SEMIANNUAL TECHNICAL REPORT

Covering Research Activity During the Period

1 March 1974 through 31 August 1974

William K. Pratt

Project Director

(213) 746-2694

Image Processing Institute

University of Southern California

University Park

Los Angeles, California 90007

30 September 1974

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Air Force Eastern Test Range under Contract No. F08606-72-C-0008, ARPA Order No. 1706

The views and conclusions in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency or the U. S. Government.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Image Processing Institute University of Southern California, University Park, Los Angeles, California 90007	UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

IMAGE PROCESSING RESEARCH

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Semiannual, 1 March 1974 to 31 August 1974

5. AUTHOR(S) (First name, middle initial, last name)

William K. Pratt (Project Director)

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
30 September 1974	126	40

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
F08606-72-C-0008 b. PROJECT NO. ARPA Order No. 1706 c. d.	USCIPI Report 540 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

Approved for release; distribution unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Advanced Research Projects Agency 1400 Wilson Boulevard Arlington, Virginia 22209

13. ABSTRACT

This technical report summarizes the image processing research activities performed by the University of Southern California during the period of 1 March 1974 to 30 September 1974 under Contract No. F08606-72-C-0008 with the Advanced Research Projects Agency, Information Processing Techniques Office.

The research program, entitled, "Image Processing Research," has as its primary purpose the analysis and development of techniques and systems for efficiently generating, processing, transmitting, and displaying visual images and two dimensional data arrays. Research is oriented toward digital processing and transmission systems. Five task areas are reported on: (1) Image Coding Projects: the investigation of digital bandwidth reduction coding methods; (2) Image Restoration and Enhancement: the improvement of image fidelity and presentation format; (3) Image Data Extraction Projects: the recognition of objects within pictures and quantitative measurement of image features; (4) Image Analysis Projects: the development of quantitative measures of image quality and analytic representation; (5) Image Processing Support Projects: development of image processing hardware and software support systems.

* *

14. Key words: Image Processing, Digital Image Processing, Image Coding, Image Enhancement, Image Restoration, Image Processing Software, Image Processing Hardware, Color Image Processing.

DD FORM 1 NOV 65 1473

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT

ABSTRACT

This technical report summarizes the image processing research activities performed by the University of Southern California during the period of 1 March 1974 to 31 August 1974 under Contract No. F08606-72-C-0008 with the Advanced Research Projects Agency, Information Processing Techniques Office.

The research program, entitled, "Image Processing Research," has as its primary purpose the analysis and development of techniques and systems for efficiently generating, processing, transmitting, and displaying visual images and two dimensional data arrays. Research is oriented toward digital processing and transmission systems. Five task areas are reported on: (1) Image Coding Projects, the investigation of digital bandwidth reduction coding methods; (2) Image Restoration and Enhancement Projects: the improvement of image fidelity and presentation format; (3) Image Data Extraction Projects: the recognition of objects within pictures and quantitative measurement of image features; (4) Image Analysis Projects, the development of quantitative measures of image quality and analytic representation; (5) Image Processing Support Projects, development of image processing hardware and software support systems.

PROJECT PARTICIPANTS

Project Director

William K. Pratt

Research Staff

Harry C. Andrews
Lee D. Davisson
Werner Frei
Ali Habibi
Ernest L. Hall
Ronald S. Hershel
Anil K. Jain
Richard P. Kruger
Nasser E. Nahi
Guner Robinson
Alexander A. Sawchuk
Lloyd R. Welch

Support Staff

Angus B. Cossey
Toyone Mayeda
Paula Meeve
James M. Pepin
Easter D. Russell
Mark A. Sanders
Ray Schmidt
Joyce Seguy
Dennis Smith
George Soen
Wai Szeto
Florence B. Tebbets

Students

Ben Britt
Steve Dashiell
Faramarz Davarian
Roy M. Glantz
Henry Glazner
Michael Huhns
Steve Hui
Mohammad Janshahi
Mohsen Khalil
San Kwok
Alan Larson
Paul Liles
Robert Liles
Eduardo Lopez
Simon Lopez-Mora
Clanton Mancill
Nelson Mascarenhas
David Miller
Firouz Naderi
Michael Patton
Mohammad Peyrovian
Stewart Robinson
Robert Wallis
Pamela Welch

TABLE OF CONTENTS

1. Research Project Overview

This report describes the progress and results of the University of Southern California image processing research study for the period of 1 March 1974 to 31 August 1974. The image processing research study has been subdivided into five projects:

Image Coding Projects
Image Restoration and Enhancement Projects
Image Data Extraction Projects
Image Analysis Projects
Image Processing Support Projects

In image coding the orientation of the research is toward the development of digital image coding systems that represent monochrome and color images with a minimal number of code bits. Image restoration is the task of improving the fidelity of an image in the sense of compensating for image degradations. In image enhancement, picture manipulation processes are performed to provide a more subjectively pleasing image or to convert the image to a form more amenable to human or machine analysis. The objectives of the image data extraction projects are the registration of images, detection of objects within pictures and measurements of image features. The image analysis projects comprise the background research effort into the basic structure of images in order to develop meaningful quantitative characterizations of an image. Finally, the image support projects include research on image processing computer languages and the development of experimental equipment for the sensing, processing, and display of images.

The next section of this report summarizes some of the research project activities during the past six months. Sections 3 to 7 describe the research effort on the projects listed above during the reporting period. A capsule description of the physical facilities of the USC Image Processing is contained in Section 8. Section 9 is a list of publications by project members.

2. Research Project Activities

Significant research project activities of the past six months are summarized below:

Summer Short Courses. One of the vehicles for the transfer of ARPA sponsored image processing research technology to the Federal and industrial communities has been through intensive one and two week short courses. For the past four years the University has offered a one week Summer short course in Mathematical Pattern Recognition and a two week course in Digital Image Processing. In addition, this Summer, a third short course of one week duration was initiated on Optical Processing. The alumni of these courses now numbers in the hundreds.

Picture Coding Symposium. Professor's William K. Pratt, Ali Habibi, and Werner Frei attended the Picture Coding Symposium held in Goslar, West Germany on 26-28 August, 1974. Approximately 150 persons from throughout the free world attended the meeting which was devoted to sessions on: human observer characteristics, intraframe image coding, interframe image coding, color image coding, and multi-spectral data coding. At the conference a contest was held to choose the best algorithms for picture coding developed during the past 18 months. Dr. Ali Habibi of USC was awarded a prize for the best coding algorithm for monochrome image coding with 1.0 bits/pixel.

Special Issues. During the past six months two special issues relating to image processing were published. Professor Harry C. Andrews of USC was editor of the IEEE Computer journal May 1974 special issue on "Computer Image Processing." The Society of Photo-Optical Instrumentation Engineers devoted its May 1974 issue of Optical Engineering to optical and digital image processing under the editorship of Professor Alexander A. Sawchuk of USC.

3. Image Coding Projects

The effort in image coding is directed toward the research and development of image coding systems providing a transmission bit rate reduction and tolerance to channel errors. Coding systems are under investigation for: monochrome and color imagery; slow scan and real time television; and information preserving and controlled fidelity operation. Results of this research study during the past six months are summarized here and presented in detail in subsequent sections.

The first report concerns a performance study of several adaptive linear predictive image coding systems. One such system which adaptively codes by DPCM or deltamodulation dependent upon picture quality has been found to provide high fidelity coding down to about 2.0 bits/pixel.

The next report describes an application of Hadamard and Slant transforms to interframe image coding. Good results are obtained at an average of much less than one bit/pixel.

In the final report simulation results are given for transform coding spectrum interpolation in which receiver post-processing is employed to reduce quantization error effects on transform coefficients. Reduction in mean square error of about 30% can be obtained by this technique.

3.1 Adaptive Dual Mode DPCM/Deltamodulation Image Coding Techniques

William K. Pratt

Standard one bit per sample deltamodulation provides a considerable bandwidth reduction, but suffers from the complementary problems of image granularity in smooth areas of an image and slope overload in image regions of rapid brightness change. These problems may be alleviated by quantizing the difference signal with more levels as in 3 bit per sample DPCM, but of course, the bandwidth reduction is sacrificed. A compromise technique has been investigated in which the image coder

operates in a dual mode: DPCM in regions of high image activity and deltamodulation in regions of low activity. The switch between modes is performed adaptively based upon the relative image activity.

In a basic deltamodulation image coding system a prediction of the next pixel to be scanned is made based upon the previously scanned pixel. The difference between the actual pixel value and its estimate is quantized to two levels and transmitted as a binary pulse (one bit pixel code). Figures 1a to 1c contain computer simulation photographs of deltamodulation coded pictures for three quantization levels (q =5%, 10%, 20%) and a previous pixel weighting of 90%. The tradeoff between slope overload error and granularity error is readily apparent from the photographs.

With DPCM encoding the difference signal between a scanned pixel and its previous pixel estimate is quantized to 8 levels (3 bit/pixel code) which are nonlinearly spaced to minimize quantization error. Figure 1d contains an example of DPCM coding when the first threshold level is set at 2.5% full scale and a 90% previous pixel weighting is employed. The resultant image quality is satisfactory for most applications.

One possibility for reducing the bit rate requirement of DPCM is to employ a dual mode coder which switches from DPCM to deltamodulation in regions of nearly constant grey level and from deltamodulation to DPCM in edge regions. A simple algorithm for the switch is as follows:

(a.) Switch from deltamodulation to DPCM if three sequential delta bits are of the same sign

(b.) Switch from DPCM to deltamodulation if DPCM quantizer shifts from smallest positive to smallest negative quantization level or vice versa.

By adjusting the quantization levels for the DPCM and deltamodulation quantizers it is possible to control the relative time division between states. One of the major advantages of the encoder is that the decoding may be performed at the receiver from the transmitted code without any explicit code bits required to designate the mode. Figure 1e contains a coded image using this algorithm coded at about 2.0 bits/pixel. In regions of little image activity, image quality is good, but mode transition

(a) deltamodulation
q = 5%, 1 bit/pixel

(b) deltamodulation
q =10%, 1 bit/pixel

(c) deltamodulation
q = 20%, 1 bit/pixel

(d) DPCM
Q_1=2.5%, 3 bits/pixel

(e) dual mode
2 bits/pixel

(f) oversampled dual/mode
2 bits/pixel

Figure 3.1-1. Examples of deltamodulation, DPCM and dual mode coded images

errors are apparent near edges.

Frei, Schindler, and Velliger [1] have suggested a dual mode system in which the original image is oversampled by a factor of three to permit a rapid detection of the delta to DPCM mode change. The coding logic is as follows:

(a.) Switch from delatamodulation to DPCM:

delta mode quantization levels

sample 1	sample 2	sample 3	Code
+q	+q	+q	1 1 1 1 0
+q	+q	-q	1
+q	-q	+q	1
-q	+q	+q	1
-q	-q	+q	0
-q	+q	-q	0
+q	-q	-q	0
-q	-q	-q	0 0 0 0 1

marker bits

After three sequential delta "ones" insert a "zero" marker bit and after three sequential delta "zeros" insert a "one" marker bit.

(b.) Switch from DPCM to deltamodulation if DPCM quantizer shifts from smallest positive to smallest negative quantization level or vice versa.

Figure 1f contains a coded image for the oversampled dual mode coding system for coding at two bits per pixel. The oversampled dual mode system is superior to the simple dual mode system described previously in terms of picture quality, however, its implementation requirements are greater.

In summary, the oversampled dual mode DPCM/deltamodulation image coding system has proven to provide good quality coded images at 2.0 bits/pixel. The disadvantage of the coder is the additional complexity as compared to conventional 3 bit/sample DPCM.

References

1. A. H. Frei, H. R. Schindler, and P. Vettiger, "An Adaptive Dual-Mode Coder/Decoder for Television Signals," IEEE Trans. on Communication Technology, Vol. COM-19, No. 6, Part I, December, 1971, pp. 933-944.

3.2 Orthogonal Transform Coding of Moving Pictures

Clifford Reader

The efficiency of coding a sequence of moving pictures is improved by techniques of interframe coding which partially remove redundancy between frames. A successful type of coding is that of conditional update operating upon the differences between successive frames. Orthogonal transform techniques have been applied to the conditional update process with additional advantages over conventional systems.

Although interframe coding may be very successful in reducing the transmission requirements for moving picture signals, the coder is usually complex and costly, involving an expensive frame memory. The object of the orthogonal transform conditional update coder is to minimize terminal costs. The coder performs a two stage process. The first stage is a conventional intraframe coder, removing spatial redundancy from single frames. This orthogonal transformation and block quantization is performed over sub-blocks within the frame. The second, interframe coding stage, then applies conditional update techniques to the differences between successive transformed and quantized sub-blocks. This is simply achieved by setting a threshold and updating a sub-block if the energy of the difference signal for that sub-block exceeds the threshold. A block diagram of the coder is shown in figure 1. It is to be noted that the updating process takes place with the data from the new frame and not the difference information. This is done to avoid problems with quantizing the difference signal which is derived from the nonlinear intraframe quantizer. An advantage is that the memory is refreshed with new data and not an estimate of the new data. This reduces the problem of multiplicative error. The principle advantage of the

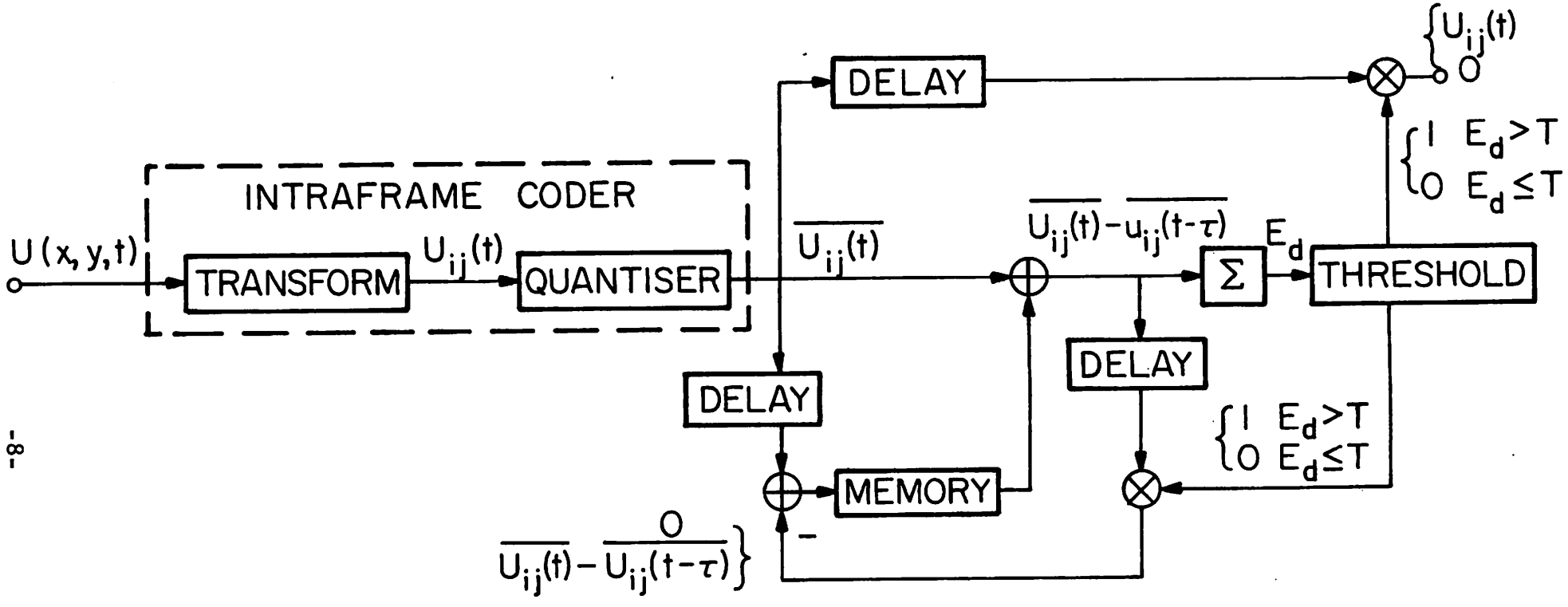


Figure 3.2-1. Block Diagram of Adaptive Interframe Coder.

interframe coder is derived from the use of intraframe coded data. This data is thus presented at reduced rate and the memory requirement of the coder is reduced. If the interframe coder produces a compression of M:1 then the size of the delays must be 1/256M frames (plus a small amount to allow for the difference signal energy thresholding to take place) and the memory size will be 1/M - 1/256M = 255/256M Frames. The overall memory (storage) requirement for the coder is thus reduced by almost a factor of M compared with the requirement for conventional coders.

Quantitative analysis of the interframe error is hampered by the nonlinearity of the updating process. Qualitatively, several observations may be made. In the absence of noise, a threshold of zero would result in the transmission of all sub-blocks posessing a finite difference signal energy and the interframe error would be zero. If the threshold were then raised, those sub-blocks which contained the least amount of motion would not be updated. Since the difference signal energy is closely related to the mean squared error made when a sub-block is not updated, it is reasonable to assume that the interframe error should be proportional to some function of the update threshold. In the presence of noise, this effect will be modified. The interframe error was examined for five pairs of frames containing different amounts of motion. Figure 2 shows mean squared error vs update threshold. The curves of figure 3 show the corresponding compression of data achieved with thresholding. The interframe error curve for medium motion is presented in full detail in figure 4. Three regions may be discerned along the curve. For the threshold set less than about 1.0, all sub-blocks which contain motion will be updated together with some of the sub-blocks which do not contain motion, depending upon the magnitude of their noise variance. For a threshold set over the range 1.0 to 4.0 approximately, all sub-blocks which contain motion will be updated except those sub-blocks which contain very little motion (for example, just a few pixels in one corner have changed) plus those sub-blocks which contain no motion, but have the largest noise variance. For a threshold greater

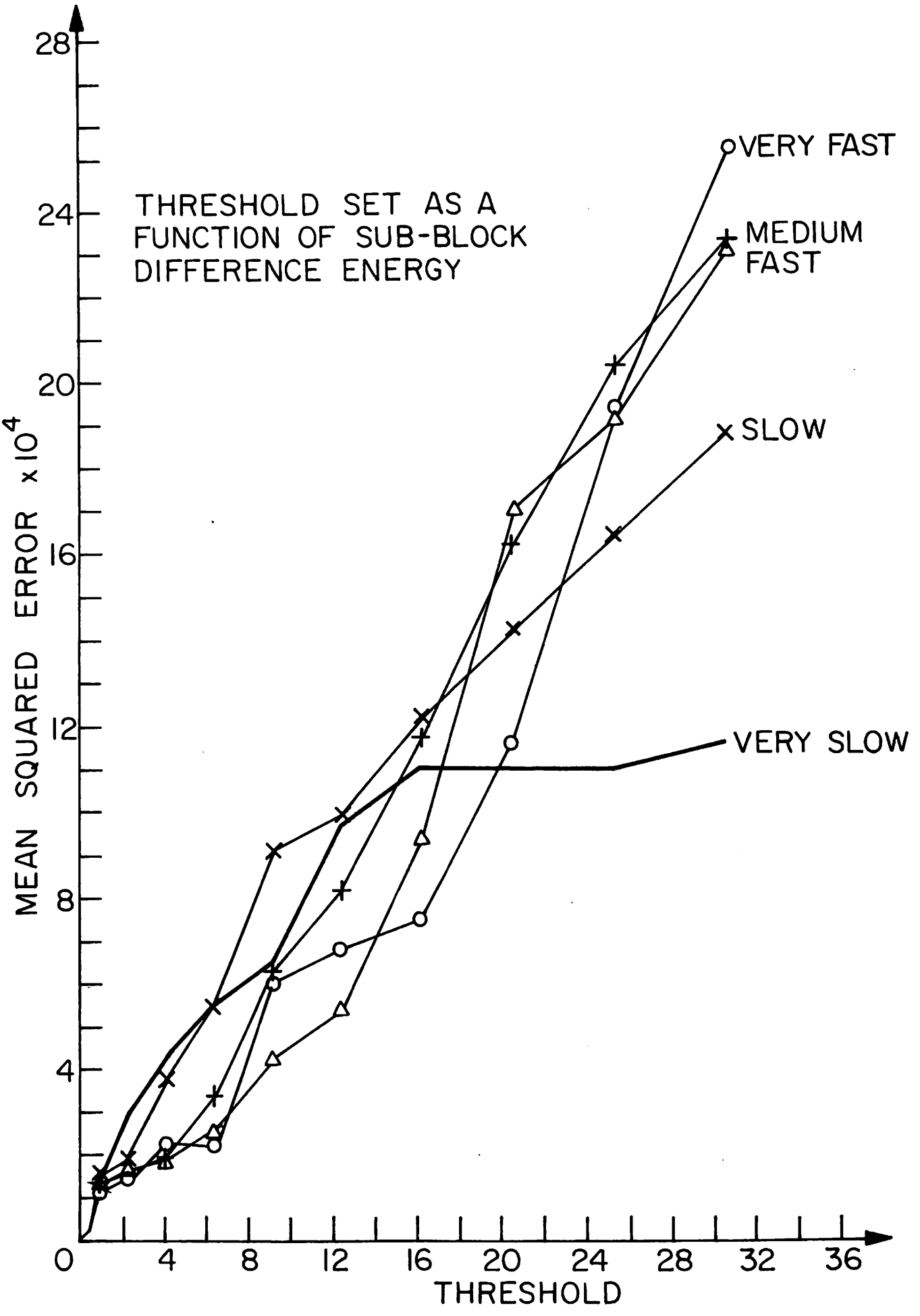


Figure 3.2-2. Graph of Mean Squared Error vs. Threshold for Five Degrees of Motion.

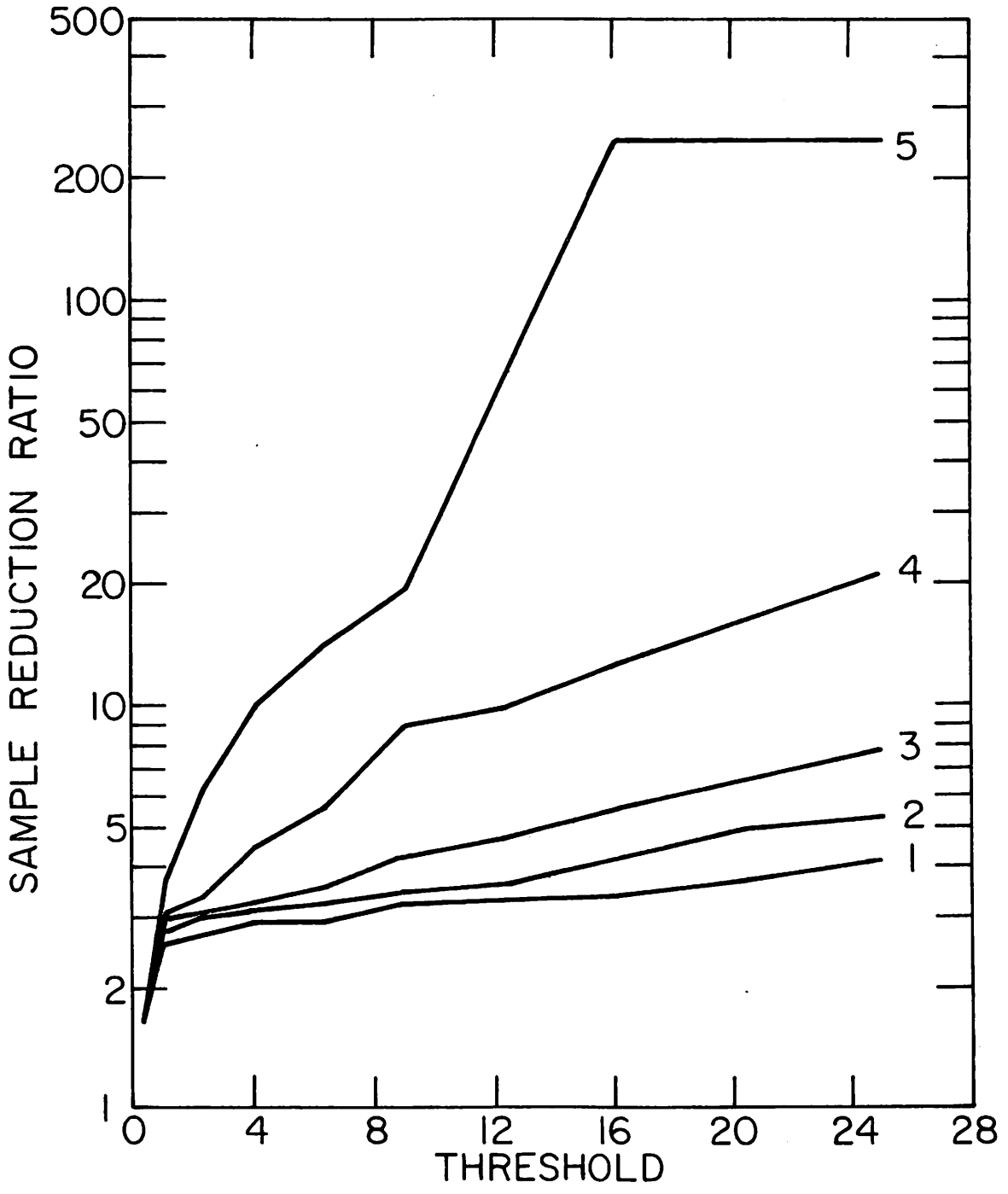


Figure 3.2-3. Graph of Interframe Sample Reduction vs. Update Threshold.

1 = very fast motion, 2 = fast motion,

3 = medium motion, 4 = slow motion,

5 = very slow motion.

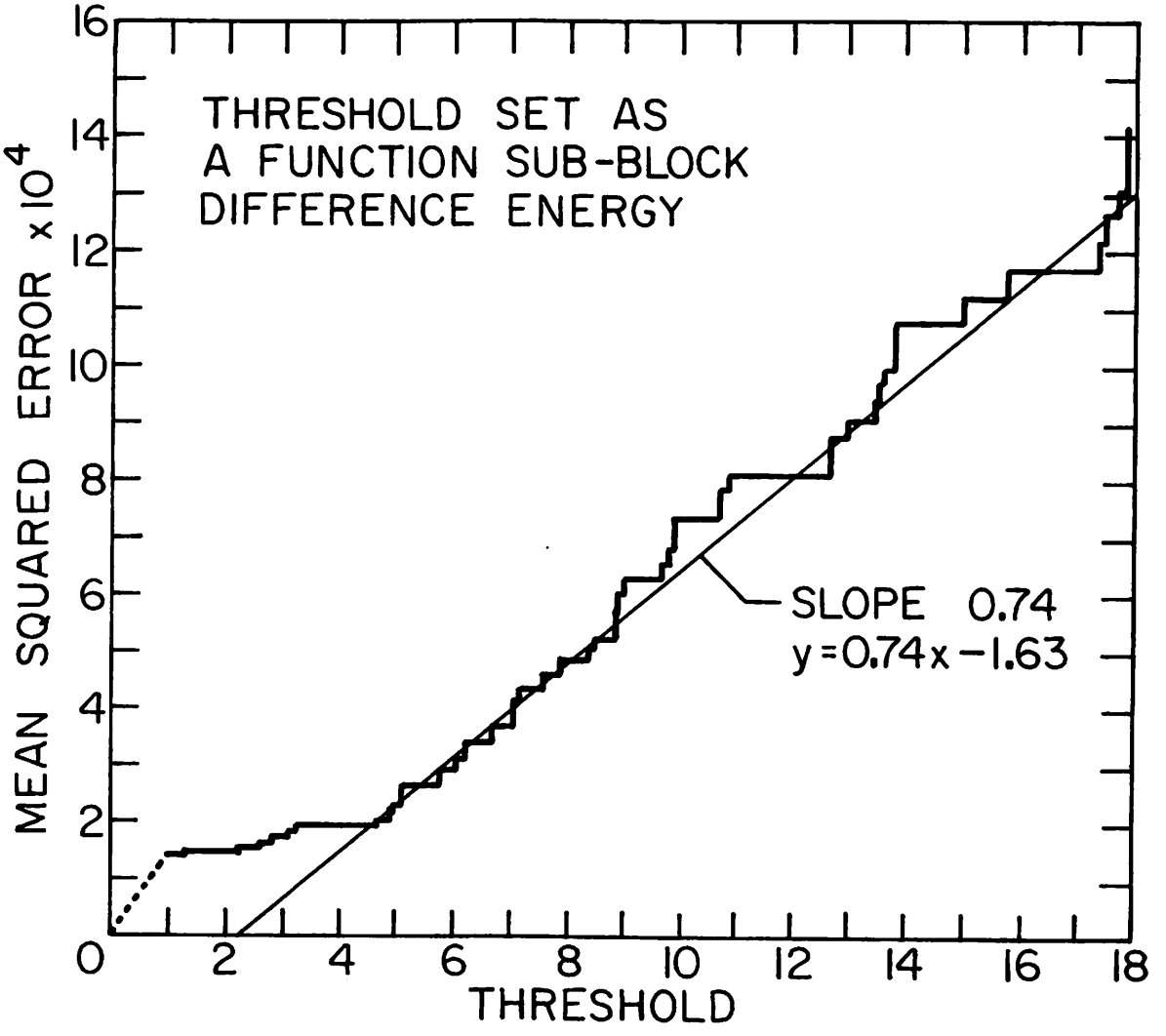


Figure 3.2-4. Graph of Mean Squared Error vs. Threshold, Medium Motion.

than about 4.0, only those sub-blocks containing the greatest amount of motion will be updated. This region of the wave is approximated by the straight line having the equation

$$\epsilon^2 = 0.74\ T - 1.63$$

where ϵ^2 is the mean squared error and T is the update threshold. This line does not fit the curves for other degrees of motion very well but must be taken as an overall measure of the error of the system assuming an average medium motion.

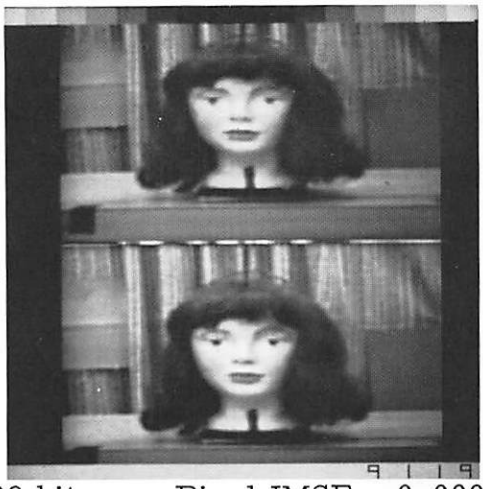
The curves for interframe mean square error and bit rate compression indicate the behavior of the system towards the different degrees of motion. The error curves for medium, fast and very fast motion are reasonably close to each other throughout the range of thresholds presented while the corresponding bit rate compression curves are nearly linear and vary from each other only in slope. In contrast, the curves for slow and very slow motion differ in being nonlinear and not aligned to the curves for faster motion. This effect is caused by the discreteness of the sub-block conditional updating mechanism, which becomes prominent when only a few sub-blocks contain motion and many of those sub-blocks overlap stationary areas. The results for very slow motion are further modified by noise which interacts with the low difference signal energy, preventing correct updating.

The data available for experiments consisted of five pairs of frames representing a range of levels of activity and a set of four frames containing very active motion. These frames were individually coded by orthogonal transformation and block quantization. The Hadamard transform was used for the five pairs of frames with quantization to 3.0 bits per pixel. The Slant transform was used for the sequence of four frames with rates of 3.0 and 1.5 bits per pixel. The result of interframe coding the five pairs of intraframe coded frames, with a threshold of 4.0 is shown in figure 5. No discrete errors are visible as a result of the interframe coding although the images are a little blurred by the intraframe coding. Raising the threshold to 9.0, figure 6 does introduce

Very fast motion

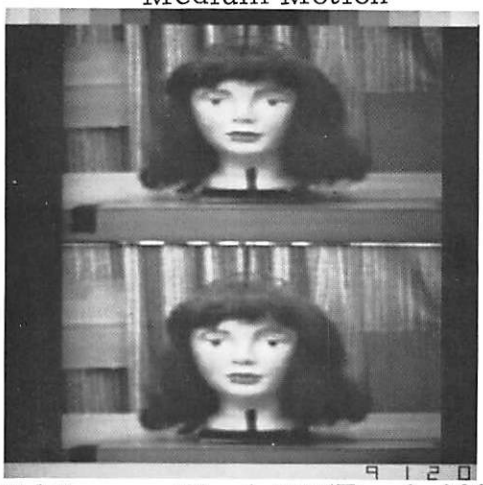
1.0 bits per Pixel IMSE = 0.000022

Fast motion



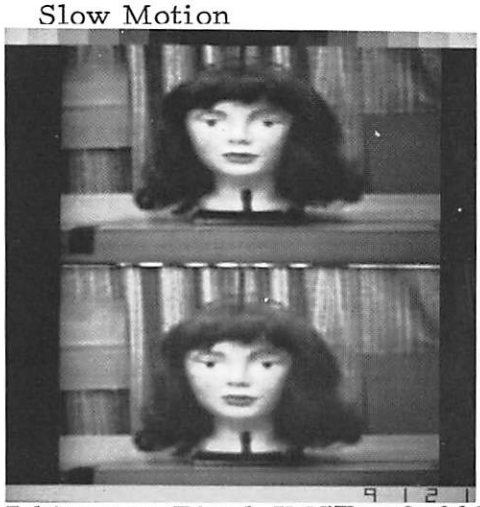
0.99 bits per Pixel IMSE = 0.000019

Medium Motion



0.99 bits per Pixel IMSE = 0.000019

Slow Motion



0.17 bits per Pixel IMSE = 0.000038

Very Slow Motion

0.30 bits per Pixel IMSE = 0.000044

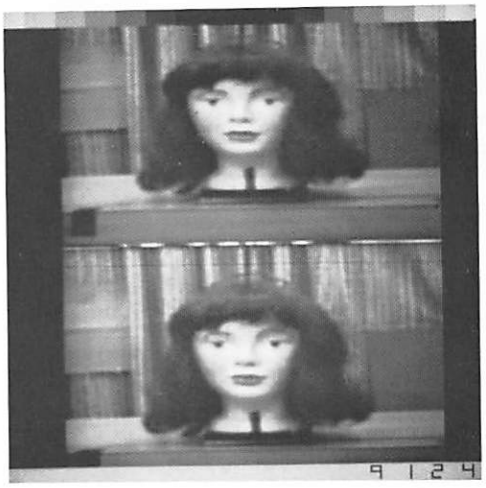
IMSE = Interframe mean square error

Figure 3.2-5. Interframe Coding with Threshold of 4.0. Reconstructions of Second Frames from five pairs of Frames Representing a Range of Activities.

Very fast motion

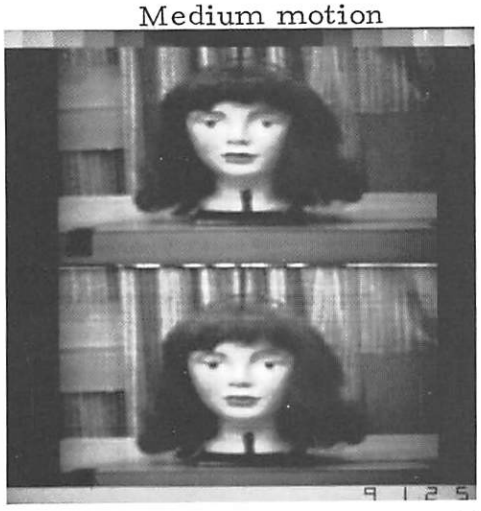
0.94 bits per Pixel IMSE = 0.000061

Fast motion



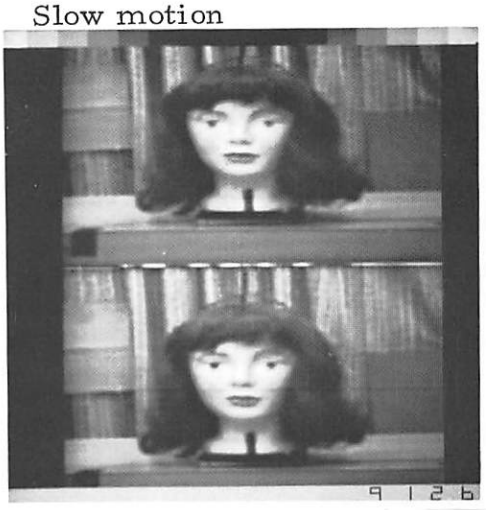
0.87 bits per Pixel IMSE = 0.000042

Medium motion



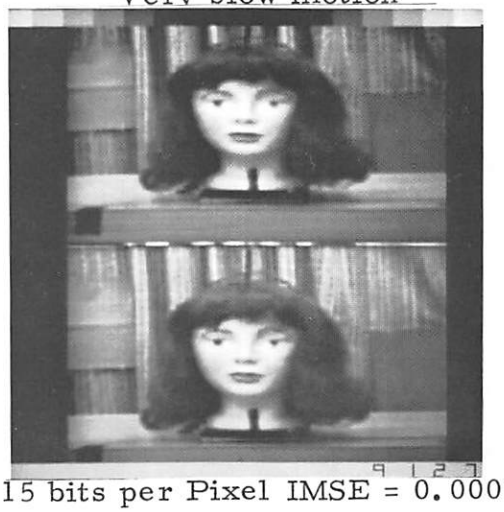
0.72 bits per Pixel IMSE = 0.000063

Slow motion



0.34 bits per Pixel IMSE = 0.000091

Very slow motion



0.15 bits per Pixel IMSE = 0.000066

IMSE - Interframe mean square error

Figure 3.2-6. Interframe Coding with Threshold of 9.0. Reconstruction of Second Frames from Five Pairs of Frames Representing a Range of Activities.

visible error (although this is partially masked by the dark hair of the subject). The discrete error appears at the edges of sub-blocks which were incorrectly updated. Two examples in frame number 9123 are the vertical edge visible on the subjects left hair edge and the L shaped nick in the edge of the hair to the right of and above the subject's left eye. Figure 7 shows the results obtained when coding the sequence of four frames and compares the effect of intraframe coding at 3.0 bits per pixel (part (i)) and 1.5 bits per pixel (part (ii)). No errors are visible in the results of part (i) but a small number may be discerned in the second set of results. The reason for these errors is that the higher intraframe compression to 1.5 bits per pixel lowers the difference signal energy. This is reflected in the higher interframe compression achieved with these results. A slightly lower threshold would alleviate the problem. The results also show that for the limited sequence of four frames no noise build up is evident.

The limited results indicate that the scheme is an effective way of coding moving pictures - the transmission rate is greatly reduced while the coder offers economy over conventional interframe coders. It is not possible to predict the effect of the sub-block update mechanism in real time - it is possible that the sub-block structure may become visible. Lack of suitable data has precluded study of this problem. The coder does produce a non-uniform data rate. Studies of the output data indicate that the bit stream may be smoothed by a buffer or handled by buffer sharing techniques with at least the same efficiency as that obtained with conventional conditional update coders.

3.3 Quantization Error Reduction for Image Coding

Michael N. Huhns

Quantization is the process of representing continuously varying quantities by discrete intervals. This process is nonlinear and some of the information about the original data is irretrievably lost. The usual restoration procedure is to choose the midpoints of each quantization interval as the estimated values of the original data. However if it is known that the original data are correlated and are non-uniformly dis-

i) Intraframe coding at 3.0 bits per pixel

ii) Intraframe coding at 1.5 bits per pixel

Frame 2

1.16 bits per pixel IMSE = 0.000052

0.52 bits per pixel IMSE = 0.000049

Frame 3

1.17 bits per pixel IMSE = 0.000051

0.54 bits per pixel IMSE = 0.000048

Frame 4

1.19 bits per pixel IMSE = 0.000048

0.55 bits per pixel IMSE = 0.000055

IMSE - Interframe Mean Square Error

Figure 3.2-7. Interframe Coding with Threshold of 4.0
Four Frame Sequence

tributed, then improved restorations are possible using this information. As shown in a previous report [1], minimum mean square error estimates of correlated data require the solution of the following equation

$$x = E\{ x | x \in D \} = \frac{\int_D \underline{x} p(\underline{x}) \, d\underline{x}}{\int_D p(\underline{x}) \, d\underline{x}} \qquad (1)$$

where $\underline{x}$ is the n-dimensional variable to be quantized, D is the particular region of n-space into which $\underline{x}$ is quantized, and $p(\underline{x})$ is the probability density function of $\underline{x}$. A partial solution to this equation has been obtained for data which have a jointly gaussian probability distribution. This solution has now been applied to the restoration of quantized one-dimensional random signals and two-dimensional transform domain zonal quantized images. The results reveal a decrease in mean square error in all cases. However, in spite of the error reduction, some images exhibit a degradation in subjective quality after restoration. Hence a nonlinear error criterion based on the human visual system and derived by Mannos and Sakrison [2] has been used in place of the mean square error function. Under this criterion a subjective image improvement as well as a numerical error reduction are obtained.

To demonstrate the utility of this restoration procedure, a randomly generated gaussian Markov signal has been quantized and restored. The results are shown in figure 1. A two bit per sample Max quantization scheme is employed to obtain the quantized approximation to the original signal. Using this quantized signal and the statistical knowledge about the original signal as inputs to the nonlinear estimator, the restoration decreases the mean square error by 33%. The average improvement in mean square error as a function of quantizing bit assignment for different correlation coefficients is shown in figure 2. It can be seen from this graph that, as the amount of correlation in the Markov process approaches zero, then the restoration provides no error improvement. There is also

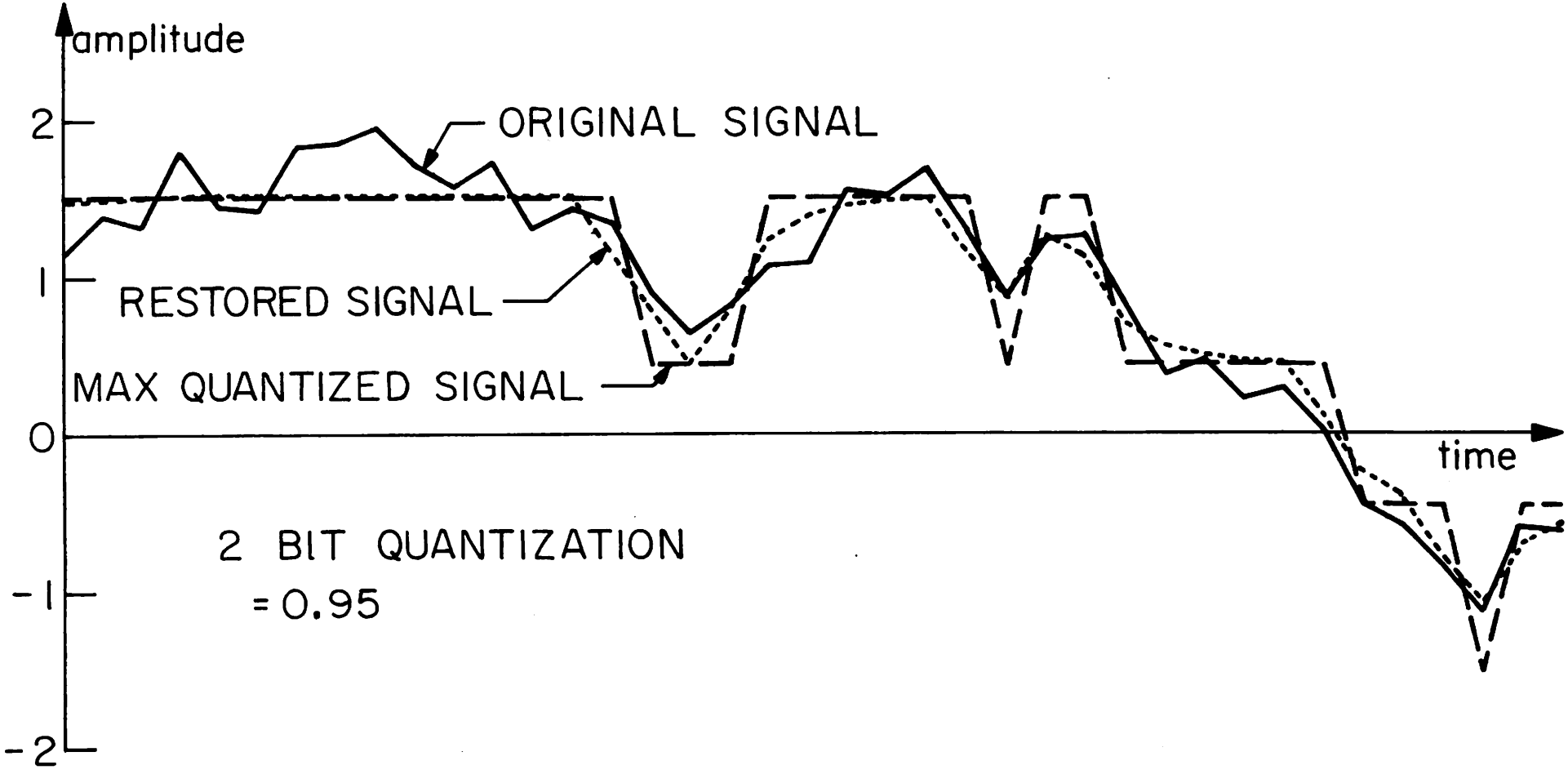


Figure 3.3-1. Restoration of a Quantized Random Markov Signal.

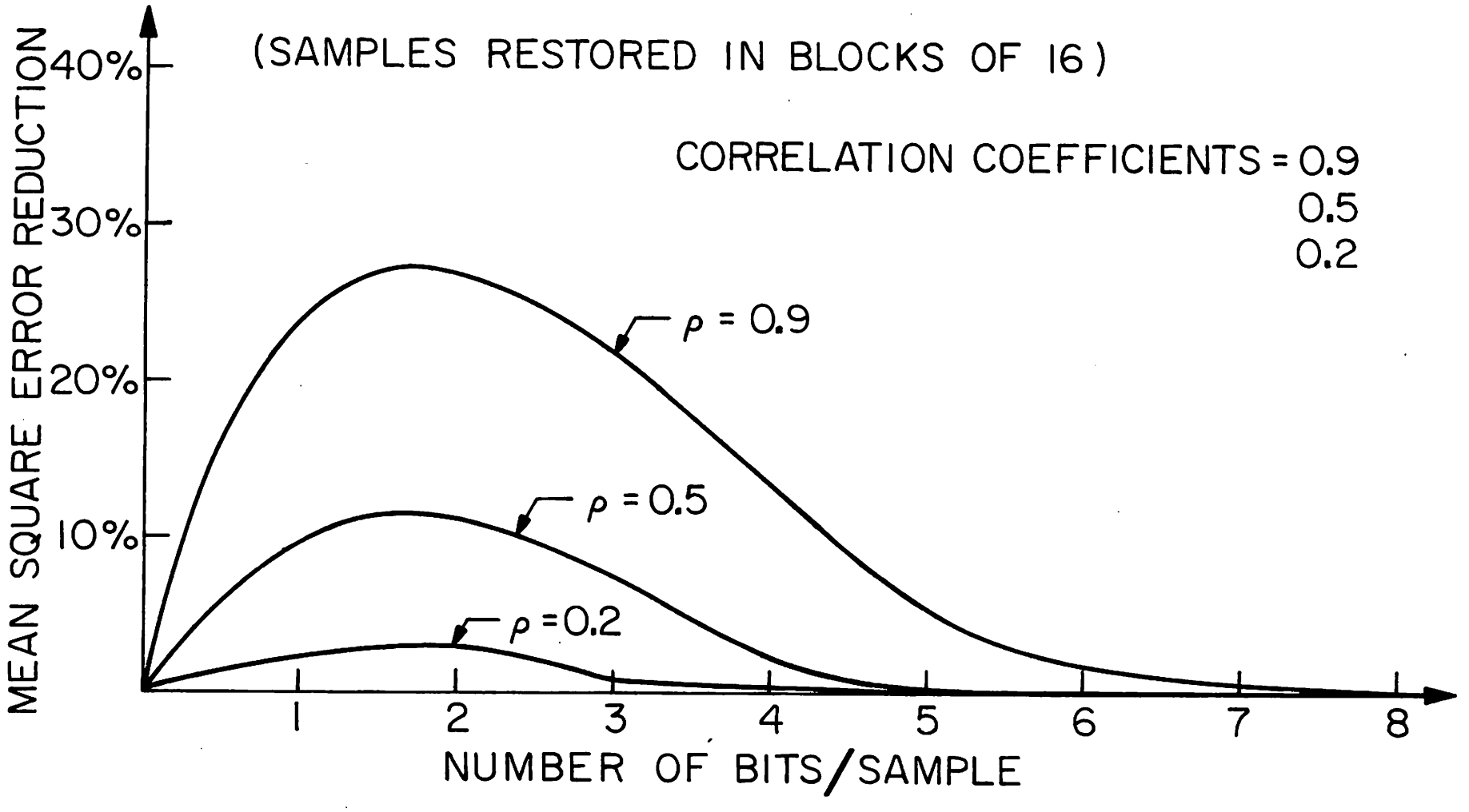


Figure 3.3-2. M.S.E. Improvement for Quantized Markov Signal.

no improvement as the number of quantizing bits becomes large and the differences between the original signal and the quantized signal vanish. Thus the restoration procedure represents a viable restoration technique when the number of quantizing bits are small and the input signals to the quantizer are correlated.

These conditions are satisfied by the zonal transform coding technique for images. In this method the transform samples have a gaussian distribution: each is the sum of a large number of random variables so that the central limit theorem can be invoked. Now these transform samples are typically quantized according to a bit assignment such as the one shown below.

8	6	5	5	4	4	4	4	3	3	3	3	3	3	3
5	4	3	3	2	2	2	2	2	2	2	2	2	2	2
4	3	2	2	1	1	1	1	1	1	1	1	1	1	1
4	3	2	2	1	1	1	1	1	1	1	1	1	1	1
4	2	1	1	1	1	1	1	0	0	0	0	0	0	0
4	2	1	1	1	1	1	1	0	0	0	0	0	0	0
4	2	1	1	1	1	1	1	0	0	0	0	0	0	0
4	2	1	1	1	1	1	1	0	0	0	0	0	0	0
3	2	1	1	0	0	0	0	0	0	0	0	0	0	0
3	2	1	1	0	0	0	0	0	0	0	0	0	0	0
3	2	1	1	0	0	0	0	0	0	0	0	0	0	0
3	2	1	1	0	0	0	0	0	0	0	0	0	0	0
3	2	1	1	0	0	0	0	0	0	0	0	0	0	0
3	2	1	1	0	0	0	0	0	0	0	0	0	0	0
3	2	1	1	0	0	0	0	0	0	0	0	0	0	0
3	2	1	1	0	0	0	0	0	0	0	0	0	0	0

The restoration technique must be applied recursively to these samples since it is only capable of restoring one sample at a time. The current best estimates of the remaining samples are then used to obtain the estimate of the sample being restored. This is repeated for each of the samples in turn and then the entire procedure is repeated until the estimates converge. This convergence has been found to be rapid and, in most cases, one iteration is sufficient.

The above procedure has been applied to the images in figure 3b and 3d. Figure 3b has been coded with an average of one bit per pixel

(a) Original Image 8 Bits/Pixel

(b) Quantized 1 Bit/ Pixel

(c) Restored 1 Bit/Pixel

(d) Quantized 0.5 Bit/Pixel

(e) Restored 0.5 Bit/Pixel

Figure 3.3-3. Restoration of Haar transform, zonal quantized images.

by using a Haar transform in 16 x 16 blocks, a zonal coding bit assignment and a Max quantizer. For the quantization and subsequent restoration, the original image samples are assumed to arise from a Markov source. Figure 3c is the restored version of this image, utilizing one iteration of the estimation procedure. The mean square error is reduced by 10% as a result of the restoration. Figures 3d and 3e, respectively, have been quantized to 0.5 bits and restored by means of the above technique. A reduction of 19% in mean square error is obtained in this case. Subjectively, the restored images appear much less noisy than the quantized images but more blurred, as is very evident in comparing figures 4b and 4c. Hence an error measure is required in which numerical results match subjective results.

This has been provided by modeling the error measure after the human visual system. It has been found that the human visual system is sensitive to approximately the cube root of incident light intensities. It is also most sensitive to middle spatial frequencies near eight cycles per degree. Hence to apply this error measure, an image is processed according to the block diagram in figure 5. The $(i, j)^{th}$ component of the filter function is chosen to be

$$T_{ij} = (.05 + .18525\ r)\ \exp\{-(.07125\ r)^{1.1}\} \tag{3}$$

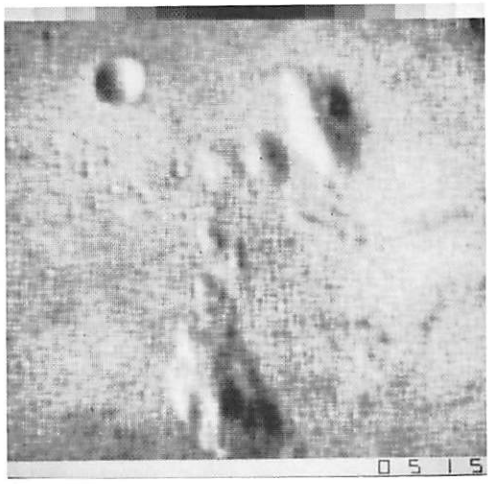
where

$$r = (i^2 + j^2)^{1/2} \tag{4}$$

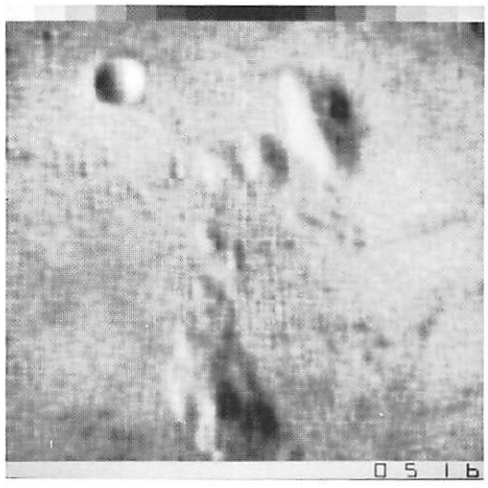
Figures 4d and 4e show the results of this procedure for a Hadamard transform, with and without the restoration step, respectively. There are both subjective and numerical error improvements after restoration.

The restoration process has also been applied to a color coding experiment. In this experiment a color image is transformed to the YIQ coordinate system and then quantized according to the bit assignment indicated on the next page for a typical block of four pixels.

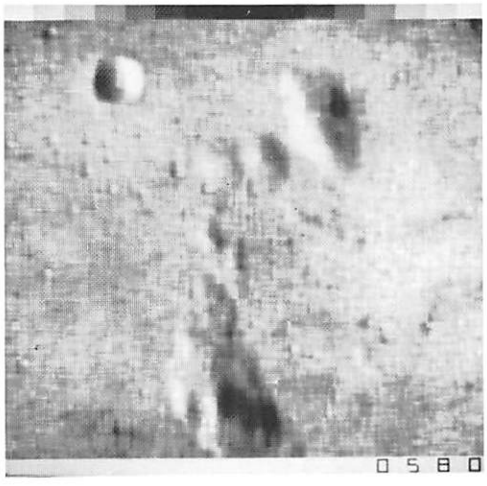
(a) Original Image 8 bits/pixel



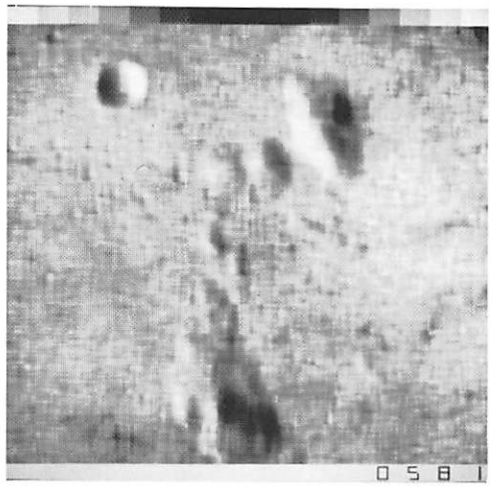
(b) Zonal Haar Quantized



(c) Zonal Haar Restored



(d) Visual Hadamard Quantized



(e) Visual Hadamard Restored

Figure 3.3-4. Restoration of 0.5 bit/pixel transform coded images for two different error criteria.

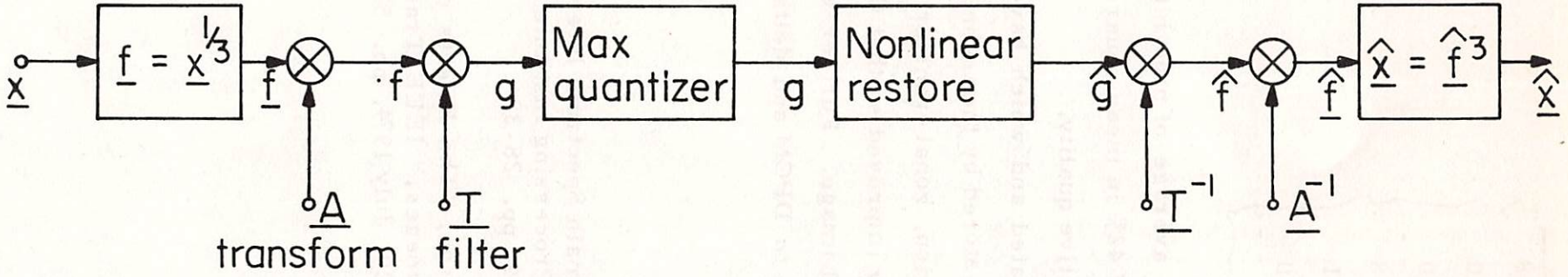
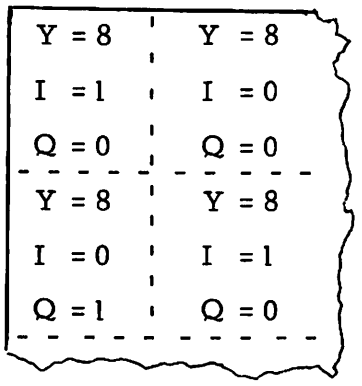


Figure 3.3-5. Coding and restoration technique for a nonlinear error criterion.



Each pixel is hence coded with an average of nine bits. Quantization restoration provides a decrease of 42% in mean square error in this case and an improvement in subjective quality.

Thus data which are correlated and which have been coarsely quantized are amenable to being restored by the techniques outlined above. By choosing a suitable error criterion, zonal transform coded images can be subjectively and analytically improved so that they more faithfully reproduce the details of an original image. Further work is expected to extend the restoration technique to DPCM and delta modulation coded images.

References

1. M. N. Huhns, "Transform Domain Spectrum Interpolation," University of Southern California Image Processing Institute Technical Report, USCIPI Report 530, March 1974, pp. 28-38.

2. J. L. Mannos and D. J. Sakrison, "The Effects of a Visual Fidelity Criterion on the Encoding of Images," IEEE Transactions on Information Theory, Vol. IT-20, July 1974, pp. 525-536.

4. Image Restoration and Enhancement

Image restoration and image enhancement are two classifications of image improvement methods. Image restoration techniques seek to reconstruct or recreate an image to the form it would have had if it had not been degraded by some physical imaging system. Image enhancement techniques have two major purposes: improvement in the visual quality of a picture to a human viewer; and manipulation of a picture for more efficient processing and data extraction by a machine. Research in both areas during the past six months is described below.

The first report addresses the general problem of restoration for space variant aberrations. Methods are presented for astigmatism restoration and restoration methods for other types of degradation are outlined.

The three reports following are concerned with pseudoinverse techniques of image restoration. In the first of the series a method is presented for performing restoration in the eigenspace of the degradation in order to avoid computational errors. The second paper describes a restoration technique which seeks to ensure that the estimated image pixel values are neither negative nor greater than an upper bound. In the third paper a computational algorithm for pseudoinverse restoration is developed. With this algorithm computation is reduced significantly compared to conventional methods and numerical stability is also improved.

The next paper discusses the use of spline function interpolation in the development of an image restoration algorithm, and in the next paper the effects of non-uniform sampling of a blurred image are analyzed. It is shown that proper non-uniform sampling actually can improve the quality of an image restoration.

Finally, a method of histogram exponentiation for image detail enhancement is described. With this process a grey level mapping is performed to produce an output image whose histogram follows an exponential density.

4.1 Restoration for Space-Variant Aberrations

Alexander A. Sawchuk and M. Javad Peyrovian

Several previous publications have detailed the difficulty of achieving image restoration for the most general case when the linear degrading system is space-variant. By systematically examining the sources of degradation and using all available a priori knowledge, considerable simplification of the problem can be achieved [1, 5].

An example of restoration for images of 100 x 100 pixels degraded by third-order astigmatic aberrations has been previously described, and descriptions of space-variant point-spread functions (SVPSF) have been given [1, 5]. Recent progress in this area has permitted images with astigmatism of up to 128 x 128 pixels in size to be restored on the USC 360/44 computer system using an SVD technique. With the addition of memory and a faster processor now underway, these capabilities should be increased in the near future. Another improvement in the simulation and testing of the astigmatism algorithm has resulted from the use of an improved quadrature formula to evaluate the space-variant integral

$$\mathcal{I}(x_1, x_2) = \int_{-\infty}^{\infty}\int h(x_1, x_2, u_1, u_2)\, \mathcal{O}(u_1, u_2)\, du_1 du_2 \qquad (1)$$

describing the degradation process. Here $\mathcal{O}(u_1, u_2)$ and $\mathcal{I}(x_1, x_2)$ are the object and image, respectively, and $h(x_1, x_2, u_1, u_2)$ is the degrading SVPSF. The new quadrature formula involves the expansion of $h(x_1, x_2, u_1, u_2)$ into spline functions, thus eliminating artifacts in the digital simulation of the imaging system and providing a better test of the restoration method.

A test of space-variant restoration for the most difficult third-order aberrations of combined astigmatism and curvature of field is now in process. The theory behind the technique is an extension of the method for astigmatism alone. The general degradation of eq. (1)

is first rewritten with a polar coordinate transformation in the form

$$\mathcal{I}(x_r, x_\theta) = \int_{-\infty}^{\infty}\int h(x_r, u_r, x_\theta, u_\theta)\, \mathcal{O}(u_r, u_\theta)\, du_r du_\theta \tag{2}$$

where (x_r, x_θ) and $u_r, u_\theta)$ are image and object polar coordinate variables. This operation is digitally performed on the degraded image by rearranging samples in memory. A simplification results because of the circular symmetry of many optical systems. In this case, the system aberration function are

$$x_r - u_r = (2C+D)\, u_r^2 \epsilon_r \cos \epsilon_\theta \tag{3a}$$

$$x_\theta - u_\theta = \tan^{-1}[D u_r \epsilon_r \sin \epsilon_\theta / 1 + (2C+D)\, u_r \epsilon_r \cos \epsilon_\theta] = g(u_r) \tag{3b}$$

where C and D are the astigmatism and curvature of field coefficients respectively, and ϵ_r and ϵ_θ are exit pupil variables. The SVPSF's derived from eq. (3) and limiting cases of C = 0 and D = 0 can be easily derived [1].

By observing that the radial blur of eq. (3a) is a function of radial coordinates only, and that the θ blur of eq. (3b) is a slowly varying function $g(u_r)$, the next step in restoration is to rewrite eq. (2) as

$$\mathcal{I}(x_r, x_\theta) = \int_{-\infty}^{\infty}\int h(x_r, u_r, x_\theta - u_\theta, g(u_r))\, \mathcal{O}(u_r, u_\theta)\, du_r du_\theta \tag{4}$$

to emphasize the dependence. If D = 0 in eq. (3) the degradation is pure astigmatism without curvature of field and the blur is entirely radial with no blurring in θ.

Defining a Fourier transform of $\mathcal{I}(x_r, x_\theta)$ in the x_θ variable by

$$\bar{\mathcal{I}}(x_r, \lambda) = \int_{-\infty}^{\infty} \mathcal{I}(x_r, x_\theta) \exp(-j2\pi\lambda x_\theta)\, dx_\theta \tag{5}$$

the transform of both sides of eq. (4) is taken to obtain

$$\bar{\mathcal{I}}(x_r, \lambda) = \int_{-\infty}^{\infty}\int \mathcal{O}(u_r, u_\theta)\, \bar{h}(x_r, u_r, \lambda, g(u_r)) \exp(-j2\pi\lambda u_\theta) du_\theta du_r \tag{6}$$

where $\bar{h}$ is the transform of h in x_θ. Grouping terms containing u_θ on the right side of eq. (6) enables a transform in this variable to be evaluated. The resulting transformed function $\bar{\mathcal{O}}(u_r, \lambda)$ is given by

$$\bar{\mathcal{O}}(u_r, \lambda) = \int_{-\infty}^{\infty} \mathcal{O}(u_r, u_\theta) \exp(-j2\pi\lambda u_\theta) du_\theta \tag{7}$$

and the reduced system equation obtained from eq. (4) is

$$\bar{\mathcal{I}}(x_r, \lambda) = \int_{-\infty}^{\infty} \tilde{h}(x_r, u_r, \lambda)\, \bar{\mathcal{O}}(u_r, \lambda)\, du_r \tag{8}$$

where

$$\tilde{h}(x_r, u_r, \lambda) = \bar{h}(x_r, u_r, \lambda, g(u_r)) \tag{9}$$

is rewritten as a function of three variables to show explicit dependence.

This derivation shows the procedure to be used for the astigmatism and curvature of field restoration. Following a polar coordinate transformation, a Fourier transform in x_θ as expressed by eq. (5) is performed to partially decouple θ blur as a slowly varying function of $g(u_r)$. The reduced system is then given by eq. (8), and an estimate of $\bar{\mathcal{O}}(u_r, \lambda)$ is produced by singular-value decomposition (SVD) techniques [1] for each separate λ by techniques similar to that used for astigmatism alone. Efforts will be made to reduce the computational effort required in this part by using the known variation of $h(x_r, u_r, \lambda)$ with λ. After the entire $\bar{\mathcal{O}}(u_r, \lambda)$ has been obtained, a series of one-dimensional inverse transforms in x is taken to find $\mathcal{O}(u_r, u_\theta)$, and an inverse polar coordinate distortion is used to get $\mathcal{O}(u_1, u_2)$ as the final restored object.

This procedure, while requiring large capabilities in computing and storage, is the only method for doing restoration on images of even moderate size. The general four-dimensional space-variant blur of eq. (1) is effectively reduced to a set of space-variant two-dimensional problems whose point-spread functions depend in a well-behaved way on λ.

Following tests of this procedure, future work on this problem will include modifications for noisy data and an analysis of sampling problems and approximations in performing a continuous space-variant restoration by computer. The analysis as presented is for a general aperture, although additional simplification may be possible for other specialized aperture shapes.

References

1. A. A. Sawchuk and M. J. Peyrovian, "Space-Variant Restoration of Astigmatism and Curvature of Field," USC Image Proc. Inst. Semi-Annual Tech. Report, USCIPI 530, 1 September 1973 - 28 February 1974, pp. 68-74.

2. A. A. Sawchuk, "Space-Variant Image Restoration by Coordinate Transformations," J. Optical Soc. Amer., Vol. 64, February, 1974, pp. 124-130.

3. A. A. Sawchuk, "Restoration of Space-Variant Aberrations by Coordinate Transformations," USC Semi-Annual Tech. Report, USCEE 459, 1 March 1973 - 31 August 1973, pp. 83-90.

4. G. M. Robbins and T. S. Huang, "Inverse Filtering for Linear Shift-Variant Imaging Systems," Proc. IEEE, Vol. 60, July, 1972, pp. 862-872.

5. A. A. Sawchuk, "Point-Spread Functions and Image Restoration with Astigmatism and Curvature of Field," J. Optical Soc. Amer., Vol. 64, No. 4, April, 1974, p. 560.

4.2 Image Restoration in the Eigenspace of the Degradation

Harry C. Andrews and Monty Adler

It has recently been noted that there are certain mathematical properties associated with point spread function blur matrices that make digital image restoration extremely attractive in the eigenspace of the degradation [1]. Consider the separable space variant imaging model of

$$\underline{G} = \underline{A}\,\underline{F}\,\underline{B} \tag{1}$$

where $\underline{G}$ is the matrix of available image data points and $\underline{F}$ is the array of object data points which are of interest for estimation. Usually both $\underline{A}$ and $\underline{B}$, the column and row blur matrices respectively, are nearly singular. Penrose [2] investigated this model, without regard to imaging systems, and suggested the pseudoinverse restoration which provides an estimate of the object as

$$\tilde{\underline{F}} = \underline{A}^{+}\underline{G}\,\underline{B}^{+} \tag{2}$$

where $\underline{A}^{+}$ and $\underline{B}^{+}$ are Moore-Penrose pseudoinverses respectively [3]. The blur matrices $\underline{A}$ and $\underline{B}$ are componentwise positive and experimentally appear to be oscillatory [4, 5]. The oscillatory property implies that the eigenvectors associated with decreasingly smaller eignvalues have increasingly larger zero crossings. Examination of the pseudoinverse solution of eq. (2) in eigenspace, reveals that the solution can be cast in the form

$$\tilde{\underline{F}} = \underline{V}_a\,\underline{\Lambda}_a^{+}\,\underline{U}_a^{T}\underline{G}\underline{V}_b\,\underline{\Lambda}_b^{T}\,\underline{U}_b^{T} \tag{3}$$

where

$$\underline{A} = \underline{U}_a\,\underline{\Lambda}_a\,\underline{V}_a \tag{4a}$$

and

$$\underline{B} = \underline{U}_b\,\underline{\Lambda}_b\,\underline{B}_b^{T} \tag{4b}$$

All matrices are orthogonal with $\underline{\Lambda}_a$ and $\underline{\Lambda}_b$ being diagonal. The matrices $\underline{\Lambda}_a^+$ and $\underline{\Lambda}_b^+$ are also diagonal with non-zero entries being reciprocals of the entries of $\underline{\Lambda}_a$ and $\underline{\Lambda}_b$ respectively, except for zeros placed on the diagonals of $\underline{\Lambda}_a^+$ and $\underline{\Lambda}_b^+$ for small (approximately zero) diagonal entries of $\underline{\Lambda}_a$ and $\underline{\Lambda}_b$. By noting that

$$\underline{\tilde{F}} = \underline{V}_a \underline{\Lambda}_a^T \underline{\alpha} \underline{\Lambda}_b^T \underline{U}_b^T \tag{5}$$

where

$$\underline{\alpha} = \underline{U}_a^T \underline{G} \underline{V}_b \tag{6}$$

it is found that the $\underline{\alpha}$ matrix is the image $\underline{G}$ represented in the eigenspace of $\underline{A}$ and $\underline{B}$. Because of the oscillatory properties of $\underline{A}$ and $\underline{B}$, contributions to $\underline{F}$ associated with large λ_i^{-1} coefficients have correspondingly large zero crossing eigenvectors. These eigenvectors are then correlated with the image $\underline{G}$ which tends to be slowly varying; and as such, the correlation tends to zero. But those eigenvalues of least confidence, due to uncorrelated noise and computational error, are associated with eigenimages of many sign changes, which in turn correlate to zero with $\underline{G}$. Consequently one encounters the happy circumstance of computational error not being as critical as could be as a result of effective zero correlation of the image $\underline{G}$ with the corresponding computationally questionable eigenvalued eigenimages. Thus greater success can be anticipated with this pseudoinverse approach for image restoration than for restoration of signals with correspondingly higher frequency content. Restoration algorithms based upon these observations are currently under investigation.

References

1. Ekstrom, M. P. "Proceedings of the Hawaii International Conference on System Science," January 1974.

2. Penrose R. "On Best Approximate Solutions of Linear Matrix

2. Penrose R., "On Best Approximate Solutions of Linear Matrix Equations," Proceedings Cambridge Philos. Soc., Vol. 52, pp. 17-19.

3. Albert, A., Regression and the Moore-Penrose Pseudo-inverse, Academic Press, New York, 1972.

4. Gantmacher, F. R., The Theory of Matrices, Vol. II, Cheleia Publishing Company, New York, New York.

5. Gantmacher, F. R. and M. G. Krein, Oscillation Matrices and Kernels and Small Vibration of Dynamical Systems, 2nd Ed. Moscow Gostekhizdat.

4.3 Pseudoinverse Method of Bounded Image Restoration

Harry C. Andrews and Monty Adler

The space variant separable imaging equation given by

$$\underline{G} = \underline{A}\,\underline{F}\,\underline{B} \tag{1}$$

has been investigated from a restoration viewpoint in previous reports. With $\underline{A}$ and $\underline{B}$ being the column and row blurs of the object $\underline{F}$ respectively, the image $\underline{G}$ will be used as a means of estimating F. Traditional solutions to this problem utilize the pseudoinverse in which

$$\hat{\underline{F}} = \underline{A}^{+}\underline{G}\underline{B}^{+} \tag{2}$$

and $\underline{A}^{+}$ and $\underline{B}^{+}$ are pseudoinverses of $\underline{A}$ and $\underline{B}$ respectively. Usually the blur matrices are nearly singular. Unfortunately, although the pseudoinverse solution provides the best estimate of $\underline{F}$ in a least squares sense for a minimum norm $\hat{\underline{F}}$, (i.e. $\hat{\underline{F}}$ minimizes $\|\hat{\underline{F}}\|$), the estimate does not take advantage of all a priori knowledge. For example, the solution ignores the fact that the elements of $\underline{F}$ are intensities, which cannot go negative (therefore positive restoration), and which cannot exceed some physically realizable total light energy (therefore bounded restoration). To utilize the positive bounded restoration model the pseudoinverse solution of eq. (2) is suggested as an initial condition to a nonlinear programming algorithm to guarantee positivity and upper

boundness. Towards this end, consider the pseudoinverse solution

$$\underline{F} = \underline{V}_a \underline{\Lambda}_a^+ \underline{U}_a^T \underline{G} \underline{V}_b \underline{\Lambda}_b^+ \underline{U}_b^T \tag{3}$$

where

$$\underline{A} = \underline{U}_a \underline{\Lambda}_a \underline{V}_a^T \tag{4a}$$

and

$$\underline{B} = \underline{U}_b \underline{\Lambda}_b \underline{V}_b^T \tag{4b}$$

The objective function for normalization purposes will be to minimize $\|\underline{G}-\hat{\underline{G}}\|^2$ subject to the bounded restoration constraint on $\hat{\underline{F}}$. Thus

$$\underline{W} = \|\underline{G}-\hat{\underline{G}}\|^2 \tag{5a}$$

$$\underline{W} = \|\underline{A}(\underline{F}-\hat{\underline{F}})\,\underline{B}\|^2 \tag{5b}$$

$$\underline{W} \leq \|\underline{A}\|^2 \|\underline{F}-\hat{\underline{F}}\|^2 \|\underline{B}\|^2 \tag{6}$$

Therefore, for analysis purposes one can choose to minimize $\|F-\hat{F}\|^2$. However, since the pseudoinverse is simply being used as an initial condition and since the eigenspace of the degradation is a useful domain for restoration processing, the variables of optimization are restricted to be diagonal in the eigenspace domain. Thus the estimate

$$\hat{\underline{F}} = \underline{V}_a \underline{\delta}_a \underline{\Lambda}_a^+ \underline{U}_a^T \underline{G} \underline{V}_b \underline{\Lambda}_b^+ \underline{\delta}_b \underline{U}_b^T \tag{7}$$

is found by adjusting the $\underline{\delta}_a$ diagonal matrices such that $\|F-\hat{F}\|^2$ is minimized and $\hat{\underline{F}}$ is componentwise positive and bounded. This reduces the restoration problem to a 2K variable problem with N^2 positive bounded boundary constraints. Here K is the number of non zero

eigenvalues retained in the $\hat{\underline{F}}$ estimate. Since the solution is in the eigenspace domain, the effect of each δ_i variable is felt throughout the entire estimated object $\underline{F}$, and as such is much more effective than simply constraining a component entry (pixel) of $\hat{\underline{F}}$ itself. A Fiacco and McCormick [1] programming algorithm has been utilized to adjust the weights in degradation eigenspace to obtain positive bounded restorations with some success. While this work is still in progress, it appears that convergence is very rapid due to the pseudoinverse starting point whereby it is meant that an iterative pseudoinversion technique is developed to obtain $\underline{F}_K$ until the positive bounded constraints are violated at which time the nonlinear programming algorithm is call to correct for the near singularity in the K^{th} pseudoinverse solution.

Reference

1. Fiacco, A. V. and G. P. McCormick (1968) Nonlinear Programming: Sequential Unconstrained Minimization Techniques, Wiley, New York, New York.

4.4 A Fast Pseudoinverse Image Restoration Algorithm

William K. Pratt and Faramarz Davarian

It is often possible to model an image degradation process by the vector equation

$$\underline{g} = \underline{B}\,\underline{f} + \underline{n} \tag{1}$$

where $\underline{g}$ denotes a column scanned M x 1 vector of physical samples of the blurred image; $\underline{f}$ is an N x 1 vector of column scanned points in the ideal image field, $\underline{B}$ is an M x N blur operator matrix representing a convolutional blur; and $\underline{n}$ denotes an M x 1 vector of observation noise or uncertainty. For this model, the ideal image vector can be estimated by pre-multiplication of the observation by the generalized inverse, $\underline{B}^-$, of the blur matrix, Thus,

$$\underline{f} = \underline{B}^{-} \underline{g} \tag{2}$$

where $\underline{B}^{-}$ is an N x M matrix that can be computed by

$$\underline{B}^{-} = (\underline{B}^{T}\underline{B})^{-1}\underline{B}^{T} \tag{3a}$$

if B is of full column rank, or can be computed by

$$\underline{B}^{-} = \underline{B}^{T}(\underline{B}\underline{B}^{T})^{-1} \tag{3b}$$

if $\underline{B}$ is of full row rank. The solution of eq. (2) is a minimum mean square error, minimum norm estimate. There are two major difficulties with pseudoinverse restoration: if the noise level is high, the solution may be unstable as a result of the usual ill conditioning of the blur matrix; and, computation of the generalized inverse and restoration by eq. (2) is usually a large task. The former problem can be avoided by restoration constraints; consideration is given here to efficient computational techniques for pseudoinverse restoration.

Fast Pseudoinverse Algorithm. As a simplification in the development of the algorithm, consideration will be initially limited to a one dimensional model in which the ideal image is represented by quadrature points at its Nyquist rate and the blurred image is sampled at the same rate. Then, let the impulse response be represented by the L x 1 vector $\underline{h}$. The blur matrix of eq. (1) then assumes the form

$$\underline{B} = \begin{bmatrix} h(L) & \cdots & h(1) & 0 & \cdots & 0 \\ 0 & h(L) & \cdots & h(1) & \cdots & 0 \\ \vdots & & & & & \vdots \\ \cdots & & 0 & h(L) & \cdots & h(1) \end{bmatrix} \tag{4}$$

In this case the number of ideal image points and the number of observed image samples are related by

$$N = M + L-1 \tag{5}$$

Now, let two vectors $\underline{f}_T$ and $\underline{g}_T$ be formed by selecting the center portions of $\underline{f}$ and $\underline{g}$, respectively. These truncated vectors are obtained by dropping L-1 elements at each end of the appropriate vector by the operations

$$\underline{f}_T = \underline{S2}_N^{(K)} \qquad K = N-2(L-1)$$

$$\underline{g}_T = \underline{S2}_M^{(R)} \qquad R = M-2(L-1)$$

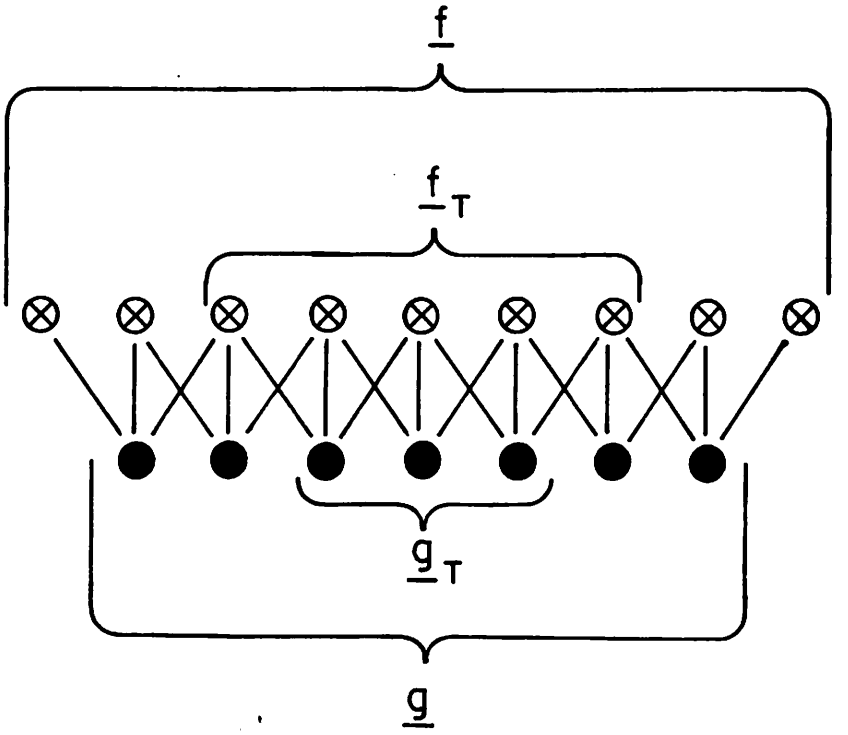
where

$$\underline{S2}_J^{(K)} = \overbrace{\begin{bmatrix} \underline{0} & \vdots & \underline{I}_K & \vdots & \underline{0} \end{bmatrix}}^{L-1 \quad K \quad L-1} \Big\} K$$

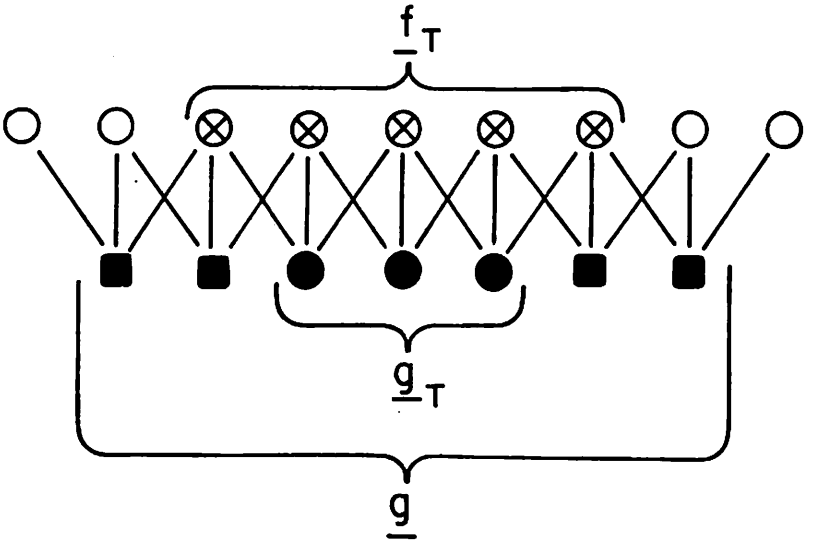
for J = K + 2(L-1). Figure 1a illustrates the relationship of all vectors for N = 9 original vector points, M = 7 observations, and an impulse response of length L = 3.

Suppose, now that the data sequence f_T is discretely convolved with the impulse response sequence n yielding the output sequence q as defined by the vector equation

$$\underline{q} = \underline{D}\,\underline{f}_T \tag{6}$$



(a) sampled continuous convolution



(b) discrete convolution

Figure 4.4-1. Examples of one dimensional sampled continuous convolution and discrete convolution.

where

$$\underline{D} = \begin{bmatrix} h(1) & 0 & \cdots & 0 \\ h(2) & h(1) & & \vdots \\ \vdots & \vdots & & \vdots \\ h(L) & h(L-1) & & 0 \\ 0 & h(L) & & h(1) \\ \vdots & \vdots & & \vdots \\ 0 & \cdots & & h(L) \end{bmatrix} \tag{7}$$

Figure 1b illustrates the relationships between the vector sequence for discrete convolution. An estimate of $\underline{f}_T$ can be obtained by premultiplying the output vector $\underline{q}$ by the generalized inverse of $\underline{D}$. That is

$$\underline{f}_T = \underline{D}^{-} \underline{q} \tag{8}$$

Referring to figure 1, it is observed that the elements of $\underline{g}_T$ are identical to the center elements of $\underline{q}$. Thus,

$$\underline{g}_T = [\underline{S2}_M^{(R)}]\,\underline{g} = [\underline{S2}_M^{(R)}]\,\underline{q}$$

while the remaining end elements differ in general. Now, let a matrix $\underline{W}$ be defined which operates on the physical sample vector $\underline{g}$ to produce an approximation to the output vector $\underline{q}$ for discrete convolution.

$$\tilde{\underline{q}} = \underline{W}\,\underline{g} \tag{a}$$

The structure of $\underline{W}$ will be derived later. Then, an estimate of $\underline{f}_T$ is formed by

$$\underline{\tilde{f}}_T = \underline{D}^- \underline{\tilde{q}} = \underline{D}^- \underline{W}\,\underline{g} \tag{10}$$

By this procedure an estimate of the center part of the image vector $\underline{f}$ can be obtained by use of the generalized inverse operator $\underline{D}^-$ rather than the generalized inverse operator $\underline{B}^-$. The advantages of the former procedure over the latter are that, in the absence of noise, the solution obtained with $\underline{D}^-$ is unique and exact, whereas there are an infinite number of feasible solutions for the $\underline{B}$ model. Also, with the $\underline{D}^-$ operator, it is possible to perform the restortion by Fourier domain processing quite efficiently.

Consideration will now be given to the structure of the weighting matrix $\underline{W}$. The objective of weighting is to express the vector $\underline{q}$ in terms of the elements of the observation $\underline{g}$. For the example of figure 1,

$$\begin{aligned} q(1) &= g(1) - h(3)f(1) - h(2)\,f(2) \\ q(2) &= g(2) - h(3)f(2) \\ q(m) &= g(m) \qquad 2 < m < M-2 \\ q(M-1) &= g(M-1) - h(1)f(N-1) \\ q(M) &= g(M) - h(2)f(N-1) - h(1)\,f(N) \end{aligned} \tag{11}$$

Since the values of $\underline{f}$ are not known, the correspondence of eq. (11) cannot be made directly. However, by making an assumption on the continuity of the original image vector that

$$f(1) = f(2) = f(3)$$

and

$$f(N-2) = f(N-1) = f(N)$$

then it is found that

$$\tilde{q}(1) = \frac{g(1)\, h(1)}{S}$$

$$q(2) = \frac{g(2)[\, h(1) + h(2)]}{S}$$

where $S \equiv h(1) + h(2) + h(3)$. Similar equations exist for $\tilde{q}(M-1)$ and $\tilde{q}(M)$. This procedure can be generalized for any size vectors. Also, more complex prediction algorithms may be employed. Figure 2 illustrates the expected mean square restoration error of $\underline{f}_T$ for various prediction algorithms as a function of the correlation of elements of $\underline{f}$ under the assumption that $\underline{f}$ is a sample of a Markov process with correlation factor ρ. A first order predictor provides a significant improvement over the zero order predictor, and for high correlation factors even surpasses the estimate for the underdetermined model.

The efficiency of the computational algorithm is based upon the use of a circulant blur operator defined as

$$\underline{C} = \begin{bmatrix} h(1) & 0 & \cdots & 0 & h(3) & h(2) \\ h(2) & h(1) & & & \cdot & \cdot \\ \cdot & h(2) & & & h(L) & \cdot \\ \cdot & \cdot & & & 0 & h(L) \\ h(L) & h(L-1) & & & \cdot & \cdot \\ \cdot & h(L) & & & \cdot & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ 0 & 0 & \cdots & 0 & h(1) & 0 \\ 0 & 0 & \cdots & 0 & h(2) & h(1) \end{bmatrix}$$

Then, for the noise free case, it can be shown that

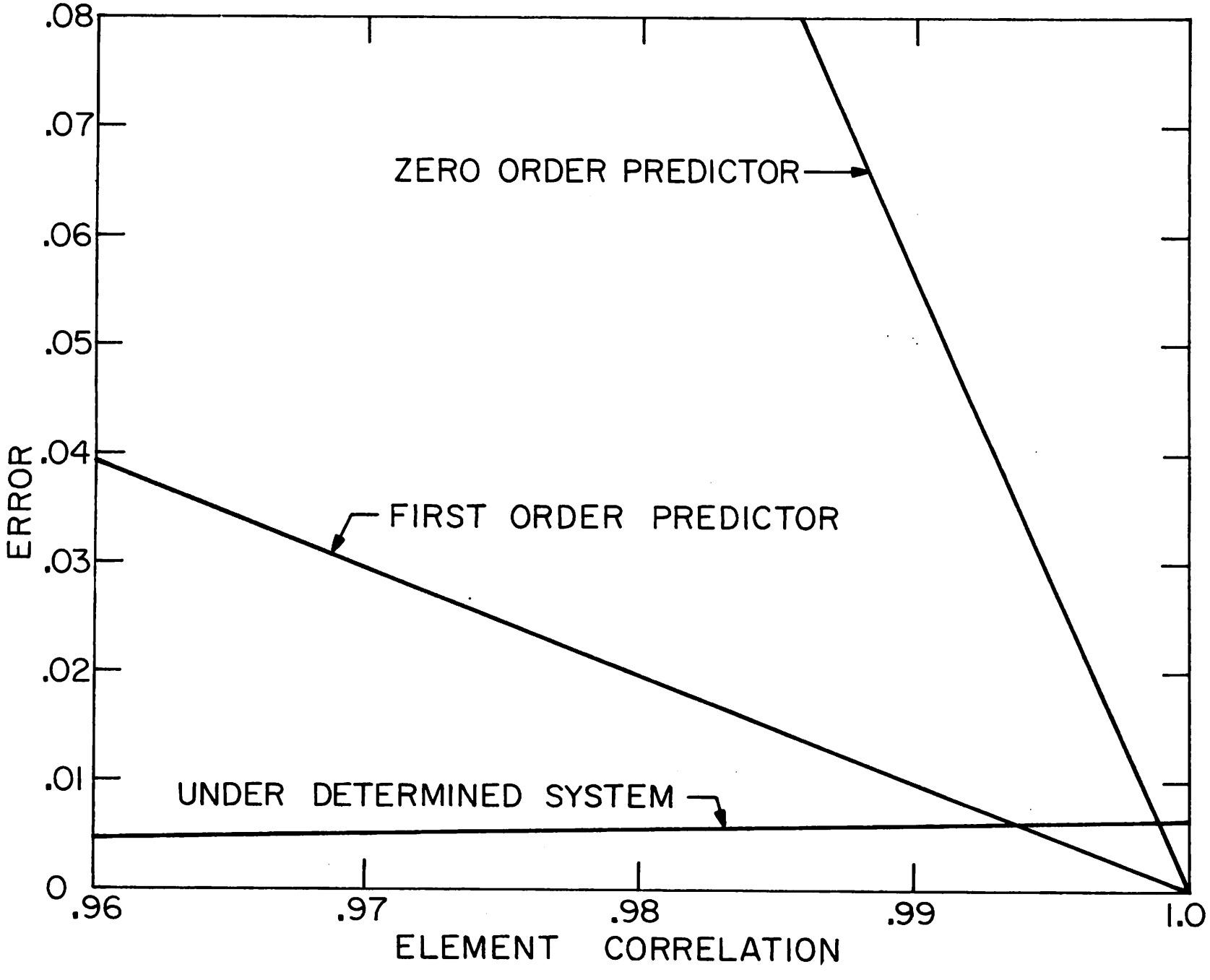
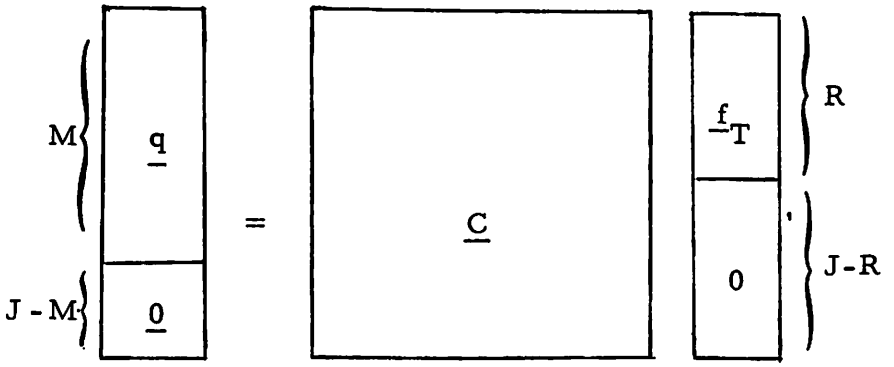
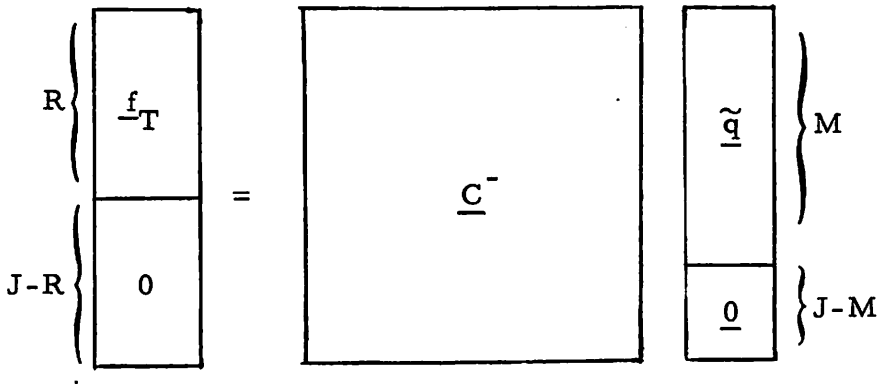


Figure 4.4-2. Restoration error for fast pseudoinverse restoration.



And hence, the image estimate may be obtained by



where $\underline{C}^-$ is the generalized inverse of $\underline{C}$, which is also circulant. Since, the two dimensional Fourier transform of a circulant matrix is of diagonal form, it becomes more efficient to perform the computations in the Fourier domain. The J^2 matrix operations associated with the pseudoinverse multiplication can then be replaced by J scalar multiplies plus 2 J $\log_2 J$ operations required for the fast Fourier transformations. While the fast pseudoinverse algorithm generalizes quite easily for two dimensional image fields. Application of the technique to image restoration is now underway.

4.5 Spline Function Image Restoration.

Steve Hou and Harry C. Andrews

For an estimated object the continuous - discrete imaging system model can be represented as the discrete matrix equation

$$(\underline{A} + \gamma \underline{B})\ \underline{c} = \underline{d} \tag{1}$$

or as

$$(\hat{\underline{A}}^T \hat{\underline{A}} + \gamma \underline{B}_1 \times \underline{B}_2)\ \underline{c} = \hat{\underline{A}}^T \underline{g} \tag{2}$$

using cubic B spline interpolation for the estimated object. In eqs. (1) and (2)

$$\underline{A} \equiv \left[\int_{-\infty}^{\infty} \int h_{ij}(\zeta, \eta)\ S_k(\zeta)\ S_\ell(\eta)\ d\zeta d\eta \right]$$

$$\underline{B}_1 \equiv \left[\int_{-\infty}^{\infty} S_k''(\zeta)\ S_m''(\zeta)\ d\zeta \right]$$

$$\underline{B}_2 \equiv \left[\int S_\ell''(\eta)\ S_n''(\eta)\ d\eta \right]$$

The size of the image is I x J and that of the interpolating grids for the estimate object is M x N. The term $S_k(\zeta)$ denotes the one-dimensional cubic B spline function centered at grid point x_k and $\underline{d}$ and $\underline{c}$ are column vectors of the image g_{ij} with unknown coefficients $c_{k\ell}$ arranged in a lexicographic order. The symbol $\otimes$ designates a Kronecker (tensor or direct) product between two matrices [1]. Both matrices

$$\underline{A} = \hat{\underline{A}}^T \hat{\underline{A}}$$

and

$$\underline{B} = \underline{B}_1 \otimes \underline{B}_2$$

are real, symmetric, non-negative definite, and the diagonal terms are larger than any off-diagonal terms. Furthermore the entries of A are positive. Making use of the following equal energy constraint on the imaging system

$$\sum_{i=1}^{I} \sum_{j=1}^{J} h_{ij}(\zeta, \varkappa) = 1$$

for all ζ and $\varkappa$, $\underline{A}$ can be shown to be a Markov matrix [1].

Ultilizing the special properties of cubic B spline function one can further show that

$$\underline{B}_1 = \frac{1}{6\Delta^3} \begin{bmatrix} 32 & -9 & 2 & 1 & & & & & 0 \\ -9 & 32 & -9 & 2 & 1 & & & & \\ 2 & -9 & 32 & -9 & 2 & & & & \\ 1 & 2 & -9 & 32 & -9 & & & & \\ & 1 & \bullet & \bullet & \bullet & \bullet & & & \\ & & \bullet & \bullet & \bullet & \bullet & \bullet & & \\ & & & \bullet & \bullet & \bullet & \bullet & \bullet & \\ & & & & \bullet & \bullet & \bullet & \bullet & \\ 0 & & & & & \bullet & \bullet & \bullet & \end{bmatrix}$$

Thus $\underline{B}_1$ is a banded matrix of seven element wide. Also, $\underline{B}_1$ is a Toeplitz, almost cyclic and strictly diagonally dominant matrix with the properties mentioned previously. The matrix $\underline{B}_2$ has the same entries as $\underline{B}_1$ except they may have different size. Hence $\underline{B}_1$ and $\underline{B}_2$ are positive definite matrices [2].

Iterative methods have been formulated for solving the unknown coefficients $\underline{c}$ in eq. (2). The matrices $\underline{A}$ and $\underline{B}$ usually have very large dimensions in an image restoration problem, hence one is faced with

the problem of inverting a matrix of very large dimension. It is known that the number of multiplications in inverting matrices grows as a nonlinear function of the dimensions of the matrices being inverted. Therefore, for large dimension, matrix inversion is a non-trivial operation even with a modern digital computer. For this reason it is assumed that the point spread function is separable i.e.

$$h_{ij}(\zeta, \varkappa) = h_i(\zeta)\, h_j(\varkappa) \tag{3}$$

for all ζ, $\varkappa$ and i, j. Now matrix $\underline{A}$ becomes

$$\underline{A} = \underline{P}_x \otimes \underline{P}_y \tag{4}$$

where

$$\underline{P}_x = \left[\int_{-\infty}^{\infty} h_i(\zeta)\, S_k(\zeta)\, d\zeta\right] \qquad \underline{P}_y = \left[\int_{-\infty}^{\infty} h_j(\varkappa)\, S_\ell(\varkappa) d\varkappa\right] \tag{4}$$

Then the following iterative equation can be derived from eq. (2) by making use of eq. (4).

$$\underline{C}^{(i)} = \underline{C}^{(0)} - \gamma(\underline{A}_1 + \underline{B}_1)\, \underline{C}^{(i-1)}(\underline{\tilde{B}}_2 \underline{A}_2^{+}) \tag{5}$$

Where $\underline{C}$ is the matrix of the unknown coefficients $C_{k\ell}$, + designates the pseudoinverse [3], and

$$\underline{C}^{(0)} = \underline{P}_x^{+}\, \underline{G}\, \underline{P}_y^{+}$$

$$\underline{A}_1 = \underline{P}_x^{T} \underline{P}_x$$

$$\underline{A}_2 = \underline{P}_y^{T} \underline{P}_y$$

$$\underline{\tilde{B}}_2 = \underline{A}_2 \underline{B}_2 \underline{A}_2$$

Work is in progress to determine the unknown parameter γ ($o < \gamma < 1$) from the iteration procedure of eq. (5) and the constrainted iterative solution of $\underline{C}$ imposed by the positive restoration and equal energy constraints.

References

1. R. Bellman, Introduction to Matrix Analysis, P. 236 and P. 265, McGraw Hill, 2nd Ed. 1970.

2. R. S. Varga, Matrix Iterative Analysis, Prentice Hall, 1962.

3. A. Albert, Regression and the Moore-Penrose Pseudoinverse, Academic Press, 1972.

4.6 Nonuniform Sampling of Observation Space

Faramarz Davarian

The fundamental model describing the image restoration problem under the assumption of a space invariant blur function is given by

$$f_o(x) = \int_{-\infty}^{\infty} f_I(\alpha) g(x - \alpha) d\alpha \qquad (1)$$

Here, $f_I(\alpha)$ represents an ideal image line and $f_o(x)$ an observed image line. To restore $f_I(\alpha)$ numerically, the above integral equation may be discretized as follows

$$f_o(x_i) = \sum_{j=1}^{N} c_{ij} f_I(j) g(x_i - j) \qquad (2)$$

where $f_I(j)$ are uniformly spaced nodes of the quadrature formula, $f_o(x_i)$ are samples of the observation and c_{ij} are quadrature coefficients.

Equation (2) can be represented in vector form as

$$\underline{f}_o = \underline{B}\,\underline{f}_I \tag{2}$$

where

$$B_{ij} = c_{ij} g(x_i - j)$$

The problem of image restoration has now been reduced to a regression problem; given the observed vector $\underline{f}_o$ and the blur matrix $\underline{B}$, a suitable estimate of $\underline{f}_I$ must be found.

If in eq. (2) matrix $\underline{B}$ has full column rank ($M \geq N$), the model is called overdetermined. Under the assumption of full column rank the pseudoinverse of B is defined to be [1]

$$\underline{B}^+ = (\underline{B}^T \underline{B})^{-1} \underline{B}^T$$

and the estimate of $\underline{f}_I$ is given by

$$\underline{f}_I = \underline{B}^+ \underline{f}_o$$

Condition Number. The condition number of a system is a measure of the affect of an input perturbation (input noise) on the output of the system. Assume that the observed image function has been perturbed by observation noise by an amount of $\Delta \underline{f}_I$. The error in the estimate of $\underline{f}_I$ is bounded by [2]

$$\frac{\| \Delta \underline{f}_I \|}{\| \underline{f}_I \|} \leq C_B \frac{\| \Delta \underline{f}_o \|}{\| \underline{f}_o \|} \tag{3}$$

where C_B is the condition number of $\underline{B}$ and is given by

$$C_B = \| \underline{B} \| \cdot \| \underline{B}^+ \|$$

From inequality (3) the importance of the condition number becomes clear. A large condition number results in numerical unstability; a small observation error will cause a large error in the estimation of the ideal image. A linear system with relatively large condition number is often called ill conditioned.

Equally Distanced Samples of Observation. Assume that the original image is sampled at points $1, 2, \ldots, N$, and the sampling interval has unit length. Suppose the sampling period for the observed image is $\Delta x (\Delta x \leq 1)$ and the length of the degrading function is L; $g(t) = 0$, if $|t| > L/2$. The relationship between the number of observed samples M and the number of original samples N is then given by

$$M = \frac{N - L}{\Delta x} + 1$$

For a given N and L, M should be chosen to minimize the condition number for $\underline{B}$. Consider the set of blur matrices $\underline{B}_M$ for which M is sequentially set equal to $N, N + 1, N + 2, \ldots$, etc. and the corresponding condition number is C_{BM}. Figure 1 contains a typical plot of the condition number as a function of M. Usually the shape of such a condition number curve depends on the variance of the blur function. In general, the curve does not decrease monotonically. The curve assumes its peak when $M = N$, and as M grows the curve tends to decrease with some periodic upward jumps. The period is related to Δx and the fact that for some integer n, $n\Delta x = k$ (in some cases such n may not exist) where k is an integer. The period of the jumps is usually a function of n and N. For very large M, the curve tends to increase, but the amplitude of the jumps decrease.

One method to improve upon the condition number of $\underline{B}$ is non-equal spacing of the observed samples. It is clear that the middle points of the ideal image line contribute to more points of observation than the points which are closer to the line boundary. If the space limited degradation function exists only on an interval of length L, a typical pixel in the middle would contribute to all the points in the observation

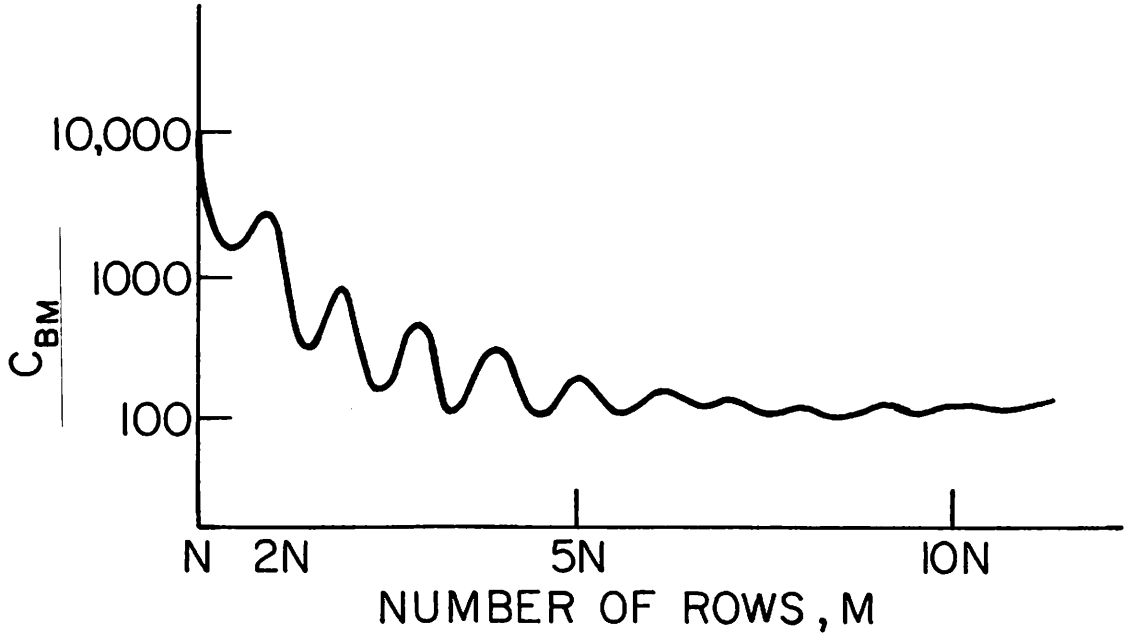


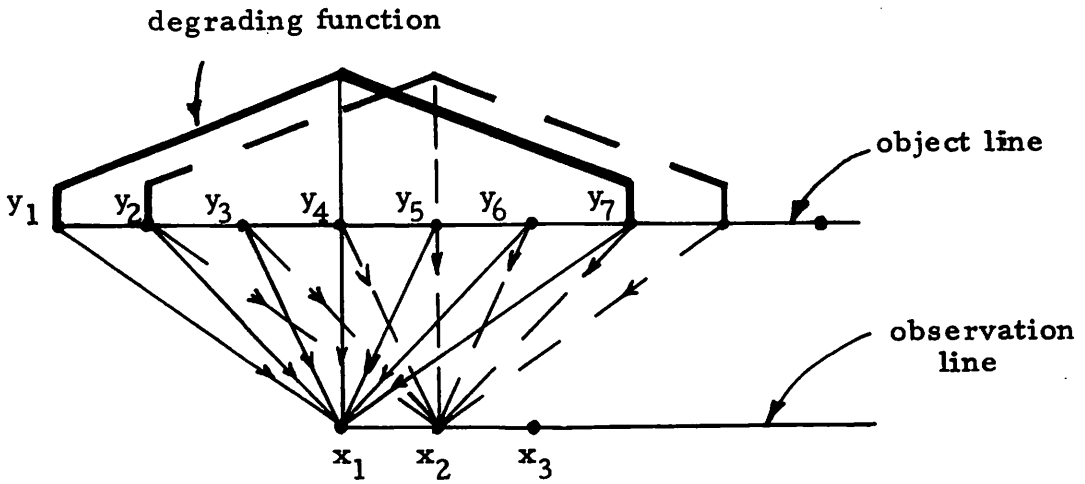
Figure 4.6-1. Typical shape of a condition number curve.

image line which lie in an interval of size L. On the other hand the two samples at the very ends of $\underline{f}_I$ only contribute to two samples at the end points of $\underline{f}_o$. This simple argument suggests that non-uniform sampling of observed image would bring out all the information needed to easily reconstruct the original image.

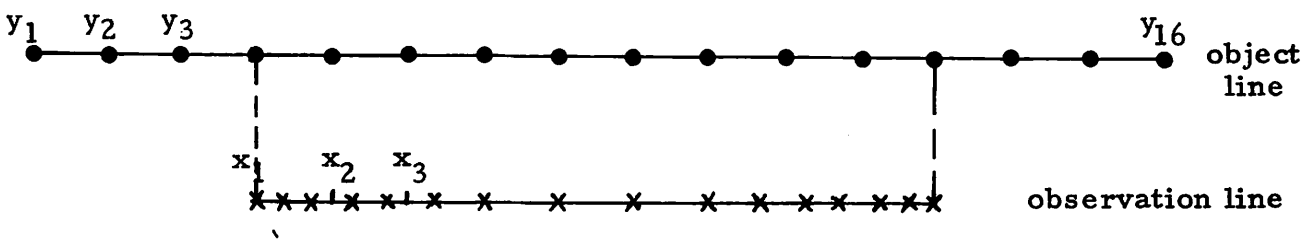
The ill-conditioning associated with uniform sampling of a blurred image have been examined experimentally. In the experimental model a Gaussian blur function has been employed with 16 equally spaced ideal image samples on an interval of length 15,. The variance of the Gaussian blur function is unity and its space limited length is 6 pixels.

For 16 equally spaced samples of the observation, the condition number is about 300,000. When the number of observation samples increases, the condition number tends to decrease. For values of M above 50 the condition number stays under 2,000 , but it never becomes much less. As mentioned before $\underline{C}_B$ is not a monotone function, i.e., for some values of M it may increase by increasing M. For a suitable non-uniform sampling of the observation, the condition number becomes about 2,000. The different methods to find the locations of the non-uniform samples are discussed in the next section. Note that there is an advantage in non-uniform sampling with M = N as compared to uniform sampling with $M \geq N$, since for the first case it is necessary only to invert matrix $\underline{B}$ itself whereas for the second case, $\underline{B}^T\underline{B}$ must be inverted.

Spacing of Non-uniform Samples. The two end points of the observation line play an important role in the sampling process; the beginning and the end points of the observed line are essential to enable the system to reconstruct the object. To simplify the explanation, consider the left half of the observation only. Starting from the beginning of the left half line, call the unit distance points of the observation $x_1, x_2, \ldots$ etc, as shown in figure 2a. In the same manner, the points of the object line are denoted as $y_1, y_2, \ldots$ etc. The problem then is to estimate points $y_1, y_2, \ldots$ etc; using N samples of the observed intervals.



(a) uniform sampling



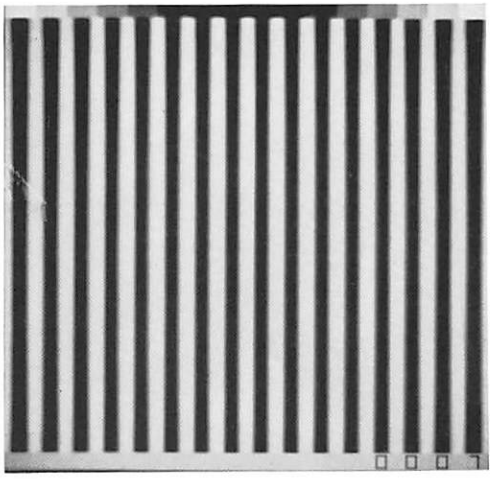
(b) nonuniform sampling

Figure 4.6-2. Relationship between object and observation line samples.

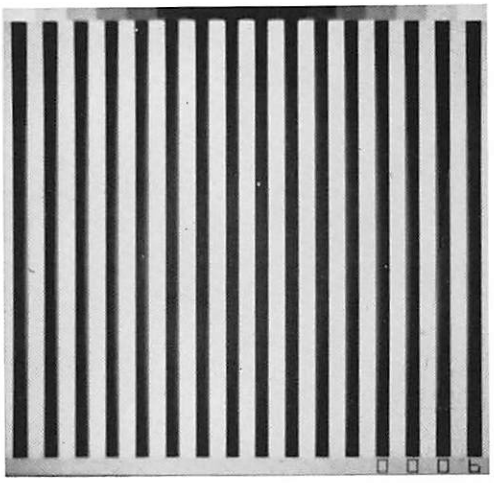
Interval $[x_1, x_2)$ is the only one containing information about $[y_1, y_2)$. The interval $[x_1, x_2)$, also, contains information about the larger interval $[y_1, y_3)$. Now, consider the interval $[y_2, y_3)$. There are only two intervals of the observed line which can contribute to the restoration of $[y_2, y_3)$; $[x_1, x_2)$ and $[x_2, x_3)$. In the interval $[y_3, y_4)$, only $[x_1, x_2)$, $[x_2, x_3)$ and $[x_3, x_4)$ of the observation contain information about $[y_3, y_4)$. This simple argument can be generalized to a typical unit size interval $[y_j, y_{j+1})$. There are at most L - 1 of the (L is the length of the degrading function) observation intervals which could contribute to reconstruction of $[y_j, y_{j+1})$. The number of intervals relating to $[y_j, y_{j+1})$ equals j if $j < L$, and equals L - 1 if $j \geq L - 1$ (assume y_j is on the left half line). Suppose for every interval of the ideal line it is necessary to have k samples of the observed line. For the first interval $[y_1, y_2)$, all these k samples must lie on $[x_1, x_2)$. For the second interval $[y_2, y_3)$, half must lie on $[x_1, x_3)$, and half on $[x_2, x_3)$. Likewise, one can consider the sample distribution for the other intervals.

Figure 2b demonstrates the method. In the table, the number of samples which are needed to reconstruct the corresponding ideal line interval is listed before each observation interval. Here, 16 samples of the object line are uniformly placed on an interval of size 15, and the size of the degrading function is 6. The length of the observation line is, therefore, 15 - 6 = 9. Three samples are placed on the first interval, two on the second, and one is placed on the third. The same distribution is considered for the last three intervals. The middle of the line is sampled exactly as the object line.

Figure 3 contains a blurred picture and its restoration using the technique described in this report. The restoration is exact since it has been performed in the absence of noise.



Observed Image



Deblurred using nonuniformly spaced samples.

Figure 4.5-3. Image reconstruction using nonuniformly spaced samples.

References

1. F. Grayhill, Introduction to Matrices with Applications in Statistics, Wadsworth, Belmont, California, 1969.

2. E. K. Blum, Numerical Analysis and Computation: Theory and Practice, Wesley, Reading, Mass., 1972.

4.7 Histogram Exponentiation

Francis Kretz

In the field of image enhancement simple nonlinear amplitude transformations are quite useful. For example, the technique of histogram equalization [1-3] has been shown to significantly improve the detail of low contrast images such as X-ray and Earth Resources Satellite (ERTS) pictures.

Examination of detailed and well-contrasted pictures reveals that their histograms are approximately exponential with a black level peak. On the other hand poor quality images usually have a non-exponentially shaped histogram. For this reason an investigation was made into methods of "histogram exponentiation" for image enhancement.

Probability Density Function Transformations. Let X denote a random variable whose range is $[0,1]$; $p(x)$, $x\in[0,1]$ be its probability density; and $F_x(x)$ represent the distribution function assumed continuous. Also, let $f(\cdot)$ be a continuous monotonically increasing function and let $Y = f(X)$ be an output random variable with distribution given by

$$F_y(y) \equiv P_r\{Y \leq y\} = P_r\{f(x) \leq y\}$$

Since $f(\cdot)$ is invertible

$$F_y(y) = P_r\{X \leq f^{-1}(y)\} \equiv F_x\{f^{-1}(y)\}$$

Thus, for a given input distribution $F_x(x)$ and a desired output distribution $F_y(y)$, it is possible to determine the necessary transfer function f(X). In the discrete case the transfer function can only be determined approximately. If the transfer function is chosen such that $F_y(y)$ is uniform the process is called histogram equalization.

Histogram Exponentiation. In the histogram exponentiation process the transfer function is selected so that the output probability distribution is of the form

$$P_r(k) = A \exp\{-a(k-1)\}$$

for k = 1, 2, . . . , M. The parameter "a" controls the shape of the exponential function. With a=0, the result is histogram equalization.

Figure 1 shows the effect of histogram equalization and exponentiation on an ERTS picture. The equalized picture in figure 3b is subjectively improved compared to the original. And the exponentiated pictures exhibits further improvement.

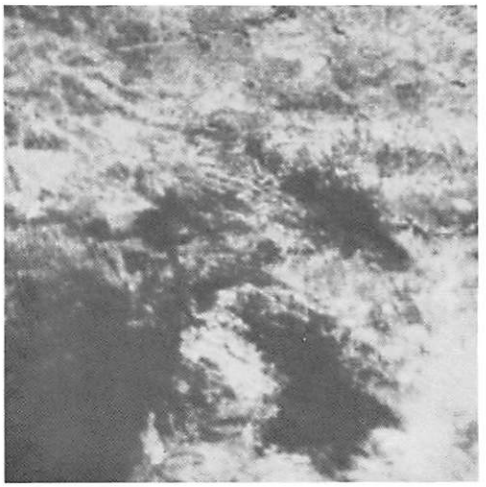
References

1. E. L. Hall et al., "A Survey of Preprocessing and Feature Extraction Techniques for Radiographic Images," IEEE Trans. Comp., Vol. c-20, pp. 1032-1044, September 1971.

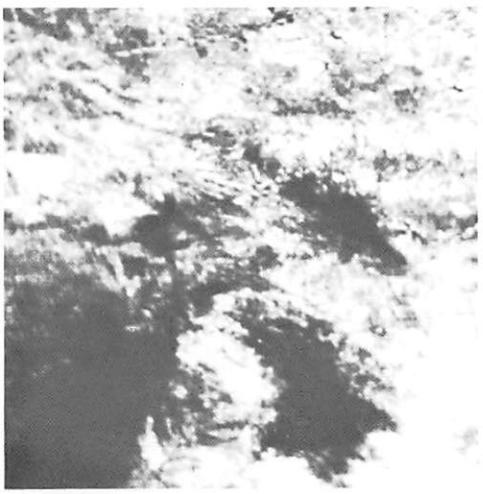
2. H. C. Andrews, A. G. Tescher, and R. P. Kruger, "Image Processing by Digital Computer," IEEE Spectrum, Vol. 9, pp. 20-32, July 1972

3. A. Rosenfeld and E. Troy, "Visual Texture Analysis," University of Maryland, Tech. Rep. 70-116, June 1970.

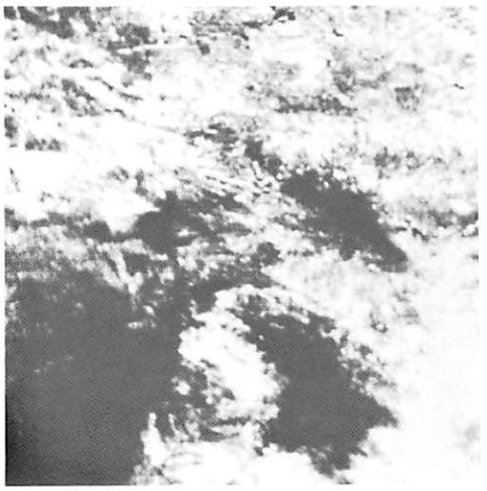
(a) Original



(b) Equalization
(M = 31.5)



(c) Exponentiation
(M = 19)



(d) Exponentiation
(M = 9.4)

Figure 4.7-1. Earth resource picture

5. Image Data Extraction Projects

Image data extraction describes the collection of projects concerned with the detection of features within an image and methods of measuring these features.

The first report describes an investigation into a method for image reconstruction from transverse-axial density projections of a solid object. The method utilizes a Fourier transform process defined on a polar raster which obviates the need for interpolation in the transform domain.

A new project described in the following report is based upon the development of nonlinear optical processing elements. These elements are constructed from halftone transparencies which are mathematically computed and recorded by a scanning microdensitometer.

5.1 Fourier-Bessel Method for Transverse-Axial Reconstruction

Dennis G. McCaughey and Richard P. Kruger

Transverse Axial Reconstruction implies the reconstruction of two dimensional cross sectional regions of an object or signal from knowledge of a discrete number of one dimensional projections. Application areas include analysis of electron microscope imagery [1], medical transverse axial tomography [2], and radar signal analysis [3]. Algorithms for this purpose may be divided into three general categories : Algebraic [4], Convolutional [2], and Fourier Transform domain [1] methods. The present discussion will be limited to the latter method.

Transform Processing Techniques. The Fourier method of reconstruction depends on the fact that the Fourier transform of the projection is identical to the corresponding central section of the Fourier transform of the density function [5]. When the Fourier transform operation is performed by a digital computer, the transform of the original density is obtained at discrete points in frequency space in polar coordinates. Since few of these points will correspond to rectangular coordinates required for the inverse transform operation, interpolation is required. Also, since the forward or inverse discrete Fourier transform operation produces the original function at this sample point to within the truncation errors of the machine used,* it would seem reasonable that some, but certainly not all, of the artifacts present in the

---

* It can be shown that the forward-inverse sequence of Fourier transform operations produce the original function exactly at the sampling points without consideration of the sampling theorem. However, this is not to imply that the sampling theorem is without importance.

reconstructed image are due to the particular interpolation process used. Other artifacts in the reconstructed cross-sectional image result from Gibbs phenomenon, high frequency components (due to edge effects) present in the estimated Fourier coefficients, and possible undersampling. While much has been said with regard to the sampling theorem concerning image degradation and resolution [18,19] more effort is necessary to determine the relative importance of the various interpolation methods in the area of image degradation. Notwithstanding the lack of knowledge concerning the relative importance of interpolation as an important contribution to image degradation, it would seem desirable to employ an algorithm that utilize the Fourier coefficients directly on polar coordinates. Crowther De Rosier and Klug proposed such an algorithm using the Fourier-Bessel transform [21]. Crowther et. al. [16] have employed this algorithm in the reconstruction of images obtained in electron microscopy with reasonable results. However, their format was such that projections were not obtained at evenly spaced angles which necessitated an interpolation process to fill the polar raster. In developing an algorithm it is reasonable to assume that projections should be available at evenly spaced angles to avoid this interpolation.

As a basis for the analysis let $F(R,\theta)$ denote the Fourier transform of the projection at angle θ and $f(r,\varphi)$ be the tomographic section in polar coordinates. The following then results [8]

$$f(r,\varphi) = \int_0^{2\pi} \int_0^{\infty} F(R,\theta)\exp(-2\pi j rR\cos(\theta-\varphi))RdRd\theta \qquad (1)$$

Note that

$$\cos(\theta-\varphi) = \sin(\theta-\varphi+\pi/2).$$

and

$$\exp(ja\sin x) = \sum_{k=-\infty}^{\infty} J_k(a)\exp(jkx)$$

where $J_k(a)$ is the k^{th} order bessel function of the first kind. By inserting these two relationships into eq. (1) and rearranging the order of summation

and integration it is possible to obtain

$$f(r,\varphi) = \sum_{k=-\infty}^{\infty} j^k \int_0^{2\pi} \int_0^{\infty} F(R,\theta) J_k(2\pi R r) \exp(jk\theta) R\,dR\,d\theta \exp(-jk\varphi) \qquad (2)$$

If $f_k(r)$ is defined as

$$f_k(r) = j^k \int_0^{2\pi} \int_0^{\infty} F(R,\theta) J_k(2\pi R r) \exp(jk\theta) R\,dR\,d\theta \qquad (3)$$

it is possible to obtain the Fourier series expression

$$f(r,\varphi) = \sum_{k=-\infty}^{\infty} f_k(r) \exp(-jk\varphi) \qquad (4)$$

Note that eq. (4) is indeed a Fourier series for

$$f_k(r) = \frac{1}{2\pi} \int_0^{2\pi} f(r,\varphi) \exp(jk\varphi)\,d\varphi \qquad (5)$$

If constraints to discrete angles evenly spaced over $(0, 2\pi)$ are assumed, eq. (4) becomes a discrete inverse Fourier transform. Knowledge of $f_k(r)$ will then permit an exact reconstruction of $f(r,\varphi)$ along radii equally spaced over $(0, 2\pi)$. The factor $F(R,\theta)$ is in general not available for all R and θ since projections are available for only a discrete set of angles and because a DFT on each projection produces the transform only at discrete frequencies. Therefore $F_s(r,\theta)$ is the 2-D sampled form of the Fourier transform $F(R,\theta)$ given by

$$F_s(R,\theta) = \frac{2\pi\Delta R}{N} \sum_{i=0}^{I-1} \sum_{i=0}^{N-1} F(i\Delta R, \frac{2\pi n}{N})\, \delta(R - i\Delta R, \theta - \frac{2\pi n}{N}) \qquad (6)$$

where N denotes the number of projections and I represents the number of points in each projection. Inserting eq. (6) into eq. (3) one obtains an estimate of $f_k(r)$ termed $\hat{f}_k(r)$ in the form of a Riemann sum

$$\hat{f}_k(r) = \frac{2\pi\Delta R}{N} \sum_{i=0}^{I-1} \sum_{n=0}^{N-1} F(R_i, \frac{2\pi n}{N}) j^k J_K(2\pi R_i r) R_i \exp(jk2\pi n/N) \qquad (7)$$

where $i\,\Delta R$ has been replaced by R_i. The estimated reconstruction is then

$$f(r,\varphi) = \sum_{k=-\infty}^{\infty} \hat{f}_k(r) \exp(-jk\varphi) \qquad (8)$$

Since $f(r,\varphi)$ is a spaced bounded continuous function, $F(R,\theta)$ is an analytic function [2] and thus eq. (7) always exists. Equation (7) is in reality rectangular integration and if ΔR and $2\pi/N$ are small, $\hat{f}_k(r)$ is an accurate estimate of $f_k(r)$. The ability to estimate the tomographic section $f(r,\varphi)$ with $\hat{f}(r,\varphi)$ is thus limited only in the accuracy of the $\hat{f}_k(r)$. It can be easily shown that eq. (8) can be expressed as

$$\hat{f}(r,\varphi) = \sum_{k=0}^{\infty} j^k \frac{2\pi\Delta R}{\epsilon_k N} \sum_{i=0}^{I-1} \sum_{n=0}^{N-1} F(R_i, \frac{2\pi n}{N})\, J_k(2\pi R_i r) R_i \exp[jk2\pi n/N] \exp(-jk\varphi)$$

$$+ \sum_{k=0}^{\infty} -1^{k} j^{-k} \frac{2\pi\Delta R}{\epsilon_k N} \sum_{i=0}^{I-1} \sum_{i=0}^{N-1} F(R_i, \frac{2\pi n}{N}) J_k(2\pi R_i r) R_i \exp[-jk2\pi n/N] \exp(jk\varphi) \qquad (9)$$

where

$$\epsilon_k = \begin{cases} 2 & k = 0 \\ 1 & k \neq 0 \end{cases}$$

By similar means it can be also shown that

$$f(r,\varphi)= 2Re\sum^{\infty}\frac{2\pi\Delta R}{\epsilon_k N}\sum_{i=0}^{I-1}\sum_{n=0}^{N-1}F(R_i,\frac{2\pi n}{N})\,J_k(2\pi R_i r)R_i\exp[jk2\pi n/N]\exp-[jk(\varphi-\pi/2)] \tag{10}$$

where j^k is replaced by $\exp[jk(\pi/2)]$. Equation 10 can be computed by two fast Fourier transform operations followed by a Bessel function weighting operation and a final fast Fourier transform operation. Current results indicate that test images can be reconstructed from 32 projections with 64 sample points in each projection in a few seconds on a standard computer. An example of a reconstructed disk is shown in figure 1.

Removing the interpolation process results in not only a significant reduction in computing time but some insight into a quantitative measurement of the accuracy of the algorithm through eq. (7). Furthermore, retaining a polar format produces an algorithm with the highest resolution near the axis of rotation. Upon first consideration this may seem a disadvantage, however, a higher resolution may be obtained in a region on interest simply by centering the axis of scan rotation in that region. This may result in the need for fewer projections.

This investigation would seem to indicate that the Fourier-Bessel method has several advantages over the more conventional transform method. It is still necessary to determine the quality of the estimate $\hat{f}_k(r)$ and to develop methods to improve this estimate. It was noted earlier that eq. (7) was in reality a rectangular integration formula. A simple method to improve the accuracy of $\hat{f}_k(r)$ would be to utilize a trapezoidal formula. This would require no increase in the number or sample points. Also if the projections are sampled at the Nyquist rate (implicit in any method), it can be shown that the function can exhibit no more than one zero crossing between the two adjacent sample points. This would imply that a trapezoidal formula would be reasonably accurate. The algorithm also may be extended to images of much higher resolution since fast Fourier transform algorithms are certainly applicable to higher dimensionality than currently used. The main effort would be directed towards calculating the Bessel functions of higher order. No particular difficulty is anticipated, since sufficiently accurate large argument approximations are available [9] along with a recursive algorithm for small arguments [10].

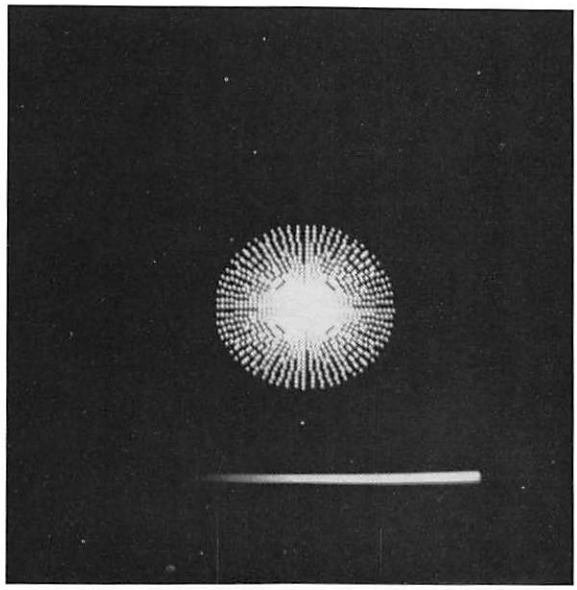


Figure 5.1-1. Disk reconstruction from 32 projections.

References

1. R. A. Crowther, D. J. DeRosier and A. Klug, "The Reconstruction of a Three Dimensional Structure from Projections and Its Application to Electron-Microscopy, "Proceedings Royal Society London, A, 317, 319-340 (1970).

2. L. A. Shepp and B. F. Logan, "Reconstructing Interior Head Tissue from X-Ray Transmissions," Special issue on Physical and Computational Aspects of 3-Dimensional Image Reconstruction, IEEE Trans. on Nuclear Science, NS-21 (3), June 1974.

3. R. N. Bracewell and A. C. Riddle, "Inversion of Fan-Beams in Radio Astronomy," Biophysical Journal 150, November 1967.

4. R. Gordon, R. Bender and G. T. Herman, "Algebraic Reconstruction Techniques for Three Dimensional Electron Microscopy and X-Ray Photography," Journal of Theor. Biology, 29, 471-481, 1970.

5. A. Papoulis, Systems and Transforms with Applications in Optics, New York, McGraw-Hill Book Company, pp. 92-93, 1968.

6. R. A. Crowther, et al, Nature 226; 421, (1970).

7. P. R. Smith, T. M. Peters and R. H. T. Bates, J. Physics A: Math. 6, March 1973.

8. G. N. Ramachandran, A. V. Lakshminarayanan, Proc. Nat. Acad. Sci. 08 (9), 2236, 1971.

9. Korn and Korn, "Mathematical Handbook for Scientists and Engineers", SEC 21-8-9.

10. M. Abramowitz and I. Stegun, Handbook of Mathematical Functions, Dover 6-385.

5.2 Nonlinear Optical Image Processing With Halftone Screens

Stephen R. Dashiell and Alexander A. Sawchuk

Under ordinary circumstances optical image processing systems are capable of performing only linear operations on input images. There is a large class of nonlinear operations which could be of great value if they could be easily performed in an optical system. Among these are homomorphic filtering [1], histogram equalization, level slicing, high intensity pass, low intensity pass, intensity band stop, and others [2].

Kato and Goodman [3] have successfully utilized commercially

available halftone screens to perform logarithmic and exponential transfer functions in an optical system, thereby allowing them to filter a multiplicative noise component with much greater success than if ordinary linear filtering had been performed.

If diffraction orders of the halftoned image higher than the on-axis zero-order diffraction component are used, and if specially made halftone screens are available, nonlinear transfer functions considerably more complex than the monotonic logarithm and exponential can be obtained.

To understand how a non-monotonic transfer function might be obtainable, consider first the Fourier transform of a rectangular array of opaque squares of side b and center to center spacing a, given by

$$F\{\underline{t}\,(x,y)\} = \delta(f_x, f_y) - \frac{a^2}{\pi^2} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \frac{1}{mn} \delta(f_x - \frac{n}{a}) \delta(f_y - \frac{n}{a}) \sin\left(\frac{\pi bn}{a}\right) \sin\left(\frac{\pi bm}{a}\right) \tag{1}$$

where m and n are numbers identifying the diffraction order impulses in the spatial frequency plane.

The transmittance function assumed by eq. (1) is typical of what would be obtained if a constant transmittance were halftoned and photocopied on high contrast copy film. The presence of the sine terms indicates that non-monotonic behavior could be expected. This becomes more evident in the special case where a (0, 1) order specified by m = 1, n = 0 is selected. In this case eq. (1) reduces to

$$F\{t(x,y)\}_{(0,1)} = -\frac{a}{\pi} \sin\left(\frac{\pi b}{a}\right) \delta\,(f_x, f_y) \tag{2}$$

Inverse Fourier transforming this expression and squaring to get the output intensity yields

$$I_{OUT} = \frac{a^2}{\pi^2} \sin^2\left(\frac{\pi b}{a}\right) \tag{3}$$

as the final intensity output.

This result is useful because the value of b is a function of the input intensity to the system, or equivalently, is a function of the density of an input transparency. In operation, the Fourier transforming and inverse transforming are instantaneously performed in a coherent optical parallel processor [5], and selection of the diffraction orders is accomplished by simple spatial filters. These filters are chosen to pass the low spatial frequency information in the original, thus desampling the halftoned picture while the nonlinear operation is performed.

To illustrate how an operation such as level slicing could be performed using this first diffraction order, a halftone screen was made by photographing crossed Ronchi rulings. The resulting screen had transparent squares on a partially transmitting background. The sides of the squares were one-half the distance between squares. The subject transparency was then photographed through this halftone screen, and three different regions were then present on the copy film after development. In the first region, where the subject was sufficiently dense that the copy film did not expose even through the clear squares, there were no dots or $b = 0$. In the second, where the subject was less dense such that the copy film did expose through the clear squares, but not through the darker background, the dot parameters were $b = a/2$. Third, where the subject was still less dense such that the copy film exposed through the clear squares and through the background, the parameters were $b = a$.

The cases $b = 0$ and $b = a$ give $I_{OUT} = 0$, however $b = a/2$ yields $I_{OUT} = a^2/\pi^2 \neq 0$. Thus, only the range of densities which gave dots appear in the output and densities above or below this range do not appear. This has been experimentally verified by slicing a picture into 6 different levels, using a single halftone screen but varying the exposure time on the copy film. The position of the sliced level is controllable by varying exposure while the width of the level is fixed by the density difference on the screen between the clear squares and the darker background. This technique for obtaining nonlinear transfer functions is not limited to level-slicing and other simple functions. Using digitally produced halftone screens with appropriately selected density profiles, more complex operations are possible.

Although these preliminary experiments in nonlinear processing involve photographic processing and coherent optical techniques, they are especially well suited to applications in hybrid digital/optical systems [4]. Several real-time optical input devices with adjustable parameters and

thresholding are now under development, and combing these with the halftoning operation would make possible a real-time nonlinear optical parallel processor. Such a system would be relatively inexpensive and simple, and would avoid problems of scanning and display. Electrically controllable selection of diffraction orders in the Fourier plane of the system would permit fast modification of the transfer functions. Immediate future work will involve the production of several types of special halftone screens in the IPI plotting microdensitometer.

References

1. A. V. Oppenheim, R. W. Schaefer and T. G. Stockham, Jr., "Nonlinear Filtering of Multiplied and Convolved Signals," Proc. IEEE, Vol. 56, No. 8, pp. 1264-1291, August, 1968.

2. H. C. Andrews, A. G. Tescher and R. P. Kruger, "Image Processing by Digital Computer," IEEE Spectrum, Vol. 9, No. 7, pp. 20-32, July, 1972.

3. H. Kato and J. W. Goodman, "Nonlinear Transformations and Logarithmic Filtering in Coherent Optical Systems," Opt. Comm., Vol. 8, No. 4, pp. 378-381, August, 1973.

4. A. A. Sawchuk, Guest Editor, special issue on Optical and Digital Information Processing, Optical Engineering, Vol. 13, No. 3, May-June, 1974.

5. J. W. Goodman, Intro. to Fourier Optics, McGraw-Hill, New York, 1968.

6. Image Analysis Projects

The image analysis projects are concerned with the background technology necessary to effectively design image coding, restoration, enhancement, and data extraction systems. Of particular interest are models of the human visual system for monochrome and color images, and the development of quantitative measures of image fidelity and intelligibility.

In the first project the continued research on a model of human color vision is described. This model is nearly completely developed and tested, with success. The first applications of the model are now introduced.

The next report considers the development of a more accurate model for human monochrome vision. This model incorporates a linear filtering element before the photoreceptors to account for the optical degradations of the eye.

6.1 A Quantitative Model of Color Vision

Werner Frei

The first part of this report summarizes recent results of an effort to describe some major properties of human color vision with a simple neuro-physiological model. The second part investigates an application of the model to the optimal quantization of color image signals.

Neuro-physiological model. The visual model in question is the result of a broad study of the neuro-anatomy of the human eye and psychophysics of vision [1, 2]. Figure 1 shows a block diagram of the model which consists of a layer of photoreceptors connected to a layer of summation cells. The photoreceptors comprise three distinct classes corresponding to Judd's fundamental sensations, with a quasi-logarithmic response to light stimulation. Three distinct patterns of connection are assumed between the two layers, such that the output of each summation cell corresponds to one of three opponent sensations "dark-light" "red-green" or "blue-yellow."

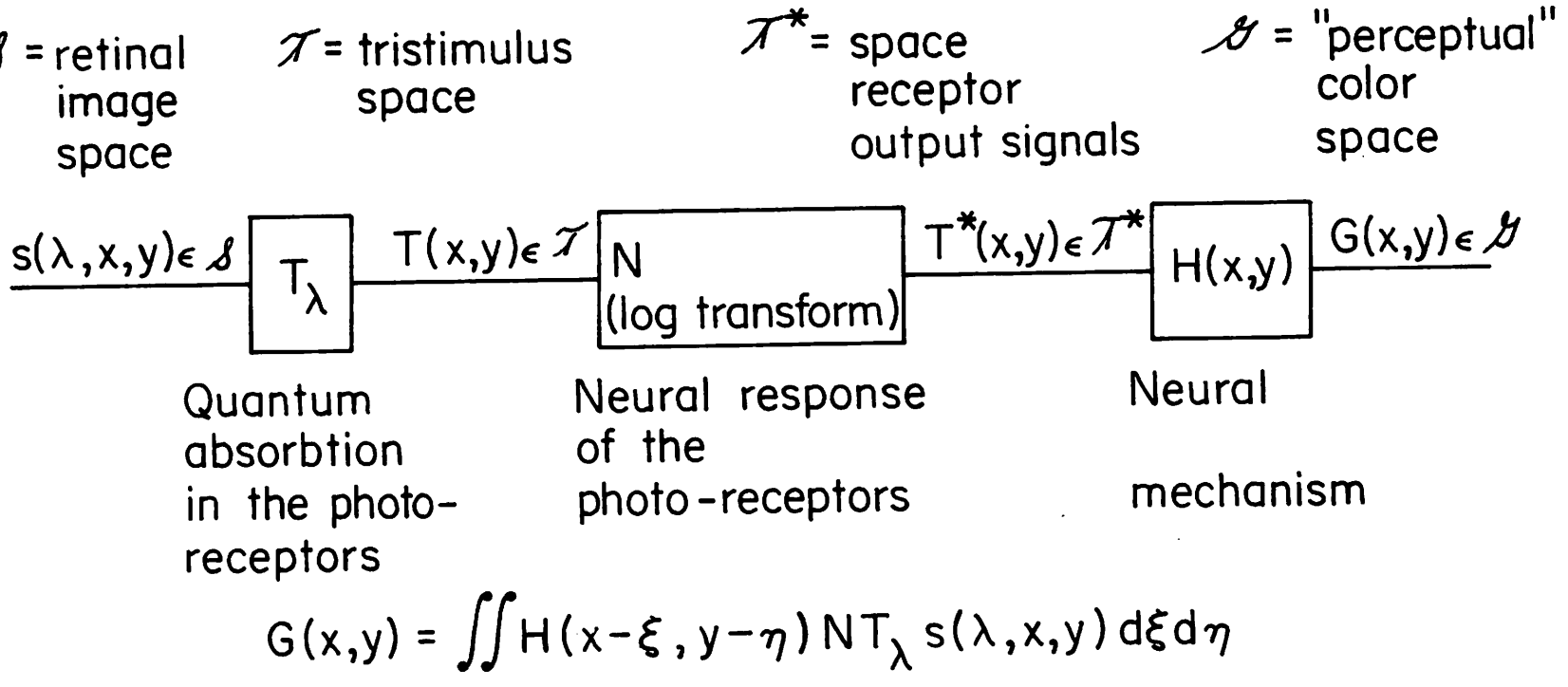


Figure 6.1-1. Neuro-physiological Model.

At high intensities and moderate saturations, the receptor response functions may be simplified to

$$t_i^* = c_i \ln(t_i) \qquad t_i > 0 \tag{5}$$

for $t_i > 0$. The spaces of color tristimulus $\mathcal{T}$ and perceptual quantities $\mathcal{G}$, respectively, are defined to be vector spaces with algebraic laws of composition defined as follows

$$T_a \perp\!\!\!\perp T_B = (t_{1A}t_{1B}, t_{2A}t_{1B}, t_{3A}t_{3B})^T \tag{6a}$$

$$r \;\underline{||}\; T = (t_1^r, t_2^r, t_3^r)^T \tag{6b}$$

$$G_A \perp\!\!\!\perp G_B = (g_{1A}+g_{1B}, g_{2A}+g_{2B}, g_{3A}+g_{3B})^T \tag{6c}$$

$$r \;\underline{||}\; G = (rg_1, rg_2, rg_3)^T \tag{6d}$$

where T_A, T_B, T are the tristimulus vectors of arbitrary colors and G_A, G_B, G are the corresponding neural quantities. The symbol $\perp\!\!\!\perp$ denotes vector addition and $\;\underline{||}\;$ multiplication by scalars r (real numbers). The nonlinear transformation defined by the model can now be easily verified to be a linear vector mapping of the above vector space $\mathcal{T}$ into $\mathcal{G}$ by

$$H \circledast N(T_A \perp\!\!\!\perp T_B) = H \circledast NT_A + H \circledast NT_B \tag{7a}$$

$$H \circledast N(r \;\underline{||}\; T) = rH \circledast NT. \tag{7b}$$

This defines a generalized law of superposition of colored sensations. For example, multiplying ($\;\underline{||}\;$) the tristimulus values of a color by a scalar r as defined by eq. (6b) yields a color appearing approximately r-times as bright, saturated, etc.

The analysis of the visual response to an arbitrary two-dimen-

$$t_i = \int_{\lambda_\ell}^{\lambda_u} S(\lambda)\, \bar{t}_{i\lambda}\, d\lambda \tag{2c}$$

where $\bar{t}_{i\lambda}$ for i=1, 2, 3 are independent linear transformations of the CIE $\bar{x}_\lambda$, y_λ, $\bar{z}_\lambda$ color matching functions; λ_ℓ and λ_u are the limits of the visible spectrum.

The mapping N describes the neural response of the receptors, assumed to be a quasi-logarithmic function of the respective light energies absorbed

$$t_i^* = c_i \ln (k_i t_i + t_{i0}) \tag{3}$$

where c_i, k_i, t_{i0} are constants. Finally the transformation H(x, y) represents linear weighted summations of the receptor outputs in the lateral inhibition process.

For simplicity three distinct shift invariant patterns of summations are assumed, which produce opponent neural signals $g_1(x, y)$, $g_2(x, y)$ $g_3(x, y)$ corresponding to "dark-light," "red-green" and "yellow-blue" perceptual quantities

$$G(x, y) = [g_1(x, y), g_2(x, y) g_3(x, y)]^T \tag{4a}$$

$$g_1(x, y) = \int_{-\infty}^{+\infty}\int h_1(x-\xi, y-\eta) c_1 \ln(t_1(\xi, y) + t_{10}) d\xi\, \eta \tag{4b}$$

$$g_2(x, y) = \int_{-\infty}^{+\infty}\int h_2(x-\xi, y-\eta) c_2 \ln\left(\frac{t_2(\xi,\eta)+t_{20}}{t_1(\xi,\eta)+t_{10}}\right) d\xi d\eta \tag{4c}$$

$$g_3(x, y) = \int_{-\infty}^{+\infty}\int h_3(x-\xi, y-\eta) c_3 \ln\left(\frac{t_3(\xi,\eta)+t_{30}}{t_1(\xi,\eta)+t_{10}}\right) d\xi d\eta \tag{4d}$$

The three opponent sensations form the basis of a "perceptual" euclidian color space. In that space "brightness" is equal to the norm of the vector sum of the three opponent signals.

The pattern of connections between the two layers also predicts spatial contrast phenomena. For intensities and color saturations of practical interest, the spaces of color tristimulus values and neural signals, respectively, are defined as algebraic vector spaces with distinct laws of composition and a linear vector mapping from the first into the second. From this, a generalized law of superposition of colored sensations is defined which enables fast computational procedures to predict the perception of color in complex visual fields. The mathematical structure of the model is now briefly outlined.

The model defines a mapping of the retinal image space $\mathcal{S}$ into a space of opponent neural signals $\mathcal{G}$, with $S(\lambda, x, y) \in \mathcal{S}$ designating the spectral energy distribution of light striking the retina and $G(x, y) \in \mathcal{G}$ a vector valued "perceptual" quantity; x, y are geometric coordinates in the retinal image plane. The mapping consists of a cascade of a linear (T_λ), a nonlinear (N) and a second linear (H(x, y)) transformation as given by

$$G(x, y) = H(x, y) \circledast N\ T_\lambda\ S(\lambda, x, y) \tag{1}$$

where $\circledast$ denotes the convolution integral. Whenever possible, the geometric coordinates (x, y) will be omitted to simplify the notation. The linear transformation T_λ maps the spectral energy distribution of light $S(\lambda)$ into a colorimetric tristimulus space $\mathcal{T}$. The factor T_λ describes the quantum absorption mechanism in three types of photopic receptors and is defined for the standard observer as

$$T_\lambda : S(\lambda, x, y) \rightarrow T(x, y) \tag{2a}$$

$$T = (t_1, t_2, t_3)^T \tag{2b}$$

sional field can now be cast, in a first approximation, into the framework of linear systems analysis. In particular, spatial contrast effects and resolution are predictable by a linear two-dimensional vector filtering $\mathcal{H}(ju, jv)$ in $\mathcal{L}$- space

$$G(x, y) = \mathcal{J}^{-1}\{\mathcal{H}(ju, jv)\ \mathcal{J}[\,CNT_{\lambda}S(\lambda, x, y)]\} \tag{8}$$

where $\mathcal{J}$ and $\mathcal{J}^{-1}$ denote the two-dimensional Fourier transformation and its inverse respectively, (u, v) are spatial frequencies in the (x,y) direction and $\mathcal{H}(ju, jv)$ is the Fourier transformation of the impulse response matrix H(x, y). The model has been previously shown [2] to be in agreement with the following psychovisual factors: a) metameric color matching, b) additivity of luminances, c) adaptation and simultaneous contrast. An additional hypothesis allows to predict the color discrimination ability of the visual system.

Quantitative Color Discrimination. The statistical errors of color matching experiments can be approximately predicted by assuming that the neural signals g_1, g_2, g_3 are corrupted by independent, additive gaussian random fluctuations.

Extensive studies of color matching errors [3] have shown that the loci of the experimental standard deviations around fixed reference colors are ellipsoids in the CIE x, y, ℓ-space, with ℓ a logarithmic function of the luminance Y ; $\ell = 0.2 \log_{10}(Y)$.

Let the probability of confusing a color $\hat{T} = T_{\lambda}\hat{S}(\lambda)$ with some reference color $T = T_{\lambda}S(\lambda)$ be equal to

$$p(G \mid \hat{T}) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2}\ [G-\hat{G})^{T}(G-\hat{G})]\} \tag{9}$$

with $G = CNT$, $\hat{G} = CN\hat{T}$.

Under this hypothesis, the experimental ellipsoids should correspond to unit diameter spheres in $\mathcal{L}$- space. Since the model has a separate luminance channel, luminance and chromaticity errors can be considered separately. (Brown and MacAdam's data [3] suggests generally small correlations of luminance and chromaticity errors.)

Using the simplified expressions for g_2 and g_3 at high radiances, and letting $k_1 = k_2 = k_3 = 1$ (see discussion on adaption below), the semi-axes of MacAdam's 25 ellipses were mapped into $\mathcal{L}$- space, adjusting the constants c_2 and c_3 to obtain a best fit with unit diameter circles in a plane g_1 = constant. Table 1 shows the lengths of the mapped semi-axes in $\mathcal{L}$- space for $c_2 = 123.2$ and $c_3 = 18.8$. Figure 2b shows the result of the inverse operation: unit diameter circles in $\mathcal{L}$- space, g_1 = constant, mapped into the CIE x-y chromaticity diagram. The constant $c_1 = 67.4$ is determined from the average standard deviations of luminances reported in the same study [3]

$$c_1^2 \left(\frac{dg_1}{dt_1} \Delta t_1\right)^2 = 60.10^4 \left(\frac{d\ell}{dY} \Delta Y\right)^2 \approx 60.10^4 \Delta\ell^2 \qquad (10)$$

Color Distance. The statistical deviations of color matches provides a measure of the visual threshold color discrimination. Practice shows that Just Noticeable Differences (JND's) are approximately equal to three times the standard deviations [4].

According to Schrodinger's hypothesis [4] the perceptual difference $|\Delta B|$ between two arbitrary colors A and C can be determined as the smallest number of JND's separating the two colors

$$|\Delta B| = \left(\frac{1}{3} \int_A^C d\right) \qquad (11)$$

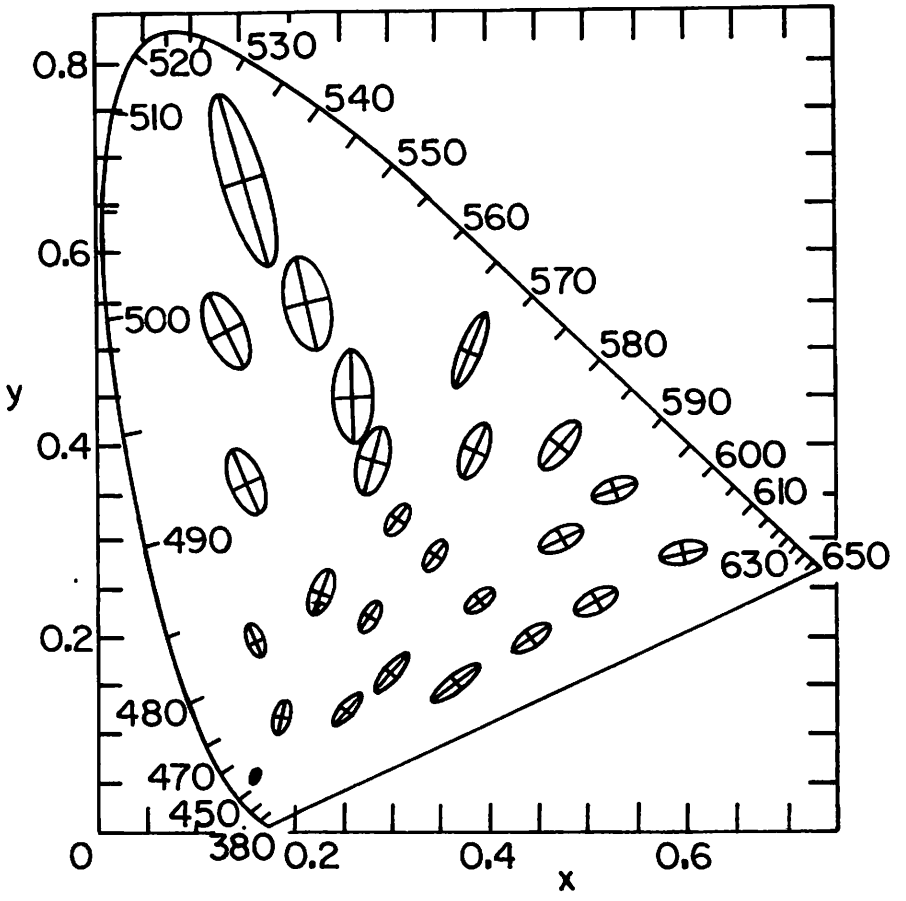
where dS is a line element and the path of integration is a geodesic line [4].

Under the assumption of neural channels perturbed by independent random fluctuations, a line element can be derived from the present model

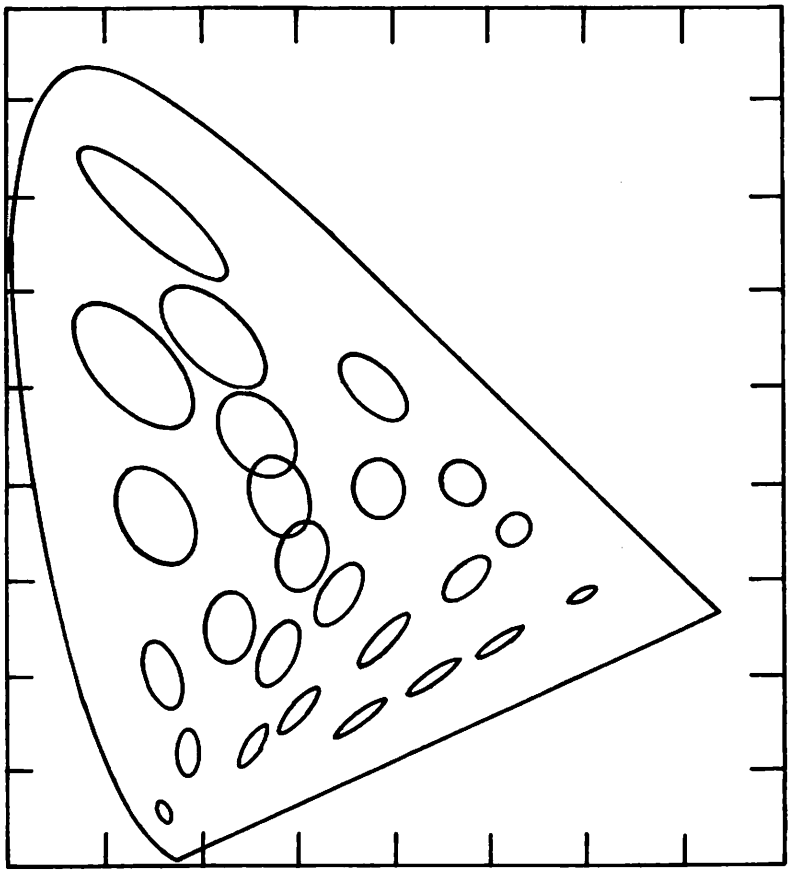
Table 6.1-1

x	y	a	b	θ	$2a_g$	$2b_g$	R
.160	.057	0.85	0.35	62.5	1.226	0.434	2.827
.187	.118	2.20	0.55	77.0	0.898	0.519	1.730
.253	.125	2.50	0.50	55.5	0.868	0.689	1.259
.150	.680	9.60	2.30	105.0	2.042	0.746	2.736
.131	.521	4.70	2.00	112.5	0.652	0.514	1.268
.212	.550	5.80	2.30	100.0	1.167	0.642	1.818
.258	.450	5.00	2.00	92.0	1.074	0.591	1.817
.152	.365	3.80	1.90	110.0	0.621	0.592	1.049
.280	.385	4.00	1.50	75.5	0.925	0.506	1.827
.380	.498	4.40	1.20	70.0	2.052	0.395	5.189
.160	.200	2.10	0.95	104.0	0.499	0.508	0.982
.228	.250	3.10	0.90	72.0	0.749	0.436	1.716
.305	.323	2.30	0.90	58.0	0.580	0.390	1.487
.385	.393	3.80	1.60	65.5	1.186	0.659	1.800
.472	.399	3.20	1.40	51.0	1.549	0.708	2.189
.527	.350	2.60	1.30	20.0	1.418	1.001	1.418
.475	.300	2.90	1.10	28.5	1.135	0.980	1.158
.510	.236	2.40	1.20	29.5	0.862	2.543	0.339
.596	.283	2.60	1.30	13.0	2.155	2.469	0.873
.344	.284	2.30	0.90	60.0	0.581	0.541	1.074
.390	.237	2.50	1.00	47.0	0.650	1.027	0.632
.441	.198	2.80	0.95	34.5	0.765	2.117	0.361
.278	.223	2.40	0.55	57.7	0.638	0.365	1.747
.300	.163	2.90	0.60	54.0	0.825	0.750	1.101
.365	.153	3.60	0.95	40.0	0.966	2.327	0.415

The mapping of MacAdam's ellipses into $\mathcal{L}$ - space. x, y are the chromaticity coordinates of the centers of the ellipses in the 1931 CIE chromaticity diagram; a and b are the lengths of the semi-axes, and θ is the angle of the major axis of the ellipses with the x-axis. $2a_g$ and $2b_g$ are the lengths of the axes 2a, 2b mapped into $\mathcal{L}$ - space and R is their ratio. The largest errors seem to occur for saturated colors, where one tristimulus component, at least, is small (see Eq. (18b) and (18c)).



a) measured standard deviations



b) predicted standard deviations

Figure 6.1-2. Visual color discrimination

$$| dS | = \left[\sum \left(\sum \frac{\partial g_i}{\partial t_j} \, dt_i \right)^2 \right]^{\frac{1}{2}} \tag{12a}$$

$$= \left[\frac{c_1^2 + c_2^2 + c_3^2}{(k_1 t_1 + t_{10})^2} \, dt_1^2 + \left(\frac{c_2}{k_2 t_2 + t_{20}} \right)^2 dt_2^2 + \left(\frac{c_3}{k_3 t_3 + t_{30}} \right)^2 dt_3^2 \right.$$

$$\left. - \frac{2c_3^2}{(k_1 t_1 + t_{10})(k_2 t_2 + t_{20})} \, dt_1 dt_2 - \frac{2c_3^2}{(k_1 t_1 + t_{10})(k_2 t_2 + t_{20})} \, dt_1 dt_3 \right]^{\frac{1}{2}} \tag{12b}$$

and, at high radiances, approximately

$$| dS | = \left[\frac{c_1^2 + c_2^2 + c_3^2}{k_1^2 t_1^2} \, dt_1^2 + \left(\frac{c_2}{k_2 t_2} \, dt_2 \right)^2 + \left(\frac{c_3}{k_3 t_3} \, dt_3 \right)^2 \right.$$

$$\left. - \frac{2c_2^2}{k_1 k_2 t_1 t_2} \, dt_1 dt_2 - \frac{2c_3^2}{k_1 k_2 t_1 t_3} \, dt_1 dt_3 \right]^{\frac{1}{2}} \tag{12c}$$

Two particular cases are of interest, namely the line elements for constant luminance and achromatic differences respectively

$$| dS |_{Y = Cst} = \left[\left(\frac{c_2}{k_2 t_2} \right)^2 dt_2^2 + \left(\frac{c_3}{k_3 t_3} \right)^2 dt_3^2 \right]^{\frac{1}{2}} \tag{12d}$$

and

$$| dS |_{\text{achromatic}} = \frac{c_1 dt_1}{k_1 t_1} \tag{12e}$$

Equation (12e) shows that the model is consistent with Weber's law.

Examination of eq. (12a) reveals that the geodesics are straight lines in $\mathcal{L}$- space. The measure of perceptual color difference thus reduces to a simple euclidian metric

$$|\Delta B| = \frac{1}{3}\int_A^C dS = \frac{1}{3}\left[(G_A - G_C)^T (G_A - G_C)\right]^{\frac{1}{2}} \tag{13}$$

The $\mathcal{L}$ - space can therefore be called an approximately euclidean color representation space.

Brightness, Lines of Constant Hue and Saturation. Assume that the sensation of brightness produced by a color A is proportional to the perceptual difference between that color and black. Brightness is then proportional to the norm of G_A

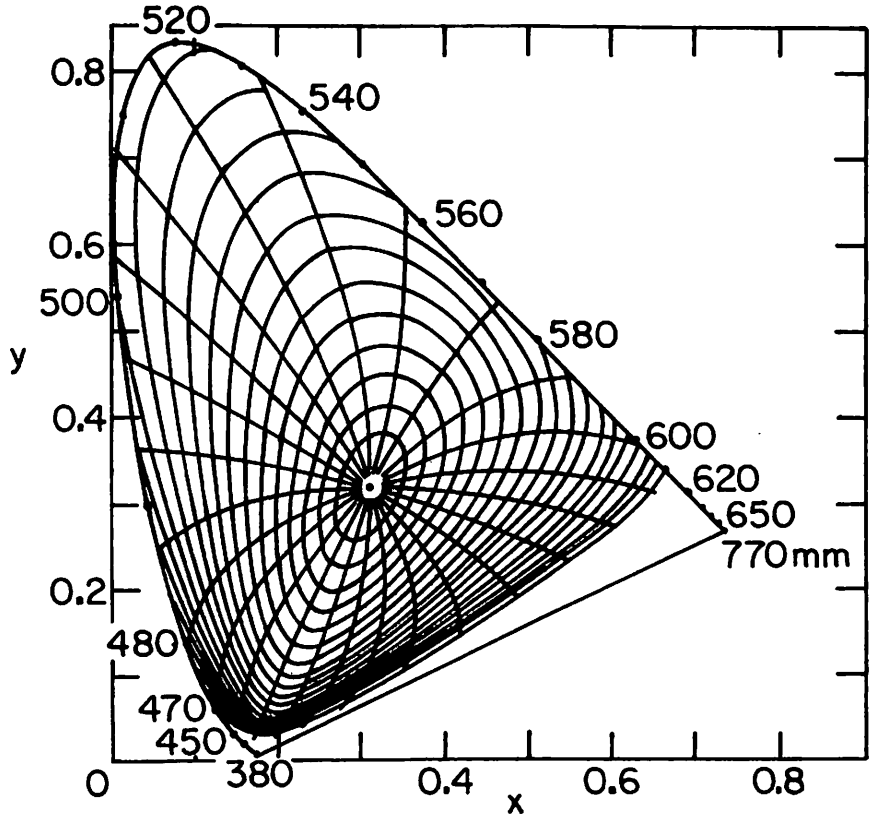
$$B \sim (G_A^T G_A)^{\frac{1}{2}} \tag{14}$$

The failure of brightness additivity in color mixtures then derives immediately from Schwartz's inequality

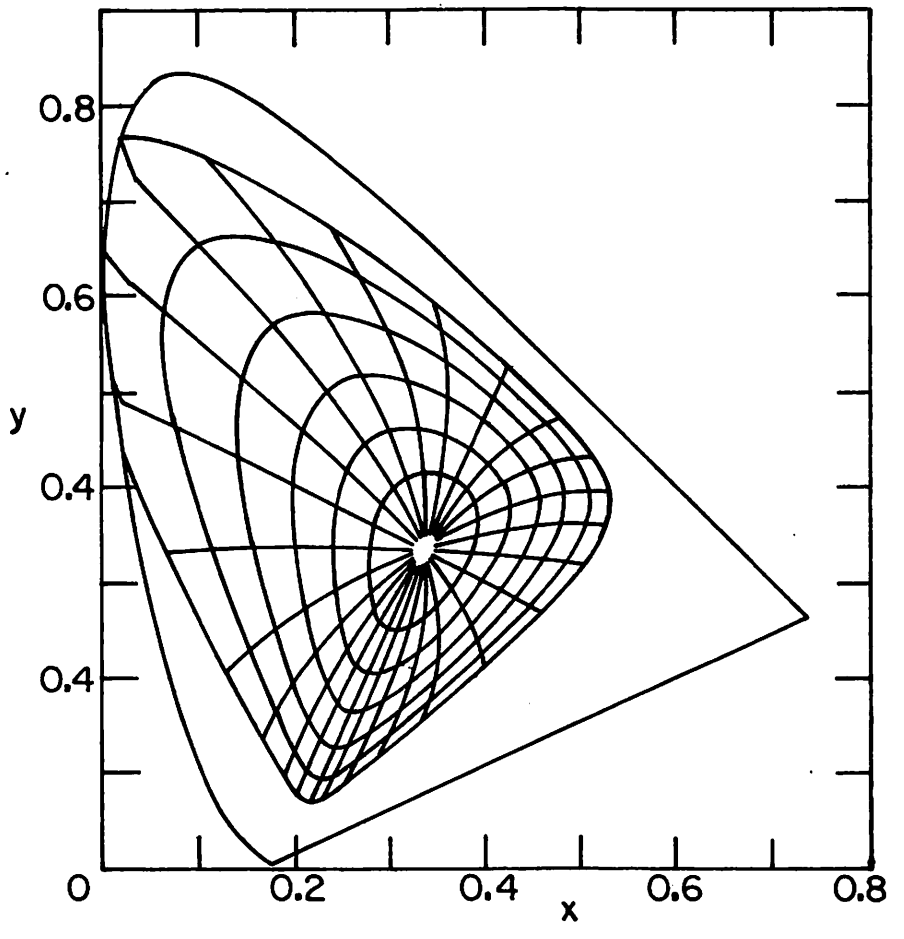
$$B_{A+C} = |G_A + G_C| \leqq B_A + B_C = |G_A| + |G_C| \tag{15}$$

The model therefore predicts that the brightness of color mixtures is always smaller or equal to the sum of the component's brightnesses.

Schrodinger's hypothesis [4] that colors of constant saturation lie on geodesic circles with constant luminance will now be explored. In $\mathcal{L}$ - space, such geodesics are ordinary circles in a plane g_1 = constant and centered at $g_2 = g_3 = 0$. The shortest distance from any point on one circle to a point on another circle is, of course, along a straight line passing through $g_2 = g_2 = 0$. Radial lines in $\mathcal{L}$ - space would then, according to Schrodinger, be projections of lines of constant hue on the plane g_1 = constant. Figure 3b shows a set of equidistant concentric circles and radial lines in $\mathcal{L}$ - space mapped into the CIE x-y diagram. These lines resemble very much the lines of constant saturation, and especially constant hue found in empirical color classifications (compare, for example, with the Munsell system [1]). It seems therefore that the perception of hue is governed by the ratio of the chromatic signals g_2 and g_3 and the sensation of color saturation would derive from the



a) Stile's live element



b) Visual Model

Figure 6.1-3. Lines of constant hue and constant saturation (Schrödinger's hypothesis)

norm of a vector sum $[0, g_2, g_3)^T (0, g_2, g_3)]^{\frac{1}{2}}$ of the same signals.

Applications. It has been shown [5] that the use of a simple model of achromatic vision can improve considerably the efficiency of coding, enhancement and restoration techniques of black and white images. This is essentially done as follows: let I(x, y) represent the intensity of an image and M a mapping into a perceptual space (here G = H * NI(x, y) is a scalar function of I). Image processing takes place in $\mathscr{G}$ - space and an inverse mapping M^{-1} is performed before displaying the processed image.

This idea has been extended to color image processing for a problem only apparently trivial, namely the efficient quantization of color image signals [6, 7]. A color image has been mapped into $\mathscr{G}$ - space according to eqs. (3) and (4) with $k_i = 1$, $t_{i0} = 0$, $h_i(x, y) = 1$, $i = 1, 2, 3$. The signals g_1, g_2, g_3 were then quantized linearly such that each quantization step corresponds to approximately three times just noticeable differences (JND). Accordingly, g_1 was quantized with 32 levels, g_2 fourteen levels and g_3 twelve levels. The quantized signals were then transformed back into red, green and blue signals and displayed on a color TV monitor. Figure 4a and b show the original and quantized picture respectively. Note that artificial contouring is almost inexistant despite the coarse quantization used. Figure 4c and 4d show the subjective effect of linear quantization on the signals g_2 and g_3 (g_1 = constant). The boundaries of the color diagram is determined by the primaries of the display device, and subjective color variations across the diagram are almost uniform (except in the neighborhood of the red primary which lies very close to the spectrum locus). The quantization of the chromatic component, in figure 4b corresponds to that of figure 4d.

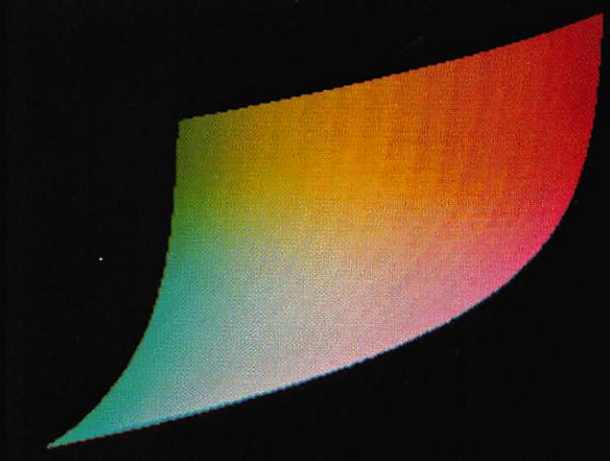
References

1. W. Frei, "Modelling Color Vision for Psychovisual Image Processing," USCEE Report 459, September, 1973, pp. 112-122.

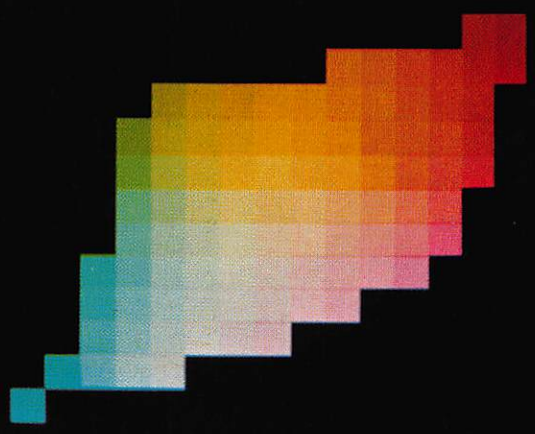
2. W. Frei, "A New Model of Color Vision and Some Practical Implication," USCEE Report 530, March, 1974. pp. 128-143

(a) Original picture

(b) Quantized picture.
Parameters:
g_1 : 32 levels
g_2 : 14 levels
g_3 : 12 levels
(no dithering!)



(c) g_2-g_3 uniform chromatic plane
(g_1 = log luminance = constant)



(d) Quantized chromatic plane.
Same parameters as in the quantized picture above.

Figure 6.1-4. Quantization of color image signals.

3. W. R. Brown and D. L. MacAdam, "Visual Sensitivities to Combined Chromaticity and Luminance Differences," T. Opt. Soc. Am. 39, 1949, pp. 808-834.

4. G. Wysecky and W. S. Stiles, Color Science, John Wiley and Sons, New York, 1957.

5. T. G. Stockham "Image Processing in the Context of a Visual Model," Proc. IEEE, Vol. 60, July, 1972, pp. 828-842.

6. W. Frei et al., "Quantization of Pictorial Color Information," (in German), Nachrichtentechnische Zeitung, Vol. 25, 1968, pp. 401-404.

7. W. Frei, "Quantization of Pictorial Color Information, Nonlinear Transforms," Proc. IEEE, 1973, pp. 465-466.

6.2 A New Look at the Visual System Model

Hideo Murakami and Ernest L. Hall

Study of the human visual system and visual perception is important in many fields especially psychology, anatomy, and physiology; however, a knowledge of visual perception is a fundamental requirement for the designer of picture processing techniques. In fact, the basic difference between the numerical analysis of matrices and computer picture processing is based on the psychovisual concept of a visual image.

A new visual system model is presented here. The structure of the model is based on the histological and anatomical structure of the human visual system. A detailed description of the human visual system may be found in [1-3]. Furthermore, the analogous optical system for the occular media is well documented [4-5]. The currently accepted spatio-temporal model is briefly presented and was constructed from a review of current literature. Then the proposed visual system model is presented and the spatial and temporal responses are explored in detail. An analysis of the combined spatio-temporal response and a conjecture about color response are also described.

Present Spatio- Temporal Model. An overall model of the visual system was not found in the literature although several references emphasizing the spatial and temporal frequency characteristics of the human visual system are available. From a careful survey of these, one may arrive at what may be called the "presently accepted visual system model." The spatial elements of this model consist of a nonlinearity followed by a bandpass spatial filter. It should also be noted that if one assumes a retinal stimulus image as does Stockham [7], the above elements do not conflict severely with the model proposed in this paper. However, in experimental work the stimulus image is presented to the occular media not to the retina. The shape of the nonlinearity is commonly taken as logarithmic although other forms are also found. The shape of the bandpass characteristics varies from strictly low pass to bandpass but generally a bandpass is used. The temporal model as well as the combined spatio-temporal model is usually considered as a

nonlinearity followed by bandpass temporal or spatio-temporal filters. This presently accepted model or subsets of it are described by Cornsweet [1], Davidson [6], Robson [8], Yasuda and Hiwatasi [10], etc. This basic model has been found useful in several experiments; however, certain experimental difficulties especially at high spatial frequencies have arisen, e.g., Davidson, Cornsweet.

The authors propose that ignoring the occular media of the eye, as is done in this representation, produces a fundamental flaw in this basic model which has led to the experimental difficulties. Furthermore, it will be shown that a new model can be developed which is no more complex than the present model, is based on the histological and anatomical structure of the visual system, predicts the physiological function of the visual system, and is useful in predicting psychophysical phenomena as shown by calculations and comparisons with well-known psychophysical measurements.

Proposed Visual System Model. Perception takes place in the brain. This non-obvious fact has been proven by observing persons who completely lost their sight from damage to the visual cortex without damage to any other part of the visual system. Therefore, a complete model of the visual system must include models of the eye, the optic tract, and the visual cortex. Although the state of development of a complete visual system model is still in its embryonic stages, the authors believe, as Cornsweet states, that "it is pleasing to consider the possiblity that the human nervous system, despite its obvious overall complexity, is really composed of a very large number of repetitions and slight variations of a few simple mechanisms" [1].

A simplified functional description of the perceptual process may be given as follows. The stimulus image is generated from a scene and transmitted to the image forming elements of the eyes. These elements focus the image on the retinal receptors. The retinal neural connections may then perform certain processing operations on the neural signals. The next element is the neural processing performed in the optic neural connections from the eyes. Cortical processing tasks include memory reference images, various types of information processing, features

extraction, decisions, or descriptions of various elements of the scene. The results of this process may be indicated in several ways by the response. The complexity of the perception process, coupled with the desire to predict even limited responses, motivate the development of simplified models of the visual system.

The proposed model of the visual system is shown in block diagram form in figure 1. The stimulus image is represented by an image function. The image forming elements of the eye are represented by an optical filter. The retinal receptor response is divided into two pathways, one for rods and one for cones. Each pathway contains a selective spectral filter, a low pass temporal filter, and a low pass spatial filter. These are followed by a threshold function, a log conversion, and a high pass spatio-temporal filter. The neural connections in each channel are represented by a spatial optical filter and summations. The neural processing in the optic nerve is again represented by a summation. The transmission link to the cortex may be represented by a set of transmission lines. Finally, the cortical computations may be modeled as summations. The stage of development of the model of the visual system is far from complete, although a large amount of work has been accomplished [1-12]. At present, certain elements such as the image forming mechanisms can be modeled very accurately; other elements, for example the cortical computations, are only modeled in a conjectural manner. Therefore, one may expect to find insight, but not a complete solution from the following description of the visual model.

Spatial Frequency Response. A fundamental difference between the "presently accepted model"and the proposed model is the location of the elements which affect the spatial frequency response. The currently accepted model is usually represented by a nonlinear response such as a logarithmetic response followed by a spatial bandpass filter. In the new model the first element encountered is a spatial low pass filter, then the nonlinearity and finally a spatial high pass filter. The consequences of these different configurations will now be explored in detail.

The spatial low pass effect before the logarithmic response in the new model comes from the optical aberration caused by the lens in the eye

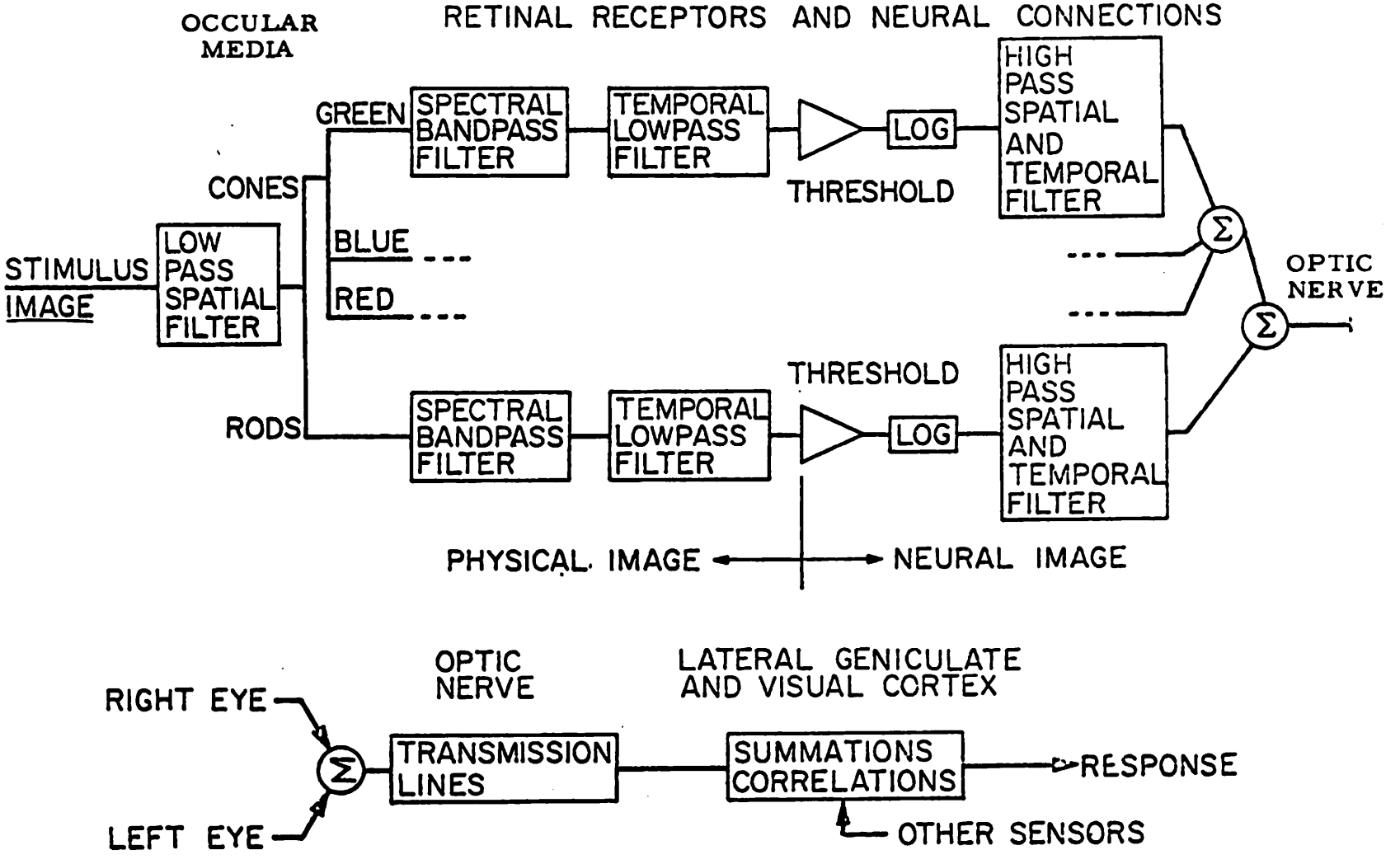


Figure 6.2-1. Backward inhibition model for two receptors.

and the light scatter within the retinal tissue. The measurement of this optical performance of the human eye has been investigated by several researchers [4-12]. Despite the lack of direct access to its image plane at the retina, Gerald Westheimer and Fergus W. Campbell [4] measured the light distribution formed by the living human eye and found that the line spread function of the human eye is approximately of the form $\exp\{-\alpha|x|\}$ where α depends mainly upon the pupil diameter. For a white light and pupil diameter of 3mm, $\alpha = 0.7$. From this line spread function one may easily compute the modulation transfer function. For simplicity one may assume the one dimensional case and calculate the Fourier transform $H_1(w)$, to determine the frequency response of the low pass filter

$$H_1(\omega) = \int_{-\infty}^{\infty} f(x)\exp\{-j\omega x\}\,dx$$

which results in

$$H_1(\omega) = \frac{2\alpha}{\alpha^2 + (\omega)^2} \quad .$$

One may see that the 3dB down cutoff frequency of this low pass filter is approximately 6.6 cycles/degree. Thus, it can hardly be disputed that anatomically, the low pass filter modifies all images information before the receptors are encountered. However, other questions naturally arise. Is the log-bandpass filter equivalent to the low pass-log-high pass combination? Is the low pass effect of the occular media of secondary importance to the low pass effect of the discrete receptor array or the low pass effect of mutual receptor excitation? To answer these questions the model of the retinal neural networks must be considered.

Retinal Neural Networks . G. G. Furman [11] proposed four types of neural networks: a forward inhibition model, a backward inhibition model, a forward shunting model, and a backward shunting model. Four of these essentially have the same effect, i.e. a spatial high pass filter

effect. The backward inhibition model is employed in our model and is explained in detail in [1].

The backward inhibition model for two receptors is shown in figure 2. In the figure, the receptors have a logarithmic response to the incoming light intensity; thus they correspond to the logarithmic operation in the proposed model. In figure 3, e_i is the frequency at which receptor i would produce pulses if the receptor i alone were illuminated (the level of excitation of the receptor i); g_i is the frequency of the pulse of the receptor i after the network (pulses/sec.); b_{ij} is the inhibitory coefficient representing the strength of the inhibition that g_j exerts on g_i. If one neglects the inhibitory threshold and the case when frequencies of pulses are below inhibitory threshold, the following equations hold

$$g_1 = e_1 - b_{12}g_2$$

$$g_2 = e_2 - b_{12}g_1 .$$

The general equation considering other receptors is

$$g_i = e_i - \sum_{j=1}^{n} b_{ij}g_j , \qquad i = 1, 2, \ldots, n.$$

The equations may also be written in concise matrix form with the definitions

$$\underline{e} = \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ \cdot \\ e_n \end{bmatrix} \qquad \underline{g} = \begin{bmatrix} g_1 \\ g_2 \\ \cdot \\ \cdot \\ \cdot \\ g_n \end{bmatrix} \qquad \text{and} \qquad \underline{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} .$$

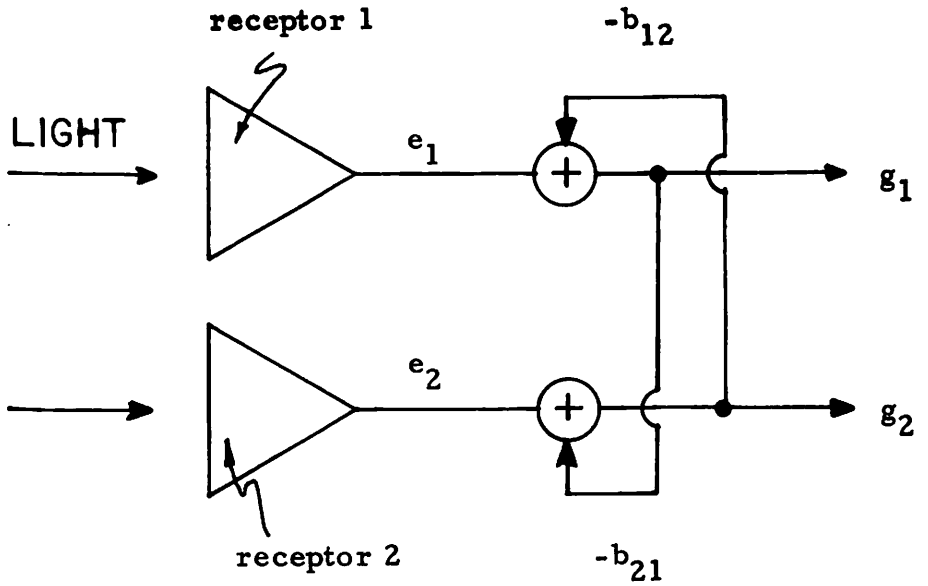
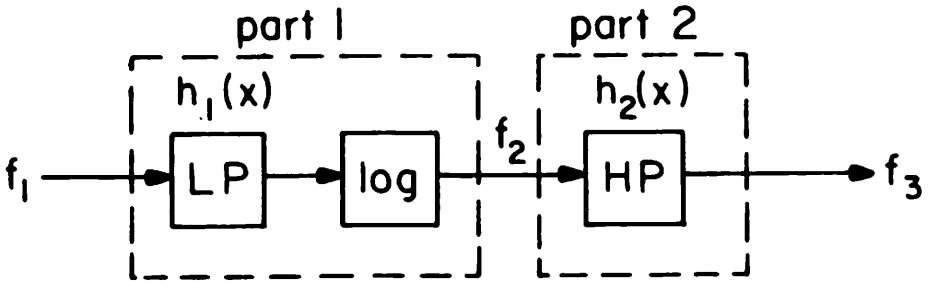
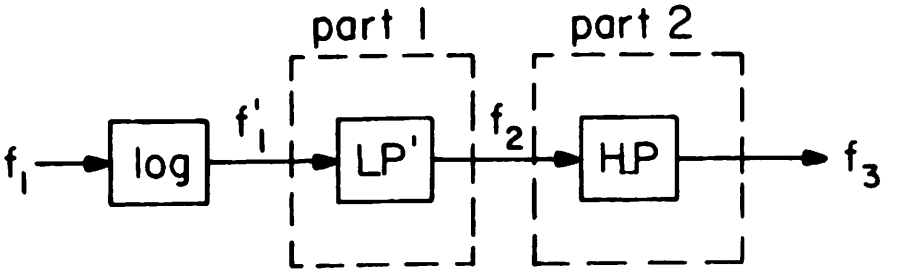


Figure 6.2-2. Backward inhibition model for two receptors.



(a) Proposed model of spatial characteristics



(b) Modified model of spatial characteristics.

Figure 6.2-3. Proposed model of spatial characteristics.

With this matrix notation the above equation will be:

$$\underline{g} = \underline{e} - \underline{\underline{B}}\underline{g}$$

Then

$$(I + B)\,\underline{g} = \underline{e}$$

or

$$\underline{g} = (I + B)^{-1}\,\underline{e}$$

Now assume that there is no self-inhibitory interaction and that the inhibitory interaction is an exponentially decreasing function of the distance of the receptors

$$b_{ij} = \begin{cases} 0 & i = j \\ a_0 \exp\{ -a\,|i-j| \}, & \end{cases}$$

Under this assumption

$$I + B = \begin{bmatrix} 1 & a_0 e^{-a} & a_0 e^{-2a} & \cdots & a_0 e^{-na} \\ a_0 e^{-a} & 1 & a_0 e^{-a} & \cdots & a_0 e^{-(n-1)a} \\ \vdots & & & & \\ a_0 e^{-na} & a_0 e^{-(n-1)a} & a_0 e^{-(n-2)a} & \cdots & 1 \end{bmatrix}$$

This matrix is Toplitz since all diagonal terms are equal, and furthermore may be interpreted as corresponding to a linear, shift invariant system. Thus, this convolutional matrix has impulse response representation

$$g(x) = (1-a_0)\,\delta(x) + a_0 \exp(-a\,|x|)$$

The Fourier transform of the impulse response is

$$G(\omega) = 2\pi(1-a_0) + a_0 \frac{2a}{a^2 + \omega^2}$$

Thus $(\underline{I} + \underline{B})^{-1}$ has the frequency characteristic

$$H_2(\omega) = \frac{1}{G(\omega)} = \frac{a^2 + \omega^2}{2a_0 a + 2\pi(1-a_0)(a^2+\omega^2)}$$

This function represents a high pass filter with $H_2(0) = \frac{a}{2a_0 + 2\pi a(1-a_0)}$

and $\lim_{w \to \infty} H_2(w) = \frac{1}{2\pi(1-a_0)}$.

From comparing this function to the experimental results of Davidson [6] it appears that suitable parameter values to use for this comparison are : $a = 0.01$ and $a_0 = 0.6$.

Davidson [6] measured approximate modulation transfer functions of the human vision at several different levels of contrast. He called these measured functions "describing functions" to emphasize the nonlinear aspects of the visual system. His experiments showed that the describing functions vary with contrast at high frequencies. This variation has been attributed to the experimental apparatus. It has been found possible to explain this variation by a simple analysis of the proposed model.

Based on the hypothesis that nonlinear characteristics of the human visual system could be traced to a simple logarithmic transformation at the earlier stage, Davidson used the stimuli of the form, $f(x,y) = f^o \exp\{A \cos(\omega y)\}$ to present to the eye in his experiments. Where f^o is a constant; A is a constant to measure the contrast of the signal; $\omega/2\pi$ is the spatial frequency of the signal. The logarithm of the stimuli will be $\log f^o + A \cos(\omega y)$, which has a desirable form to find the frequency characteristics of the linear filter after the logarithmic transformation.

The experiment was conducted as follows: the contrast of the stimuli, $f(x, y) = f^{o}\exp\{A\cos(\omega y)\}$, and, for each spatial frequency ω, A was selected so that the contrasts of the two stimuli's were seen to be the same. From these selected A's, he computed the describing functions for several chosen standard stimuli.

An analysis of the proposed model which predicts the variation of the high frequency response as a function of contrast will now be reported. For the analysis, it was convenient to divide the system into two parts as shown in figure 6(a) and investigate each separately then combined.

Consider the two relevant system configurations as shown in figure 3(a) and (b) Part 1. It should first be noted that the systems are not equivalent. For example, given a system consisting of a log operation followed by a linear filter, another system consisting of a filter followed by a log operation can be made equivalent only if the second filter function is made to depend upon the input signal. However, for the analysis, it was convenient to transform the proposed configuration into the input dependent form to obtain the exponent-log cancellation.

In the analysis it was shown that the response of the nonlinear combination of the spatial filter - log did indeed produce a variation of the response with the input signal contrast. To show this fact, the spatial filter - log system was first converted to an equivalent log - input dependent spatial filter. The response of this system was expressed analylitically and numerically approximated to obtain the "describing functions." The results for the low pass spatial filter and log operation are shown in figure 4. The combined characteristic with the high pass filter is shown in figure 15 and compares favorable with the experimental results of Davidson. Thus, the variation of the describing functions at high spatial frequencies appears not to be due to experimental equipment, but a predictable response of the visual system.

The consequences of these results also apply to several other theoretical and experimental studies of the visual system. For example, a widely used techniques in computer image processing is multiplicative filtering [7]. In experimental work with the log-bandpass model the

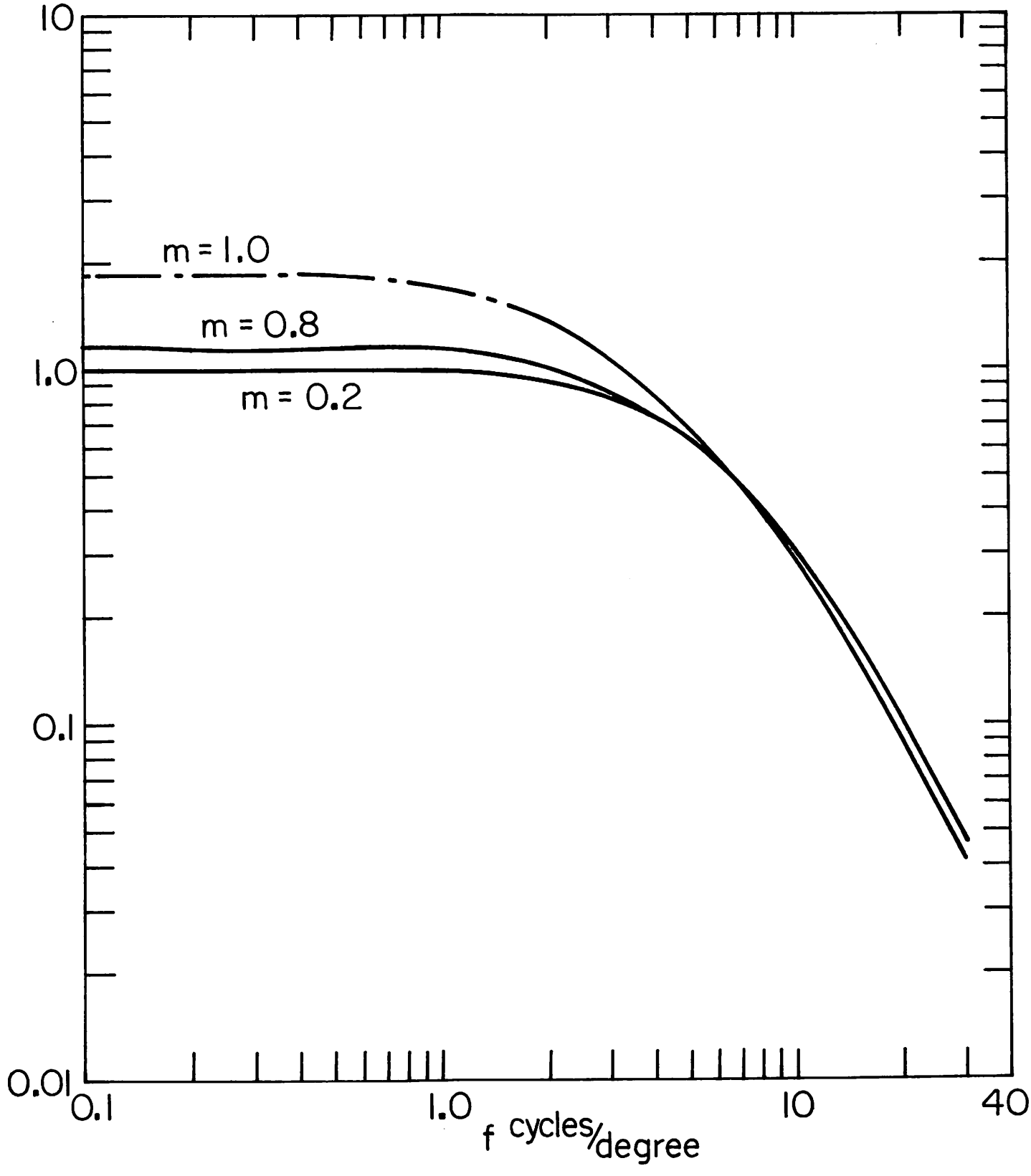


Figure 6.2-4. Characteristic spatial frequency response of Part 1 of the proposed model showing variation with contrast of input.

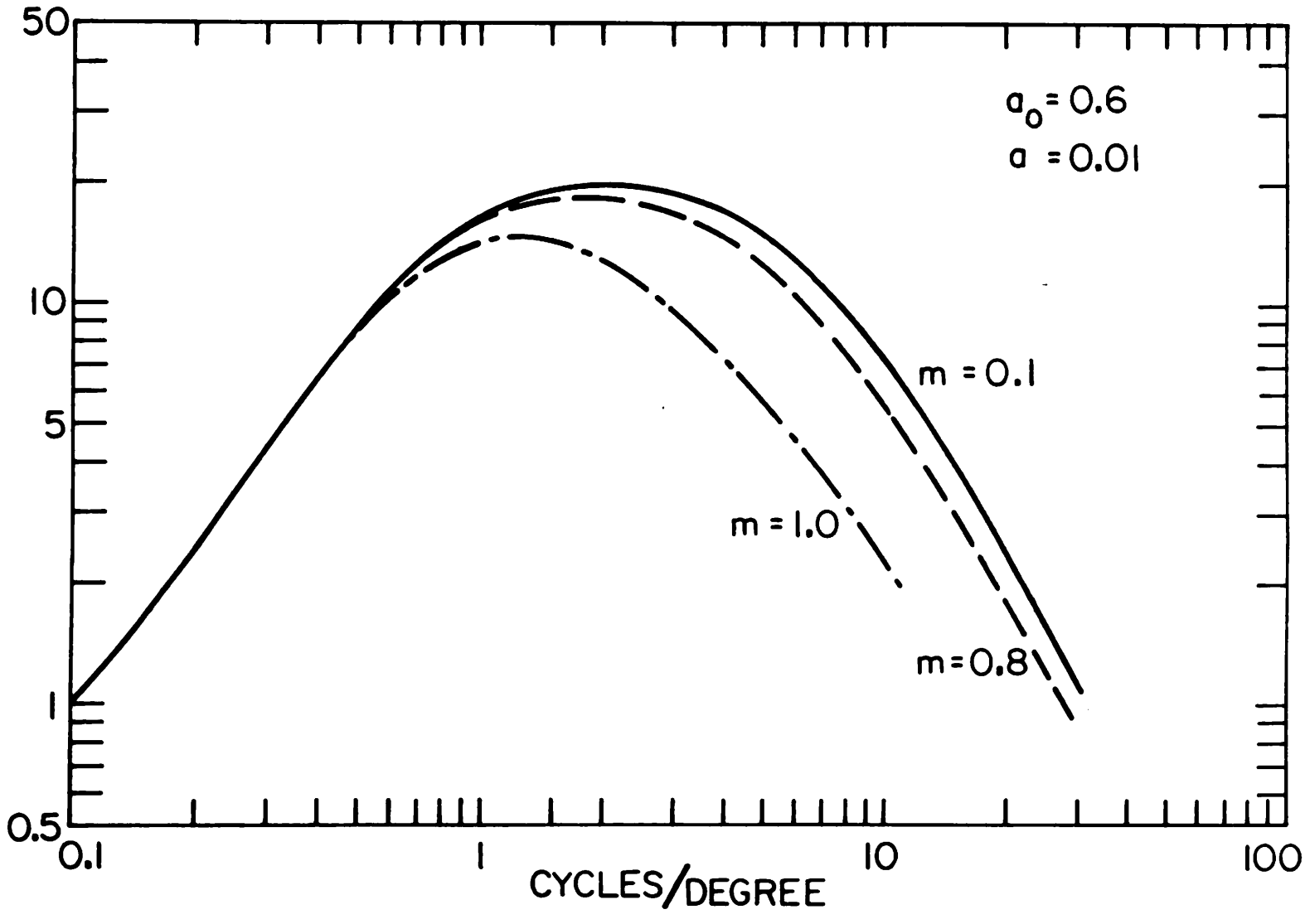


Figure 6.2-5. Combined characteristic spatial frequency response of the proposed model.

contrast variation is predictable. For example, the relatively low resolution of one lp/mm viewed at 40 cm corresponds to about 7 cycles/degree and is well above the spatial frequency at which the variation is noticable. Further the display of a square wave image display at this frequency contains even higher frequencies due to the harmonics.

Also, note that the experiments such as that performed by Davidson aimed at developing a describing function of the visual system or those conducted by Stockham which illustrated Mach band cancellation may easily be performed with the new model. Suppose that the visual system is represented by the cascade of a simple integrator, a log nonlinearity and a differentiation. Then to derive a describing function of the retinal neural network one may present a sinsoid

$$f(y) = A \cos wy$$

which is processed by the inverse of the log response giving

$$g_1(y) = \exp \{ f(y) \} = \exp \{ A \cos wy \}$$

and by the inverse of the integration

$$g_2(y) = \frac{d}{dy} g_1(y) = -A\, w \sin(wy) \exp \{ A \cos wy \}$$

Note that the image $g_2(y)$ is now weighted in amplitude proportionally to the frequency of the input. Thus, as the frequency of the input is increased, the amplitude of the processed signal is increased. A describing function derived in this manner could be called a modulation transfer function and would vary with the contrast of the input signal in a linear manner.

The preprocessing required for Mach band cancellation is also easily derived for this simple example. Again starting with the image

$$f(y) = A \cos wy$$

the first step would be an integration to negate the effects of the neural connections to produce

$$g_1(y) = \int A \cos (wy)\, dy = \frac{A}{w} \sin (wy) + C$$

Next an exponentation is required

$$g_2(y) = \exp \{ C \} \exp \{ \frac{A}{w} \sin wy \}$$

Finally, a differentiation is required

$$g_3(y) = \exp \{ C \} A \sin(wy) \exp \{ \frac{A}{w} \sin wy \}$$

The image $g_3(y)$ should appear to the observer without Mach bands.

Temporal Response. The same descrepancy between the location of elements responsible for the temporal frequency response in the currently accepted model and the proposed model may be noted. In the usual model the elements consists of a nonlinearity followed by the temporal band-pass filter. In the proposed model the first element is a low pass temporal filter, then the nonlinearity and the high pass filter.

There is again a physiological reason for placing the low pass filter before the nonlinearity. The temporal low pass filter arises from the 30 msec time required for light absorbed by the visual receptors to initiate the photochemical reaction which generates the neural pulse. Also, there is again an inconsistency encountered if one assumes the log-bandpass characteristic.

If one postulates also the existence of the excitational interaction among nerve cells as Yasuda and Hiwatashi [10], then the neural network has a band-pass characteristic. However, the difficulty in this assumption is that, in order to explain the temporal characteristics of the H. V. S., one has to assume that the response time delay of the excitational interaction is ignored whereas that of the inhibition is con-

sidered. Noticing that the excitational interaction and the inhibitory interaction produce the spatial low-pass effect and the high pass effect, respectively, the above statement means that the low pass effect does not have a time delay whereas the high pass effect does. Since the low-pass filter is produced by the optical aberration of the eye in the model, the proposed model can also overcome this difficulty.

Combined Spatio-Temporal Response. Robson [8] experimentally measured spatio-temporal frequency responses of the visual system. He used a 2.5° by 2.5° grating target in which the luminance at right angles to the bars was

$$L = L_0 [1 + m(\cos 2\pi\nu y)(\cos 2\pi ft)]$$

L_0 was kept constant at 20 cd/m^2 and the value of m at subjective disappearance of bars was measured for different frequencies ν and f. The inverse of that value of m was defined as the contrast sensitivity.

Since Robson did not vary the input contrast, the predictable variation of the resultant curves at high frequencies was of course not observed. However, another interesting fact of the proposed model may be correlated with Robson's results. This is the separability of the spatio-temporal response. As pointed out by Budrikis [9], if the human visual system response were product separable then the family of response curves of spatial or temporal frequency with one of the frequencies fixed would differ from each other by constant factors or in the logarithmic plots, by constant displacements in the vertical direction. The experimental curves do exhibit this property except at frequencies below about 5 cycles/degree and 6 cycles/second. Note that this is not predicted by the log-bandpass type model.

It can be shown that this effect is predictable from the proposed model. The basic mechanism is that if the output of the nonlinearity is at sufficiently high frequencies so as to be within the passband of the high pass spatio-temporal filter, then the effect of this high pass filter is neglibible. In this situation the spatio-temporal response is product separable since the low pass spatial and temporal filters are

independent. However, if the output of the nonlinearity contains low frequency signals which are affected by the stop band of the high pass filters then the overall response is not product separable since the high pass spatio-temporal filters are not independent.

Application to Color Vision. One possible application to color vision should also be mentioned. In previous work, a difficulty has been encountered in the study of color balance which has been attributed to different shape nonlinearities in the three color receptor channels. However, it is interesting to note that a similar effect is predictable from the proposal model. It is well known that the lens of the eye produces chromatic abberation, i.e. different focal lengths are required for different spectral signals. If a color image which may be considered as the sum of red, green and blue component images, is observed, each of the component images is seen blurred to a different extent. Thus, the low pass filters in the red, green and blue channels in the proposed model, have different cutoff frequencies. Even if the nonlinearities in the three channels were identical, (and it appears that there is no psysiological evidence to support the converse) the overall responses in each channel would appear to have different nonlinearities [12]. Thus it appears that a correction for the chromatic abberation is essential for quantitative study of color phenomena.

References

1. T. N. Cornsweet, Visual Perception, Academic, New York, N.Y., 1970.

2. G. W. Wyszecki and W. S. Stiles, Color Science, Concepts and Methods, Quantitative Data and Formulas, Wiley, New York, N.Y., 1967.

3. H. Davson, The Physiology of the Eye, Little, Brown, and Co., Boston, 1963.

4. G. Westhemier and F. W. Campbell, "Light Distribution in Image Formed by the Living Eye," Journal of the Optical Society of America, Vol. 52, No. 9, pp. 1040-1045, May 1964.

5. E. M. Lowry and J. J. DePalma, "Sine Wave Response of the Visual System, II, Since Wave and Square Wave Contrast Sensitivity," Journal of the Optical Society of America, Vol. 52, No. 8 August, 1961, pp. 740-746.

6. M. S. Davidson, "Perturbation Approach to Spatial Brightness Interaction in Human Vision," Journal of the Optical Society of America, Vol. 58, 1968, pp. 1300-1309.

7. T. G. Stockham, Jr., "Image Processing in the Context of a Visual Model," Proc. of the IEEE, Vol, 60, No. 7, July, 1972, pp. 828-1309.

8. J. G. Robson, "Spatial and Temporal Contrast - Sensitivity Functions of the Visual System," Journal of the Optical Society of America, Vol. 56, August, 1966, pp. 1141-1142.

9. Z. L. Budrikis, "Visual Fidelity Criterion and Model," Proc. of the IEEE, Vol. 60, No. 7, July, 1972, pp. 771-779.

10. M. Yasuda and K. Hiwatashi, "A Model of Retinal Neural Networks and Its Spatio-Temporal Characteristics," NHK Laboratories Note, Serial 116, January, 1968.

11. G. G. Furman, "Comparison of Models for Substructure Shunting Lateral Inhibition in Receptor Neuron Fields," Kybeonetik, Vol. 2, No. 6, 257-274, 1965.

12. R. E. Savoie, "Bezold-Brucke effect and Visual Nonlinearity," J. of Opt. Soc. Amer., Vol. 63, No. 10, Oct. 1973.

7. Image Processing Hardware and Software Projects.

The image processing hardware and software projects are developmental projects supportive of the image processing research.

The hardware projects described in the first report summarize progress on the development of real time color image displays for the ARPANET. Also included is progress on construction of the real time color television recorder and playback system.

The software projects report explains the image processing software system being implemented for access over the ARPANET.

7.1 Hardware Projects

Toyone Mayeda

A digital color television system for display of monochrome and color television pictures received over the ARPANET has been implemented and in operation since April, 1974. Development is proceeding on a second unit whose specifications are described below:

1. Receive asynchronous digital picture information from the ARPANET TIP with brightness resolution up to 256 levels and at input rates up to 19.2K baud.
2. Include a function memory which can be used to translate the 8 bit data words (from the refresh memory) with any desired transfer curve. The function memory can be remotely programmed from the TIP or by local switch control.
3. Display a 256 x 256 eight bit image, or a six or seven bit image and a one bit graphic overlay.
4. Use an alphanumeric keyboard when available to communicate with the ARPANET TIP and also to generate alphanumeric characters on the display monitor.
5. Output the monochrome video data and alphanumeric characters in composite RF format so that it can be displayed on any TV receiver using its antenna input.

The other major hardware project under development is a digital magnetic tape recorder/playback unit for transfer of real time color television signals to and from a PDP-10 computer. The tape unit records the real time imagery in a 600 ips mode and plays the data back at a 1 and 7/8 ips rate which is compatible with the PDP-10 channel capacity. The inverse process can be performed to produce real time television signals from

coded computer records. Presently, all control and signal processing circuitry has been designed and the high speed A/D has been delivered. Delivery of the Orion high speed magnetic tape recorder is expected in December, 1974.

7.2 Software Projects

James M. Pepin

The programming group has worked on several different projects in the last reporting period. These projects include support of image processing display devices, network programming and development of an image processing capability under the Tenex operating system.

The area of progress in the device support functions of the Image Processing Lab include work on support for the microdensitometer and implementing device routines for a PDP 11/10 acquired by the IPL. The support for the microdensitometer has been the major area of effort as this device is the most complex, from a programming point of view, that the IPL has ever acquired. The software has progressed to the point where it is possible to digitize and display data. The effort also required that the programming group implement diagnostic software to enable the Optronics engineers to debug the device and the IPL operations group to determine the device's status. The Optronics is presently controlled by the HP2100 in the IPL. The plan is to attach the device to a PDP 11/10 to enable the 2100 to perform other display tasks while the Optronics is in use. Another project is the development of a front-end system to the 2100 using PDP 11/10's to run the devices in IPL. This task was taken to allow more than one device to be in use simultaneously. Previously, because of device handshaking requirements, the 2100 had been only able to run one display device. With the addition of the 11/10's it will be possible to utilize many of the devices in the IPL simultaneously. The first 11/10 system has been implemented and a second is on order. This first system has been integrated into the IPL operations and is functioning in a production mode.

The largest area of effort has been the conversion of our computer usage to a KI-10 that the ECL has acquired. At the present time the KI-10 has just been installed and the checkout procedure of the code developed is just beginning. A great deal of time has been spent designing the interfaces to the HP2100 and related computers into the IPL. The 10 is expected to run many of the display devices with the users interactively. It is foreseen that a user can have a Fortran program running on the 10 in

real time interacting with any device in the IPL. This will be implemented utilizing a PDP 11/40 that will be connected by a DL10 direct memory interface to the KI-10's memory. The KI-10 will be running the Tenex operating system and much time has been spent familiarizing the staff with its operations and usage. Most of the image processing programs that were implemented for the IBM 360/44 have been converted and tested using the PDP-10's at ISI.

The last area that this report will discuss is the Front End Image Processing System (FEIS). This concept is to use the KI-10 as a front-end to the large "number crunchers" on the net. This concept has been described in the previous reports and development of this has entering the early coding stages. The staff is presently developing a set of programs of the IBM 360/91 at UCLA-CCN. This is the first set of programs being developed for FIES and it is expected that others will follow quickly. The work on the language interpreter will begin shortly on the KI-10. Also with the addition of the KI-10 the network handling functions will be designed and implemented soon.

8. Image Processing Institute Facilities

During the past year the physical facilities of the USC Image Processing Institute have been expanded considerably. The following sections contain a summary description of the present physical facilities.

The Image Processing Institute is located in Powell Hall of Information Sciences, completed in Summer 1973. Powell Hall is a six story building providing over 100 offices and laboratory facilities. The Institute laboratories shown in figure 1 include individual rooms for digital image processing equipment, image display and acquisition devices, coherent optics apparatus, and photographic processors.

8.1 Image Processing Laboratory

The Image Processing Laboratory consists of four interconnecting rooms housing image digitization and display equipment, a minicomputer system including terminals, magnetic tape units, and disks for control of the image processing devices, a well equipped darkroom, and an optical processing and holography laboratory. The Image Processing Laboratory computer system is interconnected with the Biomedical Image Processing Laboratory and the image processing computer system located in the Engineering Computer Laboratory.

Figure 2 contains a block diagram of the Image Processing Laboratory computer/controller system. This system consists of a Hewlett-Packard Model 2100 digital computer with a variety of computer peripherals accessed through an input-output controller. The I/O controller also acts as the communciations link with the image processing computer system of the Engineering Computer Laboratory and the Biomedical Image Processing Laboratory.

The Laboratory image display and digitization devices are listed in figure 3 and shown in the photographs of figures 4 to 6. With these devices it is possible to digitize monochrome and color transparencies ranging in size from 16mm up to full size X-rays and prints of up to 8 in. x 10 in. in size. Hard copy monochrome and color transparencies and prints can be produced in a wide range of sizes up to 8 in. x 10 in.

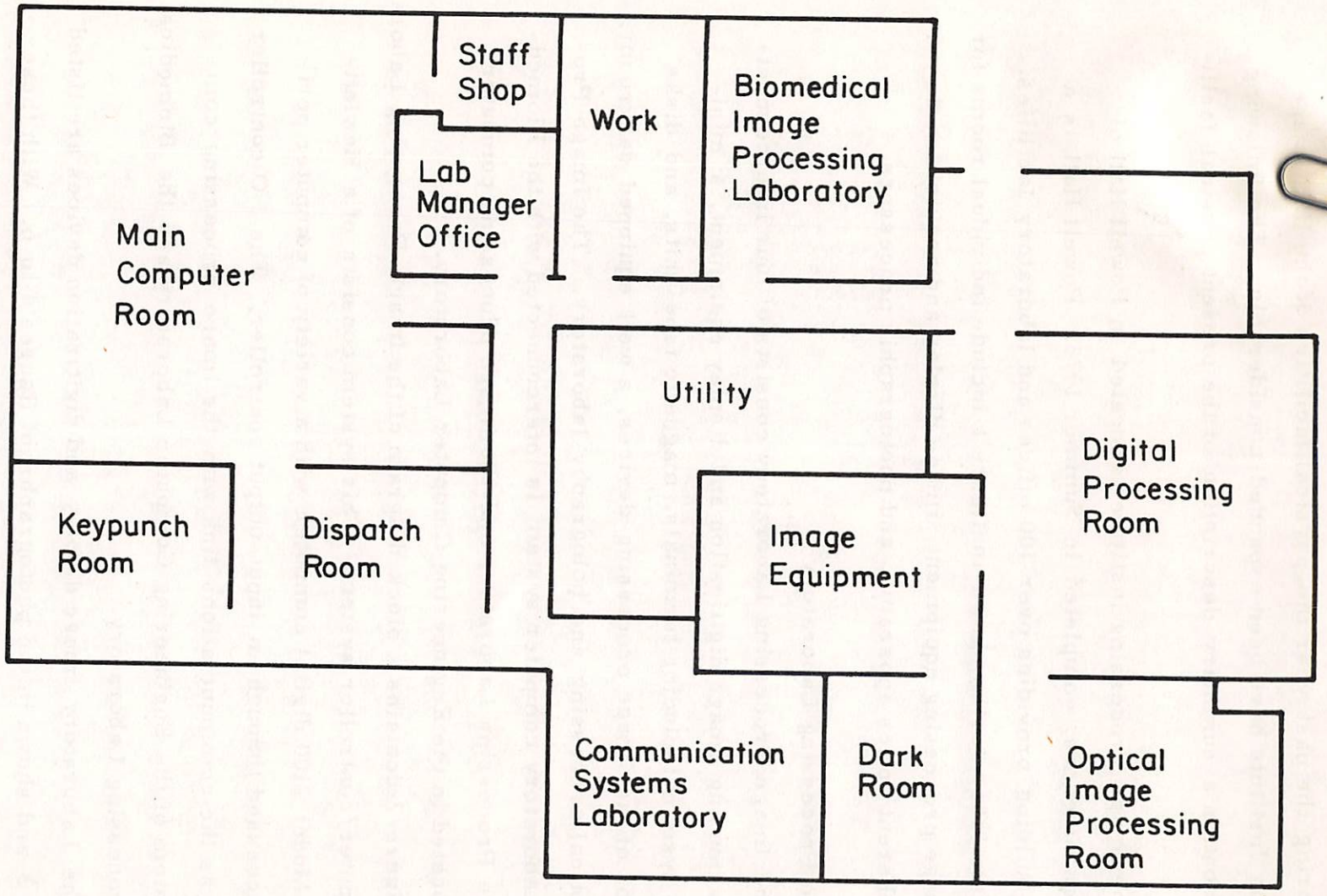


Figure 8-1 Floor plan of powell hall laboratories

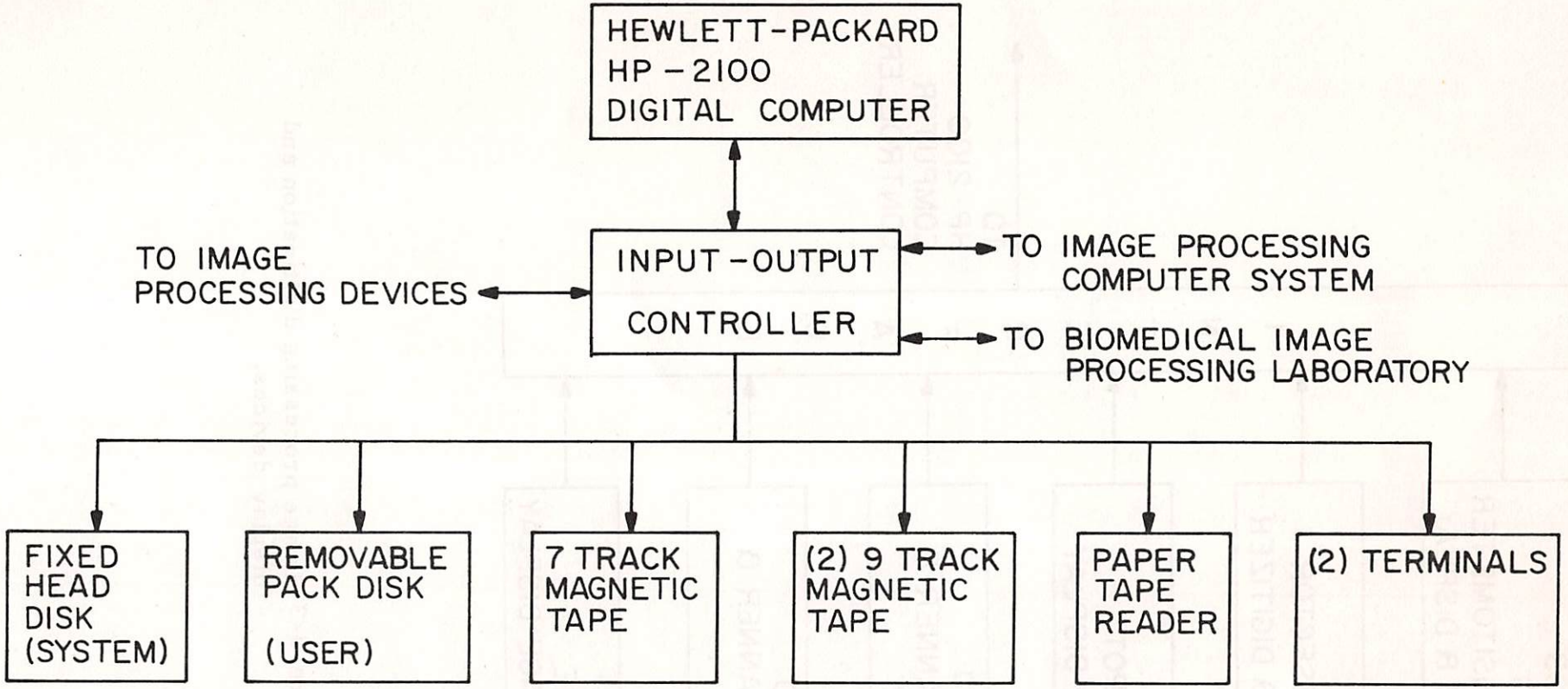


Figure 8-2 Image processing laboratory computer/controller facilities.

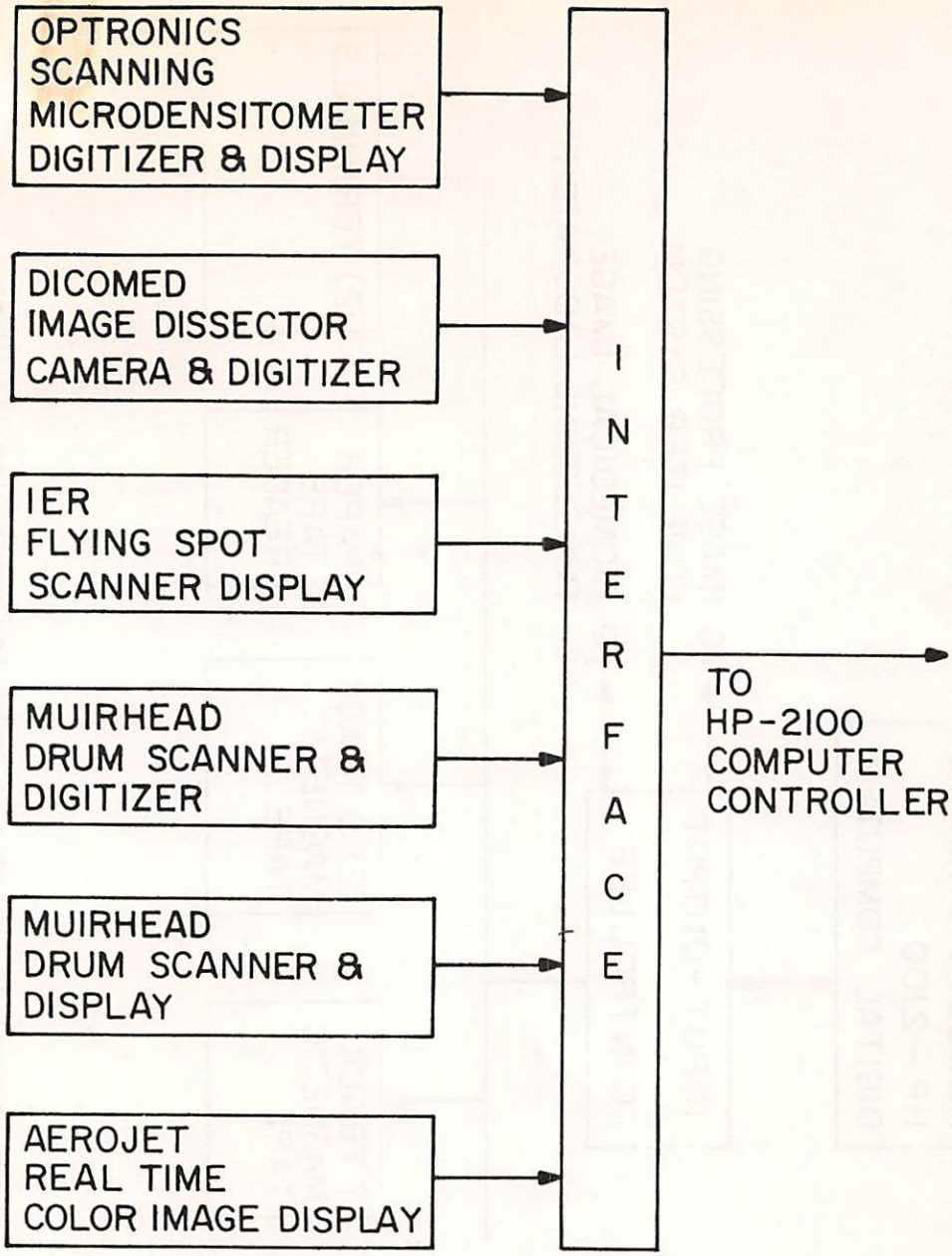
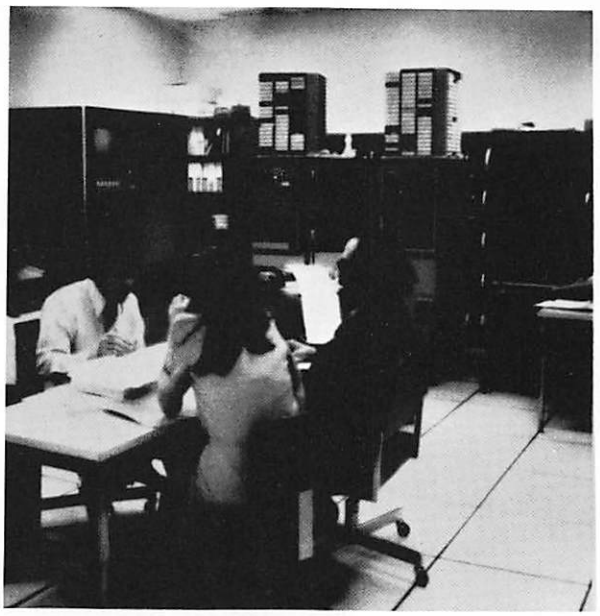
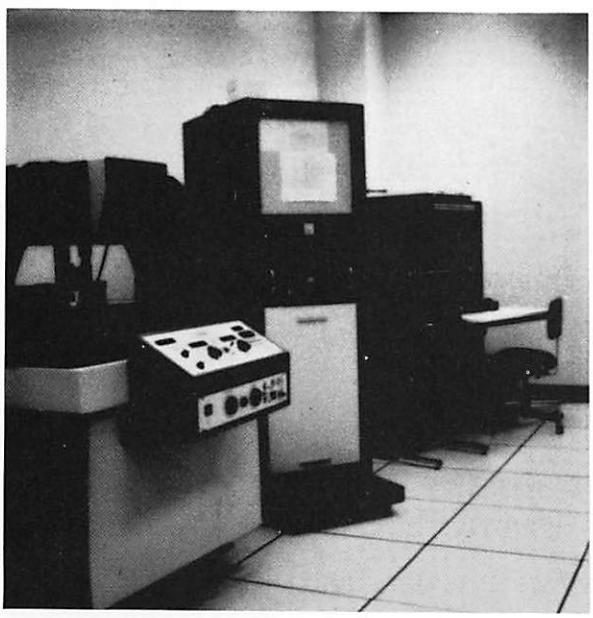


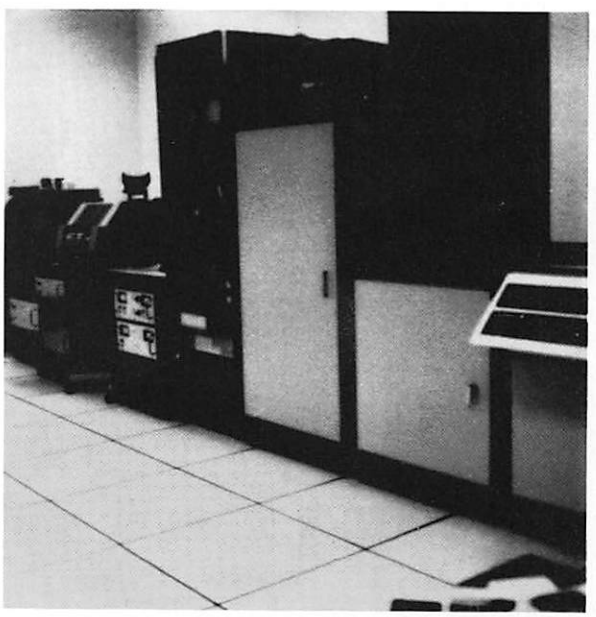
Figure 8-3 Image processing digitization and display devices.



a) Digital Process Room

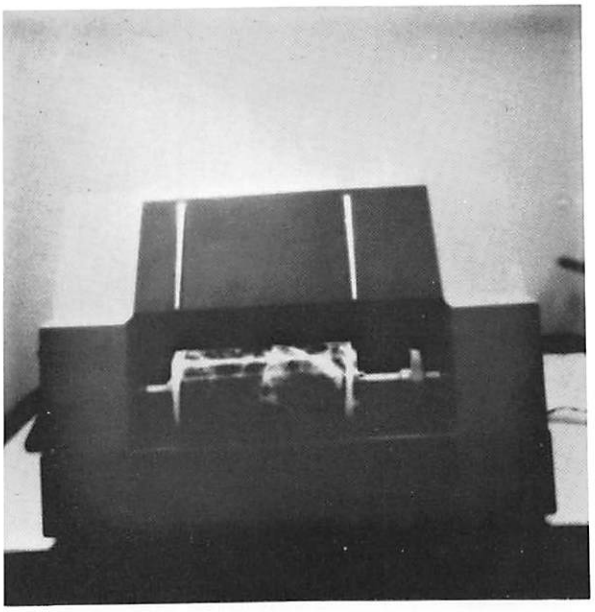


b) Left Hand View of Image Equipment Room



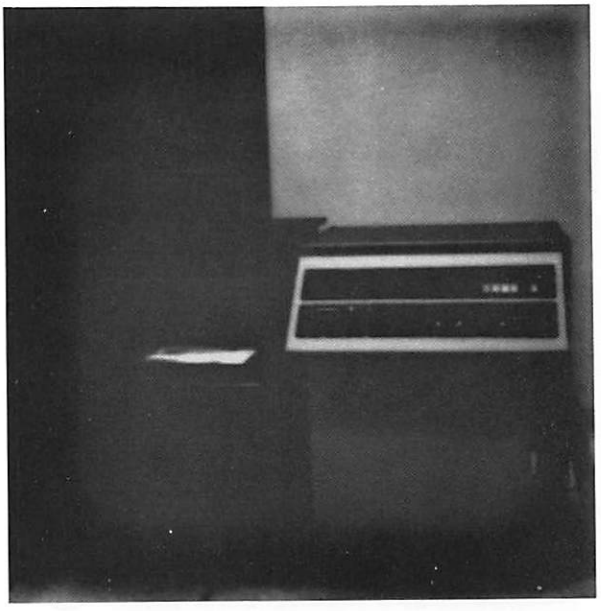
c) Right Hand View of Image Equipment Room

Figure 8-4. Facilities of the Digital Process Room and Image Equipment Room.



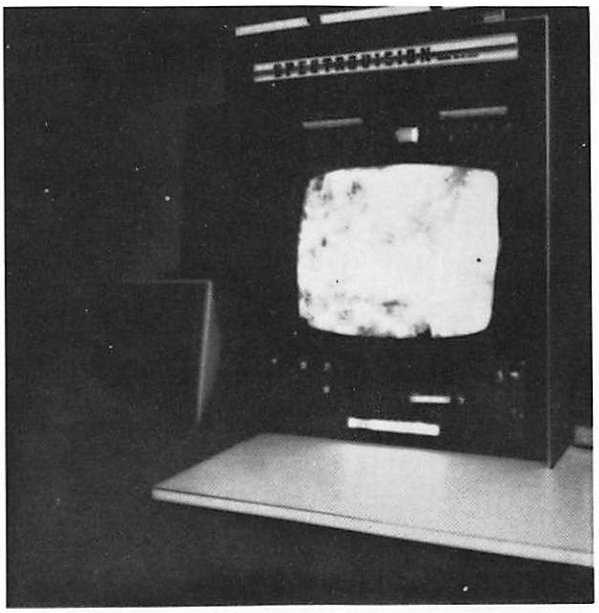
a) Muirhead Simultaneous Color Scanner

b) IER Flying Spot Scanner

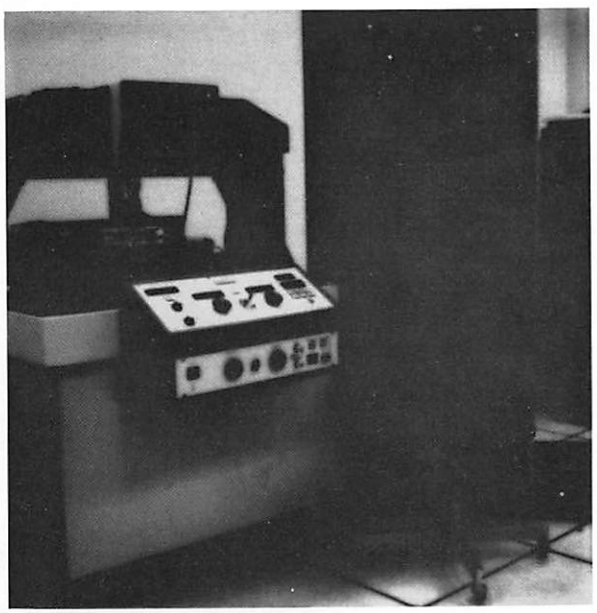


c) Dicomed 14" x 17" Scanner

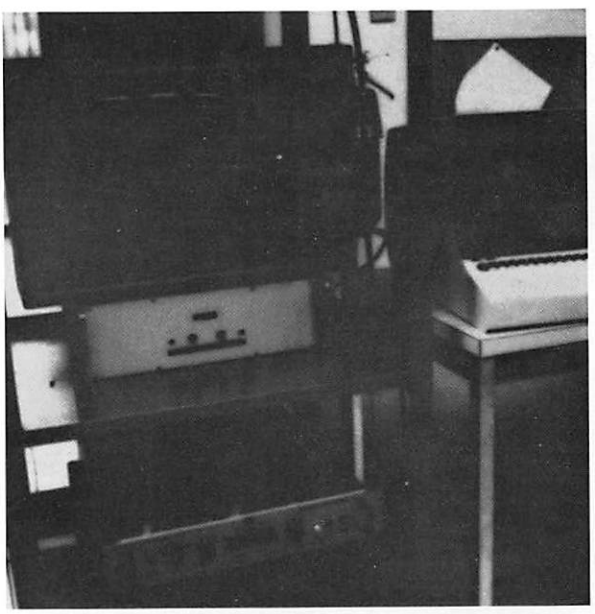
Figure 8-5. Some Devices Used for Computer-Image I/O.



a) Real Time Display



b) Flatbed Microdensitometer



c) USC Developed Display

Figure 8-6. Some Devices Used For Computer-Image I/O.

The IER flying spot scanner is used to obtain Polaroid prints and transparencies of images. Resolution is up to 1024 x 1024 pixels with 8 bits of intensity quantization. Scan times vary from 10 to 60 sec. depending upon resolution. Color photographs are obtained by sequentially scanning the white CRT phosphor through color filters.

The Muirhead color facsimile device is a two unit drum transmitter and receiver. Its capabilities include 100 line/inch resolution with simultaneous 8 bit/primary quantization. The scan time is 12 min. for an 8 in. x 10 in. print.

The Dicomed 14 in. x 17 in. X-ray and large transparency scanner is an image dissector camera mounted in a light box. It is capable of digitizing with 8 bits of accuracy and a spatial resolution of up to 2048 pixels.

The Optronics precision flat bed scanning microdensitometer scans and records 16, 35, and 70mm color or monochrome roll film on a registered transport under computer control as well as prints or transparencies up to 10 inches square. Spatial accuracy is $\pm$ 1 micron with apertures as small as 10 microns. Photographic density or transmittance is digitized with 12 bits and the display is driven by 10 bits. The scanning velocity is 8 inches/sec and the unit is capable of digitization over a 0 to 4 specular density range. Color digitization and display is performed sequentially under computer control.

Real time display of monochrome and color digital images is accomplished using an Aerojet General display device. This device utilizes a standard shadow mask CRT with 576 horizontal and 525 vertical lines of resolution. A digital disk is included with the device which makes possible a refresh rate of 60 fields/second at 64 quantization levels for each of the red, green, and blue primaries.

A real time display device designed to be plugged directly into the ARPA-TIP has been developed by the USC Image Processing Institute. A keyboard terminal controls the selection of imagery from the net to be displayed on the device.

Figure 7 contains a block diagram of a real time digital color television recording and playback system being developed by the USC Image Processing Institute. In this system the color signal output of a television camera or video tape recorder is sampled and digitized by three analog-to-digital converters at a rate compatible with U. S. television standards. The digitized signal is then recorded on the high speed Orion Titan digital tape recorder shown in figure 8. This tape recorder is capable of recording and playing back at rates of up to 10 Mbs on each of 28 data tracks for a total of 280 Mbs storage rate. A single reel of tape can record over one minute of real time television data at this rate. The recorder is also capable of playing back at a 320:1 speed reduction to match the channel rate of the PDP-11 minicomputer interface to the PDP-10. In the output mode the PDP-10 will output images for recording on the Orion at the slow rate; and the Orion will then playback at its higher speed to reproduce a reconstructed real time television signal.

8.2 Engineering Computer Laboratory

The Engineering Computer Laboratory is based upon a dual processor computer system shown in figure 9 consisting of Digital Equipment Corp. PDP Model KI10 processor for time share computing and an IBM 360/44 processor for moderate size batch computing. Both machines are host computers on the ARPA computer network. The PDP-10 utilizes the TENEX operating system while the 360/44 runs under a USC developed operating system that incorporates the VICAR image processing software system developed at the Jet Propulsion Laboratory. The PDP-10 and 360/44 are connected together and to the ARPANET-TIP through a PDP-11/40 communications controller. The controller provides character manipulation for ARPANET transmission and interfacing to the University Computer Center IBM 370/158, the Image Processing Laboratory image digitization and display devices including the real time color TV digitizer and recorder, and to the Adage AGT-10 computer graphics system.

The Image Processing Institute is developing a front end image processing system (FEIS) on the PDP-10 which will be available to any

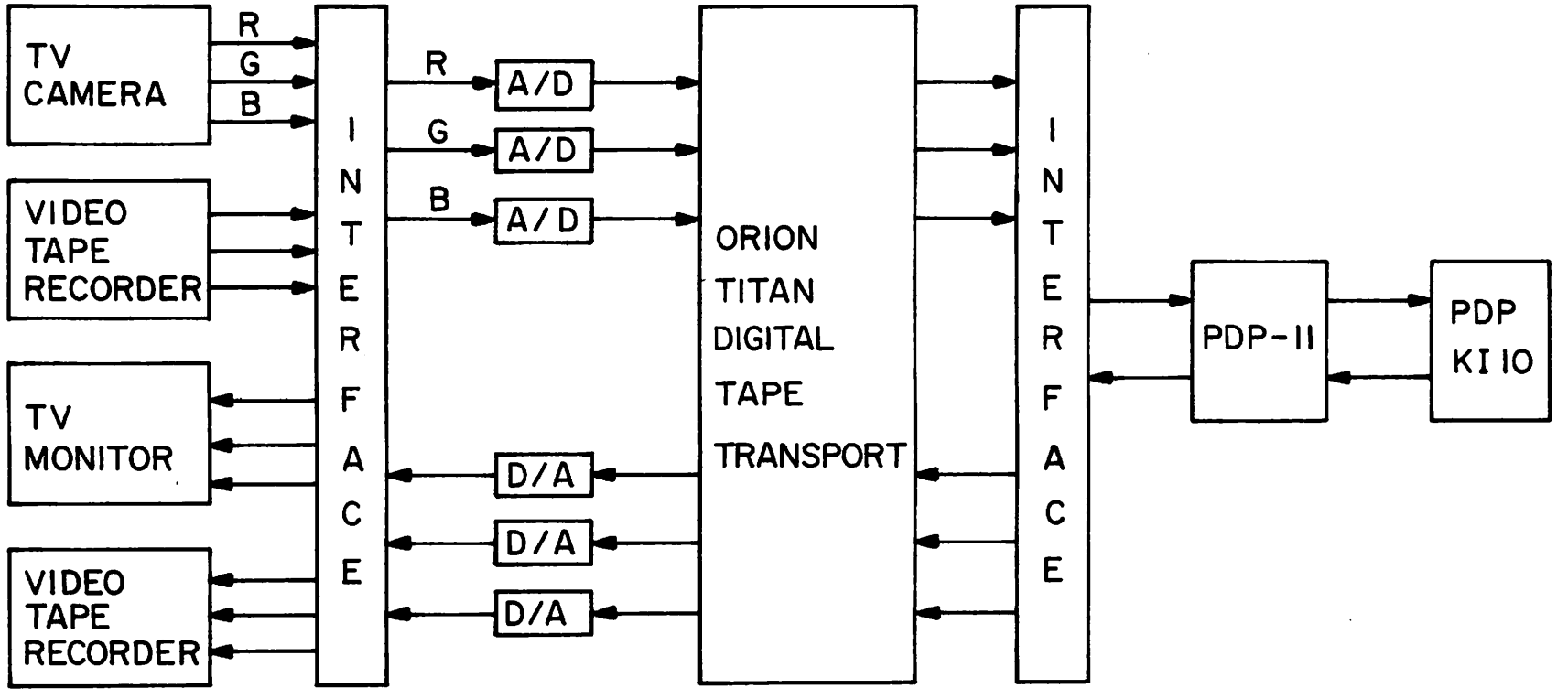


Figure 8-7. High speed television digitization and display system

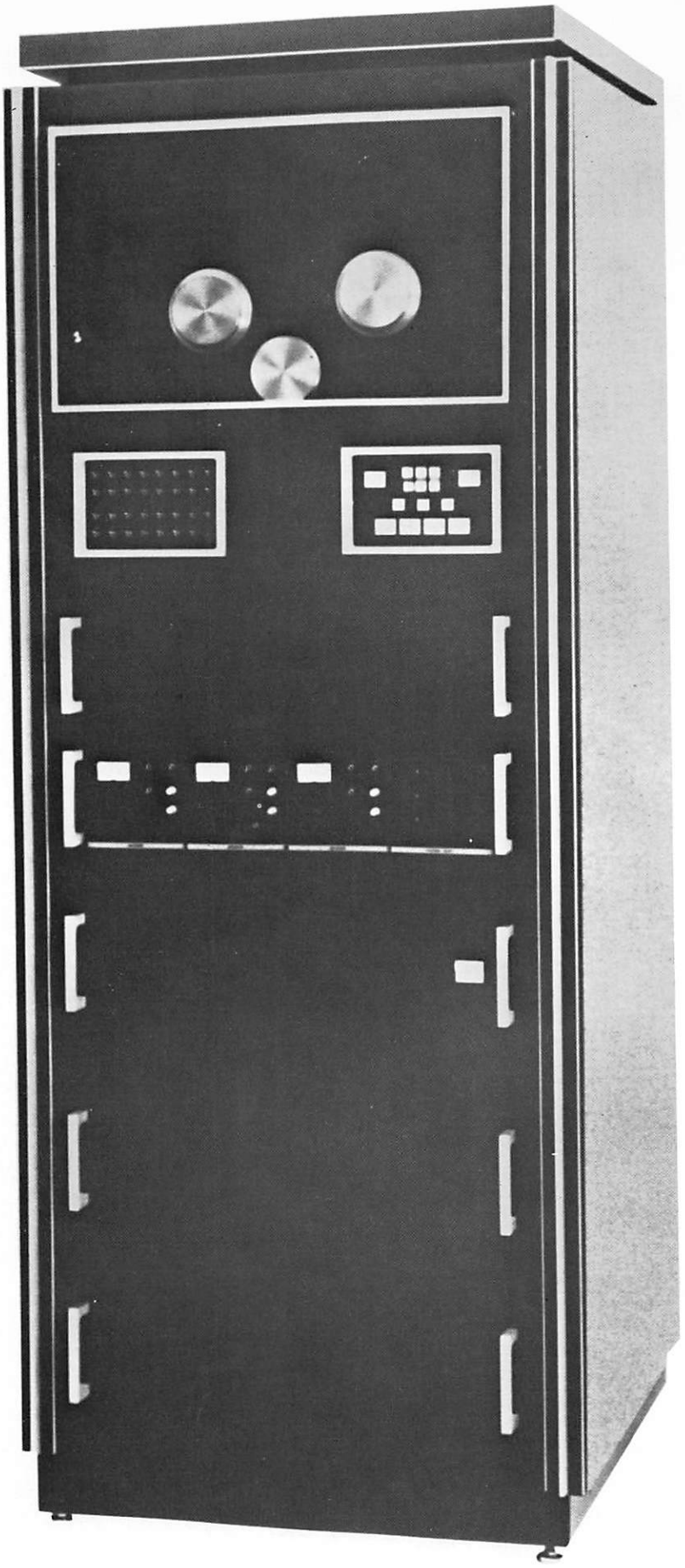


Figure 8-8. Orion Titan Digital Tape Recorder.

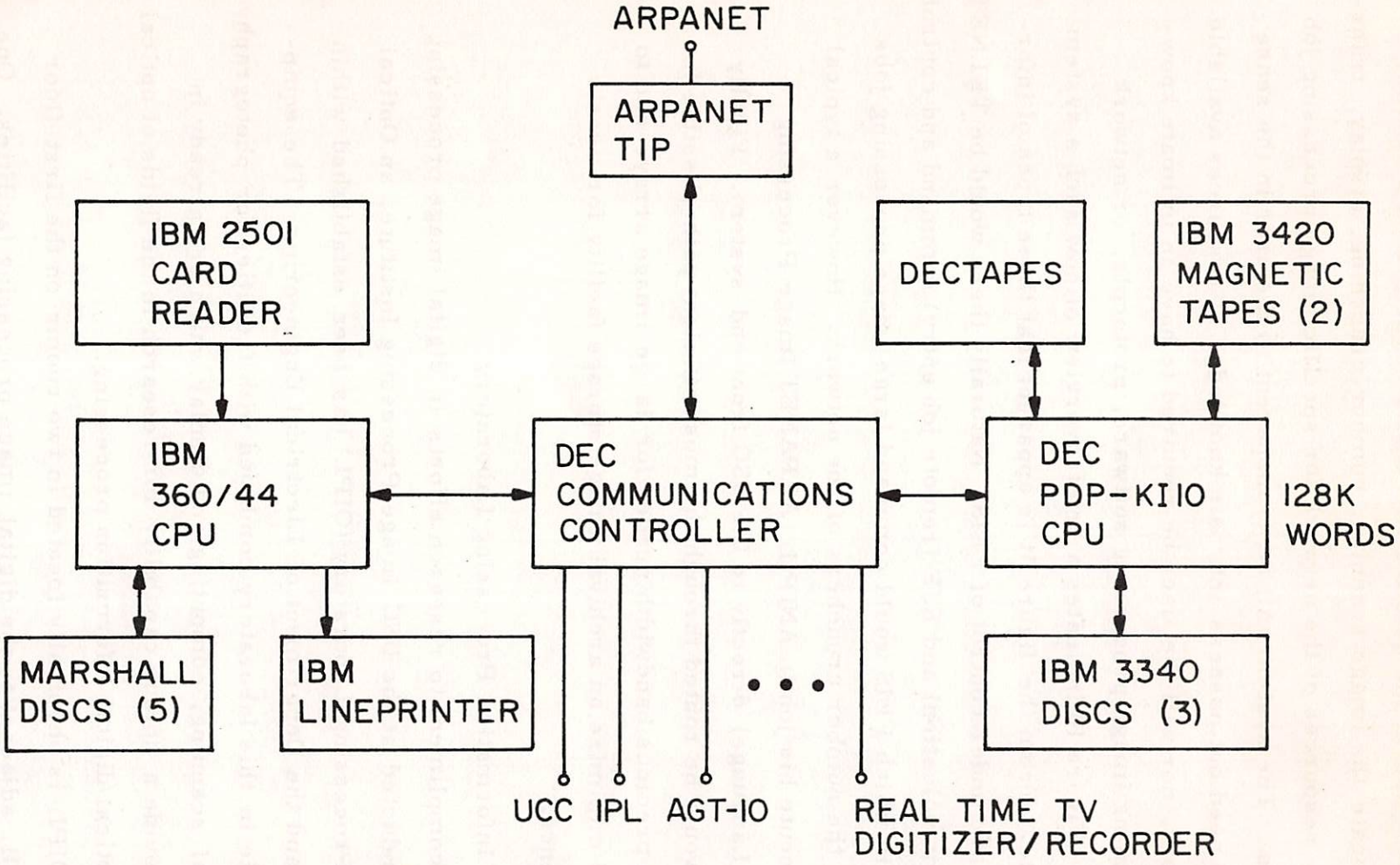


Figure 8-9. USC Engineering Computer Laboratory computational facilities.

image processing user at USC or over the ARPANET. Specifically FEIS will make use of the existing ARPANET structure on a dynamic basis to allocate the image scanning, number crunching, display, transporting, etc. resources of the network for specific image processing job requirements. The system will be transparent to the user in the sense that the user need not possess intimate knowledge of resources available on the network, nor will the user be required to have an intimate knowledge of existing image processing software, protocols, or network procedures. Figure 10 indicates a brief overview of how such a system might operate. From the figure it is apparent that three types of information flow are under control of FEIS. Naturally there would be TELNET (network communication) and RJE (remote job entry) command and control instructions in which FEIS would command large image processing jobs to be run on the number crunchers of the network. However a typical user may execute his job in ANIPIL (ARPANET Image Processing Interpretive Language) directly to the USC front end system. Finally image flow would be routed through the most efficient path possible to preserve the precious bandwidth needed for large image arrays and to establish and minimize an archival image storage facility for user's future reference.

8.3 Optical Information Processing Laboratory

As a compliment to research efforts in digital image processing now being conducted at the USC Image Processing Institute, an Optical Information Processing Laboratory (OIPL) has been established within the Institute and the Department of Electrical Engineering. The equipment available in this laboratory combined with facilities for photography and for digital scanning, computing and display which are already in operation provide a unique capability for research in the fields of optical and hybrid optical/digital information processing.

The OIPL is physically located in two rooms on the first floor of Powell Hall, adjacent to the digital image processing facilities. One room contains a steel topped honeycomb optical stable table, 4' x 10' x 8" in size, along with a computer-controlled film transport and ROSA

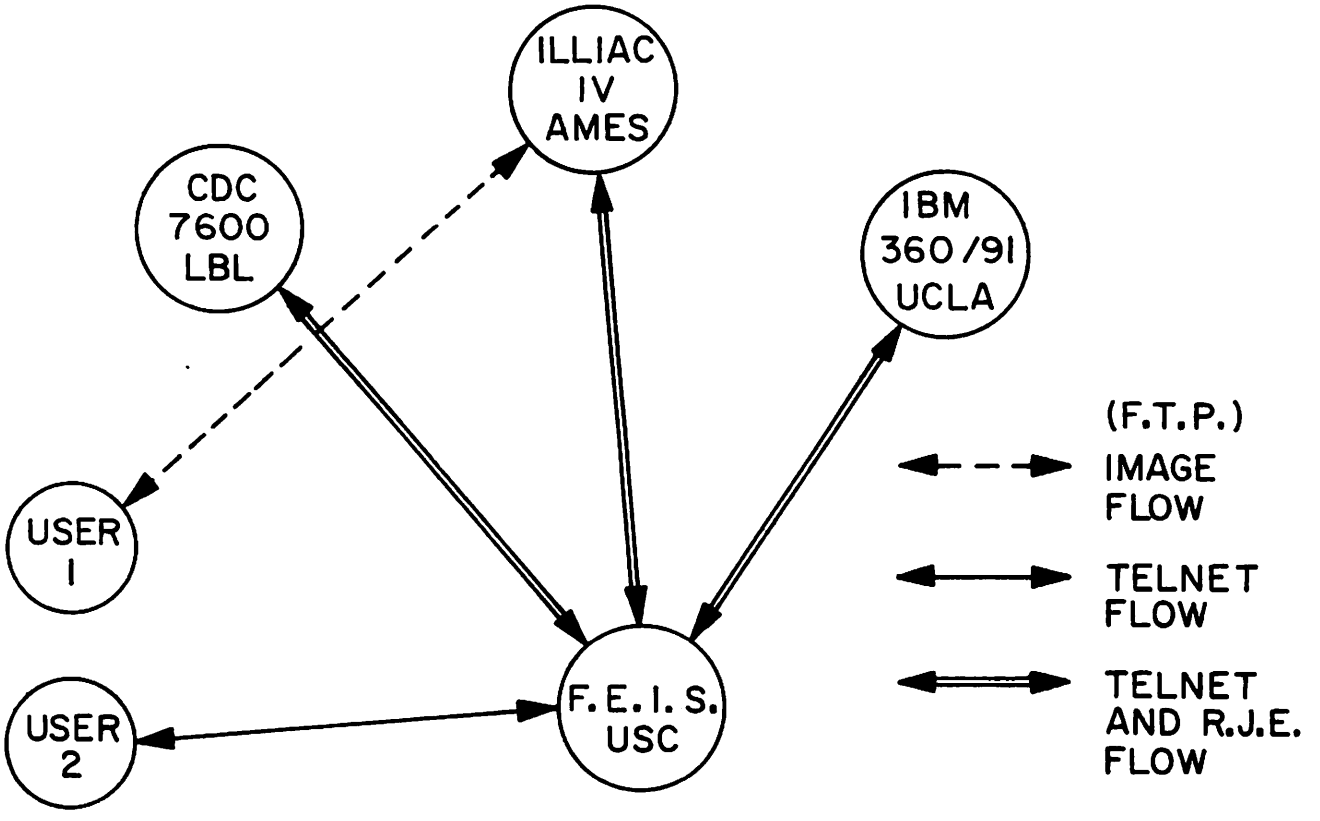


Figure 8-10. ARPA picture - typical mode of operation.

(recording optical spectrum analyzer) photodetector and data collection system. This optical/digital pattern recognition system also includes a 15 mW, He-Ne laser source, mirrors, collinating optics, pinhole spatial filter, positioners and mounts which can be rearranged into different system configurations.

Another room, located at the rear of the IPI computer room, is entirely devoted to general research in optical/digital processing. The room is light-tight, and different types of experimental arrangements are possible. Three optical tables are available; one is a 4' x 12' x 12" steel-honeycomb table; the others are granite surface plate tables 4' x 4' x 8" and 2' x 8' x 12". Both granite tables have innertube inflatable suspension systems for vibration isolation in interferometric systems. The lab has both 15 mW and a 50 mW He-He coherent laser sources, 1 m and 3 m optical rails and mounts, and a limited assortment of positioners, mirrors, lenses, flats, prisms, rulings, collinators and other optical equipment for experimentation. Measuring equipment includes transmission and reflection densitometers, a radiometer/photometer and an electrical shutter and timer. Photographic facilities for processing the high resolution plates and films generally used in optical systems are now being developed.

All these facilities can be utilized and adapted for research in real-time optical/digital systems making use of fast controllable deflectors, modulators, and detectors. Several projects now underway make use of optical masks or screen produced digitally on IPI displays. Although work in these areas is still at an early stage, research at these facilities should develop the potential of optical, digital, and hybrid information processing.

8.4 Biomedical Image Processing Laboratory

The Biomedical Image Processing Laboratory has been established in augmentations of the Image Processing Laboratory to provide research facilities in this specific area of research. Presently, the Laboratory contains an Adage Model AGT-10 computer graphics system and

several computer terminals connected to the TIP. The Adage possesses stand-alone graphics capability for interactive graphics computing, and is also connected to the IBM 360/44 computer for operations requiring greater computer power.

A Recognition Systems Incorporated computer/laser optical processing system shown in figure 11 is installed in the Laboratory. This unit contains a coherent imaging apparatus that produces a Fraunhofer diffraction light pattern of a photographic transparency. The diffraction pattern is sensed by an optical detector containing 64 elements that sample spatial frequency and amplitude components of the light pattern. The 64 outputs are then processed by analog circuitry and fed to the HP-2100 of the Image Processing Laboratory for real time data reduction. Similar equipment has been used quite successfully by USC staff members for photographic texture detection in the development of automatic equipment for the diagnosis of coal miner's pneumoconiosis and is directly applicable to reconnaissance and other DoD related imagery.

Figure 8-11. Biomedical Image Processing Laboratory Equipment.

9. Publications

The following is a list of papers, articles, and reports published or accepted for publication during the past six months and which have resulted from ARPA sponsored research.

H. C. Andrews, "Digital Image Restoration-Survey," IEEE Computer, Vol. C-23, No. 2, February, 1974.

H. C. Andrews and A. G. Tescher, "The Role of Adaptive Phase Coding in Two and Three Dimensional Fourier and Walsh Image Compression," Proc. 1974 Walsh Function Symposium, Washington, D. C., March, 1974.

H. C. Andrews and C. L. Patterson, "Digital Images as a Teaching Aid for Singular Value Decomposition," American Mathematical Monthly, to be published.

H. C. Andrews, Chapter 4, "Two Dimensional Transforms," Digital Picture Processing, T. S. Huang, ed., Springer Verlag, to be published.

H. C. Andrews, book review of "Introduction to Diffraction, Information Processing, and Holography," by F. T. S. Yu for IEEE Transactions on Communications, Vol. COM-22, No. 5, May, 1974, pp. 732-733.

L. D. Davisson and R. M. Gray "Recent Advances in Data Compression," in Advances in Communication Systems, Academic Press, to appear.

W. Frei, "A Model of Color Vision for Image Processing," Proc. 1974 Picture Coding Symposium, Goslar, West Germany, Aug. 26-28, 1974.

A. Habibi, "Hybrid Coding of Pictorial Data," IEEE Transactions on Communications, Vol. COM-22, No. 5, May 1974, pp. 614-624.

A. Habibi, R. S. Hershel, "A Unified Representation of DPCM and Transform Coding Systems," IEEE Transactions on Communications, Vol. COM-22, No. 5, May 1974, pp. 692-696.

A. Habibi, G. S. Robinson, "A Survey of Digital Picture Coding," IEEE Computer, May 1974, pp. 22-34.

A. Habibi, W. K. Pratt, et.al., "Real Time Image Redundancy Reduction Using Transform Coding Techniques", Proceedings of 1974 International Conference in Communications, June, 1974.

A. Habibi, F. Naderi, "Post Processing of Encoded Images in Fourier Domain," 1974 Picture Coding Symposium, Goslar, West Germany, August, 26-28.

E. L. Hall, R. P. Kruger and A. F. Turner, "An Optical Digital System for Automatic Processing of Chest X-Rays," Optical Engineering, Vol. 13, No. 3, May, 1974, pp. 250-257.

E. L. Hall, W. O. Crawford, Jr. and F. E. Roberts, "Computer Classification of Pneumoconiosis from Radiographs of Coal Workers," IEEE Transactions on Biomedical Engineering, to be published.

E. L. Hall, W. O. Crawford, and F. E. Roberts, "Moment Measurements for Computer Texture Discrimination in Chest X-Rays," IEEE Transactions on Computers, to be published.

A. K. Jain and E. Angel, "Image Restoration, Modelling, and Reduction of Dimensionality," IEEE Transactions on Computers, May, 1974.

N. D. A. Mascarenhas and W. K. Pratt, "Digital Image Restoration Under a Regression Model," IEEE Transactions on Circuits and Systems, to be published.

N. E. Nahi and A. Habibi, "Nonlinear Recursive Image Enhancement," to appear in IEEE Transactions on Circuits and Systems (Special Issue on Image Processing).

W. K. Pratt, "Correlation Techniques of Image Registration," IEEE Transactions on Aerospace and Electronic Systems, May, 1974.

W. K. Pratt, W. H. Chen, and L. R. Welch, "Slant Transform Image Coding," IEEE Transactions on Communications, Vol. COM-22, No. 8, August, 1974.

W. K. Pratt and M. N. Huhns, "Quantization Error Reduction for Transform Image Coding," 1974 Picture Coding Symposium, Goslar, West Germany, August 26-28, 1974.

W. K. Pratt, "Dual Mode DPCM/Deltamodulation Image Coding System for Real Time Television Transmission," 1974 Picture Coding Symposium, Goslar, West Germany, August 26-18, 1974.

A. A. Sawchuk, "Optical and Digital Information Processing," Optical Engineering, Vol. 13, No. 3, pp. 172-174, May-June, 1974.

A. A. Sawchuk, "Point-Spread Functions and Image Restoration with Astigmatism and Curvative of Field," J. Opt. Society America, Vol. 64, No. 4, p. 560, April, 1974, Spring Meeting 1974, Optical Society of America, Washington, D. C.

M. Stein and W. Frei, "Preservation of Combined Perfusion and Inhalation Lung Scan Information by Continuous Color Function Mapping," Proc. First World Congress of Nuclear Medicine, Tokyo, Japan, 1974.

A. G. Tescher, H. C. Andrews, and A. Habibi, "Adaptive Phase Coding in Two-and Three-Dimensional Fourier and Walsh Image Compression," 1974 Picture Coding Symposium, Goslar, West Germany, August 26-28, 1974.