of their discovery, see Ratliff (reference 3-33). Mach bands are associated with the border between two fields of uniform but different luminances. Assuming the conditions are proper, the observer will see a bright band on the side of the border corresponding to the field of larger luminance and a dark band on the side of the border corresponding to the weaker field. The actual luminance in the area of the bands is identical to that of the rest of the field.

The precise conditions under which Mach bands appear are not known but are a function of the luminances of the two fields and the function describing luminance versus distance within the border. Brightness matching techniques have been used to measure the variation of subjective brightness in the region of a stimulus containing Mach bands. These subjective brightness contours can then be compared for different values of a stimulus parameter, such as the slope of a linear luminance transition between the two uniform fields (references 3-33, 3-34).

Models of the retina have been proposed which account for some of these contrast effects. A presentation of these models, as well as their application to contrast effects, will be found in Chapter 4.

3.4.1.3 Brightness constancy

Brightness constancy is a non-intuitive property of human vision. Stated simply, this property implies that the brightness of an object and its surroundings will be constant regardless of the level of illumination.

Cornsweet (reference 3-18) describes an experiment used to test brightness constancy. An observer, sitting at the rear of a room, views a screen at the front of a room which is a standard shade of gray. Directly in front of the observer are several cards of different gray shades. The cards are shielded from the room illumination and are themselves illuminated by a constant source. For several values of room illumination, the

observer is asked to choose the card whose grayness matches that of the screen at the front of the room. For a wide range of room illumination the subject will choose the same matching card. In a preceding section, an experiment was described which demonstrated how varying the brightness of a surround caused the brightness of an object not to be constant. The difference between that experiment and this one is that the illumination of the object and its surround is being changed. As the illumination is changed, the physical quantity that remains constant in this case is the ratio between the light being reflected from the object and the light being reflected from any other objects in the room. This can be determined if the reflectance of all the objects in the room is known. Figure 3-15 can also be used to demonstrate brightness constancy since if L_t and L_b are increased (or decreased) by the same factor, L_m is almost constant. Points (a) and (b) in Figure 3-15 show this for a factor of 10. Cornsweet somewhat offhandedly rejects the possibility that the eye's behavior is based on reflectance estimation, either due to cognitive processes based on experience or physiological mechanisms. His explanation, instead, is based on a logarithmic stage following the photoreceptors, and prior to a lateral inhibition mechanism; the net effect of increased illumination is a small additive constant.

Another property of vision similar to brightness constancy is called color, or hue, constancy. A dramatic demonstration of this property is described by Land and

McCann (reference 3-35). They used as their test stimulus an array of rectangles of varying size, shape, and color, similar in appearance to a Mondrian painting. The Mondrian-like stimulus was illuminated by three projectors. The light coming out of each projector was filtered so that each projector emitted a unique band of spectral energy in the visable range. The three bands were chosen so that they adequately covered the entire range of visable wavelengths. The intensity of the three projectors was adjusted so that all the colors in the stimulus looked natural. This was not a critical adjustment. A particular rectangle was selected, and its luminance was measured in the three spectral bands using a telescopic photometer. The telescopic photometer was then repositioned to measure the luminances of a different rectangle. Since the hue of the second rectangle was different from that of the first, the luminance measurements were also different as would be expected. This is merely a confirmation of the well-known fact that the reflectance versus wavelength for an object of one color will differ from that for an object of another color. Now, with the photometer still aimed at the second rectangle the intensity of each projector was adjusted so that the luminance measurements were identical to those given by the first rectangle. After the adjustments, the hue of the second rectangle was unchanged and certainly was not identical to the hue of the first rectangle. In fact none of the rectangles changed hue when their luminances were adjusted to be

identical to that of the first rectangle. It was concluded that the hue of a rectangle was not related to the absolute spectral composition of the light energy reaching the eye from that rectangle. These results were not changed significantly by photochemical adaptation or flashed presentation of the stimulus for ten milliseconds. In addition, if a rectangle was moved so that it was surrounded by a different set of colored rectangles, hue constancy still resulted.

In contrast to Cornsweet's explanation for brightness constancy, Land and McCann concluded that hue constancy is related to the ability of the eye to assess the reflectance properties of an object. Further investigation indicated that the boundaries between rectangles played an important role in hue constancy. If the boundary between two different shades of gray was obscured, the difference in grayness disappeared. As a result of this observation, Land and McCann proposed an algorithm based on reflectance discontinuities at boundaries which was capable of computing the relative reflectance of different objects in a visual field. In addition, they proposed a neural network called a retinex which was capable of performing the algorithm. They proposed that hue constancy was due to the existence of three retinexes each capable of computing the relative reflectance properties for each of the three spectral sensitivity bands associated with

the three types of cone mechanisms. This model might be useful for explaining some of the "Cornsweet illusions" described previously. A detailed account of retinex theory will be found in reference 3-35.

3.4.1.4 Sine wave and square wave contrast sensitivity

The contrast sensitivity of the human eye for periodic spatially varying brightness distributions is an important psychophysical characteristic of vision. Contrast sensitivity is proportional to reciprocal contrast. Contrast is expressed in terms of the maximum luminance BMAX and the minimum luminance BMIN of the periodic spatial brightness distribution. The equation below (reference 3-3) is a conventional measure of contrast.

$$C = \frac{BMAX - BMIN}{BMAX + BMIN}$$

When the brightness distribution is sinusoidal with average value B and amplitude ΔB, the expression for contrast becomes (reference 3-4)

$$C = \frac{[(B + \Delta B) - (B - \Delta B)]}{[(B + \Delta B) + (B - \Delta B)]} = \frac{\Delta B}{B}$$

Contrast sensitivity is an important characteristic because it yields results which might be useful for describing portions of the eye in terms of a linear system. If contrast sensitivity measurements are based on a sinusoidal luminance distribution, then the results are indicative of the relative ability of the eye to perceive spatial sinusoids of different frequencies. Therefore, under certain conditions, contrast sensitivity is analogous

to the modulation transfer function of a linear system. These conditions, and the description of vision in terms of a modulation transfer function, will be discussed in Chapter 4.

At least three different methods have been used to measure contrast sensitivity. The threshold method is probably the most common technique that has been used (References 3-14, 3-36, 3-37, 3-38, 3-39, 3-40, 3-41, 3-42, 3-42). In the threshold method the observer is presented with a one-dimensional sine-wave or square-wave luminance distribution. The contrast of the brightness distribution is then adjusted until the minimum contrast is found for which the observer is able to distinguish the brightness distribution from a uniform field. This procedure is repeated for a range of spatial frequencies of the brightness distribution and the results of several trials are averaged to determine the contrast sensitivity.

Another method that has been used to measure contrast sensitivity is the brightness matching method (reference 3-44). In this method optical techniques are used to produce a visual field consisting of two contiguous half-fields. The upper half-field is uniform and its luminance can be independently controlled. The lower half-field contains a one-dimensional sinusoidal luminance distribution. The observer determines the two values of luminance for the uniform field that cause its brightness to match the minimum and maximum values of brightness in the sinusoidally varying luminance distribution. If the two values of

luminance obtained are substitued into equation 3-3, then the subjective contrast of the sinusoidal brightness distribution is obtained. This is to be compared with objective contrast which is obtained from photometric measurements of the maximum and minimum luminance in the sinusoidal luminance distribution. The ratio given by $\frac{\text{subjective contrast}}{\text{objective contrast}}$ is of interest for different values of the parameters of the sinusoidal brightness distribution. A third method that has been used to measure contrast sensitivity is the contrast matching technique (reference 3-45). Contrast matching is accomplished by comparing the perceived contrast of two one-dimensional sinusoidal brightness distributions. The average luminances of the two sinusoidal distributions are identical but their spatial frequencies are different. One of the sinusoidal distributions is used as a standard and its contrast and spatial frequency are held constant throughout the experiment. The observer compares the subjective, or perceived, contrast of the other sinusoidal distribution against the contrast of the standard and adjusts its contrast until they match. The contrast ratio, or equivalently, the luminance modulation amplitude ratio is then used as a measure of contrast sensitivity at the spatial frequency of the matching field, relative to the contrast sensitivity at the spatial frequency of the standard field. This procedure is repeated using a range of spatial frequency values for the matching sinusoidal field.

Contrast sensitivity is dependent upon the experimental conditions used for its measurement. These experimental conditions are important to any use that might be made of contrast sensitivity measurements. What follows is a discussion of some of the relevant experimental conditions.

1. Type of vision used- contrast sensitivity has been measured using both binocular and monocular vision. In addition, the vision can be either fixated or non-fixated. In non-fixated vision, the observer uses that portion of the eye that yields best results whereas in fixated vision the portion of the retina for which contrast sensitivity is being measured can be controlled. Daitch and Green (reference 3-43) used fixated vision to measure the contrast sensitivity of the peripheral retina. They used sufficiently low average luminance values that the contrast sensitivity of only rods could be measured. For a flashed 200 millisecond sinusoidal luminance distribution and a mean retinal illumination of one troland (one millimeter artificial pupil) the maximum contrast sensitivity obtained was 10 and the cut-off spatial frequency was about 6 cycles per degree. This result might be compared with those obtained by Schober and Hilz (reference 3-38) using flashed 200 millisecond square-wave luminance distributions and natural vision. For an average luminance of 110 candelas per square meter (moderate cone illumination- one troland would

be a moderate level of rod illumination) the maximum contrast sensitivity was about 25 and the cutoff frequency was about 30 cycles per degree.

2. Condition of the pupil and lens - If the pupil and lens are used in their natural state, then the pupil diameter is determined by the retinal illumination and lens accommodation is determined by the distance of the test field. Contrast sensitivity measurements have also been made using an artificial pupil placed very close to the eye. Quite often the natural pupil will be chemically dilated making it possible to artificially control pupil diameter over a wide range. In any case, retinal illumination must be based on the effective pupil diameter. Lens accommodation may be chemically neutralized in which case auxiliary lenses are used to focus the test field on the retina.

The condition of the lens will determine the degree of spherical and chromatic aberrations present. For large pupil diameters the Stiles-Crawford effect must be taken into account. Green and Campbell (reference 3-39) studied the effects of focus on contrast sensitivity. Campbell and Green (reference 3-14) studied the effects of pupil diameter and focus. They found that for small pupil diameters optimum focus for maximum contrast sensitivity was a constant independent of spatial frequency, but for large pupil

diameters (e.g. 7 mm) optimum focus was a function of spatial frequency. They concluded that spherical aberration was involved at large pupil diameters, requiring as much as 1 diopter of refractive power for correction. For a pupil diameter of 2 mm the contrast transfer ratio of the optics approaches that for a perfect diffraction limited optical system. As the pupil is increased the departure from a perfect system increases due to spherical and chromatic aberrations.

3. Viewing distance and visual angle - The distance between the observer and luminance distribution will determine the state of accommodation of the lens of the eye. This, in turn, determines the amount of aberrations that might be present. DePalma and Lowry (reference 3-37) studied the effect of viewing distance and found lower contrast sensitivity for nearer test stimuli due to increased aberrations of the lens, which is thicker for near objects. The angular subtense at the eye of the luminance distribution will determine the extent of the region on the retina illuminated by the luminance distribution. Therefore, the types of photoreceptors being illuminated as well as their density can be determined.

4. Average luminance and state of adaptation -In Figure 3-12, the visual acuity of the fovea is shown to be dependent on the luminance of the test object. Similarly, it would be expected that the eye's ability to perceive the closely spaced bars of a sinusoidal luminance distribution of high spatial frequency would be a function of the average luminance. Patel (reference 3-41) studied the effect of average luminance on contrast sensitivity and found that the contrast sensitivity for high spatial frequencies increased with increasing retinal illumination. In general, the entire contrast sensitivity curve was a function of retinal illumination. Patel's results are summarized in Figure 3-16. Shade (reference 3-36) also measured the effects of luminance and found qualitatively similar results. The dependence of acuity and contrast sensitivity on luminance is thought to be due to groups of photoreceptors pooling their signals and therefore behaving like a single photoreceptor. The number of receptors in a group is believed to increase as the retinal illumination decreases. Therefore, the very dense retinal mosaic of photoreceptors is "effectively" relatively coarse at very low levels of illumination.

The state of adaptation of the eye is important because it determines the sensitivity of the eye to luminance differences in the test field. This

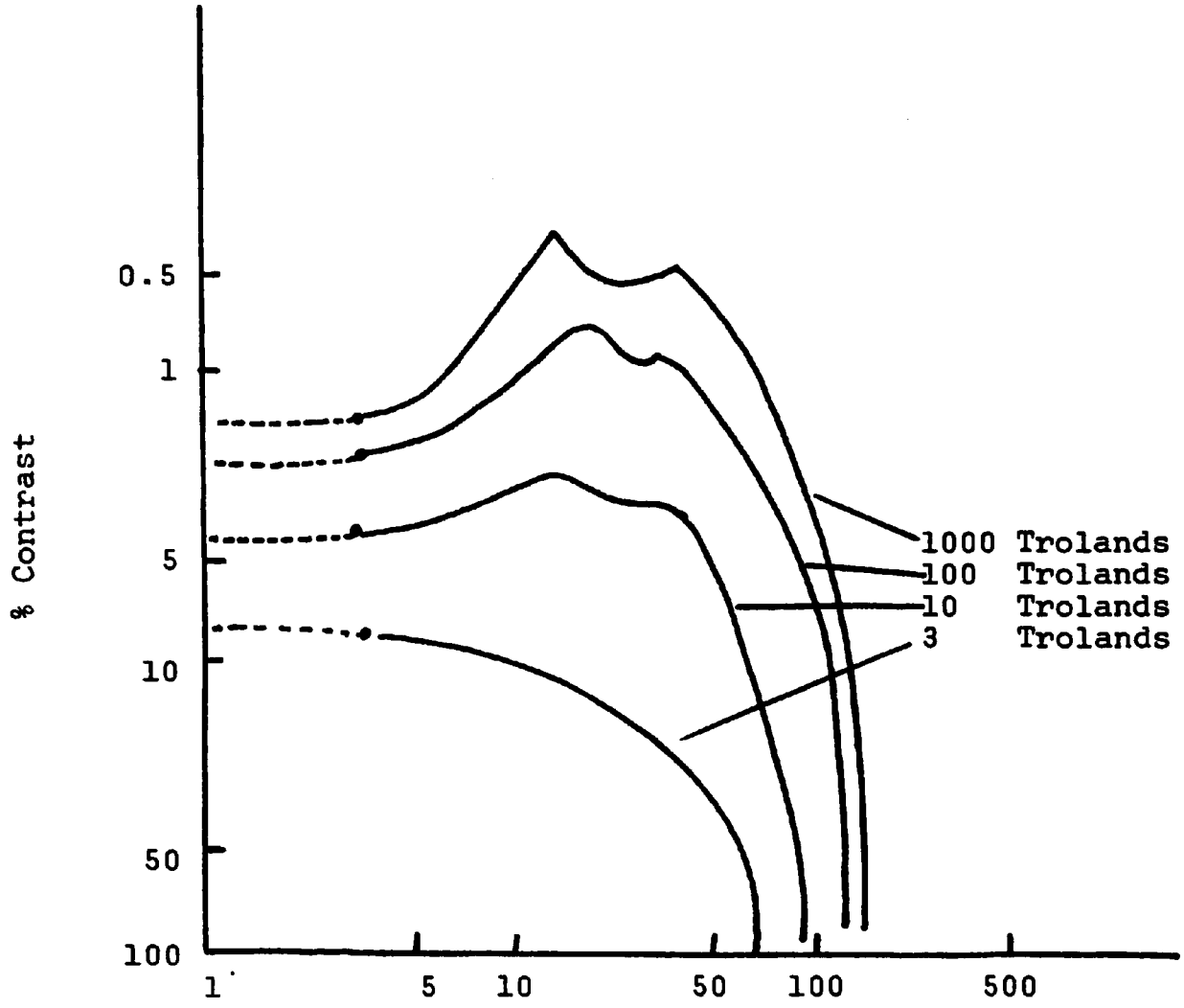


Figure 3-16. The modulation transfer functions for the visual system at four different mean retinal illuminances (reference 3-41).

dependence of incremental brightness sensitivity on adaptation is readily experienced by all of us. If we have been in a dark room long enough to dark-adapt and a light is switched on, the sensation is almost painful whereas the same light switched on during daylight has an almost unnoticeable effect. The luminance increment is the same in both cases. However the ambiant luminance and state of adaptation are quite different. Rushton (reference 3-32) has indicated that each state of adaptation is functionally equivalent to presenting the test stimulus against a background of a particular luminance. He demonstrated this equivalence by means of flashed increment threshold measurements during the course of dark adaptation. However, if his results can be applied to simultaneous contrast, then the equivalent background luminance corresponding to a particular state of adaptation would have to be added to the average luminance of the sinusoidal brightness distribution, if the observer had not been preadapted to the actual average luminance.

5. Luminance source - The spectral composition of the luminance source will determine the amount of image blur due to chromatic aberration. The degree of participation of the different types of photoreceptors in the contrast sensitivity measurements will be a function of the spectral sensitivity of the photoreceptors and the spectral composition of the source. Some common

luminance sources used in contrast sensitivity measurements are "white" lights of various color temperature, cathode ray tube phosphors in the green portion of the visual spectrum and coherent monochromatic laser sources (reference 3-14).

6. Spatial frequency range- Contrast sensitivity measurements for a sinusoidal luminance distribution show three characteristic ranges of spatial frequency. On the low end of the spatial frequency scale, going from one cycle per degree to 0.1 cycles per degree, the contrast sensitivity decreases to about one tenth of its maximum value. (references 3-37, 3-45). In the range of spatial frequencies from one cycle per degree to ten cycles per degree the contrast sensitivity reaches its maximum value. Typical maximum values of contrast sensitivity obtained using the threshold method lie in the range 100 to 500 (references 3-14, 3-41). For spatial frequencies above ten cycles per degree, the contrast sensitivity drops off very rapidly, reaching the minimum possible value of contrast sensitivity, 1.0, at about fifty cycles per degree. (references 3-14, 3-37) The typical values given above are for an assumed average luminance on the order of one hundred candelas per square meter.

7: Contrast sensitivity method and stimulus duration - The three methods used to measure contrast sensitivity, threshold, contrast matching, and brightness matching, have previously been described. In most cases, contrast sensitivity values are the average of a large number of measurements. However, because psychophysical measurements tend to be stocastic, it is possible to describe the results in terms of psychometric functions. These S-shaped curves result when, for example, the percentage of responses above threshold are plotted against contrast for fixed spatial frequency and average luminance (reference 3-42). Contrast sensitivity would then be computed from the set of contrast values obtained by taking the one point on each curve for which the percentage of responses above threshold is equal to a certain value.

The duration of a flashed spatial luminance distribution will affect the values of contrast sensitivity obtained. Most contrast sensitivity measurements have been made using stimuli presented for several seconds or even minutes. However, some measurements of the effects of stimulus duration have been made for square wave distributions (references 3-38, 3-42). In general, contrast sensitivity decreases with decreasing duration.

8. Stimulus orientation - Most contrast sensitivity measurements are based on periodic spatial brightness stimuli whose luminance varies only in the horizontal direction. Campbell et. al. (reference 3-40) studied the effect on contrast sensitivity of the orientation of one-dimensional sinusoidal luminance distributions. The results indicate that contrast sensitivity is maximum for vertical and horizontal orientations and is minimum for forty-five degree and 135 degree orientations. The difference between orientation dependent maximum and minimum contrast sensitivity increases with increasing spatial frequency.

Measurements have also been made using drifting periodic luminance distributions instead of stationary ones (reference 3-36).

Table 3-1 contains a list of authors who describe contrast sensitivity measurements and summarizes the experimental conditions under which their measurements were acquired.

The response of the eye to sinusoidal spatial brightness distributions can be divided into two parts, the response of the optics of the eye and the response of the retina-brain. The quality of the optics of the human eye, and, in particular, the response of the optics to sinusoidal gratings was discussed in section 3.2.5. The response of the optics of the eye can be summarized in the following way: (a) a sinusoidal spatial luminance distribution

Table 3-1. (Part 1) Summary of contrast sensitivity measurements and the associated experimental parameters.

	VISION	PUPIL LENS	LUMINANCE; ADAPT/READAPT.	DISTANCE; ANGULAR SUBTENSE
DePalma, Lowry (1962)	BINOCULAR FOVEAL	NATURAL	20; 300Ft-L	10,14,35,56,120 in. 6°,6°,6°,6°,1°40'
Bryngdahl (1964)	MONOCULAR FOVEAL	2 mm PUPIL	0.25, 0.60, 1.3 2.0, 0.79, 1.89 4.08, 6.28TD; 1 hr./5 min.	12.6 in.; 17.8°
Patel (1966)	MONOCULAR FOVEAL	2 mm PUPIL	3, 10, 100, 1000 TD;15-30min	39.3 in.; 2°
Green, Campbell (1965)	MONOCULAR FOVEAL	7mm, 2mm pupil	29.19 Ft-L	57 in; 2° X 1.3°
Daitch, Green (1969)	MONOCULAR 12° EXTRA FOVEAL	1 mm VERT DISPL.3 mm (STILES-CRAWFORD)	1.5, 1.0, 0.5, 0.0, -0.5, -1.0, -2.0, -2.5, LOG TD; 30 MIN	8° X 3°
Davidson (1968)	MONOCULAR FOVEAL	2.5 mm PUPIL	696Ft-L; 1 min/10 sec.	6°40'
	UNIT OF SUBJECTIVE CONTRAST WAS THE FORCED CHOICE AVERAGE THRESHOLD FOR f<3.6 cycles/degree			
Campbell, Green	MONOCULAR FOVEAL	N.A.	500 TD	N.A.;30°
	MONOCULAR FOVEAL	2, 3.8, 5.8 mm 1.5,2,2.5,3.5D	29.19Ft-L	57 in; 2° X 1.3°

Table 3-1 (Part 1) continued

	VISION	PUPIL; LENS	LUMINANCE; ADAPT/READAPT	DISTANCE ANGULAR SUBTENSE
Schober, Hilz (1965)	BINOCULAR FOVEAL	NATURAL	0.41-32.1Ft-L	39.3,122,276 in; 12° -1°
Nachmias (1967)	MONOCULAR FOVEAL	2 mm	9.3 Ft-L; 10 Min/7.5 Sec	9°
Campbell Kulikowski Levinson (1966)	--	--	--	--
Shade (1956)	BINOCULAR FOVEAL EXTRA/FOVEAL	NATURAL	0.00044-400 Ft-L;"ADEQUATE"	14.3° X 19°

Table 3-1 (Part 2) Summary etc.

	SOURCE TYPE	SPATIAL FREQ. RANGE;TARGET SHAPE	METHOD;DURATION	OTHER
DePalma, Lowry (1962)	WHITE (2850°K)	0.22-43cycles/ degree; 6° SQUARE	THRESHOLD	SINE/SQUARE-WAVE DISTANCE EFFECTS
Bryngdahl (1964) (1965)	WHITE (2800° K)	0.5-8.5 cycles/ degree; 17.8° SQUARE	BRIGHTNESS MATCHING; 3MIN	SINE; VARIABLE CONTRAST EFFECTS
Patel (1966)	GREEN (510 mm) P31 PHOSPHOR	1.1-59 cycles/ degree; 2°	THESHOLD	SINE;LUMINANCE EFFECTS
Green, Campbell	GREEN	2-50cycles/ degree;2X1.3 in.	THRESHOLD	SINE; FOCUS, SPHERICAL ABBERATION EFFECTS
Daitch, Green (1969)	GREEN(510mm) P31 PHOSPHOR FILTERED (WRATTEN #65)	0.5-10 cycles/ degree	THRESHOLD; 200 MSEC	SINE;ROD ISOLATION BY BLEACHING
Davidson (1968)	WHITE QUARTZ-IODINE W/ TUNGSTEN FILAMENT	0.336-9.6cycles/ degree; 6°40' CIRC	CONTRAST MATCH. (2) 500 msec FLASHES 1 SEC APART	EXPONENTIAL SINUSOIDS VARIABLE CONTRAST; SMALL PERTURBATION LIMIT
Campbell, Green (1965)	LASER λ=632.8mμ	4-62cycles/ degree CIRC.	THRESHOLD; A "FEW" SEC	SINE; INTERFERENCE FRINGES; RETINA/BRAIN ONLY
	GREEN(pl) λ=530mμ	2-46cycles/deg. RECT. 2X1.3IN.	THRESHOLD	SINE;PUPIL DIA. AND FOCUS VARIED

Table 3-1 (Part 2) continued

	SOURCE TYPE	SPATIAL FREQ. RANGE;TARGET SHAPE	METHOD;DURATION	OTHER
Schober, Hilz (1965)	KRYPTON LAMPS	0.18-60 cycles/ degree 30X30CM	THRESHOLD(50%) 1.5-1 SEC-∞	SQUARE; DURATION, LUMINANCE, SIZE EFFECTS
Nachmias (1967)	WHITE	0.7,1.75,17.5 cycles/degree	THRESHOLD(71%) 11-500 MSEC	SQUARE; DURATION ORIENTATION EFFECTS
Campbell Kulikowski Levinson (1966)	--	--	--	ORIENTATION EFFECTS ON SINE GRATINGS; RETINA-BRAIN, TOTAL
Shade (1956)	--	0.6-59 cycles/ degree RECT. (4:3)	THRESHOLD	DRIFTING AND STATIONARY SINE; LUMINANCE VARIED

produces a distribution of retinal illumination which is itself also sinusoidal, (b) the ratio of the luminance distribution . amplitude and the illumination distribution amplitude may be assumed constant for fixed spatial frequncy and variable contrast. The variation of the ratio between luminance amplitude and illumination amplitude with spatial frequency is given in figures 3-5 and 3-6. These figures also show the dependence of response on pupil diameter.

Some measurements have been made of the contrast sensitivity of the retina-brain only (references 3-14, 3-43). Such measurements are made using a sinusoidal illuminance distribution produced directly on the retina using interference techniques. Two uniform beams of light are derived from a coherent monochromatic laser source. The two beams are optically introduced through the pupil so that they reach the retina out of phase and interfere to produce a sinusoidal illuminance distribution directly on the retina. By adjusting the phase and the luminance of the two beams, the contrast and spatial frequency of the interference pattern can be varied. The threshold method is then used to measure the contrast sensitivity of the retina-brain only. Figure 3-17 is taken from Campbell and Green (reference 3-14) and compares the contrast sensitivity for the entire eye against the contrast sensitivity measurements for the retina-brain only.

It is difficult to say whether or not the response evoked by a sinusoidal illuminance distribution retains its sinusoidal characteristics. Bryngdahl's brightness matching

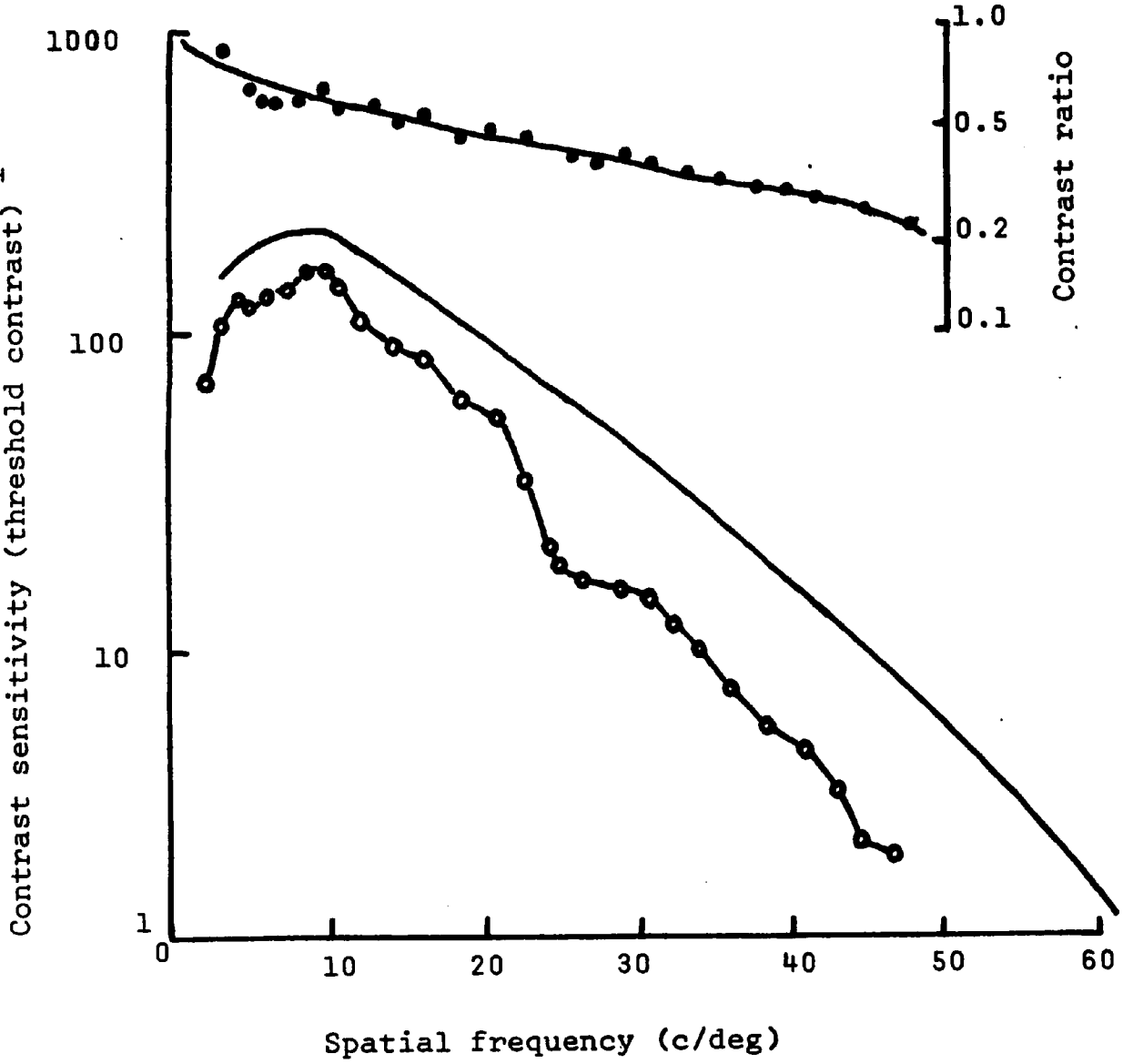


Figure 3-17. The contrast sensitivity of the composite visual pathway (open circles). Contrast sensitivity for interference fringes (smooth curve). The closed circles are the ratios of the contrast sensitivities and represent the contrast transfer function for the optics of the eye. A smooth curve has been drawn through the ratios. (reference 3-14)

results (reference 3-44) might have been useful in this regard if he had compared subjective and objective brightnesses at every point in the distribution. However, his measurements were only made at the maximum and minimum points in the distribution. He did find that the ratio of the contrast of the objective distribution to that of the subjective distribution was not a constant for a fixed value of spatial frequency. The ratio actually varied in a complex way as a function of objective contrast. His results raise doubts about the linearity of the retina-brain response.

3.4.2 Temporally varying brightness stimuli

Some of the effects described for spatially varying brightness stimuli have their counterparts when the luminance of the stimulus varies with time. For example, a field of uniform luminance for which the luminance varies sinusoidally as a function of time can be used to measure a contrast sensitivity function analogous to the one described for spatially varying distributions. As the frequency of the sinusoidal distribution is increased, a point is reached where the observer detects no change in luminance with time. This point is called the critical flicker frequency for the sinusoidal distribution. Critical flicker frequency is affected by the color, luminance, contrast, and field size of the sinusoidal stimulus. Eye movements might also

be considered under the heading of time varying stimuli. However, time varying effects will not be included in the models to be discussed and, therefore, no further discussion will be given here.

3.5 Summary

Some characteristics of the human visual system have been described. The characteristics described are reasonably separated into two categories:

(a) characteristics of the optical system of the eye,

(b) characteristics of the sensory system of the visual pathway.

The components of the eye that form the optical system are the cornea, the aqueous humour, the pupil, the lens, and the vitreous humour. A schematic model was given which described the optics of the eye in terms of an ideal, diffraction limited, optical system. The schematic model can be used to analyze accommodation and its effect on the image. Data given on pupil diameter as a function of luminance can be used in conjunction with the schematic model to estimate image quality. Experimental data was given which measured the quality of the optical system of the eye. It was shown that for certain pupil diameters the optics of the eye approaches the quality of a diffraction limited optical system. These results will be used in chapter 5 to model the optical system of the eye.

The sensory portion of the visual system begins at the retina and terminates at the brain. Functionally the sensory system begins with the photoreceptors in the retina. Some of the characteristics of the photoreceptors which were described were:

(a) receptor density as a function of retinal position,

(b) spectral sensitivity,

(c) absolute light detection threshold, and

(d) visual acuity.

Receptor density can be used to set a theoretical upper bound on acuity; it can be used to determine the relative importance of a particular type of receptor given the position of the image; it can also be used to determine whether or not the retina is spatially homogeneous.

Spectral sensitivity can be used to estimate the relative significance of experimental results using test stimuli of different spectral compositions.

The absolute light detection threshold makes it possible to determine where on the scale of luminance values a given test luminance falls.

Visual acuity as a function of retinal illuminance is relevant to the spatial sampling density of digitized pictures; it also gives an indication of the spatial organization of the photoreceptors.

The remaining characteristics described in chapter 3 were not associated with any particular physical component of the sensory system. These characteristics are based

on the subjective response of a human observer and therefore, are called psychophysical characteristics. Of particular concern is the relationship between a physical spatial luminance distribution and the resulting perceived subjective spatial brightness distribution. One of the most elementary psychophysical characteristics is incremental brightness sensitivity: the threshold luminance difference between two areas of constant luminance as a function of area geometry and spatial average luminance. These results are relevant to brightness quantization of digitized pictures. Results were described for the supra-threshold effect of adjacent areas of different luminance on perceived brightness. This effect, known as simultaneous contrast, is indicative of the complexity of the interaction of the photoreceptor responses.

The subjective response of the sensory system to sinusoidal spatial brightness distributions is of particular interest here. If the sensory system can be approximated by a linear, space invariant system, then the spatial sinusoid characteristics determine the modulation transfer function for the system.

REFERENCES

3-1 Peter C. Kronfeld, "The Gross Anatomy and Embryology of the Eye," The Eye, Vol. 1, Hugh Dawson, Ed. Academic Press, New York, 1962.

3-2 D.M. Maurice, "The Cornea and Sclera," in The Eye, ibid.

3-3 H. Dawson, "The Intra-Ocular Fluids," in The Eye, ibid.

3-4 Ruth van Heyningen, "The Lens," in The Eye, ibid.

3-5 Otto Lowenstein, Irene E. Loewenfeld, "The Pupil," in The Eye, ibid.

3-6 Antoinette Pirie, "The Vitreous Body," in The Eye, ibid.

3-7 Mathew Alpern, "Accomodation and the Pupil," in The Eye, Vol. 3, ibid.

3-8 Ragnar Granit, "The Visual Pathway," The Eye, Vol. 2, ibid.

3-9 A.G. Bennett, J.L. Francis, "Visual Optics," in The Eye, Vol. 4, ibid.

3-10 John W.T. Walsh, Photometry, Dover Publications, Inc., New York, 1958.

3-11 Joseph W. Goodman, Introduction to Fourier Optics, McGraw Hill Book Company, New York, 1968.

3-12 G. Westheimer, F.W. Campbell, "Light Distribtuion in the Image Formed by the Living Human Eye," JOSA, Vol. 52, no. 9, 1962, pp. 1040-1045.

3-13 J. Krauskopf, "Light Distribution in the Human Retinal Image," JOSA, vol. 52, no. 9, 1962, pp. 1046-1050.

3-14 F.W. Campbell, D.G. Green, "Optical and Retinal Factors Affecting Visual Resolution," J. Physiol., Vol. 181, 1965, pp 576-593.

3-15 F.W. Campbell, R.W. Gubisch, "Optical Quality of the Human Eye," J. Physiol., Vol. 186, 1966, pp. 558-578.

3-16 R.W. Gubisch, "Optical Performance of the Human Eye," JOSA, vol. 57, no. 3, March 1967, pp. 407-415.

3-17 M.H. Pirenne, "Rods and Cones," in The Eye, vol. 2, op. cit.

3-18 T.N. Cornsweet, Visual Perception, Academic Press, New York, 1970.

3-19 M.H. Pirenne, "Visual Acuity," in the Eye, vol. 2, op. cit.

3-20 M.H. Pirenne, "Spectral Luminous Efficiency of Radiation," in the Eye, op. cit.

3-21 G.S. Brindley, Physiology of the Retina and Visual Pathway, the Williams and Wilkins Co., Baltimore, 1970.

3-22 S. Hecht, S. Schlaer, M.H. Pirenne, "Energy, Quanta, and Vision," J. Gen. Physiol., vol. 25, 1942, pp. 819-840.

3-23 M. H. Pirenne, "Absolute Thresholds and Quantum Effects," in The Eye, vol. 2, op. cit.

3-24 M.H. Pirenne, "Visual Acuity," in The Eye, op. cit.

3-25 Van Nostrandt's Scientific Encyclopedia, Third Edition, D. Van Nostrand Co., Princeton, N.J. p. 859.

3-26 W.S. Stiles, "The Directional Sensitivity of the Retina and the Spectral Sensitivies of the Rods and Cones," Proc. of the Royal Society, vol.127B, 1939, pp. 64-105.

3-27 M.H. Pirenne, "Measurements of Stimulus," in the Eye, Vol. 2, op. cit.

3-28. C.H. Graham, R.H. Brown, and F.A. Mote, jr., "The Relation of Size of Stimulus and Intensity in the Human Eye: I. Intensity Thresholds for White Light," J. of Exp. Psych., Vol. 24, No. 6, June, 1939, pp 555-573.

3-29 C.H. Graham, N.R. Bartlett, "The Relation of Size of Stimulus and Intensity in the Human Eye,: III, the Influence of Area on Foveal Intensity Discrimination," J. Exp. Psych., Vol. 27, no. 2, August, 1940, pp. 149-159.

3-30 H.B. Barlow, "Temporal and Spatial Summation in Human Vision at Different Background Intensities," J. of Physiol., vol 141, 1958, pp. 3347-350.

3-31 H.Richard Blackwell, "Contrast Thresholds of the Human Eye, JOSA, vol. 36, no. 11, 1946, pp. 624-643.

3-32 W.A.H. Rushton, "Visual Adaptation," the Ferrier Lecturer, Proc. Royal Soc., Ser. B, Vol. 162, 1965, pp. 20-46.

3-33 F. Ratliff, Mach Bands: Quantitative Studies on Neural Networks in the Retina, Holden-Day, Inc., San Francisco, 1965.

3-34 J.A. Whiteside, M.L. Davidson, "Symmetrical Appearance of Bright and Dark Mach Bands from an Exponential Illumination Gradient," JOSA, Vol. 61, no. 7, 1971, pp. 958-961.

3-35 Edwin H. Land, John J. McCann, "Lightness and Retinex Theory," JOSA, Vol. 61, no. 1, 1971, ppl 1-11.

3-36 O.H. Shade, "Optical and Photoelectric Analog of the Eye," JOSA, Vol. 46, 1956, p. 721.

3-37 J.J. DePalma, E.M. Lowry, "Sine-Wave Response of the Visual System. II. Sine-Wave and Square-Wave Contrast Sensitivity," JOSA, Vol. 52, 1962, p. 328.

3-38 H.A.W. Schober, R. Hilz, "Contrast Sensitivity of the Human Eye for Square-Wave Gratings," JOSA, Vol. 55, no. 9, Sept 1965, pp. 1086-1091.

3-39 D.G. Green, F.W. Campbell, "Effect of Focus on the Visual Response to a Sinusoidally Modulated Spatial Stimulus," JOSA, Vol. 55, no. 9, 1965, pp. 1154-1157.

3-40 F.W. Campbell, J.J. Kulikowski, J. Lewinson, "The Effect of Orientation of the Visual Resolution of Gratings," J. Physiol., Vol. 187, 1966, pp. 427-436.

3-41 A.S. Patel, "Spatial Resolution by the Human Visual System, the Effects of Mean Retinal Luminance," JOSA, Vol. 56, 1966, p. 689.

3-42 J. Nachmias, "Effect of Exposure Duration of Visual Contrast Sensitivity with Square-Wave Gratings," JOSA, Vol. 57, no. 3, March 1967, pp. 421-427.

3-43 J.M. Daitch, D.G. Green, "Contrast Sensitivity of the Human Peripheral Retina," Vision Research, Vol. 9, pp. 947-952.

3-44 Olav Bryngdahl, "Characteristics of the Visual System: Psychophysical Measurements of the Response to Spatial Sine-Wave Stimuli in the Mesophic Region," JOSA, Vol. 54, 1964, p. 1152.

3-45 M. Davidson, "Perturbation Approach to Spatial Brightness Interaction in Human Vision," JOSA, Vol. 58, no. 9, pp. 1300-1308.

CHAPTER IV--PSYCHOPHYSICAL MODELS

4.0 Introduction

The purpose of this chapter is to describe some of the models which have been proposed to predict the psychophysical characteristics described in the preceding chapter. There are two sections: the first deals with a fundamental concept in the area of psychophysical modeling- the modulation transfer function--and the second describes a more general model - one that is capable of transforming a physical luminance distribution into a perceived brightness distribution. This model will be applied to the perception of pictures in the next chapter.

4.1 A model for the transformation of a luminance distribution

Fourier analysis is a familiar and useful tool in the analysis of linear time-invariant systems. In Fourier analysis, functions which satisfy certain mathematical constraints are expressable in terms of a sum of harmonically related sinusoidal components. A linear, time-invariant system is characterized by the property that a sinusoidal input will produce a sinusoidal output of the same frequency but possibly of different amplitude and phase. Also, by definition of a linear system, the resulting output is equal to the sum of outputs that would result from the individual inputs. Therefore, the response of a linear time-invariant system to an input (which is expressable in terms of a Fourier series) can be determined from the response of the linear time-invariant system to each of the individual

sinusoidal components of the input. The response of a linear time-invariant system to sinusoidal inputs for all values of frequency is known as the modulation transfer function of the system.

For each value of frequency, the modulation transfer function is the ratio between the phase and amplitude of the input and the phase and amplitude of the output. Since sinusoids of arbitrary phase and amplitude are most conveniently expressed in terms of the real part of complex numbers, the values of the modulation transfer function in general will be complex. However, in the special case where phase shifts do not occur, or are negligible, the modulation transfer function is a ratio of two real amplitudes and is therefore a real number.

Fourier analysis is also applicable in two dimensions. Equations 4-1 and 4-2 show the general form of the Fourier series for a periodic two-dimensional function.

$$f(x,y)=\sum_{0}^{\infty}\sum_{0}^{\infty} A_{nm}\cos 2\pi\left(\frac{nx}{L}+\frac{my}{M}\right)=B_{nm}\sin 2\pi\left(\frac{mx}{L}+\frac{my}{M}\right) \quad (4\text{-}1)$$

where, $$A_{nm}=\iint f(x,y)\cos 2\pi\left(\frac{nx}{L}+\frac{my}{M}\right)dxdy \quad (4\text{-}2)$$

$$B_{nm}=\iint f(x,y)\sin 2\pi\left(\frac{nx}{L}+\frac{my}{L}\right)dxdy \quad (4\text{-}2.1)$$

If any term of the series given by equation 4-1 is used as the input to a linear homogeneous isotropic system, the output will have the same form. The frequencies in the two dimensions will be identical to those of the input, and only the amplitude and phases might be changed.

Homogeneity is the two-dimensional analog of time invariance and may also be called shift-invariance. A system is isotropic if its response does not depend upon the orientation of the input. A pure rotation of the input should leave the output unchanged except for an identical rotation. In analogous fashion, the response of a linear, homogeneous, isotropic system to two-dimensional input expressable in terms of a Fourier series can be derived in terms of its response to the individual terms of that series. The response of a system to a two-dimensional sinusoids for all values of spatial frequency in the two dimensions is given by the spatial modulation transfer function of the system. Again, in general, the values of the MTF are complex, being the ratio between the complex input amplitude and the complex output amplitude.

Now we consider the application of Fourier analysis to human vision. In the case of vision, certain components of the eye would constitute the "system." It might be possible to model the eye in terms of two separate linear systems in cascade. The first of these two linear systems would represent the effect of the optics of the eye. This would include the effects of the cornea, aqueous humour, lens and vitreous humour. This is the portion of the eye that focuses the image of the object on the retina. The second linear system is intended to model the transformation performed by the retina in converting the illuminance distribution on the retina into a pattern of neural stimulation. Perceived brightness might then be deduced directly from the

distribution of neural stimulation. How linear, homogeneous, and isotropic these two regions of the eye might be, and their modulation transfer functions, will be discussed in the next two sections.

4.1.1 The modulation transfer function of the optics of the eye

The optical system of the eye consists of a single aperture, the pupil; several refracting surfaces, the cornea and the lens, and the media separating these surfaces, the aqueous and vitreous humour. Therefore, there is both diffraction and refraction involved in the optics of the eye. Diffraction occurs because of the limited pupil aperture and the limited extent of the refracting surfaces. Whether or not the optical system of the eye is linear, homogeneous, and isotropic depends on the processes of diffraction and refraction and whether they are linear.

Scalar diffraction theory can be used to estimate the linearity of the diffraction process in the eye. Two conditions are necessary in order for scalar diffraction theory to be applicable:

1) the diffracting aperture must be large compared with a wavelength, and
2) the diffracted fields must not be observed too close to the aperture (reference 4-3).

In the case of the eye, the important diffracting aperture will be the pupil. Typical pupil diameters range from two to six millimeters, whereas the wavelength of visible

radiation is on the order of 5×10^{-4} millimeters. Therefore the first condition is certainly satisfied. A typical retina to pupil distance is seventeen millimeters so the second condition is approximately satisfied.

The diffracting effects of the pupil can be approximated by a small aperture in an infinite, opaque plane. The Rayleigh-Sommerfeld formula for diffraction by an aperture in an infinite opaque plane is given by equation 4-3.

$$U(P_o)=\frac{1}{j\lambda}\iint U(P_1)\frac{e^{jkr}}{r}\cos(\bar{n},\bar{r})ds \qquad (4-3)$$

where $U(P_o)$ is the complex amplitude of the diffracted field at the point of observation. $U(P_1)$ is the complex amplitude of the incident field at a point in the aperture. $\bar{r}$ is a vector from P_0 to P_1, $\bar{n}$ is the unit normal to the aperture and λ is the wave length of the incident monochromatic radiation. $k=\frac{2\pi}{\lambda}$. The integral is only over the area of the aperture. Equation 4-3 requires that r be much greater than λ which is certainly the case.

Equation 4-3 is important because it indicates that the observed diffracted field is a linear function of the field incident on the aperture. That is, if $U(P_1)$ were a linear sum of component fields, then $U(P_0)$ would equal both the diffracted sum of the components and the sum of the diffracted components. For non-monochromatic fields, the observed diffracted field is related to the field incident on the aperture by equation 4-4, using the Rayleigh-Sommerfeld formulation.

$$u(P_o,t)=\iint \frac{\cos(\bar{n},\bar{r})}{2\pi cr}\ \frac{d}{dt}\ u(P_1,t-\frac{r}{c})ds \qquad (4\text{-}4)$$

According to Leibnitz's rule, if $u(P_1,t-r/c)$ is continuous and has a continuous derivative with respect to t over the aperture and for some range of t, then the derivative can be brought outside the integral. Now if u is a linear sum of continuous functions, then the integral of the sum is equal to the sum of the integrals. In addition, the derivative of each integral is equal to the integral of each derivative. Therefore, $u(P_0,t)$ is equal to a linear sum of the integrals. So even for non-monochromatic fields, the diffracted field is a linear function of the field incident on the aperture.

Equations 4-3 and 4-4 show that the diffracted field depends only upon values of the incident field in the plane of the aperture. Therefore, if the source field is shifted in space, in general, the value of the field in the aperture will change, and so the diffracted field will change. This would imply that the diffraction process is not strictly space invariant or homogeneous.

This is most easily illustrated by considering a point source in the object plane at P_0 emitting spherical waves of the form $U(P_0)=(A\exp(jkr_{01}))/r_{01}$ where r_{01} is the distance from P_0 to P_1, a point in the aperture. Now if P_0 is on the axis, and the aperture is small relative to r_{01}, then the phase term jkr_{01} will not vary a great deal across the aperture. Now move P_0 far away from the axis. In this case r_{01} could vary by several wave lengths from the bottom

of the aperture to the top of the aperture and therefore, the phase term could have considerable variation across the aperture. Accordingly, the diffraction pattern would be quite different in the two cases.

Analysis of the linearity of the refraction process can be simplified if certain reasonable assumptions are made. If it is possible to assume that the lens is thin, in the sense that a ray entering the front surface emerges from the rear surface at essentially the same coordinates, then such an ideal lens can be described in terms of a phase transformation. The phase transformation at any coordinate (x,y) is a function of the thickness t(x,y) of the lens at coordinates (x,y). The form of the phase transformation is given in equation 4-5 (reference 4-3).

$$\phi(x,y)=\varepsilon^{jkt_o}\,\varepsilon^{jk(n-1)t(x,y)} \quad (4-5)$$

where t_o is the axial thickness of the spherical refracting surface, $k=2\pi/\lambda$ and n is the index of refraction of the lens. The complex field in a plane immediately behind the lens $U_o(x,y)$ is related to the complex field incident on a plane immediately in front of the lens $U_i(x,y)$ by equation 4-6.

$$U_o(x,y)=\phi(x,y)U_i(x,y) \quad (4-6)$$

Since the $\phi(x,y)$ is merely a complex multiplicative constant, the transformation performed by the lens is linear. In the case of the human eye, the thin lens assumption seems reasonable, particularly for foveal vision. Essentially, only axial rays coming from an object at a comfortable distance are effective in focusing the image on the fovea.

Therefore, a ray entering the eye could undergo only slight deviations and still be focused on the fovea.

A thorough analysis of the eye, combining the effects of diffraction and refraction, would be difficult indeed. Several stages of diffraction and refraction would have to be cascaded in order to derive an equation for the eye which is the transformation of an object luminance distribution into a retinal illuminance distribution. An idealized model might consist of a real amplitude function in the object plane, representing transmissivity or reflectivity; a phase transformation in the plane of the cornea, limited by a pupil function representing the effective corneal aperture; a phase transformation in the plane of the lens, delimited by a pupil function representing the pupil of the eye; and the image plane at the position of the retina. A constant amplitude monochromatic plane wave might then be passed through the object plane in a direction perpendicular to all the planes in the model. The amplitude of the plane wave would then be multiplied by the transmissivity function for the object. The resulting plane wave would emerge from the image plane and propagate to the plane of the cornea.

Diffraction in the object plane could be ignored if the transmissivity function were defined for all values in the plane. At the plane of the cornea, the wave front undergoes a phase transformation in accordance with equation 4-6. The resultant wave front would be diffracted due to the finite corneal aperture and application of equation 4-3 would give

the resultant wave front in the plane of the lens. At the lens, another phase transformation would occur in accordance with equation 4-6. This result would then be used in equation 4-3 to give the wave front at the plane of the retina as diffracted by the pupil aperture.

Goodman (reference 4-3) has analyzed a much simpler model which nevertheless combines the effects of diffraction and refraction. This model is similar to the schematic eye with an equivalent lens (see section 3.2.1). His "object" is a monochromatic point source which produces spherical waves diverging about the point (x_o, y_o) in the object plane. At a distance d_o from the object plane is a "thin" double convex lens. The image plane is a distance d_i from the plane of the lens.

Several assumptions are contained in his derivation. It is assumed that only para-axial rays, rays at or near the axis of the optical system, are effective in producing the image. As stated above, this is certainly a reasonable assumption for foveal vision.

Goodman also assumes Fresnel diffraction. A sufficient but not necessary condition for Fresnel diffraction to be valid is given by equation 4-7.

$$z^3 >> \frac{\pi}{4\lambda} \left[(x_i - x_a)^2 + (y_i - y_a)^2\right]^2_{max} \qquad (4\text{-}7)$$

where z is the distance from aperture to image plane, λ is the wave length of the illumination (typically 5×10^{-4} millimeters), (x_i, y_i) is an image plane coordinate, and (x_a, y_a) is an aperture plane coordinate. A typical value

for z based on the principal points for the schematic eye given in Chapter 3 would be 20 millimeters. The term in brackets for a pupil diameter of 2.5 millimeters in good interior lighting would be about 36. Therefore, for the eye equation 4-7 gives 8000 56000 which is obviously false, therefore, so the condition is not satisfied. However, it is possible that the principle of stationary phase (reference 4-4) is applicable in this case and therefore, Fresnel diffraction might be valid. This principle is applicable when the oscillations of the phase factor in the diffraction integral are such that the primary contribution to the integral arises only from points near ($x_a=x_i$, $y_a=y_i$) where the rate of change of phase is minimum.

Goodman also assumes that the light detector in the image plane is sensitive only to light intensity and not phase.

A further assumption is that "the light amplitude at coordinates (x_i,y_i) must consist of contributions from only a tiny region of the object plane centered on the ideal geometrical object point."

Goodman's equation is for a particular image plane, namely the plane containing the infocus image as derived from geometrical objects. For this plane, d_o and d_i are related by the lens law discussed in chapter 3 and given in equation 4-8.

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f} \qquad (4\text{-}8) \qquad \frac{1}{f} \triangleq (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) \qquad (4\text{-}9)$$

where f is the focal length of the lens and is defined in equation 4-9. In equation 4-9, n is the index of refraction, and R_2 is the radius of curvature of the rear lens surface. For the conditions mentioned above Goodman's equation relating the object and the image is reproduced in equation 4-10.

$$h(x_i,y_i;\tilde{x}_o,\tilde{y}_o) = M\iint_{-\infty}^{\infty} P(\lambda d_i\tilde{x},\lambda d_i\tilde{y})e^{\{-j2\pi[(x_i-\tilde{x}_o)\tilde{x}+(y_i-\tilde{y}_o)\tilde{y}]\}}\,d\tilde{x}d\tilde{y} \qquad (4\text{-}10)$$

$$\tilde{x}=\frac{x_o}{\lambda d_i} \quad , \quad \tilde{y}=\frac{y_a}{\lambda d_i} \quad , \quad \tilde{x}_o=M\tilde{x}_o, \quad \tilde{y}_o=-M\tilde{y}_o \qquad (4\text{-}10)$$

where M, the magnification factor is d_i/d_o. P is the pupil function for the lens aperture and λ is the wave length of the monochromatic illumination.

Equation 4-10 is space invariant. That is, if a linear transformation is applied to the object plane coordinates, and the same linear transformation scaled by M is applied to the object plane coordinates, the result will be unchanged. The image produced by an object will be unchanged (except for position) if the object is repositioned within its plane near the axis. Since the eye has an equivalent power of about 60 diopters, an idealized equivalent model would contain a single, biconvex lens located at the principal point for rays entering the eye. Such a model is identical to the one analyzed by Goodman. The finite aperture of the equivalent lens is determined by projecting the smallest aperture in the eye, presumably the pupil, onto the plane of the equivalent lens (reference 4-3). Therefore, as a result

of Goodman's analysis, it is possible to say that an idealized model of the eye's optics is homogeneous.

The diffraction and refraction processes were shown to be linear using assumptions which are approximately satisfied by the optical system of the eye (e.g. paraxial, Fresnel diffraction, intensity detection). There optics is both linear and homogeneous.

In Chapter 3, it was pointed out that the eye's optics are not ideal in the sense that spherical and chromatic aberrations are present. However, experimental results were also shown that indicated that the eye's performance closely approximates that of the equivalent diffraction limited optical system. Therefore, as a first approximation, it does not seem unreasonable to assume the eye's optics to be linear and homogeneous.

The optics of the eye are not isotropic (references 3-18, 3-40). The contrast sensitivity of the eye is higher for vertical and horizontal sinusoidal gratings than it is for gratings with an oblique orientation. This orientational selectivity is both optical and neurological in origin.

A diffraction limited model of the eye's optics has been shown to be linear and homogeneous. Therefore, Fourier analysis can be used to analyse the transformation from object to image performed by the model.

If the modulation transfer function of the model is known, then the model is completely described. In Chapter 3, various experimental determinations of the modulation

transfer function for the optics were considered. It was determined that the results obtained by Campbell and Gubisch are the most appropriate for this research (reference 3-15). Their results are reproduced in Figures 4-1 and 4-2.

In order to properly interpret the results in Figure 4-2, it must be possible to determine the cutoff frequency of a diffraction limited optical system under the given conditions. Since the model is shift invariant, the results of Equation 4-10 will be unchanged if we restrict our attention to an impulse object positioned at the origin ($x_o=0$, $y_o=0$). The result is given in Equation 4-11

$$h(x_i,y_i)=M\int_{-\infty}^{\infty}\int P(\lambda d_i\tilde{x},\lambda d_i\tilde{y})\varepsilon^{-j2\pi(x_i\tilde{x}+y_i\tilde{y})}\,d\tilde{x}d\tilde{y} \quad (4\text{-}11)$$

$h(x_i,y_i)$ can now be seen to be the Fourier transform of the scaled pupil function. However, $h(x_i,y_i)$ is also the impulse response of the system. From linear system theory, it is well known that the Fourier transform of the impulse response is equal to the frequency response of the system. Therefore, the frequency response is given in equation 4-12.

$$H(f_x,f_y)= F\{F\{P(\lambda d_i\tilde{x},\lambda d_i\tilde{y})\}\}=P(-\lambda d_i f_x-\lambda d_i f_y) \quad (4\text{-}12)$$

Therefore, the pupil function appropriately scaled defines the frequency response of the system. The cutoff frequencies are those values which define the boundary of the pupil function in the frequency domain. For a circular pupil of diameter L, the cutoff frequencies are given by Equation 4-13 (reference 4-3).

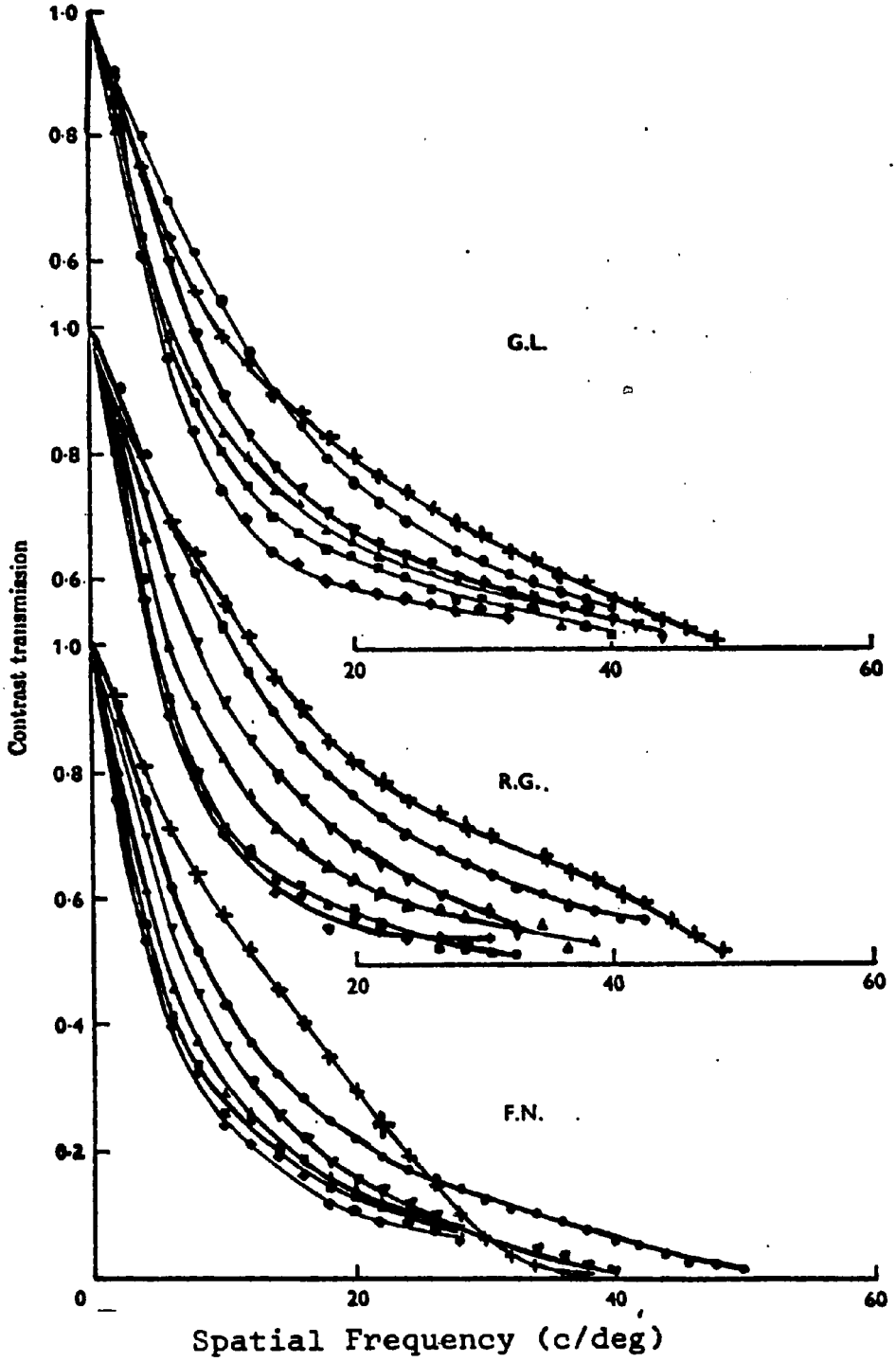


Figure 4-1. Modulation transfer functions for the eyes of three subjects. The curves display for various pupil sizes the transmission of contrast to an object imaged on the fundus; they are corrected for the double traverse of the eye made by the measured linespread function. Pupil diameters in mm are: +, 2.0; ●, 3.0; ▼, 3.8; ▲, 4.9; ■, 5.8; ◆, 6.6; (Reference 3-15).

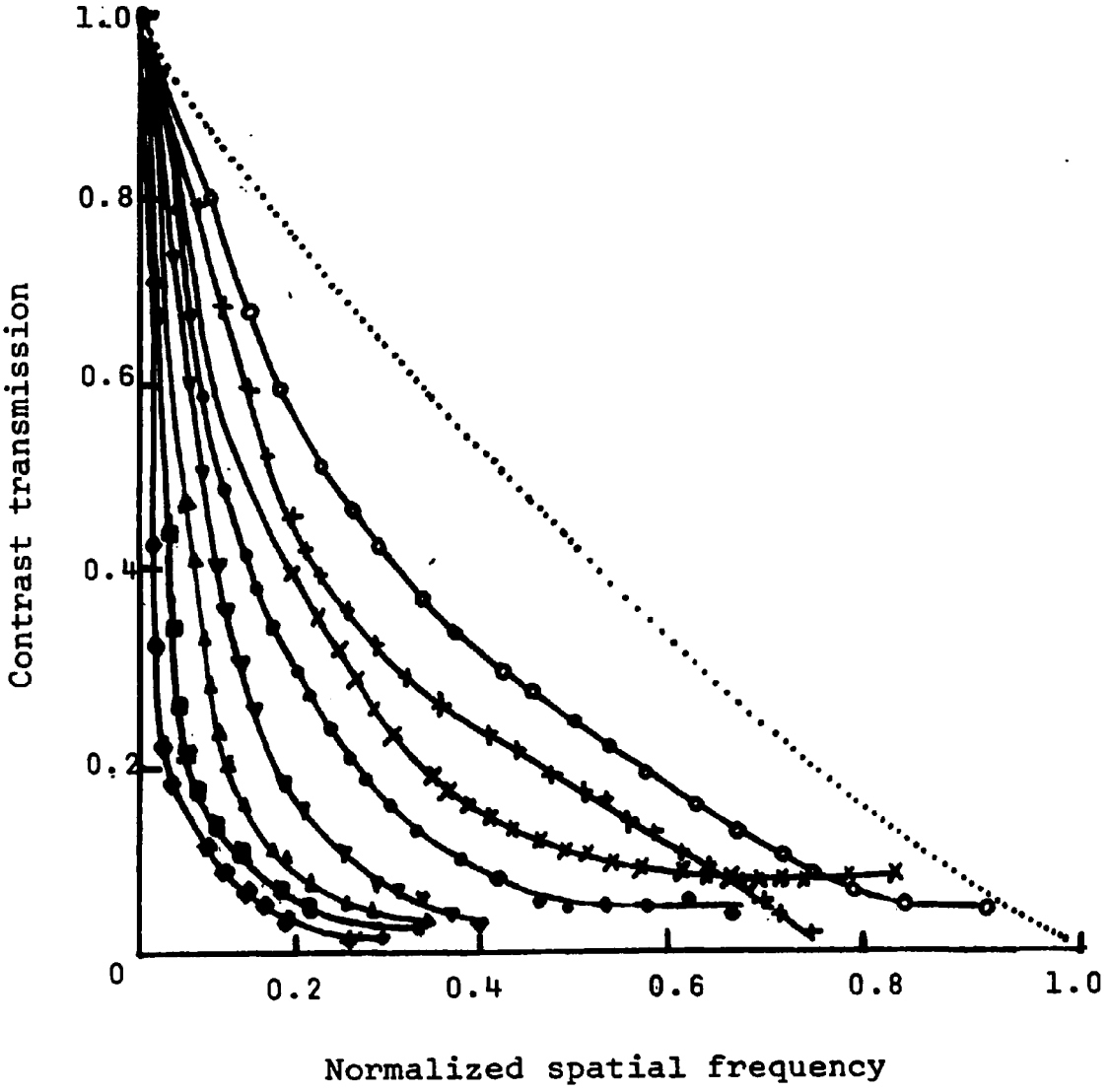


Figure 4-2. Normalized modulation transfer functions of the human eye, averaged among three subjects. The results are normalized to the highest spatial frequency transmitted by an ideal optical system with light of 570 mm wave length. Symbols are as in Figure 4-1 with: ●, 1.5 mm pupil; ✕ , 2.4 mm pupil. Dotted curve gives performance of ideal system. (Reference 3-15)

$$f_x^2 + f_y^2 = \left(\frac{L}{2\lambda d_i}\right)^2 \qquad (4\text{-}13)$$

$$[f_x]_{max} = [f_y]_{max} = \frac{L}{2\lambda d_i}$$

For the results given in Figure 4-2, λ equals 570×10^{-9} meters, d_i is the distance from the equivalent lens to the retina, about 20 millimeters. For a two millimeter pupil, equation 4-13 gives a cutoff frequency in the x and y directions of about 88 cycles per millimeter. Since each degree of visual angle corresponds to 0.291 millimeters on the retina, the cutoff frequency becomes 25.5 cycles per degree of visual angle. This result is for coherent illumination.

The results in Figure 4-1 extend beyond the coherent cutoff frequency because these results were obtained using incoherent illumination. For a circular aperture, the cutoff frequency of an incoherently illuminated optical system is twice that for the coherently illuminated system (reference 4-31). Therefore, for a two millimeter pupil and incoherent illumination, the cutoff frequency is 51 cycles per degree of visual angle.

Since the optics of the eye can be approximated by a linear homogeneous system with a modulation transfer function given in Figure 4-2, Fourier analysis can be used to determine the retinal illuminance distribution of an image from the luminance distribution in the object. The values used for the modulation transfer function will be determined by the diameter of the pupil, which in turn is determined by the average luminance of the environment.

4.1.2 Modulation transfer function of the retina - brains

The human visual pathway from retina to brain converts the physical illuminance distribution on the retina into a subjective, perceived brightness distribution. If this visual pathway can be characterized by a modulation transfer function, then it might be possible to compute perceived brightness from the retinal illuminance distribution. Whether or not the path from retina to brain can be characterized by a modulation transfer function depends upon the linearity and homogeneity of that path. The linearity and homogeneity of the optics of the eye were determined by combining the theoretical characteristics of the individual components. This same method cannot be used on the retina-brain because information about the individual components of this complex pathway is incomplete.

Ideally, it would be desirable to identify each of the functional components in the visual pathway from retina to brain, determine the inter-connections among these components, define the input-output characteristics of each components, and finally, derive the input-output characteristics of the entire system. The major components of the visual pathway from retina to brain are the retina itself, a portion of the brain described in general terms in Chapter 3, the optic nerve, the optic chiasma, and the optic tract.

The functional components of the retina are the photoreceptors and the various types of nerve cells. Figure 4-3 taken from Brindley (reference 4-5) shows the principal

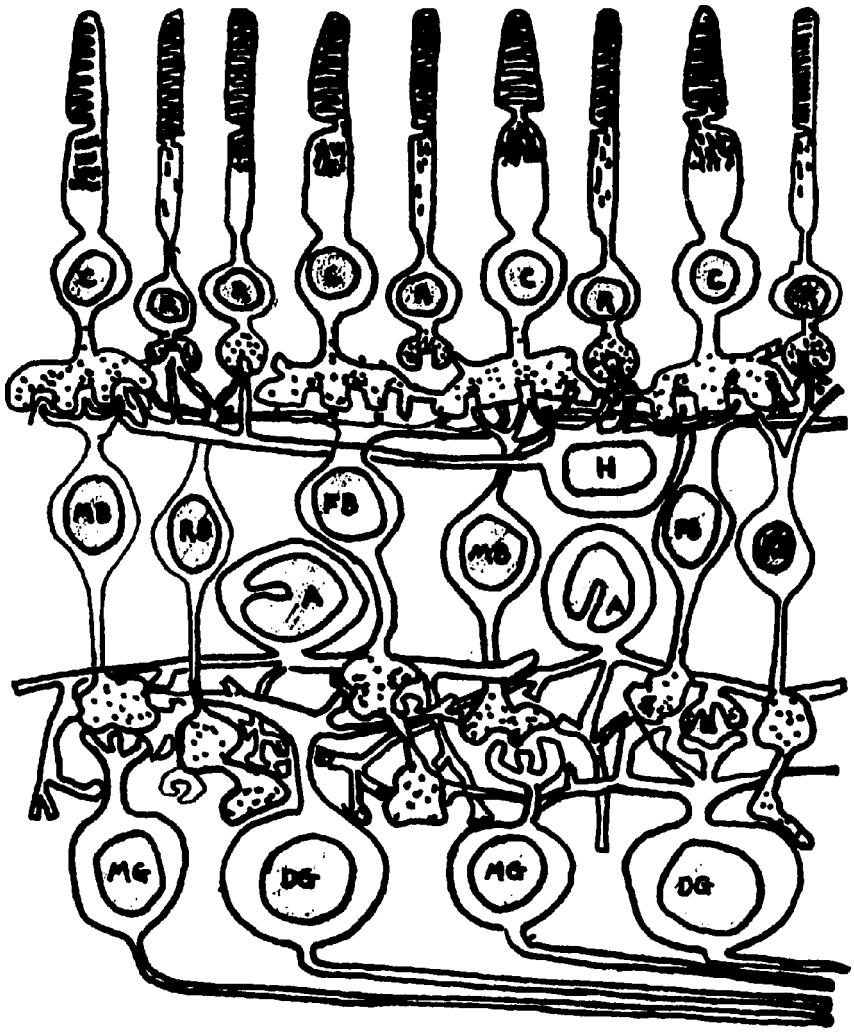


Figure 4-3. Diagram of the synaptic connexions of the primate retina (rhesus monkey and man). R, rod; C, cone; MB, midget bipolar, probably receiving only from one cone; FB, flat bipolar, receiving from several cones, but probably from no rods; RB, rod bipolar, receiving from several rods, but probably from no cones; H, horizontal cell; A, amacrine cell; MG, midget ganglion cell; DG, diffuse ganglion cell. (Reference 4-5)

functional components of the retina and illustrates recent ideas about how these components are interconnected. There is some uncertainty regarding the interconnection of the retinal components.

The terminal spherule of each rod is thought to be invaginated by the dendrites of rod bipolars (RB), horizontal cells (H), and perhaps two or three other types of cells.

The terminal pedicle of each cone (C) is invaginated by several sets of three dendrites. The middle dendrite of each set of three is thought to come from a midget bipolar (MB) and the outer two from different horizontal cells. Rods also make contact with horizontal cells, but it is uncertain whether the same horizontal cell is connected to both rods and cones. Cones also have connections with flat bipolar cells (FB), other cones, and with rod spherules.

Most bipolar cells form a large number of connections in the inner plexiform layer of the retina. Here they form connections which transmit from the bipolar cell to ganglion cells and amacrine cells (A). Another type of connection transmits from an amacrine cell to the bipolar cell.

A few bipolar cells which are thought to be rod bipolars have long axons that end in contact with the cell bodies of diffuse ganglion cells (DG).

The electrical response to light stimuli of the retina and some of its individual components has been measured. Most of these measurements have been made using non-human retinas which have a structure similar to that of the

human retina. It has been shown that all classes of vertebrate retinas are similar in a characteristic known as the electroretinogram (e.r.g.). The e.r.g. is a timewise variation of electrical potential between the cornea and the back of the eye due to light falling on the retina. An e.r.g. can also be obtained if the potential difference is measured between a microelectrode at the inner surface of the retina and a large electrode in contact with the back surface of the eye (sclera). The e.r.g. can be analyzed into three characteristic components: a cornea negative a-wave, a cornea positive b-wave, and a slowly varying c-wave. The bipolar cells are believed to be the largest contributors to the e.r.g. and in particular to its b-wave. The a-wave is believed to originate in the photoreceptors and the c-wave seems to be related to the pupillary reflex.

For light stimuli superimposed on a steady background, the amplitude and general shape of the e.r.g. is determined over a wide range by the ratio of the intensity of the incremental stimulus to that of the background. For the same ratio of intensities but different absolute values, the response is displaced in time, the displacement being shorter for larger absolute intensities (reference 4-5).

Some remarkable measurements have been made of electrical activity in individual components of the retina. For example, measurements have been made of circulating currents associated with rod photoreceptors. These currents pass through the rod membrane in the outer segment, flow axially,

and reenter the membrane as the inner segment is approached. The variation of these currents due to light stimuli and the associated potential differences resemble the a-wave component of the e.r.g. However, there is disagreement over whether or not this electrical activity in rods and cones is the signal that is transmitted to the rest of the retina. Therefore, even if measurements are available that relate stimulus intensity to electrical response, it is still not clear that that electrical response is the signal propagated to other parts of the retina. Measurements of the electrical response of the individual types of cells in the retina have also been made but here again it is not clear that these responses are the signals propagated to other parts of the retina. (reference 4-5, 4-6)

The response of ganglion cells, and consequently the optic nerve fibers to which they are connected, is characterized by a train of impulses. The onset of the stimulus will cause a burst of closely spaced impulses followed by a nearly steady discharge at a much lower frequency. At the termination of the stimulus, a brief burst of impulses precedes the cessation of response. Three types of ganglion cells contribute to this response: those which give a burst of impulses at the onset of the stimulus followed by a steady sustained discharge, those which only respond with bursts of impulses during the onset and termination of the stimulus, and those which only reppond to the termination of the stimulus. The relationship between impulse frequency and

stimulus intensity in general appears to be quite complex. For steady light, relatively few cells give a discharge whose frequency increases with increasing stimulus intensity over a wide range. (reference 4-5).

It is probably reasonable to assume that even less is known about the brain. With the realization that the above description of the retina and its components was extremely elementary, it should still be clear that a component by component analysis of the retina-brain for linearity and homogeneity is not possible.

4.1.2.1 Small perturbation characteristics of the retina-brain

Davidson (reference 4-7) studied the question of whether or not the retina-brain can be considered a linear, homogeneous, isotropic system. With respect to linearity, he states that the system in question may be assumed linear if it is recognized that linearity is "strictly valid only for infinitesimal perturbations of some fixed field." Let the illuminance distribution on the retina be given by:

$$g(u,v)=g_o(u,v)+g_p(u,v) \qquad (4\text{-}14)$$

and the corresponding subjective brightness distribution be given by:

$$h(x,y)=h_o(x,y)+h_p(x,y) \qquad (4\text{-}15)$$

where h_o is the brightness distribution resulting from g_o that is:

$$h_o(x,y)=f(x,y,g_o) \qquad (4\text{-}16)$$

$h_p(x,y)$ is the subjective brightness perturbation produced by the illuminance perturbation $g_p(u,v)$. If the first and second derivatives of f exist in every spatial location, then as the maximum absolute value of $g_p(u,v)$ approaches zero:

$$h_p(x,y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K(X,y,u,v)g_p(u,v)dudv \qquad (4\text{-}17)$$

Therefore, for small perturbations, the subjective brightness perturbation is a linear function of the illuminance perturbation. With respect to homogeneity, Davidson cites certain evidence against the retina-brain system being spatially homogeneous: the variation of visual acuity, spatial summation, receptor density and retinal cell density with retinal position, and topographical distortion of the visual field in the retina and visual cortex. However, he goes on to explain how these obstacles do not preclude the possibility of some type of homogeneity. The variation of spatial summation and visual acuity with retinal position can be overcome if a family of modulation transfer functions is assumed. Each transfer function would only be valid for a region of the retina over which the variation of spatial summation and acuity is negligible.

The variation of cone receptor density and consequently acuity is very rapid within the central one degree of the fovea. Therefore a single modulation transfer function would probably not be valid for more than ten or twenty minutes of visual angle. Davidson presupposed that the retina-brain consists of a spatial summation stage followed by a spatial

interaction stage. It is the homogeneity of the spatial interaction stage that is threatened by retina cell density variations and topographical distortions.

According to Davidson, although the spatial interaction system may be structurally inhomogeneous it might still be functionally homogeneous. One way this could happen is if the structural inhomogeneities, such as cell density, are described by a function of retinal position which contains only low spatial frequencies. Then it is only these same low spatial frequencies of illuminance that would be appreciably distorted due to inhomogeneity. But the spatial interaction system attenuates low spatial frequencies. Therefore, if the low frequencies which are most distorted are at the same time the low frequencies which are most attenuated then those spatial frequencies could be ignored and the spatial interaction system would be functionally homogeneous.

Topographical distortion in the spatial interaction system can cause each spatial frequency in the illuminance distribution to have different values of spatial frequency at different points within the spatial interaction system. However, if spatial frequencies are attenuated by the same relative amount at every point in the spatial interaction system, then the spatial interaction system is functionally homogeneous.

Davidson states that this will be true if the modulation transfer function has the following form:

$$F(f)=kf^{c} \qquad (4\text{-}18)$$

Let two spatial frequencies in the retinal illuminance distribution be f_1 and f_2. At some point in the spatial interaction system, these spatial frequencies become Af_1 and Af_2 where A is a function of retinal position. The relative attenuation of f_1 and f_2 is given by

$$\frac{k(Af_1)^c}{k(Af_2)^c} = \left(\frac{f_1}{f_2}\right) \qquad (4\text{-}19)$$

which is independent of the topographical distortion factor A and only depends upon the ratio of the two original spatial frequencies.

Davidson used the contrast matching method to determine the modulation transfer function for small perturbations. The resultant transfer function is for the complete visual system. If that transfer function is divided by the transfer function for the optics of the eye (Davidson used the results obtained by Krauskopf, and Westheimer and Campbell) then what results is the transfer function of the hypothetical spatial interaction system. For one of Davidson's subjects, the above procedure gave a transfer function which was proportional to spatial frequency, giving some support to the possibility that the spatial interaction system is functionally homogeneous. (see equation 4-18)

So, it would seem that at least in the case of small perturbations, it is possible to model the effects of the visual system by means of a modulation transfer function. In particular, it might be possible to analyze the subjective effect of errors introduced by a picture bandwidth compression

technique using the modulation transfer function for small perturbations. Most picture bandwidth compression techniques require that a trade-off be made between compression ratio and fidelity, where fidelity is a measure of the difference between the original picture and the compressed reproduced picture. Referring to equation 4-14, let g_o and g be the original and the bandwidth compressed illuminance distributions, repsectively. Therefore, g_p is the difference between these two illuminance distribution and for reasonable values of fidelity should have a small maximum absolute value. In equation 4-15, h_o would be the subjective brightness distribution corresponding to the illuminance distribution g_0. h is the subjective brightness distribution produced by the oompressed; reproduced picture. h_p is the subjective brightness perturbation produced by the illuminance perturbation g_p. Based on Davidson's analysis, the subjective brightness perturbation should be related to the illuminance perturbation by a modulation transfer function. The small perturbation modulation transfer function determined by Davidson is given in Figure 4-4. The upper curve is for the post receptor portion of the visual system, presumably the spatial interaction part of the visual system. The photoreceptors stand between the retinal illuminance distribution and the spatial interaction system. Davidson assumed that the photoreceptors perform a logarithmic transformation on the retinal illuminance distribution before presenting it to the spatial interaction system. However,

for small perturbations, the logarithmic transformation can be approximated by a linear transformation and therefore its effect is included in the modulation transfer function. Therefore, the upper curve in Figure 4-4 could be used to transform the illuminance perturbation into the subjective brightness perturbation.

4.1.2.2 Large signal characteristics of the retina-brain

There are three main objections to linearity of the retina-brain in the case of large signals: (a) The existence of contrast thresholds for spatial sine wave patterns, (b) evidence that spatial sine wave patterns do not remain sinusoidal in the subjective brightness domain, (c) some proposed models of the spatial interaction system are non-linear.

The existence of contrast thresholds implies that there is a non-linear stage somewhere in the retina-brain system. It will be shown later that it may be possible to associate these thresholds with a section of the visual system which is selectively sensitive to narrow bands of spatial frequency. Spatial frequency selectivity seems to be associated with specialized cells located in the brain. Therefore, the site of nonlinearity due to contrast thresholds may be near the termination of the visual pathway. This still leaves open the possibility of considering a linear stage between the photoreceptors and the site of spatial frequency selectivity. This is the part of the system containing the spatial summation and spatial interaction systems.

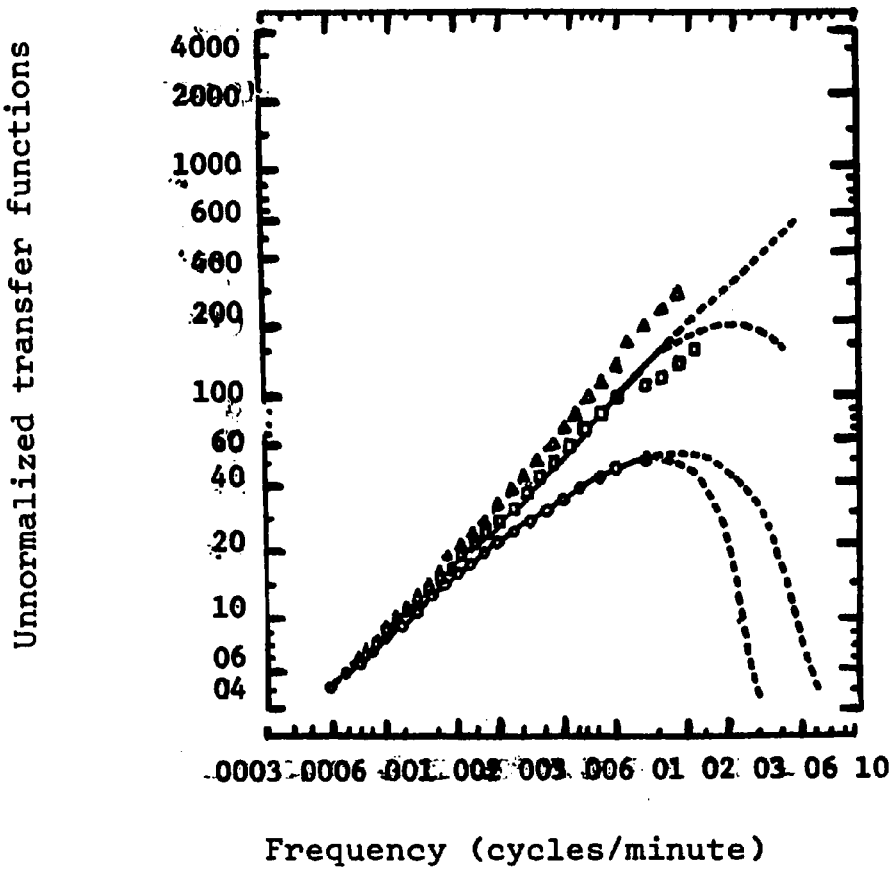


Figure 4-4. Visual system transfer functions. Lower curve (circles) is for complete visual system, upper curve for post-receptor system only. Latter obtained by using transfer functions for visual optics, only, from Westheimer and Campbell (triangles) and Krauskopf (squares). (reference 4-7)

The outcome of Bryngdahl's experiments (reference 4-8) might imply that the retina-brain system distorts sinusoidal stimuli. In the experiments, an observer views a square visual field divided into two adjacent half-fields. The lower half-field contains a one-dimensional spatial sine wave whose luminance varies in the horizontal direction. The upper half-field is of uniform luminance. The subject is asked to adjust the luminance of the upper half-field until it matches the maximum and minimum luminance respectively of the sinusoidal lower half-field. The objective contrast of the spatial sinusoid was determined by measuring its maximum and minimum luminances. The subjective contrast was based on the matching luminance values of the uniform field. The formula used for contrast is given in equation 4-20.

$$C = \frac{L_{max} - L_{min}}{L_{max} + L_{min}} \qquad (4\text{-}20)$$

With spatial frequency as a parameter, Bryngdahl plotted the ratio of subjective contrast to objective contrast as a function of objective contrast. For each spatial frequency, this ratio is found to vary in a systematic fashion, whereas for a linear system it would be expected that the ratio would be constant for each spatial frequency. However, if the "average value" of subjective brightness varied with objective contrast and the "amplitude" of subjective brightness remained constant, the contrast ratio would vary. Bryngdahl's results for the highest average objective luminance indicate a downward shift in average subjective

brightness with increasing objective contrast. This implies that the subjective contrast would increase faster than the objective contrast and therefore the contrast ratio would increase with increasing object contrast, an effect which is observed in Bryndgahl's results for the highest luminance.

Therefore, it is possible that a sinusoidal luminance distribution is transformed into a sinusoidal subjective brightness distribution but the average subjective brightness or DC response is a function of the objective contrast. Bryngdahl purposely used average luminance values in the mesopic range (0.25-2.0 candelas per square meter). In this range of luminances, the eye makes the transition from rod to cone vision. Therefore, his results for the maximum average luminance are most relevant here.

The role of spatial interaction in vision has received considerable attention (references 4-6, 4-9, 4-10, 4-11, 4-12, 4-13, 4-14, 4-15, 4-16). Spatial interaction is of particular interest in explaining a contrast phenomenon known as mach bands. Mach bands are perceived when there is an abrupt gradient between two regions of uniform luminance. What is seen is a darker band bordering the luminance gradient on the dark side of the boundary and a brighter band on the bright side of the boundary. Figure 4-5 shows a physical luminance profile and the corresponding perceived brightness profile for different values of the gradient. Some measurements have been made of the effect of the luminances of the two regions and the gradient on the perceived brightness

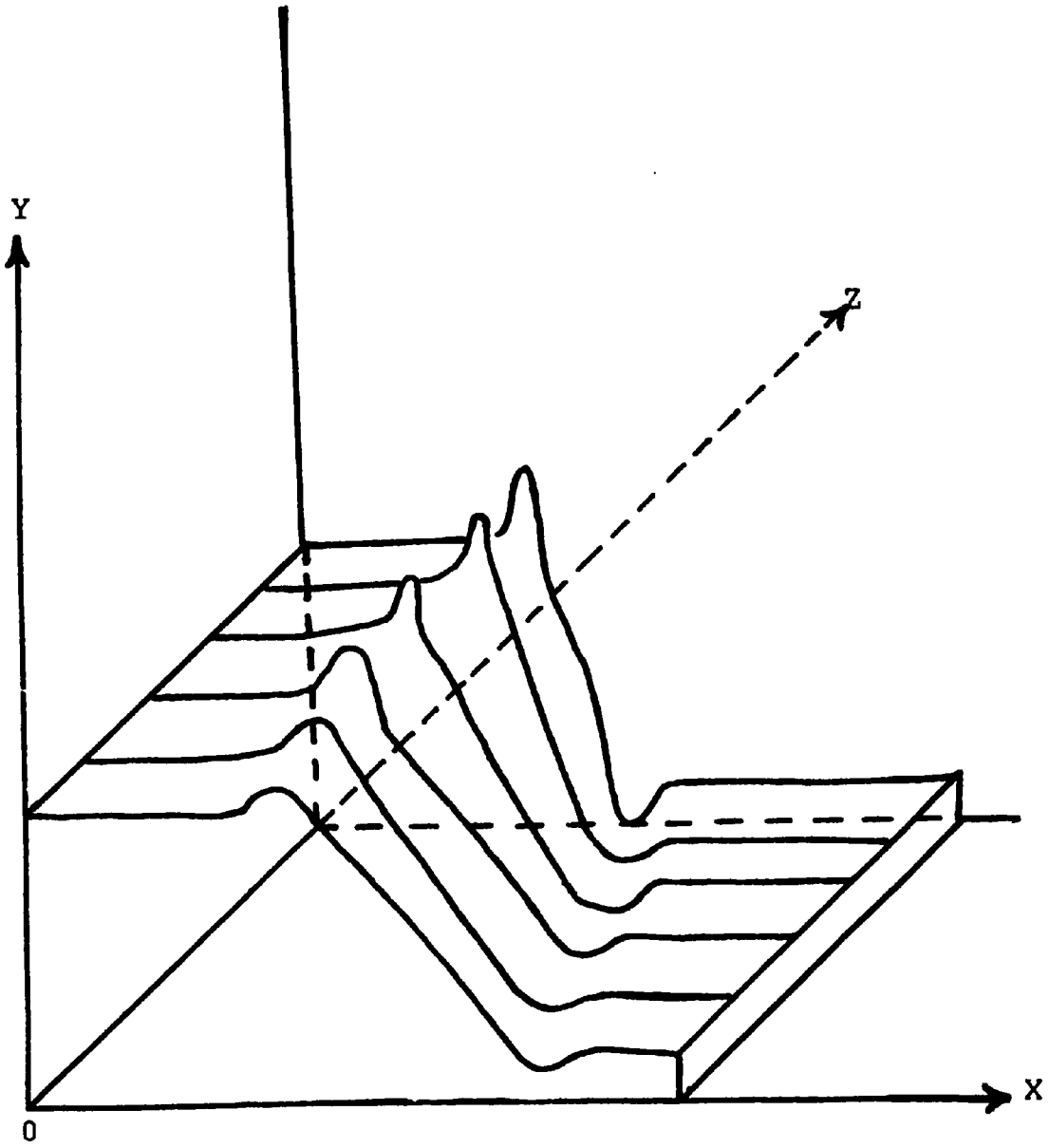


Figure 4-5. Mach bands and the influence of slope of gradient on the appearance of Mach bands. (Reference 4-1)

profiles (reference 4-9). It is apparent that the presence of the gradient has some effect on the perceived brightness some distance away from the gradient. Therefore there is some interaction between the sensory units at the gradient and those units some distance from the gradient. Figure 4-6 gives the spatial response characteristics for a common type of "sensory unit" that does produce a mach band-like response.

Several receptors contribute to the response of this sensory unit. The receptor in the center of the sensory unit makes a "excitatory" contribution to the response while the three receptors on each side of the center make an "inhibitory" contribution to the response. The contribution to the response of any given receptor is computed by multiplying the intensity of the stimulus at the location of the receptor times the inhibitory or excitatory coefficient. The total response of the sensory unit is equal to the sum of the responses of the individual receptors in its receptive field. The response given in Figure 4-6 is for a unit impulse. The response for such a sensory unit can be computed using equation 4-21.

$$r_i = \sum_j K_{i-j} r_j$$

where K_{i-j} is the excitatory or inhibitory coefficient of the receptor at position j relative to the receptor at position i. If the sensory unit shown in figure 4-6 is applied to the luminance gradient shown in figure 4-7 using equation 4-21, the response is given by the dashed curve in figure 4-7.

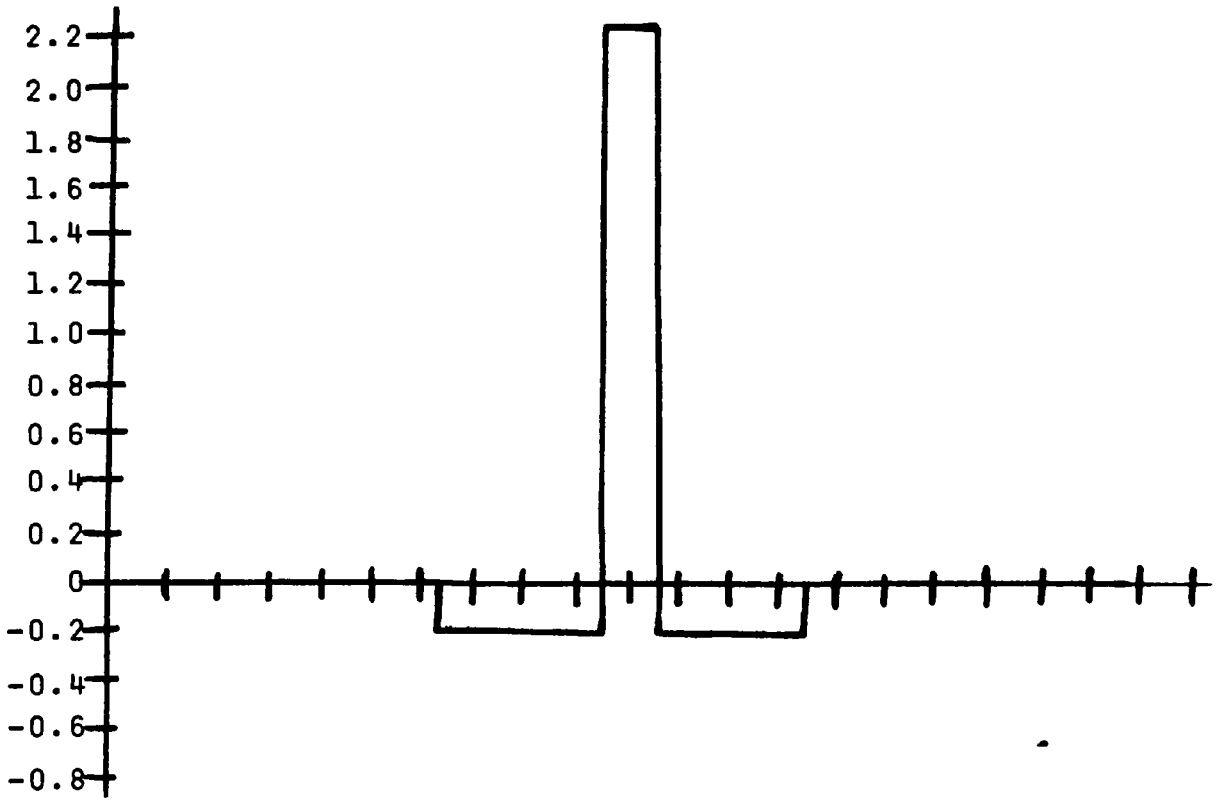


Figure 4-6 Sensory unit response due to spatial interaction (Horizontal axis is in units of inter-sensory unit spacing).

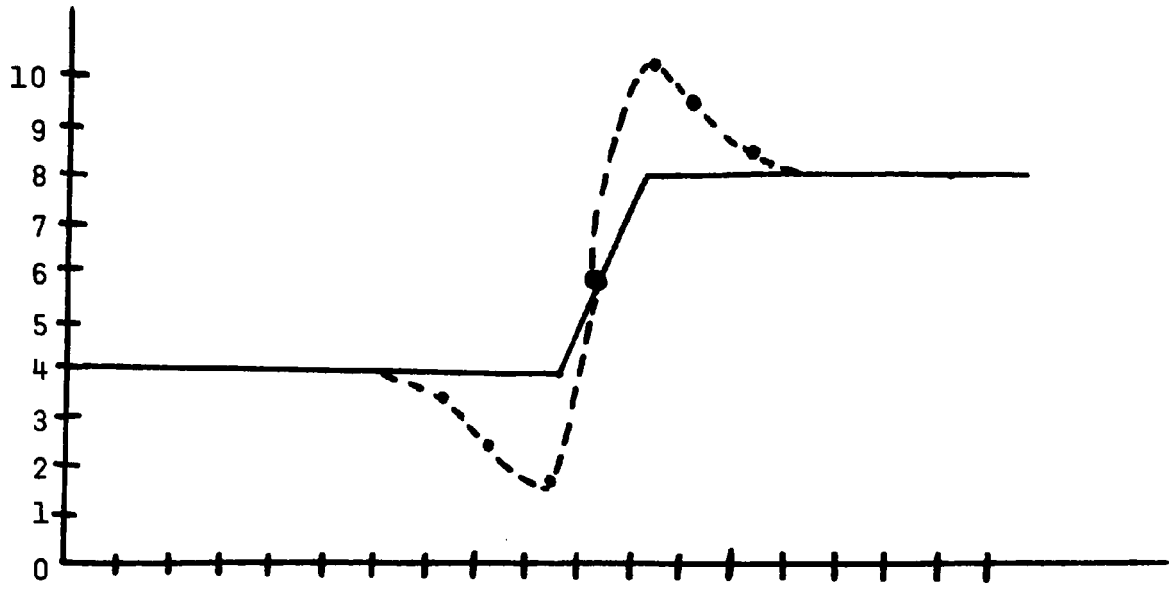


Figure 4-7. A luminance gradient and the associated response of the sensory unit given in figure (Horizontal axis is in units of inter-sensory unit spacing).

The simplified sensory unit just described is linear, however it is not an adequate model for the synthesis of Mach bands. Mach bands are asymmetrical whereas this sensory unit produces a symmetrical Mach band-like pattern. The asymmetry may be accounted for by the addition of thresholds associated with the inhibiting inputs from adjacent receptors (reference 4-9).

Many models have been proposed to explain spatial interaction. All of these models are similar in that they all make use of an excitation-inhibition function which has an excitatory center and an inhibitory surround. The things that separate these models are the spatial characteristics of the excitation-inhibition function, whether or not the interaction is recurrent, and whether or not the interaction is multiplicative or additive.

Figure 4-8, taken from reference 4-9, summarizes some of the models that have been proposed. Also illustrated there are the excitation-inhibition functions and the equations defining the response. In these models the excitation and/or inhibition function may be Gaussian, exponential, an impulse, or constant, but in every case the essential characteristics of the response are the same. Measurements made on the human eye of the extent of the inhibitory zone from the center of excitation give values ranging from ten to seventy-five minutes of visual angle depending upon the conditions. It has been shown that the existence of thresholds can be used to explain changes in the size of the inhibitory zone due to changes in stimulus magnitude.

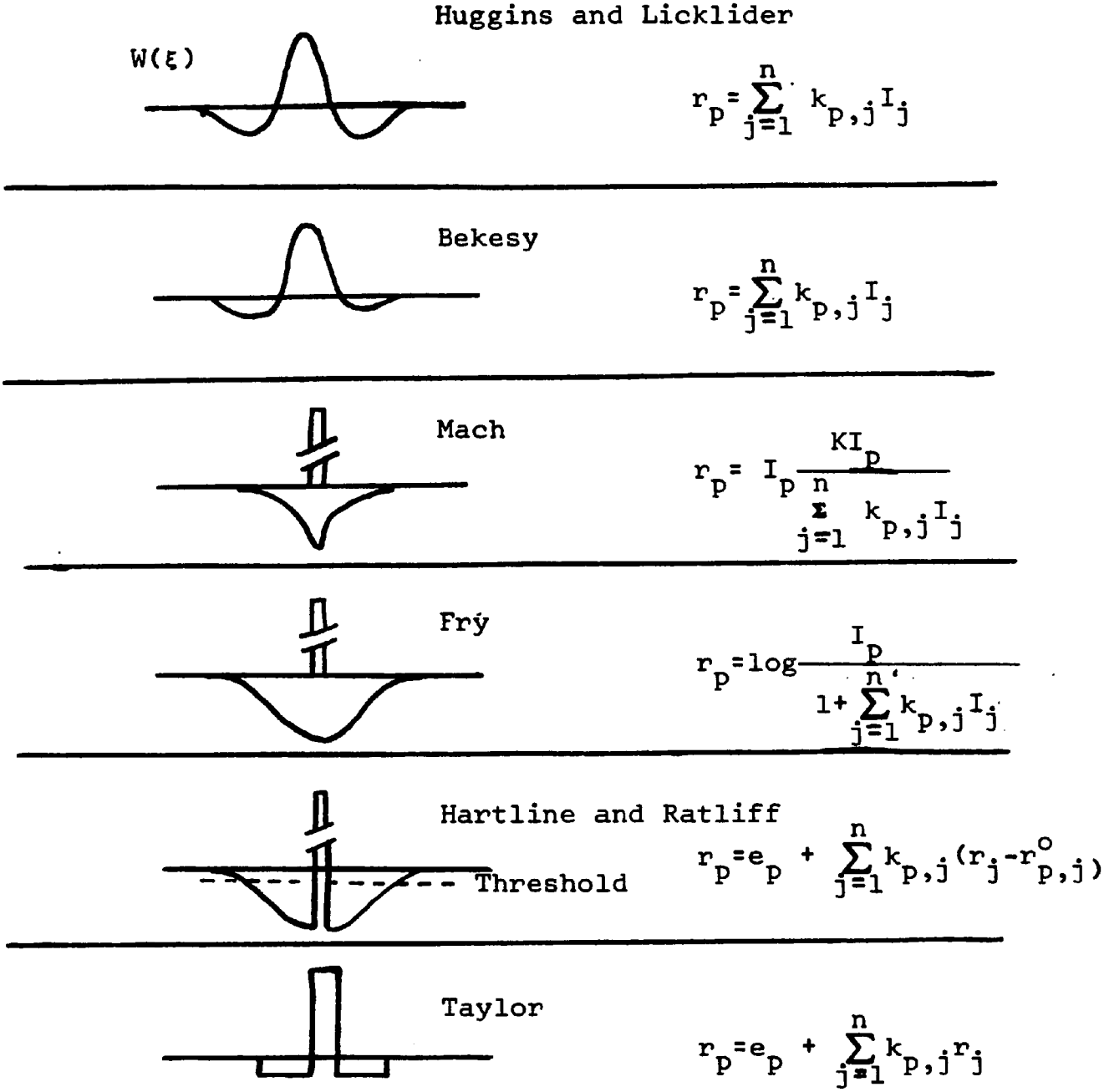


Figure 4-8 Graphical and mathematical representations of six models of inhibitory networks (reference 4-9)

A model is recurrent if the signal sent to a neighboring sensory unit is derived from a point beyond the site of spatial interaction with neighboring units rather than from a site before the point of spatial interaction. There is some evidence of recurrence in the human retina. It can be shown that a recurrent network can be approximated by several stages of nonrecurrent networks. It is also possible to approximate a recurrent network by a single stage, non-recurrent network with a $\frac{\text{sine } x}{x}$ excitation-inhibition function. The reason for this is that inhibition of a neighboring receptor disinhibits its effect on a further neighbor and the $\frac{\text{sine } x}{x}$ function, because of its alternating and decreasing excitation and inhibition has a similar effect. (reference 4-9)

If spatial interaction has a multiplicative effect, rather than an additive one, then it is a nonlinear process. Also the inclusion of threshold effects would make it non-linear. However, some of the models indicate that it may be possible to consider the spatial interaction system as piece-wise linear or at least approximately linear over a restricted range of stimuli (references 4-9, 4-17).

Lowry and DePalma (reference 4-18) investigated the linearity of the spatial interaction system in an interesting way. They were concerned about the interpretation given to sine wave contrast sensitivity measurements such as those described in Chapter 3. It is quite common to interpret reciprocal contrast sensitivity as a modulation transfer function. The underlying assumption seems to be that

contrast threshold is determined by a peak-to-peak detection device. Since the modulation transfer function is the ratio of image to object contrast at each sine wave frequency, it is a supra-threshold effect and not a threshold effect. Lowry and DePalma felt that in order to measure sine wave response correctly, it was necessary to make direct measurements of the response rather than measurements of some function of the stimulus. They used a test stimulus which was representative of normal viewing conditions in the sense that "the eye is operating in a region where both luminance and detail size are well above threshold." The stimulus was also chosen so that the eye operated over a narrow supra-threshold range of subjective brightness.

The stimulus used consisted of regions of uniform luminance with linear gradients, specifically chosen to produce Mach bands. The subjective brightness of the stimuli were measured using a slit photometer adjacent to the stimulus distribution. White light was used and the objective luminance of the gradient ranged between 3 and 20 foot-Lamberts. Foveal (cone) threshold illumination for white light stimuli 30 minutes of visual angle in diameter, expressed in terms of equivalent source intensity, is 0.00664 candelas per square meter or 0.00194 foot-Lamberts. The total dynamic range of the human eye covers about 11 log units. In foot-Lamberts this range runs from -6.5 log units to +4.5 log units. Since the cone threshold is at about -3 log units, the total dynamic range of the cones covers about 7.5 log units. with the upper limit .

being about 30,000 foot-Lamberts. Therefore, 3 to 20 foot-Lamberts is indeed a very narrow part of the supra-threshold range for foveal vision.

Lowry and DePalma began by assuming the visual system to be linear and homogeneous. They derived the modulation transfer function of the visual system by dividing the Fourier transform of the measured subjective brightness distribution by the Fourier transform of the measured objective brightness distribution. This procedure was repeated for several objective brightness distributions, all of which produced Mach bands. Some of the results they obtained are reproduced in Figure 4-9. These results for three different luminance edges agree quite well with one another except for spatial frequencies above 50 lines per millimeter, or 14.5 cycles per degree. However, it is at the high spatial frequencies that the Fourier transforms become very small in absolute value. Therefore, small absolute errors can produce large errors in the results obtained by dividing the two transforms. Lowry and DePalma used a high precision analog harmonic analyzer to perform the computations.

The quantitative similarity of the modulation transfer functions obtained for different luminance edges, as well as the qualitative similarity to modulation transfer functions obtained by other methods, might indicate that the modulation transfer function is a useful tool for the transformation of an objective luminance distribution into a subjective brightness distribution.

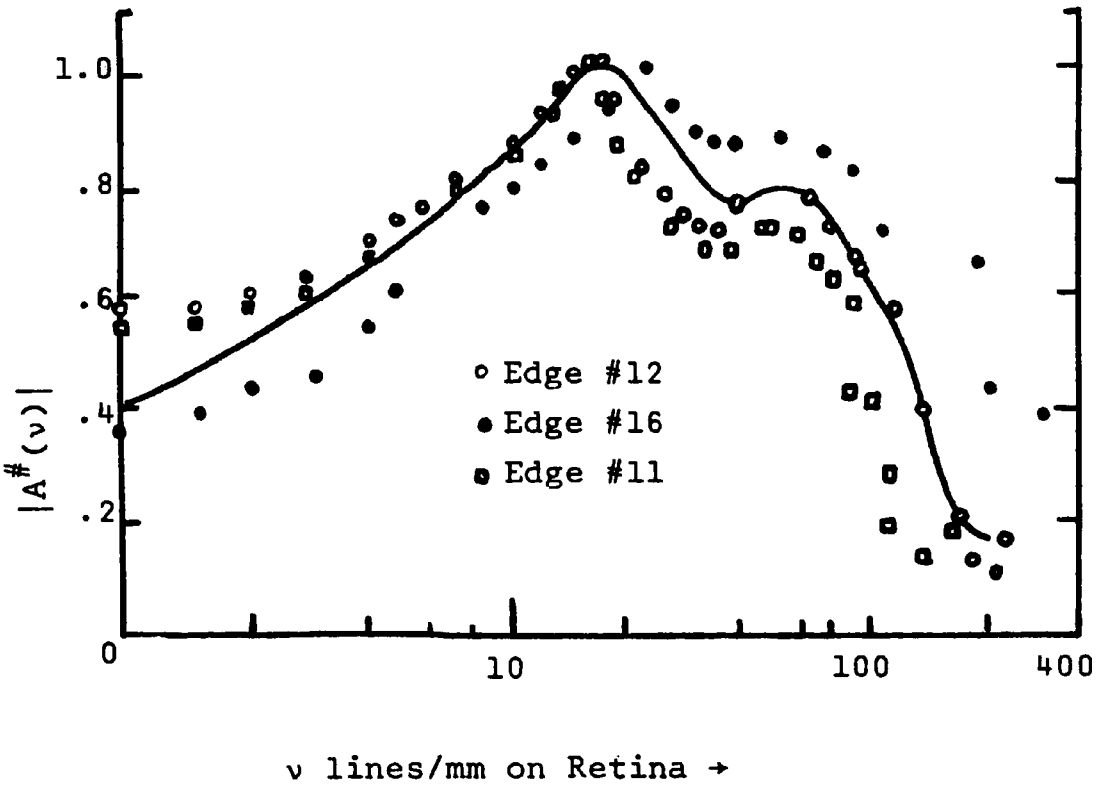


Figure 4-9. Sine wave response of the complete visual system. Derived from the Fourier transform of measured subjective and objective luminance edges.

4.1.3 A composite modulation transfer function for the human visual system

In Chapter 3, three methods were described for measuring the response of the human eye to sinusoidal luminance distributions. These were the threshold method, the contrast matching method, and the brightness matching method. In this chapter, the sinusoidal response of the eye for small perturbations was described. In addition, it was shown how the sinusoidal response might be derived from measurements of the sinusoidal brightness distributions. All of these methods give qualitatively similar results.

Invariably, the sinusoidal response is found to increase with increasing spatial frequency reaching a maximum response in the range of spatial frequencies from 1 to 20 cycles per degree. For further increases in spatial frequency, the response begins to fall off rapidly with the upper resolvable limit of spatial frequency being about 60 cycles per degree. The specific shape of the response is dependent upon pupil diameter (average luminance), lens accommodation (object distance), and the spectral content of the luminance distribution (chromatic aberration). However, for good interior lighting (30-300millilamberts), and normal viewing distances, the sinusoidal response follows the pattern described above.

If the human visual system can be modeled as a two stage linear system, then it is possible to synthesize a sinusoidal response characteristic which is in qualitative

agreement with experimental measurements. The first stage of this hypothetical linear system would represent the optics of the eye. It would be characterized by a point spread function measured for particular conditions of pupil diameter and lens accommodation. Such measurements have already been described (reference 3-15). Alternatively, the optical part of the system could be characterized by its modulation transfer function as shown in figures 4-1 and 4-2. For incoherent illumination the optical portion of the system will be linear in intensity. Therefore the modulation transfer function is the Fourier transform of the intensity of the point spread function (reference 4-3).

The second component in this hypothetical linear system would represent the spatial interaction system. The impulse response, or point spread function of the spatial interaction system was characterized by a central excitatory region and a surrounding inhibitory region. If this characteristic is idealized and represented by a central excitatory impulse and a Gaussian inhibition function centered on that impulse, then the modulation transfer function can easily be computed. It is the sum of the Fourier transform of the impulse and the Fourier transform of the Gaussian inhibition function. The Fourier transform of the impulse is a constant, and the Fourier transform of a Gaussian function is also Gaussian. These concepts, as well as the resulting modulation transfer function, are described in Figure 4-10.

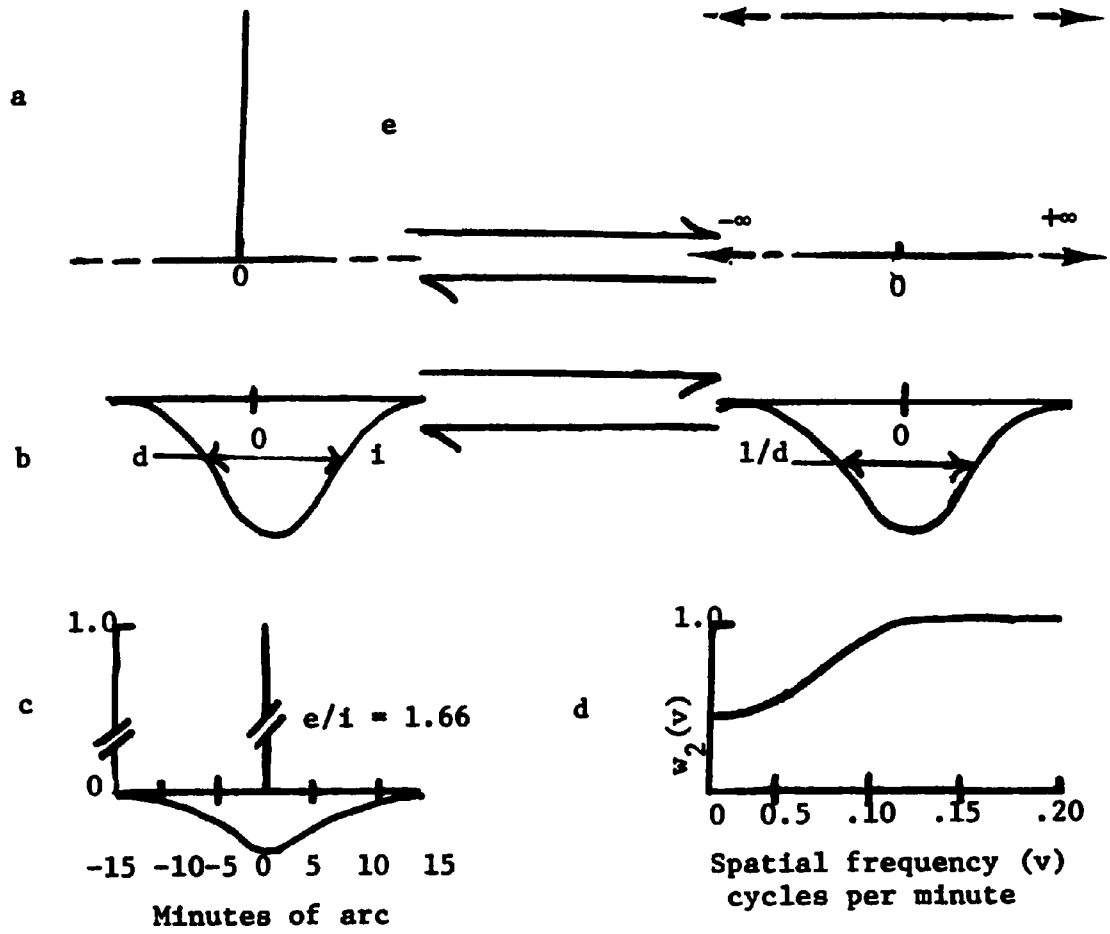


Figure 4-10. (a) The Fourier transform of a delta function. The height of the continuous spectrum of spatial frequencies in the transform on the right is proportional to the area of the delta function (unity) represented by the line on the left. (b) The Fourier transform of a Gaussian distribution with a half-width d is another Gaussian with a half width 1/d. (c) An ideal representation of an inhibitory network consisting of a delta function (excitatory) and a negative Gaussian function (inhibitory). The ratio of the areas of the two functions e/k=1.66. The half-width d of the inhibitory distribution is 10 minutes of visual angle. (d) The transform for(c) is the sum of the Fourier transforms of the excitatory component (positive delta function) and of the inhibitory component (negative Gaussian distribution) from 0 to +∞. Negative spatial frequencies are omitted since they have no physiological meaning in the present context. (reference 4-9)

The overall modulation transfer function of this two stage linear system can be found by multiplying the Fourier transforms of the two stages. In Figure 4-11, the impulse response of the optical system is shown as W_1. Its corresponding modulation transfer function is given by w_1. W_2 represents the impulse response of the spatial interaction system, and w_2 its Fourier transform. The impulse response of the overall system, W, is found by convolving the impulse responses for the individual stages. The overall modulation transfer function, w, is given by the product of w_1 and w_2 (reference 4-9).

This synthesized modulation transfer function agrees quite well qualitatively with the measured sinusoidal response characteristics of the human eye. In fact, if some reasonable assumptions are made about the half-width of the inhibition function, and if the two component modulation transfer functions are normalized to unity, the resulting composite modulation transfer function is in good quantitative agreement with observed results. For example, the maximum response is located at about five cycles per degree (see figure 4-11), when the half-width, d, in figure 4-10 is ten minutes of visual angle. Also, the response falls off to zero at about thirty cycles per degree which compares well with observed results. Normalizing the component modulation transfer functions to unity is probably not the correct thing to do. It is believed that the spatial interaction system compensates for imperfections in the optics of the eye as shown in figure 4-11. The spatial frequency response

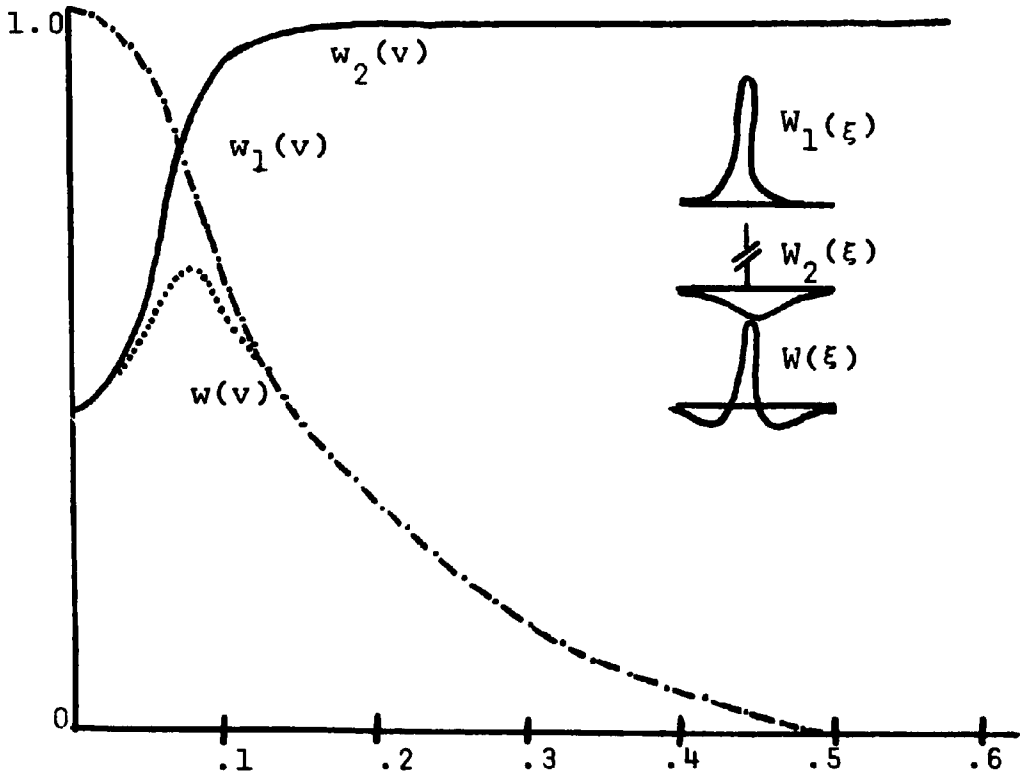


Figure 4-11. An idealized composite transfer function (dotted line) for the whole visual system. The composite function is the product of the transfer function for the dioptric apparatus (broken line) and the transfer function for the inhibitory network (solid line). The corresponding spread functions (or weighting functions) are shown at the right of the graph. (reference 4-9)

of the spatial interaction system increases with increasing frequency whereas the response of the optics decreases with increasing frequency (reference 4-9). In addition, Campbell and Greene's measurements of the sinusoidal response for the non-optical part of the visual system, as shown in figure 3018, imply that the spatial interaction system may have a relative gain greater than unity (reference 3-14). If the composite sinusoidal response curve in figure 3017 is normalized to unity, then the sinusoidal response curve for the non-optical part of the visual system will have values above unity. Another thing of interest in figure 3-17 is that the sinusoidal response characteristics for the non-optical portion of the visual system does not behave like the modulation transfer function for the spatial interaction system shown in figure 4011. The decreasing high frequency response shown in figure 3-17 may be due to the effects of interreceptor spacing and spatial summation, and a non-ideal excitation function, effects which are not taken into account in figure 4-11.

Spatial summation would cause several receptors in a given area to pool their signals to produce a response which constitutes a sample of the retinal illuminance distribution. This changes the sampling process from an ideal Dirac comb to a sequence of weighting functions. The spectrum of the spatial summation output is found by convolving the Fourier transform of the sampling process with the fransform of the input illuminance distribution. The result is a broader

spectrum with a lower high frequency response. If the excitation function is a narrow Gaussian function instead of an impulse, then its contribution to the spectrum in figure 4-11 will also be Gaussian. Since the excitation function is narrower than the inhibition function its spectrum will be broader than that of the inhibition function. The net spectrum will be the difference between a broad positive Gaussian function and a narrower negative Gaussian function. The resulting spectrum will have a response which qualitatively agrees with the measured spatial frequency characteristics described in Chapter 3. Suitable choice of the Gaussian parameters might make quantitative agreement possible.

4.2 Spatial frequency channels in human vision

In section 4.1, idealized linear models were proposed for certain components of the human visual system. It was shown that the optics of the eye can be described in terms of a modulation transfer function. A case was made for the existence of a spatial interaction network in the retina, and it was shown that this type of behavior can be approximated by a linear model with a modulation transfer function. However, it is not possible to represent the behavior of the retina in terms of a spatial interaction network of identical elements. There is evidence that the excitatory and inhibitory receptive fields of certain cells in the retina and brain vary with the type of cell and the viewing conditions. A model is described in this section which has some of the behavioral characteristics of the retina-brain portion of the human visual system.

There are currently two hypotheses which have been proposed to explain the threshold detection of visual stimuli. These are referred to as the single channel hypothesis and the multiple channel hypothesis.

The single channel model consists of a transforming device, which performs the transformation from a physical illuminance distribution to a neural stimulus distribution, and a detection device. The detection device is sensitive to spatial differences in peak to trough neural stimulation. It is not clear whether the detection device is actually composed of several detection devices each sensitive to local peak-to-trough differences or a single, wide range detection device. Which of these two detection devices should be hypothesized depends upon whether (or not) parts of a visual stimulus can be detectable or if the whole stimulus is detectable (or not). Most of the test stimuli used to investigate the single channel hypothesis are periodic and are therefore not useful in determining the nature of the detection device. The single channel model is shown schematically in figure 4-12.

The multiple channel model consists of the components shown in figure 4-13. In this model, the transformation from a physical illuminance distribution to a neural stimulus distribution is followed by several parallel channels. Each channel is selectively sensitive to a narrow band of spatial frequencies. Each band is assumed to be centered on a unique frequency and to overlap to some extent with neighboring

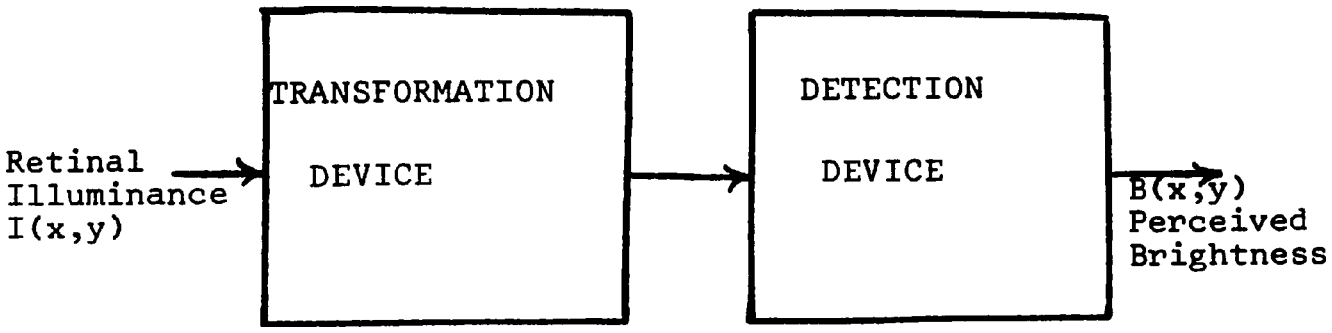


Figure 4-12. Single-channel model

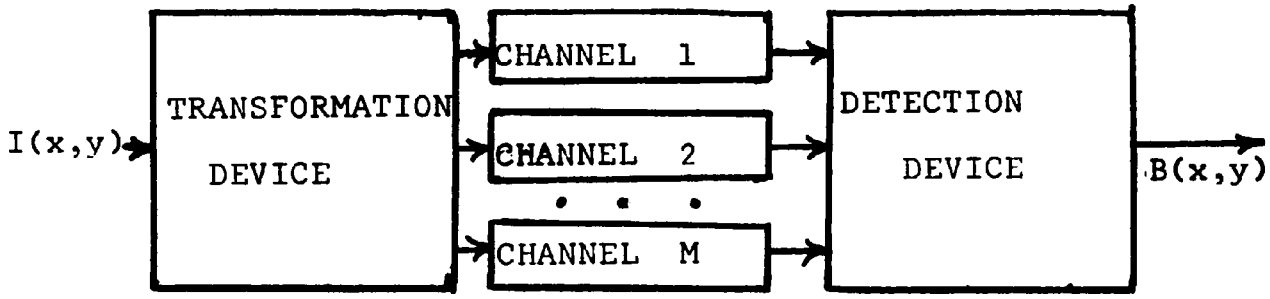


Figure 4-13. Multiple-channel model

bands. At the output of each channel is a detection mechanism. These detection mechanisms either independently signal the presence of spatial frequencies in a given band or are pooled to signal the detection of the entire pattern. Carrying this hypothesis a step further: we assume that the output from only those channels which are detected (and therefore exceed threshold) are recombined to produce a subjective perceived brightness distribution.

4.2.1 Comparative investigations of the single channel and multiple channel models

Several investigators have tried to devise psychophysical experiments to determine which of these two models best describes the observed behavior of the human visual system. A review of these investigations indicates considerable support for the multiple channel hypothesis.

4.2.1.1 Graham and Nachmias (reference 4-19)

This experiment investigates the two models by comparing the detectability of pure sinusoidal gratings with the detectability of complex gratings composed of the same pure sinusoids. The complex gratings used are formed by adding two sinusoidal components of spatial frequency f and 3f. These spatial frequencies were chosen to be harmonically related so that in-phase and out-of-phase summation would give a large difference in the peak-to-trough difference. In addition the frequencies were chosen to be far enough apart so that they would lie in different spatial frequency

channels. The experimental procedure consists of measuring the threshold contrast of the two sinusoidal gratings of spatial frequency f and 3f. Call these threshold contrasts C_f and D_{3f}. Then the detection threshold is determined for a complex grating formed by adding the two sinusoidal gratings. This is done for both in-phase and out-of-phase addition. The threshold contrast for the complex grating is determined by choosing a subthreshold value of contrast for one of the component sinusoids and then determining the contrast value for the other sinusoid which brings the complex grating to its detection threshold. Many pairs of contrast values C_f and C_{3f} are obtained.

Assume the single channel model is linear. If the single channel model is correct the following predictions can be made.

1) A pattern containing both components will be detectable even though the contrast in each component is substantially below its threshold value when presented alone, and
2) The phase in which the two components are added markedly affects the detectability of the combination.

These assumptions can be expressed in the form of an equation which can be used to predict the results of the detection threshold measurements. Let the peak-to-trough detection threshold be t. At the detection threshold of the complex grating, let the illuminance amplitude of the two sinusoids by L_f and L_{3f}. Let L_o be the average luminance. Assume that the transformation mechanism is a modulation transfer

function with values a_f and a_{3f} at the spatial frequencies of the two component sinusoids. For the case of in-phase addition, the threshold conditions are given by equation 4-22.

$$2(a_f L_f + a_{3f} L_{3f}) = t \qquad (4\text{-}22)$$

Using the following definitions for contrast and threshold contrast of the component sinusoids,

$$\bar{C}_f = \frac{t}{2a_f L_o} \qquad C_f = \frac{L_f}{L_o}$$

$$\bar{C}_{3f} = \frac{t}{2a_{3f} L_o} \qquad C_{3f} = \frac{L_{3f}}{L_o}$$

equation 4-23 results.

$$\frac{C_f}{\bar{C}_{3f}} + \frac{C_{3f}}{\bar{C}_{3f}} = 1 \qquad (4\text{-}23)$$

The measured contrast values should behave in accordance with equation 4-23 for in-phase addition of the sinusoidal components, if the visual system behaves in accordance with the single channel model. An analogous equation can be derived for out-of-phase addition of the sinusoidal components.

If the visual system behaves in accordance with the multiple channel hypothesis, and if the two component sinusoids are assumed to reside in different spatial frequency channels, then the complex grating will be detected if either or both of the component sinusoids are detected, or in terms of contrast and contrast thresholds,

$$C_f = \bar{C}_f \text{ if } C_{3f} < \bar{C}_{3f}$$
$$C_{3f} = \bar{C}_{3f} \text{ if } C_f < \bar{C}_f \qquad (4\text{-}24)$$

Two assumptions have been made here:

1) that the complex grating "is detected when at least one of the channels contains a peak-to-trough difference, at least as large as a threshold value."
2) that the noise is perfectly correlated in the two spatial frequency channels.

The results of the experiment are reproduced in figure 4-14. The dashed line represents the in-phase condition given in equation 4-23 for the single channel model. The dotted curve is for the out-of-phase condition for the same model. The solid lines represent equations 4-24 for the multiple channel model. The data points are clearly clustered around the solid lines, and therefore, strongly favor the multiple channel hypothesis. The authors cite a similar conclusion from statistical analysis of the data. Also, there was no systematic difference between the in-phase and out-of-phase results. In fact, gratings differing only in phase were equally detectable. In addition, moderately uncorrelated noise would not have adversely affected the results.

The range used for spatial frequency f was from 0.9 to 6.3 cycles per degree. The mean luminance used was 3 foot-Lamberts.

4.2.1.2 Campbell and Robson (reference 4-20)

In the first of several experiments, contrast sensitivity for a square wave grating was compared against the contrast sensitivity for a sinusoidal grating with the same period.

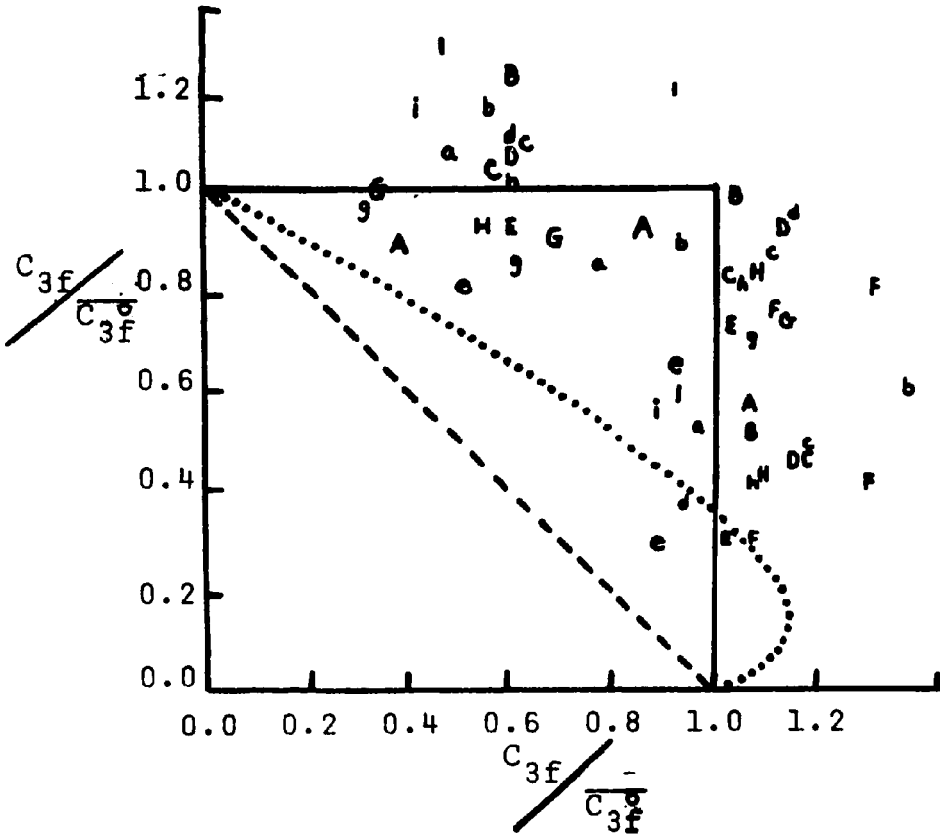


Figure 4-14. Contrast threshold for complex grating patterns containing two frequencies, f and 3f, in two different phases. The coordinate are the contrast of each component in the complex grating at threshold relative to the threshold contrast of the corresponding simple grating. Results obtained with the peaks-add form of complex gratings and with the peaks-subtract form are plotted as capital and small letters, respectively. The corresponding predictions of the single-channel model are represented by the diagonal dashed line and the dotted curve. The upper and right edges of the square represent the predictions of the multiple-channels model for both peaks-add and peaks-subtract gratings. (reference 4-19)

SYMBOLS	FREQUENCY (f) CYCLES/DEG	METHOD	SUBJECT	NUMBER OF DETERMINATIONS
A,a	0-9	Staircase	JK	6
B,b	1-8	Staircase	JK	6
C,c	2-7	Staircase	JK	7
D,d	2-7	Adjustment	JK	9
E,e	3-6	Staircase	PA	6
F,f	3-6	Adjustment	JK	10
G,g	4-5	Adjustment	JK	12
H,h	5-4	Adjustment	JK	10
I,i	6-3	Adjustment	JK	12

The comparison was made by forming the ratio of sine wave contrast sensitivity to square wave contrast sensitivity at each spatial frequency point. For a wide range of spatial frequencies, the ratio was found to be very nearly equal to $4/\pi$. The Fourier series for a unit amplitude square-wave with period X is given in equation 4-25.

$$\frac{4}{\pi}\left\{\sin\left(\frac{2\pi x}{X}\right)+\frac{1}{3}\sin 3\left(\frac{2\pi x}{X}\right)+\frac{1}{5}\sin 5\left(\frac{2\pi x}{X}\right)+\ldots\right\} \quad (4\text{-}25)$$

This equation shows that the fundamental component of a square wave has $4/\pi$ times the contrast of the square wave, or in comparing a sine wave and a square wave of equal amplitude and frequency, the fundamental of the square wave has $4/\pi$ times the contrast of the sine wave. Now if the multiple channel hypothesis is at work here, then the sine wave and the fundamental of the square wave would be detected by the same channel since they are of the same spatial frequency. At the detection threshold, the sine wave and the fundamental component of the square wave will be equal in contrast and amplitude if the amplitude of the square wave is a factor of $4/\pi$ smaller than the amplitude of the sine wave, which agrees with the contrast sensitivity ratio described above. The harmonics of the square wave would presumably lie in other spatial frequency channels and since their amplitudes are considerably below that of the fundamental, they would all be below threshold and not contribute to the detection of the square wave. On the other hand, the results expected from the single channel model are dependent upon the nature of the transformation

device, that is, whether or not the transformation device modifies the peak to trough difference of the square wave differently from the peak to trough difference of a sine-wave. If the transformation device is assumed to be linear and is characterized by a modulation transfer function such as the one shown in figure 4-11, then some predictions could be made about the behavior of the single channel model. Campbell and Robson performed the necessary calculations and obtained a predicted result which agrees in some respects with the contrast sensitivity ratio measurements. However, at low frequencies, there is a considerable deviation from the measured results. Also at about 2 cycles per degree where the third harmonic would cause a maximum reduction in the peak-to-trough difference, there seems to be no indication that this effect actually occurs.

Campbell and Robson performed another experiment in which the contrast sensitivity of a sine wave was compared against the contrast sensitivity for a rectangular wave. The comparison was made for several values of the duty cycle R of the rectangular grating. The duty cycle is the proportion of the period during which the grating is at maximum luminance. The Fourier series of a unit amplitude rectangular grating, as a function of duty cycle, is given in equation 4-26.

$$\frac{4}{\pi}\left\{(\sin\pi R)\sin\left(\frac{2\pi x}{X}\right)+\left(\frac{\sin 2\pi R}{2}\right)\sin 2\left(\frac{2\pi x}{X}\right)+\left(\frac{\sin 3\pi R}{3}\right)\sin 3\left(\frac{2\pi x}{X}\right)+\ \ldots\ \right\} \tag{4-26}$$

The contrast sensitivity ratio was measured for several values

of duty cycle and one spatial frequency (11 cycles per degree). The resulting values fitted very closely a curve given by $(4\sin\pi R)/\pi$ which is the contrast ratio between a sine wave and the fundamental component of a rectangular wave form of the same amplitude and period. The result is compatible with the multiple channel hypothesis.

A square wave has only odd harmonics, and the coefficients of these harmonics decrease according to the sequence 1/3, 1/5, 1/7, ... with increasing frequency. However, a rectangular grating with the duty cycle near zero, or near unity, not only has odd and even harmonics, but in addition the coefficients of these harmonics decrease very slowly with increasing spatial frequency. Let us assume that the eye performs detection in accordance with the multiple channel model where the transformation device is a modulation transfer function like the one shown in figure 4-11: for a square wave with a fundamental frequency on the order of 4 cycles per degree, its harmonics will be greatly attenuated relative to the fundamental. This is because the harmonic frequencies are at high multiples of the fundamental frequency and therefore fall on a high frequency portion of the modulation transfer function, and the modulation transfer function decreases rapidly for increasing spatial frequency. In addition, square wave harmonics have rapidly decreasing coefficients. For a low duty cycle rectangular grating, the harmonics are more closely spaced in frequency, and the harmonics do not decrease rapidly. Therefore, for the same fundamental spatial frequency the harmonics are

larger relative to the fundamental for a low duty cycle rectangular grating. Therefore, for decreasing fundamental frequency, the harmonics of a low duty cycle rectangular grating will begin to have an effect at higher spatial frequencies than in the case of a square wave grating. Campbell and Robson tested this hypothesis and found that for an average luminance of 500 candelas per square meter, the contrast sensitivity ratio for a low duty cycle rectangular grating was constant only above 6 cycles per degree whereas the square wave grating gave a constant ratio above 0.3 cycles per degree. The deviation from constancy at lower frequencies was attributed to the effect of harmonics.

Measurements of the contrast sensitivity ratio between a sinusoidal and a sawtooth grating for a spatial frequency of 11 cycles per degree also gave a value very near the predicted ratio.

Campbell and Robson also measured the effect on the appearance of a grating as the contrast was increased above the value necessary for threshold detection. In particular, it was of interest to know at what point in the supra-threshold region a square wave grating is distinguishable from a sinusoidal grating. The observer was presented with a square wave and sinusoidal grating simultaneously. The contrast of the two gratings was constrained to be in the ratio $4/\pi$. The amplitude of the sine wave grating was $4/\pi$ times the amplitude of the fundamental of the square wave grating. Therefore, the amplitude of the sine wave and the amplitude of the fundamental of the square wave were equal.

The contrast of the two gratings was increased simultaneously until the observer could distinguish between the square wave and the sine wave grating. These contrast values were recorded for several spatial frequencies. It was found that the resulting curve of contrast versus spatial frequency was the contrast sensitivity function for the third harmonic of the square wave grating. By shifting the contrast threshold curve for a pure sine wave grating by a factor of three on both the contrast axis and the spatial frequency axis, it was possible to duplicate the measured results. This is to be expected since the third harmonic differs by a factor of three in spatial frequency and amplitude from the fundamental component of the square wave.

It was also found that a sawtooth wave grating was distinguishable from a sine wave grating when the contrast was sufficient to raise the second harmonic above threshold. Therefore, it is possible that in both the case of the square wave and sawtooth wave gratings, the harmonics reside in separate channels from the fundamental and are separately and independently detected. In addition, when a channel component is detected, the output from that channel is added to the output from other detected channels to form the perceived signal.

4.2.1.3 Blakemore and Campbell (reference 4-21)

In these experiments, pre-adaptation of the observer's eye was used to investigate the nature of the spatial frequency channels. It was found that pre-adaptation to a

high contrast sinusoidal grating of a particular spatial frequency decreased the subsequent contrast sensitivity for that same spatial frequency. In addition, if contrast sensitivity was measured over a range of spatial frequencies after pre-adaptation, a trough occurred in the contrast sensitivity curve. This trough indicated a narrow band of decreased contrast sensitivity with the maximum decrease centered at the spatial frequency of the adapting sinusoidal grating.

A much more interesting result was obtained when a high contrast square wave grating was used for adaptation. In this case the contrast sensitivity curve showed bands of decreased contrast sensitivity at both the fundamental and the third harmonic frequencies of the adapting square wave grating.

If spatial frequency channels exist, then it would be expected that the contrast sensitivity function would be due to the effect of several overlapping narrow band filters. For spatial frequencies lying between two filter bands, it might be expected that the contrast sensitivity function would contain a slight dip. Attempts to find these points of separation between bands have not been successful. There is evidence that indicates that these bands may be very closely spaced. (reference 4-22).

4.2.1.4 Sachs, Nachmias, and Robson (reference 4-23)

A psychometric function is a relation between the frequency of occurrence of a perceptual phenomenon and the

values of a variable upon which that phenomenon depends. The frequency of detection of a sinusoidal grating versus the contrast of that grating would constitute a psychometric function. If the structure of a detection model has been defined, then it is possible to derive the corresponding psychometric function for each perceptual variable of interest. Then it is possible to perform an experiment in which the frequency of detection for a human observer is determined for each value of the perceptual variable. Depending upon whether or not the experimental results agree with the predicted psychometric function, the proposed model would be either accepted or rejected. Sachs, Nachmias, and Robson defined a single-channel model and a multiple-channel model and then derived the expected psychometric functions for both models. In this case, the perceptual phenomenon whose frequency of detection is being predicted is a periodic luminance distribution. The luminance distribution has an average value L_o. The luminance distribution is given by equation 4-27.

$$L(x)=L_o[1+m\sin(2\pi f+\phi)+m'\sin(2\pi f'+\phi')] \qquad (4\text{-}27)$$

where L_o is the average luminance, m and m' are the contrasts of the sinusoidal components of frequencies f and f' respectively. m is the perceptual variable, and m',f,f',ϕ' and ϕ' are treated as parameters. The models are shown in figures 4-15 and 4-16. These models are basically the single channel and multiple channel models discussed previously with the addition of noise sources and a guess generator.

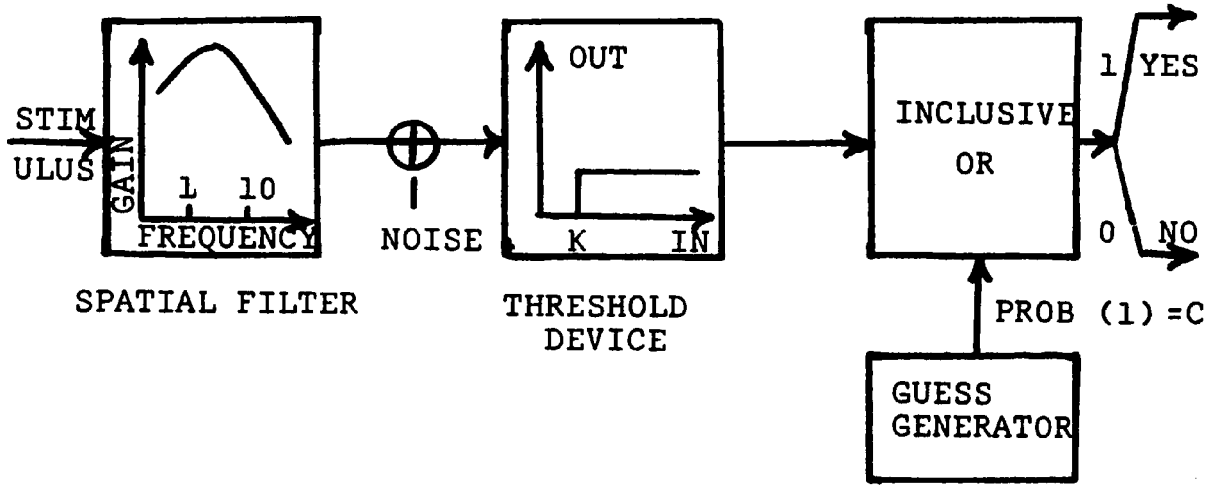


Figure 4-15. Block diagram of the single-channel model. (Reference 4-23)

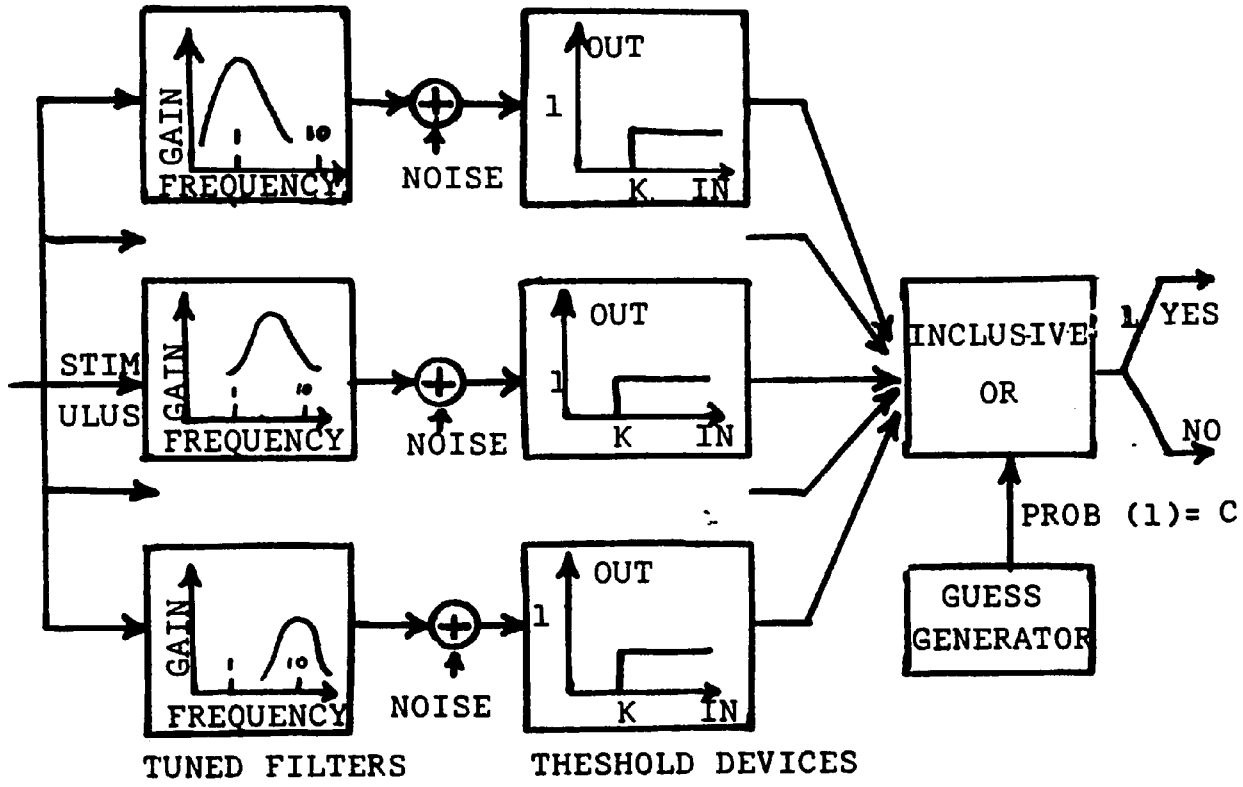


Figure 4-16. Block diagram of the multiple-channel model (Reference 4-25)

The guess generator indicates detection of the stimulus with some probability c regardless of the nature of the stimulus.

For the single-channel model, it was assumed that the probability of detecting the luminance distribution was "some function, F, of a linear combination of either the modulations, m and m', or of some transformation of these modulations," as in equation 4-28.

$$P(Y|m,m')=F[aG(m)+a'G'(m')] \qquad (4\text{-}28)$$

P is the probability of detecting the stimulus given that the sinusoidal components with frequencies f and f' have contrasts m and m'. Now if m'=0, equation 4-28 becomes equation 4-29.

$$P(Y|m,0)=F[aG(m)+a'G'(0)] \qquad (4\text{-}29)$$

This is the psychometric function for a sinusoidal luminance distribution of frequency f as a function of G. Since G'(m') is a constant for each value of m', then the psychometric function given by equation 4-28, for each value of m', can be found by shifting the psychometric function given by equation 4-29 along the G-axis. Therefore, the predicted family of psychometric functions for the single channel model should be as shown in figure 4-17.

For the multiple-channel model, assuming that f and f' lie in separate channels and assuming that the channels and the guessing mechanism are stochastically independent of one another, equation 4-30 results.

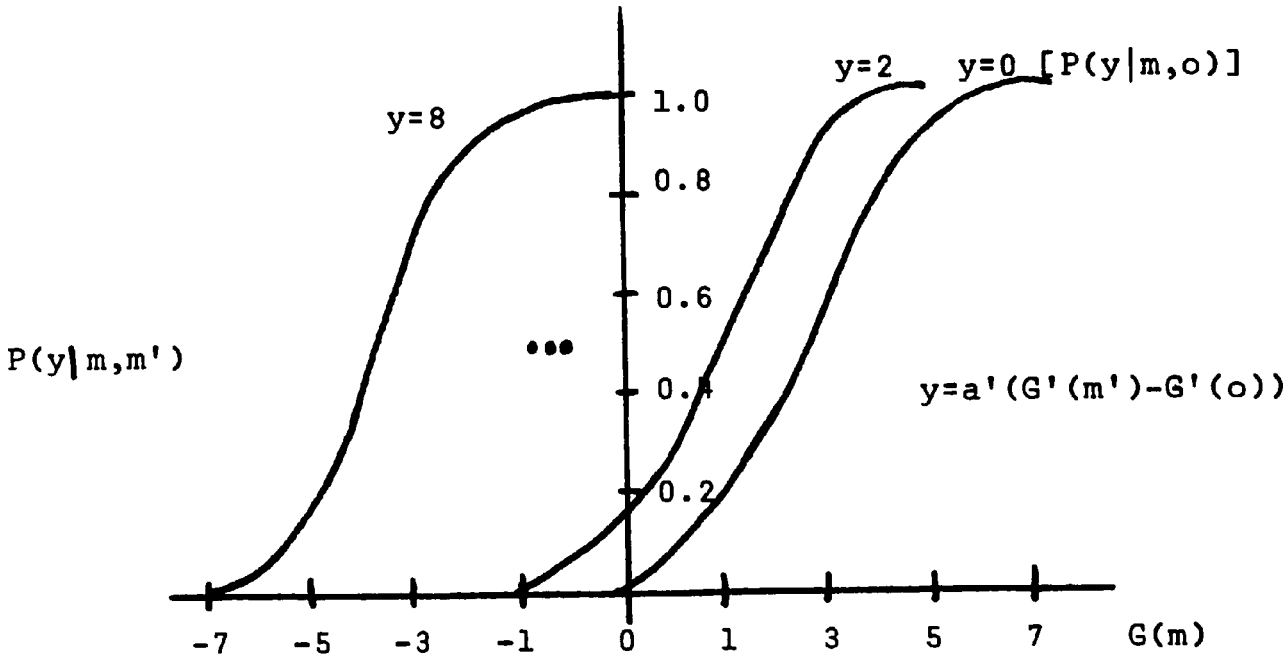


Figure 4-17. Psychometric functions for the single channel model

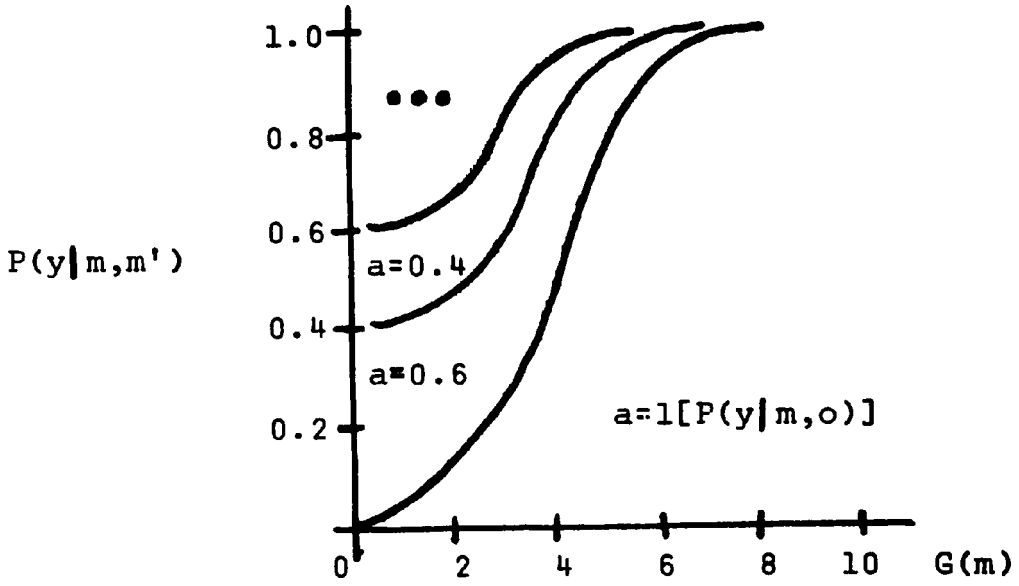


Figure 4-18. Psychometric functions for the multiple channel model

$$P(N|m,m') \propto P(N|m,0) \qquad (4\text{-}30)$$

This equation says that the probability of not detecting the complex luminance distribution is proportional to the probability of not detecting the corresponding sinusoidal luminance distribution at frequency f. The constant of proportionality is less than or equal to one. In figure 4-18 is shown a psychometric function for a sinusoidal luminance distribution and the corresponding family of psychometric functions predicted by equation 4-30.

Sachs, et al, measured the psychometric functions experimentally. Figure 4-19 shows some typical results. In some cases it can be seen that the experimental results follow very closely the expected results for the multiple-channel model. After application of statistical tests to the data, the authors concluded that the results were "clearly consistent" with the multiple-channel model. The only instances in which the results were not consistent were those cases in which f and f' were closest to one another and therefore, might not have been in separate spatial frequency channels.

4.2.1.5 The existence of spatial frequency selective cells

Experiments have been performed on cats and monkeys to determine whether or not in the visual portions of the brain there exist cells which are selectively sensitive to spatial frequency. In these experiments moving or stationary gratings are focused on the retina of the animal, and

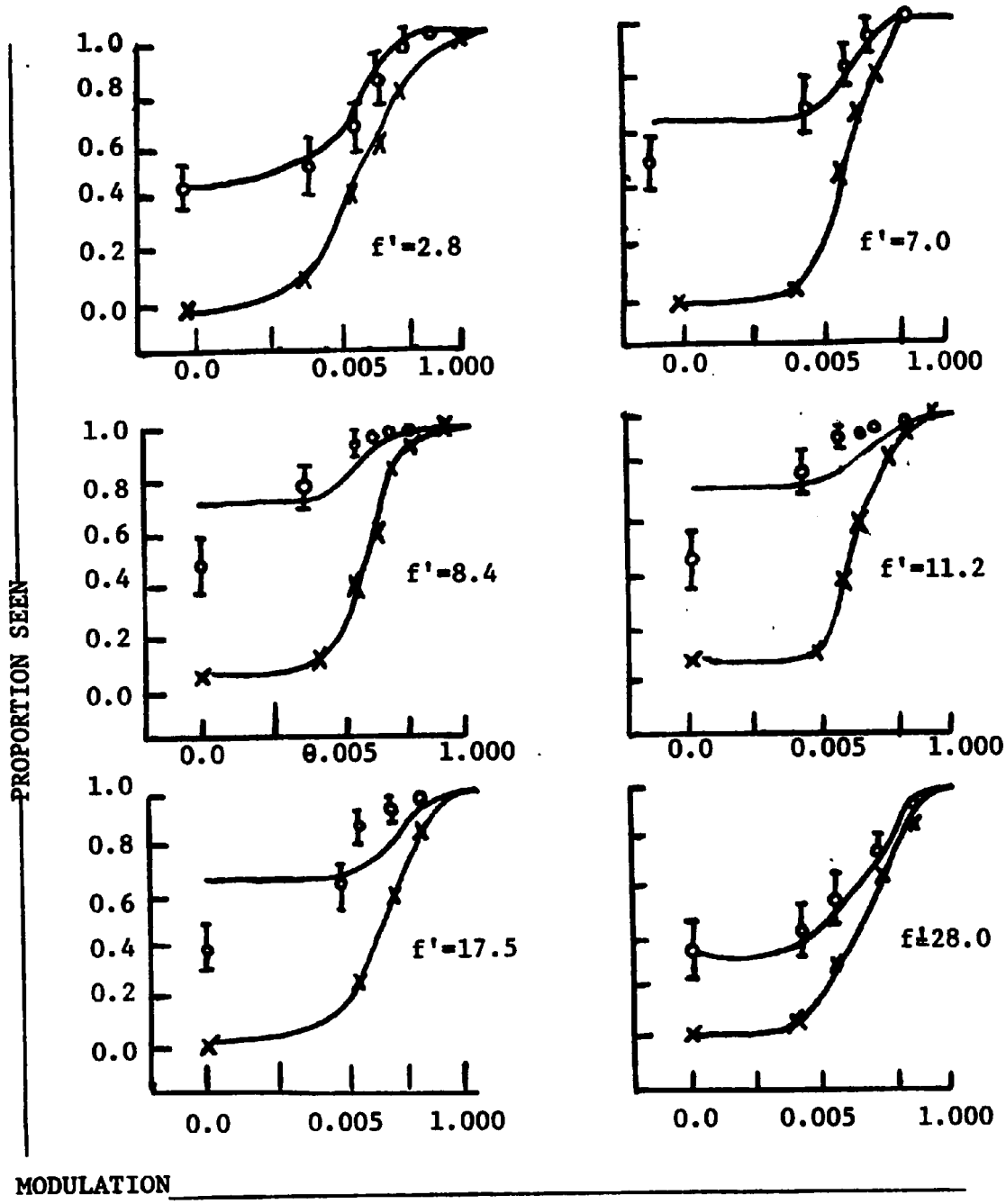


Figure 4-19. Simple and complex psychometric functions for f=14 cycles/deg and various values of f' (given in cycles/deg). The solid curves fitted to the complex psychometric functions correspond to the multiple-channel model in the case where f and f' are detected by completely separate channels (see text). Vertical bars correspond to 95% confidence intervals. Observer MR. (reference 4-23)

measurements are taken from electrodes inserted into specific fibers, or cells, in the brain. In one such experiment, (reference 4-24), electrodes were inserted into lateral geniculate fibers to measure the response to moving gratings. For uniform illumination, the measured pulse density was constant, typical of nerve fibers. The moving grating modulated the pulse density in a manner dependent on spatial frequency. The mean pulse rate was measured as a function of spatial frequency. For each of the 43 geniculate units measured, the spatial frequency was determined at which the response had fallen off by a factor of ten (1 log unit). These 43 spatial frequencies covered a range from 0.31 to 3.8 cycles per degree. Similar measurements were made of cortical cells. Their high frequency cutoff points ranged from 0. 18 to 1.6 cycles per degree. It was felt that these neurophysiological findings support the multiple-channel hypothesis and indicate the existence of spatial frequency selective neurons.

Another paper (reference 4-25) describes measurements of the response of cat retinal ganglion cells to stationary and moving sinusoidal gratings. In one part of the experiment, a stationary sinusoidal grating was switched on, and after about two seconds, was switched off. Two characteristic types of responses were observed: cells which respond with a uniform change in pulse density (X cells) and cells which characteristically respond to the onset and termination of the grating (Y cells). In addition, the response was found to vary with the phase of the grating. As the phase

of the grating was changed, pulse density was modulated symmetrically about the mean pulse density. In addition, the modulation of the pulse density had the same spatial period as the sinusoidal grating. The Y cells responded to changes in phase by changing the relative magnitude of the unset and termination responses.

In the other part of the experiment, moving gratings were employed. Since this is equivalent to a continuous change of phase, the X cells responded with a modulated pulse density. The mean pulse density was constant, independent of spatial frequency, but the amplitude of the pulse density modulation was a function of spatial frequency. The Y cell response was complex and both the mean pulse density and the response were functions of spatial frequency . For both X and Y cells, increases in contrast would cause increased response. However, the mean pulse density was constant only for the X cells. The response was proportional to the square root of contrast.

A "contrast sensitivity function" was measured for X cells by establishing a standard modulation pulse density amplitude and then measuring the contrast at each spatial frequency necessary to produce that amplitude. The contrast sensitivity functions were of different shape for different X cells. In particular, the spatial frequency corresponding to maximum contrast sensitivity was different. For a given cell, the contrast sensitivity depended upon the mean retinal illumination. All of the contrast sensitivity curves obtained were qualitatively similar to those for human vision.

The authors speculated that the receptive field for a ganglion cell consists of concentric excitatory and inhibitory zones. Using Gaussian excitation and inhibition functions, they were able to predict the contrast sensitivity functions for various X cells by suitable choice of the radii and coefficients of the excitatory and inhibitory fields. The authors concluded that the existence of a null in the X cell response for certain phase angles implied that the summation at the ganglion cells of the inhibitory and excitatory responses is linear with luminance. However, the response of the ganglion cell itself is not linear with luminance. To satisfactorily predict changes in contrast sensitivity due to changes in mean retinal illuminance, it was necessary to assume decreases in the maximum values of the Gaussian inhibitory and excitatory functions as well as increases in the radius of the central excitatory region.

The X cells are primarily located in the central retina whereas the Y cells tend to be peripherally located. Therefore, it is possible that human foveal vision is also predominated by these spatial frequency selective X cells.

Evidence of spatial frequency selective visual cells in the monkey has also been reported. (reference 4-26)

Images which have been artificially stabilized on the retina disappear rapidly and are no longer visible (reference 4-6). Therefore, it has been concluded that the image must move relative to the retina in order to be seen. In human vision it has been found that even when the observer

fixates on an object the eye is in motion causing the image to change position periodically on the retina. The effect of these eye movements is to convert a spatial retinal illuminance gradient into a temporal illuminance gradient. On this basis, the use of moving gratings in the experiments described above seems to have an effect similar to eye movements. Some of the spatial contrast characteristics of vision can be explained in terms of temporal characteristics using information about eye movements (reference 4-27, 4-28, 4-29).

4.3 A model for the transformation of a spatial luminance distribution into a perceived brightness distribution

In the preceding sections, some models were described which might be useful for simulating the behavior of different parts of the human visual system. It was shown that the optics of the eye could be described in terms of a modulation transfer function. The various neural layers of the retina were characterized in terms of a spatial interaction system. The spatial interaction system was described as consisting of an array of concentric and overlapping, excitatory, and inhibitory regions. Evidence was given for the existence of cells in the retina and the visula cortex whose response could be described in terms of the elements of a spatial interaction system. These cells were shown to behave like spatial frequency band-pass filters. The evidence indicated the strong possibility

that human vision can be described in terms of multiple spatial frequency selective channels.

In order to evaluate the effect of a picture bandwidth compression technique on the quality of a reconstructed picture, it would be useful to be able to compare perceived brightness of the original and reconstructed pictures. What has been described here is probably inadequate for the synthesis of a general model for the prediction of perceived brightness. However, there is a model that seems to be a natural consequence of the material in the preceding sections, and therefore is worthy of investigation. The model is shown in figure 4-20.

This model combines the modulation transfer function for the optics with multiple spatial frequency channels for the retina and visual cortex. A logarithmic transformation is shown for the conversion of retinal illuminance to photoreceptor response. There is some evidence that this is the mechanism whereby the total dynamic luminance range of the eye (11 log units) is scaled by the photoreceptors of the eye (reference 4-6). The parameters used in the model are dependent upon the visual environment. The ambient luminance level will determine the pupil diameter which in turn determines the modulation transfer function for the optics. The range of retinal illuminance will determine the operating range on the logarithmic transformation. Since the characteristics of the receptive field of ganglion cells depend upon the ambient luminance, the band-pass characteristics of the individual spatial frequency channels will also be

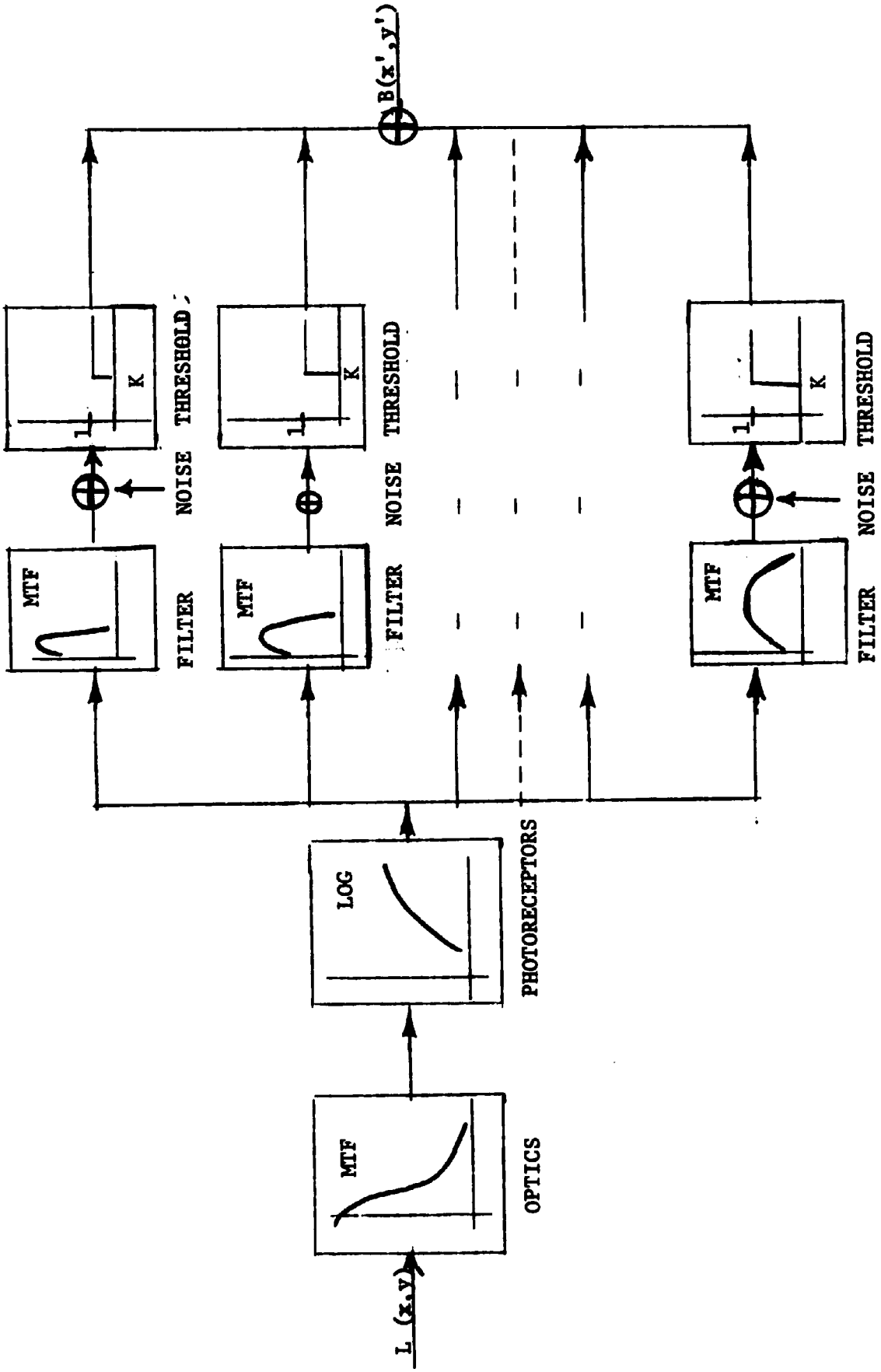


Figure 4-20 A MODEL FOR THE PREDICTION OF HUMAN BRIGHTNESS PERCEPTION

dependent upon ambient luminance. Another manifestation of this dependence is the change in contrast sensitivity with average luminance. The noise sources represent random fluctuations in the number of quanta striking the retina as well as noise internal to the cells of the retina and visual cortex.

REFERENCES

4-1 R.V. Churchill, Fourier Series and Boundary Value Problems, McGraw-Hill, New York, 1941.

4-2 W.M. Brown, Analysis of Linear Time-Invariant Systems, McGraw-Hill, New York, 1963.

4-3 J.W. Goodman, Introduction to Fourier Optics, McGraw-Hill, New York, 1968.

4-4 M. Born, E. Wolf, Principles of Optics, 2nd revised edition, Pergamon Press, New York, 1964.

4-5 G.S. Brindley, Physiology of the Retina and Visual Pathway, Williams and Williams, Baltimore, 1970.

4-6 T.N. Cornsweet, Visual Perception, Academic Press, N.Y., 1970.

4-7 M. Davidson, "Perturbation Approach to Spatial Brightness Interaction in Human Vision," JOSA, vol. 58, no. 9, Sept 1968, p. 1300.

4-8 O. Bryngdahl, "Characteristics of the Visual System: Psychophysical Measurements of the Response to Spatial Sine-Wave Stimuli in the Mesopic Region," JOSA, vol. 54, no. 9, 1964, p. 1152.

4-9 F. Ratliff, Mach Bands, Holden-Day Inc., San Francisco, 1965.

4-10 J. Bliss, W. Macurdy, "Linear Models of Contrast Phenomena," JOSA, vol. 51, no. 12, 1961, p. 1373.

4-11 R.B. Barlow, Jr., "Inhibitory Fields in the Limulus Lateral Eye," J. Gen. Physiol, vol. 54, 1969, p. 383.

4-12 F., Ratliff, H.K. Hartline, W.H. Miller, "Spatial and Temporal Aspects of Retinal Inhibitory Interaction," JOSA, vol. 53, pp. 110-120, 1963.

4-13 H.K. Hartline, F. Ratliff, W.H. Miller, "Inhibitory Interaction in the Retina and its Significance in Vision," Nervous Inhibition, E. Florey, ed. Pergamon Press, N.Y., 1961, p. 241.

4-14 F. Ratliff, H.K. Hartline, "The Response of Limulus Optic Nerve Fibers to Patterns of Illumination on the Retinal Mosaic," J. Gen. Physiol. vol. 42, pp. 1241-1255.

4-15 H.K. Hartline, F. Ratliff, "Inhibitory Interaction of Receptor Units in the Eye of Limulus," J. Gen. Physiol. vol. 40, pp. 357-376.

4-16 H.K. Harline, H.G. Wagner, F. Ratliff, "Inhibition in the Eye of Limulus," vol. 39, J. Gen. Physiol., 1956, pp. 651-673.

4-17 J.C. Trinder, "Effects of Image Blur and Lateral Inhibition In the Visual System on Visual Performance," Optica Acta, vol. 18, no. 6, June, 1971, pp. 461-476.

4-18 E.M. Lowry, J.J. DePalma, "Sine-Wave Response of the Visual System, the Mach Phenomena," JOSA, vol. 51, 1961, p.740.

4-19 N. Graham , J. Nachimas, "Detection of Grating Patterns Containing Two Spatial Frequencies: a Comparison of Single Channel and Multiple Channel Models," Vision Research, 1971, vol. 11, no. 3, pp. 251-260.

4-20 F.W. Campbell, J.G. Robson, "Application of Fourier Analysis to the Visibility of Gratings," J. Physiol., vol. 197, (1968), pp. 551-566.

4-21 C. Blakemore, F.W. Campbell, "Adaptation to Spatial Stimuli," J. of Physiol., vol. 200, no. 1, Jan., 1969, pp. 11p-13p.

4-22 F.W. Campbell, J. Nachmias, J. Jukes, "Spatial Frequency Discrimination in Human Vision," JOSA, vol. 50, pp. 555-559, 1970.

4-23 M.B. Sachs, J. Nachmias, J.G. Robson,"Spatial Frequency Channels in Human Vision," Vo. 61, no. 9, JOSA, pp. 1176-1186.

4-24 F.W. Campbell, G.F. Cooper, Christina Enroth-Cugell, "The Spatial Selectivity of the Visual Cells of the Cat," J. Physiol. (1969), vol. 203, pp. 223-235.

4-25 Christina Enroth-Cugell, J.G. Robson, "The Contrast Sensitivity of Retinal Ganglion Cells, of the Cat," J. Physiol., Vol. 187, 1966, pp. 517-552.

4-26 F.W. Campbell, G.F. Cooper, J.G. Robson, M.B. Sachs, "The Spatial Selectivity of Visual Cells of the Cat and the Squirrel Monkey," J. of Physiol., vol. 204, no. 2, Oct 1969, pp. 120P-121P.

4-27 L. Ronchi, G. Toraldo di Francia, "On the Response of the Human Eye to Light Stimuli Presenting a Spatial or Temporal Gradient of Luminance," JOSA, vol. 47, 1957, p. 639.

4-28 J. Kulikowski, "Some Stimulus Parameters Affecting Spatial and Temporal Resolution of Human Vision," Vision Research, vol. 11, pp. 83-93.

4-29 J. Kulikowski, "Effect of Eye Movements on the Contrast Sensitivity of Spatio-Temporal Patterns," Vision Research, vol. 11, pp. 261-273.

CHAPTER V

APPLICATION OF THE PERCEPTION MODEL TO PICTURE BANDWIDTH COMPRESSION

5.0 Introduction

In order to properly apply the characteristics of human vision to picture bandwidth compression, it is necessary to determine the physical parameters of the picture viewing environment. Two important viewing parameters are luminance measured at the eye and viewing distance. Some of the characteristics of vision which are dependent on luminance are the modulation transfer function for the optics, visual acuity, and contrast sensitivity. Viewing distance not only affects lens accommodation, but also determines the spatial frequency content of the image on the retina. For a given set of viewing conditions, it will be possible to set bounds on the number of brightness quantization levels and the spatial sampling density for digitized pictures. This will establish a basis for the determination of the efficiency of a picture bandwidth compression technique.

Once the characteristics of the viewing environment have been determined, it will be possible to establish the parameters for the model shown in figure 4-20. The model in figure 4-20 will not be used directly because of its complexity. The approach taken here will be to make considerable simplifying assumptions in order to obtain a model from which results can be obtained easily. Then the model will be increased in complexity in order to bring it into closer

agreement which hypothesize characteristics of vision. At each level of complexity, the results will be compared. This will not only elucidate changes introduced by added complexity but will also be indicative of the utility gained by added complexity.

In general terms, the model will be used to determine which spatial frequency components in the original picture are not perceptible. The imperceptible spatial frequency components will be eliminated, and the remaining components will be used to reconstruct the picture. A comparison will then be made between the compressibility of the original and the reconstructed picture.

5.1 The picture viewing environment

The parameters of the typical viewing environment will now be described. The relationship between viewing distance and spatial frequency will be determined, and a basis for compression efficiency will be discussed. A particular set of sample digitized pictures were available for this work. The characteristics of these pictures will also be described in this section.

5.1.1 Luminance, reflectance, and pupil diameter

In order to understand the concept of luminance and the units that are used for its measurement, some terms need to be defined (reference 3-10).

Radiant flux (Watts): radiant energy being transferred from a source to a receiving surface per unit time.

Luminous flux (lumen): radiant flux in the frequency range which produces visual sensation, evaluated according to the values of relative luminous efficiency for the light adapted eye (adopted by the Commision Internationale de L'Eclairage). At the wavelength of maximum photopic visual sensitivity, 555 nm, one watt of radiant energy equals 621 lumens. At other wavelengths, correspondingly fewer lumens are associated with one watt in accordance with the photopic relative luminous efficiency given in table 5-1.

Luminous intensity (candela or lumen/steradian): the quotient of the luminous flux emitted from a source or from an element of a source in an infinitesimal cone containing the direction of measurement, by the solid angle subtended by the cone. By definition, a uniform point source of intensity one candela emits one lumen of luminous flux per unit solid angle or a total of 4π lumens.

$$I=\lim_{\Omega\to 0}\frac{F_\Omega}{\Omega} \qquad (5\text{-}1)$$

Luminance (candela per square meter): the quotient of the luminous intensity of an infinitesimal element of the source in the desired direction, by the area of the orthogonal projection of the source surface element on a plane perpendicular to that direction.

$$L=\lim_{s\to 0}\frac{I_\theta}{s\cos\theta} \qquad (5\text{-}2)$$

where θ is the angle between the normal to the area s and the direction of measurement.

In illumination engineering, the luminance of a surface in any given direction is sometimes expressed in terms of the luminance of a "uniform diffuser" emitting one lumen per unit area. A uniformly diffusing surface is one for which the luminous intensity in any direction varies as the cosine of the angle between that direction and the normal to the surface so that it appears equally bright from whatever direction it is viewed. For such a surface,

$$I_\theta = I\cos\theta \qquad (5\text{-}3)$$

$$L = \lim_{s\to 0} \frac{I\cos\theta}{s\cos\theta} = \lim_{s\to 0} \frac{I}{s} \qquad (5\text{-}4)$$

If the luminance of a uniform diffuser is I candelas per unit area, then the total flux emitted per unit area is given by equation 5-5,

$$F = \int_0^{\pi} I\cos\theta \, d\Omega = \int_0^{\pi} (I\cos\theta)\frac{ds}{r^2} = \int_0^{\pi} \frac{(I\cos\theta)(2\pi r^2 \sin\theta)}{r^2} d\theta = \pi \cdot I \qquad (5\text{-}5)$$

where F is in lumens per unit area. Therefore, if I equals I/π candelas per unit area, the uniform diffuser will emit one lumen per unit area. This establishes a unit of luminance relative to a uniform diffuser emitting one lumen per unit area ($1/\pi$ candelas/ft^2 = 1 foot-Lambert, $1/\pi$ candelas/cm^2 = 1 lambert = 1000 millilamberts).

Illumination (lumens per unit area): illumination at a point of a surface is the quotient of the luminous flux incident on an infinitesimal element of surface containing the point by the area of this element.

$$E = \lim_{s\to 0} \frac{F}{s} \qquad (5\text{-}6)$$

The illumination corresponding to a luminous intensity I is given by equation 5-7 for an area s which is small relative to the distance d from the source, and s normal to the direction incidence.

$$E = \left(\frac{I}{s}\right)\left(\frac{s}{d^2}\right) = \frac{I}{d^2} \qquad (5\text{-}7)$$

Retinal illumination: for vision experiments, a special unit has been devised which accounts for the variation in retinal illumination with pupil size for a fixed value of source luminance. This unit is called the troland. One troland is the retinal illumination produced by a surface having a constant luminance of 1 candela per square meter when the pupil area is one square millimeter.

TABLE 5-1

THE RELATIVE PHOTOPIC LUMINOUS EFFICIENCY OF RADIATION (V_λ)

(Ref. 3-10)

M_μ	0	1	2	3	4	5	6	7	8	9
400	0004	0005	0006	0006	0007	0008	0009	0009	0010	0011
10	0012	0013	0015	0017	0019	0022	0025	0028	0031	0035
20	0040	0046	0052	0058	0065	0073	0081	0089	0098	0107
30	0118	0126	0136	0146	0157	0168	0180	0192	0204	0217
40	023	024	026	027	028	029	031	032	034	036
450	038	040	042	044	046	048	050	053	055	058
60	060	063	065	068	071	074	077	080	084	087
70	091	095	099	103	108	112	117	122	128	133
80	139	144	150	156	163	169	176	183	191	199
90	208	217	227	237	247	258	270	282	295	309
500	323	339	355	372	389	407	426	445	464	483
10	503	524	545	566	587	608	629	650	670	690
20	710	728	745	762	778	793	808	822	836	849
30	862	874	885	896	906	915	924	932	940	947
40	954	960	966	971	975	980	984	987	990	993
550	995	997	998	999	1	1	1	999	998	997
60	995	993	990	987	983	979	974	969	964	958
70	952	945	938	931	923	915	907	898	889	880
80	870	860	849	839	828	817	805	793	781	769
90	757	745	732	720	707	695	682	670	657	644
600	631	618	605	593	580	567	554	541	528	515
10	503	490	478	465	453	441	429	417	405	393
20	381	369	357	345	333	321	309	298	287	276
30	265	255	245	235	226	217	208	199	191	183
40	175	167	160	152	145	138	132	125	119	113
650	107	101	096	091	086	081	077	073	069	065
60	061	057	054	051	048	045	042	039	037	034
70	032	030	028	026	025	023	022	020	019	018
80	017	0159	9148	0138	0128	0119	0111	0103	0095	0088
90	0082	0076	0071	0066	0061	0057	0053	0050	0047	0044
700	0041	0038	0033	0031	0029	0028	0026	0024	0024	0022
10	0021	00196	00183	00171	00159	00148	00138	00129	00120	00112
20	00105	00098	00091	00085	00079	00074	00069	00064	00060	00056
30	00052	00048	00045	00042	00039	00036	00033	00031	00029	00027
40	00025	00023	00022	00020	00019	00017	00016	00015	00014	00013
750	00012	00011	00010	00009	00009	00008	00008	00007	00007	00006
60	00006									

For pupil diameters less than 2-2.5 millimeters, the retinal illumination is proportional to pupil area. However, for large pupil diameters, the reduced physiological effectiveness of light which eccentrically enters the pupil (Stiles-Crawford effect), complicates the conversion to trolands. Figure 5-1 illustrates the relationship between relative brightness and pupil diameter under photopic conditions (cone vision).

The Stiles-Crawford Effect

Pupil Diam. (mm)	Relative Area	Relative Apparent Brightness
2	1	1.00
3	2.25	2.14
4	4	3.53
5	6.25	4.99
6	9	6.34
7	12.25	7.46
8	16	8.26

Figure 5-1 Relative brightness versus pupil diameter. (reference 3-10)

A source of illumination can be characterized by the distribution of its radiant flux as a function of wave length and spatial coordinates. The distribution of radiant flux versus wave length is described by the spectral density curve for the source. For a point source, this would be the amount of radiant flux in watts emitted at each wavelength W_λ. If W_λ is multiplied by the relative luminous efficiency V_λ for each wave length, the luminous flux distribution for the source is obtained. If W_λ in a given direction is known for the source then the luminous intensity is given by equation 5-8,

$$I=K_m\int V_\lambda W_\lambda d\lambda \qquad (5\text{-}8)$$

K_m=680 lumens per watt. K_m is determined by the standard that the luminance of a full radiator at the temperature of solidifying platinum (2042.1°K) is 60 candelas/cm^2.

In the picture viewing environment, the luminous intensity of the source of illumination is important, but it is the luminance of the picture itself that determines the operating characteristics of the eye. Certainly light enters the eye from sources other than the picture. It will be assumed that this stray light is not as effective as the light from the picture focused on the retina. The amount of source illumination reflected by the picture into the eye depends upon the reflective properties of the picture surface. Reflection at the surface is basically of two types: <u>specular</u> and <u>diffuse</u>. Specular reflection is governed by by Fresnel's law and is dependent on the polarization of the

illumination. Diffuse reflection is ideal if the surface is uniformly diffusing. In effect, a fraction of the incident illumination is absorbed and then re-emitted in accordance with the cosine law of emission. The surface appears uniformly bright in all directions. In practice, the reflecting properties of a surface are neither entirely specular nor entirely uniformly diffuse. Specular reflection can be ignored because it is of little value in "seeing" the reflecting surface. Specularly reflected rays entering the eye form an image of the illuminating source and not of the reflecting surface. In addition, specular reflection has appreciable intensity only at the angle complementary to the angle of incidence of the illuminating source. It is diffuse reflection which makes an object visible. If an object, for diffuse reflection, behaves like a uniformly diffusing surface, then if the illumination is known, the illuminance of that surface can be determined. Let the illumination be E lumens per unit area. If a proportion r of the total illumination is diffusely reflected, then the total luminous flux reflected per unit area will be rE. In equation 5-4 it was shown that the total flux emitted per unit area of a uniformly diffusing surface is πL where L is the luminance in candelas per unit area. Therefore, $rE=\pi I$ and the luminance due to diffuse reflection is equal to rE/π candelas per unit area, or rE foot-Lamberts.

Table 5-2 indicates that 10 to 50 lumens per square foot is a reasonable range of illumination for viewing pictures.

Table 5-2 lists some typical values of luminance and illumination.

Luminance Source	Millilamberts (millilumens/cm^2)	CD/M^2 (X3.183)
Direct sunlight	8250	26100
Daylight outside	$0.3X10^3$-$0.3X10^4$	$0.95X10^3$-$0.95X10^4$
Good interior lighting	30-300	95-950
Moderate interior lighting	3-30	9.5-95
Feeble interior lighting	0.3-3	0.95-9.5
Outdoor night lighting	0.03-0.3	0.095-0.95
Night vision	$0.3X10^{-4}$-0.03	$0.95X10^{-4}$-0.095

Illumination Source	Lumens/ft^2
Direct sunlight plus skylight	10,000
Direct sunlight	8,000
Dark day	100
Drafting room, art gallery (on pictures)	50
Classroom, laboratory, library	30-50
Offices	10-50
Full moon	0.03

Table 5-2. Typical values of luminance and illumination (reference 5-1)

Table 5-3 lists specular and diffuse reflection factors of some common papers and inks (reference 5-1). These are materials which would commonly contain text, pictures or photographs. Equations 5-9 and 5-10 describe the meaning of these reflectance factors,

$$B_s = r_s B_t \qquad (5\text{-}9)$$

$$B_d = r_d E \qquad (5\text{-}10)$$

where B_s is the specularly reflected luminance (foot-Lamberts), B_t is the source luminance (foot-lamberts), r_s is the specular reflection factor for a particular angle of view, B_d is the diffusely reflected luminance (foot-lamberts), E is the source illumination (foot-candles), and r_d is the diffuse reflection factor. The value given in Table 5-2 as a standard for the illumination of pictures and drawings is 50 foot-candles (lumens per square foot). Therefore, let E be 50 foot-candles. The source luminance required to produce a surface illumination of 50 foot-candles is dependent upon the geometrical relationship between the source and the reflecting surface. According to equation 5-7, for a small source distant from the reflecting surface, the illumination per unit area of the source is inversely proportional to the distance between the source and the surface. The luminance of a one square foot uniform source required to produce an illumination of 50 foot candles on a surface ten feet from the source would be $5{,}000\pi$ foot-lamberts. On the other extreme, the same illumination could be obtained in a room with a very large uniformly luminous

TYPES OF PAPERS AND INKS	REFLECTION FACTOR	
	Specular	Diffuse
Matte black paper	.005	.04
Matte white paper	.0030	.77
Paper from popular magazine	.0035	.73
Paper from scientific journal	.0070	.75
Drawing paper	.0045	.80
Newspaper	.0065	.68
Medium glossy paper	.0055	.78
Fairly glossy paper	.0080	.75
Semi-matte white photographic paper	.014	.83
Very glossy white photographic paper	.048	.83
Metallic papers		
Silver	.042	.49
Gold	.079	.26
Copper	.11	.28
Inks on appropriate papers		
Dull black	.006	.045
Ordinary black	.009	.027
Glossy black	.017	.019
Super-gloss black	.039	.016

Table 5-3. Reflection Factors of Common Papers and Inks (reference 5-1).

ceiling with a luminance of 25 foot-lamberts (reference 3-10). Good interior lighting is generally considered to produce a luminance in the range 30 to 300 foot-lamberts. Let the luminance B_t be 50 foot-lamberts. This is the same as the luminance of a perfectly uniformly diffusing surface with an illumination of 50 foot-candles. For these values of illumination and luminance, table 5-4 shows the computed values of specular and diffuse luminance of some typical paper reflecting surfaces.

Specular luminance ranges between 0.025 and 2.4 foot-lamberts. If the light flux associated with the specular luminance entered the eye, it would not be in focus on the retina and would form an average veiling luminance which would tend to decrease the contrast. This is demonstrated in equation 5-11:

$$\frac{B_1-B_2}{B_1+B_2} > \frac{(B_1+B_s')-(B_2-B_s')}{(B_1+B_s')+(B_2+B_s')} \qquad (5\text{-}11)$$

where B_s' is the specular veiling luminance. B_1 and B_2 are contrasting values of diffuse luminance focused on the retina.

The diffuse luminance ranges between 0.8 and 38.5 foot-lamberts. The range of diffuse luminance will be used to establish the typical luminance operating range of the eye for a picture surface illumination.

Pupil diameter depends upon field luminance, lens accomodation, and age of the observer. In general, it is not possible to say precisely what the pupil diameter will be if all of these factors are known, especially for a

Type of Paper	r_s	r_d	B_s(ft-L)	B_d(ft-L)
Matte black paper	.0005	.04	.025	2
Matte white paper	.0030	.77	.15	38.5
Very glossy white photographic paper	.048	.83	2.4	41.5
Super gloss black ink	.039	.016	1.95	0.8
Newspaper	.0065	.68	.325	34
Magazine	.0035	.73	.175	36.5

Table 5-4

The specular and diffuse luminance (foot-lamberts) of various reflecting surfaces for B_s = 50 ft-L and E = 50 foot-candles.

complex luminance distribution. If we assume an average field luminance of 20 foot-lamberts (68.5 candelas per square meter), figure 3-4 gives a pupil diameter of 3.1 millimeters. Reference 5-1 contains a graph which gives a pupil diameter of 3.7 millimeters for both eyes exposed to a large uniform field of luminance 20 foot-lamberts. For purposes of defining the picture viewing environment, an average pupil diameter of 3.4 millimeters will be used.

The retinal illumination in trolands can be approximately found by multiplying the field luminance in candelas per square meter by the pupil area in square millimeters. The average retinal illumination for an average field luminance of 68.5 candelas per square meter and an average pupil diameter of 3.4 millimeters (9.07 square millimeters) is approximately 622 trolands. The retinal illumination in terms of luminous flux per unit area depends upon the solid angle subtended at the pupil by a point on the reflecting surface as well as the magnification factor of the eye. For picture to eye distances d, large compared to the pupil diameter, the solid angle is approximately given by a/d^2 where a is the pupil area. A common criteria used for picture viewing distance is 4 times picture height. If we assume that the distance between the anterior and posterior nodal points of the eye is negligible compared to the picture distance (see figure 3-2), and we assume that the posterior nodal point is 17 millimeters in front of the retina, then the magnification of the eye can be determined from figure 5-2.

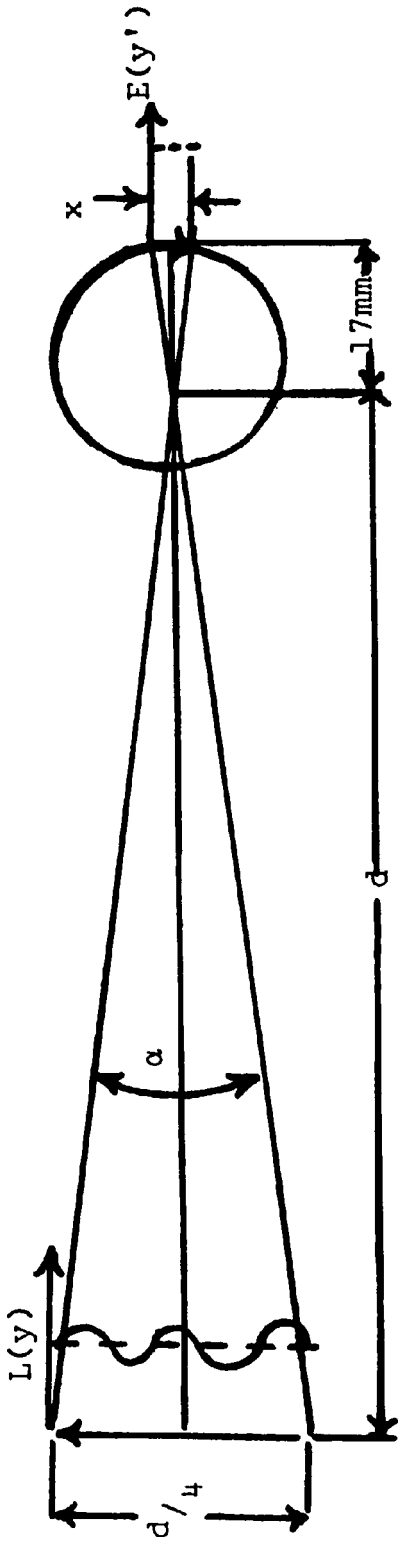


Figure 5-2. Magnification and Visual Angle of a Sine Wave Grating.

$$\frac{(d/8)}{d} = \frac{(x/2)}{17}; \quad x = \frac{17}{4}\ ; \quad M = \frac{(17/4)}{(d/4)} = \frac{17}{d} \qquad (5\text{-}12)$$

The retinal illumination is now given by equation 5-13.

$$E_r = B_d\left(\frac{a}{d^2}\right)\left(\frac{1}{M}\right) = (68.5)\left(\frac{9.07}{d^2}\right)\left(\frac{d}{17}\right) = \frac{36.6}{d} \text{ lux} \qquad (5\text{-}13)$$

where d is in millimeters, and E_r is in lumens per square meter (lux).

5.1.2 Viewing distance, visual angle, and spatial frequency

The perceptual model in figure 4-20 has, as some of its components, spatial frequency filters. Thus, the concept of spatial frequency is of central importance. Some common units of spatial frequency are cycles per degree of visual angle and cycles per millimeter on the retina. Figure 5-2 illustrates how spatial frequency can be computed. The object is a sinusoidal grating, which means that the luminance L(y) measured at the eye varies sinusoidally about a mean luminance. The object subtends a visual angle α relative to the eye. The image on the retina also subtends an angle α relative to the posterior nodal point. It will be assumed that the anterior and posterior nodal points are approximately identical for usual values of distance d. The image is a retinal illuminance distribution $E_r(y')$ which is also sinusoidal about some mean value of illuminance. If the number of cycles in the object luminance distribution is n, then the spatial frequency of the sinusoidal grating is n/α cycles per degree. This would also be the spatial frequency of the retinal illuminance distribution. If the posterior

nodal point is assumed to be 17 millimeters from the retina, then it is possible to compute the visual angle subtended by one millimeter on the retina. This is given by $\tan^{-1}$ (1/17) which gives 3 degrees, 22 minutes of visual angle per millimeter. Therefore, one cycle per degree equals 3.367 cycles per millimeter on the retina. For a complex luminance distribution, in general the spatial frequency of each sinusoidal component will have to be computed. However, if the object is assumed to be periodic in space, then the Fourier components will be harmonically related and so will the spatial frequency components.

Analysis of a luminance distribution into its component sinusoidal frequency components will result in spatial frequencies which are expressed in terms of cycles per unit length along the picture surface. If the fundamental component of a spatially periodic luminance distribution has frequency f, then the period of that component $T = 1/f$ unit lengths per cycle. The visual angle subtended by one cycle of the fundamental is therefore given by $\alpha=\tan^{-1}(T/d)$. The spatial frequency of the fundamental is $1/\alpha$ cycles per degree, and the spatial frequency of the harmonics is n/α cycles per degree, for $n=1,2,\ldots$.

The method described for the computation of spatial frequency is only useful for small visual angles, angles for which $\tan^{-1}(n\cdot\tan\alpha)\underline{\underline{\cong}}n\alpha$. It is for large visual angles that this relation is not satisfied, and the mapping from cycles per unit length on the picture to cycles per degree is non-linear.

There are several ways in which the spatial frequency content of a picture can be determined. The problem is to determine whether or not all of these methods are equivalent, and if they are not, which method is most appropriate to human vision. For example, one of the most straightforward ways to determine the spatial frequency components of a picture is to treat it as a spatially periodic function. If the luminance distribution of the original picture is given by L(x,y) for 0>x>X and 0>y>Y, then the periodically extended picture would be given by Figure 5-3. Such a function can be analyzed in terms of its Fourier series if the conditions stated in the following theorem hold.(reference 5-2)

Theorem 5-1: let f(x+2L)=f(x) for all x, and let f(x) be sectionally continuous in the interval (-L,L). Then at any point where f(x) has a right and left hand derivative, it is true that

$$\frac{1}{2}[f(x+0)+f(x-0)]=\frac{1}{2}a_o\sum_{n=1}^{\infty}\left(a_n\cos\frac{n\pi x}{L}+b_n\sin\frac{n\pi x}{L}\right) \qquad (5\text{-}14)$$

$$a_n=\frac{1}{L}\int_{-L}^{L}f(x)\cos\frac{n\pi x}{L}\,dx \qquad (5\text{-}15)$$

$$b_n=\frac{1}{L}\int_{-L}^{L}f(x)\sin\frac{n\pi x}{L}\,dx \qquad (5\text{-}16)$$

Luminance functions are continuous and have derivatives at every point. Therefore, they can be analyzed in terms of a Fourier series. This is true for both one dimensional and two dimensional functions. In general, at each spatial

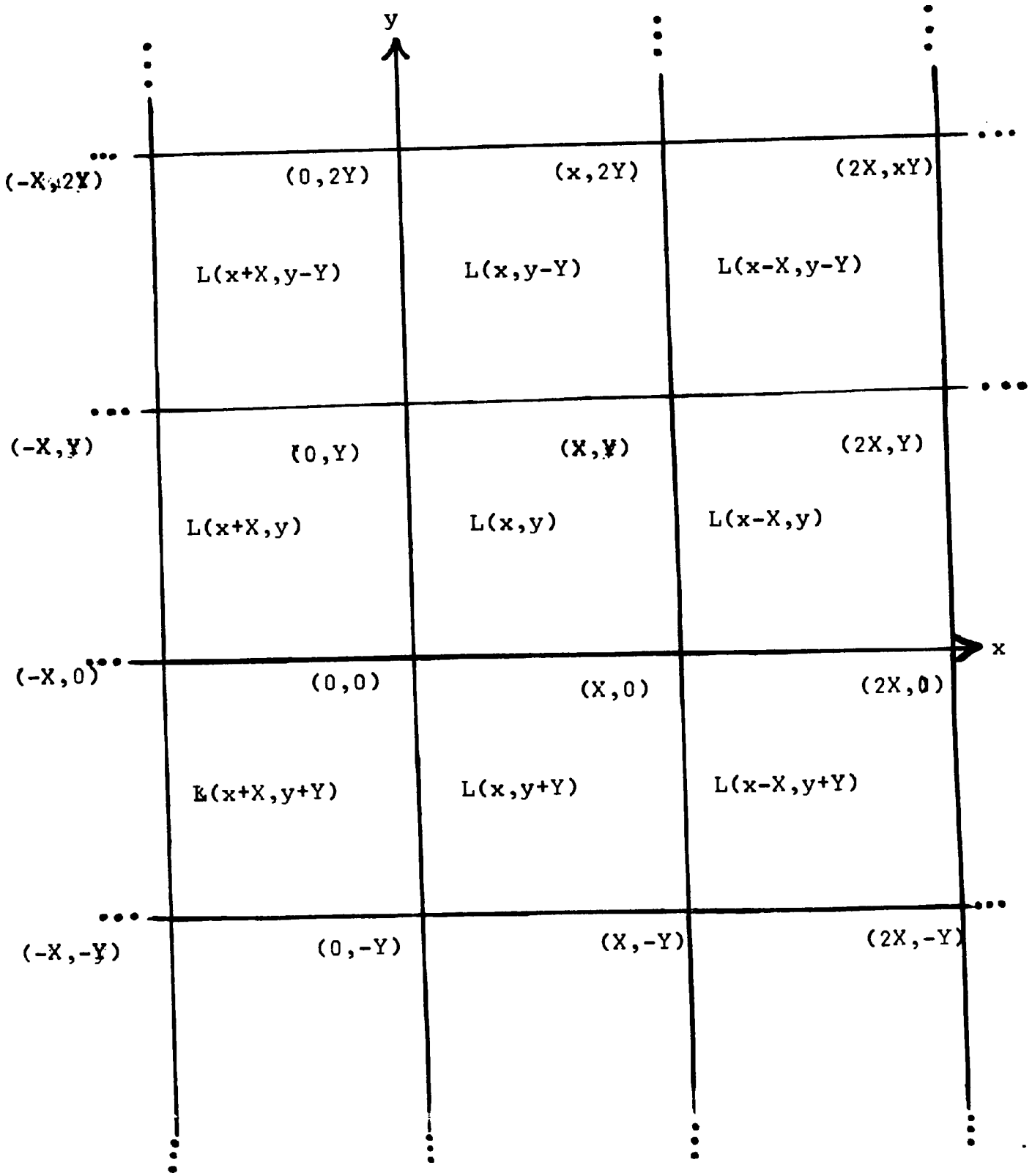


Figure 5-3 Periodic extension of a two-dimensional picture luminance function

frequency, a function has both a sine and a cosine component but in the application of modulation transfer functions and spatial frequency channels, it is the total amplitude of each spatial frequency component that is of interest. At each spatial frequency, the sine and cosine components can be combined to form a single, sinusoidal component and its associated phase as shown in equations 5-17.

$$A\sin\left(\frac{n\pi x}{L}+\phi\right)=A\sin\phi\cos\frac{n\pi x}{L}+A\cos\phi\sin\frac{n\pi x}{L} \qquad (5\text{-}17)$$

$$=a_n\cos\frac{n\pi x}{L}+b_n\sin\frac{n\pi x}{L} \qquad (5\text{-}17.1)$$

therefore,

$$a_n=A\sin\phi \qquad b_n=A\cos\phi \qquad (5\text{-}17.2)$$

$$a_n^2+b_n^2=A^2 \qquad A=\sqrt{a_n^2+b_n^2} \qquad (5\text{-}17.3)$$

$$\tan\phi=a_n/b_n \qquad (5\text{-}17.4)$$

The complex form of the one-dimensional Fourier series is given in equation 5-18.

$$f(x)=c_o+\sum_{m=1}^{\infty}\left(c_n e^{j\frac{n\pi x}{L}}+c_{-n}e^{-j\frac{n\pi x}{L}}\right) \qquad (5\text{-}18)$$

$$c_n=\frac{1}{2L}\int_{-L}^{L} f(x)e^{-j\frac{n\pi x}{L}}\,dx \qquad (5\text{-}18.1)$$

c_n and c_{-n} are related to a_n and b_n by equation 5-19.

$$c_n=\frac{a_n-jb_n}{2} \qquad c_{-n}=\frac{a_n+jb_n}{2} \qquad (5\text{-}19)$$

$$a_n=c_n+c_{-n} \qquad b_n=\frac{c_{-n}-c_n}{j} \qquad (5\text{-}19.1)$$

therefore the amplitude and phase in terms of c_n and c_{-n} is given by equations 5-20.

$$A = 2\sqrt{c_n c_{-n}} \tag{5-20}$$

$$\tan\phi = j\left(\frac{c_n + c_{-n}}{c_{-n} - c_n}\right) \tag{5-20.1}$$

Since the functions here are real functions, a_n and b_n are real numbers. Therefore, c_n and c_{-n} are complex conjugates.

The equations given above are all for one-dimensional functions. For two-dimensional functions, the real and complex Fourier series are defined in equations 5-21. The definitions of frequency component amplitude and phase are also given. These equations are for a function that has a period 2X in the x-direction and 2Y in the y-direction.

The definition of the two dimensional Fourier series is given by

$$f(x,y)=\sum_{p=o}^{\infty}\sum_{q=o}^{\infty}\left\{a_{pq}\cos\left(\frac{p\pi x}{X}\right)\cos\left(\frac{q\pi y}{Y}\right)+b_{pq}\sin\left(\frac{p\pi x}{X}\right)\cos\left(\frac{q\pi y}{Y}\right)\right.$$
$$\left.+c_{pq}\cos\left(\frac{p\pi x}{X}\right)\sin\left(\frac{q\pi y}{Y}\right)+d_{pq}\sin\left(\frac{p\pi x}{X}\right)\sin\left(\frac{q\pi y}{Y}\right)\right\} \tag{5-21}$$

where,

$$a_{pq}=\frac{1}{XY}\int_0^{2X}\int_0^{2Y} f(x,y)\cos\left(\frac{p\pi x}{X}\right)\cos\left(\frac{q\pi y}{Y}\right)dxdy \tag{5-21.1}$$

$$b_{pq}=\frac{1}{XY}\int_0^{2X}\int_0^{2Y} f(x,y)\sin\left(\frac{p\pi x}{X}\right)\cos\left(\frac{q\pi y}{Y}\right)dxdy \tag{5-21.2}$$

$$c_{pq}=\frac{1}{XY}\int_0^{2X}\int_0^{2Y} f(x,y)\cos\frac{p\pi x}{X}\sin\frac{q\pi y}{Y}\,dxdy \qquad (5\text{-}21.3)$$

$$d_{pq}=\frac{1}{XY}\int_0^{2X}\int_0^{2Y} f(x,y)\sin\frac{p\pi x}{X}\sin\frac{q\pi y}{Y}\,dxdy \qquad (5\text{-}21.4)$$

An alternate form for the terms of the series is given by

$$A_{pq}\sin\left[\pi\left(\frac{px}{X}+\frac{qy}{Y}\right)+\phi_{pq}\right]=A_{pq}\left\{\sin\left(\frac{p\pi x}{X}\right)+\phi_{pq}\cos\left(\frac{q\pi y}{Y}\right)\right.$$
$$\left.+\cos\left(\frac{p\pi x}{X}+\phi_{pq}\right)\left(\sin\frac{q\pi y}{Y}\right)\right\} \qquad (5\text{-}21.5)$$

$$=A_{pq}\left\{\left[\sin\left(\frac{p\pi x}{X}\right)\cos\phi_{pq}+\cos\left(\frac{p\pi x}{X}\right)\sin\phi_{pq}\right]\cos\left(\frac{q\pi y}{Y}\right)\right.$$
$$\left.+\left[\cos\left(\frac{p\pi x}{X}\right)\cos\phi_{pq}-\sin\left(\frac{p\pi x}{X}\right)\sin\phi_{pq}\right]\sin\left(\frac{q\pi y}{Y}\right)\right\} \qquad (2\text{-}21.6)$$

where the relationship between a_{pq}, b_{pq}, c_{pq}, d_{pq}, and A_{pq}, ϕ_{pq} is given by

$$a_{pq}=A_{pq}\sin\phi_{pq} \qquad (5\text{-}21.7)$$

$$b_{pq}=A_{pq}\cos\phi_{pq} \qquad (5\text{-}21.8)$$

$$c_{pq}=A_{pq}\cos\phi_{pq} \qquad (5\text{-}21.9)$$

$$d_{pq}=A_{pq}\sin\phi_{pq} \qquad (5\text{-}21.10)$$

$$A_{pq}=\sqrt{a_{pq}^2+b_{pq}^2}=\sqrt{c_{pq}^2+d_{pq}^2} \qquad (5\text{-}21.11)$$

$$\tan\phi_{pq}=\frac{a_{pq}}{b_{pq}}=\frac{-d_{pq}}{c_{pq}} \qquad (5\text{-}21.12)$$

The complex form of the two dimensional Fourier series is given by

$$f(x,y)=\sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty}c'_{nm}\,e^{j\left(\frac{n\pi x}{X}+\frac{m\pi y}{Y}\right)} \qquad (5\text{-}21.13)$$

where

$$c'_{nm}=\frac{1}{4XY}\int_0^{2X}\int_0^{2Y} f(x,y)e^{-j\left(\frac{n\pi x}{X}+\frac{m\pi y}{Y}\right)} \qquad (5\text{-}21.14)$$

The relationship between the real and complex coefficients is given by

$$c'_{nm}=\frac{\left(a_{nm}-d_{nm}\right)-j\left(b_{nm}+c_{nm}\right)}{4} \qquad (5\text{-}21.16)$$

From equation 5-21.14 complementing both sides gives

$$\bar{c}'_{nm}=c'_{-n-m} \qquad \text{for } f(x,y) \text{ real} \qquad (5\text{-}21.16)$$

Since $a_{nm}=-d_{nm}$ and $b_{nm}=c_{nm}$,

$$c'_{nm}=\frac{a_{nm}-jb_{nm}}{2} \qquad (5\text{-}21.17)$$

$$a_{nm}=c'_{nm}+c'_{-n-m} \qquad (5\text{-}21.18)$$

$$b_{nm}=\frac{c'_{-n-m}-c'_{nm}}{j} \qquad (5\text{-}21.19)$$

$$A_{nm}=2\sqrt{c'_{nm}c_{-n-m}} \qquad (5\text{-}21.20)$$

$$\tan\phi_{nm}=j\left(\frac{c'_{nm}+c'_{-n-m}}{c'_{-n-m}-c'_{nm}}\right) \qquad (5\text{-}21.21)$$

If $m=0$, then the equations for the two-dimensional case reduce to those for the one-dimensional case as they should.

It is of interest to consider the nature of two-dimensional spatial sinusoids as defined in equation 5-21.5. For each value of p and q, this equation describes what amounts to a one-dimensional sinusoid whose lines of constant phase are at some angle with respect to the x-y axis. Referring to figure 5-4, it is of interest to determine the period T

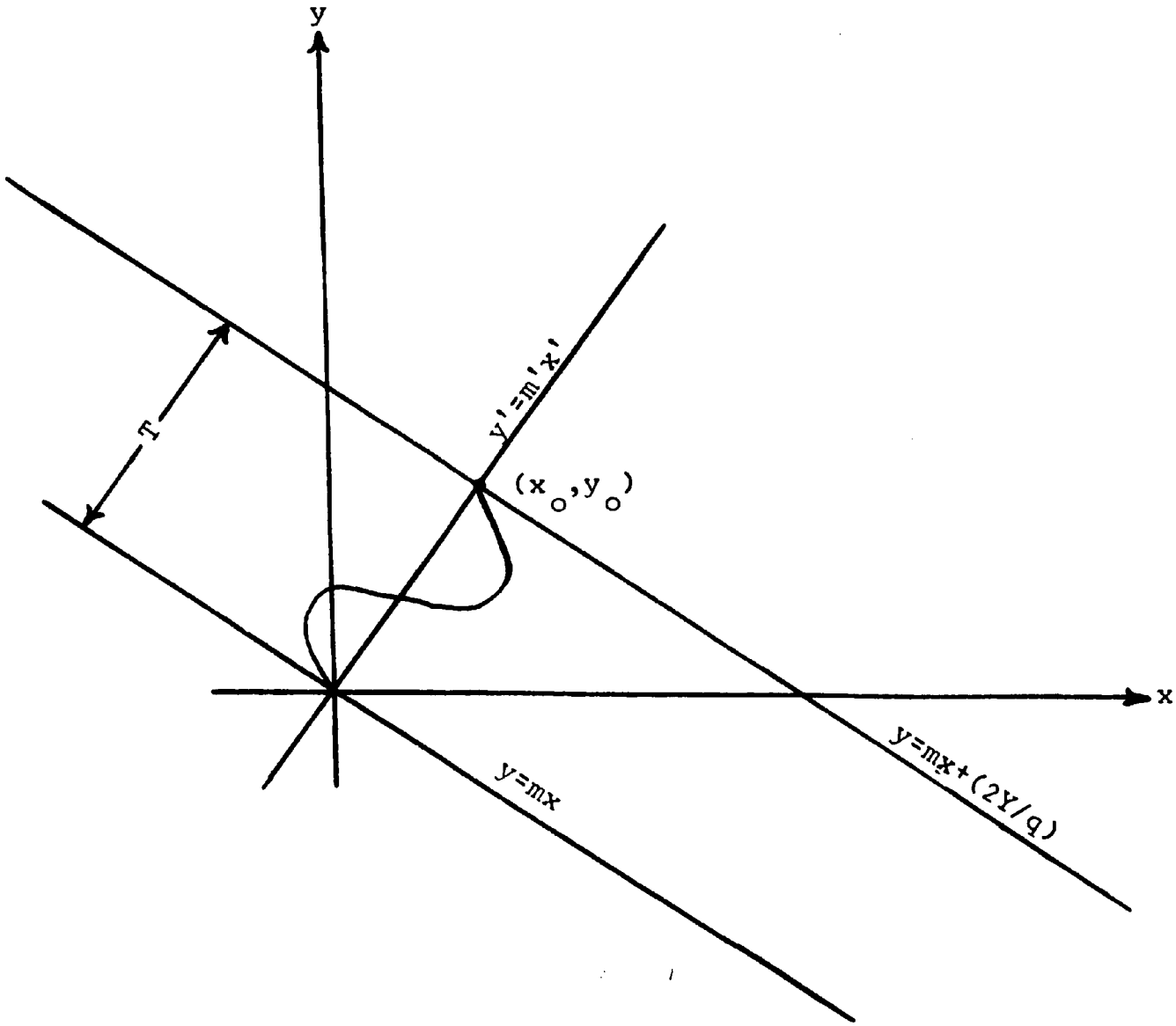


Figure 5-4. Slope and period of two-dimensional sinusoids.

and the slope m of the lines of constant phase. From 5-21.5, the line of constant phase passing through the origin is given by

$$\frac{px}{X} + \frac{qy}{Y} = 0 \qquad (5\text{-}22)$$

Therefore,

$$m = \frac{-pY}{qX} \qquad (5\text{-}22.1)$$

The line of constant phase which is exactly one period away from the line given by 5-22 is given by

$$\frac{px}{X} + \frac{qy}{Y} = 2 \qquad (5\text{-}22.2)$$

The line perpendicular to 5-22.2 and passing through the origin has a slope of -1/m and is given by

$$y' = \frac{qX}{pY} x' \qquad (5\text{-}22.3)$$

the distance between the origin and the intersection (x_o, y_o) of 5-22.2 and 5-23.3 is the period T,

$$x_o = \frac{2\left(\frac{p}{X}\right)}{\left(\frac{p}{X}\right)^2 + \left(\frac{q}{Y}\right)^2} \qquad y_o = \frac{2\left(\frac{q}{Y}\right)}{\left(\frac{p}{X}\right)^2 + \left(\frac{q}{Y}\right)^2} \qquad (5\text{-}22.4)$$

$$T_{pq} = \sqrt{x_o^2 + y_o^2} = \frac{1}{\sqrt{\left(\frac{p}{2X}\right)^2 + \left(\frac{q}{2Y}\right)^2}} \qquad (5\text{-}22.5)$$

When the period of the fundamental is the same in the two dimensions, e.g. $2 \cdot X = 2 \cdot Y$, then the frequency is given by

$$f_{pq} = {}^{1}/_{T_{pq}} = \frac{1}{2X}\sqrt{p^2 + q^2} \qquad (5\text{-}22.6)$$

Therefore, the Fourier series representation of a two-dimensional function consists of the superposition of an infinite number of one-dimensional sinusoids, each having a unique and discrete slope of constant phase as given in equation 5-22.1.

One difficulty in treating the picture luminance function as a two-dimensional periodic function, is the discontinuities introduced at the edge of a luminance function. In general, L(0) is not equal to L(2X). Therefore, in the periodic luminance function between the end of one cycle and the beginning of the next, there is an abrupt change in the value of the luminance function. This has the effect of introducing some large amplitude high spatial frequency components that might not otherwise be present. In the case of the actual picture, it is certainly possible to have luminance discontinuities near the edge of the picture. These discontinuities could be due to a picture frame or background which has a luminance quite different from the luminance of the picture at its edge. In general, these natural discontinuities would not be accurately represented by a periodic continuation of the picture. Of course, the luminance function used to determine one cycle of the periodic function could include not only the luminance of the picture, but also the luminance of a portion of the region surrounding the picture. The eye has a limited field of view in terms of its ability to perceive detail. Therefore, the entire surround could perhaps approach a single mean value of

of luminance. The periodic extension of such a luminance function would not introduce artificial discontinuities and the associated high spatial frequency components.

Another way to treat the picture for purposes of determining its spatial frequency content, is to treat the picture as a luminance impulse on an infinite background of uniform luminance. If the extent of the uniform surround included in the picture luminance function is increased without limit, what results is a luminance impulse. It is of interest to know what effect the inclusion of increasing areas of the surround has on the amplitude of the individual spatial frequency components. If increasing the period of the luminance function toward infinity has the effect of changing the contrast of the individual spatial frequency components, and if the multiple channel hypothesis is correct, changing the contrast would affect the visibility of these components. The problem can be analyzed in terms of the periodic pulse shown in figure 5-5 (reference 543). The complex Fourier series components c_n for this periodic pulse are given by equation 5-23.

$$c_n = \frac{Ea}{2L} \frac{\sin \frac{n\pi a}{2L}}{\frac{n\pi a}{2L}} \qquad (5\text{-}23)$$

The magnitude of c_n is shown in figure 5-6. The envelope containing the c_n and its maximum value is also shown there. Now we can observe what happens when a and E are held constant and the period 2L is increased. There are two things that happen. The maximum value of the envelope decreases and in

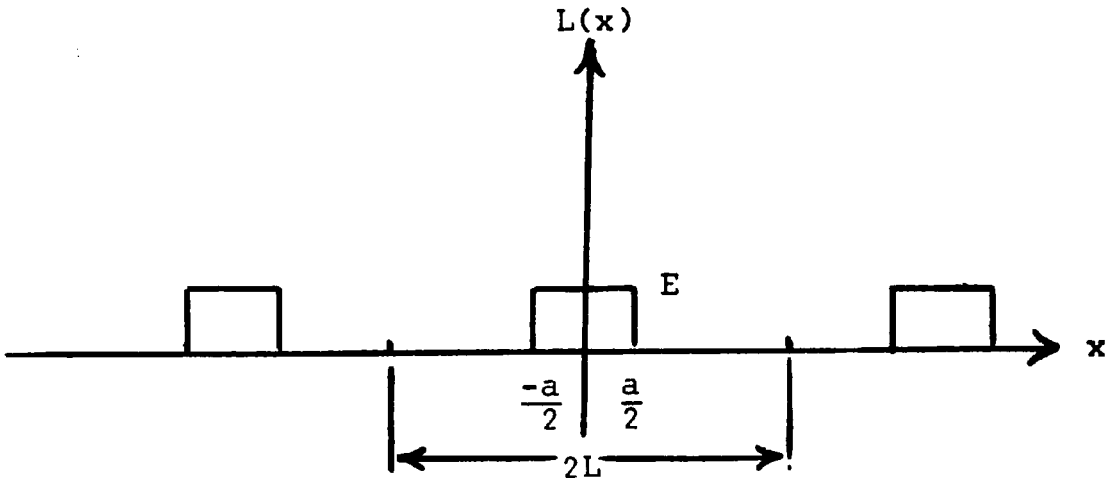


Figure 5-5. A periodic pulse of duration a and period 2L.

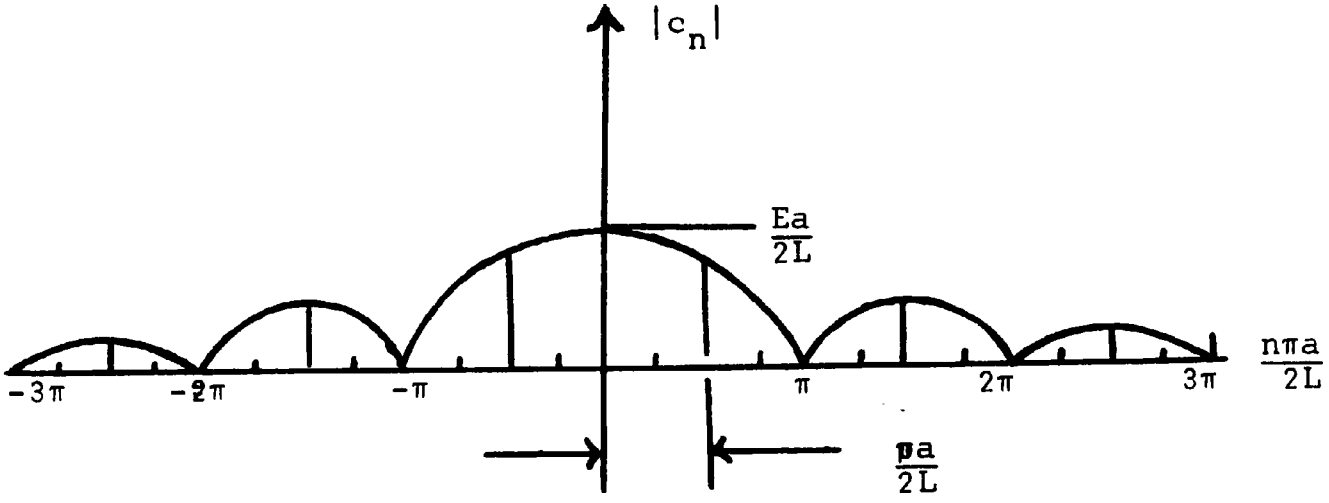


Figure 5-6. $|c_n|$ and the envelope of $|c_n|$

the limit approaches zero. Also the spacing of the spatial frequency components descreases and more components are contained in each lobe of the envelope. It can be concluded then that increasing the period does decrease the contrast of the individual spatial frequency components.

In the limit of infinite period, what results from this process is the continuous Fourier transform for the impulse. The definition of the continuous Fourier transform is given in equation 5-24. The two-dimensional continuous Fourier transform is given in equation 5-25.

$$g(w)=\frac{1}{2}\int_{-\infty}^{\infty} f(x)\, e^{-jwx}dx \qquad (5\text{-}24)$$

$$g(u,v)=\frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, e^{-j(ux+vy)}dxdy \qquad (5\text{-}25)$$

The frequency spectrum of an impulse as determined by the Fourier transform is a continuous function. In this case, it is not possible to determine the amplitude of a discrete frequency component since, in general, from a mathematical standpoint, individual components have zero amplitude. Therefore, for a function which has a continuous Fourier transform, it is not meaningful to determine the contrast of discrete frequency components.

However, the contrast of individual spatial frequency components is only relevant if the components are so widely separated that each individual component lies in a separate spatial frequency channel. As the spacing between spatial frequency components decreases, more and more components

occupy a given spatial frequency channel, and it is the response of the threshold mechanism to the composite output of each channel that becomes relevant. In the work that follows, it will be assumed that each spatial frequency component lies in a separate channel.

There are both digital and analog picture bandwidth compression techniques, as well as techniques which are combinations of the two. The perceptual model and its effect on picture bandwidth compression was studied by means of computer simulation. Therefore, the pictures used in the simulation will be in digital form, and a digital compression technique will be used. This also implies that the spatial frequency analysis will also have to be performed digitally. There are digital techniques available for the computation of the Fourier series and computation of the discrete Fourier transform. The discrete Fourier transform is really equivalent to the digital complex Fourier series, and is therefore a representation of the periodic extension of the digital function. Calculation of the discrete Fourier transform has become very convenient and efficient in recent years since the introduction of the various fast Fourier transform algorithms (references 5-4, 5-5, 5-6, 5-7). The discrete Fourier transform for one and two dimensions is given in equations 5-26 and 5-27.

$$\hat{c}_n = \frac{1}{N} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} f\left(\frac{2kL}{N}\right) e^{-j\frac{2\pi kn}{N}} \qquad (5\text{-}26)$$

for $n=-\frac{N}{2}, -\frac{N}{2}+1, \ldots, -1, 0, 1, \ldots, \frac{N}{2}-1$

$$\hat{c}_{pq}=\frac{1}{NM}\sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1}\sum_{l=-\frac{M}{2}}^{\frac{M}{2}-1} f\left(\frac{2kX}{N},\frac{2lY}{M}\right)e^{-j2\pi\left(\frac{pk}{N}+\frac{ql}{M}\right)} \qquad (5\text{-}27)$$

for $p=-\frac{N}{2}, \ldots, -1, 0, 1, \ldots, \frac{N}{2}-1$

$q=-\frac{M}{2}, \ldots, -1, 0, 1, \ldots, \frac{M}{2}-1$

With reference to 5-27, the function f is assumed to have dimensions 2X in the x-direction and 2Y in the y-direction. The function is sampled and consists of m uniformly spaced rows, 2Y/M units of length apart, and each row consists of N uniformly spaced samples, 2X/N units of length apart. Thus the picture is a regular M by N rectangular array of samples. The spatial frequency corresponding to each value of p and q is given by equations 5-22.5. The magnitude of each spatial frequency component is given by equation 5-21.20, and the phase is given by equation 5-21.21.

5.1.3 A basis for the evaluation of picture compression efficiency

When a picture is digitized for processing or transmission, the digitization parameters must be chosen so that when the picture is reconstructed, the degradation relative to the original picture is minimized. This implies that some kind of subjective or objective fidelity criterion is being used. It is possible to establish some objective fidelity criteria based on the characteristics of human vision.

The two parameters of interest in the digitization of pictures are sampling density and the discretization of gray scale. It is possible to take the view that the sampling parameters should be based purely on the resolution limits of the picture media and/or the resolution limits of the digitization and reconstruction equipment. On the basis of such a criterion, the sampling density might be based on the grain size of the film emulsion or the resolution limit of the digitizer optics. The gray scale quantization would probably be based on the signal-to-noise ratio of the digitizer photoelectric circuitry and analog to digital converter.

Another way to view the choice of the picture digitization parameters is to say that if the picture viewing environment is known beforehand, then an economic choice of parameters should be based on the characteristics of the eye for that environment.

This is most certainly the kind of criterion that was used to choose the parameters associated with the picture constructed by a television receiver (reference 3-36). The frame rate of 30 frames per second is based on the critical flicker frequency characteristics of the eye. A TV picture frame is made up of 525 interlaced lines. If a standard viewing distance of 4 times picture height is assumed, then the maximum spatial frequency that can be present in the picture is 18.4 cycles per degree. If we assume a cathode ray tube with a P4 phosphor (white), the

luminance of the TV picture will be between 5 and 10 foot-lamberts (reference 5-8). Shade (reference 3-36) found that for this range of luminances, the contrast sensitivity of the eye to an 18.4 cycle per degree vertical spatial sinusoid was 20% of its maximum contrast sensitivity. Now consider the maximum spatial frequency present in the horizontal direction. The bandwidth set aside for the picture luminance signal is 4 megacycles. Therefore, the period associated with the highest allowable frequency in the luminance signal is 0.25 microseconds. Since the entire 525 line raster is traced every 30th of a second, each raster line requires 63.5 microseconds. Allowing for the retrace time, the time to produce a raster line is 56.5 microseconds. This gives a maximum horizontal spatial frequency of 226 cycles in one screen width (reference 5-9). Assuming a 5:4 aspect ratio, this gives a maximum spatial frequency in the horizontal direction of 19.7 cycles per degree, a value not too different from the maximum vertical spatial frequency.

Consider now how the picture digitization parameters might be chosen in accordance with the picture viewing conditions. Let all other parameters of a viewing environment be held constant and only consider the effect of variable mean luminance. For each value of mean luminance, there are two psychophysical, visual characteristics which are of interest. These are incremental brightness sensitivity and sine-wave contrast sensitivity. Measurements of incremental

brightness sensitivity were described in section 3.4.1.1. It was shown there that increment threshold varies considerably depending upon the parameters of the experiment. In particular, some results were given there for white light and natural vision. In general, incremental brightness sensitivity increases with the duration of observation and with the size of the object. Figure 3-14 indicates that for long observation times, the incremental brightness sensitivity approaches an asymptotic value as the object size is increased. Therefore, it is possible to set an upper bound on contrast sensitivity for a given value of mean luminance. Figures 3-14 indicates that for mean luminance values between 1 and 1,000 foot lamberts, the log of the luminal contrast is about -2.5. Therefore, the luminal brightness increment is equal to 0.316% of the mean luminance. In a typical picture viewing environment, it was shown that the luminance measured at the eye can vary between 0.8 and 38.5 foot lamberts. The average luminance would be about 20 foot-lamberts. If this range of luminance is not quantized properly, the eye might see an edge between large areas of slightly different luminance. The incremental brightness sensitivity measurements tell us that if the luminance difference between the two areas is less than 0.316% relative to either of the two luminance values, then the eye will not see an edge. According to Figure 3-14, the incremental brightness sensitivity can be assumed constant at 0.316% for the luminance range 0.8 to 38.5 foot lamberts. In equation 5-28, it is shown that the relationship between

the luminance levels and the quantization levels is logarithmic if the incremental brightness sensitivity is a constant.

$$\frac{L_i-L_{i-1}}{L_i} = \frac{\Delta L_i}{L_i} = C = a(Q_i-Q_{i-1})=a\,\Delta Q_i \qquad (5\text{-}28)$$

$$\frac{\Delta L}{L} = a\Delta Q;\ Q\cong\frac{1}{a}\log L+b \qquad (5\text{-}28.1)$$

where L_{i-1} is a given luminance level, L_i is the liminal luminance level relative to L_{i-1}. Q_{i-1} is the quantization level assigned to L_{i-1} and Q_i is the quantization level assigned to L_i. The range from 0.8 to 38.5 foot lamberts would require approximately 1,230 quantization levels. This is shown in equations 5-29.

$$0.8\ (1+0.00316)^n=38.5 \qquad (5\text{-}29)$$

$$(1+0.00316)=\sqrt[n]{48.2} \qquad (5.29.1)$$

$$\log(1.00316)=\frac{1}{n}\log(48.2) \qquad (5.29.2)$$

$$0.00137n=1.683 \qquad (5\text{-}29.3)$$

$$n\cong1230 \qquad (5\text{-}29.4)$$

If linear quantization about a mean luminance value of 20 foot-lamberts were used, then equation 5-30 shows that about 600 quantization levels would result. It the quantization levels were chosen linearly, relative to 0.8 foot-lamberts then 14,900 quantization levels would result.

$$0.8+\bar{L}(0.00316)n=38.5 \qquad (5\text{-}30)$$

$$n\cong600 \text{ for } \bar{L}=20 \qquad (5\text{-}30.1)$$

$$n\cong14900 \text{ for } \bar{L}=0.8 \qquad (5\text{-}30.2)$$

At this point it is interesting to note that it has become standard to use 128 or 256 quantization levels. According to equations 5-31, these quantization levels correspond to log liminal contrasts of -1.51 and -1.82 respectively.

$$(1+x)^{y}=48.2 \qquad (5\text{-}31)$$

$$\log x=-1.51,\ x=0.0307 \text{ for } y=128 \qquad (5\text{-}31.1)$$

$$\log x=-1.82,\ x=0.152 \text{ for } y=256 \qquad (5\text{-}31.2)$$

From figure 3-14, these values would correspond to an assumed mean luminance of between 10^{-2} and 10^{-3} foot-lamberts. Either the luminance range has been overestimated, or the false contour problem is less critical, because pictures of good quality can generally be obtained with 256 linear quantization levels. Another explanation is that the average critical size of objects that produce false contours is small. Using the results given in equation 5-31.2 and an average luminance of 20 foot-lamberts in figure 3-14 gives a critical size of about 5 minutes of visual angle.

The other digitization parameter, which can be chosen on the basis of viewing conditions, is sampling density. If it is assumed that contrast sensitivity measurements are indicative of the bandpass characteristics of the spatial frequency channels, then it is possible to estimate the limiting high frequency characteristics of the eye. In section 3.4.1.4, the various conditions were described under which contrast sensitivity measurements have been obtained. Of particular interest were measurements made using white

light and binocular natural vision. For these conditions, Shade (reference 3-36) provides us with threshold contrast sensitivity measurements for a range of luminance values from 0.00044—400 foot-lamberts. The viewing distance was four times picture height. DePalma and Lowry (reference 3-37) measured the threshold contrast sensitivity for different distances of observation. The changing modulation transfer function of the eye's optics due to lens accommodation is provided by these measurements. The luminance used by DePalma and Lowry was 20 foot-lamberts. For each value of viewing distance, DePalma and Lowry extrapolated the contrast sensitivity curve to unity contrast. This is a measure of the sine-wave resolution limit for human vision. All spatial frequencies above the spatial frequency corresponding to unity contrast would require a contrast greater than unity to be perceptible. Unity contrast implies the maximum possible luminance deviations about the mean luminance. Therefore, spatial frequencies above the resolution limit are imperceptible. Table 5-5 lists the spatial frequencies at the resolution limit for each viewing distance. Shade's results for a mean luminance of 40 foot-lamberts imply an approximate resolution limit of 148 cycles per millimeter on the retina. Patel (reference 3-41) also measured sine-wave contrast thresholds for different values of retinal illuminance. His values are not for a natural pupil, however, since a two millimeter artificial pupil was used. For a mean retinal illuminance of 622 trolands, Patel set

Viewing distance (inches)	Limit of resolution for sine-wave test objects in cycles or lines per millimeter on the retina
10	155
14	193
35	215
56	225
120	250

Table 5-5. Effect of viewing distance on the limit of resolution, constant luminance, 20 ft-L. (reference 3-37).

an upper bound on the resolution limit at 235 cycles per millimeter. This was for a retinal illuminance of 100,000 trolands. These measurements indicate that the resolution limit lies somewhere in the range 148-250 cycles per millimeter on the retina. In cycles per degree, that range is 43.9—74.2.

These results can be related to sampling density in the following way: let the number of samples given in a vertical column of a digitized picture be M. If the sampling density has been chosen in accordance with the Nyquist criterion, then the sampling density is equal to two times the highest spatial frequency present in the original picture. However, if the original picture were a physical luminance distribution from a real life scene, the spatial frequency bandwidth could essentially be unbounded. If the original picture were derived from a 35 mm transparency, the resolution of the film itself could be anywhere from 50—120 lines/mm. The range of sampling densities implied by this range of resolution is 100—240 samples per millimeter. Each row or column of samples across the 35 mm length would therefore contain 3500-8400 samples.

Now consider an alternate approach in which it is assumed that the picture will be viewed from a distance equal to four times the picture height and in an environment in which the mean luminance measured at the eye is 20 foot-lamberts. For these conditions, it was determined that the resolution limit of the eye ranged between 44 and 74 cycles per degree. This implies a sampling density somewhere

between 88 and 148 samples per degree of visual angle. Since the picture would subtend a visual angle of 14.25 degrees, any row or column of samples in the picture would have between 1250 and 2100 samples. This is to be compared with the common use of a 256 X 256 sampling raster which, for the same viewing distance, implies a maximum spatial frequency of 9 cycles per degree. Shade's measurements indicate that the response of the eye to 9 cycles per degree is about 35% below its maximum sensitivity. If the bandwidth of the eye were based on its half power point, the spatial frequency where the contrast sensitivity is 0.707 of the maximum contrast sensitivity, then 9 cycles per degree would be a reasonable choice.

There appears to be a discrepancy between current practice in picture digitization and the measured psychophysical characteristics of human vision. This is particularly true in regard to the selection of quantization levels. Nevertheless, given the viewing environment, it is possible to set upper bounds on the number of quantization levels and the sampling density. If the sampling density exceeds that which would be implied by the resolution limit of the eye, based on luminance and viewing distance, then the picture has been over-sampled. In addition, it would probably be erroneous to base a measure of compression efficiency on the number of samples in the oversampled picture.

There appears to be an interrelationshp between the choice of sampling density and the choice of brightness quantization levels. Consider the following "worst case

picture", a picture which consists of one-half frame of one brightness quantization level and the other half frame consisting of the next higher, or next lower, quantization level. If the picture has been sampled to match the resolution limits of the eye for the viewing environment, then it is possible to say how many degrees of visual angle are subtended by each half frame. If we assume a resolution limit for the eye of 59 cycles per degree, and assume that the picture contains 256 X 256 samples, then a half-frame is 128 samples wide. This would correspond to 64 cycles of the highest spatial frequency, or approximately one degree of visual angle. For a background luminance of 20 foot-lamberts and an object subtending a visual angle of one degree, Blackwell's data gives a liminal contrast of 0.00315. Now consider the case where the region of the picture represented by one of the quantization levels occupies one-tenth of the picture width and the region at the other quantization level occupies the other nine-tenths of the width. For this configuration, the liminal contrast becomes 0.0048. In the limit of only one sample being one quantization level different from the rest of the picture, the liminal contrast is 0.78.

If the sampling density were chosen on some other basis, such as the half-power point of contrast sensitivity, then the visual angle subtended in each of the above cases would be different, causing the liminal contrast values to change. This, in turn, would change the required number of quantization levels.

It should be pointed out that it has been decided here to use Blackwell's liminal contrast data for 15 second observation times and localized targets. If, instead, Blackwell's data for 6 second observation times had been used, the liminal contrast values stated in every case would be lower by about one-half log unit.

5.1.4 The digitized pictures used in the computer simulation

Two separate sets of digitized pictures, distinguished on the basis of the number of quantization levels, were used in the computer simulation of the perceptual model. The first set of pictures used 64 levels of linearly quantized gray scale.

(1) a picture of the head and shoulders of a man containing 128 X 128 samples which will be given the title "Programmer",

(2) a picture of Rodin's "Thinker". This picture contained 256 X 256 samples,

(3) a picture of some solid geometrical "objects"-- a sphere, a cube, and a tetrahedron. This picture also contained 256 X 256 samples.

A second set of pictures of much higher quality was later obtained. This set of pictures contains 256 linearly quantized levels of gray scale. All three of these pictures contain 256 X 256 samples.

(1) a picture of the "face" of a young woman,

(2) a full length picture of a man using a movie camera on a tripod in an outdoor setting entitled "camera man",

(3) a picture of a "crowd" of young people sitting outdoors on the grass.

This second set of pictures is reproduced in appendix A. It can be seen that the pictures within this set represent a diversity of visual stimuli.

5.2 A simplified model of the eye

The perceptual model used in the computer simulation is shown in figure 5-7. This model treats the eye as a spatial frequency filter. The model will be used to determine the spatial frequency components in the original luminance distribution which are imperceptible. The basis for the determination of the perceptibility of spatial frequency components will be the contrast sensitivity function for the retina-brain portion of the visual system.

The contrast sensitivity function for the retina-brain can be thought of as a simplified expression of the multiple channel hypothesis. The assumption upon which this simplification is based is that all spatial frequency components lie in separate channels. For such a model, the spatial frequency component in each channel first passes through the filter mechanism and then through the threshold mechanism. The filter attenuates (or amplifies) the spatial frequency component based on the value of the modulation transfer function for that channel. The threshold mechanism permits the spatial frequency component to pass through to the output of the channel if its amplitude is sufficient for detection. It is difficult to determine the exact nature of the threshold mechanism. However, the multiple channel hypothesis presupposes that

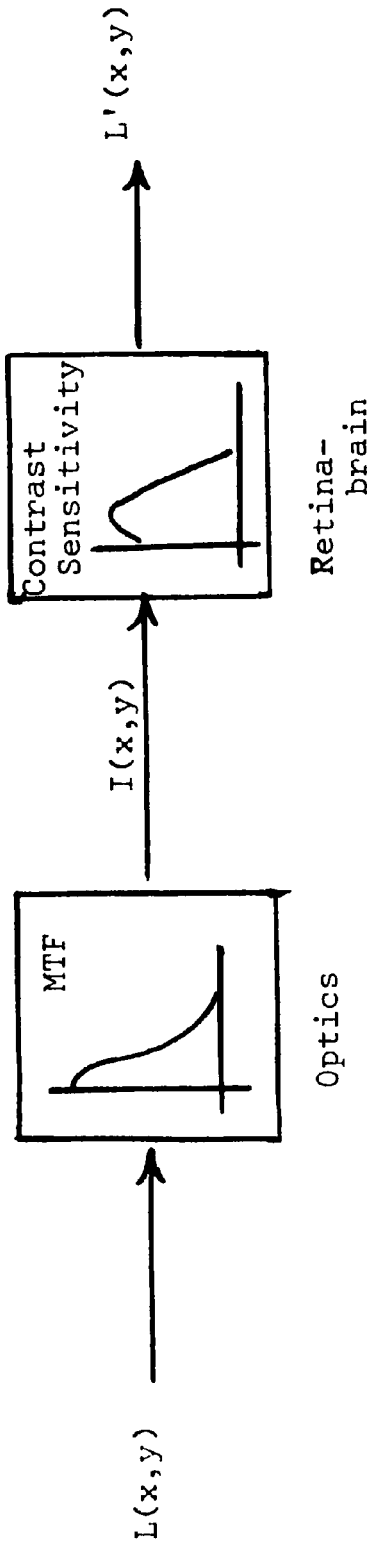


Figure 5-7. A model for the determination of imperceptible spatial frequency components,

the threshold mechanism is based on a contrast threshold. In such a case, it is possible to associate a contrast sensitivity with each channel. Since each channel is assumed to be sensitive to a unique spatial frequency component, the contrast sensitivities of all channels can be combined into a composite contrast sensitivity function. The net effect of the spatial frequency channels is to eliminate spatial frequency components which have insufficient contrast to be perceptible.

The model in Fig. 5-7 would function in the following way: the picture luminance distribution L(x,y) would pass through the optics of the eye producing a retinal illuminance distribution I(x,y). The modulation transfer function for the optics is used to derive the illuminance distribution I from the luminance distribution L. The retinal illuminance distribution is now passed through the contrast sensitivity function of the retina-brain. This has the effect of removing imperceptible spatial frequency components and producing a new luminance distribution L'(x,y). L'(x,y) is composed of those spatial frequencies in L(x,y) which have been determined to be perceptible by the model.

The net effect of the optics is to decrease the contrast of spatial frequency components in accordance with the modulation transfer function of the optics. For spatial frequency f, let the contrast threshold measured at the retina be given by $t_r(f)$. The retina-brain contrast sensitivity curve is given by $1/t_r(f)$. Let M(f) be the modulation

transfer function for the optics. Now if a picture luminance distribution has a spatial frequency component of frequency f, and of contrast c_f, then the contrast of that spatial frequency component measured at the retina will be $M(f)c_f$ which at threshold will equal t(f). Therefore, at threshold the contrast of the spatial frequency component measured external to the eye will be t(f)/M(f). Thus, it is possible to combine the effects of the optics with the contrast sensitivity function for the retina-brain into an effective contrast sensitivity function for the composite visual system. Most of the contrast sensitivity measurements described in Chapter III are based on the threshold contrast of a sinusoidal luminance distribution measured external to the eye and are therefore measurements of the composite contrast sensitivity function. On the other hand, the method based on the synthesis of interference fringes directly on the retina measures the retina-brain contrast sensitivity function.

The computer simulation procedure illustrated in figure 5-8 is based on the use of the composite contrast sensitivity function as a spatial frequency filter. More specifically, the composite contrast sensitivity function is used as a threshold device sensitive to contrast. A spatial frequency component passes through the device unaltered if its contrast is above the contrast threshold for that spatial frequency component. Otherwise, that spatial frequency component is infinitely attenuated.

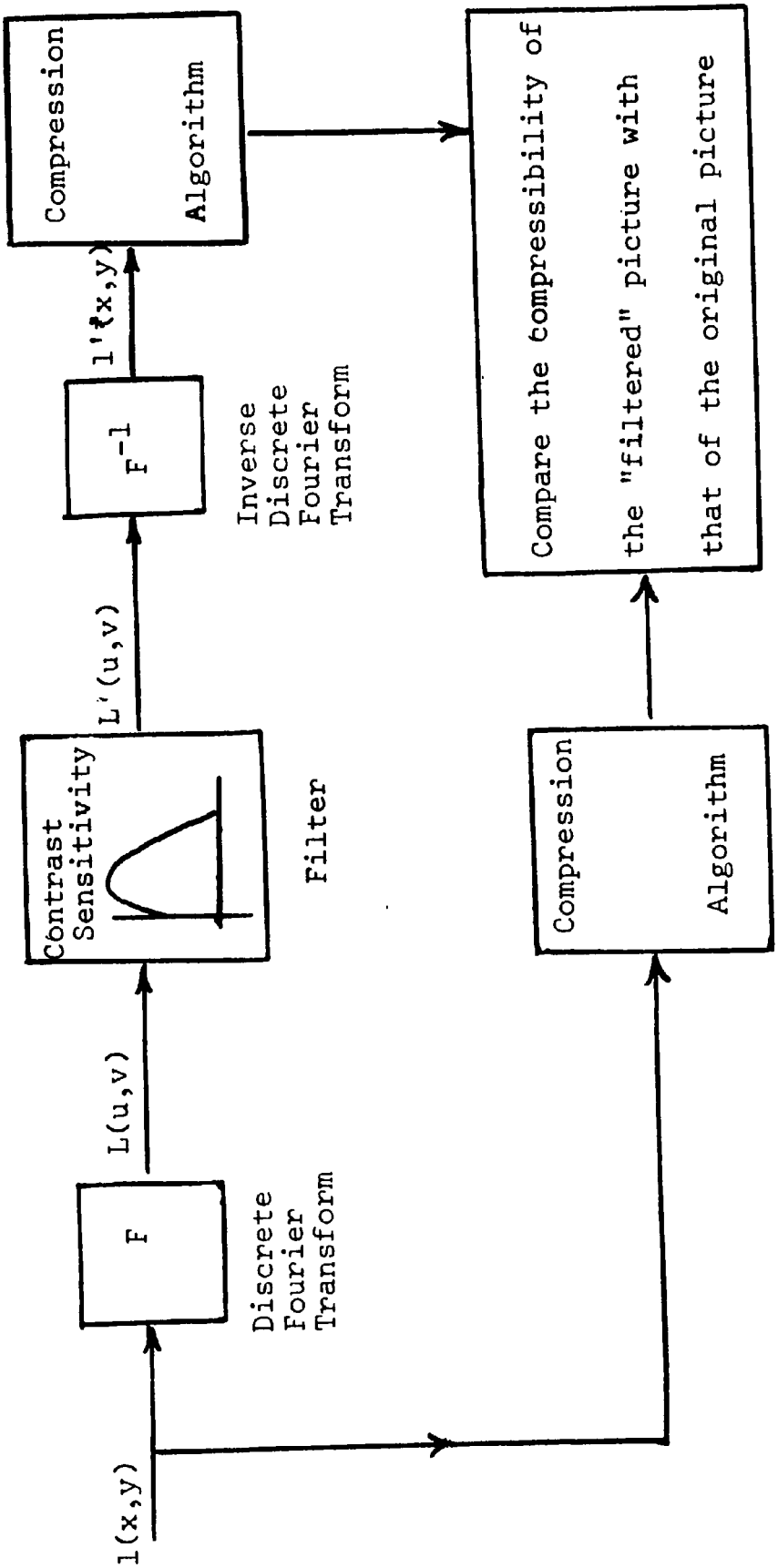


Figure 5-8 Computer Simulation Procedure

The purpose of the computer simulation is to determine if picture bandwidth compression can be improved by eliminating imperceptible spatial frequency components from the picture. The composite contrast sensitivity function is used as a basis for the elimination of spatial frequency components, and this, in turn, is based on a simplification of the multiple channel hypothesis. As shown in figure 5-8, the procedure begins by taking the discrete Fourier transform of the digitized picture. The result of this procedure is a set of harmonically related Fourier components. The magnitude and contrast of each spatial frequency component is then determined and compared against the contrast sensitivity function. Only those components which are above threshold are retained and inverse transformed to form a filtered version of the original picture. A compression algorithm is then used to compress the original picture and to compress the filtered picture. The compression results for the two cases are then compared to see if the elimination of spatial frequency components has had any effect. The comparison is made by comparing the bit compression ratios in the two cases. It is also of interest to consider the composite contrast sensitivity function as a criterion for the selection of spatial frequency components in Fourier domain picture coding.

The compression algorithm used in the computer simulation is a zero-order floating aperture interpolator. The FORTRAN program for the compression algorithm is given in Appendix B. The algorithm is a standard one, and can be used on

sampled data with integer or real values. For a given error tolerance Δ, the algorithm forms a single run from all consecutive data points which lie within a horizontal corridor of width 2Δ or less. The algorithm has been modified for two dimensional data by causing it to zigzag down the picture on a line by line basis. This method has a slight advantage over a strictly left to right line by line scan since samples one above the other at the ends of lines have a higher probability of equality than two samples at opposite ends of adjacent lines (see Appendix B).

The computer simulation will also be used to investigate the effects of viewing distance. When the viewing distance is changed, the angle subtended by the picture also changes. This in turn changes the spatial frequency value of each spatial frequency component of the picture. The magnitude and contrast of a given spatial frequency component does not change, but its value in cycles per degree changes. Since the position of a given spatial frequency component on the scale of spatial frequency changes with viewing distance, its contrast threshold will also change in accordance with the contrast sensitivity function for the eye. Therefore a spatial frequency component which may be visible at one viewing distance might become imperceptible at another viewing distance. The effect of viewing distance on perceptible spatial frequency content is investigated using the computer simulation model.

Simulated viewing distance is also investigated as a potential basis for the systematized elimination of spatial

frequency components for purposes of bandwidth compression in the Fourier domain.

As the viewing distance is changed, the portion of the picture that can occupy the high acuity foveal region of the retina changes. It is possible that the eye only performs spatial frequency analysis on the portion of the picture focused on the fovea, ignoring the remainder of the picture. This effect is incorporated into the model by processing the picture in segments where each segment approximately occupies one degree of visual angle based on the viewing distance.

5.2.1 One-dimensional simulation

The first method used in the computer simulation was to perform one-dimensional filtering on each line of the picture individually. Each picture line is transformed, filtered, and inverse transformed. After all lines of the picture have been filtered, the filtered picture is compressed and the results obtained are compared with those obtained by compressing the original picture.

For the computer simulation, it was necessary to determine an appropriate contrast sensitivity function. Based on the results contained in the papers listed in Table 3-1, a conservative contrast sensitivity function was synthesized. An attempt was made to use only contrast sensitivity results which were obtained under conditions similar to those

discussed earlier in this chapter. These conditions are characterized by an average luminance of about 20 foot-lamberts and a pupil diameter of about 3.4 millimeters. DePalma and Lowry (reference 3-37) used an average luminance of 20 foot-lamberts. The contrast sensitivity functions that they obtained had a minimum contrast threshold of about 0.002. This minimum occurred at about 3 cycles per degree. The maximum limit of resolution obtained was 73 cycles per degree. Campbell and Green (reference 3-14) used an average luminance of 29 foot-lamberts for a pupil diameter of 3.8 millimeters, they obtained a contrast sensitivity function with a minimum threshold contrast of 0.006. This minimum occurred at about 10 cycles per degree. The resolution limit was about 52 cycles per degree. Since the DePalma and Lowry results set a lower bound on contrast threshold than those of Campbell and Green, and also have a higher resolution limit, their results form the basis for a conservative estimate of a contrast threshold function. A functional approximation was determined which would adequately represent the contrast sensitivity function over a wide range of spatial frequencies. This functional approximation is given in Equations 5-31.

$$C(f)=533(0.258f^{0.788}) \text{ for } 0\leq f<5.6 \text{ cycles/degree} \quad (5\text{-}31)$$

$$C(f)=533 \text{ for } 5.6\leq f<10 \text{ cycles/degree} \quad (5\text{-}31.1)$$

$$C(f)=533(3.0(0.896)^{f}) \text{ for } 10\leq f<\infty \text{ cycles/degree} \quad (5\text{-}31.2)$$

In Figure 5-9, the contrast sensitivity function given by the functional approximation is compared with that obtained by DePalma and Lowry, and Campbell and Green. It can be seen that the contrast sensitivity function given by the functional approximation is at its maximum over a broader range of spatial frequencies, that range roughly beginning were the DePalma and Lowry results are at a maximum, and ending where the Campbell and Green results are at a maximum.

The Computer simulation program performs the following operations:

(1) Requests a compression error criterion and compresses the original picture according to that criterion.

(2) Requests a picture height/viewing distance ratio and computes the implied corresponding range of spatial frequencies that will be produced by the discrete Fourier transform.

(3) Computes the values of the contrast sensitivity function for the previously determined range of discrete frequencies.

(4) Calculates the discrete Fourier transform of the original picture and zeroes out the complex amplitude of those spatial frequency components which are below threshold.

(5) Inverse transforms the new set of discrete Fourier components to produce a filtered picture.

(6) Requests an error criterion (Δ) and compresses the filtered picture according to that criterion.

(7) Produces a graph on the line printer of sample brightness versus sample position for selected lines of the picture. Each graph of a single picture line consists of two superimposed graphs; one of the original picture line and another of the corresponding filtered picture line.

(8) Produces a pictorial representation of the Fourier domain making it possible to distinguish those Fourier components which are above threshold from those which are below threshold.

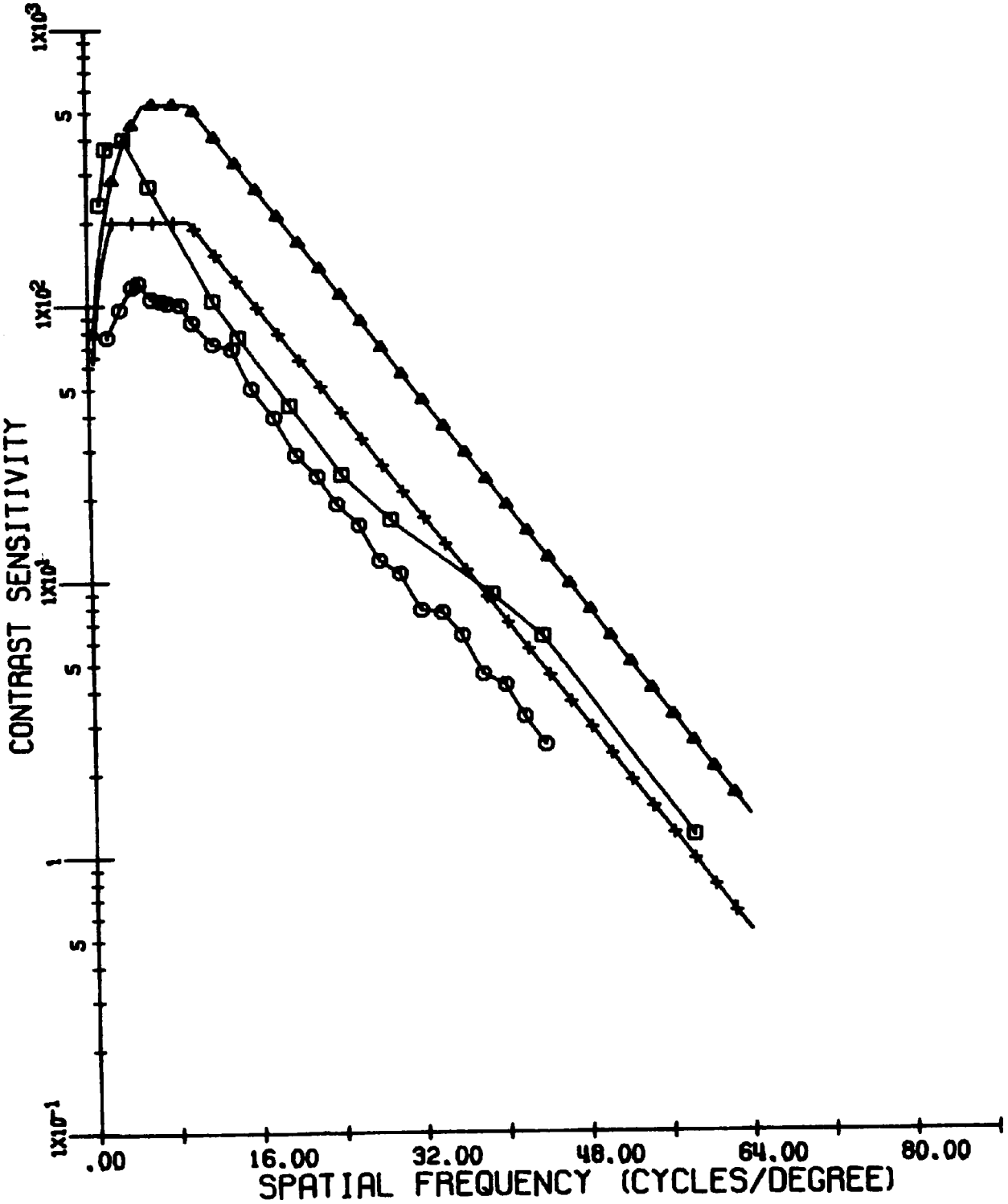


Figure 5-9 Composite sine-wave contrast sensitivity $1/C_t$ where $C_t=(B_{max}-B_{min})/(B_{max}+B_{min})$. Campbell and Green's (reference 3-14) measurements for 29.2 ft-L, 3.8 mm artificial pupil, artificial focus for 57 in. (circles). DePalma and Lowry's (reference 3-37) measurements for 20 ft-L, natural pupil, natural focus for 56 in. (squares). Functional approximation given by equations 5-31 (triangles) and 5-33 (crosses).

Therefore, two interesting features of the simulation program are the ability to simulate the effects of viewing distance and the ability to visually compare the effects of spatial frequency filtering on the appearance of a picture line. This latter capability also makes it possible to judge, in a given instance, whether or not the effects of filtering were intuitively reasonable.

In order to verify that the computer simulation program was operating properly, a program was written to produce sample sine waves of any desired sampling density, number of gray scale levels, average value, amplitude, frequency, and phase. Using sampled sine waves of different contrast, frequency, and phase, it was verified that the computer simulation program properly computed Fourier component amplitudes, calculated contrast properly, and correctly eliminated sub-threshold components.

5.2.1.1 Results for 64 levels of gray scale

Preliminary to the actual computer simulation, the picture "Programmer" was compressed in its original form. Statistics derived from the compression process made it possible to locate lines in the picture that were representative of extremes in picture detail. Two scan lines were selected near the top border of the picture which had a noise-like quality. The difference between the maximum and minimum brightness across both picture lines was about 5 quantization levels. Visually, they appeared uniformly bright.

A second pair of lines were chosen which contained what appeared to be a large amount of essential picture detail. These lines were also characterized by a broad range of brightness values. The maximum and minimum brightness across a line were separated by about 45 quantization levels. Both pairs of lines were about equally compressible in their original form. The noise-like pair had a sample compression ratio of 1.778 and a bit compression ratio 0.899. The pair containing picture detail had a sample compression ratio of 1.571 and a bit compression ratio of 0.875. In both cases, an error criterion of zero brightness units was used.

It was anticipated that two different types of results would be obtained when the filtering technique was applied to these two pairs of lines. Since the noise-like lines have a small dynamic range of brightness, it would be expected that these lines would be composed of spatial frequency components of low amplitude and contrast. Hopefully many of these components, especially those at high spatial frequencies would be below threshold and therefore would be eliminated. This would cause these lines to become more uniform, and, therefore, more compressible. On the other hand, the lines containing picture detail have a wide dynamic range of brightness, and therefore contain spatial frequency components of large amplitude and contrast. These components would not be eliminated by filtering and essentially all of the picture detail should be present after the filtering process.

The computer simulation program was applied to the two pairs of picture lines. In one respect, the anticipated results were obtained; that is, for the pair of lines containing picture detail, filtering did not change these lines at all. Sample point by sample point, the filtered lines were identical to the original lines. In terms of Fourier components, out of 128 Fourier components produced by the discrete Fourier transform, only eight components were filtered out by the simulation program. However, in the case of the noise-like picture lines, the results were only partially anticipated. As expected, more than half the Fourier components were filtered out. Fifty-four remained out of a total of 128. But, in comparing the original picture lines with the reproduced filtered lines, it was found that the filtered lines only partially resembled the original lines. In addition, the filtered lines were less compressible than the original lines. The major discrepancy between the original and the filtered lines seemed to be due to the loss of local average values. This can be explained by the fact that mostly low spatial frequencies were filtered out by the simulation program. The retention of mostly high spatial frequencies accounts for the lower compressibility.

5.2.1.2 Effects of different contrast sensitivity parameters

The viewing distance used to obtain the computer simulation results was four times picture height. For this

viewing distance and 128 X 128 sampling raster, the discrete Fourier transform produces spatial frequency components in the range 0.07-4.491 cycles per degree. This range of spatial frequencies lies on the portion of the contrast sensitivity function which is increasing toward maximum with increasing spatial frequency. Therefore, in this range, the threshold contrast starts out high for the low spatial frequency components and decreases with increasing spatial frequency. So, for a noise-like function, that is, one having equal amplitude spatial frequency components, it would be expected that the higher spatial frequency components would be retained. In fact, the Fourier coefficients for the noise-like picture lines are all very similar in amplitude except for a few low spatial frequency components. However, although these components were larger, they were not of sufficient amplitude to be retained.

It is the low spatial frequency portion of the contrast sensitivity function which is least accurately known. This could account for part of the problem. Another possible problem could be an excessively high estimate of the maximum contrast sensitivity.

The contrast of the first few low spatial frequency components for the noise-like lines ranged between 1/30 and 1/80 whereas most of the other spatial frequency components had a contrast of 1/200 or less. Therefore, if a lower value were used for the maximum contrast sensitivity (533 was used in the computer simulation), and if the contrast sensitivity for low spatial frequencies was increased,

then more favorable results would probably be obtained. Campbell and Green obtained a maximum contrast sensitivity of 200. Therefore, it is not unreasonable to use a minimum threshold contrast of 1/200. Davidson's measurements of small perturbation contrast sensitivity (reference 3-45) if extrapolated to 0.07 cycles per degree give a contrast sensitivity of 0.08 relative to the maximum contrast sensitivity. For a maximum contrast sensitivity of 200, this would give a contrast sensitivity at 0.07 cycles per degree of 16 which is not sufficiently low to capture some of the low spatial frequency components.

Equation 5-31 was modified to approximate Davidson's results for low spatial frequencies. Equations 5-31 have the general form given in Equations 5-32.

$$C(f)=MTFMAX\cdot(Af^{B}) \text{ for } 0\leq f<F1 \text{ cycles/degree} \qquad (5\text{-}32)$$

$$C(f)=MTFMAX \text{ for } F1\leq f<F2 \text{ cycles/degree} \qquad (5\text{-}32.1)$$

$$C(f)=MTFMAX\cdot(C\cdot D^{f}) \text{ for } F2\leq f<\infty \text{ cycles/degree} \qquad (5\text{-}32.2)$$

Equation 5-32 is a straight line approximation in log-log coordinates. The slope of this line is given by B. In Equation 5-31, B = 0.788. Davidson's contrast sensitivity for low spatial frequencies decreases less rapidly for decreasing spatial frequency. A straight line approximation for his results gives A = 0.535, B = 0.713, C = 3.0, D = 0.896, F1 = 2.4, F2 = 10.0. These parameters give the lower bound approximation to the human contrast sensitivity function shown in equations 5-33.

$C(f)=200(0.535f^{0.713})$ for $0 \leq f < 2.4$ cycles/degree (5.33)

$C(f)=200$ for $2.4 \leq f < 10$ cycles/degree (5-33.1)

$C(f)=200(3.0 \cdot 0.896^{f})$ for $10 \leq f < \infty$ cycles/degree (5-33.2)

The results were as expected. The pair of lines containing picture detail were only very slightly changed, and essentially all the spatial frequency components were above threshold. For the noise-like picture lines, low and high spatial frequency components disappeared; out of 128 spatial frequency components, 22 were above threshold. The reproduced filtered picture lines were not improved in compressibility and again bore little resemblance to the original noise-like picture lines.

If a maximum contrast sensitivity value of 400 were used with Davidson's results, then the contrast sensitivity at low spatial frequencies would be sufficiently high to retain a few low spatial frequency components. When the computer simulation was performed using this larger maximum contrast sensitivity value, filtering left all four picture lines essentially unchanged. As expected, for the noise-like lines, many low spatial frequency components were retained, but so were many high spatial frequency components. There was no indication that the filtering process tended to smooth out the noise-like lines and increase their compressibility.

The preceding results were generally inconclusive. For the noise-life picture lines, the results obtained were strongly dependent upon the parameters used for the contrast sensitivity function. The lines containing picture detail were insensitive to the contrast sensitivity parameters and filtering had a negligible effect.

5.2.2 One dimensional simulation using segments

At this point, it is pertinent to compare the method used in the simulation program to analyze a picture into its spatial frequency components with how the eye might perform this same function. In the simulation process, the entire line, regardless of its spatial extent, is used to determine spatial frequency content. It is unlikely that the eye would operate in the same way. The region of the eye that is able to perceive the full range of visible spatial frequencies is quite limited. In Chapter III, it was shown that the region of high visual acuity is restricted to the fovea which is the central one degree of vision. The eye cannot see an entire picture in detail all at one time. The eye actually scans the picture fixating on those points where the details are of interest. Therefore, each area of the picture is in effect seen through a window. This window does not have sharp edges, but gradually makes a transition from the region of high acuity to a region of low acuity. Therefore, it was decided to investigate the effect of processing the picture in segments where each segment subtends a visual angle of one degree. The one-dimensional discrete Fourier transform is used, but rather than processing an entire picture line as a unit, the picture line is divided into segments where each segment is approximately one degree of visual angle in width. For reasons of efficiency of the fast Fourier transform, the number of samples in each segment was adjusted to be the nearest power of 2.

5.2.2.1 Results for 64 levels of gray scale

The two pairs of lines from "programmer" were processed using one-dimensional segments subtending approximately one degree of visual angle. For a viewing distance of four times picture height, each picture line was divided up into sixteen segments. Each segment contained 8 samples and subtended a visual angle of about 0.89 degrees. The discrete Fourier transform of each segment produced a DC term and four spatial frequency components ranging in spatial frequency from 1.123-4.491 cycles per degree. Both contrast sensitivity functions were used in the computer simulation; the one described by Equations 5-31, where MTFMAX = 533 and the one given by equations 5-33 with MTFMAX = 200. In both cases, the filtered picture lines were identical to the original picture lines. The two contrast sensitivity functions used are representative of the upper and lower extremes of human contrast sensitivity. Therefore, for this method, the external appearance of the picture lines did not seem to be contrast-sensitivity dependent.

However, in the Fourier domain, the results obtained were different for the two contrast sensitivity functions. The contrast sensitivity function with the lower maximum value eliminated more Fourier components (44 versus 22). This is to be expected, but it is not expected that the appearance of the picture line would be unchanged by the elimination of additional Fourier components. Upon closer examination though, it is not too surprising in this case that the eliminated Fourier components do not affect the

appearance of the picture lines. For a viewing distance of four times picture height, the spatial frequencies are near to the maximum contrast sensitivity for the human eye. As a result, only spatial frequency components with a contrast of 1/116 or less were eliminated, but the minimum contrast between any two sample points for 64 levels of gray scale is 1/63. Therefore, the eliminated spatial frequency components had an amplitude of 1/2 quantization level or less. So it is not surprising that eliminating these spatial frequency components did not affect the appearance of the picture lines.

5.2.2.2 Effects of increased viewing distance

In order to observe the effects of the model when spatial frequency components of appreciable amplitude are eliminated, it is necessary to view the picture from a distance which results in spatial frequencies for which the eye has a low contrast sensitivity. For example, if the entire picture subtended a visual angle of one degree, then the resultant spatial frequencies would be in the range 1-64 cycles per degree, for a picture containing 128 samples per line. This range of spatial frequencies essentially covers the whole range of human contrast sensitivity. The computer simulation was applied to the two pairs of lines from the picture "programmer" for a viewing distance of 60 times picture height.

The qualitative effect that this viewing distance had on the picture lines was to smooth them. This is true of both the noise-like lines and the lines containing picture detail. Qualitatively the filtered picture lines were very similar to the original picture lines. In the case of the noise-like picture lines, filtering significantly increased the compressibility. The original pair of noise-like picture lines required 144 runs, giving a sample compression ratio of 1.778 and a bit compression ratio of 0.889. The filtered noise-like lines required 84 runs and had a sample compression ratio of 3.048 and a bit compression ratio of 1.524. These results were for MTFMAX = 533. For MFTMAX = 200, the filtered noise-like lines required only 57 runs. The sample compression ratio was 4.491 and the bit compression ratio was 2.246.

Filtering did not increase the compressibility of the lines containing picture detail. For this pair of lines, smoothing seemed to be confined to the central portion of the picture line. This smoothing was offset by an increased number of fluctuations near the edges, although this effect was not present for MTFMAX = 200. The increase in compressibility was very small. It is possible that the fluctuations near the edge are due to high spatial frequency components introduced by the discrete Fourier transform. As mentioned previously, these are caused by a difference between the brightness of the right and left edges of the picture. This difference causes an edge to occur when the picture is periodically extended into space.

A more dramatic result is obtained in the Fourier domain for a viewing distance of 60 times picture height. The original four picture lines produced 256 Fourier components. After filtering, 84 Fourier components were retained for MTFMAX = 533 and 60 Fourier components were retained for MTFMAX = 200.

The results obtained from the computer simulation for a large viewing distance are qualitatively what would be expected. The smoothing obtained was a manifestation of the eye's decreased ability with increased distance to discern fine detail. The basic characteristics of the picture lines were retained, and consequently, for lines containing a large amount of picture detail, compressibility was not increased. For the same level of picture detail and increased sampling density, further increases in compressibility would be expected.

5.2.2.3 Results for 256 levels of gray scale

The computer simulation results that have been described so far were obtained using a relatively low quality picture. It is of interest to determine the effects of increased sampling density and gray scale on the computer simulation results. For this purpose, two pairs of lines were selected from the picture called "face". The first pair of lines were selected from the top border of the picture which appears to be relatively uniform in brightness. The actual

gray scale values of this pair range between 66 and 142. The other pair of lines taken from the middle of the picture have a much wider dynamic range of brightness. For these lines, the numerical brightness ranges between 18 and 254. These pairs of lines were processed in segments subtending approximately one degree of visual angle.

5.2.2.4 Results for a viewing distance of four times picture height

The computer simulation program was applied to the four selected picture lines. A viewing distance of four times picture height was simulated. Therefore, each picture line was divided into sixteen segments, each segment containing sixteen samples. Each segment subtended a visual angle of 0.89 degrees. The discrete spatial frequencies produced for each segment ranged from 1.123 to 8.982 cycles per degree. The original four picture lines had a sample compression ratio of 1.364 and a bit compression ratio of 0.682.

First the simulation was performed using the parameters for contrast sensitivity which approximate an upper bound to the human contrast sensitivity function as given in Equations 5-31. The total number of Fourier components came to 512. This does not include the average value terms for each segment. 487 Fourier components were above threshold. The resulting filtered picture lines were virtually identical to the original picture lines. The original picture lines required 751 runs whereas the filtered lines required 758.

The sample compression ratio of the filtered lines was 1.351, and the bit compression ratio was 0.675. The mean squared error between the original and filtered picture lines was 0.037 quantization units squared.

The simulation was also performed for the lower bound approximation to the human contrast sensitivity function as given in Equations 5-33. Of the 512 Fourier components, 429 were retained. In this case, the differences between the original and the filtered picture lines are quite noticeable. Examining the differences on a point by point basis shows that essentially all the differences are on the order of one quantization level and are confined to a small percentage of the total number of sample points. In terms of compressibility, the original four picture lines required 751 runs. The filtered lines require 798 runs and have a sample compression ratio of 1.283 and a bit compression ratio of 0.642. Therefore, the compressibility is essentially unchanged. The mean squared error was 0.235.

Qualitatively, the results obtained for picture lines with 256 levels of gray scale are identical to those obtained for the picture lines with 64 levels of gray scale. Namely, for a viewing distance of four times picture height, no increase in compressibility is obtained. This is true for compression in the spatial domain and in the Fourier domain. The number of subthreshold Fourier components that are eliminated is insufficient to form the basis for a compression scheme in the Fourier domain. If the computer simulation

model is an accurate analog of human brightness perception, then the preceding conclusions imply that, for a viewing distance of four times picture height, the eye perceives all of the detail which can be represented using 256 samples and 256 levels of gray scale.

It is difficult to determine whether or not these conclusions are correct. In section 5.1.3, a case was made for a sampling raster at least 1250 X 1250. This was based on the resolution limit of human contrast sensitivity and a viewing distance of four times picture height. It was also shown there that under worst case circumstances 600 linear quantization levels could be justified based on the usual range of luminances encountered in a typical picture viewing environment. Therefore, on the basis of that analysis, it could be argued that it is not unreasonable for the eye to perceive all the detail representable in a 256 X 256 sampling array using 256 levels of gray scale. On the other hand, it is known that high quality pictures can be obtained if errors are intentionally introduced into the original picture for purposes of increased compressibility. But picture quality is not judged on the basis of a point by point comparison of the two pictures, but rather by comparing the two pictures as a whole. It is indeed possible that if the observer were told the specific samples that are different in the two pictures that he would then be able to perceive the difference, yet perception of these differences probably would not affect his judgment of picture quality.

5.2.2.5 The effects of increased viewing distance

For a viewing distance of four times picture height, the spatial frequencies for the picture line segments range between 1.123 and 8.982 cycles per degree. For this range of spatial frequencies, the contrast sensitivity of the human eye is at or near its highest value. Therefore, it would be expected that most of the spatial frequency components would be above threshold, and therefore, the filtered picture lines would be nearly identical to the original picture lines. Now results will be obtained for the 256 sample picture lines using a viewing distance that will produce a much broader range of spatial frequency components. The expected result will be that more hypothetically imperceptible spatial frequency components will be eliminated. Most of the eliminated spatial frequency components will be at the high end of the spatial frequency scale, and therefore, high frequency fluctuations in the picture will be smoothed out.

A viewing distance of 60 times picture height was used. This caused the entire picture to subtend a visual angle of 0.955 degrees. The four picture lines from the picture "face" were again used. For each picture line, 128 spatial frequency components were produced plus an average value term. These spatial frequencies ranged between 1.047 and 134.044 cycles per degree. This range is roughly double the contrast threshold for human vision. Therefore, nearly half the spatial frequency components would be eliminated at the outset.

First the simulation was performed using the upper bound approximation to the human contrast sensitivity function given in Equations 5-31. Qualitatively the filtered picture lines were a smoothed version of the original picture lines. Sharp details with a width smaller than four or five sample points and an amplitude of ten or more quantization levels were completely eliminated. Whereas, broader, lower amplitude detail was smoothed and decreased in amplitude. However, most of the basic characteristics of the original picture lines were preserved. The sample compression ratio was 1.298 and the bit compression ratio was 0.649. Comparable statistics for the original picture were 1.364 and 0.682 respectively. In the Fourier domain, 91 out of a total of 512 Fourier components were retained. The mean square error was 16.849.

Next, the lower bound approximation to the human contrast sensitivity function given by Equations 5-33 was used. As would be expected, there was a further increase in the loss of detail in portions of the picture lines containing high frequency fluctuations. But again, the basic characteristies of the picture lines were preserved. In this case, the number of runs required was 773. The sample compression ratio was 1.325 and the bit compression ratio was 0.662. In the Fourier domain, 65 out of a total of 512 Fourier components were retained. The mean square error was 23.342.

In summary, it can be said that the effects of simulated distance are those that would be expected; namely, a systematic loss of high frequency detail. There is one very apparent

difference between the results obtained with the 128 X 64 picture lines and those obtained with the 256 X 256 picture lines: increased viewing distance increased compressibility in the former case, whereas with the 256 X 256 picture lines, using a viewing distance of 60 times picture height did not aid compressibility in the spatial domain. With an increased number of quantization levels, the probability that two adjacent samples will have the same quantization level is decreased. Each quantization level in a 64 quantization level scheme is equal to four of the quantization levels in a 256 quantization level scheme. If the error criterion used to compress the 256 level lines was four times that used to compress the 64 level lines, compressibility differences might disappear. However, such a scheme would either make the 64 level lines incompressible, because the error criterion would be too small to allow all real picture values produced by the inverse Fourier transform to straddle some discrete level of gray scale, or it would result in a loss in effective gray scale for the 256 level lines because the error criterion would straddle several gray scale levels.

The choice of compression algorithm seems to be the main cause of the poor compressibility characteristics of the 256 X 256 pictures lines. The zero order interpolator was chosen for its simplicity and economy of computation. However, a more complex algorithm such as a first-order interpolator would probably yield better compression results. This is because increased gray scale resolution seemed to

produce more regions of constant slope than of constant gray scale. In addition, when simulated distance was increased, the length of constant slope regions seemed to also increase. The effect of the perceptual model on first order interpolation compression results would, in itself, be another subject worth investigating.

5.2.2.6 The use of simulated distance for compression in the Fourier domain

The results obtained from the one-dimensional simulation of the perceptual model indicate that compression using a zero-order interpolator is not enhanced by the elimination of hypothetically imperceptible spatial frequency components. However, some of the statistics obtained for the Fourier domain from the simulation results indicate that the model might be useful for compression in the Fourier domain. For a viewing distance of four times picture height, essentially all the Fourier components were retained, and consequently, there was no physical degradation of the original picture whatsoever. This was true for both the 128 X 64 picture lines and the 256 X 256 picture lines. When a viewing distance of 60 times picture height was used, a large percentage of the Fourier components were eliminated even when an upper bound approximation to human contrast sensitivity was used. For the 128 X 64 picture lines, 172 out of 256 components were eliminated. For the 256 X 256 picture lines, 421 out of 512 components were eliminated. In Chapter II, it was shown that systematic degradation of a

picture in both the spatial and Fourier domains enhanced compressibility without a serious loss of picture quality. Anderson and Huang (reference 2-29) showed that the Fourier domain could be run-length encoded using, on the average, between 4.78 and 5.6 bits per Fourier coefficient. Using their Fourier domain threshold technique, they were able to represent pictures with, on the average, between 0.928 and 1.583 bits per picture point. This is to be compared with 8 bits per picture point for the original picture. Therefore, they were able to achieve a bit compression ratio between 5.05 and 8.62.

Simulated distance in conjunction with the human contrast sensitivity function constitutes a form of systematic picture degradation. It is a unique method which can be used to eliminate any desired number of Fourier components. It is possible that this method yields a better quality picture for a given number of retained spatial frequency components, because the threshold for each spatial frequency component is weighted in accordance with the relative ability of the eye to perceive that spatial frequency.

5.2.2.7 Composite results for some entire pictures

In order to assess the potential of this method for Fourier domain compression, the three 256 X 256 X 256 pictures were processed for several values of simulated distance. The lower bound approximation for the human contrast sensitivity function was used in the simulation.

The simulated distances were chosen so that each picture line subtended visual angles of one degree, two degrees, four degrees, and eight degrees. This caused each picture line to be processed as one, two, four, and eight segments respectively. For each value of simulated distance the mean square error between the original and the filtered picture was computed, and a reproduction of the filtered picture was obtained. In all cases, the error criterion used in the zero-order interpolator compression algorithm was 0.5 brightness quantization units. The results obtained are summarized in Tables 5-6, 5-7, and 5-8.

Some of the captions used in these tables will now be defined. D/H is the ratio between simulated viewing distance and picture height. SCR is the sample compression ratio. This is the number of samples in the original picture divided by the number of samples in the compressed filtered picture. FCR is the Fourier compression ratio; this is the number of Fourier components in the original picture divided by the number of Fourier components which are above threshold in the filtered picture. The number of Fourier components in the original picture is based on the number of Fourier components produced by the discrete Fourier transform. This in turn is based on the number of segments required to represent each picture line. For viewing distances of 8, 16, 32, 64 times picture height, the total number of Fourier components in the original picture are 34816, 33792, 33280, and 33020 respectively. MSE is the mean squared error between the original and the filtered picture. FBCR

TABLE 5-6 ONE-DIMENSIONAL SIMULATION RESULTS FOR "FACE"

D/H	SCR	FCR	MSE	FBCR
8:1	1.301	1.123	0.534	1.403
16:1	1.144	1.649	11.916	2.135
32:1	1.171	3.29	21.528	4.279
64:1	1.246	6.33	29.843	8.361

TABLE 5-7 ONE-DIMENSIONAL SIMULATION RESULTS FOR "CAMERAMAN"

D/H	SCR	FCR	MSE	FBCR
8:1	1.387	1.295	0.874	1.623
16:1	1.206	1.747	10.239	2.246
32:1	1.208	3.02	44.068	3.964
64:1	1.340	5.97	105.913	7.909

TABLE 5-8 ONE-DIMENSIONAL

SIMULATION RESULTS FOR "CROWD"

D/H	SCR	FCR	MSE	FBCR
8:1	1.307	1.032	0.065	1.295
16:1	1.118	1.335	5.534	1.727
32:1	1.093	2.39	29.606	3.144
64:1	1.104	4.32	87.360	5.718

is the Fourier bit compression ratio; this is the number of bits required to represent the original picture divided by the number of bits required to represent the Fourier components which are above threshold and form the filtered picture. This computation is based on research results obtained by Anderson and Huang, and Andrews. Their results make it possible to estimate the average number of bits required per Fourier component for run-length encoding of the Fourier domain. For each Fourier component, Andrews found that if the phase was linearly quantized with six bits, and the magnitude was statistically quantized according to a Gaussian or a Rayleigh distribution using an average of six bits, then high quality reproduced pictures were obtained. This implies an average of twelve bits per Fourier component not including the run length data. Anderson and Huang obtained favorable results using between 9.56 and 11.2 bits per Fourier coefficient including the run length data. Therefore, the value used here to compute FBCR is 12 bits per Fourier component. The FBCR is based on the number of nonredundant Fourier components that must be transmitted. This number excludes all those components which can be obtained by computing the conjugate of the components that were actually transmitted.

Table 5-6 contains the one-dimensional compression results for the picture "face". The sample compression ratio for the original unfiltered picture is 1.361 for an error criterion of 0 brightness quantization units.

The number of nonredundant Fourier components retained with increasing viewing distance was 31136, 20459, 10208, and 5225 respectively.

Table 5-7 gives the one-dimensional simulation results for the picture "cameraman". The sample compression ratio for the original picture was 1.506 for an error criterion of 0. The number of retained Fourier coefficients decreased with increasing viewing distance in the following way: 26910, 19447, 11021, 5524.

Table 5-8 gives the one-dimensional simulation results for the picture "crowd". The sample compression ratio for the original picture was 1.313. The number of retained nonredundant Fourier components for increasing distance were 33730, 25292, 13896, 7640. The computer printout obtained from the one-dimensional simulation program using "crowd" is reproduced in appendix C.

The computer simulation results need to be evaluated from the standpoint of two objectives. The first objective is to investigate the utility of the perceptual models for bandwidth compression in the Fourier domain. The second objective is to evaluate the capabilities of the model for simulating the effects of distance on spatial frequency rejection of the human visual system. The one-dimensional SCR statistics imply that the model does not enhance bandwidth compression results in the spatial domain when the compression algorithm is a zero-order floating aperture interpolator. Therefore, that objective will not be pursued any further.

Evaluation of the utility of simulated distance for compression in the Fourier domain can be done both in an objective and in a subjective manner. The objective criteria that might be used are mean squared error and line by line comparison. The subjective method would be to compare photographs of the original and the filtered pictures.

On the basis of mean squared error, it is safe to conclude that "face" and "cameraman" are not preceptibly degraded for a simulated viewing distance of 8 times picture height. But these pictures would probably contain perceptible errors for viewing distances greater than 16 times picture height. The mean squared error for "crowd" however, is reasonably low for viewing distances as large as 16 times picture height so it would be expected that simulated distances as large as this value would not introduce perceptible degradation. On this basis, it would be predicted that simulated distance would be useful for achieving Fourier bit compression ratios of something less than 2.0. Basing the evaluation on a low mean squared error, less than 1.0 for example, can be deceiving because, if the model introduced large errors only in regions containing high spatial frequencies, then a large mean squared error would not imply picture degradation. This is due to the low incremental brightness sensitivity of eye for small backgrounds.

There are inherent limitations in the use of line-by-line comparisons just on the basis of volume of data that

would be generated. In the course of the computer simulation a few selected picture lines were graphed for purposes of line by line comparison. A comparison of selected lines from "face" with their filtered counterparts indicates that for viewing distances up to 16 times picture height, large errors are confined to regions of high activity. Therefore it would be expected that, for these distances, the degradation would be imperceptible. For "face" and viewing distances of 32 times picture height and 64 times the picture height, line-by-line comparison indicates a definite blurring of picture detail which would certainly be perceptible. "Camera has degradation characteristics as a function of simulated distance similar to those for "face". In fact, line by line comparison indicates that "cameraman" would be of reasonably good quality for a simulated distance of 32 times picture height. A phenomenon is found in "cameraman" whereby narrow regions of contatant brightness are replaced by low amplitude oscillations at or near the highest possible spatial frequency whose average value is near the value of constant brightness. An analysis of selected lines from "crowd" indicate quite good fidelity all the way up to 32 times picture height. This is to be expected, since for this picture, more Fourier components were retained at each distance than were retained for each of the other two pictures. In summary, then, line-by-line comparison indicates that Fourier bit compression ratios as high as 3.0 might be feasible with only slightly perceptible picture degradation.

Since a subjective evaluation is only valid for the current observer, this exercise is the responsibility of the reader. Photographs of the original and filtered pictures will be found in Appendix F.

The model which has been simulated eliminates spatial frequency components which are hypothetically imperceptibly at a given viewing distance. The utility of the model in this regard can be evaluated by comparing the original picture with a processed version of the original which has been filtered for a particular simulated viewing distance. When the two pictures are viewed from that distance, they should appear to be nearly identical. Such a comparison is inherently subjective and would probably depend upon the contrast sensitivity characteristics of the observer. It is the opinion of the author that this function is performed reasonably well by the model, especially for large viewing distances such as 64 times picture height. In such a case, the two photographs are clearly distinguishable close up but lose their dissimilarity at the appropriate distance.

5.2.3 Two-dimensional simulation

The preceding results were obtained by treating a picture as a sequence of independent, one-dimensional lines. Each line was segmented in accordance with the simulated viewing distance, and each individual one-dimensional line segment was transformed, filtered and inverse transformed. The filtered segments were then reconstituted to form the

filtered picture. The contrast sensitivity function used to filter the line segments was derived from psychophysical observations using one-dimensional sinusoidal grating whose lines of constant phase were generally vertical. Since the largest portion of the picture being processed at any given time could not exceed one picture line or 256 bytes, the computer simulation program was relatively unsophisticated and inexpensive to run (the actual cost to filter and compress a single 256 X 256 X 256 picture for one simulated viewing distance was about $13.00.)

In this section, results will be described which were obtained when the picture was processed as a two-dimensional entity. The simulation will be based on the assumption that the retina analyzes the two-dimensional image focused on the fovea in terms of its two-dimensional spatial frequency components and that the spatial frequency channels are distributed in some manner over the two-dimensional spatial frequency plane. Therefore, as in the one-dimensional case, the perceived image is the result of filtering and threshold mechanisms in the individual spatial frequency channels. It is the threshold mechanism of the individual channels that will be utilized here to eliminate hypothetically imperceptible two-dimensional spatial frequency components.

In the computer simulation of the perceptual model, the analysis of the picture in terms of fovea-size images will be approximated by dividing the picture into two-dimensional

segments which subtend approximately one square degree of visual angle. This segmentation is performed under the constraint that efficient operation of the two-dimensional fast Fourier transform algorithm requires that the two dimensions be powers of 2.

The contrast sensitivity function for human vision is assumed to be a psychophysical measurement of the composite threshold characteristics of the spatial frequency channels. Using the contrast sensitivity function as a threshold mechanism implies that the individual spatial frequency components lie in separate channels. This is a questionable assumption, especially for close viewing distances where the spatial frequency components are not widely separated in the spatial frequency plane. However, little is known about the bandwidth of these hypothetical spatial frequency channels.

Comprehensive data does not exist on the contrast sensitivity function for two-dimensional spatial frequencies. Equation 5-22.6 indicates that the period of a two-dimensional sinusoid is rotationally symmetric with respect to the x and y frequency numbers. Therefore, if the contrast sensitivity function is independent of rotational transformations of the one-dimensional sinusoid, then the contrast sensitivity function is rotationally symmetric. This implies that the optics of the eye and the retina-brain are isotropic. In reality, measurements of contrast sensitivity dependence on sine wave orientation indicate a significant dependence.

Campbell et al. (reference 3-40) obtained results which show a 20% decrease in the maximum resolvable spatial frequency when the lines of constant phase are changed from a vertical to a diagonal orientation. Orientational specificity appears to be primarily neurological in origin. However, a comparison of the results they obtained using sinusoidal gratings produced on a CRT with those obtained by producing interference fringes directly on the retina imply the possibility of some optical orientational sensitivity. Nevertheless, for purposes of computational simplicity, rotational symmetry will be assumed as a first approximation. Therefore, in Equation 5-33, f will be replaced by f_{pq} given by 5-22.6. The picture will be divided into square segments. Therefore the fundamental frequency in the x and in the y directions will be identical.

The computer program used to perform the two-dimensional simulation is significantly more complex and expensive to run than that used for the one-dimensional simulation. Dynamic storage allocation is used to provide buffer space for picture segmentation and picture reconstitution. Temporary files hold picture segments while they are not being processed. Depending upon the simulation parameters, up to 263,000 bytes of virtual memory may be in use by the program. The cost of processing a single 256 X 256 X 256 picture for a single viewing distance averaged about $21.00. The two dimensional simulation program is reproduced in appendix D.

5.2.3.1 Summary of the simulation results

Computer simulation of the two-dimensional perceptual model was performed using the three 256 X 256 X 256 pictures. The rotationally symmetric contrast sensitivity function was used with the parameters described by Equations 5-33. Each of the three pictures was processed for five simulated viewing distances. The viewing distances used were 4, 8, 16, 32, and 64 times picture height. These viewing distances were chosen because they each represent a different level of picture segmentation. With increasing viewing distance, the picture is divided up into segments containing 16 X 16, 32 X 32, 64 X 64, 128 X 128, and 256 X 256 sample points.

The Fourier compression ratio (FCR) is the total possible number of nonredundant Fourier components divided by the number of nonredundant Fourier components after filtering. Now, for an N X N picture segment, Equation 5-27 shows that the total possible number of Fourier components produced by the discrete Fourier transform is equal to N^2, but some of these components are redundant in the sense that they can be derived from the others. The discrete Fourier transform is a discrete form of the complex Fourier series and Equation 5-21.16 indicates that the coefficients with index (n,m) are the complement of the coefficients with index (-n,-m). After elimination of these redundant Fourier coefficients, the total possible number of nonredundant coefficients is given by

$$\left(\frac{N}{2} + 1\right) \cdot N - 2\left(\frac{N}{2} - 1\right) \qquad (5\text{-}34)$$

This is the number per segment. Therefore, the total number is given by this number times the number of picture segments. The data on segment size, number of segments, and total number of nonredundant Fourier coefficients for each value of simulated distance is summarized in Table 5-9.

In order to verify that the computer simulation program properly processed two-dimensional sinusoids, a program was written which could produce a two-dimensional sampled sinusoid of any desired average value, amplitude, x and y frequency numbers, and phase. It was also possible to specify the x and y sampling densities and the number of levels of gray scale. The simulation program was tested on selected two-dimensional sinusoids, particularly those with extreme values of spatial frequency. It was verified that the simulation program properly calculated the amplitude of those components and was cognizant of the true coordinates of those components in the Fourier domain.

The results obtained from the computer simulation for each of the three pictures are summarized in Tables 5-10, 5-11, and 5-12. Table 5-13 summarizes the Fourier coefficient data for each of the three pictures and each of the simulated viewing distances. This Fourier coefficient data forms the basis for the calculation of the Fourier compression ratio (FCR) and the Fourier bit compression ration (FBCR).

TABLE 5-9 SEGMENT SIZE, NUMBER OF SEGMENTS, AND TOTAL NUMBER OF NON-REDUNDANT FOURIER COEFFICIENTS AS A FUNCTION OF SIMULATED DISTANCE.

D/H	SEGMENT SIZE	NUMBER OF SEGMENTS	NUMBER OF FOUTIER COEFFICIENTS
4:1	16 X 16	256	33280
8:1	32 X 32	64	32896
16:1	64 X 64	16	32800
32:1	128 X 128	4	32776
64:1	256 X 256	1	32770

TABLE 5-10 TWO-DIMENSIONAL

SIMULATION RESULTS FOR "FACE"

D/H	SCR	FCR	MSE	FBCR
4:1	1.248	1.90	3.181	2.50
8:1	1.149	3.419	13.668	4.541
16:1	1.138	11.460	52.231	15.265
32:1	1.201	39.632	112.564	52.829
64:1	1.216	112.611	166.013	150.139

TABLE 5-11 TWO-DIMENSIONAL

SIMULATION RESULTS FOR "CAMERAMAN"

D/H	SCR	FCR	MSE	FBCR
4:1	1.277	1.075	0.230	1.412
8:1	1.096	1.675	7.665	2.225
16:1	1.082	5.264	57.905	7.011
32:1	1.096	17.726	145.550	23.628
64:1	1.130	61.597	270.254	82.124

TABLE 5-11 TWO-DIMENSIONAL

SIMULATION RESULTS FOR "CROWD"

D/H	SCR	FCR	MSE	FBCR
4:1	1.457	2.025	3.326	2.659
8:1	1.353	3.518	13.289	4.672
16:1	1.254	10.857	86.298	14.462
32:1	1.227	32.291	191.591	43.044
64:1	1.307	117.877	314.195	157.158

TABLE 5-13 NUMBER OF NON-REDUNDANT FOURIER COEFFICIENTS RETAINED FOR EACH PICTURE AND EACH SIMULATED DISTANCE.

D/H	SCR	FCR	MSE	FBCR
4:1	1.248	1.90	3.181	2.50
8:1	1.149	3.419	13.668	4.541
16:1	1.138	11.460	52.231	15.265
32:1	1.201	39.632	112.564	52.829
64:1	1.216	112.611	166.013	150.139

5.2.3.2 Analysis of results

There are two aspects to the analysis of the results. One aspect is to evaluate the two-dimensional simulation results relative to the objectives, those objectives being Fourier domain compression and distance simulation. The other aspect is a comparison of the two-dimensional simulation results with the one-dimensional simulation results.

In a strict sense, Fourier domain compression is being achieved if an FBCR greater than 1.0 can be obtained without perceptible degradation of the picture relative to the original. Mean squared error, line by line comparison, and photographic comparison will be used to evaluate this objective.

In only one case was the mean squared error such that the average error per picture point was less than one brightness quantization level. This was for the picture "crowd" for a simulated viewing distance of 4 times picture height. In all other instances, due to the values of mean squared error, there are picture points that are guaranteed to have an error greater than one brightness quantization level. Incremental brightness sensitivity data indicates that if the errors are grouped in such a way that they form an area subtending several picture elements in a uniform part of the picture, then the area in error would be perceived, even if the error were on the order of only one or two brightness quantization levels. However, if the errors are randomly scattered throughout the picture, and in

particular, occur in very nonuniform areas of the picture, they would be less likely to be perceived. Again, this assumes that the magnitude of the errors are on the order of one or two brightness quantization units. So, on the basis of mean squared error, an FBCR of 2.0 or less is implied.

Line by line comparison of a pair of lines from "face" for a simulated viewing distance of four times picture height indicates that most of the errors are on the order of one or two brightness quantization units. Larger errors are generally confined to edges of brightness in regions where, over a distance of a few samples, the brightness changes magnitude by a large amount. At eight times picture height, the same observations hold, but in addition, low amplitude "spikes" of brightness are both eliminated and introduced. These "spikes" are one sample wide and less than ten brightness quantization units in amplitude. At sixteen times picture height, these "spikes" increase in amplitude and width and would most certainly be visible. For 32 and 64 times picture height, the effect is one of smoothing or blurring of detail. The same general observations apply to "cameraman". In addition, for large viewing distances, uniform areas are replaced by low amplitude, medium frequency "ringing."

This high frequency "ringing" may be due to the artificial edges created by periodic extension of the picture segments. It would be of interest to know if this effect disappears when steps are taken to eliminate the artifical edges. This could be accomplished by adding samples to the periphery

of a segment in such a way as to cause all edges of the picture to gradually transition to some common value. Because the discrete fast Fourier transform algorithm used in this simulation required that the segment dimensions be a power of two, it was not possible to simply extend the size of the segment by a few samples. Instead, it would have been necessary to use, in effect, overlapping picture segments. This would have required a considerable modification of the simulation program.

The picture "crowd" is considerably more detailed than either "face" or "cameraman". This is reflected in Table 5-13 where the number of Fourier components retained at each viewing distance for "crowd" is roughly double the number retained for either of the other two pictures. Line by line comparison of "crowd" confirms the very low mean squared error obtained for four times picture height. The errors are very few in number, and when they occur, they amount to one brightness quantization unit. For eight times picture height, large errors are confined to regions of large brightness slope. In addition narrow, low amplitude "spikes" of brightness are both introduced and deleted. For sixteen times picture height, the "spikes" increase in amplitude and there is some loss of detail. For 32 and 64 times picture height, the degradation is unmistakeable. Brightness contours in the filtered picture lines deviate considerably from their original counterparts. Some sample results obtained for "crowd" are given in appendix E.

Therefore, in general, line by line comparison indicates that serious picture degradation occurs for viewing distances greater than eight times picture height. On this basis, a Fourier bit compression ratio (FBCR) on the order of 4.5 would be achievable for "face" and "cameraman . For "crowd" this number would be on the order of 2.2.

In order that the reader may subjectively evaluate the utility of this perceptual model for Fourier domain compression and distance simulation, photographs of the processed pictures are given in Appendix F.

If the model is useful for simulating the spatial frequency filtering effect of viewing distance, then the model can be used to determine the number of Fourier coefficients required to reconstruct a picture for a given viewing distance. The results indicate that for large viewing distances the amount of data required can be dramatically reduced. The same effect would be observed for dimmer viewing conditions, because for decreased mean luminance the contrast sensitivity function decreases in maximum value, and the maximum resolvable spatial frequency also decreases.

For purposes of comparison, the Fourier coefficient data resulting from the one-dimensional simulation is reproduced in Table 5-14. Comparing this data with that in Table 5-13 for the two- dimensional simulation, it is seen that many more Fourier coefficients were eliminated as a results of two-dimensional filtering. The mean squared error data agrees with this observation in that for the same viewing distance-two dimensional filtering resulted in

a larger mean squared error. In the one-dimensional case, line by line comparison resulted in the conclusion that unacceptable degradation probably occurred above sixteen times picture height whereas eight times picture height was the cutoff level for two-dimensional filtering. Nevertheless because so many more Fourier coefficients are eliminated for the same viewing distance, by two-dimensional filtering, the final estimates of potential compression in the Fourier domain were higher.

TABLE 5-14 FOURIER COEFFICIENTS

RETAINED IN THE ONE-DIMENSIONAL SIMULATION

D/H	"FACE"	"CAMERAMAN"	"CROWD"
8:1	31136	26910	33730
16:1	20459	19447	25292
32:1	10208	11021	13896
64:1	5225	5224	7640

REFERENCES

5-1 IES Lighting Handbook, 2nd ed., Illuminating Engineering Society, N.Y., 1952.

5-2 R.V. Churchill, Fourier Series and Boundary Value Problems, McGraw Hill, New York, 1941.

5-3 M.E. Van Valkenburg, Network Analysis, Prentice-Hall Englewood Cliffs, New Jersey, 1959.

5-4 G.D. Bergland, "A Guided Tour of the Fast Fourier Transform," IEEE Spectrum, vol. 6, no. 7, July 1969, pp. 41-52.

5-5 J.A. Glassman, "A Generalization of the Fast Fourier Transform," IEEE Trans. on Computers, vol. C19, no. 2, Feb. 1970, pp. 105-116.

5-6 H.C. Andrews, E.L. Caspari, "A Generalized Technique for Spectral Analysis," IEEE Trans. of Computers, Vol. C19, no. 1, Jan, 1970, pp. 16-25.

5-7 H. Andrews, "A High-Speed Algorithm for the Computer Generation of Fourier Transforms," Short note, IEEE Trans. on Computers, vol. C17, no. 4, April, 1968, pp. 373-375.

5-8 S. Davis, Computer Data Displays, Prentice Hall, Engle-Wood Cliffs, N. Jersey, 1969.

5-9 L.B. Arguibau, Vacuum-Tube Circuits and Transistors, John Wiley and Sons, New York, 1959.

CHAPTER VI
SUMMARY AND COMMENTS

The purpose of this research was to determine a model for human vision suitable for the prediction of imperceptible components of visual data. The derived model was then used as a prefilter in a picture bandwidth compression system. The objective was to determine whether or not the use of such a prefilter would enhance the results that could be obtained from standard redundancy reduction techniques.

A study of the physiology and psychophysics of human vision provides evidence that mechanisms exist in the human visual system which are sensitive to spatial frequency. The detection threshold of a sinusoidal luminance distribution depends upon its spatial frequency. This characteristic of vision is well known and forms the basis for dissimilar treatment of the low frequency and high frequency portions of a picture, a method employed in several redundancy reduction techniques. A lesser known characteristic of vision is the multiple channel hypothesis. This theory is based on evidence that the eye detects signals by analyzing them into bandlimited components which are separately and independently detected. The contrast sensitivity function for the retina-brain portion of the visual system is assumed to be a composite representation of the detection thresholds of the spatial frequency channels.

Since the maximum contrast sensitivity and the maximum resolvable spatial frequency are functions of the viewing environment, typical environmental parameters were determined in order to select a suitable contrast sensitivity function. In the computer simulation, the contrast sensitivity function was used as a threshold detection function in the discrete Fourier domain. This technique implies that the Fourier coefficients produced by the discrete Fourier transform lie in separate spatial frequency channels. This is a questionable assumption since little is known about the probable bandwidth of the individual channels. However, it does make it practical to obtain approximate results. In addition, it is assumed that only the foveal portion of the retina participates in the spatial frequency analysis. Therefore, the visual data is processed in segments which subtend approximately one degree of visual angle. The number of samples in a segment depends upon the viewing distance which is a parameter in the computer simulation.

In terms of the original objective, analysis of the computer simulation results indicates that, at least for a zero-order interpolator, bandwidth compression results are not enhanced by the simulated prefilter. However, it was determined that the Fourier domain threshold perception model was useful for compression in the Fourier domain. A systematized tradeoff between picture degradation and bandwidth compression could be achieved by means of the spatial frequency filtering effect of simulated viewing distance.

In addition, it was found that the model was useful for determining the relationship between Fourier domain compression ratio and viewing distance in a case where picture fidelity must be maintained.

On the basis of line by line comparison and mean squared error, it was estimated that a bit compression ratio on the order of 4.5 could be achieved in the Fourier domain for pictures of low complexity, e.g. "face", "cameraman". For highly complex pictures, a bit compression ratio of 2.2 would be expected. Line-by-line comparison and mean squared error are admittedly crude and inaccurate methods for judging fidelity. Andrews (reference 2-27) found that there was a good correspondence between matched filter correlation and image fidelity. Unfortunately, such a technique would have added considerable cost to an already very expensive computer simulation.

For monochrome still pictures containing at least 256 X 256 samples and at least 64 levels of gray scale, bit compression ratios as low as 2:1 and as high as 22:1 have been reported. PCM techniques are at the low end of the scale while the midrange is occupied by Fourier encoding techniques, interpolators and dual mode methods. The only technique that seems capable of achieving bit compression ratios greater than 10:1 is contour encoding, however, this method is even more complex than Fourier encoding techniques. The Fourier domain threshold detection model investigated here gives results which compare favorably

with other Fourier domain techniques. In addition, the model can be used to achieve considerable bandwidth compression for large viewing distances without a serious loss of fidelity.

There are several aspects of the multiple channel threshold model which might be suitable for further investigation. No attempt was made to remove the effects of artificial edges introduced by periodic continuation of the picture segments. This could be investigated by using overlapped segments and a data window which attenuates the segment edges down to some common value of intensity. An interesting data window would be one that corresponds to the profile of visual acuity as a function of retinal position. Line-by-line comparison indicated that in some instances the model tended to make the slope of brightness more uniform, especially for large viewing distances. This implies that the model might be used to enhance first-order interpolation results. The spatial frequency channels need to be investigated in terms of overlapping narrow band filters. In general, for high resolution pictures, several Fourier components would lie within the passband of a given channel. A threshold criterion would have to be determined for the output of each channel in addition to the position and shape of each passband. Another interesting idea would be to apply the threshold model to the difference between the original picture and its compressed counterpart in order to assess the perceptibility of the difference. The effects

of including a logarithmic transformation in the model as well as the effects of noise should also be investigated.

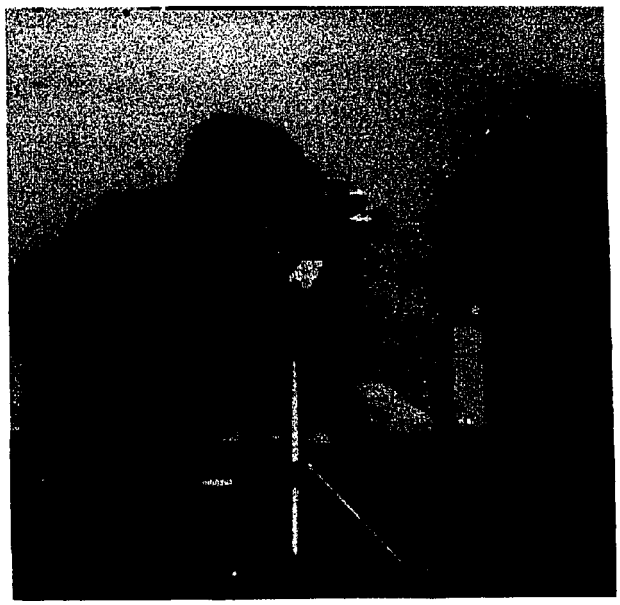
APPENDIX A

THE DIGITIZED PICTURES USED IN THE COMPUTER SIMULATION

These digitized pictures were obtained from the Massachusetts Institute of Technology Research Laboratory for Electronics. They contain 256 X 256 samples where each sample is linearly quantized to 8 bits.



"FACE"



"CAMERAMAN"



"CROWD"

APPENDIX B

THE FORTRAN PROGRAM FOR THE ZERO-ORDER FLOATING APERTURE INTERPOLATOR

The following pages contain a compiler listing of the FORTRAN program used to simulate the spatial domain compression algorithm. This program accepts as input a two-dimensional array of short floating point numbers, e.g., a sampled picture or the result produced by the two-dimensional discrete Fourier transform. Two results are returned by the program: (1) a two-dimensional array of integer valued short floating point numbers and (2) a file containing the run-length encoded picture.

The algorithm is constrained to produce an integer valued array because the results will ultimately be compared with an original sampled quantized picture. Therefore the error tolerance, one of the calling parameters, must be of sufficient magnitude to produce an error corridor for each run which straddles an integer quantization level. Otherwise an error return occurs.

The run-length encoded data, which is optional, consists of 2-byte pairs. The first byte is the length of the run and the second byte is the value of the run.

The picture may be fed to the program line-by-line in which case intermediate results are saved between lines, e.g., a run which is in progress but not terminated is continued when the subroutine is reentered with the next picture line.

The program begins by interpolating the top line of the picture from left-to-right. Upon reaching the end of the line, interpolation continues with the rightmost sample of the next scan.line down, and moves from right-to-left. This "zig-zag" pattern continues until the end of the picture.

```
                C                                                                       1.000
                C                                                                       2.000
                C           THIS PROGRAM USES ZERO-ORDER FLOATING APERTURE              3.000
                C           INTERPOLATION TO COMPRESS A PICTURE. THE ORIGINAL           4.000
                C           PICTURE MUST BE IN SHORT FLOATING POINT FORMAT              5.000
                C           USING 4 BYTES PER PICTURE ELEMENT. THE RETURNED             6.000
                C           PICTURE IS ALSO IN SHORT FLOATING POINT FORMAT,             7.000
                C           WITH EACH ELEMENT IN A RUN REPLACED BY THE VALUE            8.000
                C           OF THE RUN. THE VALUE OF A RUN MUST BE AN                   9.000
                C           INTEGER, THEREFORE THE ERROR TOLERANCE MUST BE              10.000
                C           SUFFICIENTLY LARGE SO THAT THE ERROR CORRIDOR               11.000
                C           FOR EACH RUN STRADDLES AN INTEGER QUANTIZATION              12.000
                C           LEVEL. IF MORE THAN ONE INTEGRAL LEVEL IS IN THE            13.000
                C           CORRIDOR THE RUN VALUE IS THE INTEGER NEAREST               14.000
                C           THE AVERAGE VALUE OF THE CORRIDOR.                          15.000
                C                                                                       16.000
                C           THE RUN-LENGTH ENCODED DATA OUTPUT ON I/O UNIT              17.000
                C           #4 USES 2 BYTES PER RUN, ONE BYTE FOR THE LENGTH            18.000
                C           AND THE FOLLOWING BYTE FOR THE VALUE.                       19.000
                C                                                                       20.000
                C           CALLING ARGUMENTS :                                         21.000
                C                                                                       22.000
                C           LABEL - VECTOR CONTAINING 4 HALF WORDS (XMAX,YMAX,          23.000
                C                   BMAX,BMIN)                                          24.000
                C           PICT - TWO DIMENSIONAL VECTOR CONTAINING THE PICTURE        25.000
                C           ERROR - ERROR TOLERANCE (SHORT FLOATING POINT)              26.000
                C           RESULT - RESULTING PICTURE                                  27.000
                C           CODE - ENTRY CODE (0=CONTINUATION,2=TERMINATION,            28.000
                C                  OTHER=INITIALIZATION)                                29.000
                C           IOCODE - I/O CODE (0=DO NOT FORM AND OUTPUT RUN             30.000
                C                    LENGTH ENCODED DATA,NONZERO=DO OUTPUT)             31.000
                C                                                                       32.000
                C           THIS SUBROUTINE CAN BE USED TO COMPRESS A PICTURE           33.000
                C           LINE-BY-LINE. BETWEEN CALLS FOR EACH LINE RUN               34.000
                C           LENGTH DATA IS SAVED AND ON THE NEXT CALL THE               35.000
                C           SUBROUTINE PICKS-UP WHERE IT LEFT OFF. A ZIG-ZAG            36.000
                C           SCAN IS USED INSTEAD OF TREATING EACH LINE IN               37.000
                C           A STRICTLY LEFT-TO-RIGHT SCAN.                              38.000
                C                                                                       39.000
                C           I/O UNITS USED :                                            40.000
                C                                                                       41.000
                C           UNIT #4 - OUTPUT UNIT FOR THE RUN-LENGTH ENCODED PICTURE    42.000
                C                                                                       43.000
0001                    SUBROUTINE PRDICT(LABEL,PICT,ERROR,RESULT,CODE,IOCODE)          44.000
0002                    INTEGER*2 LABEL(20),LLEN                                        45.000
0003                    INTEGER XMAX,YMAX,CODE,RUNLEN                                   46.000
0004                    REAL PICT(256,256),RESULT(256,256)                              47.000
0005                    REAL PEL,LGNUM,LRECRD,LESAV                                     48.000
0006                    LOGICAL*1 SWITCH,SWITC                                          49.000
0007                    LOGICAL*1 CPICT(254),LGNUM(4)                                   50.000
                C                                                                       51.000
0008                    EQUIVALENCE (LGNUM,LGCNUM)                                      52.000
                C                                                                       53.000
                C                   INITIALIZE VARIABLES                                54.000
                C                                                                       55.000
0009                    IF (CODE.EQ.0) GO TO 50
```

```
0019              IF (CODE.EQ.2) GO TO 480                                    56.000
0011              RUNSUM=0.0                                                  57.000
0012              RUNNUM=0.0                                                  58.000
0013              IP=0                                                        59.000
0014              RUNLEN=0                                                    60.000
0015              SWITCH=.FALSE.                                              61.000
0016              SWITC=.FALSE.                                               62.000
0017              LS1=1                                                       63.000
0018              LS2=1                                                       64.000
0019              UBOUND=PICT(1,1)+ERROR+0.001                                65.000
0020              LBOUND=PICT(1,1)-ERROR-0.001                                66.000
                  C                                                         67.000
                  C                 COMPRESS USING ALTERNATE LEFT-TO-RIGHT AND 68.000
                  C                 RIGHT-TO-LEFT SCANS                     69.000
                  C                                                         70.000
                  C                                                         71.000
0021              50  BOUNUM=RUNNUM                                         72.000
0022                  XMAX=LABEL(1)                                         73.000
0023                  YMAX=LABEL(2)                                         74.000
0024                  DO 400 LOOP1=1,YMAX                                   75.000
0025                  DO 300 LOOP2=1,XMAX                                   76.000
0026                  INDEX=LOOP2                                           77.000
0027                  IF (SWITCH) INDEX=XMAX-LOOP2+1                        78.000
0028                  PEL=PICT(INDEX,LOOP1)                                 79.000
0029                  UPLIM=PEL+ERROR                                       80.000
0030                  LOWLIM=PEL-ERROR                                      81.000
0031              100 UBSAV=UBOUND                                          82.000
0032                  LBSAV=LBOUND                                          83.000
0033                  IF (UPLIM.LT.LBOUND) GO TO 500                        84.000
0034                  IF (LOWLIM.GT.UBOUND) GO TO 500                       85.000
0035                  IF (UPLIM.LT.UBOUND) UBOUND=UPLIM+0.001               86.000
0036                  IF (LOWLIM.GT.LBOUND) LBOUND=LOWLIM-0.001             87.000
0037                  IF (AINT(UBOUND).LT.LBOUND) GO TO 500                 88.000
0038              300 RUNLEN=RUNLEN+1                                       89.000
0039              400 SWITCH=.NOT.SWITCH                                    90.000
                  C                                                         91.000
                  C                 PRINT RESULTS                           92.000
                  C                                                         93.000
0040              450 AVGRUN=RUNSUM/RUNNUM                                  94.000
0041                  CMPRAT=RUNSUM/(RUNNUM*2.0)                            95.000
0042                  DLTNUM=RUNNUM-EGNNUM                                  96.000
0043                  CODE=0                                                97.000
0044                  WRITE (6,910) CMPRAT,AVGRUN,RUNNUM,DLTNUM,RUNSUM      98.000
0045                  RETURN                                                99.000
                  C                                                         100.000
                  C                 END-OF-PICTURE (CODE=2)                 101.000
                  C                                                         102.000
0046              480 IF (RUNLEN.EQ.0) GO TO 800                            103.000
0047                  GO TO 510                                             104.000
                  C                                                         105.000
                  C                 END-OF-RUN DETECTED                     106.000
                  C                                                         107.000
0048              500 IF (RUNLEN.EQ.0) GO TO 700                            108.000
0049              510 RUNSUM=RUNSUM+RUNLEN                                  109.000
0050                  RUNNUM=RUNNUM+1.0                                     110.000
0051                  SPEL=(UBSAV+LBSAV)/2.0
```

```
0052              SPEL1=AINT(SPEL)                                                  111.000
0053              SPEL2=AINT(SPEL+1.0)                                              112.000
0054              IF (SPEL1.LT.LBSAV) GO TO 650                                     113.000
0055              IF (SPEL2.GT.UBSAV) GO TO 520                                     114.000
0056              IF ((SPEL-SPEL1).GT.(SPEL2-SPEL)) GO TO 650                       115.000
0057        520   SPEL=SPEL1                                                        116.000
0058              IF (IOCODE.EQ.0) GO TO 529                                        117.000
                  C                                                               118.000
                  C              FORM AND OUTPUT THE RUN-LENGTH ENCODED DATA       119.000
                  C                                                               120.000
0059        525   IF (IOCODE.EQ.0) GO TO 529                                        121.000
0060              LOGNUM=RUNLEN-1                                                   122.000
0061              IF (LOGNUM.GT.255) LOGNUM=255                                     123.000
0062              RUNLEN=RUNLEN-256                                                 124.000
0063              CPICT(IP+1)=LOGNUM(4)                                             125.000
0064              LOGNUM=SPEL                                                       126.000
0065              CPICT(IP+2)=LOGNUM(4)                                             127.000
0066              IP=IP+2                                                           128.000
0067              IF (IP.LT.254) GO TO 527                                          129.000
0068              LLEN=IP                                                           130.000
0069              CALL WRITE(CPICT,LLEN,0,LNUM,4)                                   131.000
0070              IP=0                                                              132.000
0071        527   IF (RUNLEN.GT.0) GO TO 525                                        133.000
                  C                                                               134.000
                  C              FORM THE COMPRESSED PICTURE LINE                 135.000
                  C                                                               136.000
0072        529   RUNLEN=0                                                          137.000
0073              LSTOP1=LOOP1                                                      138.000
0074              LSTOP2=XMAX                                                       139.000
0075              IF ((LOOP2-1).EQ.0) LSTOP1=LOOP1-1                                140.000
0076              DO 600 LOOP3=LS1,LSTOP1                                           141.000
0077              IF (LOOP3.NE.LS1) LS2=1                                           142.000
0078              IF (LOOP3.NE.LSTOP1) GO TO 545                                    143.000
0079              IF ((LOOP2-1).NE.0) LSTOP2=LOOP2-1                                144.000
0080        545   DO 550 LOOP4=LS2,LSTOP2                                           145.000
0081              INDEX=LOOP4                                                       146.000
0082              IF (SWITC) INDEX=XMAX-LOOP4+1                                     147.000
0083        550   RESULT(INDEX,LOOP3)=SPEL                                          148.000
0084        600   SWITC=.NOT.SWITC                                                  149.000
0085              IF (CODE.EQ.2) GO TO 800                                          150.000
0086              IF((LSTOP1.GE.YMAX).AND.(LSTOP2.GE.XMAX)) GO TO 450               151.000
0087              LS1=LOOP1                                                         152.000
0088              LS2=LOOP2                                                         153.000
0089              USOUND=UPLIM+0.001                                                154.000
0090              LBOUND=LOWLIM-0.001                                               155.000
0091              SWITC=SWITCH                                                      156.000
0092              GO TO 100                                                         157.000
                  C                                                               158.000
0093        650   SPEL=SPEL2                                                        159.000
0094              GO TO 525                                                         160.000
                  C                                                               161.000
                  C              PICTURE CANNOT BE COMPRESSED WITH GIVEN ERROR ARGUMENT  162.000
                  C                                                               163.000
0095        700   PRINT 900                                                         164.000
0096              CODE=1                                                            165.000
```

```
0097              RETURN                                                        166.000
                                                                                167.000
                C                                                               168.000
                C             WRITE LAST LINE OF RUN-LENGTH DATA                169.000
                C                                                               170.000
0098        600   IF (IP.EQ.0) RETURN                                           171.000
0099              LLEN=IP                                                       172.000
0100              CALL WRITE(CPICT,LLEN,0,LNUM,4)                               173.000
0101              RETURN                                                        174.000
                C                                                               175.000
                C             FORMAT STATEMENTS                                 176.000
                C                                                               177.000
0102        900   FORMAT('0**** PICTURE CANNOT BE COMPRESSED - ERROR TOO SMALL')178.000
0103        910   FORMAT('0**** CMPRAT,AVGRUN,RUNNUM,DLTNUM,RUNSUM:',5F10.3)   179.000
0104              END
```

APPENDIX C

SAMPLE RESULTS FROM

THE ONE-DIMENSIONAL SIMULATION

The first few figures in this appendix contain line-by-line comparisons from the picture "CROWD". The continuous curves are the original picture lines. The superimposed symbols represent the sample points of the same picture line after being prefiltered for the specified viewing distance.

The last figure contains a representation of the Fourier domain for the picture "CROWD". The X's represent the position of Fourier components which are above threshold for a simulated viewing distance of 32 times picture height. The structure of the Fourier domain is explained by the use of line-by-line transformation, and the segmentation of each line into four parts based on the viewing distance and associated visual angle.

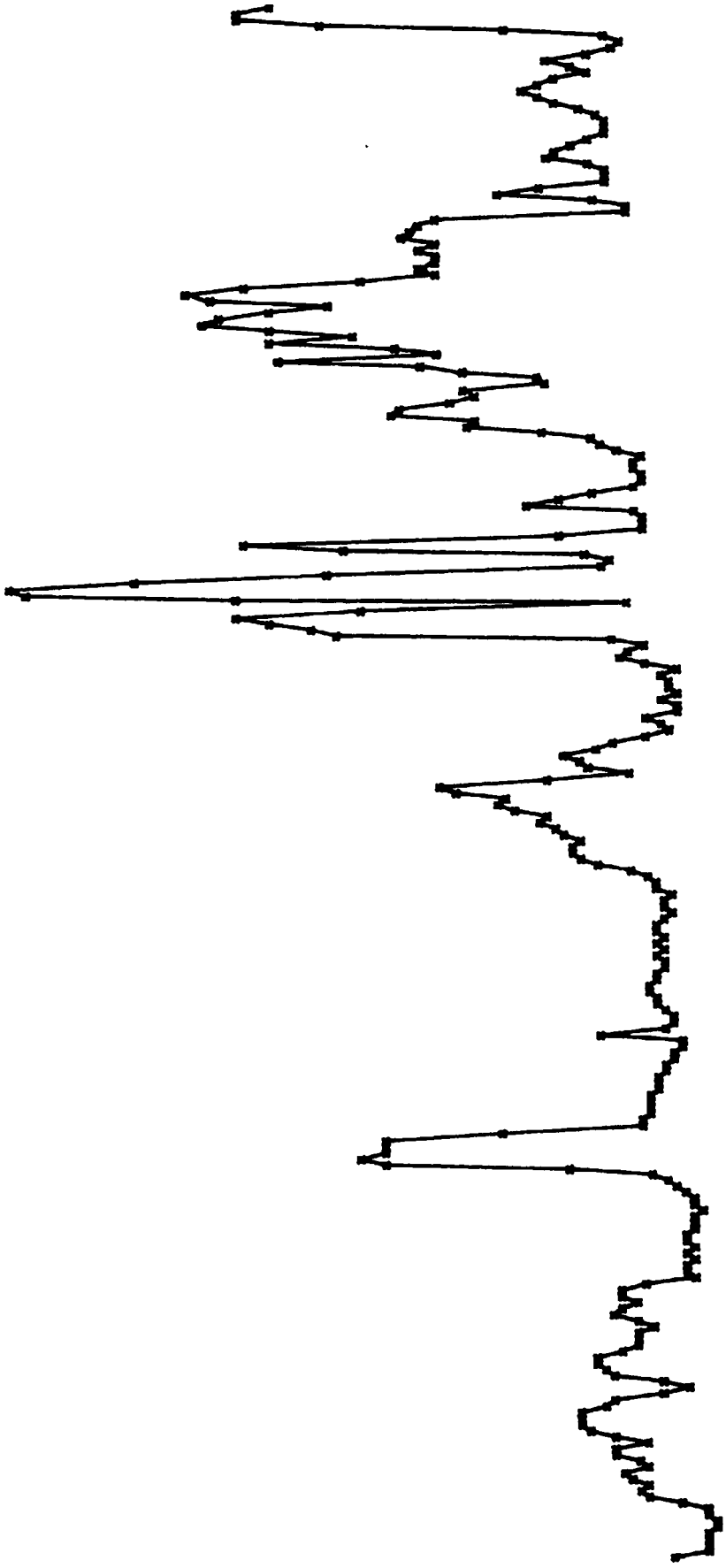


Figure C-1 Line # 100 of "CROWD" before and after being prefiltered using the one-dimensional simulation program for a simulated viewing distance of 8 times picture height.

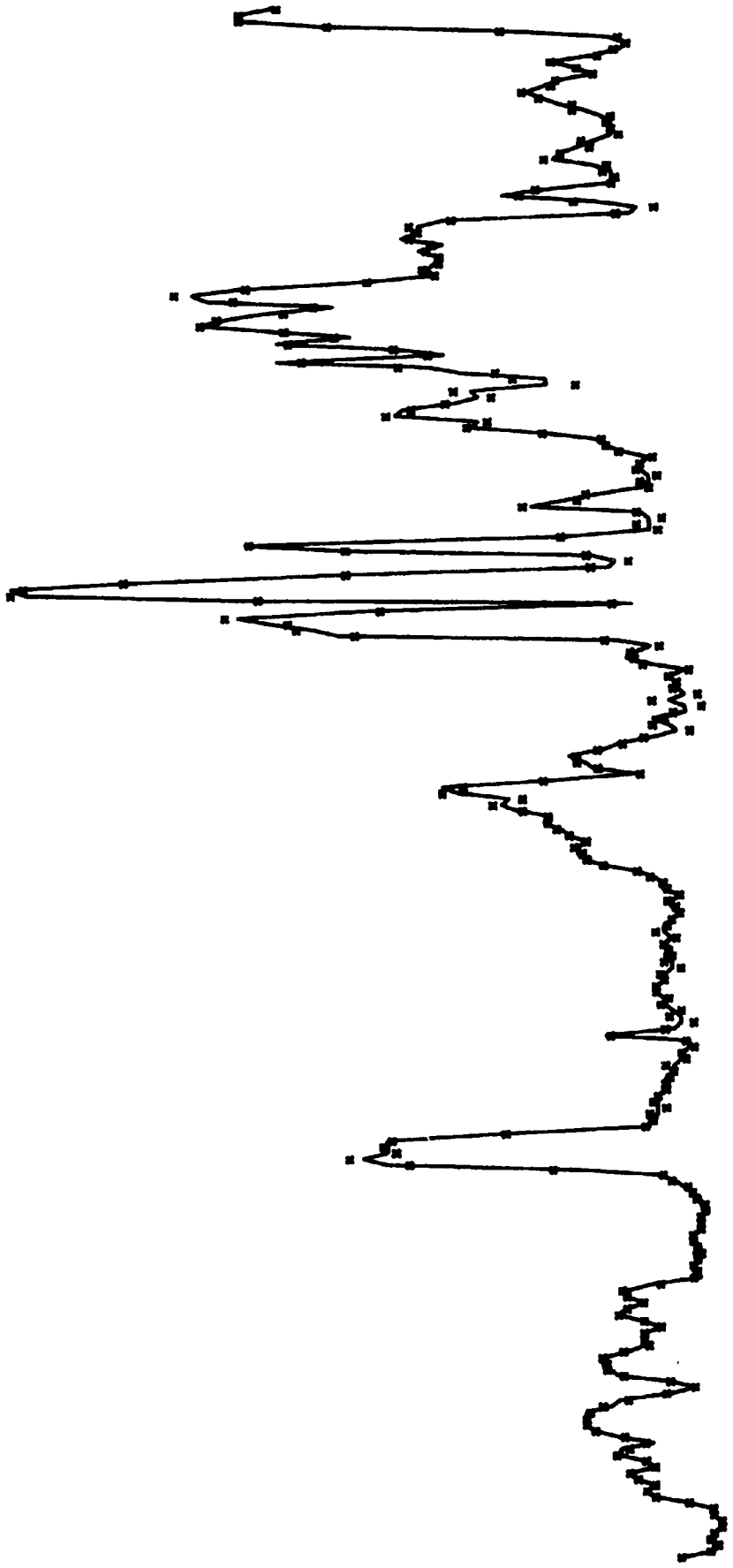


Figure C-2 Line # 100 of "CROWD" before and after being prefiltered using the one-dimensional simulation program for a simulated viewing distance of 16 times picture height.

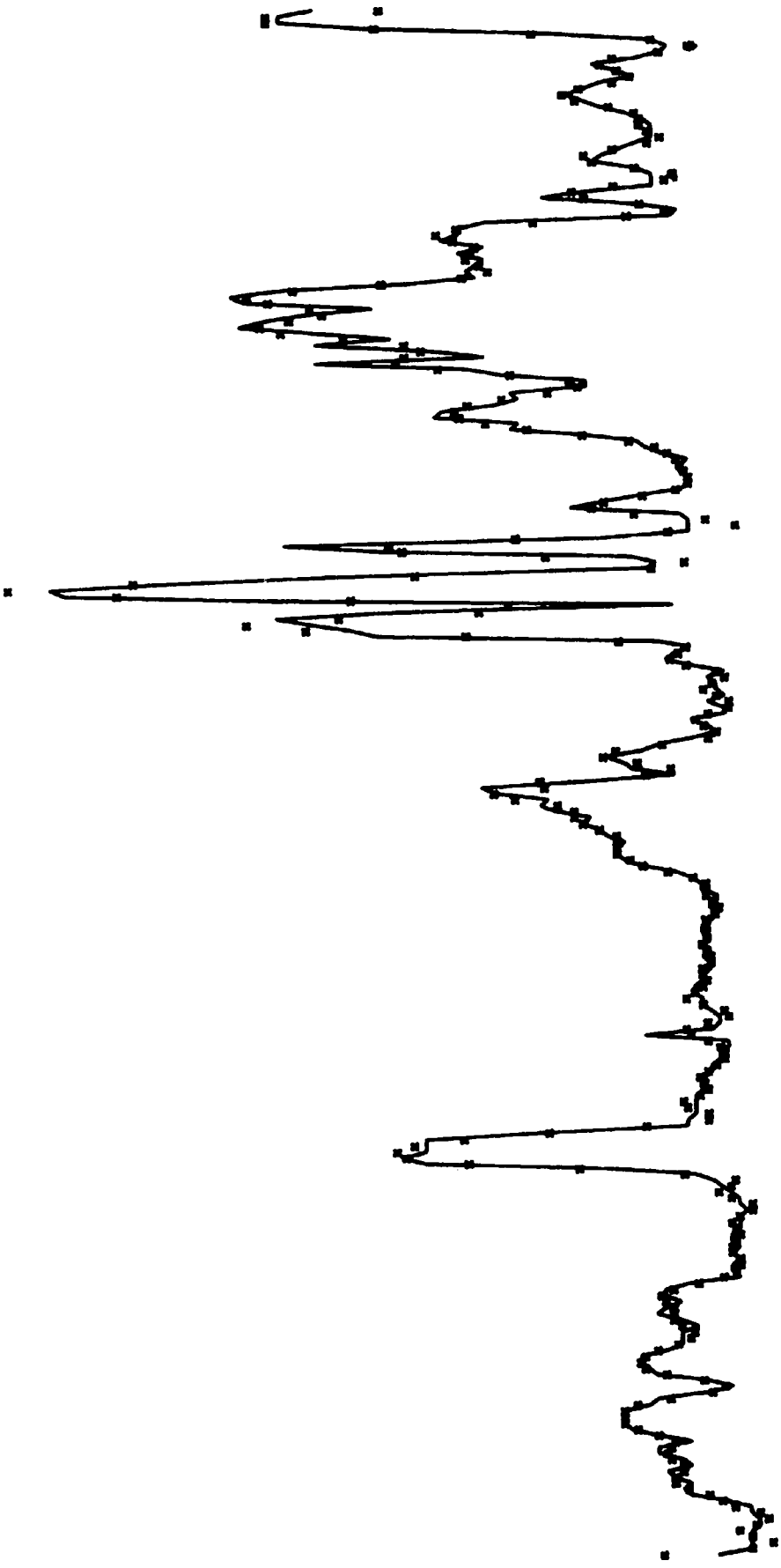


Figure C-3 Line # 100 of "CROWD" before and after being prefiltered using the one-dimensional simulation program for a simulated viewing distance of 32 times picture height.

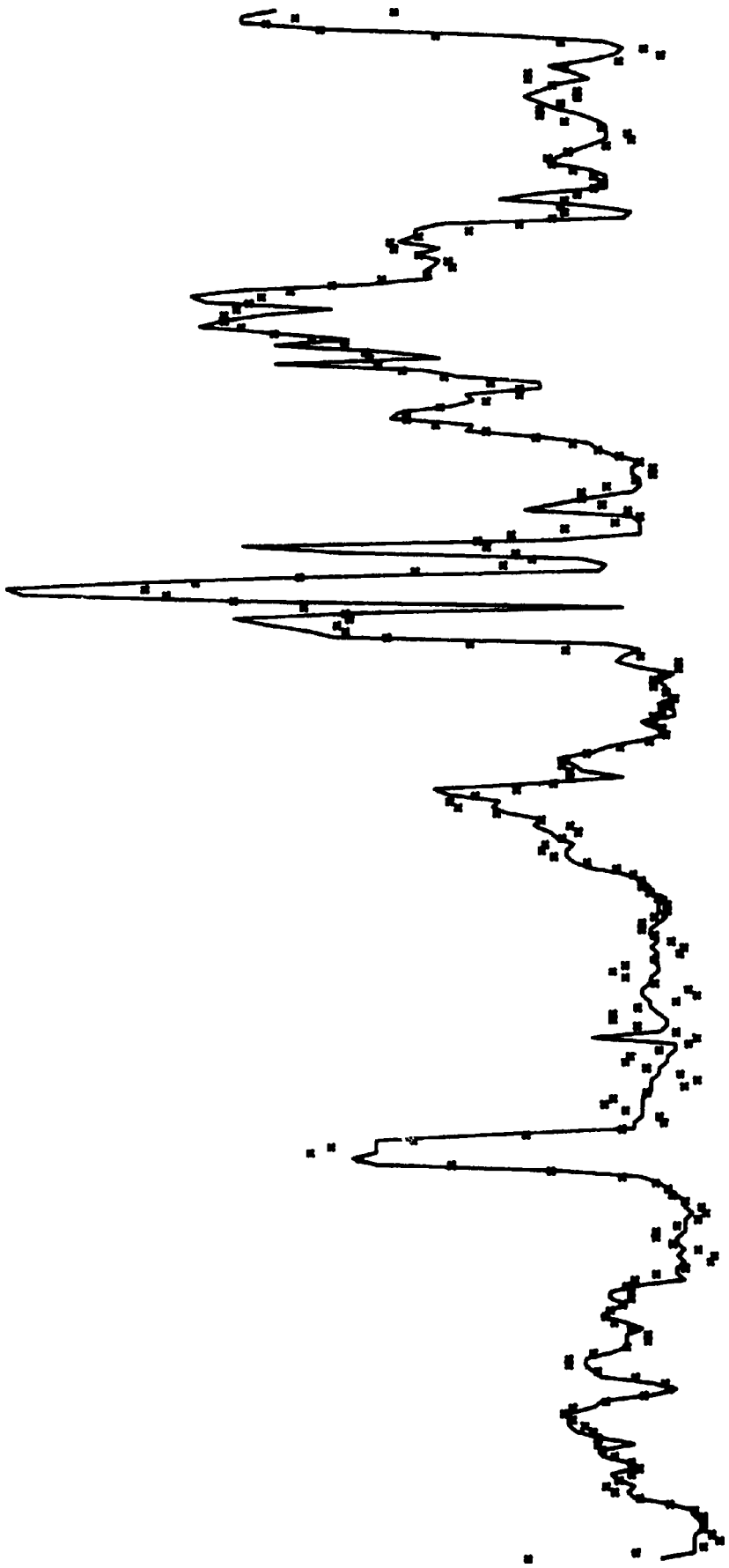


Figure C-4 Line # 100 of "CROWD" before and after being prefiltered using the one-dimensional simulation program for a simulated viewing distance of 64 times picture height.

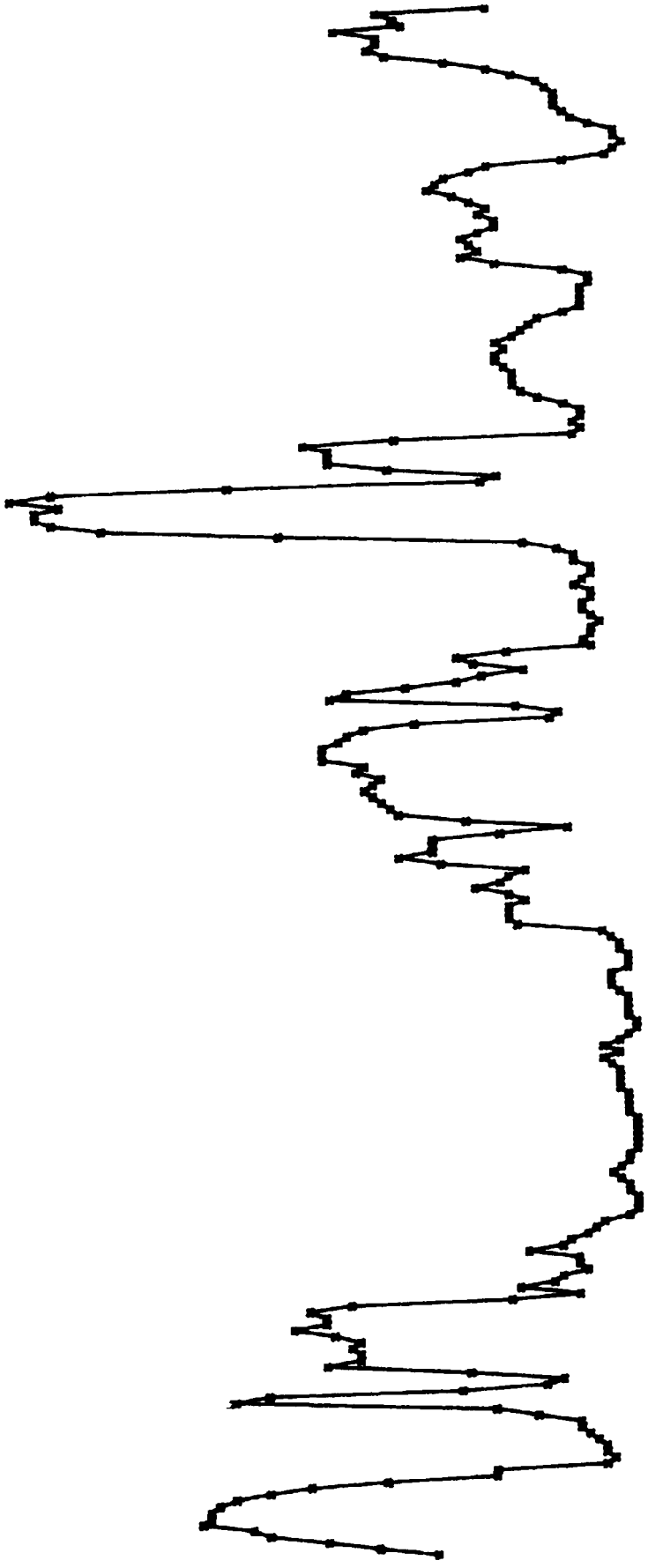


Figure C-5 Line # 200 of "CROWD" before and after being prefiltered using the one-dimensional simulation program for a simulated viewing distance of 8 times picture height.

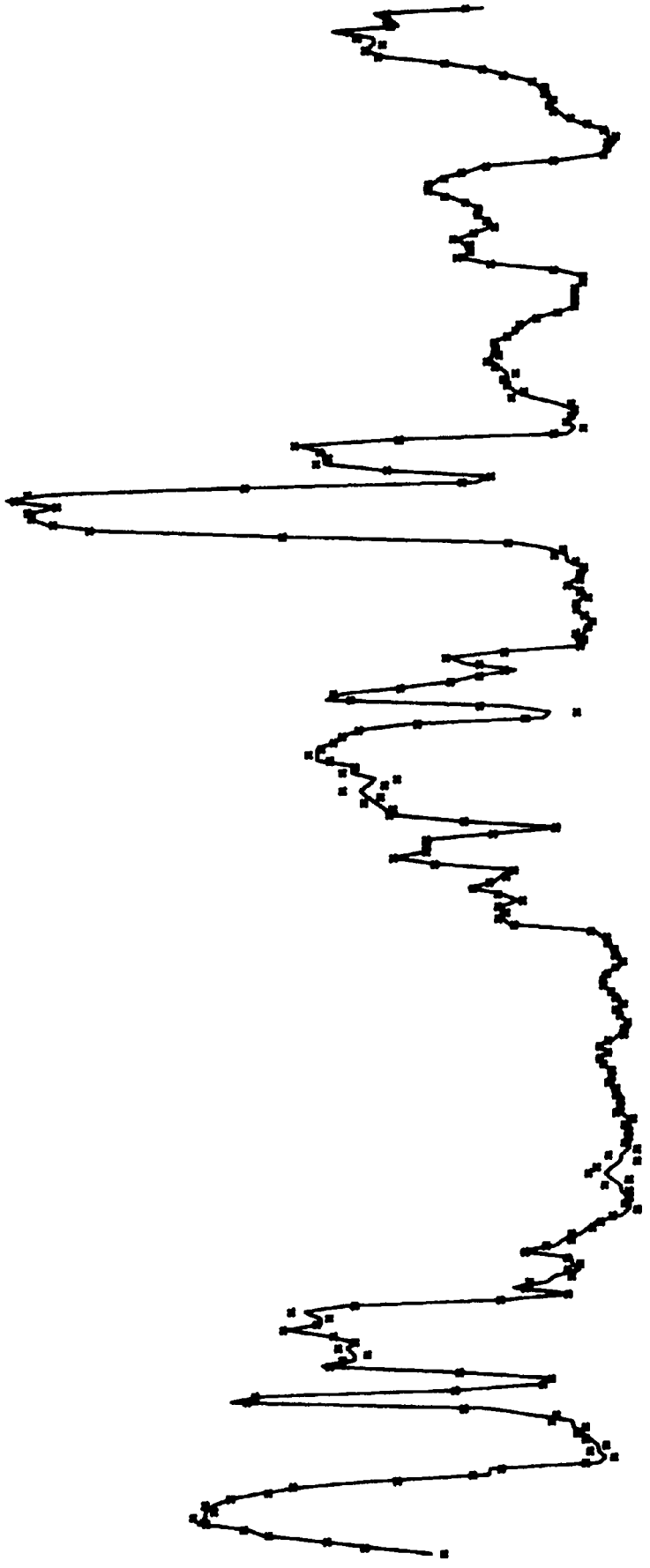


Figure C-6 Line # 200 of "CROWD" before and after being prefiltered using the one-dimensional simulation program for a simulated viewing distance of 16 times picture height.

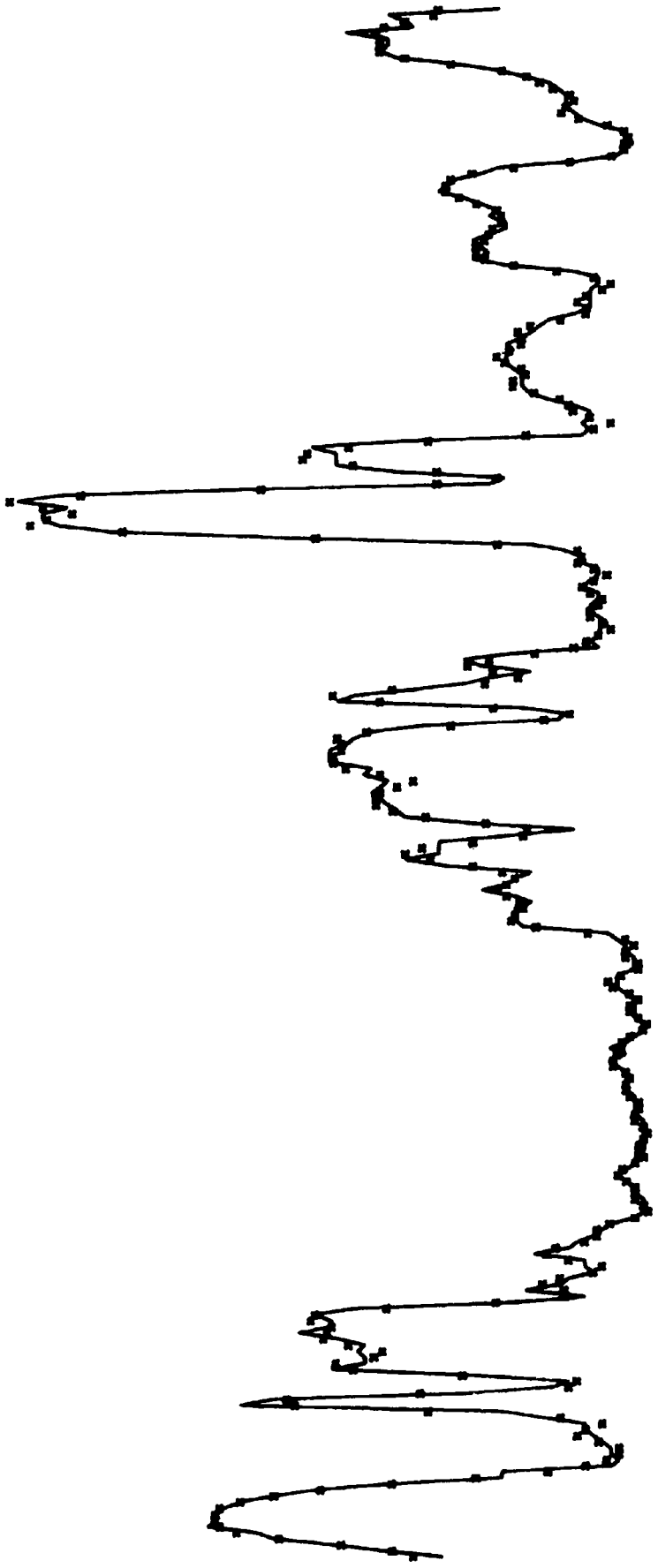


Figure C-7 Line # 200 of "CROWD" before and after being prefiltered using the one-dimensional simulation program for a simulated viewing distance of 32 times picture height.

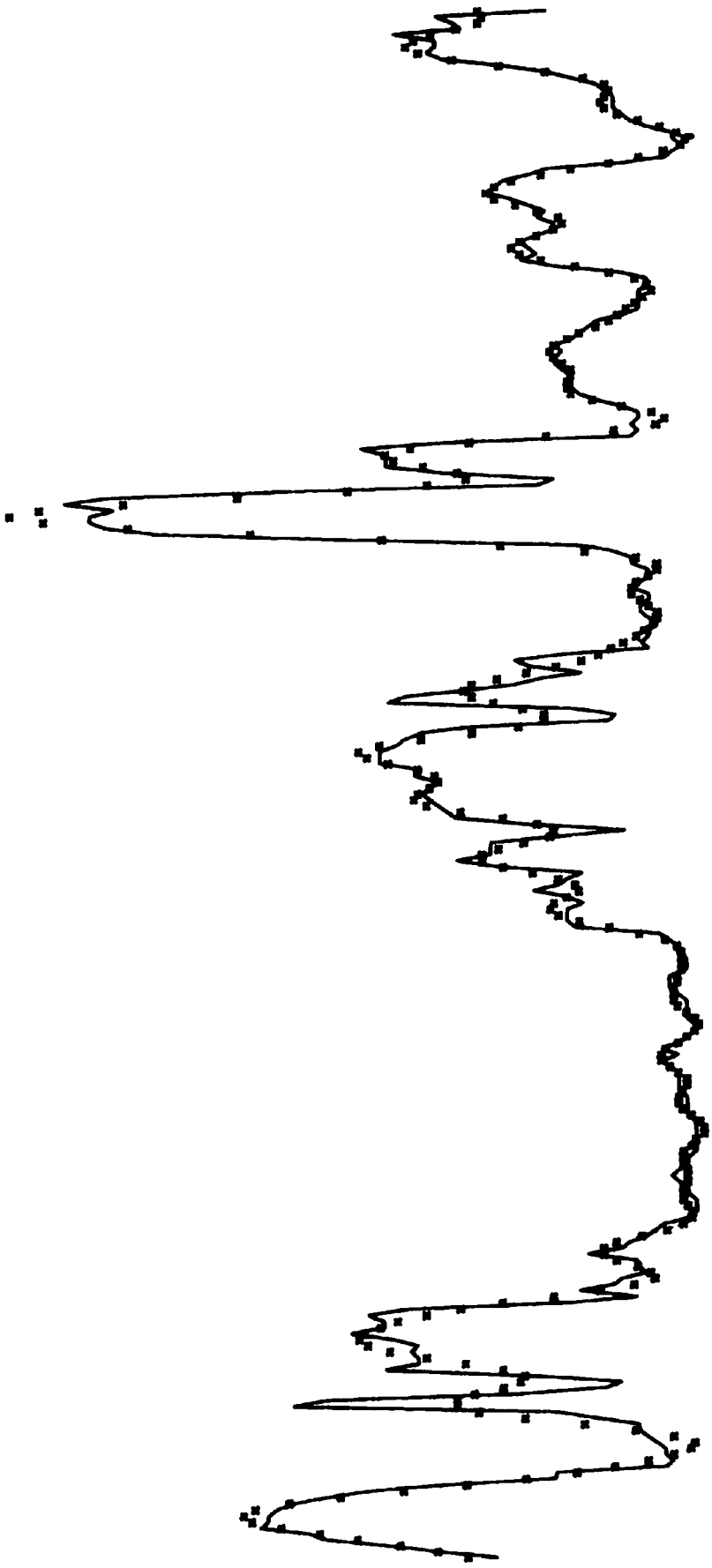


Figure C-8 Line # 200 of "CROWD" before and after being prefiltered using the one-dimensional simulation program for a simulated viewing distance of 64 times picture height.

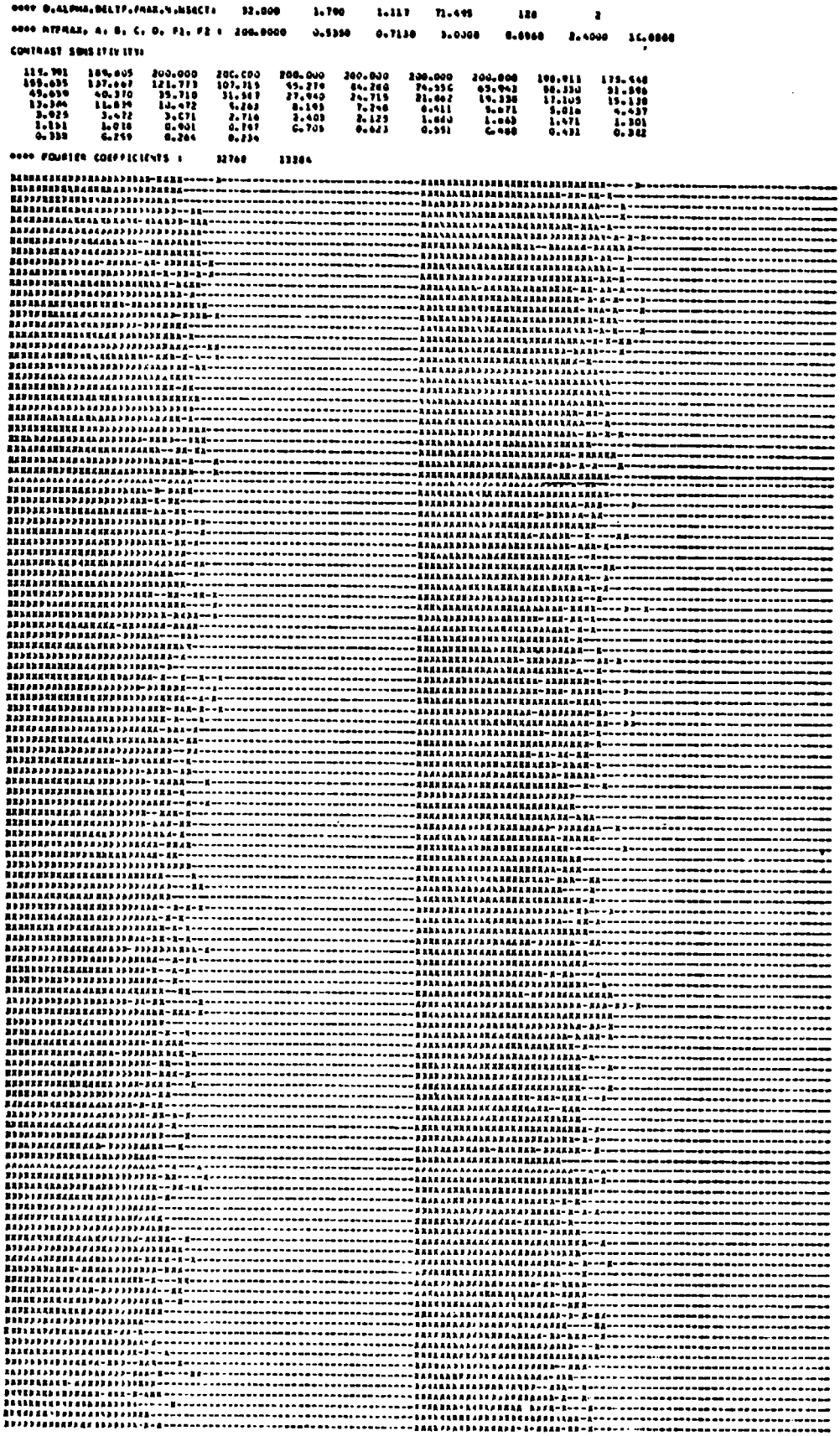


Figure C-9 Fourier domain for "CROWD" after being prefiltered using the one-dimensional simulation program for a simulated viewing distance of 32 times picture height.

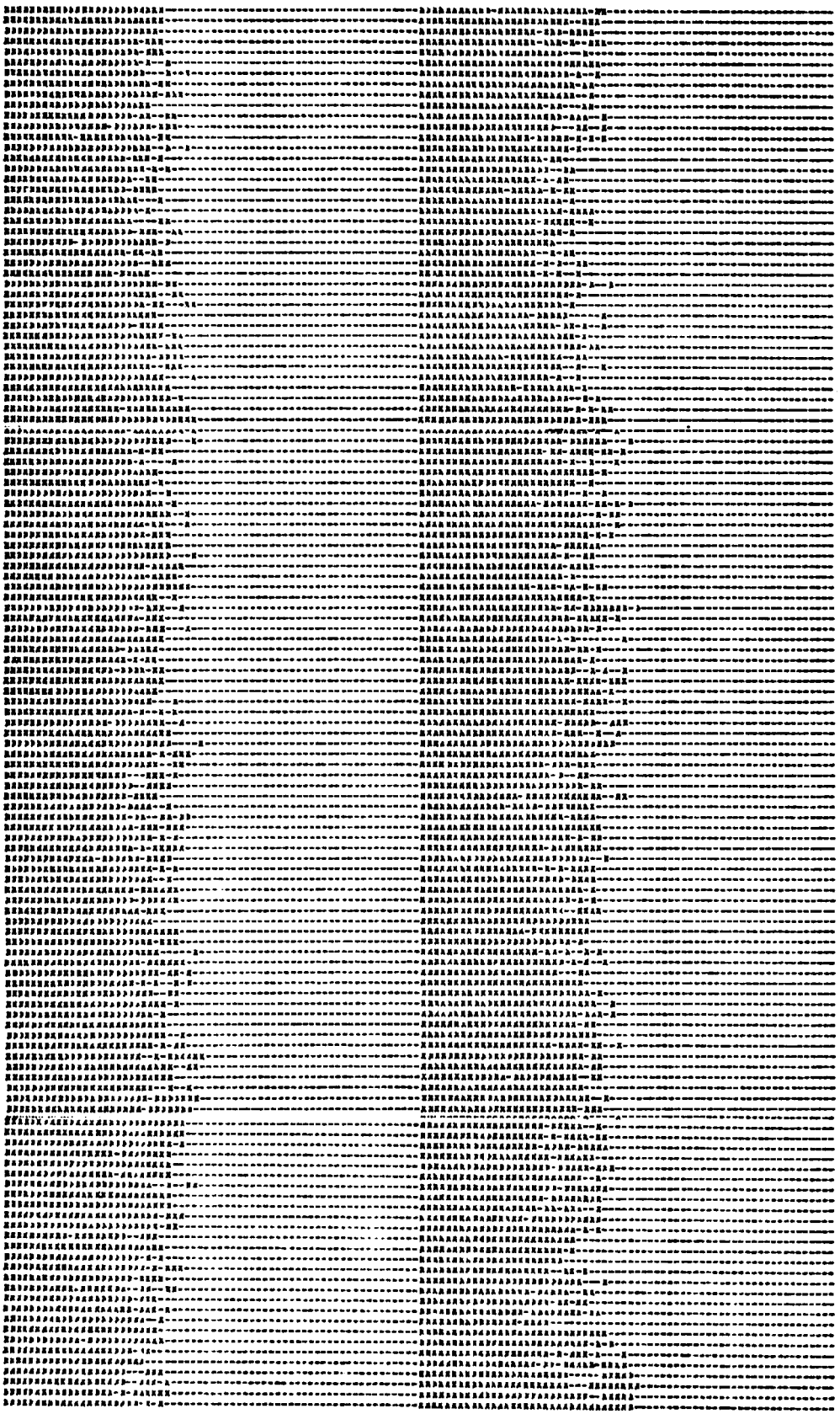


Figure C-9, continued.

APPENDIX D

The computer program for
the two-dimensional simulation

```
                C                                                                1.000
                C                                                                2.000
                C           THIS PROGRAM COMPUTES THE TWO-DIMENSIONAL FOURIER    3.000
                C           TRANSFORM OF PICTURE SEGMENTS ONE DEGREE SQUARE IN   4.000
                C           VISUAL ANGLE. THE RESULT IS FILTERED IN THE FOURIER  5.000
                C           DOMAIN USING THE SINE-WAVE GRATING CONTRAST SENSITIVITY 6.000
                C           FUNCTION FOR HUMAN VISION. THE RESULT IS INVERSE     7.000
                C           TRANSFORMED AND COMPRESSED USING ZERO-ORDER           8.000
                C           INTERPOLATION.                                       9.000
                C                                                                10.000
                C           INPUT PARAMETERS :                                   11.000
                C                                                                12.000
                C           ERROR TOLERANCE FOR COMPRESSION OF THE ORIGINAL      13.000
                C           PICTURE VIEWING DISTANCE                             14.000
                C           ERROR TOLERANCE FOR COMPRESSION OF THE FILTERED PICTURE 15.000
                C                                                                16.000
                C           I/O UNITS USED :                                     17.000
                C                                                                18.000
                C            UNIT #0 - THE SOURCE CONTAINING THE ORIGINAL PICTURE 19.000
                C            UNIT #1 - THE DESTINATION FOR THE FILTERED PICTURE  20.000
                C            UNIT #2 - A TEMPORARY FILE(F.P. SEGMENTS OF ORIGINAL) 21.000
                C            UNIT #3 - A TEMPORARY FILE(F.P. SEGMENTS OF FILTERED) 22.000
                C            UNIT #4 - THE DESTINATION FOR THE RUN-LENGTH ENCODED 23.000
                C                      FILTERED PICTURE                          24.000
                C            UNIT #5 - A TEMPORARY FILE(F.P. FILTERED PICTURE)   25.000
                C                                                                26.000
0001                    EXTERNAL GETSPA,FREESP                                   27.000
                C                                                                28.000
0002                    LOGICAL*1 PICT(256),FGRAPH(136,256)                      29.000
0003                    LOGICAL*1 STAR/'*'/,MARK/'X'/,DASH/'-'/                  30.000
0004                    LOGICAL*1 LDIGIT(4),NINE/'9'/                            31.000
0005                    LOGICAL*1 LPICT(256)                                     32.000
                C                                                                33.000
0006                    INTEGER IGRAPH(34,256),DIM(2)                            34.000
0007                    INTEGER CODE,YSECT,FRQMAX,FRQCNT,XMAX,YMAX               35.000
0008                    INTEGER SWITCH,LENGTH,JUNK,ADDR,ADRGF,ADDRSS             36.000
0009                    INTEGER ZERO/240/,DIGIT/0/,IBLANK/'    '/,BMAX           37.000
0010                    INTEGER SEGMNT/Z06000000/                                38.000
                C                                                                39.000
0011                    INTEGER*2 LABEL(18),LABLEN,LLEN                          40.000
                C                                                                41.000
0012                    REAL CONSNS(129),RPICT(2)                                42.000
                C                                                                43.000
0013                    COMPLEX LINE(2)                                          44.000
                C                                                                45.000
0014                    EQUIVALENCE (RPICT,LINE)                                 46.000
0015                    EQUIVALENCE (IGRAPH,FGRAPH)                              47.000
0016                    EQUIVALENCE (DIGIT,LDIGIT)                               48.000
                C                                                                49.000
0017                    PI=3.1415927                                             50.000
                C                                                                51.000
                C               READ THE PICTURE PARAMETERS                      52.000
                C                                                                53.000
0018                 50 CALL READ(LABEL,LABLEN,0,LNUM,0,&195)                    54.000
0019                    WRITE (8,200) (LABEL(I),I=1,4)                           55.000
0020                    XMAX=LABEL(1)
```

```
0021            YMAX=LABEL(2)                                                    56.000
0022            BMAX=LABEL(3)                                                    57.000
0023            MAX=BMAX+1                                                       58.000
0024            SWITCH=0                                                         59.000
                                                                                 60.000
                C
                C               READ ERROR TOLERANCE FOR ORIGINAL PICTURE        61.000
                C                                                                62.000
0025        20  PRINT 201                                                        63.000
0026            READ (9,202,END=100) ERROR1                                      64.000
0027            WRITE (8,203) ERROR1                                             65.000
                                                                                 66.000
                C
                C               ALLOCATE STORAGE FOR ONE PICTURE LINE            67.000
                C                                                                68.000
0028            IF (SWITCH.EQ.1) GO TO 30                                        69.000
0029            SWITCH=1                                                         70.000
0030            LENGTH=SEGMNT+4*XMAX                                             71.000
0031            CALL RCALL(GETSPA,2,0,LENGTH,2,JUNK,ADDRSS,&191)                 72.000
0032            ADDR=(ADDRSS-ADROF(RPICT))/4                                     73.000
0033            GO TO 40                                                         74.000
                                                                                 75.000
                C
                C               READ PICTURE LINE-BY-LINE AND CONVERT TO FLOATING POINT  76.000
                C                                                                77.000
0034        30  CALL REWIND(0)                                                   78.000
0035            CALL READ(LABEL,LABLEN,0,LNUM,0,&190)                            79.000
0036        40  CODE=1                                                           80.000
0037            IOCODE=0                                                         81.000
0038            LABEL(2)=1                                                       82.000
0039            DO 70 I=1,YMAX                                                   83.000
0040            CALL READ(PICT,LLEN,0,LNUM,0,&190)                               84.000
0041            DO 60 J=1,XMAX                                                   85.000
0042            LDIGIT(4)=PICT(J)                                                86.000
0043        60  RPICT(J+ADDR)=DIGIT                                              87.000
                                                                                 88.000
                C
                C               COMPRESS THE PICTURE LINE-BY-LINE                89.000
                C                                                                90.000
0044            CALL PRDICT(LABEL,RPICT(ADDR+1),ERRCR1,RPICT(ADDR+1),CODE,IOCODE)  91.000
0045            IF (CODE.GT.0) GO TO 193                                         92.000
0046        70  CONTINUE                                                         93.000
0047            GO TO 20                                                         94.000
                                                                                 95.000
                C
                C               READ D, COMPLTE SPATIAL FREQUENCIES              96.000
                C                                                                97.000
0048        100 IF (SWITCH.EQ.0) GO TO 105                                       98.000
0049            CALL RCALL(FREESP,2,0,ADDRSS,0,&192)                             99.000
0050            SWITCH=0                                                        100.000
0051        105 PRINT 205                                                       101.000
0052            READ (9,202,END=50) DIST                                        102.000
0053            ALPHA=ATAN(1.0/(2.0*DIST))*(180.0/PI)*2.0                       103.000
0054            R=XMAX/ALPHA                                                    104.000
0055            IF (ALPHA.LE.1.0) R=XMAX                                        105.000
0056            N=0                                                             106.000
0057        110 IF (R.LT.2.0) GO TO 115                                         107.000
0058            R=R/2.0                                                         108.000
0059            N=N+1                                                           109.000
0060            GO TO 110                                                       110.000
```

```
                C                                                                111.000
0061            115   N=2**N                                                     112.000
0062                  NSECT=XMAX/N                                               113.000
0063                  DELTF=NSECT/ALPHA                                          114.000
0064                  FMAX=N*DELTF/2.0                                           115.000
0065                  DIM(1)=N                                                   116.000
0066                  NY=N                                                       117.000
0067                  IF (N.GT.YMAX) NY=YMAX                                     118.000
0068                  YSECT=YMAX/N                                               119.000
0069                  IF (YSECT.EQ.0) YSECT=1                                    120.000
0070                  DIM(2)=NY                                                  121.000
0071                  WRITE (8,220) DIST,ALPHA,DELTF,FMAX,N,NSECT,NY,YSECT       122.000
                                                                                 123.000
                C                                                                124.000
                C             COMPUTE AND FORMAT MTF                             125.000
                C                                                                126.000
0072                  JMAX=N/2                                                   127.000
0073                  JMAXY=NY/2                                                 128.000
0074                  CALL EYEMTF(CONSNS,JMAX,DELTF,DELTF)                       129.000
0075                  WRITE (8,240) (CONSNS(I),I=1,JMAX)                         130.000
                C                                                                131.000
                C             TRANSFORM THE PICTURE SEGMENT-BY-SEGMENT           132.000
                C                                                                133.000
0076                  CALL REWIND(2)                                             134.000
0077                  IMAX=JMAX+1                                                135.000
0078                  IMAXY=JMAXY+1                                              136.000
0079                  FRQMAX=(IMAX*NY-2*(NY-IMAXY))*NSECT*YSECT                  137.000
0080                  FRQCNT=FRQMAX                                              138.000
                C                                                                139.000
                C             ALLOCATE SPACE FOR A ROW OF SEGMENTS               140.000
                C                                                                141.000
0081                  LENGTH=SEGMNT+(XMAX/2+1)*NY*8                              142.000
0082                  CALL RCALL(GETSPA,2,0,LENGTH,2,JUNK,ADDRSS,&191)           143.000
0083                  ADDR=(ADDRSS-ADRCF(RPICT))/4                               144.000
0084                  CALL REWIND(0)                                             145.000
0085                  CALL READ(LABEL,LABLEN,0,LNUM,0,&190)                      146.000
                C                                                                147.000
                C             READ IN EACH ROW OF SEGMENTS AND CONVERT TO REAL   148.000
                C                                                                149.000
0086                  DO 139 LOOP1=1,YSECT                                       150.000
0087                  DO 137 I=1,NY                                              151.000
0088                  L=(I-1)*XMAX                                               152.000
0089                  CALL READ(PICT,LLEN,0,LNUM,0,&190)                         153.000
0090                  DO 137 J=1,XMAX                                            154.000
0091                  DIGIT(4)=PICT(J)                                           155.000
0092                  K=L+J                                                      156.000
0093            137   RPICT(K+ADDR)=DIGIT                                        157.000
                C                                                                158.000
                C             WRITE THE SEGMENTS IN SEQUENCE ON UNIT 2           159.000
                C                                                                160.000
0094                  LLEN=N*4                                                   161.000
0095                  DO 139 I=1,NSECT                                           162.000
0096                  L=(I-1)*N+1                                                163.000
0097                  DO 139 J=1,NY                                              164.000
0098                  K=L+(J-1)*XMAX                                             165.000
0099            139   CALL WRITE(RPICT(K+ADDR),LLEN,0,LNUM,2,&194)
```

```
                C                                                                166.000
                C               DEALLOCATE SCRE STORAGE - KEEP ENOUGH            167.000
                C               FOR ONE SEGMENT                                  168.000
                C                                                                169.000
0100                    IF (NSECT.EQ.1) GO TO 141                                170.000
0101                    LENGTH=LENGTH-(N/2+1)*NY*8                               171.000
0102                    JUNK=ADDRSS+(N/2+1)*NY*8                                 172.000
0103                    CALL RCALL(FREESP,2,LENGTH,JUNK,0,&192)                  173.000
                C                                                                174.000
                C               READ IN EACH SEGMENT AND TRANSFORM IT            175.000
                C                                                                176.000
0104            141     CALL REWIND(2)                                           177.000
0105                    CALL REWIND(3)                                           178.000
0106                    PRINT 28o                                                179.000
0107                    READ (9,282) IOCODE                                      180.000
0108                    JJJ=2-IMAX                                               181.000
0109                    DO 160 LOOP1=1,NSECT                                     182.000
0110                    DO 160 LOOP2=1,NSECT                                     183.000
0111                    JJJ=JJJ+IMAX                                             184.000
0112                    DO 142 I=1,NY                                            185.000
0113                    K=(I-1)*N+1                                              186.000
0114            142     CALL READ(RPICT(ADDR+K),LLEN,0,LNUM,2,&190)              187.000
0115                    CALL FOUR2(RPICT(ADDR+1),DIM,2,-1,0)                     188.000
0116                    IF (IOCODE.LT.2) GO TO 143                               189.000
0117                    K=2*(N/2+1)*NY                                           190.000
0118                    WRITE (8,270) (RPICT(ADDR+J),J=1,K)                      191.000
                C                                                                192.000
                C               COMPUTE RELATIVE AMPLITUDES OF SPATIAL FREQUENCIES 193.000
                C               AND FILTER OUT SUB-THRESHOLD COMPONENTS          194.000
                C                                                                195.000
0119            143     ADDR=(ADDRSS-ADROF(LINE))/8                              196.000
0120                    FGRAPH(JJJ,IMAXY)=MARK                                   197.000
0121                    B=REAL(LINE(ADDR+1))                                     198.000
0122                    CONSNS(1)=B/(N*NY)                                       199.000
0123                    L=2                                                      200.000
0124                    DO 155 I=1,NY                                            201.000
0125                    II=I+IMAXY-1                                             202.000
0126                    K=I                                                      203.000
0127                    IF (I.LT.IMAXY) GO TO 144                                204.000
0128                    II=I-IMAXY+1                                             205.000
0129                    K=2*IMAXY-I                                              206.000
0130            144     JSTART=1                                                 207.000
0131                    IF (I.EQ.1) JSTART=2                                     208.000
0132                    DO 152 J=JSTART,IMAX                                     209.000
0133                    JJ=JJJ+J-1                                               210.000
0134                    FGRAPH(JJ,II)=MARK                                       211.000
0135                    C=CABS(LINE(ADDR+L))*2.0                                 212.000
0136                    IF ((I.EQ.1).AND.(J.EQ.IMAX)) C=C/2.0                    213.000
0137                    IF ((J.EQ.1).AND.(K.EQ.IMAXY)) C=C/2.0                   214.000
0138                    IF ((I.EQ.IMAXY).AND.(J.EQ.IMAX)) C=C/2.0                215.000
0139                    CONSNS(J)=C/(N*NY)                                       216.000
0140                    IF (C.LT.0.00001) GO TO 150                              217.000
0141                    D=B/C                                                    218.000
0142                    R=SQRT((K-1)**2.0+(J-1)**2.0)*DELTF                      219.000
0143                    CALL EVENTF(E,1,R,R)                                     220.000
```

```
0144              IF (C.LE.E) GO TO 152                                                  221.000
0145        150   IF ((I.LE.IMAXY).OR.((J.NE.1).AND.(J.NE.IMAX))) FRQCNT=FRQCNT-1       222.000
0146              FGRAPH(JJ,II)=DASH                                                     223.000
0147              LINE(ADDR+L)=(0.0,0.0)                                                 224.000
0148        152   L=L+1                                                                  225.000
0149              IF (IOCODE.EQ.0) GO TO 155                                             226.000
0150              WRITE (8,270) (CONSNS(J),J=1,IMAX)                                     227.000
0151        155   CONTINUE                                                               228.000
                  C                                                                    229.000
                  C              COMPUTE INVERSE TRANSFORM                             230.000
                  C              WRITE RESULT ON I/O UNIT 3                            231.000
                  C              (THE RESULT IS THE FILTERED PICTURE)                  232.000
                  C                                                                    233.000
0152              CALL FOUR2(LINE(ADDR+1),DIM,2,1,-1)                                    234.000
0153              K=N*NY*4                                                               235.000
0154              ADDR=ADDR*2                                                            236.000
0155              NREC=0                                                                 237.000
0156              KK=1                                                                   238.000
0157        1552  LLEN=K                                                                 239.000
0158              IF (K.GT.32764) LLEN=32764                                             240.000
0159              CALL WRITE(RPICT(ADDR+KK),LLEN,0,LNUM,3,&194)                          241.000
0160              NREC=NREC+1                                                            242.000
0161              K=K-LLEN                                                               243.000
0162              KK=KK+LLEN/4                                                           244.000
0163              IF (K.GT.0) GO TO 1552                                                 245.000
                  C                                                                    246.000
                  C              GRAPH THE FOURIER COEFFICIENTS                        247.000
                  C                                                                    248.000
0164              IF (LOOP2.EQ.NSECT) GO TO 156                                          249.000
0165              IF ((JJ+IMAX).LT.133) GO TO 160                                        250.000
0166        156   LLEN=JJ                                                                251.000
0167              WRITE (8,213) FRQMAX,FRQCNT                                            252.000
0168              DO 157 J=1,NY                                                          253.000
0169              FGRAPH(1,J)=NINE                                                       254.000
0170        157   CALL WRITE(FGRAPH(1,J),LLEN,0,LNUM,8)                                  255.000
0171              JJJ=2-IMAX                                                             256.000
0172        160   CONTINUE                                                               257.000
0173              MIN=1                                                                  258.000
0174              IF (MAX.LT.NY) MAX=NY                                                  259.000
                  C                                                                    260.000
                  C              DEALLOCATE SEGMENT STORAGE                            261.000
                  C                                                                    262.000
0175              CALL RCALL(FREESP,2,0,ADDRSS,0,&192)                                   263.000
                  C                                                                    264.000
                  C              RECONSTRUCT PICTURE FROM SEGMENTS - ALLOCATE STORAGE  265.000
                  C                                                                    266.000
0176              LENGTH=SEGMNT*4*XMAX*NY*4*XMAX                                         267.000
0177              CALL RCALL(GETSPA,2,0,LENGTH,2,JUNK,ADDRSS,&191)                       268.000
0178              ADDR=(ADDRSS-ADROF(RPICT))/4                                           269.000
                  C                                                                    270.000
                  C              READ IN A ROW OF SEGMENTS                             271.000
                  C                                                                    272.000
0179              CALL REWIND(3)                                                         273.000
0180              CALL REWIND(5)                                                         274.000
0181              DO 1606 LOOP1=1,YSECT                                                  275.000
```

```
0182              DO 1602 I=1,NSECT                                                  276.000
0183              LLEN=0                                                             277.000
0184              DO 1602 J=1,NREC                                                   278.000
0185              K=(I-1)*NY*N+1+(J-1)*LLEN/4                                        279.000
0186        1602  CALL READ(RPICT(K+ADDR),LLEN,0,LNUM,3,&190)                        280.000
                  C                                                                281.000
                  C              FORM PICTURE LINE-BY-LINE                         282.000
                  C              WRITE RESULTS ON I/O UNIT 5                       283.000
                  C                                                                284.000
0187              NN=XMAX*NY+ADDR                                                    285.000
0188              LLEN=XMAX*4                                                        286.000
0189              A=N*NY                                                             287.000
0190              DO 1606 I=1,NY                                                     288.000
0191              L=(I-1)*N+ADDR                                                     289.000
0192              DO 1604 J=1,NSECT                                                  290.000
0193              K=(J-1)*N*NY                                                       291.000
0194              LL=(J-1)*N+NN                                                      292.000
0195              DO 1604 M=1,N                                                      293.000
0196        1604  RPICT(LL+M)=RPICT(L+K+M)/A                                         294.000
0197        1606  CALL WRITE(RPICT(NN+1),LLEN,0,LNUM,5,&198)                         295.000
                  C                                                                296.000
                  C              DEALLOCATE ALL BUT ONE PICTURE LINE                297.000
                  C                                                                298.000
0198              LENGTH=4*XMAX*NY                                                   299.000
0199              JUNK=ADDRSS+4*XMAX                                                 300.000
0200              CALL RCALL(FREESP,2,LENGTH,JUNK,0,&192)                            301.000
0201              SWITCH=1                                                           302.000
                  C                                                                303.000
                  C              READ IN ERROR TOLERANCE FOR FILTERED PICTURE       304.000
                  C                                                                305.000
0202        161   PRINT 210                                                          306.000
0203              READ (9,202,END=100) ERROR2                                        307.000
0204              WRITE (8,203) ERROR2                                               308.000
                  C                                                                309.000
                  C              COMPRESS THE FILTERED PICTURE                     310.000
                  C                                                                311.000
0205              CALL REWIND(0)                                                     312.000
0206              CALL READ(LABEL,LABLEN,0,LNUM,0,&190)                              313.000
0207              CALL WRITE(LABEL,LABLEN,0,LNUM,1,&198)                             314.000
0208              CALL REWIND(5)                                                     315.000
0209              PRINT 280                                                          316.000
0210              READ (9,282) NPLOTT                                                317.000
0211              NPLOT=NPLOTT                                                       318.000
0212              IOCODE=0                                                           319.000
0213              LABEL(2)=1                                                         320.000
0214              CODE=1                                                             321.000
0215              DO 176 LOOP1=1,YMAX                                                322.000
0216              CALL READ(RPICT(ADDR+1),LLEN,0,LNUM,5,&190)                        323.000
0217              CALL PRDICT(LABEL,RPICT(ADDR+1),ERROR2,RPICT(ADDR+1),CODE,IOCODE)  324.000
0218              IF (CODE.GT.0) GO TO 180                                           325.000
                  C                                                                326.000
                  C              OUTPUT THE COMPRESSED PICTURE                     327.000
                  C                                                                328.000
0219              DO 164 LOOP2=1,XMAX                                                329.000
0220              DIGIT=RPICT(ADDR+LOOP2)+0.001                                      330.000
```

```
0221              IF (DIGIT.LT.0) DIGIT=0                                          331.000
0222              IF (DIGIT.GT.8MAX) DIGIT=8MAX                                    332.000
0223        164   PICT(LOOP2)=LDIGIT(4)                                            333.000
0224              LLEN=XMAX                                                        334.000
0225              CALL WRITE(PICT,LLEN,0,LNUM,1,&198)                              335.000
            C                                                                      336.000
            C               GRAPH EACH PICTURE LINE                                337.000
            C                                                                      338.000
0226              CALL READ(LPICT,LLEN,0,LNUM,0,&195)                              339.000
0227              IF (NPLOT.NE.LOOP1) GO TO 176                                    340.000
0228              NPLOT=NPLOT+NPLOTT                                               341.000
0229              DO 174 LOOP5=1,XMAX,128                                          342.000
0230              WRITE (8,211)                                                    343.000
0231              DO 170 I=MIN,MAX                                                 344.000
0232              DO 170 J=1,34                                                    345.000
0233        170   IGRAPH(J,I)=IBLANK                                               346.000
0234              MIN=256                                                          347.000
0235              MAX=0                                                            348.000
0236              LEN=LOOP5+128-1                                                  349.000
0237              IF (LEN.GT.XMAX) LEN=XMAX                                        350.000
0238              DO 172 LOOP8=LOOP5,LEN                                           351.000
0239              I=LOOP8-LOOP5+6                                                  352.000
0240              LDIGIT(4)=PICT(LOOP8)                                            353.000
0241              INDEX=DIGIT+1                                                    354.000
0242              IF (INDEX.GT.MAX) MAX=INDEX                                      355.000
0243              IF (INDEX.LT.MIN) MIN=INDEX                                      356.000
0244              FGRAPH(I,INDEX)=MARK                                             357.000
0245              LDIGIT(4)=LPICT(LOOP8)                                           358.000
0246              INDEX=DIGIT+1                                                    359.000
0247              IF (INDEX.GT.MAX) MAX=INDEX                                      360.000
0248              IF (INDEX.LT.MIN) MIN=INDEX                                      361.000
0249        172   FGRAPH(I,INDEX)=STAR                                             362.000
0250              LLEN=133                                                         363.000
0251              DO 174 I=MIN,MAX                                                 364.000
0252              FGRAPH(1,I)=NINE                                                 365.000
0253              J=I/100                                                          366.000
0254              DIGIT=ZERO+J                                                     367.000
0255              FGRAPH(2,I)=LDIGIT(4)                                            368.000
0256              K=(I-J*100)/10                                                   369.000
0257              DIGIT=ZERO+K                                                     370.000
0258              FGRAPH(3,I)=LDIGIT(4)                                            371.000
0259              J=I-J*100-K*10                                                   372.000
0260              DIGIT=ZERO+J                                                     373.000
0261              FGRAPH(4,I)=LDIGIT(4)                                            374.000
0262        174   CALL WRITE(FGRAPH(1,1),LLEN,0,LNUM,8)                            375.000
0263        176   CONTINUE                                                         376.000
0264              CODE=2                                                           377.000
0265              CALL PRDICT(LABEL,RPICT(ACCR+1),ERROR2,RPICT(ADDR+1),CODE,IOCODE) 378.000
0266              LLEN=3                                                           379.000
0267              CALL CNTRL('WTM',LLEN,1)                                         380.000
0268              GO TO 161                                                        381.000
            C                                                                      382.000
0269        180   CALL REWIND(0)                                                   383.000
0270              GO TO 161                                                        384.000
            C                                                                      385.000
```

```
                C               LABEL DATA INCORRECT                                    386.000
                C                                                                       387.000
                C                                                                       388.000
0271            190   PRINT 235                                                       389.000
0272                  GO TO 100                                                       390.000
                C                                                                       391.000
                C               STORAGE COULD NOT BE ACQUIRED FROM GETSPACE             392.000
                C                                                                       393.000
0273            191   PRINT 231                                                       394.000
0274                  STOP 1                                                          395.000
                C                                                                       396.000
                C               ERROR RETURN FROM FREESPACE                             397.000
                C                                                                       398.000
0275            192   PRINT 232                                                       399.000
0276                  STOP 1                                                          400.000
                C                                                                       401.000
                C               ORIGINAL PICTURE WILL NOT COMPRESS                      402.000
                C                                                                       403.000
0277            193   PRINT 236                                                       404.000
0278                  STOP 1                                                          405.000
                C                                                                       406.000
                C               ERROR RETURN FROM WRITE ON UNIT 3                       407.000
                C                                                                       408.000
0279            194   PRINT 233                                                       409.000
0280                  STOP 1                                                          410.000
                C                                                                       411.000
                C               ALL INPUT HAS BEEN PROCESSED                            412.000
                C                                                                       413.000
0281            195   PRINT 209                                                       414.000
0282                  STOP 0                                                          415.000
                C                                                                       416.000
                C               ERROR DURING OUTPUT OF COMPRESSED PICTURE               417.000
                C                                                                       418.000
0283            198   PRINT 260                                                       419.000
0284                  STOP 1                                                          420.000
                C                                                                       421.000
                C               FORMAT STATEMENTS                                       422.000
                C                                                                       423.000
0285            200   FORMAT('0PICTURE PARAMETERS:'/4I10)                             424.000
0286            201   FORMAT(//'0ENTER COMPRESSION ERROR FOR ORIGINAL PICTURE:')      425.000
0287            202   FORMAT(F6.3)                                                    426.000
0288            203   FORMAT('0ERROR ARGUMENT=',F6.3,' BRIGHTNESS UNITS'/)            427.000
0289            205   FORMAT(//'0ENTER VIEWING DISTANCE:')                            428.000
0290            209   FORMAT('0**** ALL INPUT HAS BEEN PROCESSED')                    429.000
0291            210   FORMAT(//'0ENTER COMPRESSION ERROR FOR FILTERED PICTURE:')      430.000
0292            211   FORMAT('0')                                                     431.000
0293            213   FORMAT('0**** FOURIER COEFFICIENTS :',2I10/)                    432.000
0294            214   FORMAT('9',I2,1X,128A1)                                         433.000
0295            215   FORMAT('9',129A1)                                               434.000
0296            220   FORMAT('0**** D,ALPHA,DELTF,FMAX,N,NSECT,NY,YSECT:',4F10.3,4I10) 435.000
0297            230   FORMAT('0M(1)=',I1)                                             436.000
0298            231   FORMAT('0**** ERROR RETURN FROM GETSPACE')                      437.000
0299            232   FORMAT('0**** ERROR RETURN FROM FREESPACE')                     438.000
0300            233   FORMAT('0**** ERROR RETURN FROM WRITE ON UNIT 3')               439.000
0301            235   FORMAT('0**** LABEL DATA INCORRECT')                            440.000
0302            236   FORMAT('0**** ORIGINAL PICTURE WILL NOT COMPRESS')
```

```
0303        240   FORMAT('-EYEMTF:'//(1CF10.3))                                  441.000
0304        250   FORMAT('0CONTRAST SENSITIVITY:'//(10F10.3))                    442.000
0305        260   FORMAT('0**** WRITE ERRCR CN I/O UNIT')                        443.000
0306        270   FORMAT(/(12F10.3))                                             444.000
0307        280   FORMAT(//'0EVERY N-TH LINE WILL BE PLOTTED. ENTER N :')        445.000
0308        282   FORMAT(I4)                                                     446.000
0309        286   FORMAT(//'0ENTER ZERO TC SUPPRESS FOURIER COEFFICIENTS :')     447.000
0310              END                                                            448.000
 *OPTIONS IN EFFECT*   IC,EBCDIC,SCURCE,ACLIST,NODECK,LOAC,NOMAP
 *OPTIONS IN EFFECT*   NAME = MAIN     , LINECNT =      57
 *STATISTICS*     SOURCE STATEMENTS =      310,PROGRAM SIZE =    45576
 *STATISTICS*   NC DIAGNJSTICS GENERATED

NO STATEMENTS FLAGGED IN THE ABOVE COMPILATIONS.
```

APPENDIX E

SAMPLE RESULTS FROM THE TWO-DIMENSIONAL SIMULATION

The first few figures in this appendix contain line-by-line comparisons from the picture "CROWD". The continuous curves are the original picture lines. The superimposed symbols represent the sample points of the same picture line after being prefiltered for the specified viewing distance.

The last figure contains a representation of the Fourier domain for the picture "CROWD". The X's represent the position of Fourier components which are above threshold for a simulated viewing distance of 32 times picture height. The structure of the Fourier domain is explained by the use of two-dimensional transformation, and the segmentation of the picture into four parts based on the viewing distance and associated visual angle.

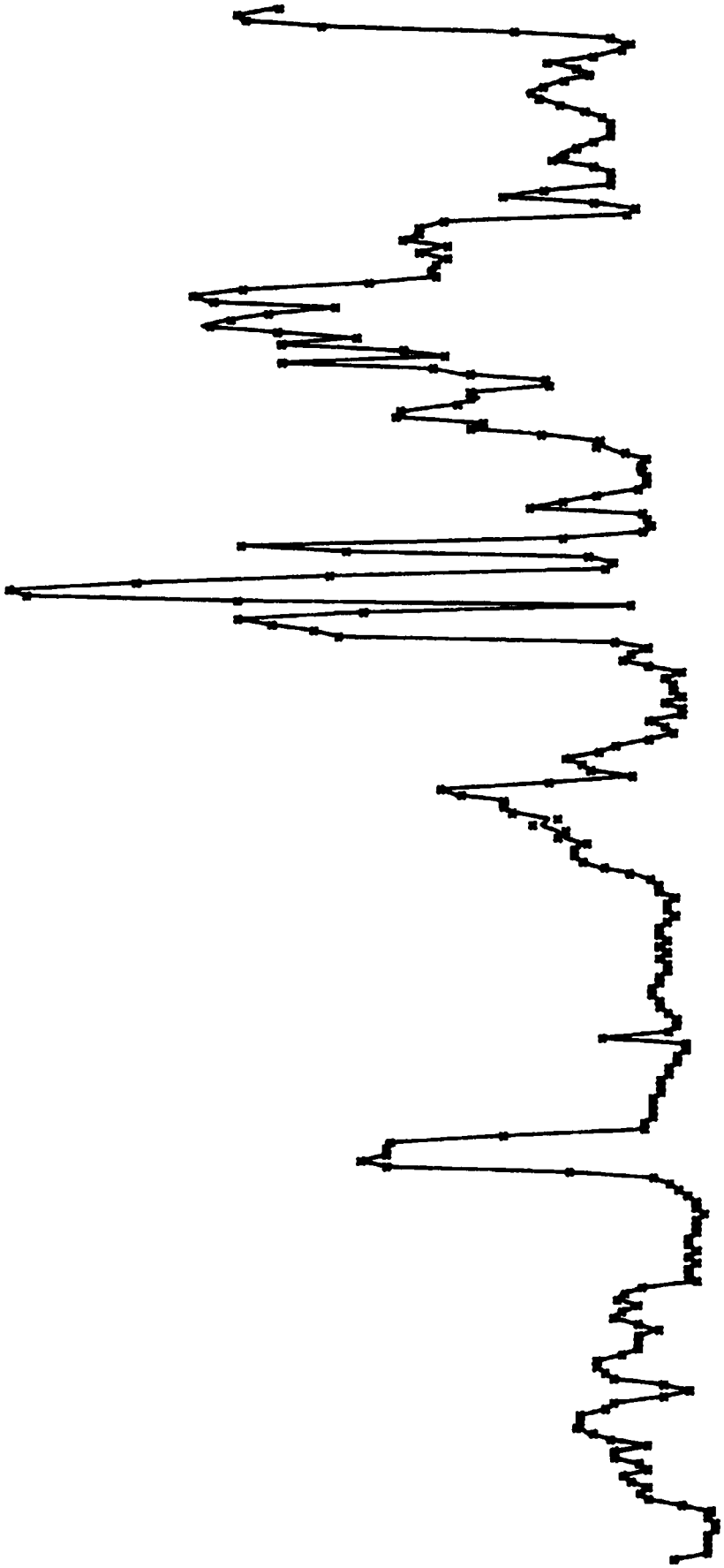


Figure E-1 Line # 100 of "CROWD" before and after being prefiltered using the two-dimensional simulation program for a simulated viewing distance of 4 times picture height.

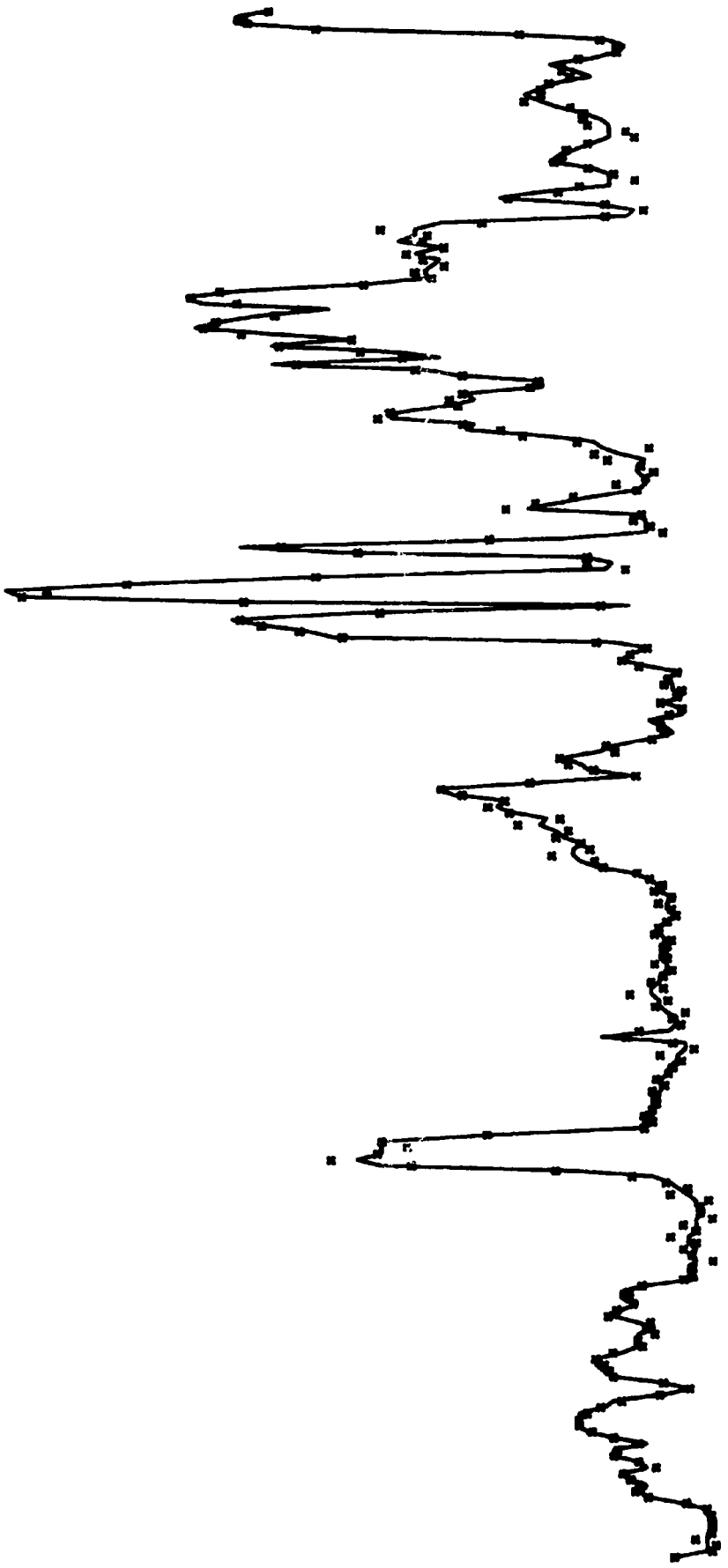


Figure E-2 Line # 100 of "CROWD" before and after being prefiltered using the two-dimensional simulation program for a simulated viewing distance of 8 times picture height.

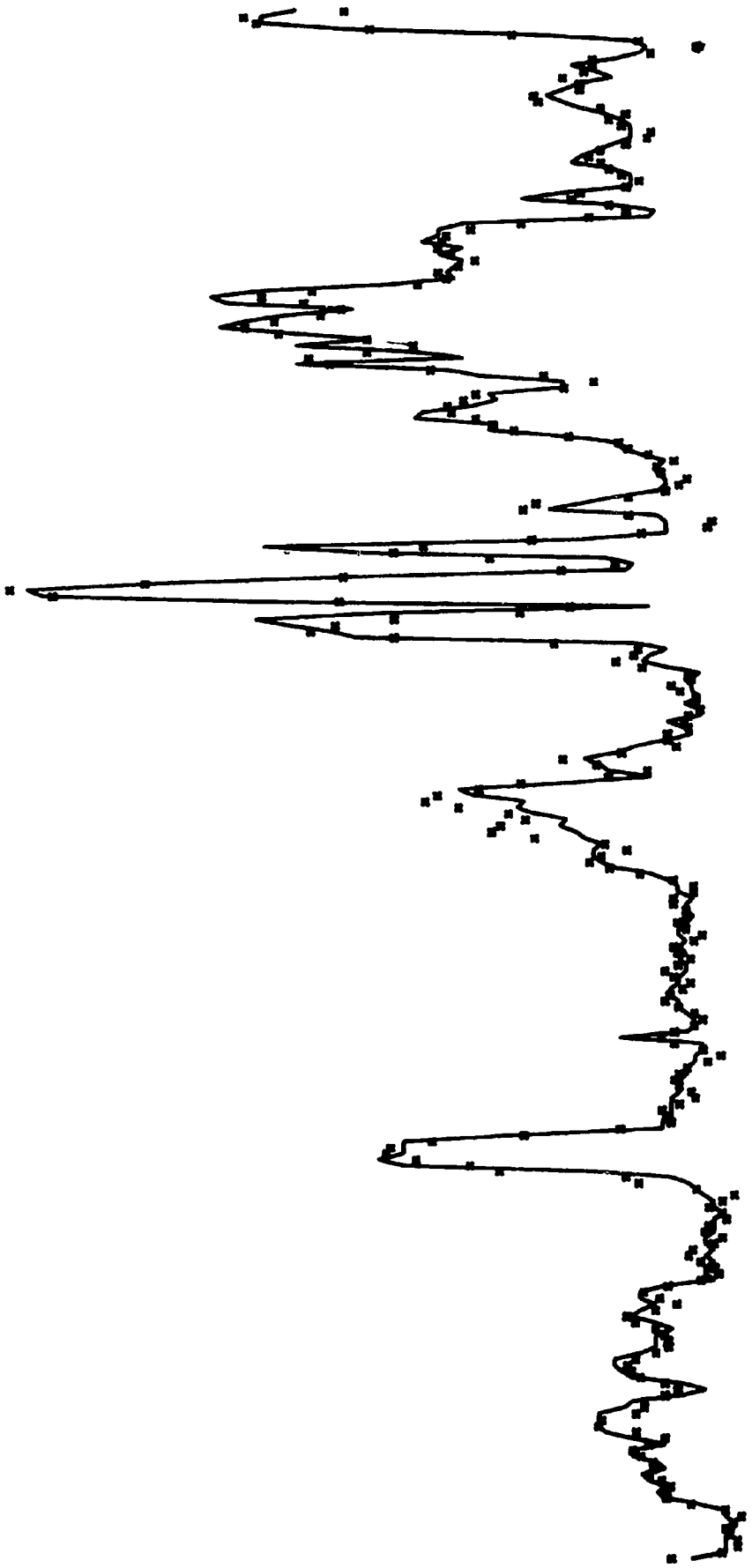


Figure E-3 Line # 100 of "CROWD" before and after being prefiltered using the two-dimensional simulation program for a simulated viewing distance of 16 times picture height.

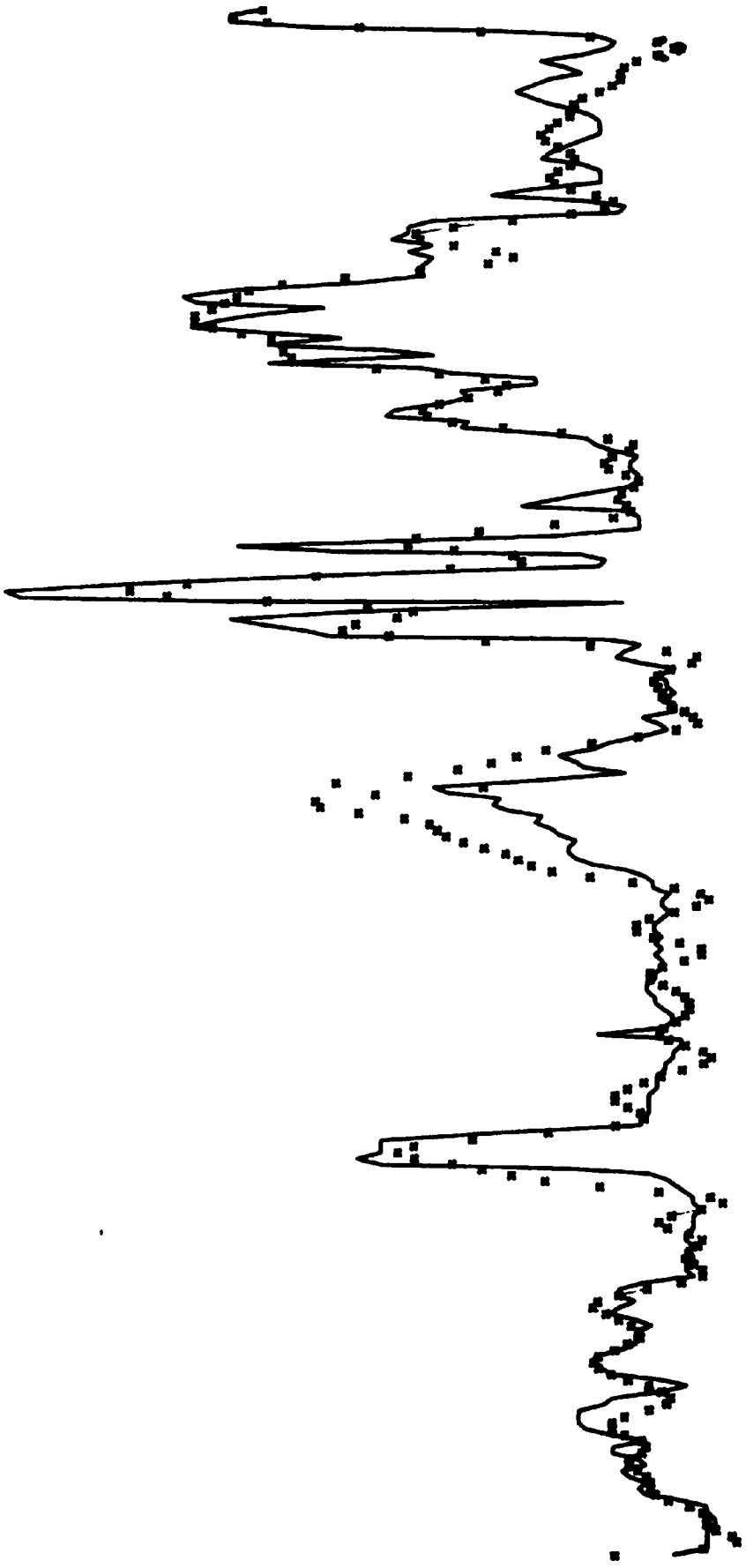


Figure E-4 Line # 100 of "CROWD" before and after being prefiltered using the two-dimensional simulation program for a simulated viewing distance of 32 times picture height.

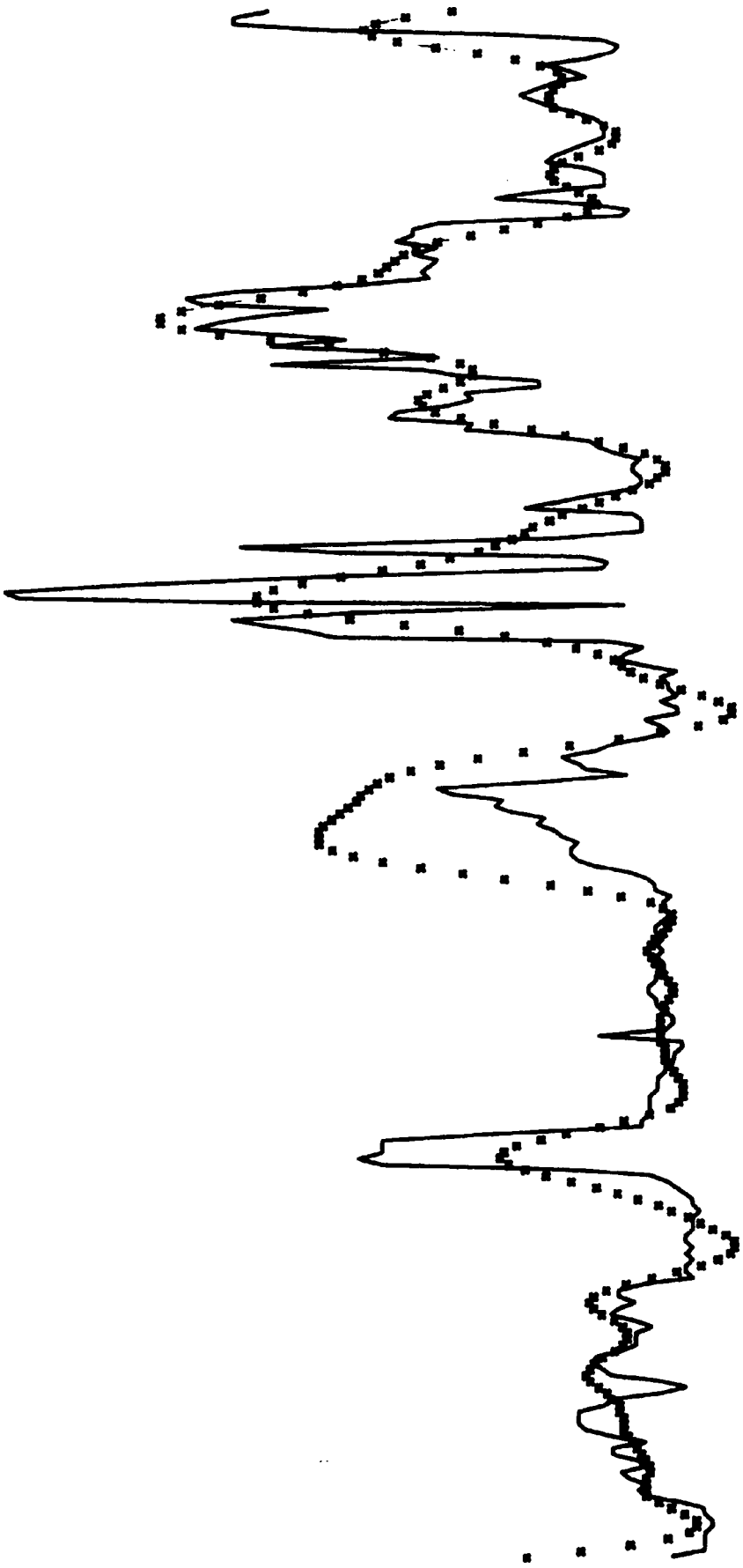


Figure E-5 Line # 100 of "CROWD" before and after being prefiltered using the two-dimensional simulation program for a simulated viewing distance of 64 times picture height.

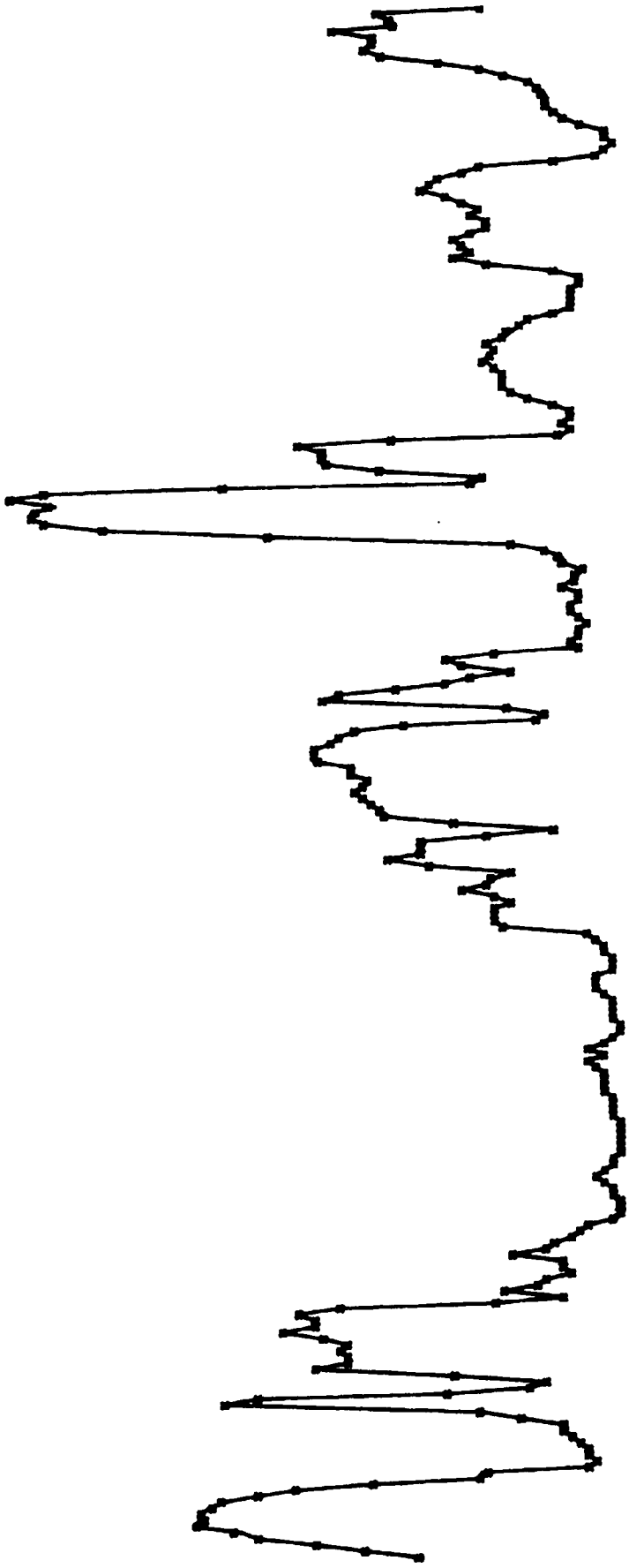


Figure E-6 Line # 200 of "CROWD" before and after being prefiltered using the two-dimensional simulation program for a simulated viewing distance of 4 times picture height.

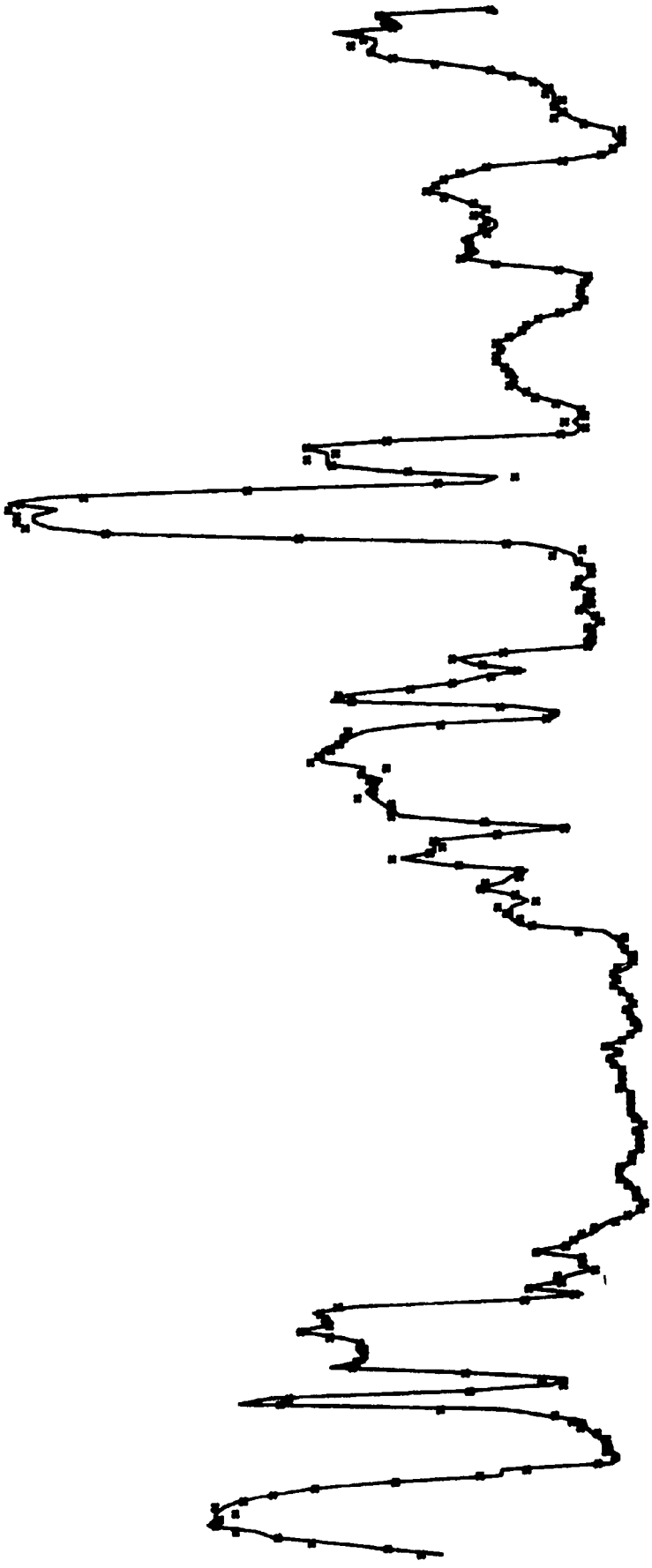


Figure E-7 Line # 200 of "CROWD" before and after being prefiltered using the two-dimensional simulation program for a simulated viewing distance of 8 times picture height.

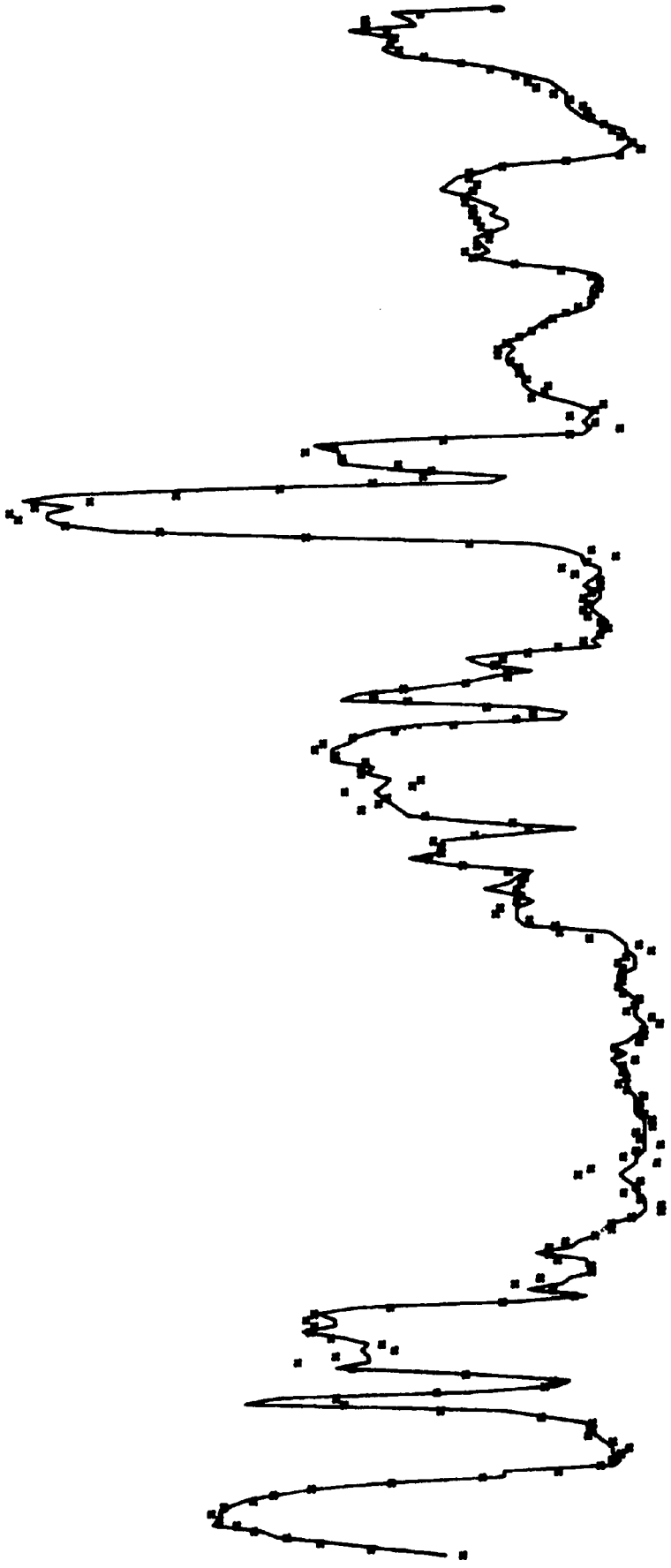


Figure E-8 Line # 200 of "CROWD" before and after being prefiltered using the two-dimensional simulation program for a simulated viewing distance of 16 times picture height.

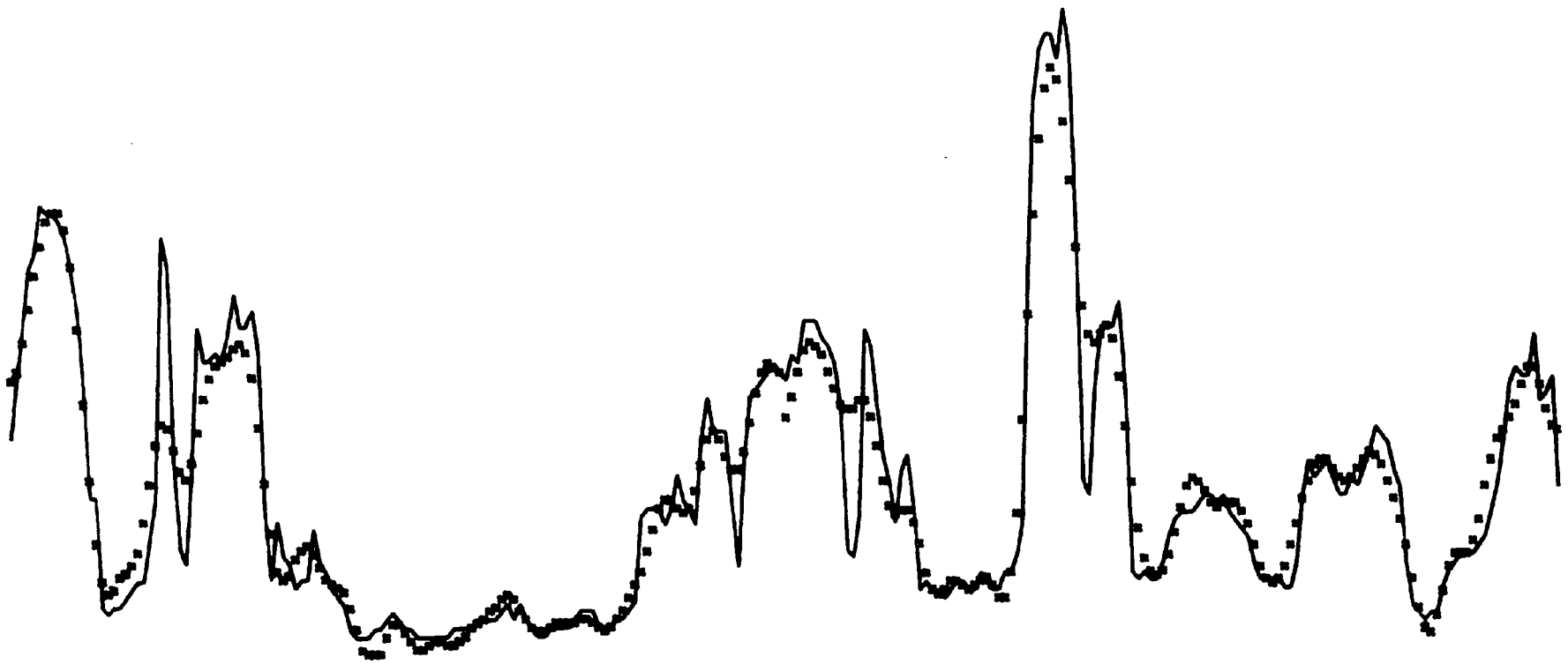


Figure E-9 Line # 200 of "CROWD" before and after being prefiltered using the two-dimensional simulation program for a simulated viewing distance of 32 times picture height.

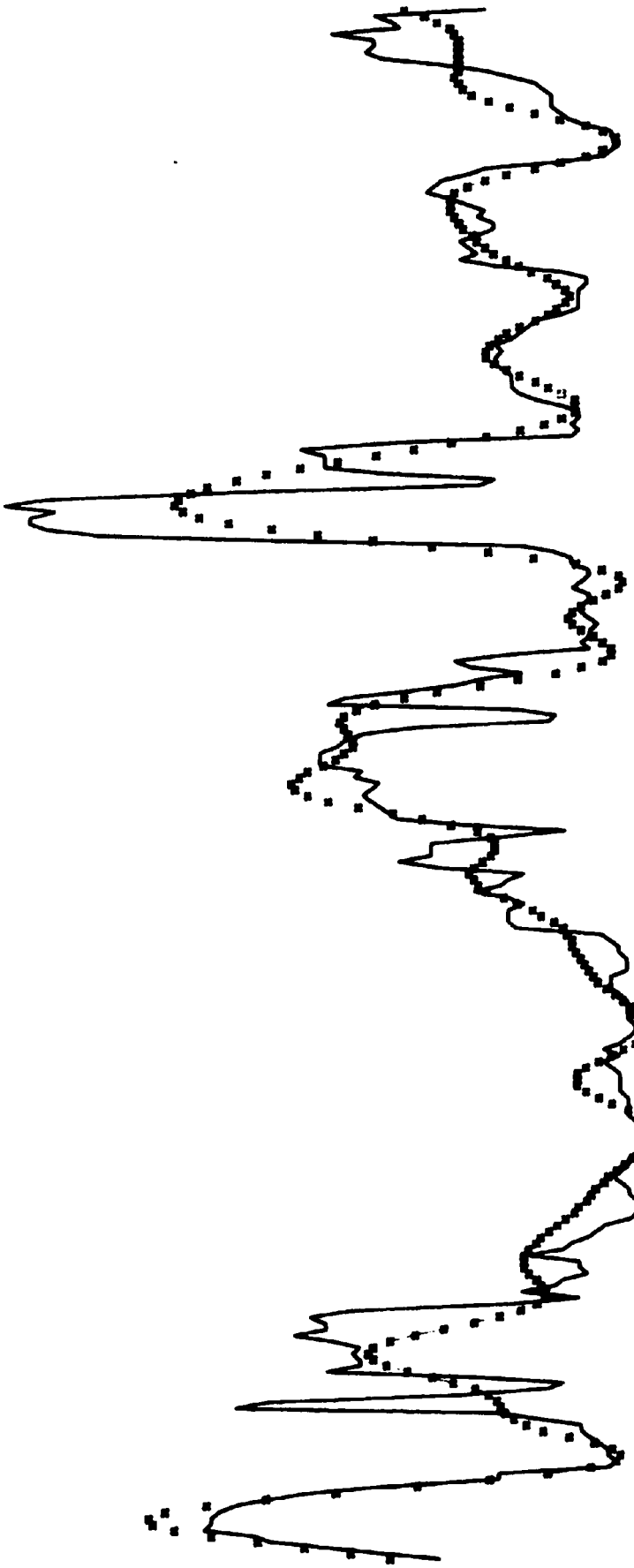


Figure E-10 Line # 200 of "CROWD" before and after being prefiltered using the two-dimensional simulation program for a simulated viewing distance of 64 times picture height.

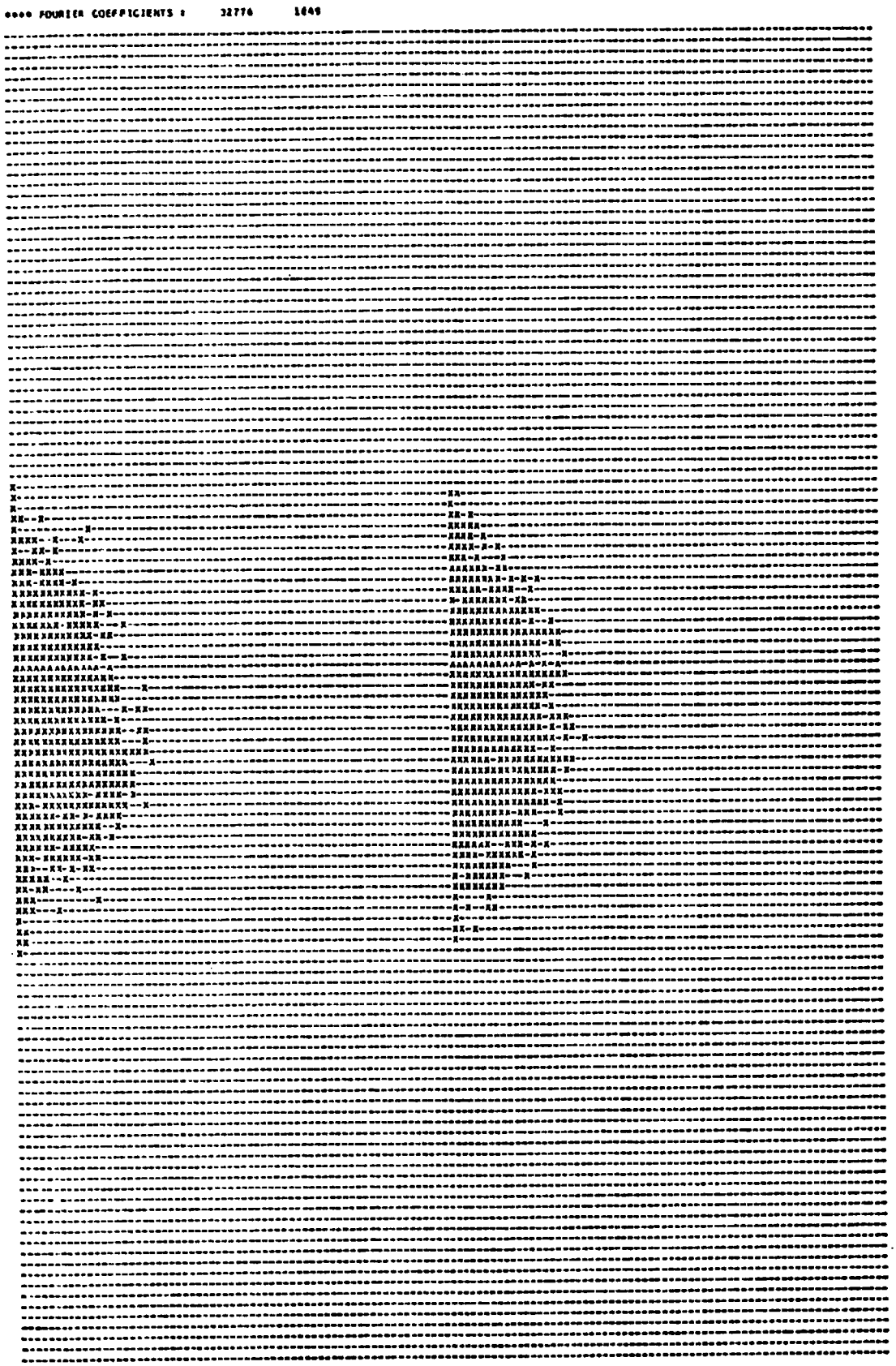


Figure E-11 Fourier domain for "CROWD" after being prefiltered using the two-dimensional simulation program for a simulated viewing distance of 32 times picture height.

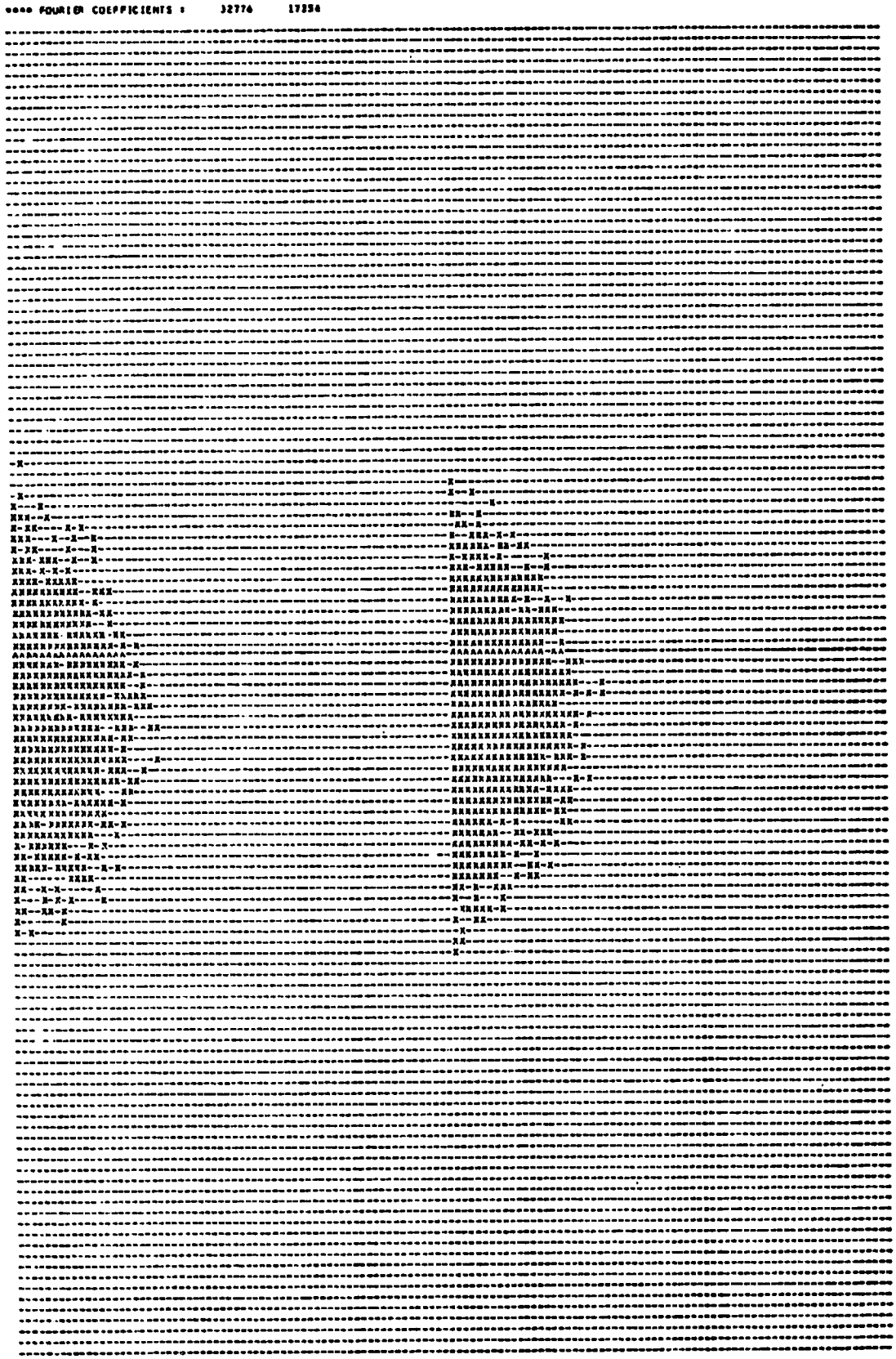
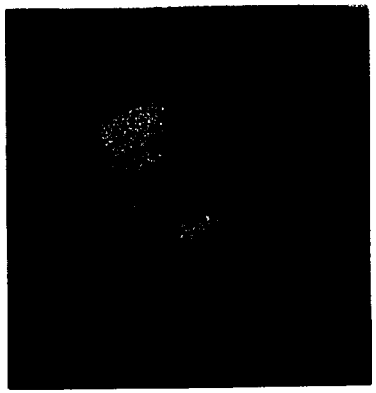


Figure E-11 Continued

APPENDIX F

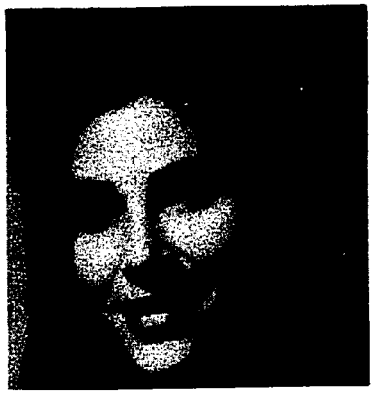
Photographs of pictures
that have been prefiltered
using the one-dimensional
and two-dimensional simulation
programs and various values of
simulated viewing distance.



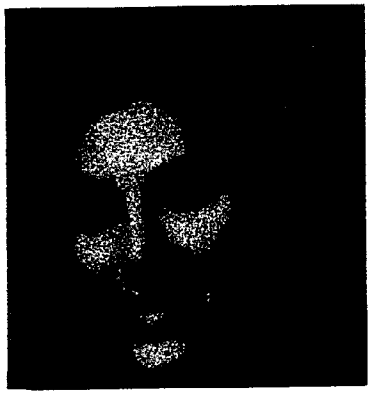
ORIGINAL
"FACE"



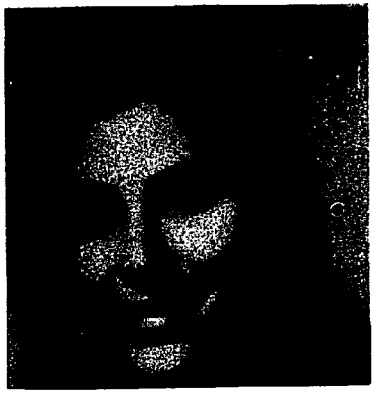
1 Dim. 8:1
1.403



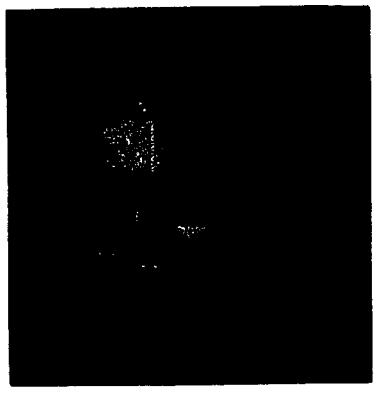
1 Dim. 16:1
2.135



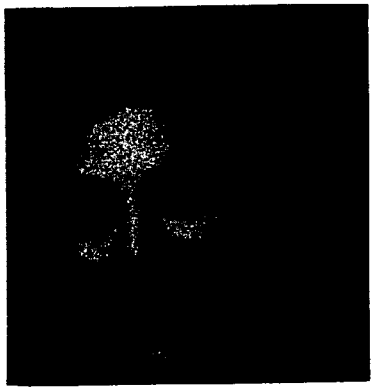
1 Dim. 32:1
4.279



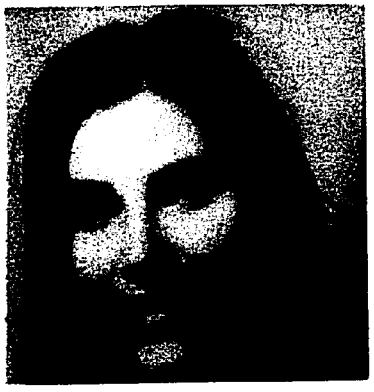
1 Dim. 64:1
8.361



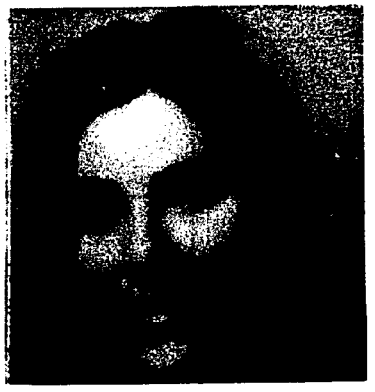
2 Dim. 4:1
2.50



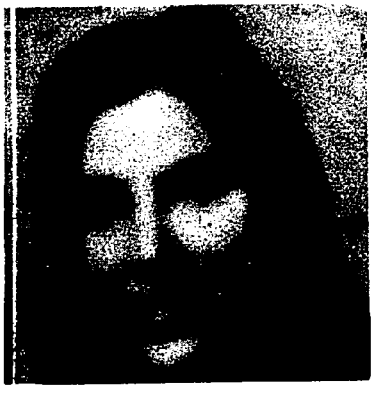
2 Dim. 8:1
4.541



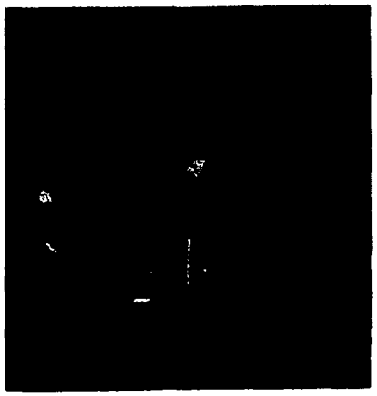
2 Dim. 16:1
15.265



2 Dim. 32:1
52.829



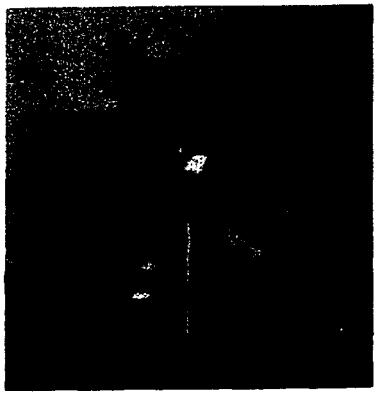
2 Dim. 64:1
150.139



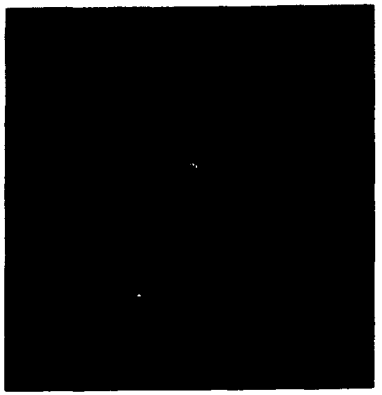
ORIGINAL
"CAMERAMAN"



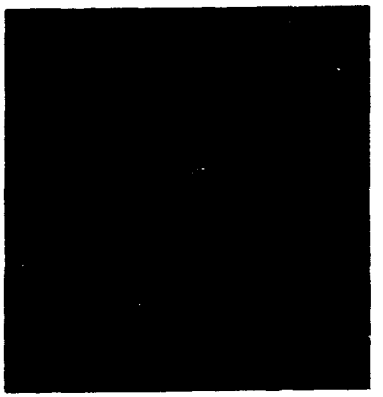
1 Dim. 8:1
1.623



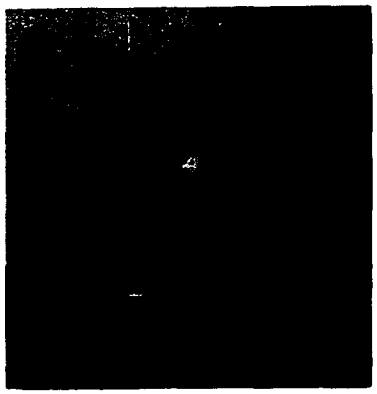
1 Dim. 16:1
2.246



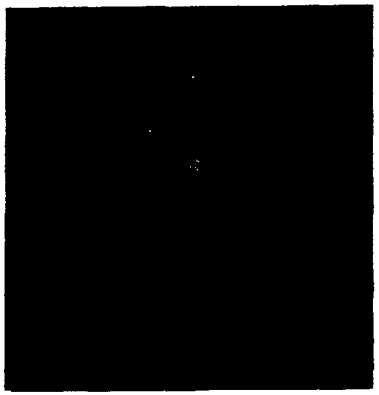
1 Dim. 32:1
3.964



1 Dim. 64:1
7.909



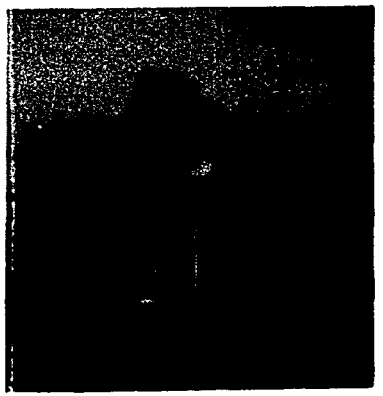
2 Dim. 4:1
2.659



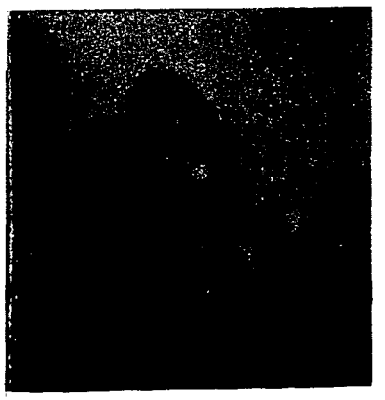
2 Dim. 8:1
4.672



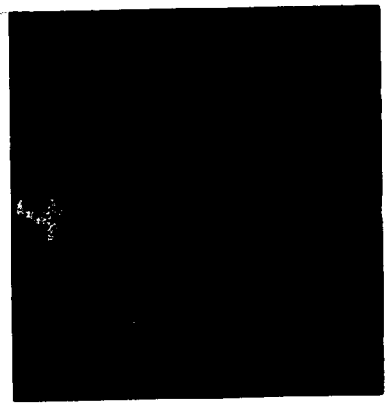
2 Dim. 16:1
14.462



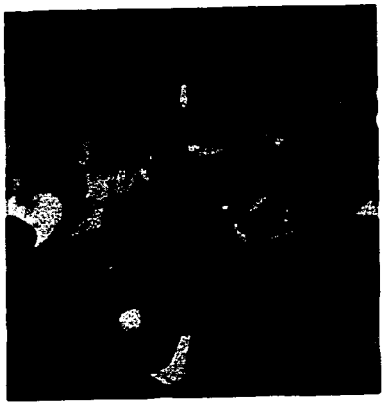
2 Dim. 32:1
43.044



2 Dim. 64:1
157.158



ORIGINAL
"CROWD"



1 Dim. 8:1
1.295



1 Dim. 16:1
1.727



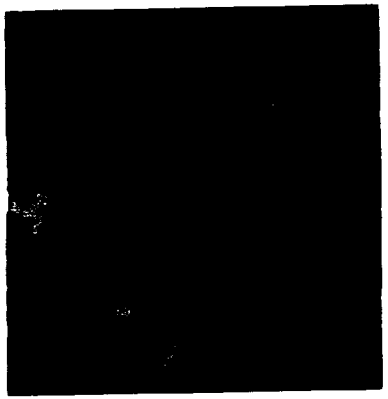
1 Dim. 32:1
3.144



1 Dim. 64:1
5.718



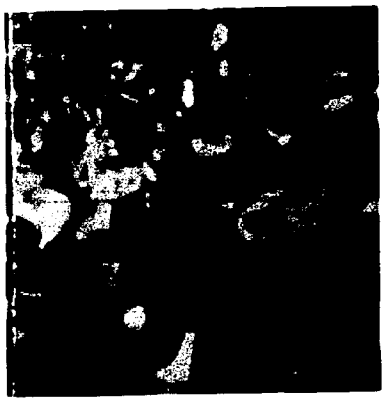
2 Dim. 4:1
1.412



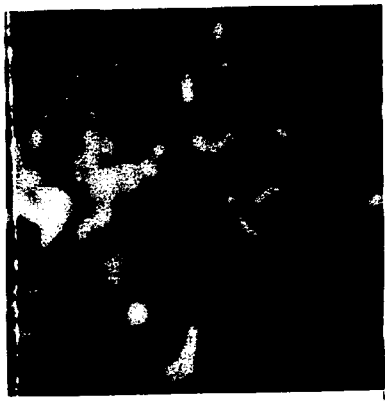
2 Dim. 8:1
2.225



2 Dim. 16:1
7.011



2 Dim. 32:1
23.628



2 Dim. 64:1
82.124

BIBLIOGRAPHY

Brainard, R.C., "Subjective Evaluation of PCM Noise-Feedback Coder for Television," Proc. IEEE, vol. 55, no. 3, pp. 346-353.

Bruce, R.A., "Optimum Pre-Emphasis and De-Emphasis Networks for Transmission of Television by PCM," IEEE Trans. on Communications Technology, vol. COM-12, pp. 91-96.

Capon, J., "Bounds to the Entropy of Television Signals," Technical Report 296, MIT Research Lab. of Electronics, Cambridge, Mass. 1955.

Carbrey, R.L., "Video Transmission over Telephone Cable Pairs by Pulse Code Modulation," Proc. IRE, vol. 48, September, 1960, vol. 48, pp. 1546-1561.

Cheng, G., R. Ledley, D. Pollock, A. Rosenfeld, eds., Pictorial Pattern Recognition, Thompson Books.

Cherry, E.C., G.G. Gouriet, "Some Possibilities for the Compression of Television Signals by Recoding," Proc. IEEE, Paper No. 1401R, January 1953, (100, part III, p.9).

Cherry, C., M.H. Kubba, D.E. Pearson, and M.P. Barton, "An Experimental Study of the Possible Bandwidth Compression of Visual Image Signals," Proc. IEEE, vol. 51, Nov, 1963, pp. 1507-1517.

Cherry, E., A.H. Robinson, "Results of a Prototype Television Bandwidth Compression Scheme," Proc. IEEE, vol. 55, no. 3, March 1967, pp. 356-363.

Davenport, W.B., jr., W.L. Root, An Introduction to the Theory of Random Signals and Noise, McGraw-Hill, 1958.

Deutsch, S., "A Note on Some Statistics Concerning Typewritten or Printed Material," IRE Trans. on Information Theory, vol. IT-3, June, 1957, pp. 147-148.

Elias, R., "Predictive Coding," IRE Trans. on Information Theory, vol. IT-1, March, 1955, pp. 16-33.

Foy, W.H., jr. "Entropy of Simple Line Drawings," IEEE Trans. on Information Theory, vol. IT-10, April, 1964, pp. 165-167.

Freeman, "On the Encoding of Arbitrary Geometric Configurations," IRE Trans. on Elec. Computers, vol. EC-10, no. 2, June 1961, pp. 260-268.

Golomb, S.W., "Run-Length Encodings," IEEE Trans. on Information Theory, vol. IT-13, no. 4, July, 1966, pp. 399-401.

Goodall, W.M., "Television by Pulse Code Modulation," Bell System Tech. J., vol. 30, January, 1951, pp. 33-49.

Graham, R.E., "Communications Theory applied to Television Coding," Proc. International Symposium on Physical Problems of Color TV, ACTA Electronica, vol. 2, 1957-58.

_____"Predictive Quantizing of Television Signals," IRE WESCON Conv. Records, (Electronic Computers), August 1958, p. 147.

_____"Snow Removal--A Noise-Stripping Process for Picture Signals," IRE Trans. on Information Theory, vol. IT8, February, 1962, pp. 129-144.

_____"Subjective Experiments in Visual Communications," IRE National Conv. Record, vol. 6, pt. 4, 1958, pp. 100-106.

Grossman, M., "On the Nature of Information--An Application of Entropy," IEEE Spectrum, vol. 2, no. 12, Dec 1965, pp. 70-80.

C.W. Harrison, "Experiments with Linear Prediction in Television," Bell Sys. Tech. J., 1952, 31, p. 764.

Horowitz, I., Synthesis of Feedback Systems, Academic Press, New York, p. 596.

Huang, T.S., "PCM Picture Transmission," IEEE Spectrum, vol. 2, no. 12, Dec 1965, pp. 57-63.

Huffman, D.A., "A Method for the Construction of Minimum-Redundancy Codes," Proc. of IRE, vol. 40, no. 9, Sept 1952, pp. 1098-1101.

Kirsch, R.A., "Computer Interpretation of English Text and Picture Patterns," IEEE Transactions on Electronic Computers, August 1964, vol. EC-13, no. 4, pp. 363-376.

Kirsch, R.A., L. Cahn, C. Ray, G.H. Urban, "Experiments in Processing Pictorial Information with a Digital Computer," Proc. EJCC, 1957, pp. 221-229.

Kovasznay, L.S.G. Kovasznay, H.M. Joseph, "Image Processing," Proc. of the IRE, vol. 43, no. 5, May 1955, pp. 560-570.

Kretzmer, E.R., "Reduced Alphabet Representation of Television Signals," IRE Conv. Record, Part IV, 1956, p. 140.

_____, "Statistics of television Signals," Bell Sys. Tech. J., vol. 31, July, 1952, pp. 751-763.

Kubba, M.H., "Automatic Picture Detail Detection in the Presence of Random Noise," Proc. IEEE, vol. 51, Nov 1963, pp. 1517-1523.

Lacmmel, A.E., "Coding Processes for Bandwidth Reduction in Picture Transmission," Microwave Research Inst., Polytech. Inst. of Brooklyn, N.Y., Report R-246-51, PIB-187, August, 1951.

Loeb, J., "Communications Theory of the Transmission of Simple Line Drawings," Communications Theory, Willis, Jackson, ed., Butterworth Scientific Publications, London, England,1953, pp. 323-325.

Macleod, I.D.G., "Pictorial Output with a Line Printer," IEEE Trans. on E.C., vol. C-19, no. 2, Feb 1970, pp. 160-162.

Mertz, P., and F. Gray, "A Theory of Scanning," Bell Sys. Tech. J., vol. 13, pp. 464-515.

Michel, W.S., "Statistical Encoding for Text and Picture Communications," AIEE Conference Paper CP 57-723, May 1, 1957.

Michel et al, "A Coded Facsimile System," IRE Wescon Conv. Record, Pt. 2, 1957, p. 84.

Montgomery, W.D., "Reconstruction of Pictures from Scanned Records," IEEE Trans. on Information Theory, vol. IT-11, no. 2, April 1965, pp. 204-206.

Nagy, G., "Preliminary Investigation of Techniques for Automated Reading of Unformatted Text," Comm. ACM, vol. 11, no. 7, July 1968, pp. 480-487.

Nishikawa, S., et al, "Area Properties of Television Pictures," IEEE Trans. on Information Theory, vol. IT11, no.3, pp. 348-352.

Nyquist, H., "Certain Topics in Telegraph Transmission Theory," Trans. AIEE, vol. 47, April, 1928, pp. 617-644.

Oliver, B.M., "Efficient Coding," Bell Sys. Tech. J., July 1952, vol. 31, no. 4, pp. 724-750.

Oliver, B.M., J.R. Pierce, and C.E. Shannon, "The Philosophy of PCM," Proc. IRE, vol. 36, Nov 1948, pp. 1324-1331.

O'Neal, J.B., jr., "A Bound on Signal-to-Quantizing Noise Ratios for Digital Encoding Systems," Proc. IEEE, vol. 55, no.3, March 1967, pp. 287-292.

_____"Deltamodulation Quantizing Noise Analytical and Computer Simulation Results for Gaussian and TV Input Signals," Bell Sys. Tech. J., vol. 45, Jan 1966, pp. 117-142.

_____"Predictive Quantizing Systems for the Transmission of TV Signals," Bell Sys. Tech. J., vol. 45, pp. 689-722.

Oppenheim, A.V., R.W. Schafer, T.G. Stockham, jr., "Nonlinear Filtering of Multiplied and Convolved Signals," Proc. IEEE, vol. 56, no. 8, August 1968, pp. 1264-1290.

Pearson, D.E., "A Realistic Model for Visual Communications Systems," Proc. IEEE, Vol. 55, no. 3, pp. 380-389.

Powers, K.H., "A Prediction Theory Approach to Information Rates," IRE Convention Record, pt. 4, 1956, pp. 132-139.

Pratt, F., Secret and Urgent, Blue Ribbon Books, New York, New York, 1939.

Pratt, W.K., "A Bibliography on Television Bandwidth Reduction Studies," IEEE Trans. on Information Theory, vol. IT-13, no.1, Jan 1967, pp. 114-115.

Russell, J.K., "A Visual Image Processor," IEEE Transactions on Computers, vol. 317, no. 7, July 1968, pp. 635-639.

Schreiber, W.F., "The Measurement of Third-Order Probability Distribution of Television Signals," Trans. of IRE, 1956, p. 94-105.

_____, "The Mathematical Foundation of the Synthetic Highs System," Research Lab. of Electronics, MIT, Cambridge, Mass., Quarterly Report 68, p. 140.

_____, "Probability Distributions of Television Signals," Ph.D. Thesis, Harvard University, 1953, IRE Conv. Record, Part 4, 1953, p. 35.

Schroeder, M.R., "Images from Computers and Microfilm Plotters," Comm. of the ACM, vol. 12, no. 2, Feb 1969, pp. 99-101.

Schuckmann, L, "Dither Signals," IEEE Trans. on Comm. Tech., December 1964, pp. 162-165.

Schwartz, J.W., R.C. Barker, "Bit Plane Encoding: A Technique for Source Encoding," IEEE Trans. AES-2, July 1966, pp. 385-392.

Sekey, A., "Communications at Maximum Rate Without Error Correction," IEEE Tran. on Information Theory, vol. IT-10, Oct 1964, p. 381.

_____, "Detail Detection in Television Signals," Proc. IEEE, Vol. 53, January 1965, pp. 75-76.

Seyler, A.J., "The Coding of Visual Signals to Reduce Channel Capacity Requirements," Proc. IEEE (London), vol. 109, pt. C, July 1962, pp. 676-684.

Shannon, C.E., "Coding Theorems for a Discrete Source with a Fidelity Criterion," in Information and Decision Processes, R.E. Machol, ed., McGraw Hill, New York, 1960, pp. 93-126.

_____, "Prediction and Entropy of Printed English," Bell System Technical Journal, New York, N.Y., Vol. 30, Jan 1951, pp. 50-64.

Shaw, Alan C., "Parsing of Graph-Representable Pictures," JACM, Vol. 17, no. 3, pp. 453-481.

Simpson, R.S., et al, "Compendium of Redundancy Removal Processes," IEEE Trans. AES-2, July 1966, pp. 471-474.

Stara, H., and K.H. Powers, "Some Relations Between Television Picture Redundance and Bandwidth Requiremnets," Trans. AIEE, vol. 1957, vol. 76, pt. 1, p. 492-496.

Stoddard, J.C., "Measurement of Second-Order Probability Distributions of Pictures by Digital Means," MIT Research Lab. for Electronics, Tech. Report, No. 302, 1955.

Stucki, P., "Generation of Grey Tones by Computer for Simulation of Visual Information Systems," IEEE Trans. on Computers, vol. C-18, no. 7, July, 1969, pp. 642-643.

Swerling, P., "Statistical Properties of the Contours of Random Surfaces," IRE Trans. on Information Theory, vol. IT-8, July 1962, pp. 315-321.

Tallman, O.H., "A Model for the Classification of Visual Images," Proc. of the 21st. National Aerospace Electronics Convention, (NAECON), Dayton, Ohio, May 1969, pp. 3-1-370.

Unger, S.H., "Pattern Detection and Recognition," Proc. IRE, Vol. 47, no. 10, October 1959, pp. 1737-1752.

Wyle, H., T. Erb, R. Banow, "Reduced-Time Facsimile Transmission by Digital Coding," IRE Trans. on Communications Systems, vol. CS-9, pp. 215-222.

Youngblood, W.A., "Picture Processing," Quarterly Progress Report 48, MIT Research Lab. for Electronics, 1958, pp. 95-100.