USCIPI REPORT 530

SEMIANNUAL TECHNICAL REPORT

Covering Research Activity During the Period
1 September 1973 through 28 February 1974

William K. Pratt
Project Director
(213) 746-2694

Image Processing Institute
University of Southern California
University Park
Los Angeles, California 90007

31 March 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

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)*
Image Processing Institute
University of Southern California, University Park,
Los Angeles, California 90007

2a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

2b. GROUP

3. REPORT TITLE
IMAGE PROCESSING RESEARCH

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Technical Semiannual, 1 September 1973 to 28 February 1974

5. AUTHOR(S) *(First name, middle initial, last name)*
William K. Pratt (Project Director)

6. REPORT DATE
31 March 1974

7a. TOTAL NO. OF PAGES
157

7b. NO. OF REFS
50

8a. CONTRACT OR GRANT NO.
F08606-72-C-0008

b. PROJECT NO.
ARPA Order No. 1706

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)
USCIPI Report 530

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 February 1973 to 28 February 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 September 1973 to 28 February 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
Dennis Smith
George Soen
Wai Szeto
John Tahl
Florence B. Tebbets

Students

Ben Britt
Steve Dashiell
Faramarz Davarian
Roy M. Glantz
Henry Glazner
Michael Huhns
Steve Hui
Mohammad Jahanashahi
Mohammad Khalil
San Kwoh
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
Carl Wedberg
Pamela Welch

TABLE OF CONTENTS

1. Research Project Overview 1
2. Research Project Activities 2
3. Image Coding Projects 3
    - 3.1 Image Coding via Two Dimensional Interpolative Representations 4
    - 3.2 Optimal Logarithmic Quantization for Picture Processing 11
    - 3.3 Image Data Compression Using Spline Functions 19
    - 3.4 Positive Extrapolation of Signals and Images 25
    - 3.5 Transform Domain Spectrum Interpolation 28
    - 3.6 Variable Rate Image Coding for Sources with Unknown Probabilities 38
4. Image Restoration and Enhancement 42
    - 4.1 Space Variant Point Spread Function Pseudoinversion 43
    - 4.2 Pseudoinverse Image Restoration by Transform Domain Processing 50
    - 4.3 Modelling Superposition Integrals for Image Restoration 57
    - 4.4 Spline Function Restorations 64
    - 4.5 Space-Variant Restoration of Astigmatism and Curvature of Field 68
    - 4.6 Fast-Suboptimal Wiener Filtering of Markov Processes 75
    - 4.7 Color Image Restoration by Linear Estimation Methods 80
    - 4.8 Film Recording of Color Images from a Television Monitor 85
    - 4.9 Psycho-Visual Pseudo-Color Mapping of Imagery 91
5. Image Data Extraction Projects 95
    - 5.1 An Optical-Digital System for Feature Extraction From Images 95
    - 5.2 A Terminal/Timeshare Based Device for Interactive Boundary Tracing and Analysis 103
    - 5.3 Reconstruction of Three-Dimensional Object Arrays From Their Projections 107
    - 5.4 Application of Kalman Filtering to the Reconstruction of Images from Noisy Projections 118
    - 5.5 Boundary Estimation of Statistical Objects in Noise 122

6. Image Analysis Projects 128
    6.1 A New Model of Color Vision and Some Practical Implications 128
7. Image Processing Hardware and Software Projects 144
    7.1 Real Time ARPANET Digital Image Display 144
    7.2 Real Time Color Image Digital Recorder and Display 147
    7.3 Software Progress Report 149
8. Publications 151

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 September 1973 to 28 February 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. Section 8 is a list of publications by project members.

2. Research Project Activities

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

Alexander A. Sawchuk has been appointed editor of the May-June 1974 issue of Optical Engineering, the journal of the Society of Photo-Optical Instrumentation Engineers. This is a special issue devoted to optical and digital image and information processing, and will contain more than eleven papers discussing theory and applications of such systems. Topics in this issue include: hybrid optical/digital systems using real-time input/output devices; digital holograms for optical processing; and image processing of synthetic aperture data.

Harry C. Andrews has been appointed guest editor of the May 1974 issue, of Computer, the journal of the IEEE Computer Society. The issue, entitled "Computer Image Processing," consists of seven papers spanning picture coding, image restoration, and digital image processing facilities at Aerospace, EG&G, and AEC facilities. The issue will include photographic examples of computer processed images both in black and white as well as in color. In addition to normal circulation, the issue is being overprinted for additional distribution at the NCC conference in Chicago.

3. Image Coding Projects

The research effort in image coding has been directed toward a wide variety of applications. 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.

In the first report an interpolative data representation is utilized to develop three image coding algorithms. One algorithm is based upon spatial domain coding, another upon transform domain coding, and the third is a hybrid coding scheme. Computational requirements for the algorithms are specified and the image performance is evaluated for several pictorial examples.

The next report describes the analysis of logarithmic quantization scales for monochrome image quantization. Quantization errors are evaluated in terms of a model of the human visual process.

Spline functions, which are a special class of truncated polynomials, are known to be quite accurate for the approximation of one dimensional functions. Their use in image approximation for purposes of bandwidth reduction is explored in the next report.

In the following two reports, the concept of transform domain spectrum extrapolation and interpolation for image coding is investigated. With these techniques, transform coefficient quantization error can be reduced by post-processing at the coder with a significant improvement in image quality.

The last report considers extensions to the universal coding concept. In particular, a rate distortion bound is established for coding image sources with unknown probabilities.

3.1 Image Coding via Two Dimensional Interpolative Representations

Anil K. Jain

For finite discrete signals, non-causal "interpolative" representations may be used for coding. These non-causal representations lead to three different coding algorithms in the spatial, hybrid and frequency domains.

Interpolative Modeling For simplicity in presentation, only the first order stationary Markov signal will be considered. Let $\{u_i\}$, $i = 0, 1, \ldots, N, N+1$ represent such a signal with zero mean and autocorrelation given by

$$E[u_i u_j] = \rho^{|i-j|} \tag{1}$$

Representation of a sequence $\{u_i\}$ denotes a relationship

$$\mathcal{L}\{u_i\} = v_i \tag{2}$$

such that the sequence $\{u_i\}$ can be reconstructed from v_i. For example

$$u_{i+1} - \rho u_i = v_i \tag{3}$$

is the Markov representation of eq. (1) with

$$\beta_i^2 \overset{\Delta}{=} E[v_i^2] = (1-\rho^2) \tag{4}$$

It can be shown that the linear minimum mean square representation of eq. (1) is given by

$$u_i - \frac{\rho}{1+\rho^2}(u_{i+1}+u_{i-1}) = v_i \qquad i = 1, \ldots, N \tag{5a}$$

$$u_0 - \rho u_1 = v_0 \tag{5b}$$

$$u_{N+1} - \rho u_N = v_{N+1} \tag{5c}$$

This representation is such that $E[v_i^2]$ is minimum compared to all other linear representations and is given by

$$E[v_i^2] \triangleq \beta_2^2 = \frac{(1-\rho^2)}{(1+\rho^2)} \tag{6}$$

(For a discussion of the correlation properties of v_i and generalization of the above statement to non-stationary case see [1].) Also observe that $\beta_2^2 < \beta_1^2$ since $0 < \rho < 1$.

For a two dimensional zero mean discrete signal u_{ij} with

$$E[u_{ij}u_{i+n,j+m}] = \rho^{|n|+|m|} \tag{7}$$

"horizontal" and "vertical" representations are defined as

$$\epsilon_h(i,j) = u_{ij} - a(u_{i,j+1}+u_{i,j-1}) \tag{8a}$$

$$\epsilon_v(i,j) = u_{ij} - a(u_{i+1,j}+u_{i-1,j}) \tag{8b}$$

where $a = \rho/(1+\rho^2)$. Then the representation

$$\epsilon_{ij} = u_{ij} - \frac{a}{2}(u_{i+1,j}+u_{i-1,j}+u_{i,j+1}+u_{i,j-1}) \tag{9}$$

is such that $E[\epsilon_h^2+\epsilon_v^2]$ is minimized, and

$$E[\epsilon_{ij}^2] = \frac{(1-\rho^2)}{(1+\rho^2)^2} \tag{9a}$$

This is in contrast with the Markov representation of eq. (7) given by

$$u_{ij} = \rho u_{i-1,j} + \rho u_{i,j-1} - \rho^2 u_{i-1,j-1} + e_{ij} \tag{9b}$$

with

$$E[e_{ij}^2] = (1-\rho^2)^2 \tag{9c}$$

Comparison between eqs. (9a) and (9c) shows that $E[\epsilon_{ij}^2] < E[e_{ij}^2]$ for $0 < \rho \leq 0.786$. For values of $\rho \simeq 1$, the two values have small mean square difference. Also, in eq. (9), the coefficient 'a' is relatively insensitive to changes in image statistics (ρ parameter). In fact

$$\frac{da}{a} = \frac{(1-\rho^2)}{(1+\rho^2)}\left(\frac{d\rho}{\rho}\right) \qquad (10)$$

Therefore in the vicinity of $\rho = 1$, small changes in image statistics will not alter the performance of the interpolative representation.

Coding and Reconstruction Algorithms It should be recognized that if the ϵ_{ij} become observables in eq. (9), then this equation has to be solved (for reasons of stability) as a boundary value problem with end conditions $u_{0,j}$, $u_{N+1,j}$, $u_{i,0}$ and $u_{i,N+1}$ known. For simplicity, it will be assumed here that these quantities are zero, i.e., the picture under consideration is imbedded in a dark background. One might question the validity of this assumption in view of the stationary statistics of eq. (7); however, this assumption is actually non-essential for the coding algorithms below. Equation (9) can then be rewritten in vector form as

$$\frac{a}{2}u_{j+1} - Qu_j + \frac{a}{2}u_{j-1} = -\epsilon_j , \qquad (11)$$

where u_j and ϵ_j now represent $N \times 1$ column vectors of elements $(u_{1j}, \ldots, u_{Nj})$ and $(\epsilon_{1j}, \ldots, \epsilon_{Nj})$ respectively. The matrix Q is a symmetric tridiagonal matrix given by the elements

$$Q_{ij} = \begin{cases} 1 & i = j \\ -a/2 & |i-j| = 1 \\ 0 & \text{otherwise} \end{cases} \qquad (12)$$

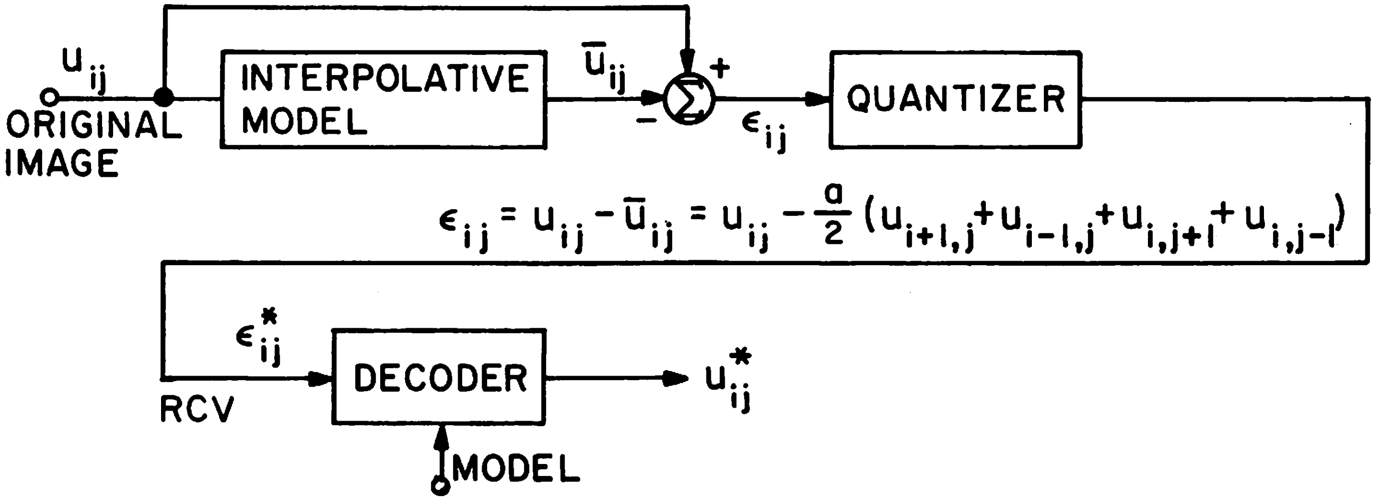
Algorithm A1: (Spatial Domain Coding, Figure 1a)

1. Quantize and code ϵ_i, after obtaining it through eq. (9). Let $\epsilon^*_{i,j}$ denote the received signal.
2. The reconstructed sequence u^*_j is obtained by solving

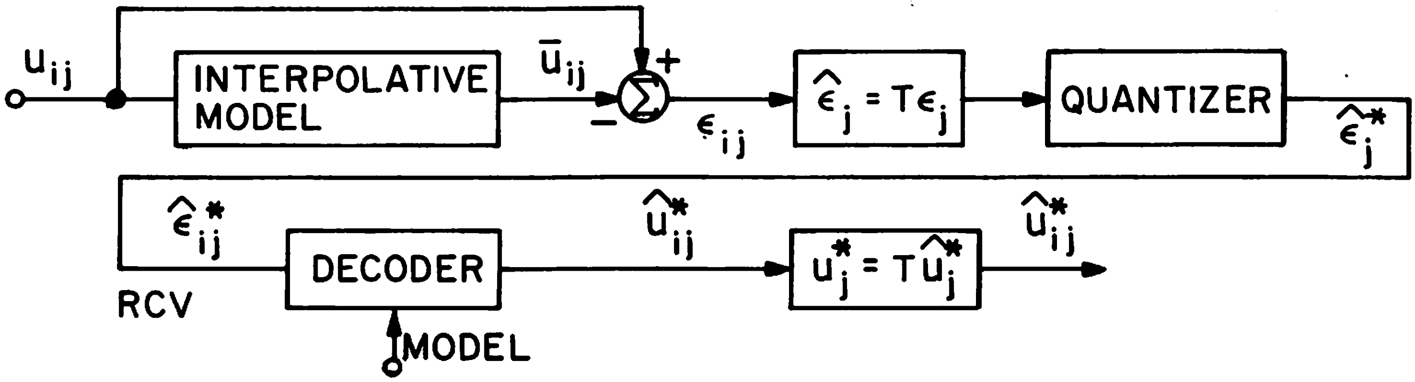
$$u^*_{j+1} = R_j u^*_j + s_j \qquad u_0 = 0 \qquad (13a)$$

$$R_{j-1} = \frac{a}{2}\left(Q - \frac{a}{2}R_j\right)^{-1} \qquad R_N = 0 \qquad (13b)$$

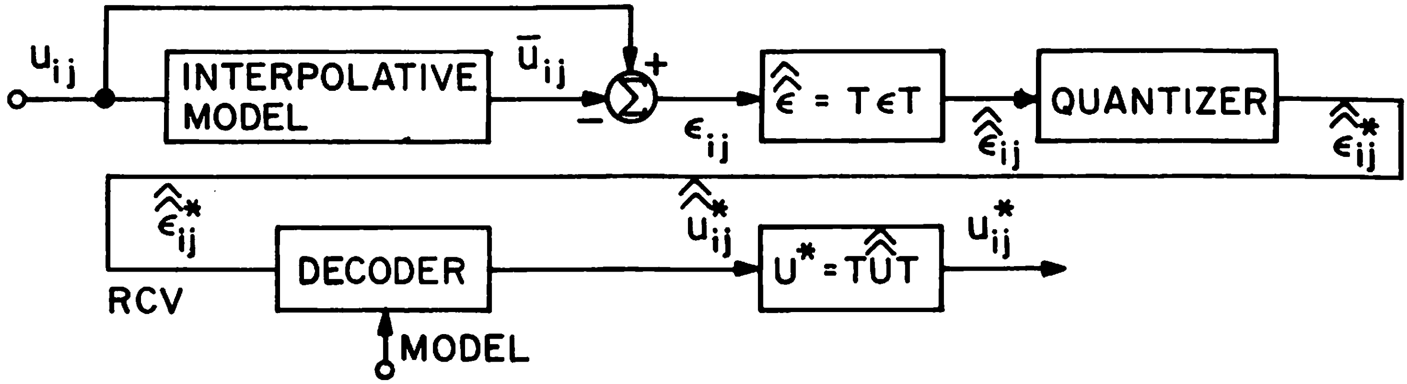
$$s_{j-1} = R_{j-1}\left(\frac{2}{a}\epsilon^*_j + s_j\right) \qquad s_N = 0 \qquad (13c)$$



a) Spatial domain coding algorithm A1



b) Hybrid domain coding algorithm A2



c) Transform domain coding algorithm A3

Figure 3.1.1 Image coding algorithms.

Eq. (13) can be solved in about $(N^2 \log_2 N)$ computations, employing the structure of matrix Q. [1, 2, 3]

Algorithm A2: (Hybrid Domain Coding, Figure 1b)

1. Find $\hat{\epsilon}_j \overset{\Delta}{=} T\epsilon_j$, where T is an N × N matrix containing $\sqrt{\frac{2}{N+1}}\,\sin\left(\frac{ij\pi}{N+1}\right)$ terms. It can be shown [3] that T is idempotent and diagonalizes Q i.e.

$$TQT = \Lambda = \left\{1 - a\cos\frac{i\pi}{N+1}\right\} \tag{14}$$

2. Quantize $\hat{\epsilon}_{ij}$ with the number of bits allocated for the i^{th} row being proportional to $\log\left(1 - \cos\frac{i\pi}{N+1}\right)^{-1}$, to obtain $\hat{\epsilon}^*_{ij}$.
3. The reconstructed image u^*_{ij} is obtained by

$$u^*_j = T\hat{u}^*_j \tag{15a}$$

$$\hat{u}_{ij+1} = r_{ij}\hat{u}_{ij} + s_{ij} \qquad \hat{u}_{i,0} = 0 \tag{15b}$$

$$r_{ij-1} = \frac{a}{2}\left(\lambda_i - \frac{a}{2}r_{ij}\right)^{-1} \qquad r_{i,N} = 0 \tag{15c}$$

$$s_{ij-1} = r_{i,j-1}\left(\frac{2}{a}\hat{\epsilon}^*_{ij} + s_{ij}\right) \qquad s_{iN} = 0 \tag{15d}$$

and $\qquad \lambda_i \overset{\Delta}{=} 1 - a\cos\frac{i\pi}{N+1}$

Again eq. (15) requires about $(N^2 \log_2 N)$ computations.

Algorithm A3: (Transform Domain Coding, Figure 1c)

1. If ϵ is N × N matrix of element ϵ_{ij}, then first find

$$\hat{\hat{\epsilon}} = T\epsilon T \tag{16}$$

2. Quantize $\hat{\hat{\epsilon}}_{ij}$ by allocating n_{ij} bits to it such that

$$n_{ij} \alpha \log(\mu_i + \mu_j)^{-1} \tag{17}$$

where $\qquad \mu_i \overset{\Delta}{=} 1 - 2a\cos\left(\frac{i\pi}{N+1}\right) \qquad i = 1, \ldots, N\,. \qquad (17a)$

3. The reconstructed image is given by

$$\{u^*_{ij}\} = U^* = T\hat{\hat{U}}^* T \tag{18}$$

and

$$\hat{u}_{ij} = \frac{2\hat{\hat{\varepsilon}}^*_{ij}}{\mu_i+\mu_j} \tag{18a}$$

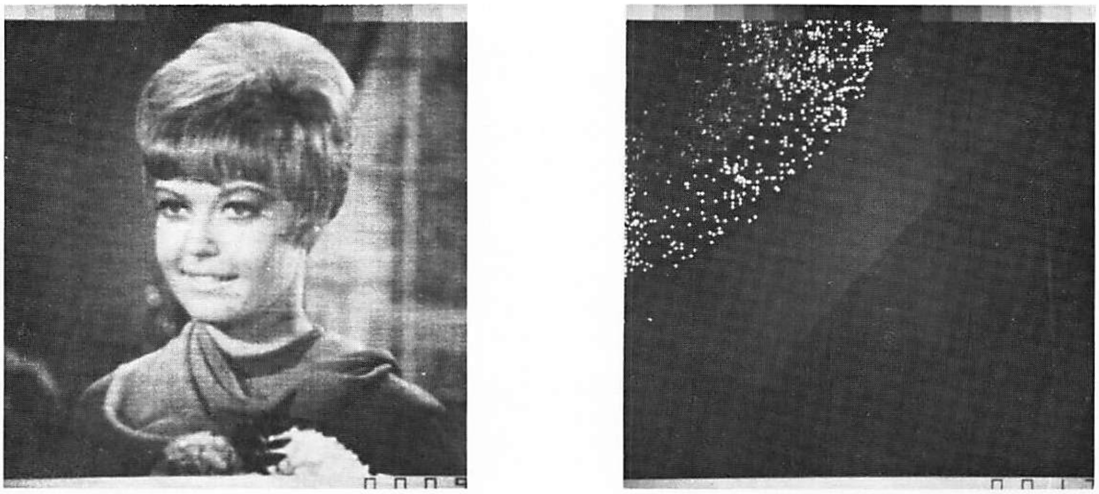
References [1, 2] provide details and generalizations.

Examples, Implementation and Computational Considerations Algorithms A2 and A3 have been simulated for the 256 × 256 pixel girl image of fig. 2a. The average value of ρ for this picture is 0.96. In all the simulations the actual value used was $\rho = 1$. The difference in ε_{ij} for these two values, visually and quantitatively (in terms of $\sum \varepsilon^2_{ij}$) both was insignificant. Figure 2b shows the display of $|\varepsilon_{ij}|$. Figure 2c contains the encoded image according to algorithm A2 with 3 bits/pixel on the average using a uniform quantizer. The entropy of the quantized signal $\hat{\varepsilon}^*$ was actually 2.35 bits, so that a variable length Huffman code could be employed to obtain the same image with 2.35 bits/pixel. Figure 2d shows the quantized $|\hat{\hat{\varepsilon}}^*_{ij}|$ signal according to algorithm A3 and bit rate of 1 bit/pixel. Figures 2e and 2f show the encoded images for 1 and 1.47 bits/pixel.

It can be shown that the total computational load in each algorithm is of the same order. In algorithm A1, the memory and computational requirements on the transmitter are minimal and the major computational burden is at the receiver. In algorithm A2, the total processing burden is roughly divided in a 1:2 ratio (the transmitter needs to take a one sided transform and the receiver solves scalar interpolative equations and takes an inverse transform). In algorithm A3, the processing load is roughly equally divided between the transmitter and the receiver. Thus the three algorithms obtained via a single representation spell out three different communication system architectures, and their relative use therefore depends on the particular application. In this sense the interpolative representation leads to a unification of some of the different methods of image coding employed currently.

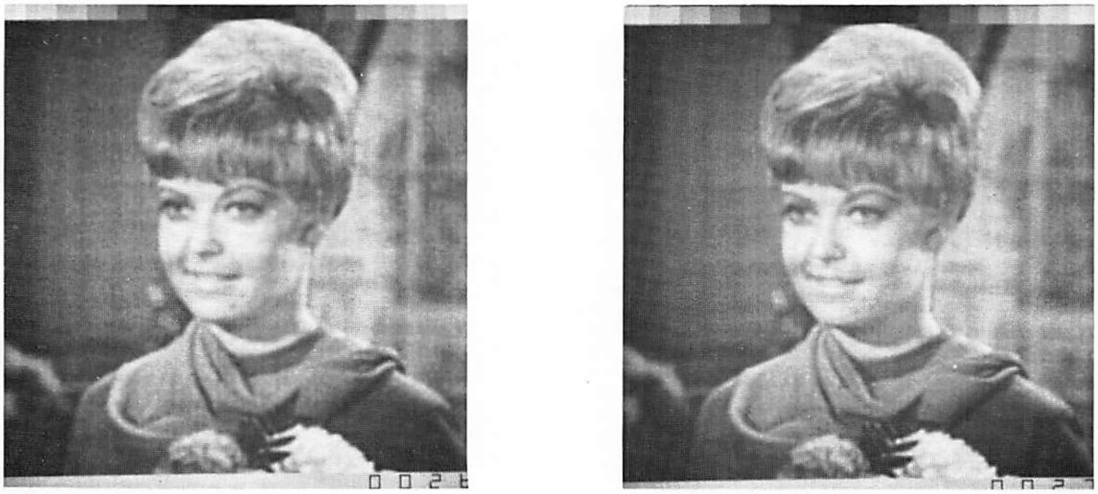
(a) original image

(b) interpolative signal $|\epsilon|$



(c) hybrid domain encoded image

(d) quantized signal $|\hat{\hat{\epsilon}}^{*}|$



(e) encoded image 1 bit/pixel

(f) encoded image 1.47 bits/pixel

Figure 3.1-2. Image coding results via interpolative model.

Finally, it should be remarked that the representation used here is but one member of a class of similar non-causal and semicausal representations. The representation reported here corresponds to a discrete version of the Poisson equation ($\nabla^2 u=\epsilon$). The relative merits of other representations is currently under study and will be reported in the future.

References

1. A. K. Jain, "A Theory of Non-Causal Representations for Finite Discrete Signals," (to appear).

2. A. K. Jain, "Image Modelling for Unification of Transform and DPCM Coding of Two-Dimensional Images," National Electronics Conference, Chicago, Illinois, Oct. 1973.

3. A. K. Jain and E. Angel, "Image Restoration, Modelling and Reduction of Dimensionality," (to appear in the IEEE Transactions on Computers).

3.2 Optimal Logarithmic Quantization for Picture Processing

Francis Kretz and Werner Frei

An optimal quantization law for image intensities for television monitor display of digitized and processed images has been considered. Also, the effects of "brightness" adjustments for television displays has been analyzed.

Subjective Criterion A possible subjective criterion for quantization distortion is to postulate that the decision and reconstruction levels should be perceptually equi-distant. Since it is well known that the perception of intensity is a concave, monotomically increasing function of light intensity $\Lambda(I)$, an optimal quantizer in the above sense can be derived from $\Lambda(I)$ as

$$Q(I) = \text{int}[(2^N-1)\Lambda(I/I_{max})] \qquad (1)$$

where $\text{int}[\cdot]$ denotes the nearest integer ($0 \leq Q \leq 2^N-1$) of the argument, N is the number of bits of the quantizer, and $\Lambda(I)$ is normalized so that

$$\Lambda(0) = 0, \quad \Lambda(1) = 1$$

Various functions for $\Lambda(I)$ such as the square and cube root, polynomials and logarithms have been proposed to fit experimental perceptual data [1]. The parameters and range of validity of these functions depend very much upon the experimental conditions under which the data was obtained. In particular, background illumination has a strong influence on $\Lambda(I)$.

In the case of television displays, the range of intensities is fairly well defined (about two orders of magnitude), but the background illumination for each pixel is a complex, more or less random field (the image itself). Since one desires to design one quantizer for all pixels of an image, the parameters of $\Lambda(I)$ will be the result of some compromise.

Experiments have been carried out to determine the optimal slope at the origin of the Λ-function

$$\Lambda(I) = b \log_{10}(1 + I/a) \qquad (2)$$

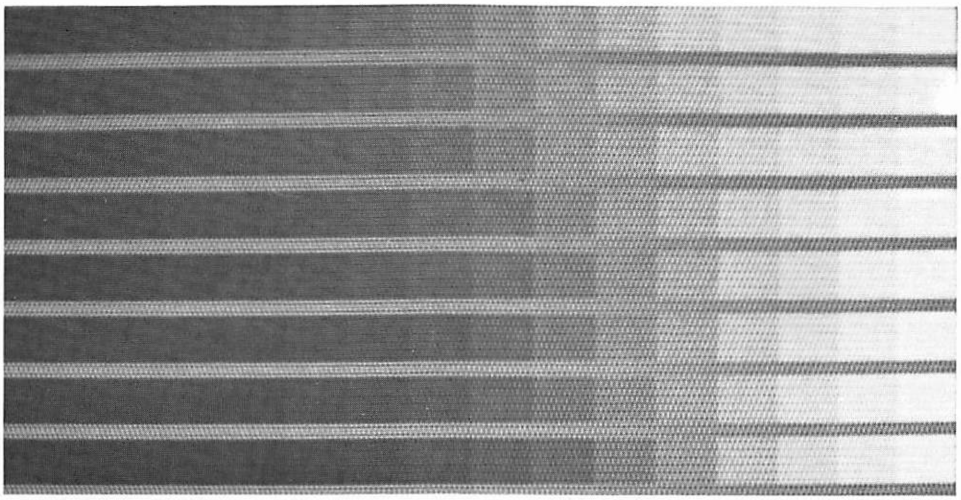
where

$$b = (\log_{10}(1 + a))^{-1}$$

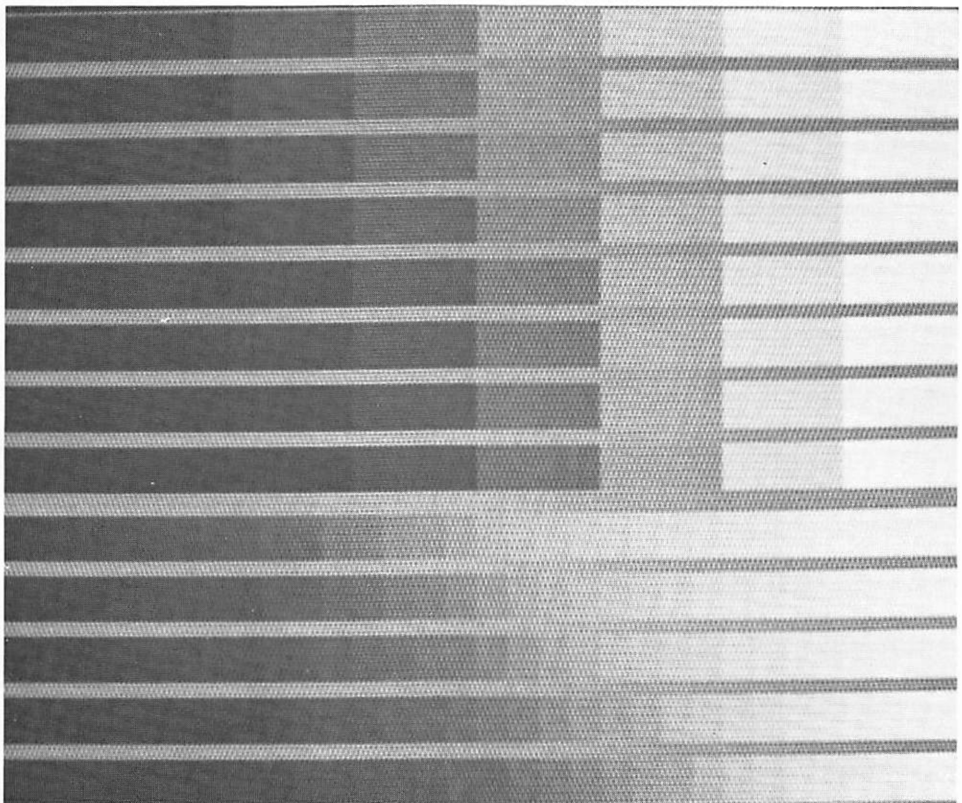
with $\Lambda(0) = 0$ and $\Lambda(1) = 1$. For this Λ-function the slope at the origin is

$$\Lambda'(0) = \frac{b}{a \log_e(10)}$$

In the first set of experiment, a ramp of intensities was quantized according to eqs. (1) and (2), generating a set of grey scales with $\Lambda'(0) =$ 4.4, 6.3, 8.0, 11.3, 16.1, 19.7, 22.9, 29.9 (from top to bottom, respectively in figure 1a (N = 4) and figure 1b (N = 3 and 5 bits). Then the parameter $\Lambda'(0)$ was chosen corresponding to the scale with the most uniform subjective spacing of intensities over the entire range. It should be pointed out that the figures reproduced here have been subjected to a series of distortions inherent to the lithographic process. Several observers viewing the TV monitor preferred $\Lambda'(0) = 16$ (slightly steeper than the Richter scale, see [1]).



(a) 16 - level logarithmic grey scale



(b) 8- and 32- level logarithmic grey scales

Figure 3.2-1. Logarithmic grey scale.

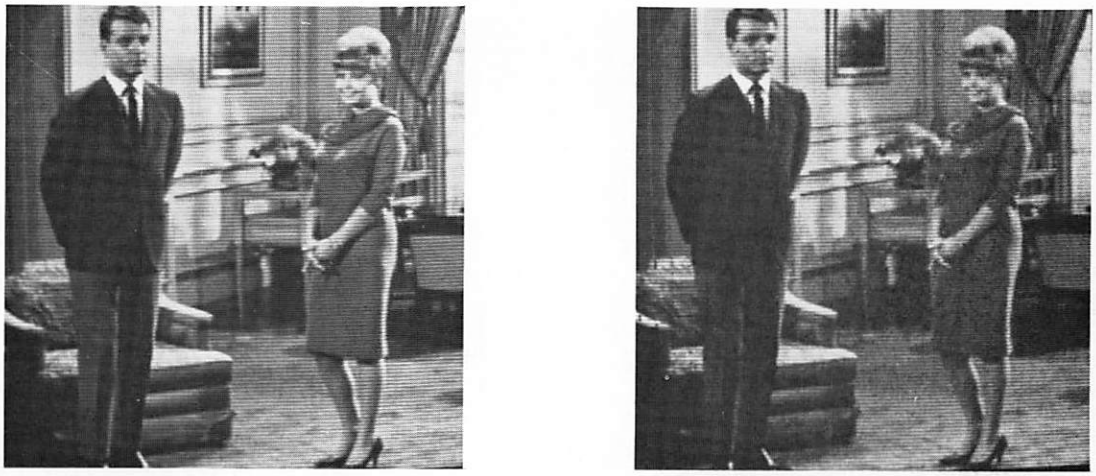
In the second experiment, two images (the SMPTE "couple" and "girl") were coarsely quantized (N = 4 bits per pixel) in order to verify both the postulate regarding the quality criterion and the optimal slope determined in the previous experiment. Figures 2 and 3 show the respective original images and a set of logarithmically quantized versions with different parameters $\Lambda'(0)$. For comparison, linearly quantized versions with four and five bits are included.

The second experiment shows that the optimal parameter $\Lambda'(0)$ depends on the picture content. Note that a small value of $\Lambda'(0)$ tends to create large subjective increments in dark areas and vice-versa. It is observed that the very small optimal value of $\Lambda'(0)$ for the "girl" picture is a consequence of the unusually rare occurrence of low intensities. In fact, the histogram of the "girl" picture has a maximum at about 20% intensity. The "couple" picture has a more typical negative exponential-like histogram and the optimal slope $\Lambda'(0) = 16$ is the same as determined in the first experiment.

Minimum square error quantization laws for sources with given amplitude probabilities have been studied [2, 3]. Assuming a probability density function of the form $p(I) = k \exp(-\alpha I)$, the MSE quantization law follows a concave monotomically increasing function of I quite similar to $\Lambda(I)$. The agreement between both criteria is presently being studied in more detail.

Practical Considerations In typical computer image processing environments, images are usually scanned and quantized linearly. It obviously makes little sense to coarsely requantize. The results of this study are therefore primarily relevant to scanner or coding hardware.

Comparing the quality of linearly and logarithmically quantized images, one sees that one bit at least out of five, possibly two out of eight can be saved with appropriate quantization, which represents a 20-25% bandwidth reduction, or storage saving, whichever is relevant.

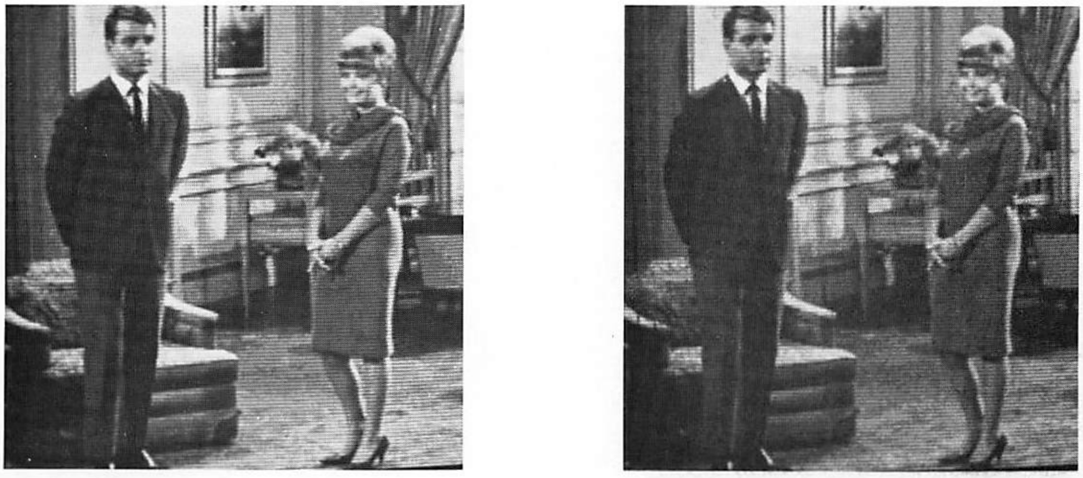


(a) original

(b) PCM 5 bits linear

(c) PCM 4 bits linear

(d) PCM log. 4 bits $\Lambda'(0)=6.3$



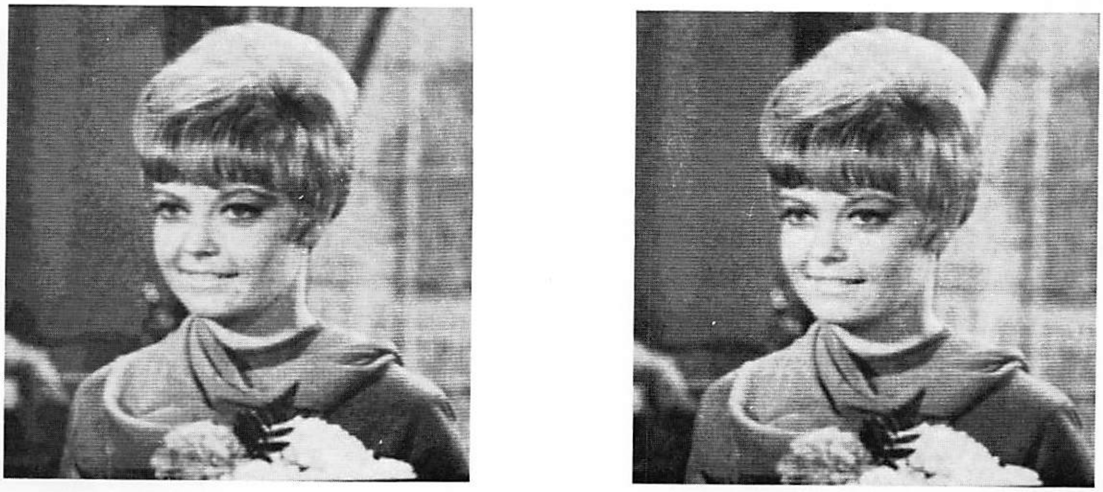
(e) PCM log. 4 bits $\Lambda'(0)=16.1$

(f) PCM log. 4 bits $\Lambda'(0)=22.9$

Figure 3.2-2. Examples of grey scale quantization of "couple" picture.

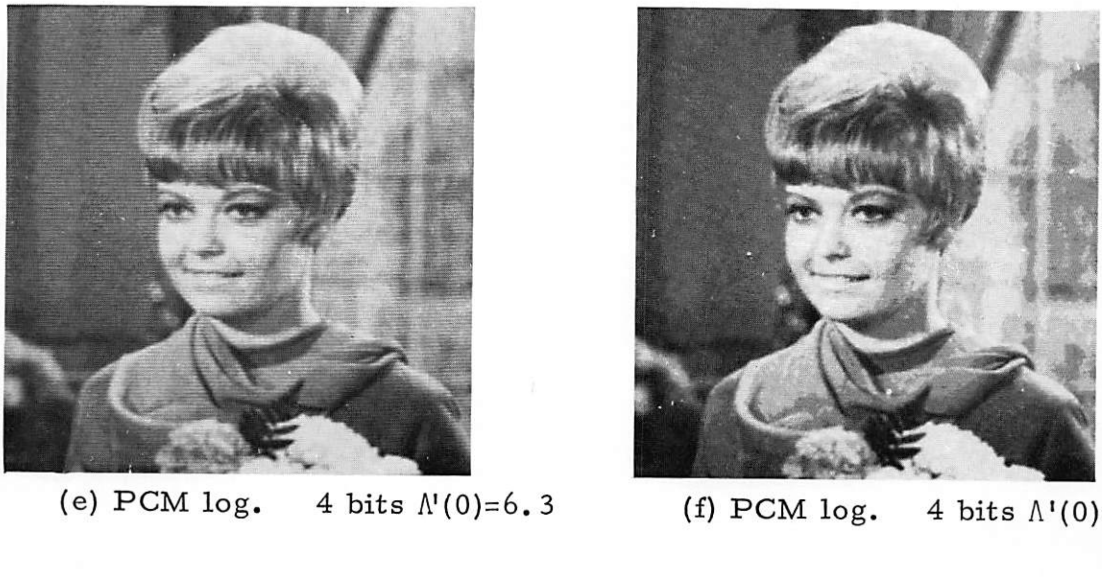
(a) original

(b) PCM 5 bits linear



(c) PCM 4 bits linear

(d)PCM log. 4 bits $\Lambda'(0)=4.4$



(e) PCM log. 4 bits $\Lambda'(0)=6.3$

(f) PCM log. 4 bits $\Lambda'(0)=16.1$

Figure 3.2-3. Examples of grey scale quantization of "girl" picture.

In the context of hardware considerations, the influence of the "brightness" control of a television monitor has been analyzed. Let u be the signal applied to the grid of the cathode ray tube (CRT), U_0 the CRT cut-off voltage, U_1 the bias voltage (controlled by the "brightness" setting) and v the intensity to be displayed. The light intensity of the CRT is proportional to

$$I = (u + U_0 + U_1)^{\gamma} \qquad u+U_0+U_1 > 0$$

$$I = 0 \qquad u+U_0+U_1 \leq 0$$

The γ correction is carried out by letting

$$u = k_1 v^{1/\gamma}$$

where k_1 is a constant. An incorrect setting of the brightness control $U_0+U_1 = \Delta U \neq 0$ does not simply reduce the useful dynamic range of the display. It also upsets the linearity of the gamma-corrected transfer function

$$I = (k_1 v^{1/\gamma} + \Delta U)^{\gamma}$$

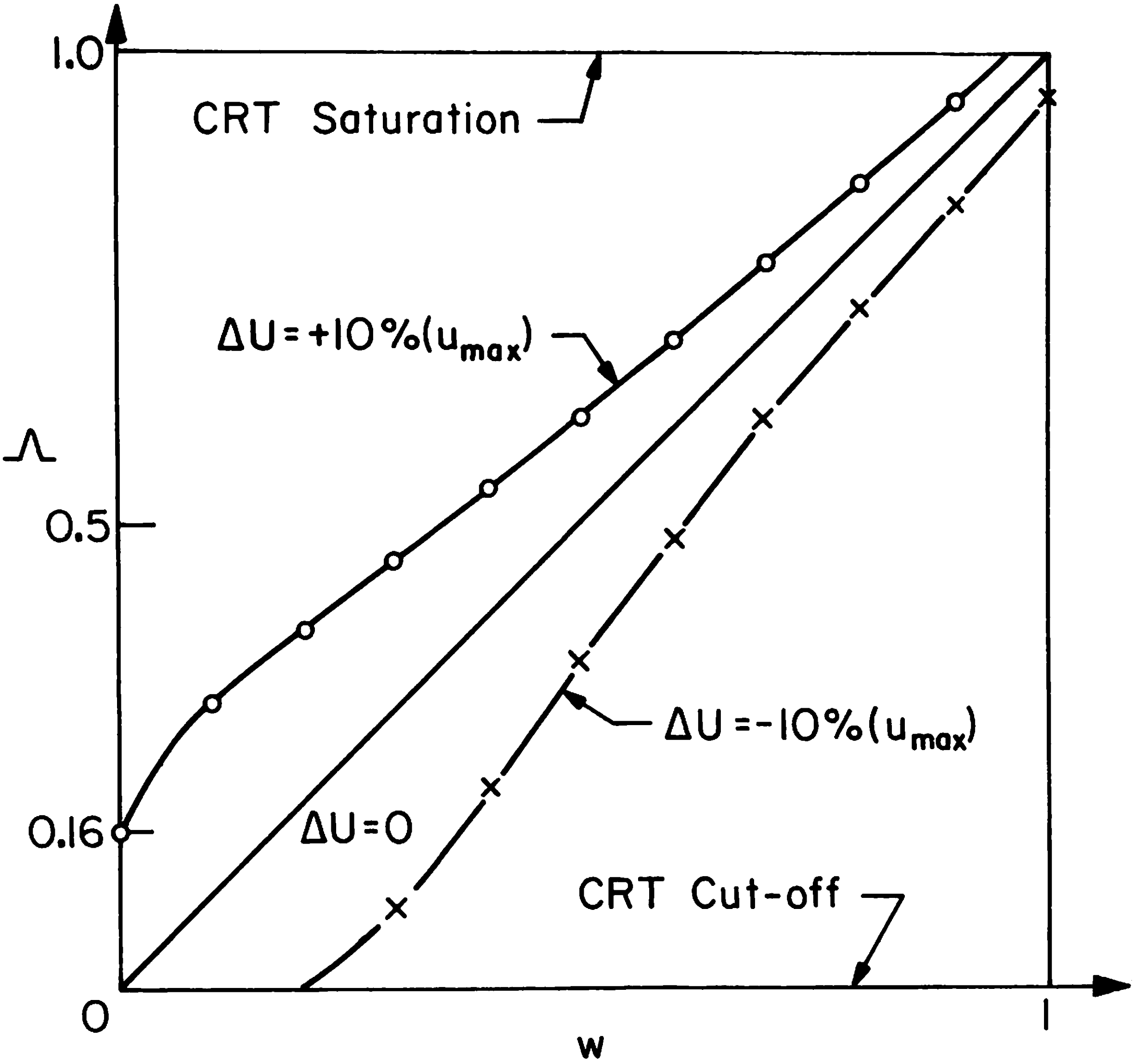
Since the eye is very sensitive to errors at low intensities, the effect of ΔU is quite severe for the lower levels. Figure 4a shows the overall transfer function $\Lambda(w)$ where w represents the physiologically companded intensities

$$v = a \left[\exp\left(\frac{w}{b}\right) - 1\right]$$

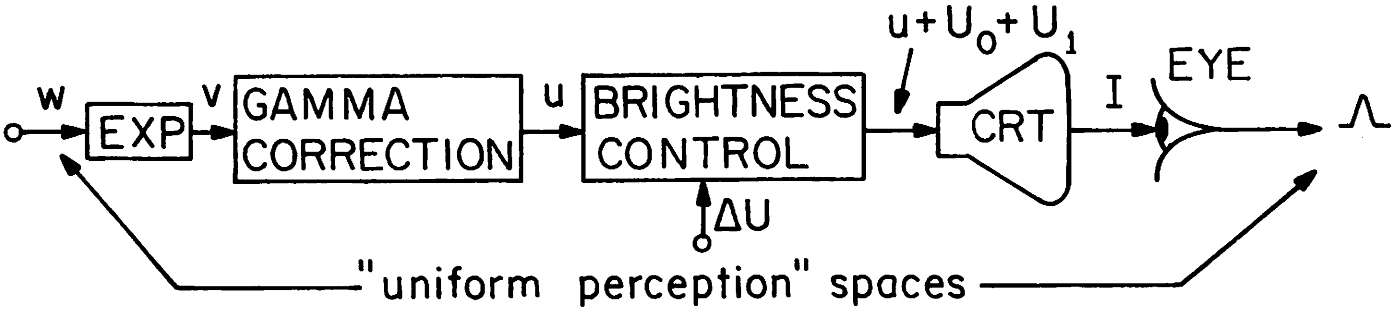
and

$$\Lambda = b \log(1 + I/a)$$

as shown in figure 4b. These results indicate the importance of brightness adjustments. A well designed hardware system should have an automatic video clamping circuit controlled by the D/A converter, such that ΔU is set to zero when a digital zero is read.



(a) influence of brightness control on overall perceptual system.



(b) perceptual system block diagram.

Figure 3.2-4. Subjective effect of brightness control.

References

1. G. Wysecky and W. S. Stiles, Color Science, Wiley and Sons, 1967.
2. Max, "Quantization Distortion of PCM with Non-uniform Spacing of Levels," Proc. IRE, Vol. 39, No. 1, pp. 44-48, Jan. 1951.
3. Algazi, "Useful Approximation to Optimal Quantization," IEEE Trans. on Communications, Vol. COM-14, pp. 297-301, June 1966.

3.3 Image Data Compression Using Spline Functions

Faramaz Davarian

Spline functions can be used to simultaneously compress and interpolate a given set of data. Among different sets of bases spanning the space of splines, the B-splines are most suitable, since they possess a local basis property and result in matrices which are easily invertable.

A dimensionality and subsequent bandwidth reduction can be achieved by a least squares fit of n points using m ($m \leq n$) basis functions. In essence, the data compression method is simply a transformation of the n-dimensional space of data points into a smaller m-dimensional spline space. It is noteworthy that elements of the m-dimensional spline space will directly generate the continuous estimate of the original signal rather than the sampled estimate. The method is described below along with a study of the statistical properties of the transform domain.

Definition of Spline Functions Given a strictly increasing sequence of real numbers, $t_1, t_2, \ldots, t_n$, a spline function $S(x)$ of degree m with knots $t_1, t_2, \ldots, t_n$ is a function defined on the real line having the following two properties:

i) In each interval (t_i, t_{i+1}) for $i = 0, 1, \ldots, n-1$, $S(x)$ is given by some polynomial of degree m or less.

ii) $S(x)$ and its derivatives of order $1, 2, \ldots, m-1$ are continuous on $[t_1, t_n]$.

Thus, a spline function is a piecewise polynomial function satisfying certain conditions regarding continuity of the function and its derivatives.

It is generally believed that in many circumstances a spline function is a more adaptable approximating function than a polynomial. This is based in part on actual numerical experience, and in part on mathematical demonstrations that solutions of a variety of problems of best approximation turn out to be spline functions. A spline function may be defined in terms of a truncated power function

$$(x-t)_+^{\ell} = \begin{cases} (x-t)^{\ell} & (x > t) \\ 0 & (x \leq t) \end{cases}$$

It is easily seen that [1] any spline of degree 1 with knots $t_1, t_2, \ldots, t_n$ has a unique representation of the form

$$S(x) = P(x) + \sum_{j=1}^{n} c_j (x - t_j)_+^{\ell}$$

where P(x) is a polynomial of degree ℓ or less. The above representation of a spline function normally results in an ill conditioned set of linear equations. To overcome this deficiency it is possible to introduce a new set of local basis functions for the space of splines [2].

The preceeding considerations lead to the B-splines studied by Schoenberg [3], which are in a sense, the splines of minimal support for a given degree (consisting of the smallest possible number of intervals between knots). Figure 1 illustrates the typical shape of a B-spline function.

Data Compression by Least Squares Method Given a set of data pairs (t_i, y_i) for $i = 1, 2, \ldots, n$, which can be interpolated as digitized values of the points of the graph $y = f(t)$, let the unknown function f(x) be approximated by a linear combination of suitably chosen functions $M_1(t), M_2(t), \ldots, M_m(t)$ which are the basis splines.* Then

$$f(t) = c_1 M_1(t) + c_2 M_2(t) + \ldots + c_m M_m(t)$$

---

*$M_1, M_2, \ldots, M_m$ form a complete set of bases for the space of m data points (m dimensional). This basis can interpolate m data elements exactly.

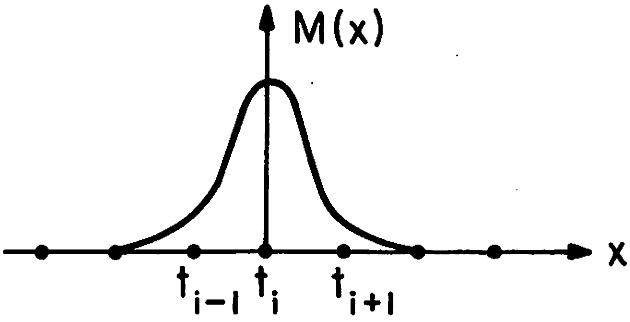


Figure 3.3-1. Typical shape of a B-spline.

where the unknown coefficients $c_1, c_2, \ldots, c_m$ are independent parameters to be determined, and $m \leq n$. To minimize the mean square approximation error

$$Q = \sum_i [F(t_i) - y_i]^2 = \left[\sum_i \sum_j c_j M_j(t_i) - y_i\right]^2$$

differentiate Q with respect to a_k and set the result to zero to obtain

$$\frac{\partial Q}{\partial c_k} = 2\sum_i \left\{\sum_j c_j M_j(t_i) - y_i\right\} M_k(t_i) = 0 \qquad k = 1, 2, \ldots, m$$

or

$$\sum_j c_j \sum_i M_j(t_i) M_k(t_i) - \sum_i y_i M_k(t_i) = 0$$

In matrix form

$$\left[\sum_i M_i M_k\right] [c_k] = \left[\sum_i y_i M_k(t_i)\right]$$

or

$$\underline{B}^T \underline{B}\,\underline{C} = \underline{B}^T \underline{y}$$

where

$$\underline{y} = [y_1, y_2, y_3, \ldots y_n]^T$$

$$\underline{C} = [c_1, c_2, \ldots, c_m]^T$$

$$\underline{B}^T = \begin{bmatrix} M_1(t_1) & M_1(t_2) & \ldots\ldots\ldots\ldots & M_1(t_n) \\ M_2(t_1) & M_2(t_2) & \ldots\ldots\ldots\ldots & M_2(t_n) \\ \vdots & & & \\ \vdots & & & \\ M_m(t_1) & \ldots\ldots & \ldots\ldots\ldots\ldots & M_m(t_n) \end{bmatrix}$$

with $M_i(t_{i+j}) = 0$, for $|j| \geq \rho$. Thus, $\underline{B}$ has many zero value off diagonal entries. The vector of weighting coefficients is then

$$\underline{C} = \underline{B}^{+}\underline{y}$$

where

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

is the pseudoinverse matrix of $\underline{B}$. Note that n elements of data vector $\underline{y}$ are mapped into m elements of the $\underline{C}$ vector, which, represents the coefficients of m spline basis functions.

The estimate of f(t) is then

$$\hat{f}(t) = c_1 M_1(t) + c_2 M_2(t) + \ldots + c_m M_m(t)$$

$$\hat{f}(t_i) = \sum_j c_j M_j(t_i) \qquad i = 1, \ldots, n$$

Let

$$\hat{\underline{y}} = [\hat{F}(t_i) \ldots \hat{F}(t_n)]^T$$

$$\hat{\underline{y}} = \underline{B}\underline{C} = \underline{B}(\underline{B}^T\underline{B})^{-1}\underline{B}^T\underline{y}$$

The error vector $\underline{e}$ can then be expressed as

$$\underline{e} = \underline{y} - \hat{\underline{y}} = [\underline{I}_{n \times n} - \underline{B}(\underline{B}^T\underline{B})^{-1}\underline{B}^T]\underline{y}$$

Statistical Properties of B-Spline Coefficients If the data vector $\underline{y}$ is modeled as a sample of a vector random process with known mean, E{y}, and known covariance, $\underline{K}_y$, the B-spline coefficients given by

$$\underline{C} = \underline{B}^{+}\underline{y}$$

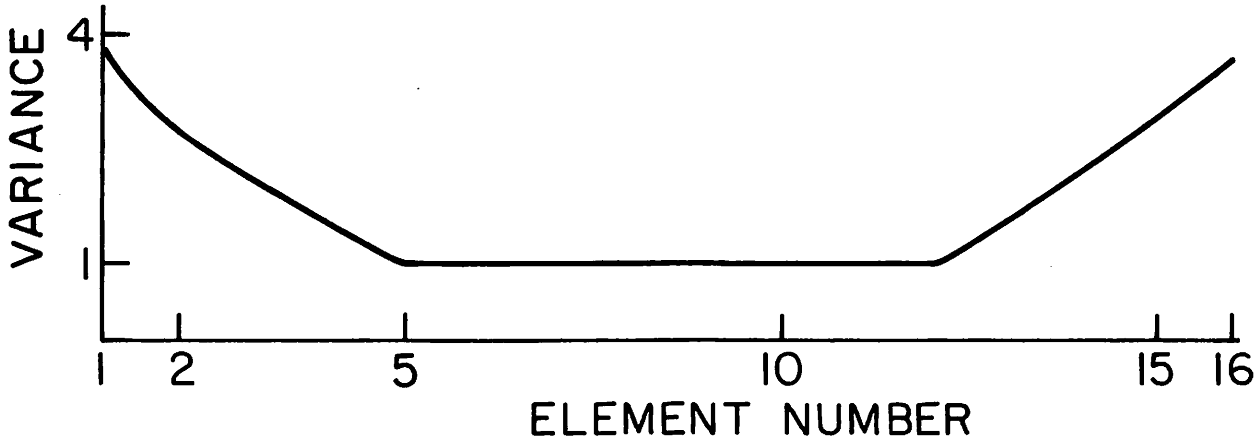
are also random. Their mean and covariance are

$$E\{\underline{C}\} = E\{\underline{B}^{+}\underline{y}\} = \underline{B}^{+}E\{\underline{y}\}$$

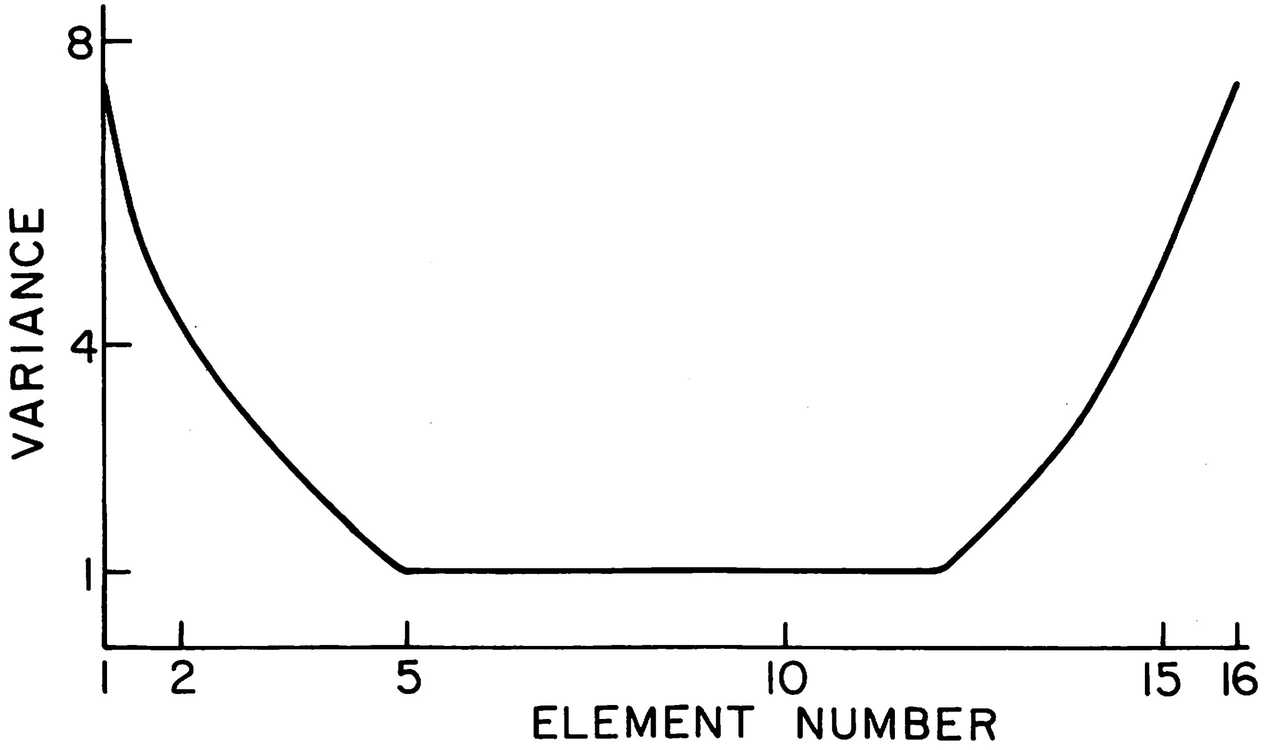
and

$$E\{\underline{C}\,\underline{C}^T\} = \underline{B}^{+}\underline{K}_y(\underline{B}^{+})^T$$

If the data vector $\underline{y}$ is considered a sample of Markov process with a correlation coefficient of ρ $(0 < \rho < 1)$ between each adjacent pixels and



b) N=16, ELEMENT CORRELATION = .96



a) N=16, ELEMENT CORRELATION =.95

Figure 3.3-2. Variance of B-spline coefficients.

self correlation coefficient of unity, then

$$\underline{K}_y = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{N-1} \\ \rho & 1 & \rho & \dots & \rho^{N-2} \\ \vdots & & & & \vdots \\ \rho^{N-1} & \dots & \dots & \dots & 1 \end{bmatrix}$$

Figures 2a and 2b contain two plots of the variance function of c_i as a function of i, where c_i is the i^{th} element of $\underline{C}$. The plots are obtained with N = M = 16.

References

1. T. N. E. Greville, "Data Fitting by Spline Functions," Trans. 12th Conf. Army Mathematicians, Report 67-1, U. S. Army Research Office-Durham, Durham, N.C., 1967, pp. 65-90.
2. T. N. E. Greville, Theory and Applications of Spline Functions, Academic Press, 1969.
3. I. J. Schoenberg, "Contributions to the Problem of Approximation of Equi-Distant Data by Analytic Functions," Quart. Appl. Math., Vol. 4, 1946, part A, pp. 45-99; part B, pp. 112-141.
4. J. F. Steffensen, Interpolation, Williams and Wilkins, Baltimore, 1927.

3.4 Positive Extrapolation of Signals and Images

Ali Habibi, Firouz Naderi

In a transform coding system a bandwidth reduction is achieved by discarding a number of transform coefficients of a natural image. Those coefficients possessing a small variance are of low information content, and replacing them by zeros at the receiver results in a rather small degradation in the quality of the encoded signal. The customary approach in designing transform coding systems has been to substitute the missing coefficients at the receiver with zeros. However, the quality of the coded signal improves by extrapolation of the missing coefficients from those which have been transmitted. The problem is analogous to one encountered in spectral estimation of data where the covariance function is first estimated for a number of lag values, then the covariance function at the available lag values is extrapolated for additional lag values prior to taking its Fourier transform. The extrapolation problem as applied to transform coding is more complicated. In transform coding only the quantized values of the coefficients are available at the receiver; thus, one is forced to estimate the missing components of the transformed data from the available possibly noisy components.

This problem has been analyzed using two different approaches. The first approach is statistical and is based upon the correlation among the transformed components using suboptimal transforms such as the Fourier, Hadamard, and Slant transforms. The second approach is called positive extrapolation since it is based upon the positiveness of the video data and the fact that the Toeplitz matrix constructed from the Fourier coefficients of a positive, real signal is always positive definite.

Extrapolation of Signals Let F(0), F(1), . . . , F(N) refer to the first N+1 components of vector $\underline{F}$, the Fourier transform of one line of a video data, which is composed of M points. To make $\underline{F}$ real one must generate an even function by first reflecting the video data about the t = 0 axis and then taking a Fourier transform of the even signal.

Now consider the Toeplitz matrix T(N+1) defined as

$$\underline{T}(N+1) = \begin{pmatrix} F(0) & F(1) & \cdots & F(N) & F(N+1) \\ F(1) & F(1) & \cdots & F(N-1) & F(N) \\ \vdots & \vdots & & \vdots & \vdots \\ F(N+1) & F(N) & \cdots & F(1) & F(0) \end{pmatrix} \qquad (1)$$

Since the modified video data is positive and even, and the real matrix $\underline{T}$(N+1) is semipositive definite, it follows that the determinant of $\underline{T}$(N+1) as a function of $\underline{F}$(N+1) has a single maximum. Hence the allowable values of F(N+1) are those that make the determinant of $\underline{T}$(N+1) equal to zero and all values in between. Expanding the determinant of $\underline{T}(n+1)$ in terms of the last row and the last column gives an expression for the determinant of $\underline{T}(N+1)$ in terms of $F(0), F(1), \ldots, F(N+1)$. Since $F(0), F(1), \ldots, F(N)$ are known, this is an expression for the determinant of $\underline{T}(N+1)$ in terms of $F(N+1)$. Choosing $F(N+1)$ to maximize the determinant of $T(N+1)$ gives a recursive algorithm to estimate $F(N+1)$ from $F(0), F(1), \ldots, F(N)$. The recursive algorithm can be used further to estimate $F(N+2)$ from $F(1), \ldots, F(N)$ and the estimated value of $F(N+1)$ i.e.

$$F(j) = \sum_{k=1}^{N} A(k)F(|j-k|) \quad \text{for } j = N+1, \ldots M-1 \tag{2}$$

where $A(k)$, $k = 1, \ldots, N$, are a set of fixed constants specified by matrix $\underline{T}(N+1)$.

Extrapolation of Images The positive extrapolation technique discussed for one-dimensional signals in the previous section can be generalized to extrapolate two-dimensional spectral density functions as well as two-dimensional Fourier transform of images. This is achieved by extending eq. (2) to functions of two variables by letting

$$F(i,j) = \sum_{k=1}^{N} \sum_{\ell=1}^{N} A(k,\ell)F(|i-\ell|, |j-k|) \qquad \text{for } i,j = 0, 1, \ldots, M-1 \tag{3}$$

where $F(i,j)$ is the two-dimensional discrete Fourier transform of the image and consists of M^2 elements. At the receiver site $(N+1)^2$ elements are available and these $(N+1)^2$ elements are used to extrapolate the missing elements prior to taking the inverse Fourier transform to obtain a reconstruction of the original image. Analogous to the one-dimensional system, the original picture is first folded along the $x = 0$ and $y = 0$ axes to generate an even two-dimensional array. This is required to make $F(i,j)$ an array of real elements. Solving eq. (3) for $A(k,\ell)$ is straightforward since $(N+1)^2$ values of $F(i,j)$ are known.

Experimental Results The performance of the positive extrapolation methods described above has been considered for a number of examples. The one-dimensional example is a discrete signal of 32 samples that consists of a pulse superimposed over a slowly varying background. The discrete Fourier transform of this signal is calculated and all but the first eight samples substituted by zeros. The inverse Fourier transform of the truncated signal and the original signal are shown on Figure 1. The truncated signal in the transform domain is then extrapolated to recover all 32 components. The inverse transform of the extrapolated signal is also shown on Figure 1. The extrapolated signal remains positive as expected and tracks the original signal significantly closer than the unextrapolated signal.

Figures 2 and 3 contain examples of positive extrapolation for a discrete (sin x/x)(sin y/y) signal and a block pulse signal, respectively. In each case the Fourier transform has been truncated preserving the transform samples in a 3 x 3 low frequency block out of a total of 15 x 15 coefficients. The positive extrapolation process is seen to provide a significant improvement over the reconstruction without extrapolation.

3.5 Transform Domain Spectrum Interpolation

Michael N. Huhns

Quantization occurs whenever continuous physical properties are represented numerically. A quantizer is a zero-memory nonlinear device which restricts an input variable to a finite number of possible output regions. This process is irreversible and information is invariably destroyed since only the region containing the input is known at the output. However this output data can be combined with a priori knowledge about the input to reduce the amount of information lost by interpolating between the discrete outputs.

In transform image coding a block of image pixels undergoes a two dimensional transformation using a unitary transform such as the Fourier,

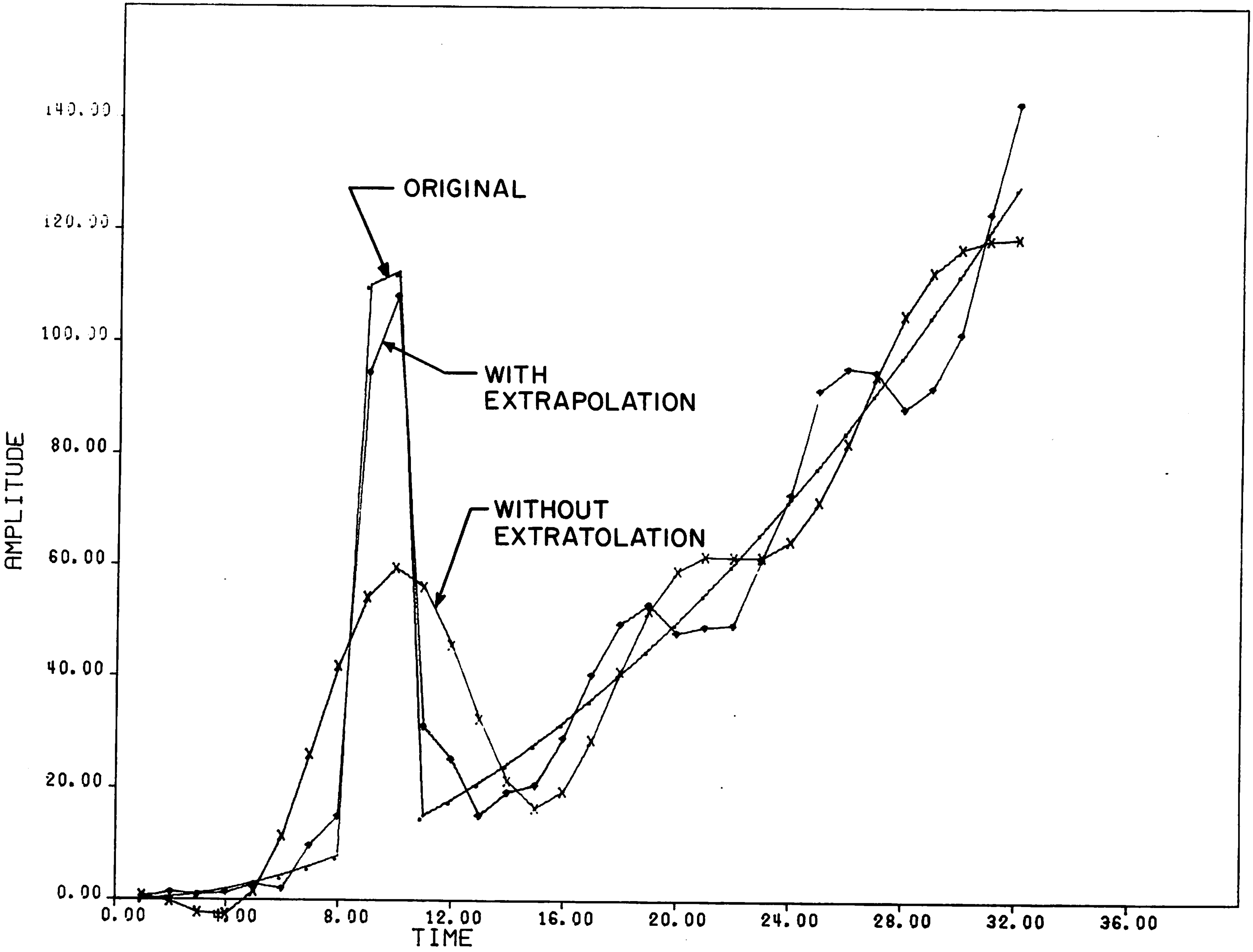
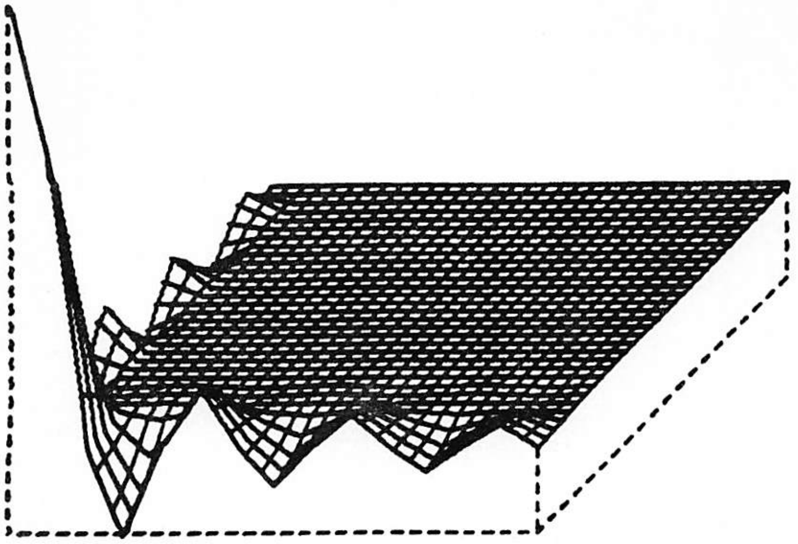
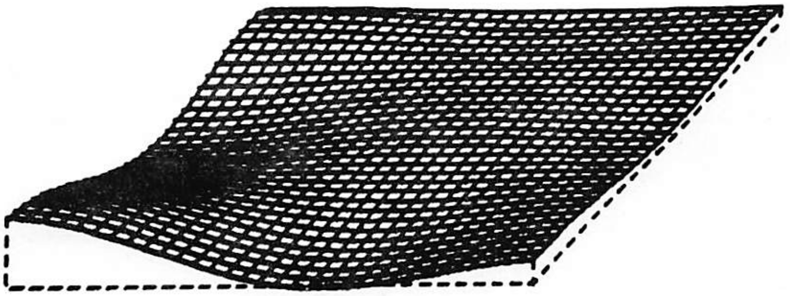


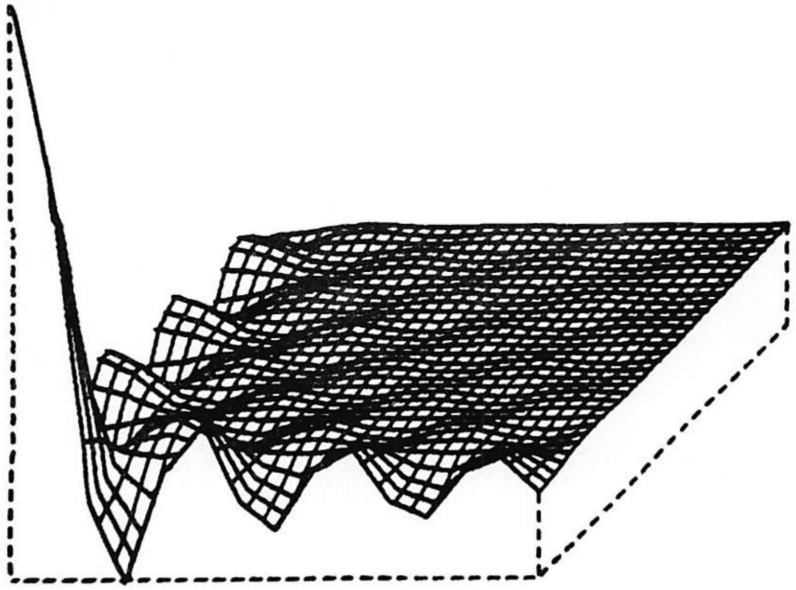
Figure 3.4-1. Positive extrapolation example.



(a) Original

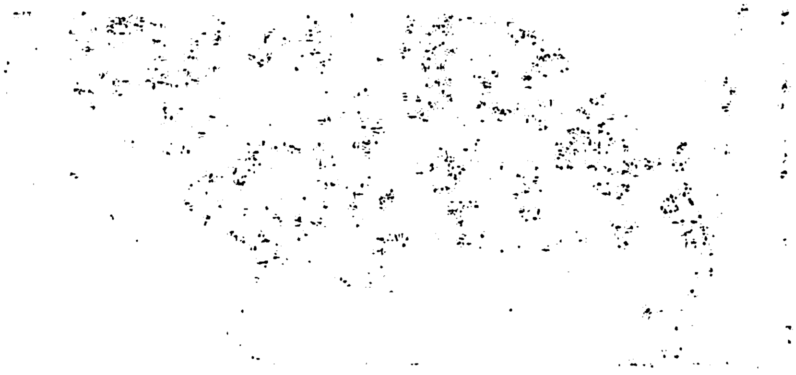


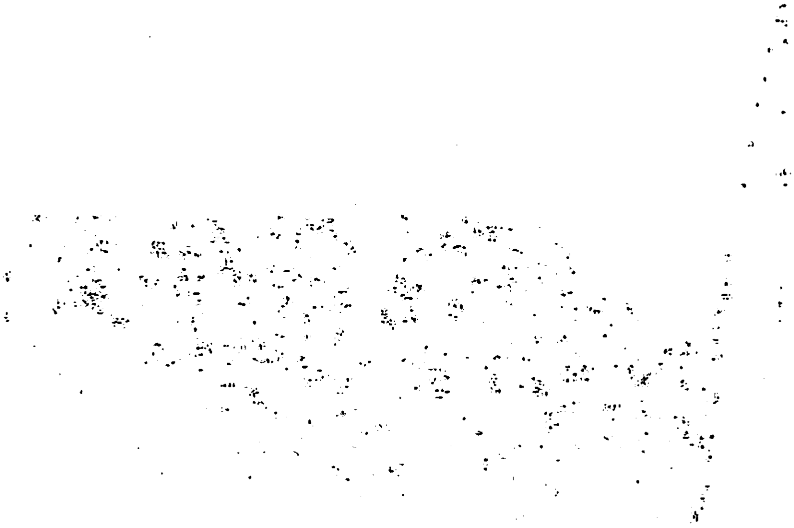
(b) Un-extrapolated

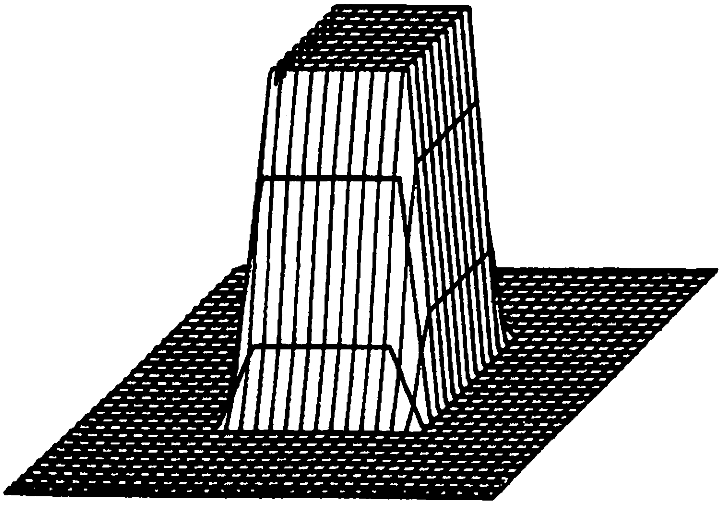


(c) Extrapolated

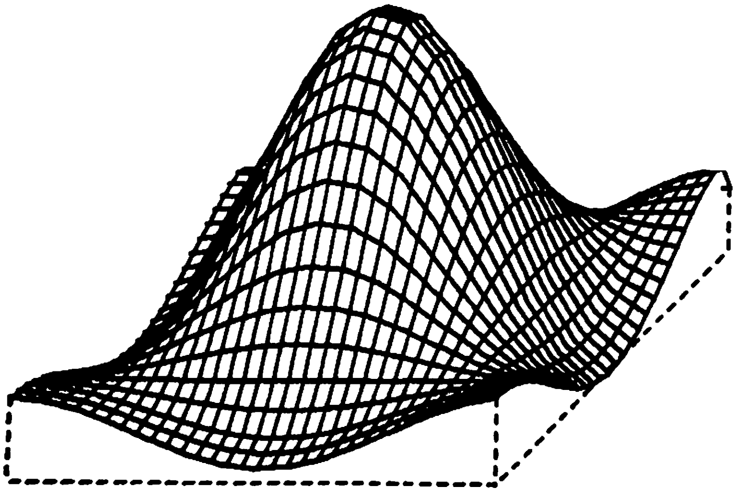
Figure 3.4-2. The original, un-extrapolated and the extrapolated two-dimensional Signals. Reduction ratio is $\frac{225}{9}$.



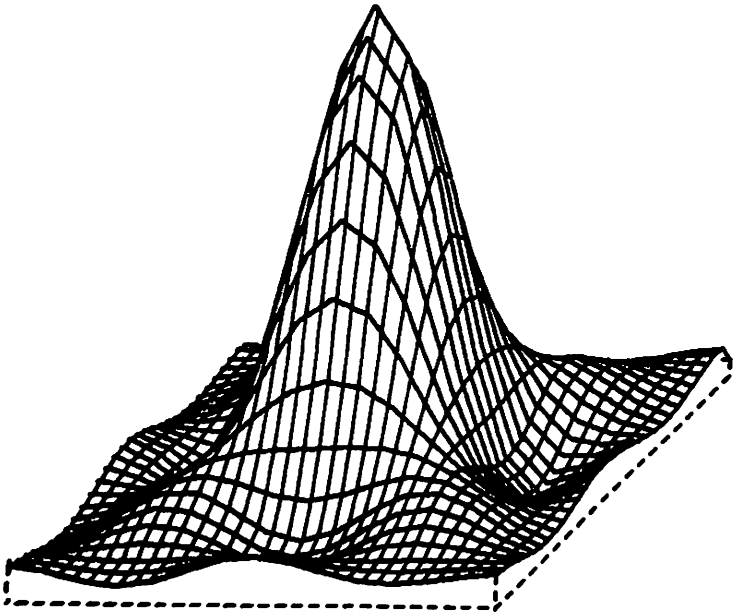




(a) Original



(b) Un-extrapolated



(c) Extrapolated

Figure 3.4-3. The original, un-extrapolated and the extrapolated two-dimensional Signals. Reduction ratio is $\frac{225}{9}$.

Hadamard, or Slant transform. Next, the transform coefficients are quantized and coded for transmission. Figure 1 illustrates a typical bit assignment for a zonal quantization and coding algorithm. The number of quantization levels assigned to the coefficient at coordinate (u, v) is

$$M(u,v) = 2^{b(u,v)} \tag{1}$$

where b(u, v) denotes the bit assignment. At the receiver, the quantized coefficients are reconstructed and an inverse transformation is performed to obtain an image estimate.

If a transform coefficient is quantized to zero bits, then its restoration is equivalent to a spectrum extrapolation as outlined by Pratt [1]. Those coefficients that are quantized to two or more levels can also be restored by a technique called spectrum interpolation.

Analysis Let the N element column vector $\underline{x}$ with probability density $p_x(\underline{x})$ denote a vector of input data samples. For two-dimensional data arrays, $\underline{x}$ is formed by column scanning the data array. Each data sample is quantized into one of M output regions, denoted by D_i, $i = 1, 2, \ldots, M$. The estimated value of $\underline{x}$ based upon the observed D_i regions is the quantizer output vector $\underline{y}_i$. The average error in this estimate is then defined as

$$\mathcal{E} = \sum_{i=1}^{M} \int_{D_i} e(\underline{x}-\underline{y}_i) p_x(\underline{x}) d\underline{x} \tag{2}$$

where e(·) is an arbitrary error weighting criterion. The vector of estimates $\underline{y}_i$ should be chosen to minimize the average error. This choice can be determined by utilizing the principles of calculus to find the stationary points of the error surface $\mathcal{E}$ with respect to each $\underline{y}_i$. Hence

$$\frac{\partial \mathcal{E}}{\partial \underline{y}_i} = -\int_{D_i} \frac{\partial}{\partial \underline{y}_i} [e(\underline{x}-\underline{y}_i)] p_x(\underline{x}) d\underline{x} \qquad i = 1, 2, \ldots, M \tag{3}$$

8	8	8	8	8	8	7	7	6	6	6	6	5	5	5	5
8	8	8	7	6	6	4	4	4	4	4	4	3	3	3	3
8	8	7	5	5	4	3	3	2	2	2	2	2	2	2	2
8	7	5	3	3	3	2	2	1	1	1	1	1	1	1	1
8	6	5	3	3	2	1	1	1	1	1	1	1	1	1	1
8	6	4	3	2	2	1	1	1	1	1	1	1	1	1	1
7	4	3	2	1	1	1	1	0	0	0	0	0	0	0	0
7	4	3	2	1	1	1	1	0	0	0	0	0	0	0	0
6	4	2	1	1	1	0	0	0	0	0	0	0	0	0	0
6	4	2	1	1	1	0	0	0	0	0	0	0	0	0	0
6	4	2	1	1	1	0	0	0	0	0	0	0	0	0	0
6	4	2	1	1	1	0	0	0	0	0	0	0	0	0	0
5	3	2	1	1	1	0	0	0	0	0	0	0	0	0	0
5	3	2	1	1	1	0	0	0	0	0	0	0	0	0	0
5	3	2	1	1	1	0	0	0	0	0	0	0	0	0	0
5	3	2	1	1	1	0	0	0	0	0	0	0	0	0	0

Figure 3.5-1. Typical transform domain quantizing bit assignment.

with the assumption that the error function e(·) is differentiable. Solving eq. (3) for the quadratic error criterion

$$e(\underline{x}-\underline{y}_i) = \mathrm{Tr}\{(\underline{x}-\underline{y}_i)(\underline{x}-\underline{y}_i)^T\} \tag{4}$$

one obtains

$$\frac{\partial e}{\partial \underline{y}_i} = -2(\underline{x} - \underline{y}_i) \tag{5}$$

which implies that

$$\int_{D_i} (\underline{x}-\underline{y}_i)p_x(\underline{x})d\underline{x} = 0 \qquad i = 1, 2, \ldots, M \tag{6}$$

Rearrangement reveals

$$\underline{y}_i = \frac{\int_{D_i} \underline{x}p_x(\underline{x})d\underline{x}}{\int_{D_i} p_x(\underline{x})d\underline{x}} \qquad i = 1, 2, \ldots, M \tag{7}$$

or

$$\underline{y}_i = E\{\underline{x}|\underline{x}\epsilon D_i\} \tag{8}$$

This is an expression for the best nonlinear mean square estimate of $\underline{x}$, given that $\underline{x}$ lies within region D_i.

Now assume that $\underline{x}$ is distributed according to a Gaussian probability density function

$$p_x(\underline{x}) = K \exp\{-\tfrac{1}{2}\underline{x}^T \underline{C}_x^{-1} \underline{x}\} \tag{9}$$

where $\underline{C}_x$ is the covariance matrix of $\underline{x}$ and the mean is assumed to be zero. Also let

$$D_i = \{x_i | x_j \epsilon [a_j, b_j)\} \qquad j = 1, 2, \ldots, N \tag{10}$$

Then

$$\underline{y}_i = \frac{\int_{\underline{a}}^{\underline{b}} \underline{x} K \exp\{-\tfrac{1}{2}\underline{x}^T \underline{C}_x^{-1} \underline{x}\} d\underline{x}}{\int_{\underline{a}}^{\underline{b}} K \exp\{-\tfrac{1}{2}\underline{x}^T \underline{C}_x^{-1} \underline{x}\} d\underline{x}} \tag{11}$$

Curry [2] has solved this equation for finely quantized values of x_j, i.e.

$$b_j - a_j < \sigma_j \qquad j = 1, 2, \ldots, N \tag{12}$$

where σ_j is the standard deviation of the j^{th} component of $\underline{x}$. His approach is to approximate the Gaussian density by the first three terms of its Taylor series expansion about the midpoint of the region D_i. The integration can then be performed, with the result that

$$E\{\underline{x} | \underline{x} \epsilon D_i\} = (\underline{I} - \underline{\Delta} \underline{C}_x^{-1}) \frac{\underline{b} + \underline{a}}{2} \tag{13}$$

where

$$\underline{\Delta} = \left\{ \frac{(b_j - a_j)^2}{12} \delta_{kj} \right\} \qquad k, j = 1, 2, \ldots, N \tag{14}$$

An exact solution can be obtained when the components of $\underline{x}$ are uncorrelated. In this case the covariance matrix can be expressed as

$$\underline{C}_x = \{\sigma_j^2 \delta_{kj}\} \qquad k, j = 1, 2, \ldots, N \tag{15}$$

and much computation reveals that

$$\underline{y}_i = \sqrt{\frac{2}{\pi}} \begin{bmatrix} \dfrac{\sigma_1 (e^{-b_1^2/2\sigma_1^2} - e^{-a_1^2/2\sigma_1^2})}{\operatorname{erf} \dfrac{b_1}{\sqrt{2}\sigma_1} - \operatorname{erf} \dfrac{a_1}{\sqrt{2}\sigma_1}} \\ \vdots \\ \dfrac{\sigma_N (e^{-b_N^2/2\sigma_N^2} - e^{-a_N^2/2\sigma_N^2})}{\operatorname{erf} \dfrac{b_N}{\sqrt{2}\sigma_N} - \operatorname{erf} \dfrac{a_N}{\sqrt{2}\sigma_N}} \end{bmatrix} \tag{16}$$

Gaussian variables which had been decorrelated by means of a Karhunen-Loeve transformation and then quantized could be restored according to a minimum mean square error criterion by utilizing this last equation.

An exact analytical solution to eq. (11) also exists when an estimate of a single vector component, x_N, is desired based upon two types of information -- (a) the other components, $x_1, x_2, \ldots, x_{N-1}$, which are known completely (quantized with an infinite number of bits); (b) the quantizer output which nonlinearly specifies the interval containing x_N. To derive this, consider

$$\underline{y}_i = E\{\underline{x} \mid x_1 = a_1, x_2 = a_2, \ldots, x_{N-1} = a_{N-1}; a_N \leq x_N < b_N\} \quad (17)$$

or

$$\underline{y}_i = \frac{\int_{D_i} \begin{pmatrix} a_1 \\ \vdots \\ a_{N-1} \\ x_N \end{pmatrix} \exp\left\{ -\tfrac{1}{2}(a_1 \ldots a_{N-1} x_N)\underline{C}_x^{-1} \begin{pmatrix} a_1 \\ \vdots \\ a_{N-1} \\ x_N \end{pmatrix} \right\} d\underline{x}}{\int_{D_i} \exp\left\{ -\tfrac{1}{2}(a_1 \ldots a_{N-1} x_N)\underline{C}_x^{-1} \begin{pmatrix} a_1 \\ \vdots \\ a_{N-1} \\ x_N \end{pmatrix} \right\} d\underline{x}} \quad (18)$$

or

$$\underline{y}_i = \frac{\int_{a_N}^{b_N} \begin{pmatrix} a_1 \\ \vdots \\ a_{N-1} \\ x_N \end{pmatrix} \exp\left\{ -\tfrac{1}{2}(a_1 \ldots a_{N-1} x_N)\underline{C}_x^{-1} \begin{pmatrix} a_1 \\ \vdots \\ a_{N-1} \\ x_N \end{pmatrix} \right\} dx_N}{\int_{a_N}^{b_N} \exp\left\{ -\tfrac{1}{2}(a_1 \ldots a_{N-1} x_N)\underline{C}_x^{-1} \begin{pmatrix} a_1 \\ \vdots \\ a_{N-1} \\ x_N \end{pmatrix} \right\} dx_N} \quad (19)$$

Now denote the elements of $(\underline{C}_x)^{-1}$ by

$$(\underline{C}_x)^{-1} = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1N} \\ r_{21} & r_{22} & \cdots & r_{2N} \\ \vdots & \vdots & & \vdots \\ r_{N1} & r_{N2} & \cdots & r_{NN} \end{pmatrix} \tag{20}$$

Then performing the one-dimensional integrations in eq. (19) yields

$$\underline{y}_i = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{N-1} \\ \dfrac{-\exp\left\{-\dfrac{1}{2r_{NN}}\left(r_{NN}x_N+\displaystyle\sum_{j=1}^{N-1} a_j r_{jN}\right)^2\right\}\Bigg|_{a_N}^{b_N}}{\sqrt{\dfrac{\pi r_{NN}}{2}}\;\mathrm{erf}\left(\dfrac{r_{NN}x_N+\displaystyle\sum_{j=1}^{N-1} a_j r_{jN}}{\sqrt{2r_{NN}}}\right)\Bigg|_{a_N}^{b_N}} - \dfrac{1}{r_{NN}}\displaystyle\sum_{j=1}^{N-1} a_j r_{jN} \end{pmatrix} \tag{21}$$

If x_N is quantized to an infinite number of bits, then $y_i^N = a_N = b_N$, as expected. If x_N is quantized to zero bits, its interval is the real line ($-a_N = b_N = \infty$), and then its estimate, y_i^N, is

$$y_i^N = -\frac{1}{r_{NN}} \sum_{j=1}^{N-1} a_j r_{jN} \tag{22}$$

This result is identical to that obtained by Pratt [1] in estimating an unknown spectral value based on known spectral components. However eq. (20) is a more general result in that it can be utilized to estimate components that have been quantized to any number of bits by an arbitrary quantization scheme.

Transform Domain Spectrum Interpolation The above solution is applicable to the mean square restoration of zonal coded transform samples. In this case, the transform samples have a Gaussian distribution, since each is the sum of a large number of random variables so that the central limit theorem can be invoked. These transform samples are typically quantized according to a bit assignment such as the one shown in Figure 1. For such a quantizing scheme, only eq. (16) can be utilized directly for restoration; however this equation ignores the known correlation existing between the samples. Curry's method of eq. (13) is unable to restore samples quantized to fewer than two bits. However, for greater bit assignments, it has the advantage of providing a simultaneous solution utilizing all the available information. The technique developed in eqs. (17) to (21) avoids the above difficulties, but requires a recursive solution which may be only asymptotically optimal (further analysis is expected to establish this). Therefore the best restoration, on the basis of optimality and ease of implementation, is obtained from a combination of the solutions presented above and must be adapted to the particular quantizer used. This technique will soon be applied to zonal transform coded images. It is anticipated that the resultant image will have a lower mean square error and improved subjective quality.

References

1. W. K. Pratt, "Transform Image Coding Spectrum Extrapolation," Proc. Seventh Hawaii International Conference on System Sciences, January, 1974, pp. 7-9.

2. R. E. Curry, Estimation and Control with Quantized Measurements, The M.I.T. Press, Cambridge, Massachusetts 1970.

3.6 Variable Rate Image Coding for Sources with Unknown Probabilities

Lee D. Davisson

The average distortion of image encoding at a fixed rate subject to a fidelity criterion depends upon an actual statistical source index θ, in effect, the actual stationary ergodic source model for the image to be encoded. Thus distortion is a random variable over the ensemble with distribution given by the distribution of the parameter θ, i.e. the class of possible images. In many applications it may be more desirable to allow the coding rate to depend on θ while holding the average distortion fixed over the ensemble. This is the case, for example, when the image is to be stored or a variable transmission rate exists due to the multiplexing of many messages, e.g. the ARPANET.

A coding theorem has been established for the special case of a finite number of subsources, $\theta = 1, 2, \ldots, K$, in the ensemble. The theorem holds for noncountable ensembles as well, but the proof is considerably more involved. In addition, it is assumed that there is a maximum distortion value, ρ_M.

For each value of k, generate a set of codewords according to the usual coding theorem for stationary, ergodic sources. If $R_k(D)$ is the rate distortion function in bits of the k^{th} subsource, and D is the desired value of average distortion, each code will contain L_k codewords where

$$\log L_k = (N(R_k(D-\epsilon) + \epsilon) \tag{1}$$

and ϵ is an arbitrary positive constant with the blocksize chosen large enough so that the average distortion is $D-\epsilon$ for all θ, and so that the probability that there is no codeword with distortion less than $D-\epsilon/2$ is less than $\epsilon/2\,\rho_M$.

The coded representation of each of the L_k codewords generated for each k consists of two parts. The first part is the fixed length binary number equal to $k-1$, $k = 1, 2, \ldots, K$ using at most $\log K+1$ bits to identify the codeword. The second part is the location of the codeword in a list for each k of length at most $\log L_k+1$ bits. Thus the rate of any codeword for a given θ is at most

$$r_k = \frac{\log K + 2}{N} + R_k(D-\epsilon) + \epsilon \text{ bits} \tag{2}$$

Obviously by choosing N large enough and ϵ small enough, the rate can be made arbitrarily close to $R_\theta(D)$ for $\theta = 1, 2, \ldots, K$.

The achievement of D and $R_\theta(D)$ for the combined supercode depends upon the actual choice of a codeword out of the

$$L = \sum_{i=1}^{K} L_j$$

total codewords. The coding rule is as follows: Upon observing an output block of length N, among the codewords of distortion less than $D-\epsilon/2$, find the one of minimum rate, if any. If there is no codeword with distortion less than $D-\epsilon/2$, make a random choice. The average rate for $\theta = k$ then is

$$\hat{R}_k(d) \le r_k + (\sup_j r_j)\text{Prob}[\text{no codeword of distortion less than D in the kth code}]$$

or

$$\hat{R}_k(D) \le r_k + (\sup_j r_j)\epsilon/2\,\rho_M = \frac{\log K+2}{N} + R_k(D-\epsilon) + \epsilon + (\sup_j r_j)\,\epsilon/2\rho_M \tag{3}$$

which is arbitrarily close to $R_k(D)$ for small enough ϵ and large enough N. The average distortion for $\theta = k$ is

$$D_k \le D-\epsilon/2 + \rho_M \text{ Prob}[\text{no codeword of distortion less than D in the } k^{th} \text{ code}]$$

$$D_k \le D-\epsilon/2 + \rho_M\, \epsilon/2\rho_M = D \tag{4}$$

The result described is largely an existence theorem and therefore does not prescribe a specific method of synthesizing data compression systems. It does, however, provide figures of merit and optimal performance bounds that can serve as an absolute yardstick for comparison with real systems. Furthermore, the results provide theoretical justification for overall design philosophies that have proved useful in practice. In practice, one could, for example have a coding scheme for each of several classes of pictures, e.g. classed according to "busyness." The appropriate coder would be switched in for each encoded image with the coder identity sent as a prefix to the encoded picture information as suggested above. The approximate minimum number of bits would then be used for each image depending upon the average allowed distortion.

4. Image Restoration and Enhancement

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. Both techniques are subjects of continuing study; results of this effort during the past six months are summarized below.

The first report deals with methods of restoration utilizing the matrix pseudoinverse of a blur matrix which models a space variant point spread function. Computation of the pseudoinverse is obtained indirectly by a singular value decomposition of the blur matrix. The following report considers another approach to pseudoinverse image restoration. Computational techniques are developed for pseudoinverse restoration by processing in the transform domain.

An analysis is presented in the next report of the effects of discrete modelling of the superposition integral for image restoration. The effect of modelling errors and the ill-conditioning of the blur matrix are quantified for typical image blur models.

The use of spline functions in imaging models is explored next. It is shown that the B-splines offer computational promise for image restoration with smoothness constraints.

Astigmatism and curvature of field image degradations are characterized by spatially variant imaging models. It is shown that the imaging model can be decomposed into a cascade of a geometric distortion, a spatially invariant linear system, and another geometric distortion. Methods of inverting this cascaded model are developed.

In the next report, the concept of transform domain Wiener filtering for image restoration is extended to include lower triangular transformations. For a Markov process image model, it is found that an extremely efficient computational algorithm can be obtained.

The two following reports are concerned with two aspects of color image restoration. Once report considers techniques for estimating the tristimulus values of a color from spectral observations of the color. The other report presents methods for compensating for film nonlinearities in a photograph of a color television display.

The final report presents a quantitative development of pseudocolor image enhancement techniques to map a grey scale image into a color display. Mappings are found that are psychophysically relevant, yet computationally efficient.

4.1 Space Variant Point Spread Function Pseudoinversion

Monty Adler and Harry C. Andrews

Many complex forms of image degradation cannot be modelled by a spatially invariant point spread function; consequently, Fourier techniques are not applicable to the restoration process. For these spatially variant point spread function systems, restorations can be achieved by a matrix formulation in which point spread function matrices are inverted. For singular blur matrices, the inverses must be replaced by pseudoinverses to achieve the least squares approximation to the original. For computational simplification an assumption of a separable space variant point spread function allows the following analysis.

Let $\underline{F}$ be a matrix representing a perfect image which is acted upon by a separable blur function to produce a blurred image $\underline{G}$ as modelled by

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

where $\underline{B}$ blurs the rows of $\underline{F}$ and $\underline{A}$ blurs the columns of $\underline{F}$. Usually $\underline{A}$ and $\underline{B}$ are singular or almost singular. One would like to find an approximation to $\underline{F}$, $\hat{\underline{F}}$ such that $\|\hat{\underline{F}} - \underline{F}\|$ is minimized. The analysis and results which follow do not depend on any of the matrices being square but for simplicity it will be assumed that all matrices are N by N. If $\underline{A}$ and $\underline{B}$ are nonsingular and square the solution is, of course, $\hat{\underline{F}} = \underline{A}^{-1}\,\underline{G}\,\underline{B}^{-1} = \underline{F}$.

Using appropriate $\underline{A}$ and $\underline{B}$ matrices one can model space variant blur, linear motion blur, and in general any separable blur function.

It is known that $\underline{A}$ and $\underline{B}$ can be expressed as a sum of matrices of rank one as follows

$$\underline{A} = \sum_{i=1}^{N} \eta_i^a\, \underline{U}_i^a\, \underline{V}_i^{aT} \qquad \text{for} \quad \eta_i^a \geq \eta_{i+1}^a \geq 0$$

$$\underline{B} = \sum_{i=1}^{N} \eta_i^b\, \underline{U}_i^b\, \underline{V}_i^{bT} \qquad \text{for} \quad \eta_i^b \geq \eta_{i+1}^b \geq 0$$

where η_i^a, η_i^b are scalars, and $\underline{U}_i^a$, $\underline{U}_i^b$, V_i^a, $\underline{V}_i^b$ are column vectors of length N, and the sets $\{\underline{U}_i^a\}$, $\{\underline{U}_i^b\}$, $\{\underline{V}_i^a\}$, $\{\underline{V}_i^b\}$ are orthonormal. In an ideal computational environment, $\eta_i^a = 0$ for $i > R_a$ where R_a is the rank of $\underline{A}$; similarly for $\underline{B}$.

The ideal pseudoinverse of $\underline{A}$ can be expressed as

$$\underline{A}^+ = \sum_{i=1}^{N} \eta_i^{a+}\, \underline{V}_i^a\, \underline{U}_i^{aT}$$

where

$$\eta_i^{a+} = 1/\eta_i^a \quad \text{if} \quad \eta_i^a \neq 0$$

$$\eta_i^{a+} = 0 \quad \text{if} \quad \eta_i^a = 0$$

A model is needed for the computed SVD. Using a computer the calculated SVD will be modeled by

$$\underline{A} = \sum_{i=1}^{N} \lambda_i^a\, \underline{U}_i^a\, \underline{V}_i^{aT} \quad \text{for} \quad \lambda_i^a \geq \lambda_{i+1}^a \geq 0.$$

The λ_i^a will not be zero for $i > R_a$ due to computational inaccuracy; and $\underline{U}_i^a$,

$\underline{V}_i^a$ will in general be identical to the ideal case. Note that the rank of $\underline{A}$, R_a cannot be accurately determined since the λ_i^a do not go to zero. Defining

$$\underline{A}_K^+ = \sum_{i=1}^{K} 1/\lambda_i^a \, \underline{V}_i^a \, \underline{U}_i^{aT}$$

then as soon as K goes beyond R_a, the algorithm divides by $1/\lambda_i^a$ for an inaccurate λ_i^a and $\underline{A}_K^+$ blows up. It is desirable to "stop K at R_a" but R_a is not known.

The problem is further complicated if $\underline{A}$ and $\underline{B}$ are non-singular but have small λ_i^a, λ_i^b which cannot be computed accurately. Even if R_a and R_b were known, the algorithm would not want to divide by the inaccurate λ_i^a, λ_i^b values. Unfortunately the typical situation is that $\underline{A}$ and $\underline{B}$ are either singular or non-singular with small λ_i^a, λ_i^b; but, of course, it is not known which is the case.

For simplicity $\underline{A}$ and $\underline{B}$ will be assumed to have the same or close to the same rank. The methods described below are trivially modifieable if the ranks of $\underline{A}$ and $\underline{B}$ differ.

Define

$$\hat{\underline{F}}_{JK} = \underline{A}_J^+ \, \underline{G} \, \underline{B}_K^+$$

and

$$\hat{\underline{G}}_{JK} = \underline{A} \, \hat{\underline{F}}_{JK} \, \underline{B} = \underline{A} \, \underline{A}_J^+ \, \underline{G} \, \underline{B}_K^+ \, \underline{B}$$

Since under the above assumption the best J equals the best K it is possible to drop the double subscript.

$$\hat{\underline{F}}_K = \underline{A}_K^+ \underline{G} \, \underline{B}_K^+$$

$$\hat{\underline{G}}_K = \underline{A} \, \hat{\underline{F}}_K \, \underline{B}$$

It is desirable to minimize $\| \underline{F} - \hat{\underline{F}}_K \|$ over all values of K. Three methods are proposed, one of which can be shown to find the best K.

Method 1 - Human Intervention. Successively compute $\hat{\underline{F}}_1, \hat{\underline{F}}_2 \ldots$ and display to a human. Although the best $\hat{\underline{F}}_K$ will not be apparent, it will be clear when $\hat{\underline{F}}_K$ is blowing up. In figure 1, the estimate $\hat{F}_{51}$ is seen to blow up, and this fact is obvious to a viewer. For a display purposes all $\hat{\underline{F}}_K$ values have been rounded to the nearest integers, all negative values have been set to zero, and all values greater than 63 have been set to 63. This explains the black and white blocks in the figures, which are out of range values.

Method 2 - Track $\underline{A}\,\underline{A}_K^+$ and $\underline{B}_K^+\,\underline{B}$. It can be shown that

$$\|\underline{I} - \underline{A}\,\underline{A}_K^+\|^2 = N - K$$

where N is the size of $\underline{A}$. Then one would expect $E_K(\underline{A}) = \|\underline{I} - \underline{A}\,\underline{A}_K^+\|^2 - (N-K)$ to deviate from zero when $\underline{A}_K^+$ blows up. Trials on a computer show that $E_K(\underline{A})$ remains at a low constant value until K approaches R_a at which point $E_K(\underline{A})$ increases by a factor of 10^3 around R_a. It thus appears that tracking $\underline{A}\,\underline{A}_K^+$ or $\underline{B}_K^+\,\underline{B}$ could be a method of find the best K. Table 1 contains the results of two a typical computer trial where

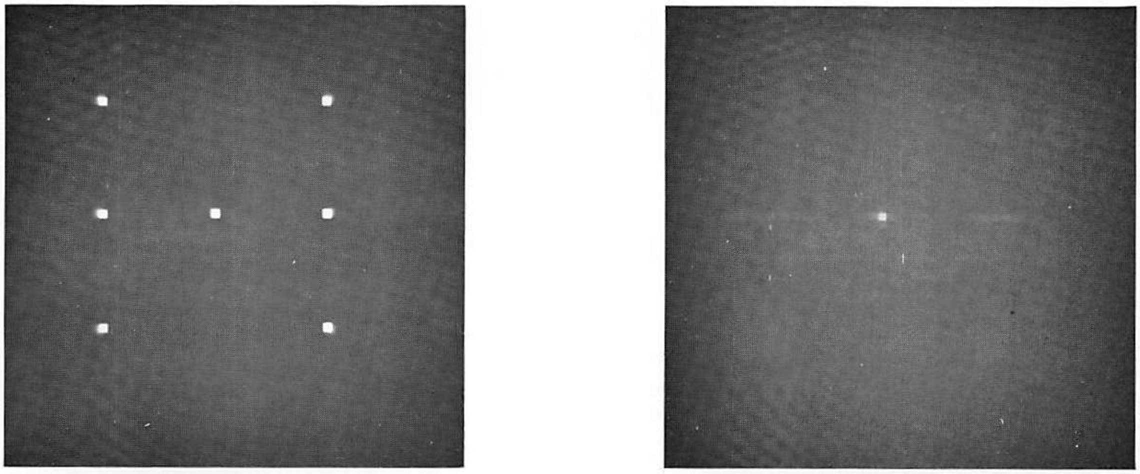
$$E(\underline{A}) = 10^{10}\,(\|\underline{I} - \underline{A}_K^+\underline{A}\|^2 - (N-K)\,)$$

$$E(\underline{B}) = 10^{10}\,(\|\underline{I} - \underline{B}_K^+\,\underline{B}\|^2 - (N-K)\,)$$

$$E(\underline{F}) = \|\underline{F} - \hat{\underline{F}}_K\|$$

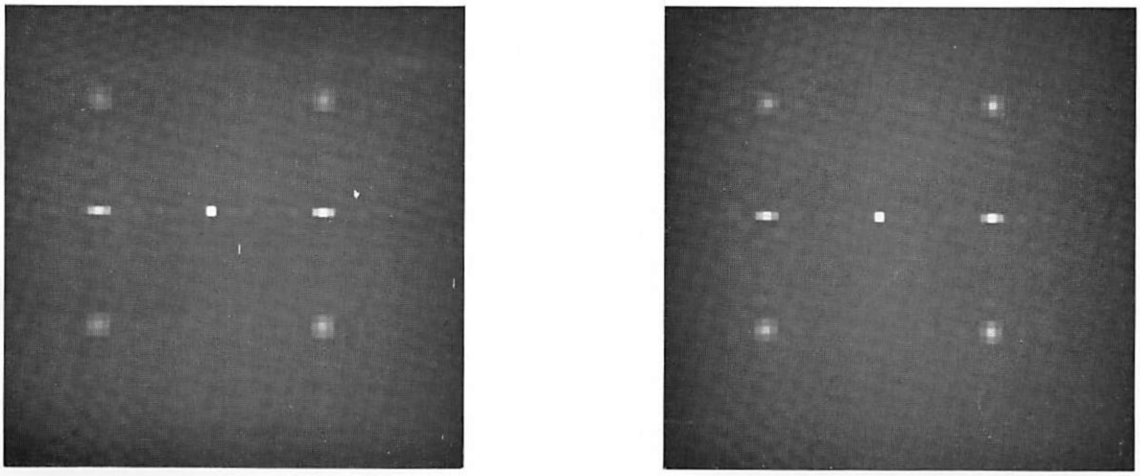
$$E(\underline{G}) = \|\underline{G} - \hat{\underline{G}}_K\|$$

Method 3. Track $\|\underline{G} - \hat{\underline{G}}_K\|$ It can be shown that min $\|\underline{F} - \hat{\underline{F}}_K\|$ occurs at the same value of K as $\min_K \|\underline{G} - \hat{\underline{G}}_K\|$ and therefore the best K can be found simply by tracking $\min_K \|\underline{G} - \hat{\underline{G}}_K\|$. Figures 1 to 3 show the results of three simulation cases. In each case the best reconstruction occurs at the K for which $\|\underline{G} - \hat{\underline{G}}_K\|$ is a minimum.



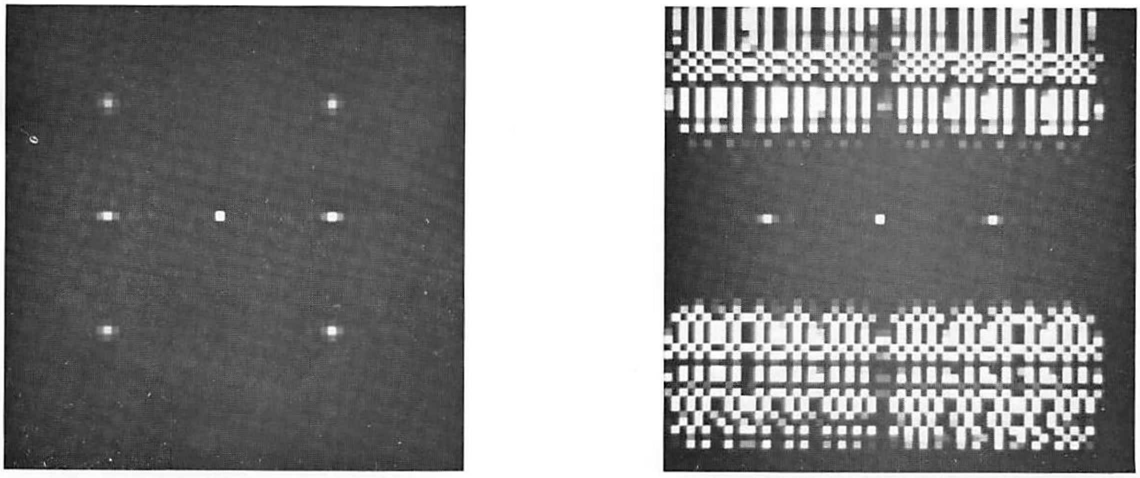
(a) original

(b) space variant blur



(c) $\hat{F}_{36}$

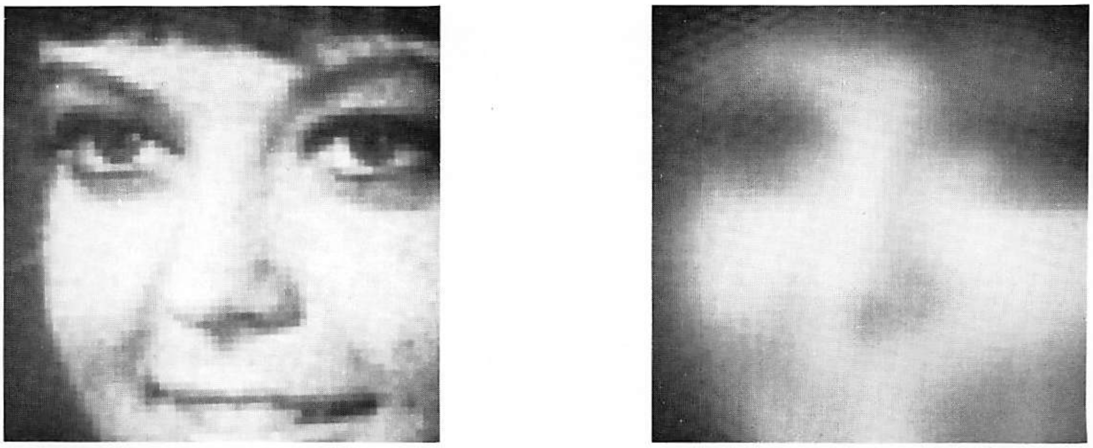
(d) $\hat{F}_{42}$



(e) $\hat{F}_{48}$ (best restoration)

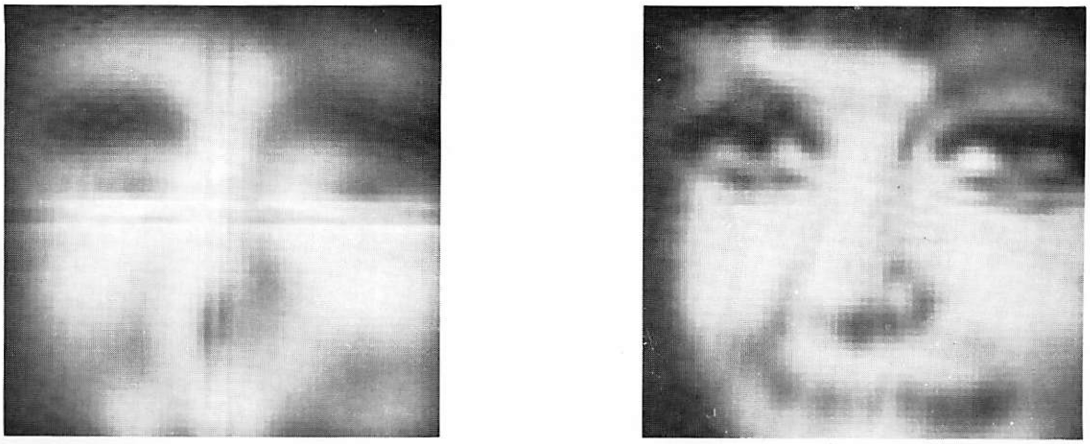
(f) $\hat{F}_{51}$ (note space variant singularity)

Figure 4.1-1. Restoration of space variant point sources, N=60 (Gaussian blur best at center) (Computer Generated Image).



(a) original

(b) space variant blur



(c) $\hat{F}_{12}$

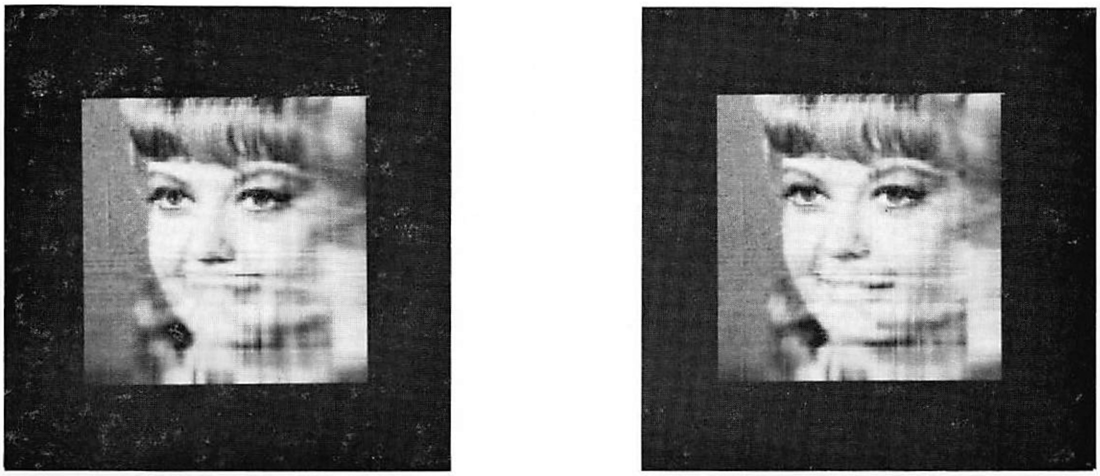
(d) $\hat{F}_{24}$

(e) $\hat{F}_{46}$(best restoration)

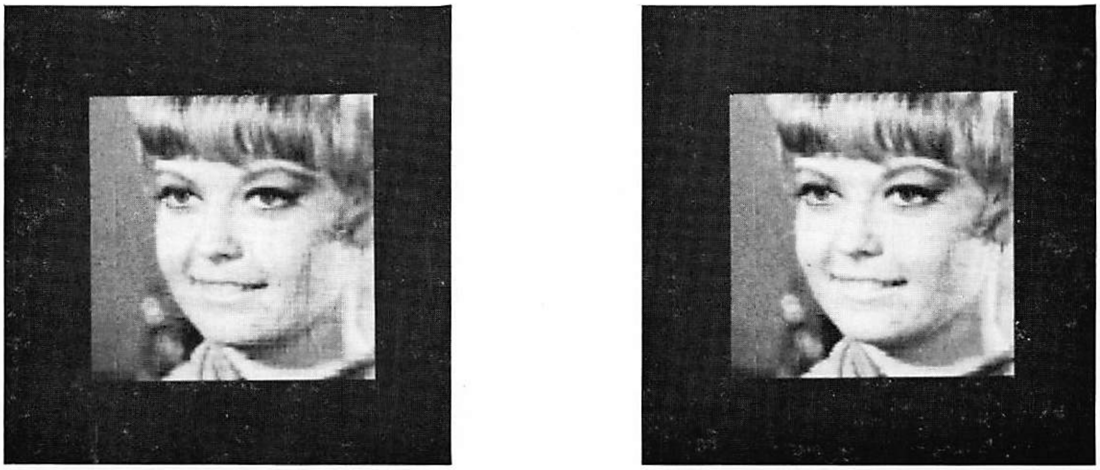
(f) $\hat{F}_{48}$(note space variant singularity)

Figure 4.1-2. Restoration of space variant point sources, N=60 (Gaussian blur best at center).

(a) original (b) space variant blur



(c) $\hat{F}_{71}$ (d) $\hat{F}_{76}$



(e) $\hat{F}_{96}$ (f) $\hat{F}_{103}$(minimum error)

Figure 4.1-3. Restoration from space variant point spread function, N-112 (blur best at center)

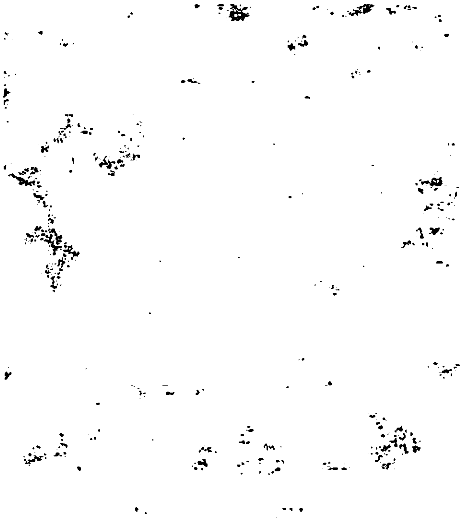
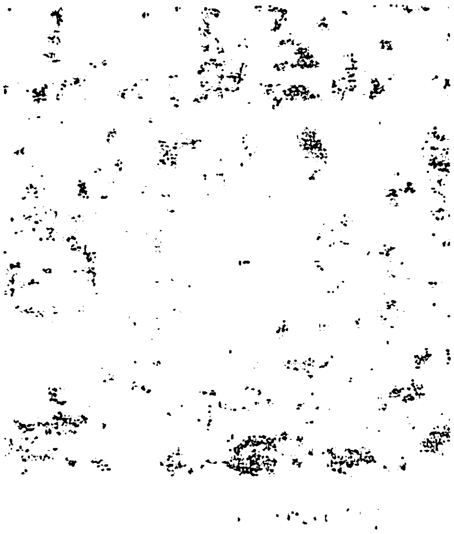





Table I Pseudoinverse and Reconstruction Computation Error - N = 60

I	E(F)	E(G)	E(A)	E(B)
0	156.9			
6	166.3	25.6	2.5	2.4
12	158.9	12.8	2.2	2.2
18	147.2	1.9	2.0	2.0
24	141.6	0.3	1.7	1.7
30	134.8	2.4E-2	1.3	1.3
36	125.9	1.1E-3	1.1	1.1
42	114.0	1.6E-5	1.2	0.13
43	109.7	7.6E-6	-0.1	2.5
44	108.4	2.2E-6	-4.5	2.9
45	106.5	1.6E-6	6.3	-3.8
46	104.3	3.1E-7	8.8	-16.
47	101.1	2.589E-7	58.	-33.
48*	99.2	2.583E-7	-100.	94.
49	177.8	1.9E-6	-67.	510.
50	382.9	2.9E-6	-420.	-250.
51	2.2E+4	2.3E-5	-8800.	7200.

* minimum

Experimental Results Figures 1, 2, and 3 show an image, its blur and the reconstruction for different values of K. Note that once the minimum is found, the reconstructed picture quickly blows up. Table 1 shows $\| \underline{G} - \hat{\underline{G}}_K \|$ and $\| \underline{F} - \hat{\underline{F}}_K \|$ for different K. Although the minimum occurs at the same point, $\| \underline{G} - \hat{\underline{G}}_K \|$ remains small after the minimum by $\| \underline{F} - \hat{\underline{F}}_K \|$ blows up. Naturally, in the real world one would not be able to track $\| \underline{F} - \hat{\underline{F}} \|$ but $\| \underline{G} - \hat{\underline{G}} \|$ is available.

4.2 Pseudoinverse Image Restoration by Transform Domain Processing

William K. Pratt

Linear operations on data can often be performed more efficiently by indirect techniques which involve projection of the data to another vector space through a unitary transformation of the data. This concept has been applied with success for discrete Wiener filtering of images. The following outlines the extension of the concept to pseudoinversion for image restoration.

Imaging Model Let $F(x, y)$ denote an ideal, continuous, infinite extent image field, and let $\tilde{F}(x,y)$ represent an observed image field which is also continuous and of infinite extent. For a large class of imaging systems the observed image is related to the ideal image by the convolution integral

$$\tilde{F}(x, y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\alpha, B)g(x-\alpha, y-\beta)d\alpha dB$$

where $g(x-\alpha;y-\beta)$ represents the impulse response of the imaging system. In the discrete model of the imaging system, the observed image is represented by physical samples spaced evenly over a unit grid, and the continuous integration is approximated by a quadrature formula resulting in

$$F(m_1, m_2) = \sum_{n_1=m_1}^{L+m_1-1} \sum_{n_2=m_2}^{L+m_2-1} F(n_1, n_2)H(m_1-n_1+L, m_2-n_2+L) \qquad (1)$$

for $1 \leq m_i \leq M$ and $1 \leq n_i \leq N$ where the array H, assumed to be zero outside its range of indices, represents the sampled impulse response and incorporates all quadrature factors. The impulse response is also truncated to an L×L array. In order to prevent serious modelling errors at the image boundary, it is necessary that

$$N \geq M + L - 1$$

It is computationally convenient to represent the data arrays $\underline{F}$ and $\underline{\tilde{F}}$ as column vectors, $\underline{f}$ and $\underline{\tilde{f}}$, respectively by column scanning. Then the sampled superposition operation can be described by

$$\underline{\tilde{f}} = \underline{B}\,\underline{f} \qquad (2)$$

where $\underline{B}$ is an $M^2 \times N^2$ matrix which can be partitioned as

$$\underline{B} = \begin{bmatrix} \underline{B}_{1,1} & \underline{B}_{1,2} & \cdots & \underline{B}_{1,L} & 0 & 0 & \cdots & 0 \\ 0 & \underline{B}_{2,2} & \cdots & \underline{B}_{2,L} & B_{2,L+1} & 0 & \cdots & 0 \\ \cdot & \cdot & & & & & & \cdot \\ \cdot & \cdot & & & & & & \cdot \\ \cdot & \cdot & & & & & & \cdot \\ 0 & 0 & \cdots\cdots & & \cdots\cdots & & \cdots & \underline{B}_{M,N} \end{bmatrix} \qquad (3)$$

The general term of $\underline{B}$ is then given by

$$B_{m_1,n_1}(m_2,n_2)=H(m_1-n_1+L,\, m_2-n_2+L) \qquad (4)$$

for

$$1 \le m_1 \le M \qquad 1 \le m_2 \le M$$

$$m_1 \le n_1 \le L+m_1 -1 \qquad m_2 \le n_2 \le L+m_2 -1$$

Pseudoinverse The imaging model of eq. (2) can be "inverted" by a pseudoinverse operator $\underline{B}^+$ in the sense that an estimate $\hat{\underline{f}}$ of the ideal image vector $\underline{f}$ can be computed by

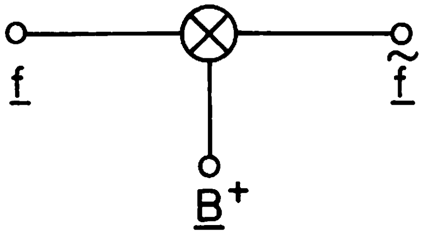
$$\hat{\underline{f}} = \underline{B}^+ \tilde{\underline{f}} = \underline{B}^+ \underline{B}\, \underline{f} \qquad (5)$$

If $\underline{B}$ is of rank M^2, then the pseudoinverse is equal to

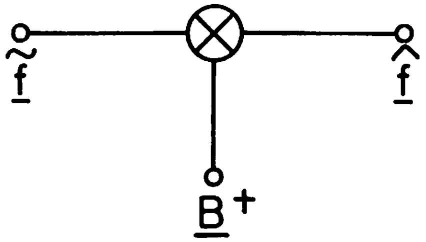
$$\underline{B}^+ = \underline{B}^T (\underline{B}\, \underline{B}^T)^{-1} \qquad (6)$$

It should be noted that since $M^2 < N^2$, that is, there are fewer observations than points on the ideal image to be estimated, the estimate will not be exact even in the absence of observational error. The pseudoinverse, however, does provide a minimum mean square error, minimum norm estimate.

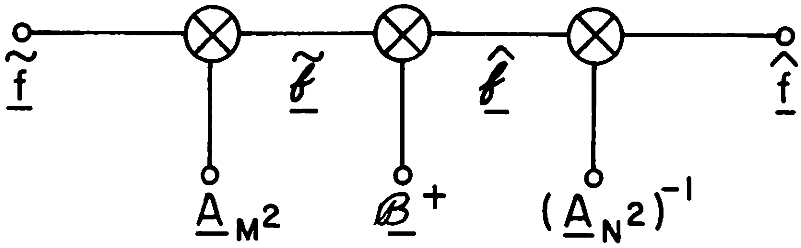
Transform Domain Pseudoinverse Figure 1 illustrates the computational steps involved in direct pseudoinversion of an image vector, and in



a) IMAGING MODEL



b) DIRECT PSEUDOINVERSE PROCESSING



c) TRANSFORM DOMAIN PSEUDOINVERSE PROCESSING

Figure 4.2-1. Direct and transform domain pseudoinverse processing.

the transformation processing approach. In transform processing a unitary transformation is performed on the observed vector $\underline{\tilde{f}}$ prior to multiplication by the transform domain pseudoinverse matrix $\underline{\mathcal{B}}^{+}$. An inverse transform reconstructs $\hat{f}$. From figure 1, since

$$\underline{\hat{f}} = \underline{B}^{+} \; \underline{\tilde{f}} \tag{7a}$$

and

$$\underline{\hat{f}} = (\underline{A}_{M^2})^{-1} \; \underline{\mathcal{B}}^{+} \; (\underline{A}_{N^2}) \; \underline{\tilde{f}} \tag{7b}$$

then clearly

$$\underline{\mathcal{B}}^{+} = (\underline{A}_{M^2}) \, \underline{B}^{+} \; (\underline{A}_{N^2})^{-1} \tag{8}$$

It is easy to show that

$$\underline{\mathcal{B}}^{+} = \underline{\mathcal{B}}^{*T} \; (\underline{\mathcal{B}} \; \underline{\mathcal{B}}^{*T})^{-1} \tag{9}$$

where

$$\underline{\mathcal{B}} = (\underline{A}_{M^2}) \, \underline{B} \, (\underline{A}_{N^2})^{-1} \tag{10}$$

is the transform domain representation of the blur matrix of eq. (3).

Computational efficiencies in transform domain pseudoinversion result from the sparseness and structure of the blur matrix $\underline{B}$. As an example, consider the Fourier transform representation of $\underline{B}^{+}$. In this case the transformation matrix is of the form

$$\underline{A}_{K^2} = \underline{A}_K \otimes \underline{A}_K \tag{11}$$

where

$$A_K \equiv \frac{1}{K} \, W^{(x-1)(y-1)} \qquad\qquad W \equiv \exp\left\{\frac{-2\pi i}{K}\right\}$$

for $x, y = 1, 2, \ldots, K$. Now, let $\underline{H}_E^{(N)}$ denote the extended impulse response obtained by imbedding the L × L impulse response matrix in the upper left corner of an N × N matrix of zeros. The two dimensional Fourier transform of the extended impulse response matrix is obtained from

$$\underline{\mathcal{H}}_E^{(N)} = \underline{A}_N \underline{H}_E^{(N)} \underline{A}_N \tag{12}$$

These transform components are then column scanned and inserted as the diagonal elements of the $N^2 \times N^2$ matrix

$$\underline{\mathcal{H}}^{(N)} = \text{diag}\ \{\underline{\mathcal{H}}_E(1,1), \underline{\mathcal{H}}_E(2,1), \ldots, \underline{\mathcal{H}}_E(N,N)\} \tag{13}$$

Then, it can be shown, after considerable manipulation, that the Fourier transform blur matrix is

$$\underline{\mathcal{B}} = [\underline{P}_B \otimes \underline{P}_B]\, \underline{\mathcal{H}}^{(N)} \tag{14}$$

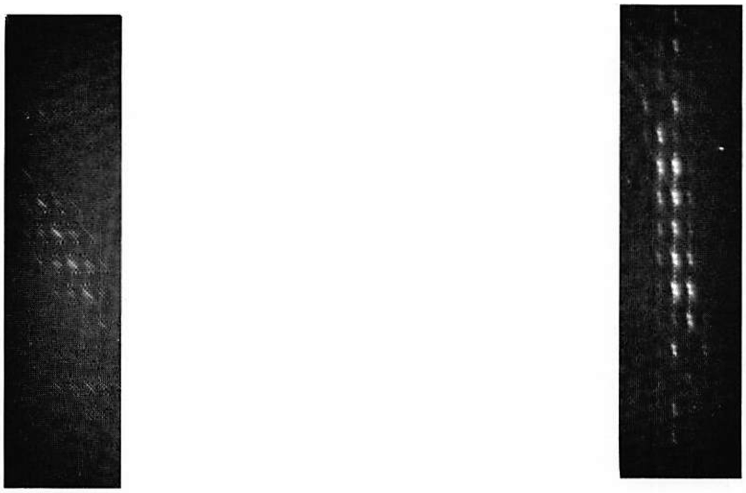
where

$$\underline{P}_B(u,v) = \frac{1}{\sqrt{M}}\ \frac{1 - W_N^{-(v-1)(L-1)}}{1 - W_M^{(u-1)} W_N^{-(v-1)}} \tag{15}$$

Thus, the $\underline{\mathcal{B}}$ matrix operator consists of a scalar weighting matrix, $\underline{\mathcal{H}}^{(N)}$, and a matrix $[\underline{P}_B \otimes \underline{P}_B]$ that performs the dimensionality conversion (an interpolation operation) between an N^2 element input vector and an M^2 element output vector. The dimensionality matrix is extremely sparse, and therefore, savings can be obtained in the computation of $\underline{\mathcal{B}}$ and subsequently, $\underline{\mathcal{B}}^+$. As an example figure 2 contains displays of the blur matrix $\underline{B}$, the pseudoinverse matrix $\underline{B}^+$, and their transform domain representations, $\underline{\mathcal{B}}$ and $\underline{\mathcal{B}}^+$ for a 9 × 9 Gaussian shaped impulse response, and M=8 and N=16. The relative sparseness of $\underline{\mathcal{B}}$ and $\underline{\mathcal{B}}^+$ are apparent from the figures.

(a) spatial domain, $\underline{B}$

(b) Fourier domain, $\underline{\mathcal{B}}$



(c) spatial domain, $\underline{B}^{+}$

(d) Fourier domain, $\underline{\mathcal{B}}^{+}$

Figure 4.2-2. Examples of blur matrix and its pseudoinverse in spatial and Fourier transform domain.

4.3 Modelling Superposition Integrals for Image Restoration

Nelson D. A. Mascarenhas

The process of image blurring and addition of noise in an incoherent optical system can often be described by the equation

$$y(\alpha,\beta) = \int_a^b \int_a^b x(\xi,\eta)h(\alpha,\xi;\beta,\eta)d\xi d\eta + \eta(\alpha,\beta) \qquad -\infty < \alpha,\beta < \infty$$

where $x(\alpha,\beta)$ denotes an ideal image, $y(\alpha,\beta)$ is the observed image, $\eta(\alpha,\beta)$ is an additive noise field, and $h(\alpha,\xi;\beta,\eta)$ is the image system spread function. When a restoration problem is to be solved with a digital computer to estimate $x(\alpha,\beta)$, a discretization has to be performed. By using a lexicographic ordering, it is possible to reduce the resulting two dimensional data arrays into vector format. The following equation describes the discrete model

$$\underline{y} = \underline{B}\,\underline{x} + \underline{n}$$

where

$\underline{y}$ = $(M^2 \times 1)$ vector of observed values
$\underline{B}$ = $(M^2 \times N^2)$ blur matrix
$\underline{x}$ = $(N^2 \times 1)$ vector of original pixel values
$\underline{n}$ = $(M^2 \times 1)$ vector of noise components.

In general the entries of the blur matrix depend on both the kernel of the integral equation and the weights of quadrature integration. In the simulation experiments described in this section these weights have been assumed to have unity value.

Figure 1 describes the data arrays involved when an overdetermined model for restoration is used. This leads to the following structure for the blur matrix $\underline{B}$

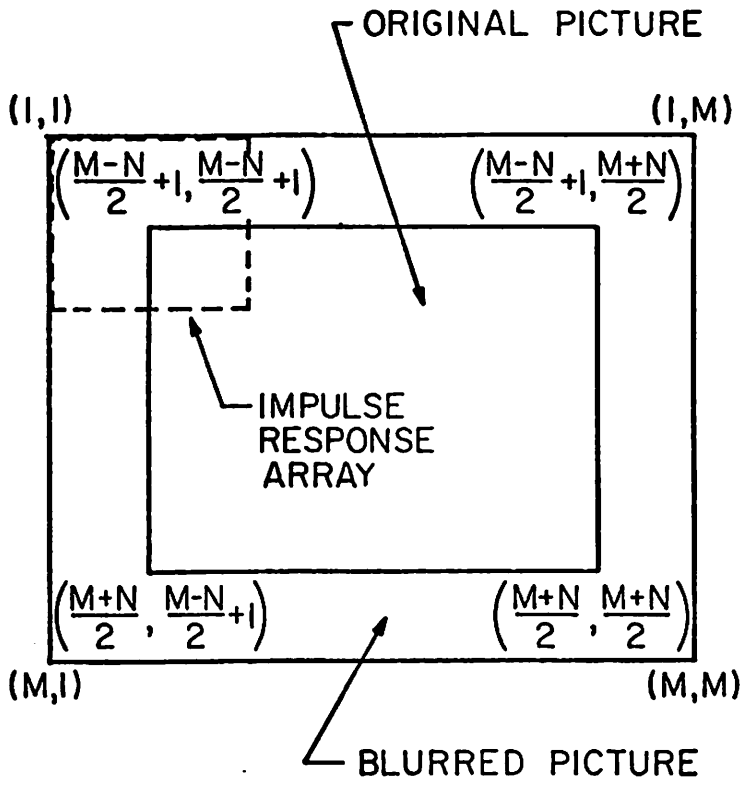


Figure 4.3-1. Data arrays in the overdetermined model

$$\underline{B} = \begin{bmatrix} \underline{\underline{B}}_{1,1} & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ \underline{\underline{B}}_{2,1} & \underline{\underline{B}}_{2,2} & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & & & & & \\ \cdot & \cdot & & & & & \cdot \\ \cdot & \cdot & & & & & \cdot \\ \underline{\underline{B}}_{L,1} & \underline{\underline{B}}_{L,2} & & & & & \\ 0 & \underline{\underline{B}}_{L+1,2} & & & & & \\ 0 & 0 & & & & & 0 \\ \cdot & \cdot & & & & & \underline{\underline{B}}_{N,N} \\ \cdot & \cdot & & & & & \cdot \\ \cdot & \cdot & & & & & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & \underline{\underline{B}}_{M,N} \end{bmatrix}$$

First $\underline{B}$ is partitioned into submatrices $\underline{\underline{B}}_{i,j}$ of size (M × N). Then each submatrix has a similar structure, being composed by a nonzero diagonal band of elements.

Two expressions for the blur have been used. The first simulates the effect of atmospheric turbulence over a long exposure. The spread function is given by

$$h(\alpha_i, \xi_m; \beta_j, \eta_\ell) = \exp\left\{ -\frac{(\alpha_i - \beta_j)^2}{b_V} + \frac{(\xi_m - \eta_\ell)^2}{b_H} \right\}$$

where the coefficients b_V and b_H control the amount of blur imposed on the vertical and horizontal directions, respectively. The second blur function, also space invariant and in separable form, simulates the effect of a diffraction limited optical system as given by

$$h(\alpha_i, \xi_m; \beta_j, \eta_\ell) = \left[\frac{\sin\left(\frac{\alpha_i - \beta_j}{b_V}\right)}{\left(\frac{\alpha_i - \beta_j}{b_V}\right)} \right]^2 \left[\frac{\sin\left(\frac{\xi_m - \eta_\ell}{b_H}\right)}{\left(\frac{\xi_m - \eta_\ell}{b_H}\right)} \right]^2$$

For white noise, the best linear unbiased estimator of the original picture is obtained as

$$\underline{\hat{x}} = (\underline{B}^T \underline{B})^{-1} \underline{B}^T \underline{y}$$

and the covariance matrix of this estimator is

$$\underline{V}_{\underline{\hat{x}}} = \sigma^2 (\underline{B}^T \underline{B})^{-1}$$

where σ^2 is the variance of the noise. The amount of uncertainty on the estimators will depend on the degree of singularity of $(B^T B)$. A possible measure of this is given by the condition number of the matrix $\underline{B}$, which can be expressed by the ratio of the largest to the smallest singular value of the matrix $\underline{B}$ [1].

Figures 2 and 3 illustrate curves of condition number versus blur coefficient, for a given number of original pixel values (N=8) [2]. In the curves the number of observed values M is varied while maintaining the structure of the blur matrix described previously. The existence of a maximum of the condition number curves can be explained in terms of the truncation of the point spread function displayed in Figure 4. In fact, for increasing M, the number of points where this function can be nonzero is increased, and the effect of the truncation starts only for higher blur coefficients. Consequently, the curves for different values of M have essentially a common ascending branch and the descending part starts at varying parts for different values of blur coefficients. If there were no truncation, the curve would approach infinity very fast, the asymptotic value being obtained for the smoothest possible kernel, with constant value unity, implying a blur matrix with rank one. With the truncation, the curves show a descending branch that begins at the point where the increasingly wider kernel starts to be cut down substantially. For greater amounts of blur, the curves tend to a finite value that depends on M.

The curves of figs. 2 and 3 can be used as a guide for the choice of the number of sampling points, once the number of quadrature nodes is fixed.

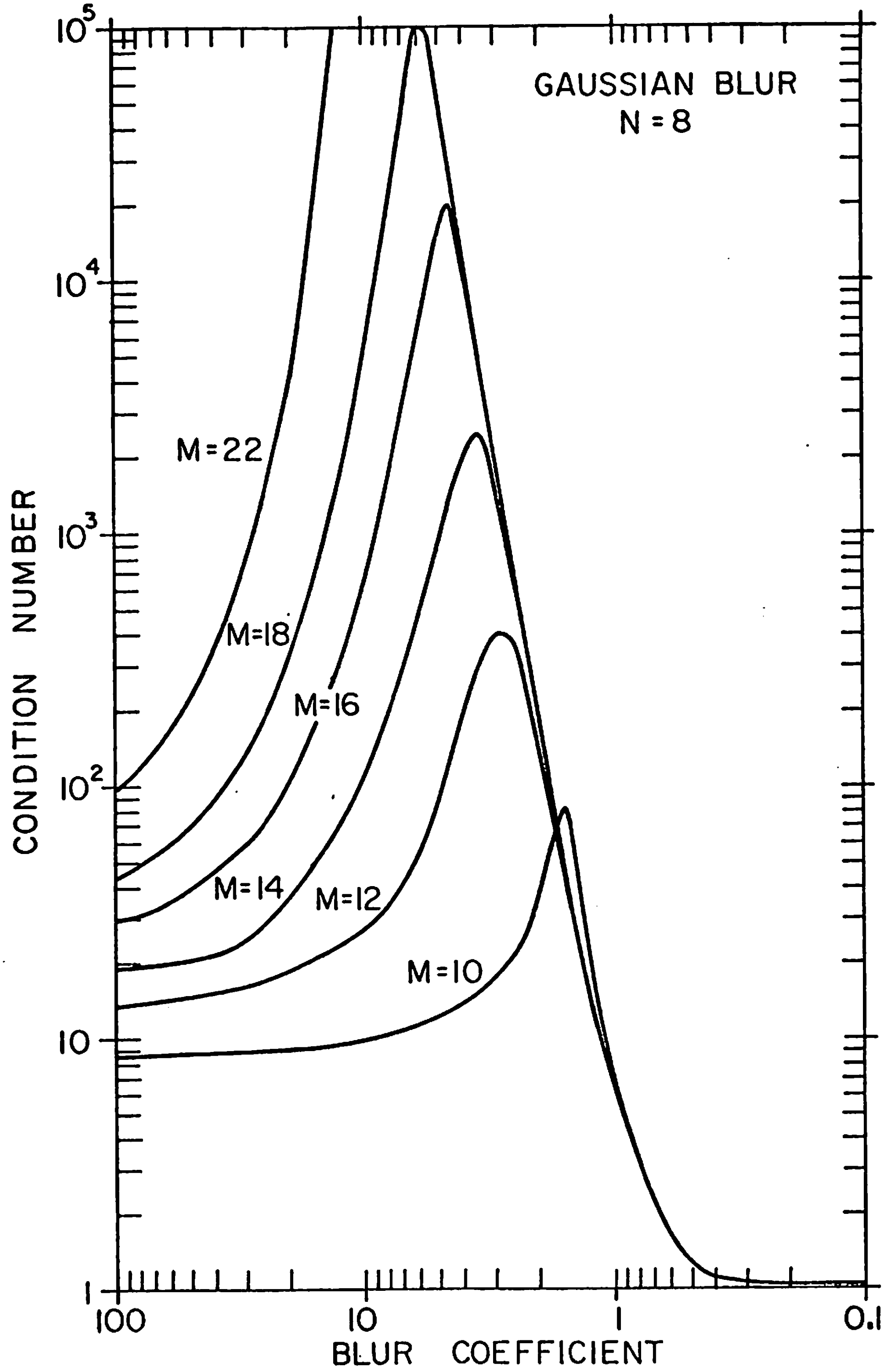


Figure 4.3-2. Condition number curves for different number of sampled values - Gaussian blur.

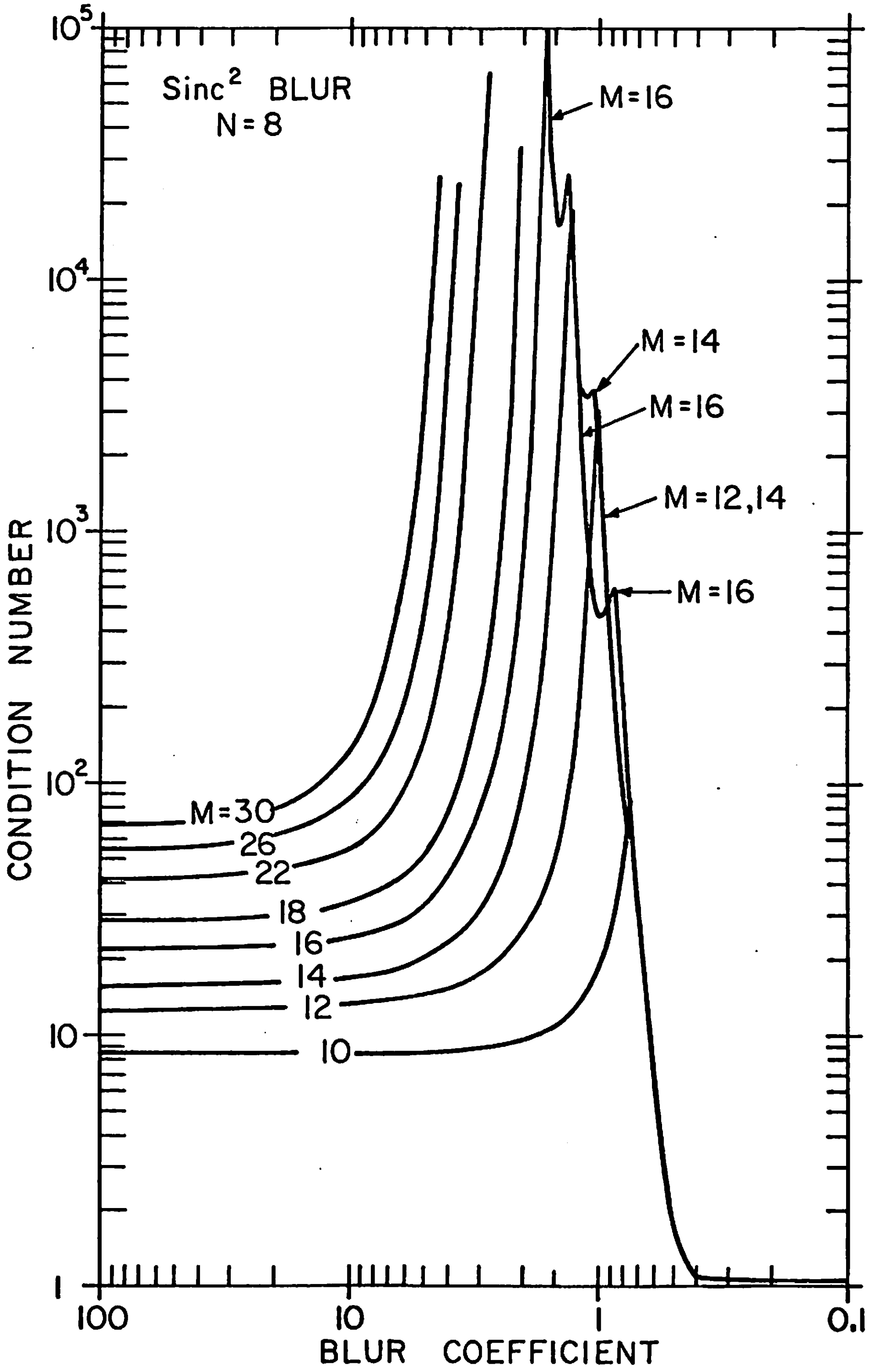


Figure 4.3-3. Condition number curves for different number of sampled values - Sinc ^{2} Blur

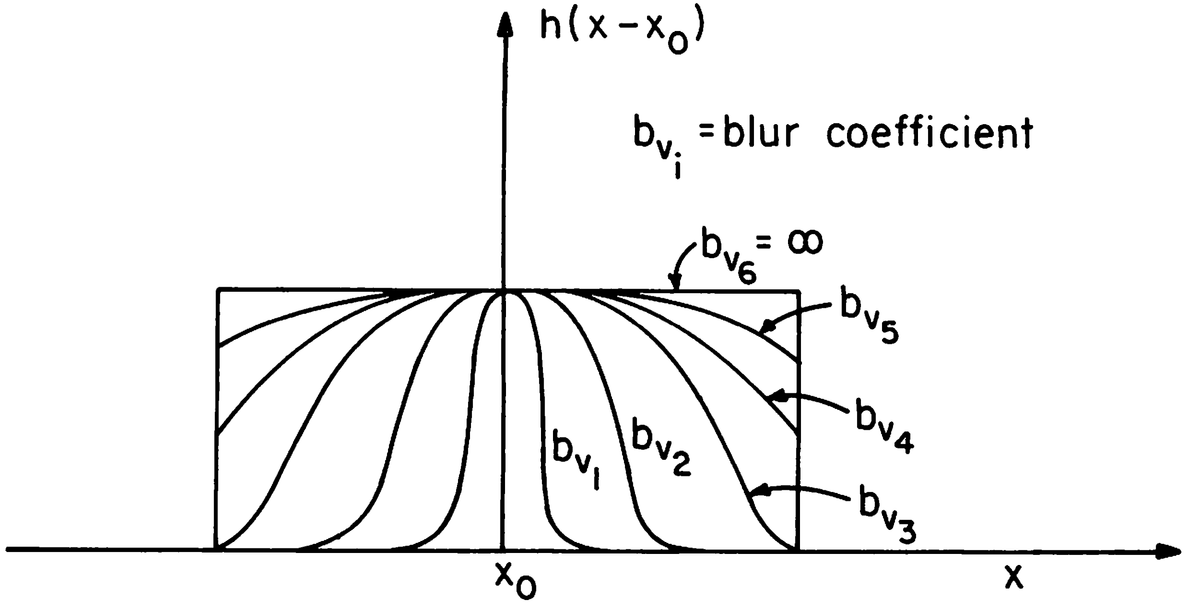


Figure 4.3-4. Truncation of the point spread function.

For a very small amount of blur all curves coincide so that the designer may choose M = N with almost no error. In this case blur plays no role, only noise affects the restoration. With increasing blur, different numbers of sampling points give different values of condition number. If a curve on an ascending branch is chosen, truncation has no effect on the kernel, but a high condition number imposes high variances on the estimators. If a curve on a descending branch is selected, lower variances of the estimators are obtained, at the price of error on the estimation of the continuous function due to the truncation error in the discrete model. Therefore, a trade-off between the variance of the estimators and the modeling error can be characterized.

Although these conclusions are drawn based on the particular model discussed in this section, they are more general. Since the inverse of the integral operator that describes the blur is unbounded, therefore, the closer the discrete model follows the continuous one, the more ill conditioned the former model tends to be. A move in the opposite direction reduces singularity but imposes modeling errors. This inevitable dilemma can only be broken with correct a priori knowledge about the solution.

References

1. B. W. Rust and W. M. Burrus, Mathematical Programming and the Numerical Solution of Linear Equations, American Elsevier, New York, 1972.

2. N. D. A. Mascarenhas, "Digital Image Restoration Under a Regression Model - The unconstrained, Linear Equality and Inequality Constrained Approaches," Ph.D Dissertation, University of Southern California, Department of Electrical Engineering, January 1974.

4.4 Spline Function Restorations

Steve Hou and Harry C. Andrews

A variety of models for linear imaging systems have been postulated for processing by a general purpose digital computer. Three of these models are:

Continuous-continuous. Let

$$g(x,y) = \iint f(\zeta,\eta)h(x,y,\zeta,\eta)d\zeta d\eta$$

where the image, g, object, f, and point spread functions, h, are all described in continuous notation.

Discrete-discrete. Let

$$\underline{g} = \underline{H}\,\underline{f}$$

where the image and object have been scanned or stacked into $N^2 \times 1$ column vectors and $\underline{H}$ is an $N^2 \times N^2$ matrix.

Continuous-discrete. Let

$$\underline{G} = \iint \underline{H}(\zeta,\eta)f(\zeta,\eta)d\zeta d\eta$$

where the image is a matrix of entries g_{ij} and the object is continuous and passes through a matrix of continuous point spread functions. Thus

$$g_{ij} = \iint h_{ij}(\zeta,\eta)f(\zeta,\eta)d\zeta d\eta$$

Figure 1 indicates the geometry associated with this model. Possibly this model is the most realistic of the above alternatives in the sense that the object is continuous before being sampled and the image (once in the computer) is inherently discrete.

In attempting a restoration of the discrete image $\underline{G}$ to a better estimate $\hat{f}(\zeta,\eta)$ of the continuous object $f(\zeta,\eta)$ it is clear that the degrees of freedom in the object are infinite while those in the image are finite. Thus it is necessary to provide a translation of the approximation of a continuous function to the estimation of a finite representation of a functional form of that continuous object. The mechanism of spline functions are suggested for this purpose. If $\hat{f}(\zeta,\eta)$ denotes the restored object, then

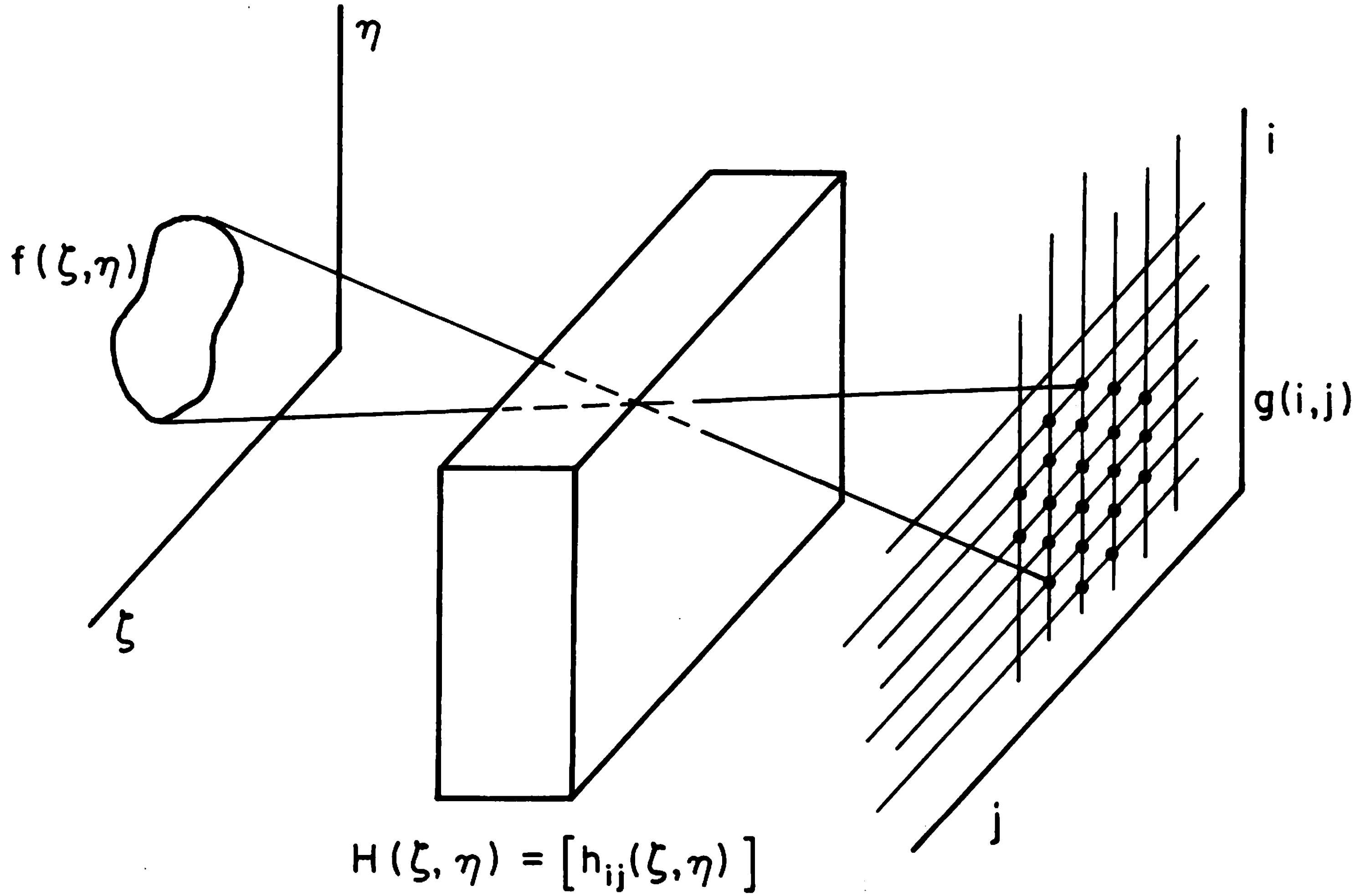


Figure 4.4-1. Continuous discrete model.

let

$$\underline{P} = \iint \underline{H}(\zeta,\eta)\hat{f}(\zeta,\eta)d\zeta\, d\eta$$

It is then possible to formulate the restoration expression in terms of the minimization of an objective function

$$W(f) = tr\{[\underline{G}-\underline{P}]^T[\underline{G}-\underline{P}]\}$$

In this case, the estimation is a least squares restoration. Additional constraints on the objective function may be desired. One such constraint might be that the restored object not oscillate wildly; in particular that the second derivative of that object be minimum. Then using the method of Lagrange multipliers the objective function becomes

$$W(f) = tr\{[\underline{G}-\underline{P}]^T[G-P]\}+\lambda\iint(\nabla^2\hat{f}(\zeta,\eta))^2d\zeta d\eta$$

Now expanding the restored object into a set of nonorthogonal basis functions determined by bi-cubic B splines on a uniform two dimensional grid, one obtains

$$\hat{f}(\zeta,\eta) = \sum_{\alpha}\sum_{\beta} c_{\alpha\beta}S_{\alpha}(\zeta)S_{\beta}(\eta)$$

where $S_\alpha(\zeta)$ is the cubic B spline. Such an expansion allows a priori knowledge of the second derivative of $\hat{f}(\zeta,\eta)$ in terms of the coefficients $c_{\alpha\beta}$.

When the objective function is minimized, the matrix equation

$$(\underline{A} + \lambda\underline{B})\,\underline{c} = \underline{d}$$

is obtained where $\underline{A}$ is defined by the blurred spline functions, $\underline{B}$ is the differential spline functions and highly banded, $\underline{c}$ is the unknown coefficient vector and $\underline{d}$ is defined by the image $\underline{G}$.

Techniques for solving $\underline{c}$, the equality constrained least squares problem described above, are being investigated in the framework of spline functions and image arrays. In addition inequality constraints on the $\underline{c}$ vector such that $\hat{f}(\zeta, \eta)$ is positive and bounded are being considered. The approach looks fruitful as the B-spline basis functions provide a unique set of non-orthogonal basis functions which are inherently positive and whose coefficients provide second derivative (smoothness) criterion control.

4.5 Space-Variant Restoration of Astigmatism and Curvature of Field

Alexander A. Sawchuk and M. Javad Peyrovian

As discussed in previous work, the general problem of image restoration with a space-variant point-spread function (SVPSF) degradation is exceedingly difficult to solve due to the high system dimensionality [1, 2]. To get a practical solution, it is necessary to completely exploit the degrading system symmetry and structure in order to reduce the effective dimensionality. The approach has wide applications in the restoration of motion blur [1, 3] and certain geometrical optics aberrations [2, 5, 6]. This section describes an extension of the technique to astigmatism and curvature of field.

A great deal of insight into the restoration problem is obtained by careful examination and derivation of the degrading system point-spread function (PSF). The aberration function [1, 2, 4] of geometrical optics for astigmatism and curvature of field are given by

$$x_1 - u_1 = \Delta_1 = (2C+D)u_1^2 \varepsilon_r \cos \varepsilon_\theta \qquad \text{(1a)}$$

$$x_2 - u_2 = \Delta_2 = D\, u_1^2 \varepsilon_1 \sin \varepsilon_\theta \qquad \text{(1b)}$$

where (x_1, x_2) and (u_1, u_2) are the rectangular image and object coordinates respectively, ε_r and ε_θ are ray intercepts in the exit pupil of the optical system, and C and D are constant coefficients describing the degree of astigmatism and curvature of field, respectively. Using previous results [1,2] an equivalent system SVPSF $h(\underline{x}, \underline{u})$ which describes the degradation can be derived. For a circular exit pupil of radius R this SVPSF is

$$h(x_1,x_2,u_1,u_2=0)=\begin{cases}\dfrac{1}{D(2C+D)u_1^4}\ , & \dfrac{x_2^2}{R^2R^2u_1^4}+\dfrac{(x_1-u_1)^2}{(2C+D)^2u_1^4\,R^2}\le 1\\ 0\ , & \text{elsewhere}\end{cases}\qquad(2)$$

with the input object impulse function at $(u_1,u_2=0)$. The function of eq. (2) is shown in figure la for the impulses at various locations in the (u_1,u_2) plane. The regions of nonzero response are defined by ellipses which increase in size as the square of the radial distance. At the same time, the amplitude of the response decreases inversely with u_1^4. Because of the inherent circular symmetry, the amplitude and shape of the response is a function of radial distance only and not a function of angle θ, as shown in figures la and lb. Thus one reduction of system complexity is obtained by rewriting the aberration functions of eq. (1) in polar coordinate form as

$$x_r-u_r = (2C+D)u_r^2\,\varepsilon_r\cos\varepsilon_\theta \qquad (3a)$$

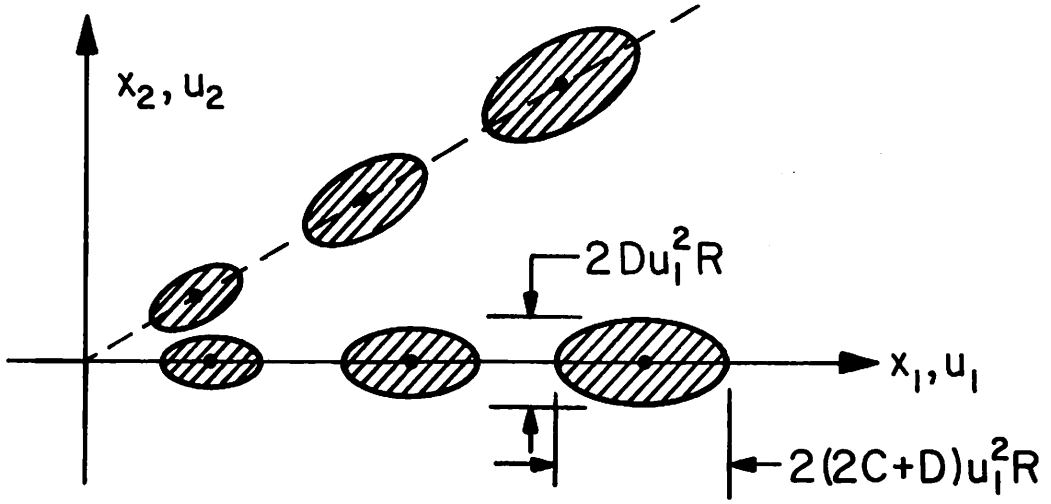
$$x_\theta-u_\theta = \tan^{-1}(Du_r\varepsilon_r\sin\varepsilon_\theta/1+(2C+D)u_r\varepsilon_r\cos\ \varepsilon_\theta) \qquad (3b)$$

where (x_r,x_θ) and (u_r,u_θ) are image and object polar coordinate variables. This procedure suggests that the coordinate transformation restoration (CTR) techniques used for other kinds of space-variant degradations are also valuable here. Performing this transformation to the system described by eq. (3) effectively converts the problem to a space-invariant blur in θ which changes slowly with u_r(eq. (3b)) and a purely radial two-dimensional space-variant blur (eq. (3a)).

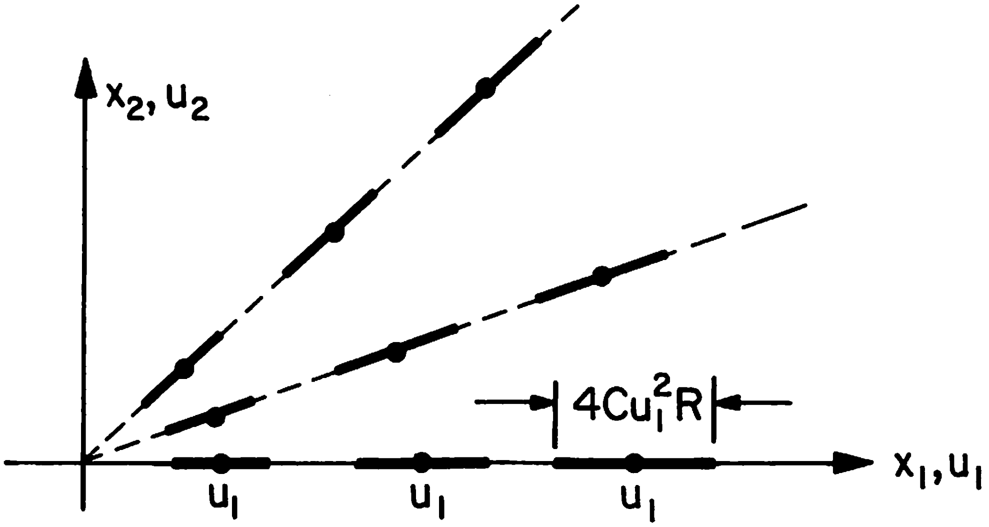
For the case of pure curvature of field with no astigmatism, C = 0 in eq. (2) becomes singular. The SVPSF for astigmatism only has no θ blurring and collapses to a radial space-variant blur $h_c(x_1,u_1,u_2=0)$ obtained by evaluating

$$h_c(x_1,u_1,u_2=0) = \int_{-\infty}^{\infty} h(x_1,x_2,u_1,u_2=0)\,dx_2 \qquad (4)$$

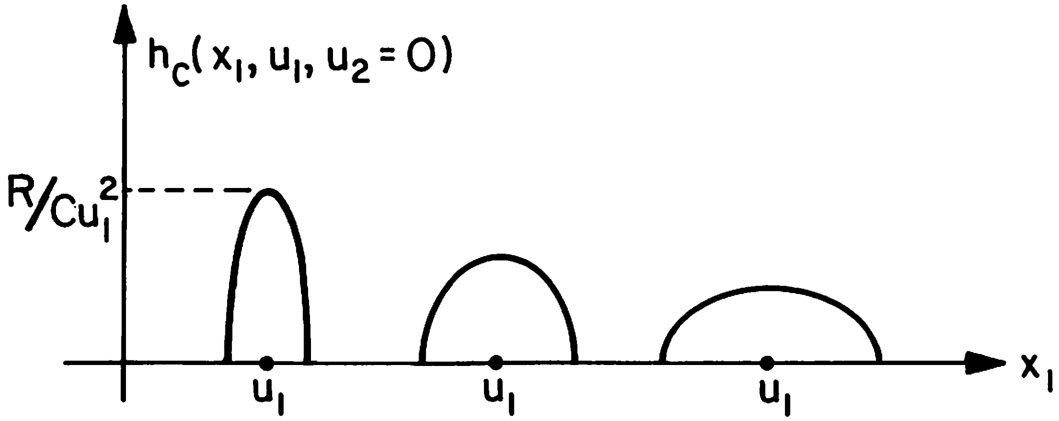
and letting D approach 0. The result is



a) SVPSF for astigmatism (C) and curvature of field (D).



b) SVPSF for astigmatism only with inputs at radial distance u_1.



c) cross-section of astigmatism SVPSF.

Figure 4.5-1. Space variant point spread functions.

$$h_c(x_1, u_1, u_2=0) = \begin{cases} \dfrac{(4C^2R^2u_1^4 - (x_1-u_1)^2)^{\frac{1}{2}}}{2C^2u_1^4} , & u_1-2Cu_1^2\, R \leq x_1 \leq u_1+2Cu_1^2 R \\ 0 , & \text{elsewhere} \end{cases} \qquad (5)$$

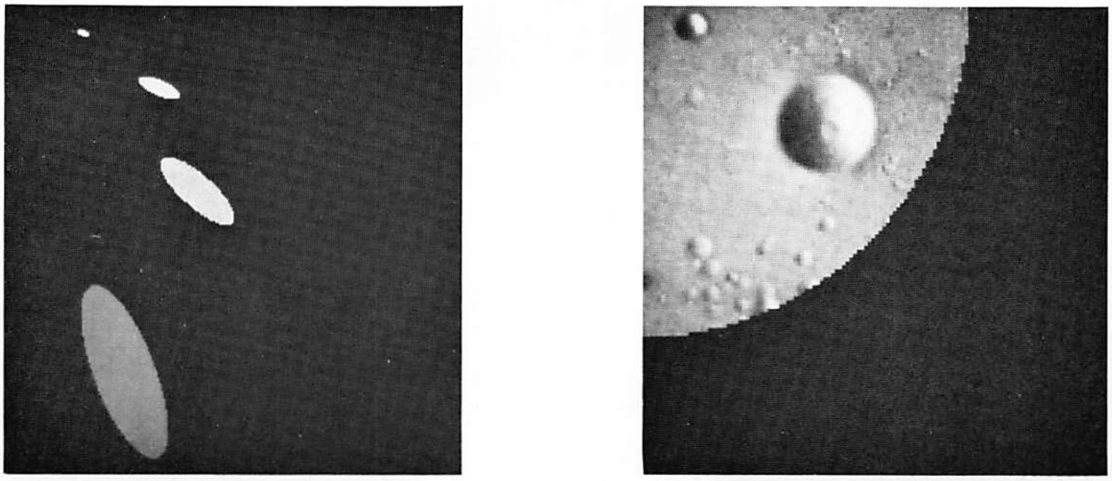
which is shown in figures 2a and 1c. For astigmatism only, there is no blurring in the θ direction and one-dimensional space-variant blur in the radial direction.

The reduced dimensionality of the astigmatism degradation following the polar coordinate transformation makes it possible to write the degradation operation in matrix form as

$$\underline{I}(x_\theta, x_r) = \underline{H}(x_r, u_r)\underline{\mathcal{O}}(x_\theta, u_r) \qquad (6)$$

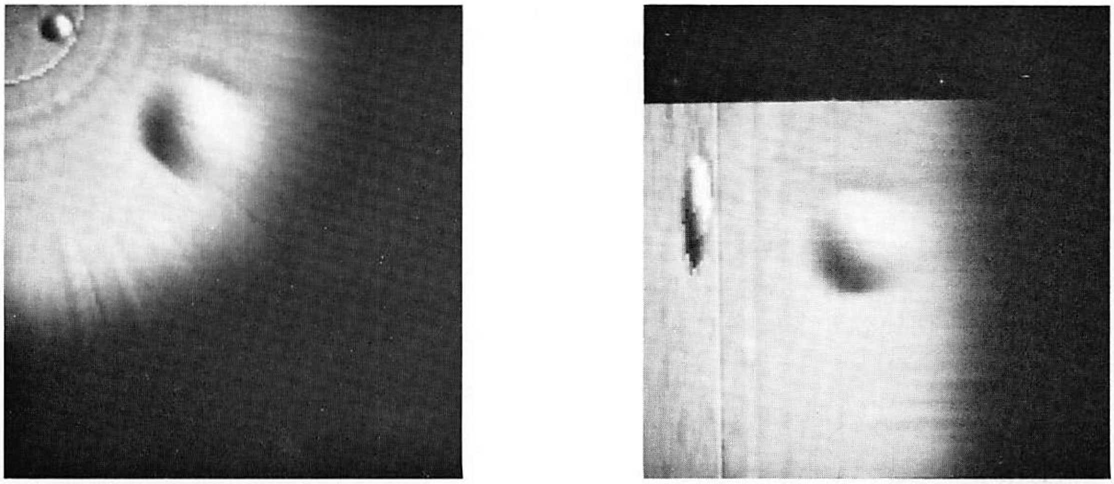
where $\underline{I}(x_\theta, x_r)$ and $\underline{\mathcal{O}}(x_\theta, u_r)$ are one-dimensional line images for each x_θ and $\underline{H}(x_r, u_r)$ is the discrete blur matrix obtained applying a quadrature formula or other approximation to the continuous space description of eq. (5). The degradation due to astigmatism and the restoration process is shown in a series of figures that follow. Figure 2b shows an original undegraded object scene, and fig. 2c shows the effect on the original of space-variant astigmatism described by eq. (5). Note that the amount of blurring increases proportional to the square of the distance from the origin.

One method of image restoration for the space-variant system of eq. (6) is to use a discrete pseudo-inversion technique. As a result of the information reducing aspects of eq. (6), $H(x_r, u_r)$ is usually singular so that pseudo-inversion must be used carefully while avoiding the system noise usually associated with ill-conditioned systems. The analogous technique for space-invariant systems is to use Fourier techniques to diagonalize $\underline{H}$, thus simplifying the inversion if a careful choice of spectral cutoff is made to reduce system noise. For inversion, singular-value decomposition (SVD) techniques [7-10] have been used to obtain a unique pseudo-inverse $\underline{H}^+(u_r, x_r)$ which has then used in the restoration operation



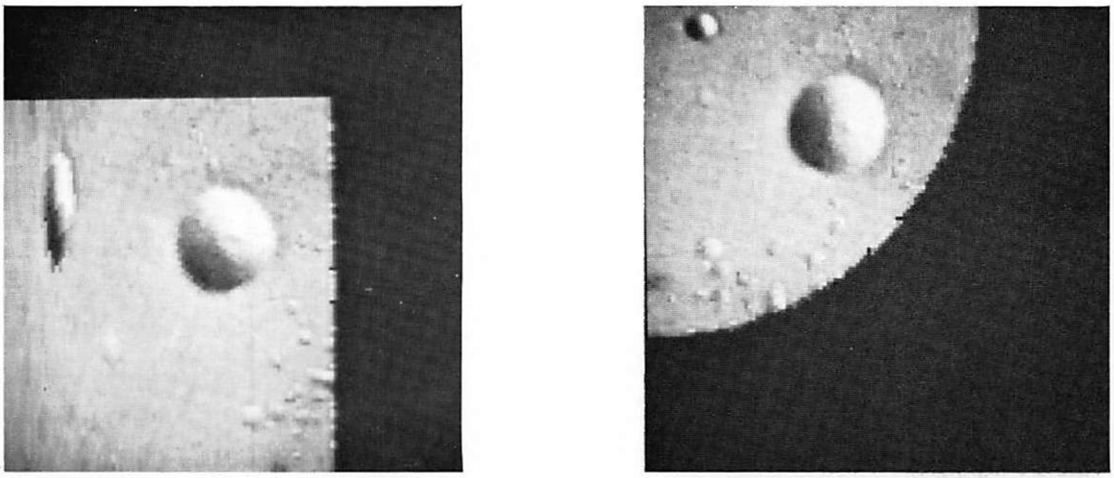
(a) space variant
point spread function

(b) original image



(c) degraded image

(d) polar transformation of (c)



(e) restoration of (d)

(f) restored image

Figure 4.5-2. Example of astigmatism restoration.

$$\hat{\underline{O}}(x_\theta, u_r) = \underline{H}^{+}(u_r, x_r) I(x_\theta, x_r) \qquad (7)$$

The estimate $\hat{\underline{O}}(x_\theta, u_r)$ produced by this procedure minimizes the functional $\| \underline{H}\hat{\underline{O}} - \underline{I} \|$ with the measured image $\underline{I}$ and known $\underline{H}$, and at the same time finds the $\hat{\underline{O}}$ with minimum length $\| \hat{\underline{O}} \|_{min}$ (here the arguments x_r, u_r and x_θ are omitted). This is accomplished by the SVD method in which the matrix $\underline{H}$ is expressed as

$$\underline{H} = \sum_{i=1}^{N} \lambda_i^{\frac{1}{2}} \underline{u}_i * \underline{v}_i^{T} \qquad (8)$$

where the λ_i are the eigenvalues of $\underline{H}\,\underline{H}^T$ (and of $\underline{H}^T\underline{H}$), the $\underline{u}_i$ are the orthonormal eigenvectors of $\underline{H}\,\underline{H}^T$, and the $\underline{v}_i$ are the orthonormal eigenvectors of $\underline{H}^T\underline{H}$. It can be shown that the pseudo inverse $\underline{H}^+$ is given by

$$\underline{H}^{+} = \sum_{i=1}^{K} \lambda_i^{\frac{1}{2}} \underline{v}_i * \underline{u}_i^{t} \qquad (9)$$

where λ_i, $i = 1, K$ are the nonzero eigenvalues. The SVD computer routine of Golub and Reinsch [10] computes these in a numerically stable way; by a judicious choice of the eigenvalue cutoff λ_k a system inversion described by eq. (7) is obtained.

This procedure has been implemented on the 100 X 100 point image of figure 2d blurred by astigmatism. Figure 2e shows the result of a polar coordinate transformation applied to figure 2d; this operation converts the degradation to a space-variant blur in one direction. Figure 2f shows the restoration by SVD, in which the 5 eigenvalues out of 100 whose magnitudes were less than 10^{-5} were set to zero. Following restoration of each line in the (x_r, x_θ) system, an inverse polar coordinate transformation [1] is used to produce the final restoration of figure 2f. Future modifications to this method may use the SVD for Wiener filtering and estimation under various constraints and noise statistics.

This procedure can also be extended for restoration of the general astigmatism and curvature of field case. A brief description of the procedure involves: (a) a polar coordinate distortion;(b) a Fourier transform in the θ direction to decouple space-invariant blur as a function of u_r as expressed in eq. (3b);(c) a space-variant estimate in the radial direction by pseudo-inverse methods to obtain $\hat{\mathcal{O}}(u_r, \lambda)$ for each θ spatial frequency variable λ;(d) an inverse Fourier transform in the θ variable to obtain $\hat{\mathcal{O}}(u_r, u_\theta)$; and (e) a reverse polar coordinate distortion. Although complicated, this procedure enables the general four-dimensional problem of restoration to be converted to a family of two-dimensional space-variant problems with point-spread functions depending on λ.

References

1. A. A. Sawchuk, "Space-Variant Image Restoration by Coordinate Transformations," J. Optical Society of America, Vol. 64, February 1974, pp. 124-130.
2. A. A. Sawchuk, "Restoration of Space-Variant Aberrations by Coordinate Transformations," USC Semiannual Technical Report, USCEE 459, 1 March 1973-31 August 1973, pp. 83-90.
3. A. A. Sawchuk, "Space-Variant System Analysis of Image Motion," J. Optical Society of America, Vol. 63, September, 1973, pp. 1052-1063.
4. M. Born and E. Wolf, <u>Principles of Optics</u>, Fourth Ed. Pergamon Press, London 1970.
5. G. M. Robbins, Mass. Instr. Tech. Res. Lab. Electronics Quart. Prog. Report No. 93, 1969, pp. 243-251.
6. G. M. Robbins and T. S. Huang, "Inverse Filtering for Linear Shift-Variant Imaging Systems, "Proc. IEEE, Vol. 60, July, 1972, pp. 862-872.
7. M. Adler and H. C. Andrews, "Image Data Compression Using Singular Value Decomposition, " USC Semiannual Tech. Report, USCEE 459, 1 Mar. 1973 - 31 August 1973, pp. 36-43.
8. W. K. Pratt and N. D. A. Mascarenhas, "Application of the Matrix Pseudo-inverse to Image Restoration," USC Semiannual Tech. Report, USCEE 459, 1 Mar. 1973 - 31 August 1973, pp. 49-58.
9. H. C. Andrews and C. L. Patterson, "Outer Product Expansions and their Uses in Digital Image Processing," submitted to IEEE Trans. Computers.
10. G. H. Golub and C. Reinsch, "Singular Value Decomposition and Least Squares Solutions," Numer. Math. Vol. 14, 1970, pp. 403-420.

4.6 Fast-Suboptimal Wiener Filtering of Markov Processes

Ali Habibi

Wiener filtering, the classical technique of estimating signals from noisy observations, has been developed primarily for continuous signals. In reference [1], Wiener filtering of discrete signals using various unitary transformations such as the Karhunen-Loeve, Fourier, and the Hadamard transforms has been considered in terms of computational efficiencies. It has been shown that optimal results can be obtained with any unitary transformation of the data, as opposed to operating on the data directly, but optimal transform domain processing requires more computations. However if one is willing to accept a small degradation in performance, the unitary transformations provide considerable computational savings. The Wiener filter transformation concept is extended here to the use of lower-triangular transformations.

Consider an N-dimensional data vector $\underline{X} = (x_1, x_2, \ldots, x_N)^T$ with a covariance matrix $\underline{C}_X$ which has been corrupted by a white noise vector $\underline{V} = (v_1, v_2, \ldots, v_N)^T$ with a covariance matrix $\sigma_V^2 \underline{I}$ resulting in a noisy signal $\underline{S} = \underline{X} + \underline{V}$. Wiener filtering $\underline{S}$ consists of premultiplying $\underline{S}$ by filter matrix

$$\underline{G} = \underline{C}_X (\underline{C}_X + \underline{C}_V)^{-1} \qquad (1)$$

Using any linear transformation $\underline{A}$ on the observation, and the inverse of the transform $\underline{A}^{-1}$ on the filtered signal does not change the estimated signal if the filter response is altered accordingly, that is, if instead of a filter matrix $\underline{G}$, $\underline{G}_A = \underline{A}\,\underline{G}\,\underline{A}^{-1}$ is substituted. If $\underline{A}$ is a unitary operator that transforms $\underline{X}$ to an uncorrelated vector $\underline{Y}$ then vector $\underline{Y}$ possesses a diagonal covariance matrix. If $\underline{A}$ is unitary it follows that $\underline{C}_W$, the covariance of the transformed noise vector $\underline{W}$ remains diagonal. Thus $\underline{G}_A$ is diagonal and Wiener filtering by it requires only N multiplications. However $2N^2$ multiplications are needed to perform the transformation of the input and the output of the filter by $\underline{A}$ and $\underline{A}^{-1}$.

Wiener Filtering by Lower-Triangular Transformations Given any covariance matrix $\underline{C}_X$ one can find lower-triangular matrices $\underline{L}$ and $\underline{L}^{-1}$ such that $\underline{L}\,\underline{C}_X\,\underline{L}^T$ is diagonal [2]. This implies that $\underline{Y} = \underline{L}\,\underline{X}$ is an uncorrelated vector and transforming by $\underline{L}$ and $\underline{L}^{-1}$ requires a total of $2(N^2/2 - N)$ multiplication operations. However, using this transformation on the noisy signal $\underline{S}$ prior to filtering requires filtering $\underline{L}\,\underline{S}$ by $\underline{G}_L$ where

$$\underline{G}_L = \underline{L}\,\underline{G}\,\underline{L}^{-1} \tag{2}$$

Denoting the covariance matrix of the signal vector $\underline{Y}$ and the noise vector $\underline{W}$ by $\underline{C}_Y$ and $\underline{C}_W$ in the transformed domain, respectively,

$$\underline{C}_Y = \underline{L}\,\underline{C}_X\,\underline{L}^T \tag{3}$$

and

$$\underline{C}_W = \underline{L}\,\underline{C}_V\,\underline{L}^T = \sigma_V^2\,\underline{L}\,\underline{L}^T \tag{4}$$

Using eqs. (3) and (4) in eq. (2) gives

$$\underline{G}_L = \underline{C}_Y(\underline{C}_Y + \sigma_V^2\,\underline{L}\,\underline{L}^T)^{-1} \tag{5}$$

Since $\underline{L}$ is non-unitary, unlike the Karhunen-Loeve transformation, $\underline{G}_L$ is not diagonal. Therefore, the total number of operations using the $\underline{L}$ transformation requires $2[(N^2/2)-N] + N^2$ operations, which is still less than the number of multiplications required using the Karhunen-Loeve transform. However, it is more than that required for direct Wiener filtering. Table I summarizes the number of multiplications required for Wiener filtering of a vector of N components for various transformations. For Markov processes, the lower-triangular transformation is used to design a suboptimal Wiener filter that reduces the number of multiplications significantly. This is possible since the $\underline{L}$ matrix for a Markov process is banded. This result will be demonstrated for a first-order Markov process; extension to an mth order Markov process is straightforward.

When x_i, $i = 1, \ldots, N$ is a first-order Markov process, operator L is lower-triangular, banded with one off-diagonal band, i.e.,

$$\underline{L} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ -\alpha_1 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & -\alpha_1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & -\alpha_1 & 1 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ 0 & 0 & 0 & 0 & \cdots & -\alpha_1 & 1 \end{bmatrix} \qquad (6)$$

Transformations by $\underline{L}$ and $\underline{L}^{-1}$ are accomplished recursively as shown on figure 1 each using N multiplications. This also implies that $\underline{L}\,\underline{L}^T$ is a banded matrix of one off-diagonal band which in turn implies that $\underline{G}_L^{-1}$ is a banded matrix of one off-diagonal band.

Filtering $\underline{L}\,\underline{S}$ by $\underline{G}_L$ is equivalent of solving the matrix equation

$$\underline{G}_L^{-1}\hat{\underline{Y}} = (\underline{L}\,\underline{S}) \qquad (7)$$

for the estimated transformed signal $\hat{\underline{Y}}$. An exact solution to eq. (7) for known $\underline{G}_L^{-1}$ requires N^2 operations, but if $\underline{G}_L^{-1}$ is positive definite and posseses only m off-diagonal bands it is shown that eq. (7) can be solved using only 2(m+1) N multiplications [3]. Matrix $\underline{G}_L^{-1}$ can be approximated by a positive definite matrix if $\underline{C}_Y$ in eq. (5) is approximated as

$$\underline{C}_Y \cong (1-\alpha_1^2)\,\underline{I} \qquad (8)$$

For a first order Markov process this approximation corresponds to changing only the first element of $\underline{C}_Y$ from unity to $(1-\alpha_1^2)$ which in turn corresponds to ignoring the transient state of the process. Naturally the effect of the approximation reduces as N increases or as α_1 (correlation of adjacent elements in the process) decreases. Figure 2 shows the effect

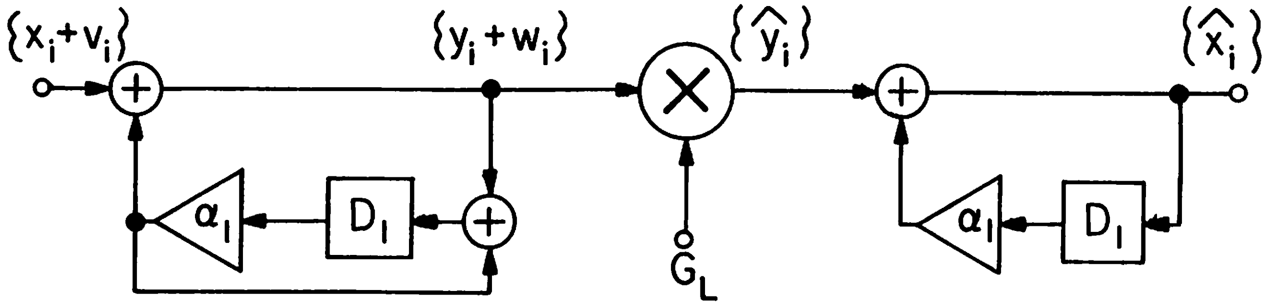


Figure 4.6-1. Wiener filtering of a first order Markov Process using lower-triangular transformation.

Table I

Wiener Filtering Computation Requirements for Various Transformations	
Transform	Number of Multiplication Operations
Identity	N^2
Karhunen-Loeve	$2N^2 + N$
Fourier/Hadamard	$N^2 + 2N \log_2 N$
Lower-Triangular	$2N^2 - 2N$
Lower-Triangular for m-th order Markov "suboptimal"	$2(2m+1)$

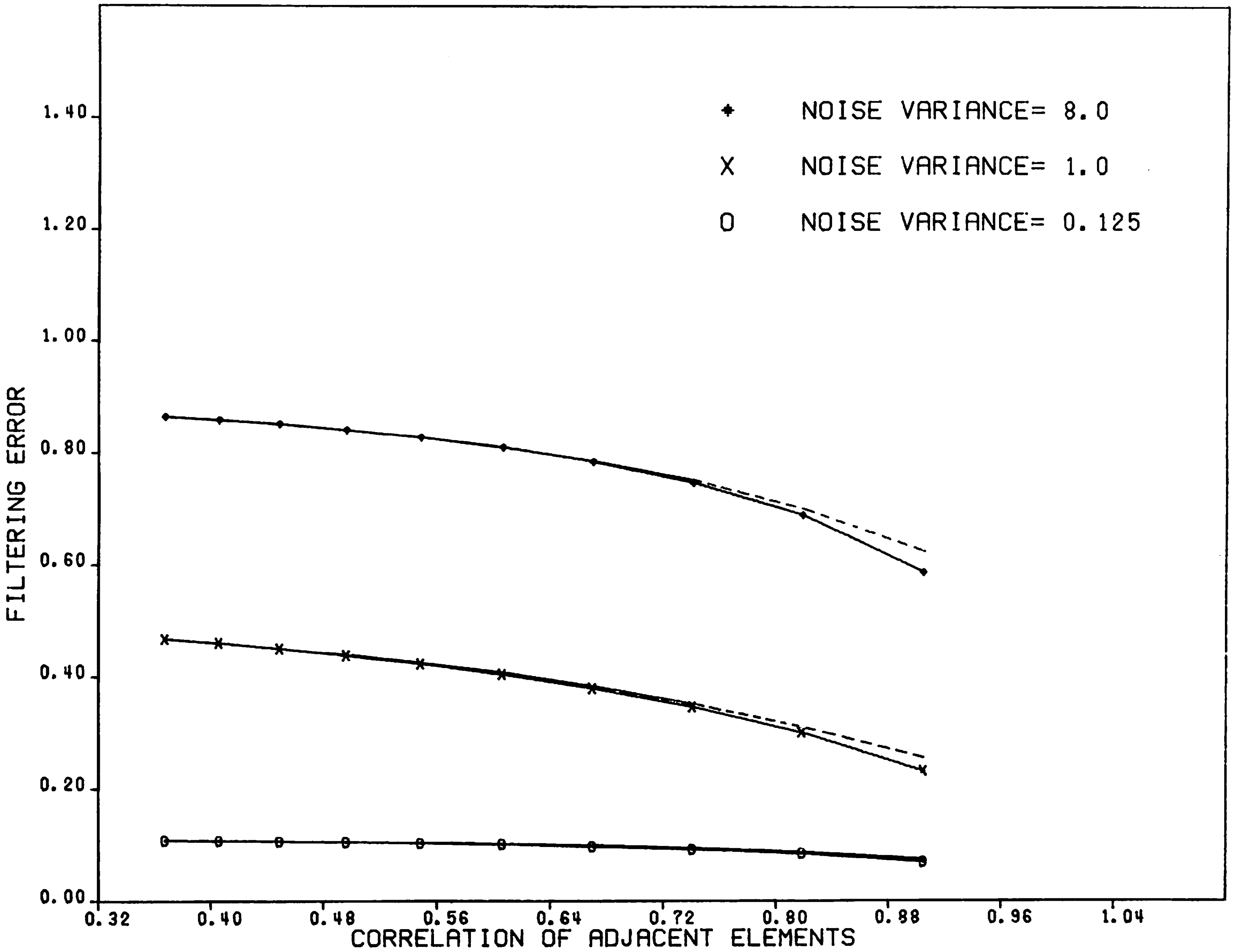


Figure 4.6-2. A comparison of the theoretical performance of the suboptimum and the optimum Wiener filtering of first-order Markov processes.

of this approximation on the theoretical value of the estimation error for various values of α_1 and the signal to noise ratios for a first order Markov process. The effect of this approximation will be more pronounced for higher order Markov processes. However, the required number of computations for this suboptimal Wiener filtering is 2(2m+1)N for an m^{th} order Markov process which is even smaller than the number of operations needed for Kalman filtering of the same signal.

References

1. W. K. Pratt, "Generalized Wiener Filtering Computational Techniques," IEEE Trans. on Computers, Vol. C-21, No. 7, July, 1972, pp. 636-641.
2. A. Habibi and R. S. Hershel, "A Unified Representation of DPCM and Transform Coding Systems," IEEE Trans. on Communications, Vol. COM-22, No. 5, to appear, May 1974.
3. R. S. Martin and J. H. Wilkinson, "Symmetric decomposition of positive definite band matrices," Numerische Mathematic 7, pp. 355-361, 1965.

4.7 Color Image Restoration by Linear Estimation Methods

Clanton E. Mancill

Trichromatic color sensing systems, such as color film or a color television camera, reduce the spectral intensity $C(\lambda)$ at an image point to a triple of numbers (x_1, x_2, x_3) through a set of three integral equations of the form

$$x_i = \int_{\lambda_L}^{\lambda_U} C(\lambda) S_i(\lambda) d\lambda \qquad i = 1, 2, 3 \tag{1}$$

The $S_i(\lambda)$ functions are the spectral sensitivity functions of the three color sensors. The colorimetric model of human color vision describes the sensation obtained by observing the color $C(\lambda)$ in terms of three numbers (t_1, t_2, t_3), called tristimulus values, which are given by the set of equations

$$t_i = \int_{\lambda_L}^{\lambda_U} C(\lambda)\, T_i(\lambda)\, d\lambda \qquad \text{for } i = 1, 2, 3 \tag{2}$$

The color matching functions $T_i(\lambda)$ describe the spectral response of the human eye under the colorimetric model. Since the $T_i(\lambda)$ functions can rarely be expressed as linear weighted sums of $S_i(\lambda)$, the tristimulus values t_i cannot be determined exactly by observing the sensor outputs x_i.

Linear Tristimulus Estimation The purpose of the work described in this section is to estimate the tristimulus values of a color after observing the sensor outputs. The approach used is to discretize the linear integral eqs. (1) and (2) and then to solve the resulting matrix equations for an estimate of the tristimulus values using several well known linear estimation methods. The matrix equations corresponding to eqs. (1) and (2) are

$$\underline{x} = \underline{S}\,\underline{c} + \underline{\varepsilon} \tag{3}$$

$$\underline{t} = \underline{T}\,\underline{c} \tag{4}$$

where the $3 \times n$ matrices $\underline{S}$ and $\underline{T}$ are the discretized sensor and color matching functions (including quadrature formulas for numerical integration over n mesh points), $\underline{t}$, $\underline{x}$, and $\underline{\varepsilon}$ are 3×1 vectors corresponding to tristimulus values, sensor outputs, and observation errors, respectively, and $\underline{c}$ is the $n \times 1$ vector representing the color $C(\lambda)$. The matrices $\underline{S}$ and $\underline{T}$ are known, the vectors $\underline{c}$ and $\underline{\varepsilon}$ are unknown, the vector $\underline{x}$ is observed, and $\underline{t}$ is to be estimated.

Estimation Methods The linear estimation methods which have been considered thus far are enumerated below.

(1) Least Squares Estimate: The rows of the color matching matrix $\underline{T}$ can be approximated (in the least squares sense) by a weighted sum of the rows of the sensor matrix $\underline{S}$. If the norm in n-space is chosen to be the Euclidean distance norm,

$$\|\underline{c}\| = (\underline{c}^T \underline{c})^{\frac{1}{2}} \tag{5}$$

then the least squares tristimulus estimate is given by

$$\hat{\underline{t}}_{LS} = \underline{T}\,\underline{S}^{T}\,(\underline{S}\,\underline{S}^{T})^{-1}\,\underline{x} \tag{6}$$

This estimate neglects the observation noise and is, strictly speaking, a deterministic solution.

2. Minimum Norm Estimate: Another deterministic solution is provided when the equation $\underline{x} = \underline{S}\,\underline{c}$ is solved for $\underline{c}$ under a linear constraint. Since the equation is underdetermined, there are many solutions. The solution which minimizes a norm of the form

$$\|\underline{c}\|_{N} = (\underline{c}^{T}\,\underline{N}\,\underline{c})^{\frac{1}{2}} \tag{7}$$

is chosen, where $\underline{N}$ is a positive definite matrix chosen so that $\hat{\underline{c}}$ maximizes some known a priori property of spectral functions, such as smoothness. For example, $\underline{N}$ could be selected so that the solution to $\underline{x} = \underline{S}\,\underline{c}$ is chosen which has the smallest average squared second difference. The minimum norm tristimulus estimate is given by

$$\hat{\underline{t}}_{MN} = \underline{T}\,\underline{N}^{-1}\,\underline{S}^{T}\,(\underline{S}\,\underline{N}^{-1}\,\underline{S}^{T})^{-1}\,\underline{x} \tag{8}$$

3. Minimum Bias - Minimum Variance Estimate: Assume that the observation error $\underline{\varepsilon}$ has mean and covariance given by

$$E(\underline{\varepsilon}) = \underline{0} \qquad E(\underline{\varepsilon}\,\underline{\varepsilon}^{T}) = \underline{R}_{\varepsilon\varepsilon} \tag{9}$$

An unbiased linear estimator of $\underline{t}$ or $\underline{c}$ is not possible since the equation $\underline{x} = \underline{S}\,\underline{c}$ is underdetermined. The estimator is restricted to the class of minimum bias estimators of minimum variance. The resulting tristimulus estimator is then given by

$$\hat{\underline{t}}_{MBMV} = \underline{T}\,\underline{M}\,\underline{S}^{T}\,(\underline{S}\,\underline{M}\,\underline{S}^{T})^{-1}\,\underline{x} \tag{10}$$

where the M-norm $\|\underline{c}\|_M = (\underline{c}^T \underline{M}\,\underline{c})^{\frac{1}{2}}$ is used in the n-dimensional space and the Euclidean norm is used in three dimensional space. Note that the estimate does not contain $\underline{R}_{\varepsilon\varepsilon}$. If $\underline{M}$ is replaced by the identity matrix, eq. (10) becomes eq. (6), the least squares estimate.

The three estimates considered thus far are of the form

$$\underline{\hat{t}} = \underline{T}\,\underline{\hat{c}} = \underline{T}\,\underline{S}^+\underline{x} \tag{11}$$

where $\underline{S}^+$ is a generalized inverse of $\underline{S}$.

4. Wiener Estimate: If the color vector $\underline{c}$ is itself a random vector whose first two moments are known, then it is possible to derive a Wiener estimator which gives a solution of eq. (3) that minimizes the expected squared error

$$Q = E[(\underline{c} - \underline{\hat{c}})^T(\underline{c} - \underline{\hat{c}})] \tag{12}$$

The moments of $\underline{c}$ and the observation error $\underline{\varepsilon}$ are given by

$$E(\underline{c}) = \underline{\mu}_c \qquad E[(\underline{c} - \underline{\mu}_c)(\underline{c} - \underline{\mu}_c)^T] = \underline{R}_{cc} \tag{13}$$

$$E(\underline{\varepsilon}) = \underline{0} \qquad E[\underline{\varepsilon}\,\underline{\varepsilon}^T] = \underline{R}_{\varepsilon\varepsilon}$$

The resulting $\underline{\hat{c}}$, when multiplied by $\underline{T}$, gives the Wiener tristimulus estimate

$$\underline{\hat{t}}_w = \underline{T}(\bar{\underline{\mu}}_c + \underline{R}_{cc}\underline{S}^T(\underline{S}\,\underline{R}_{cc}\underline{S}^T + \underline{R}_{\varepsilon\varepsilon})^{-1}\underline{x}) \tag{14}$$

Simulation Results The four types of estimates described have been tested using as inputs a set of ten test colors. These represent spectral distributions of natural colors (sky, grass, flesh tones, etc.) seen in daylight. The $\underline{S}$ matrix represents the layer sensitivities of a common reversal color film, corrected for lens absorption. The $\underline{T}$ matrix consists of the color matching curves of the Uniform Chromaticity Scale (UCS) system. Eighty mesh points have been used in the discretization of $S(\lambda), T(\lambda)$, and $C(\lambda)$. The minimum norm estimate utilized a smoothing matrix $\underline{N}$ which minimizes the average squared second difference of $\underline{\hat{c}}$. The minimum-bias, minimum variance estimate contains $\underline{M} = \underline{I}$ giving an estimate equivalent to the least squares estimate. The Wiener estimate utilizes a first order Markov covariance with an adjacent element correlation of 0.95, and assumes a zero observation error. The mean vector $\underline{\mu}_c$ has been set equal to the illumination curve times a constant reflectivity of 0.3. The results are shown in Table 1. The chrominance error is the RMS error distance in uniform chromaticity space averaged over the ten test colors, and the luminance error is the normalized RMS value averaged over the ten colors. It is clear that specifying more and more a priori information on $\underline{c}$ improves the estimate of $\underline{t}$. Further improvement might be made by imposing a nonlinear boundedness constraint on $\underline{c}$, since spectral reflectivities must lie between zero and one at all wavelengths.

Type of Estimate	Chrominance Error (ten color RMS average)	Luminance Error (ten color RMS avg.
Least Squares	.023	.046
Minimum Bias-Minimum Variance, M = I		
Minimum Norm (smoothness norm)	.022	.034
Wiener	.012	.005

Table 1.

Linear Tristimulus Estimator Luminance and Chrominance Errors

4.8 Film Recording of Color Images from a Television Monitor

Robert Wallis

The end result of computer image processing is generally a photographic image. Unfortunately, the gross distortions of grey level and color information that are inherent in film recording are typically ignored. The result is often an image which has been processed to a greater degree by the film than the computer. The goal of this research is the capability of controlling the photographic process not by chemical means, but by an appropriate pre-distortion of the image in order to neutralize the film's distortions.

Color display The color monitor is an additive three color device which generates a spectral distribution given by

$$S(\lambda) = \sum_{i=1}^{3} p_i P_i(\lambda)$$

where $P_i(\lambda)$ $i = 1, 2, 3$ are the spectral distributions of the red, green, and blue phosphors and the p_i, $i = 1, 2, 3$ are the weights given to each phosphor. If the nonlinearities of the CRT are accounted for (gamma correction) the weights in the summation correspond to the drive signals of the monitor. These weights are often denoted as R, G, and B.

Color film Color film [2] is difficult to analyze for two reasons. First, it is a subtractive process. That is, it generates colors by blocking certain wavelengths from white light, whereas an additive system superimposes various spectra. The dyes used as subtractive primaries are cyan, magenta, and yellow. These can be considered notch filters which reject red, green and blue respectively. The use of red, green, and blue primaries for a subtractive system would be unfeasible. This is because red, green, and blue are generated by passing narrow spectral portions of white light. Thus, if any two such dyes were to be mixed, the resultant dye would block nearly all wavelengths, since their passbands would not overlap.

A second difficulty is the nonlinear response of the photographic emulsion to light. The relationship between exposure and the resulting optical density is known as the Hurter-Driffeld curve [2]. If the image intensity at each point of the image to be photographed is predistorted by the inverse of this curve, the film can be forced to yield a linear response.

Tristimulus values The basis of colorimetry is the three color theory of human vision, which states that human perception of color is intrinsically a three dimensional phenomenon. Specifically, for each spectral distribution $S(\lambda)$ color perception can be characterized by three tristimulus values given by

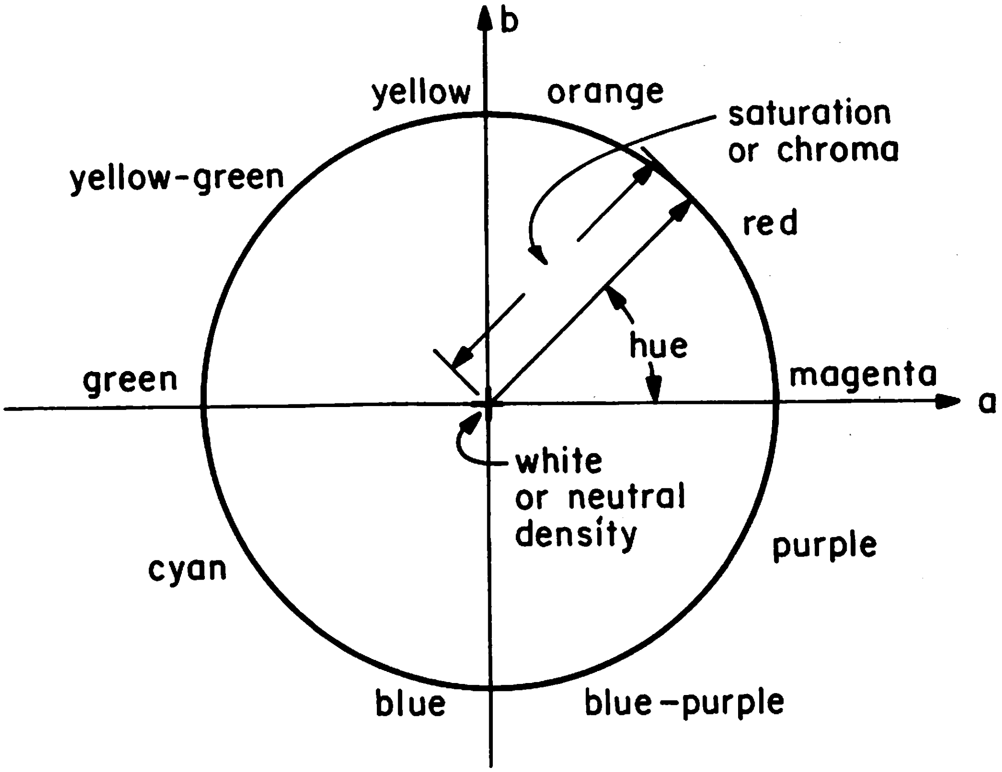
$$t_i = \int T_i(\lambda)S(\lambda)d\lambda \qquad i = 1, 2, 3$$

where the $T_i(\lambda)$ are known as color matching functions. The t_i are scaled such that $t_i = 1$ for a reference white. In order for two colors to match, it is sufficient that their tristimulus values match, but it is unnecessary that their spectral distributions match. This of course presupposes that the colors are compared under similar circumstances. Tristimulus values can be transformed into a number of different coordinate systems in order to facilitate interpretation. One system which has been found useful is the cube root coordinate system [3], which is described in figure 1. The cube root color space has the advantage that equal distances in the space very nearly correspond to equal distances perceptually. Thus, it can be utilized to judge colorimetric errors between pairs of colors.

Computer simulation of color film Using the model of figure 2, a simulation of color transparency film has been performed. The hypothetical inputs are the three individual color monitor phosphors exposed over a wide range of "f-stops". (Each f-stop corresponds to a doubling of the exposure.) The results are given in figure 3. The loops are the paths in color space that result as the exposure is incremented in one-half f-stop intervals. In an attempt to verify the accuracy of the model, red, green and blue fields of a Conrac color monitor have been photographed using Ektachrome-X film over a similar range of exposures. The resulting colors of the transparencies have been measured using a colorimeter. The agreement between theory and experiment is good, as can be judged from figure 3. It is expected that more careful modelling will yield even better correlation. Evidently, the exposure

$$L = 25.29 \left[(100Y)^{1/3} - 18.38\right]$$
$$a = 106.0 \left[(102X)^{1/3} - (100Y)^{1/3}\right]$$
$$b = 42.34 \left[(100Y)^{1/3} - (84.7Z)^{1/3}\right]$$

where X, Y, Z are the C.I.E. tristimulus values



the 3rd coordinate, L (or lightness), is orthogonal to both a and b (i.e. comes out of the page)

Figure 4.8-1. Cube root color coordinate system.

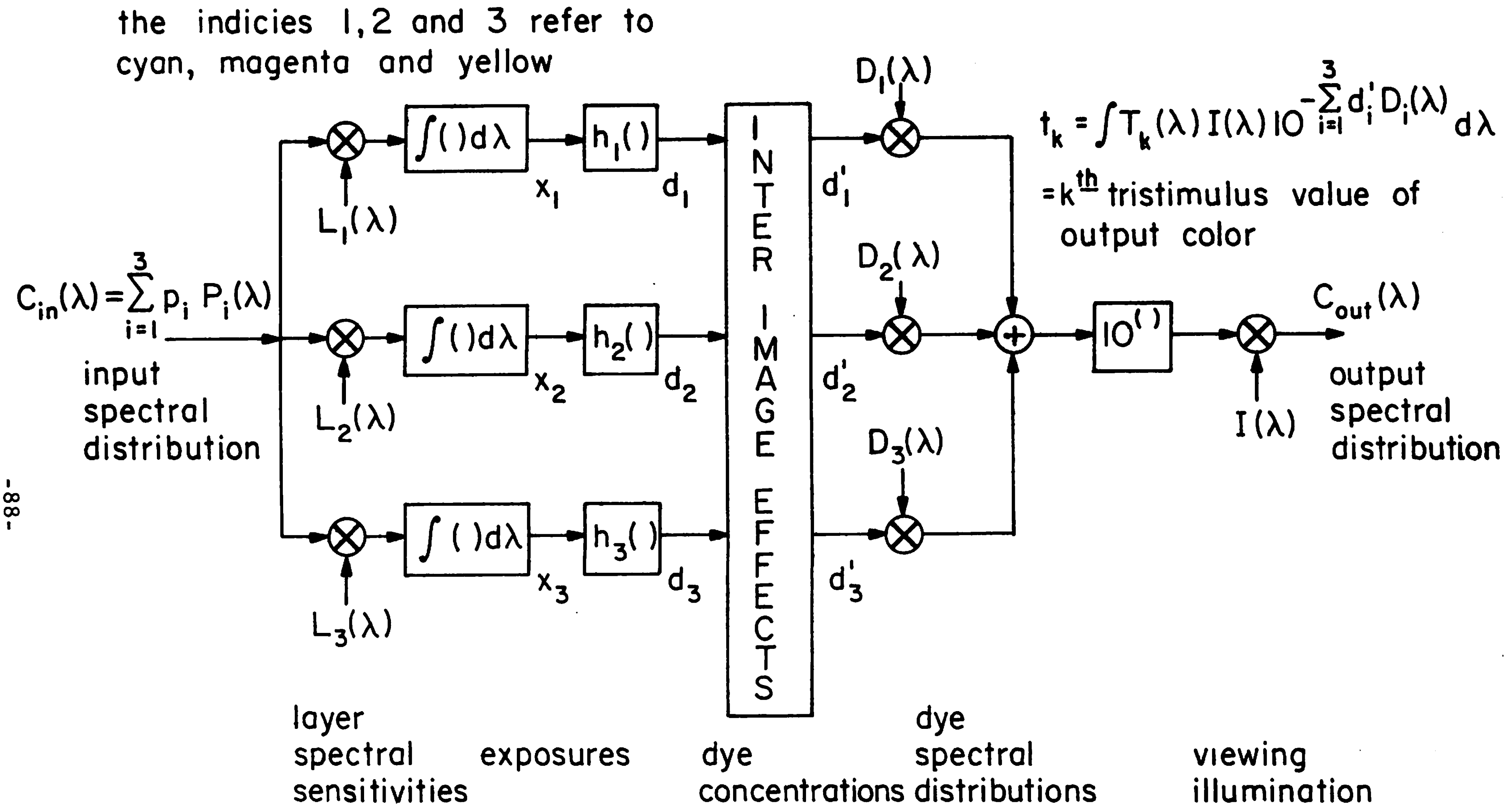
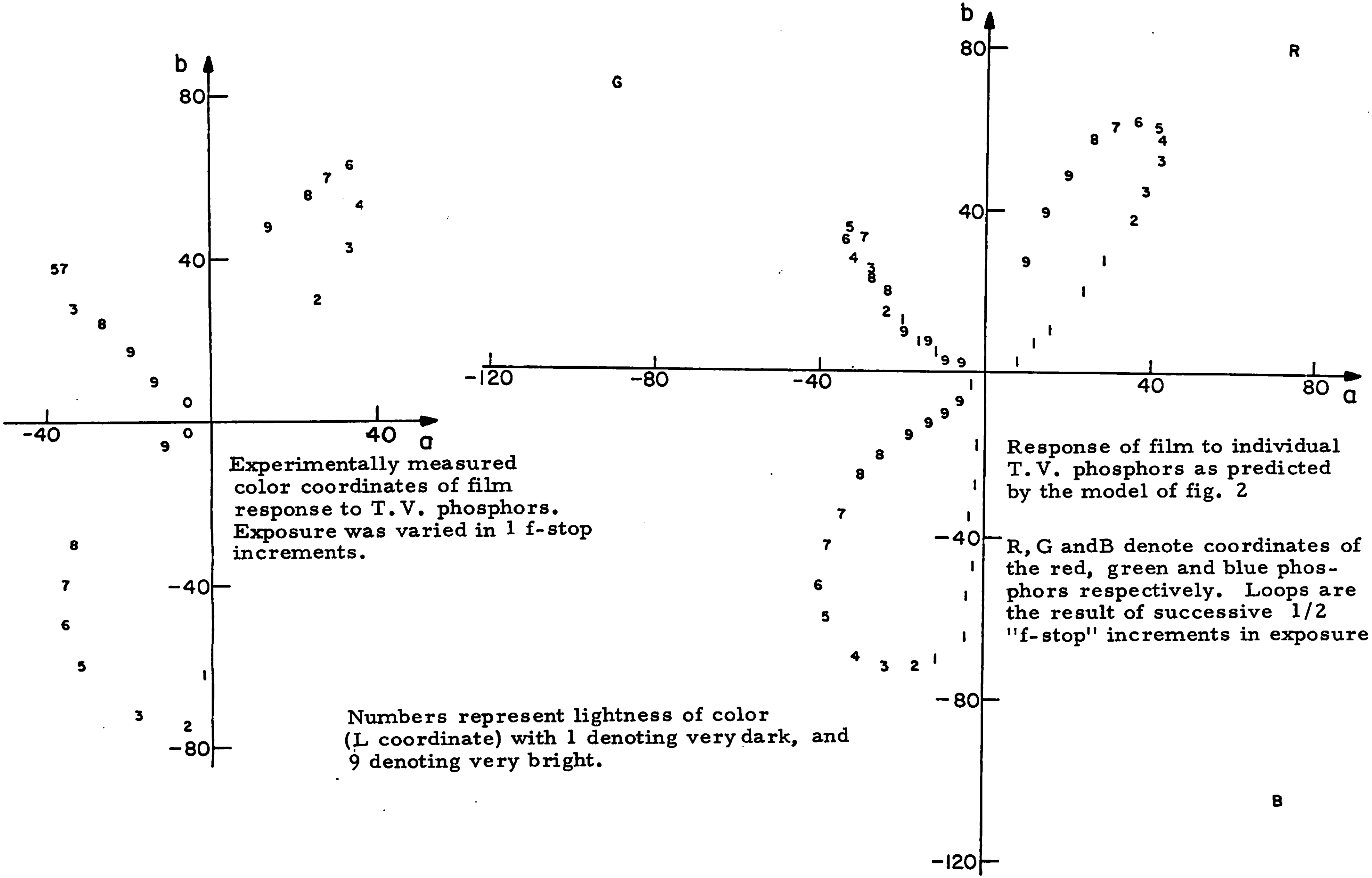


Figure 4.8-2. Mathematical model of color transparency film.



Experimentally measured color coordinates of film response to T.V. phosphors. Exposure was varied in 1 f-stop increments.

Numbers represent lightness of color (L coordinate) with 1 denoting very dark, and 9 denoting very bright.

Response of film to individual T.V. phosphors as predicted by the model of fig. 2

R, G and B denote coordinates of the red, green and blue phosphors respectively. Loops are the result of successive 1/2 "f-stop" increments in exposure

Figure 4.8-3. Theoretical and experimental results.

is a parameter of great concern, in that large shifts in hue and saturation result from even a one-half f-stop change in exposure. For instance, the red shifts towards orange and the blue shifts towards cyan as exposure is increased. Another general result is the loss of saturation as compared with the original color photographed. Note the greater saturation of the original red, green and blue primaries of figure 3. A simple algorithm providing a first order correction to this loss of saturation has been developed. It essentially subtracts white (or neutral density) from each color to be photographed, but maintains the original luminance. Thus, the loss of saturation suffered by film recording is somewhat offset by the pre-saturation. The algorithm is

$$Y = .241 R + .682 G + .077 B$$
$$\alpha = k Y$$

where Y is the luminance of the color and k is a constant which determines the degree of presaturation desired. The R, G and B are then transformed as follows

$$R' = \left(\frac{Y}{Y - \alpha}\right) (R - \alpha)$$
$$G' = \left(\frac{Y}{Y - \alpha}\right) (G - \alpha)$$
$$B' = \left(\frac{Y}{Y - \alpha}\right) (B - \alpha)$$

Note that the luminance remains invariant, but that the differences between R, G and B are exaggerated. The algorithm has been tested on an actual color image, and has been found to result in a marked improvement. An algorithm providing a colorimetrically <u>exact</u> correction is possible, assuming the desired color is within the film's gamut of color reproducibility. That is, given a desired output color, the required input color can be found exactly by iterative techniques. However, for a large image, the required computation time would be prohibitive. A compromise algorithm employing look up tables is being developed which will combine speed and accuracy.

References

1. G. Wyszecki and W. S. Stiles, Color Science, John Wiley and Sons, New York, 1967.
2. R. M. Evans, W. T. Hanson and W. L. Brewer, Principles of Color Photography, John Wiley and Sons, New York, 1953.
3. L. G. Glasser, A. H. McKinney, C. D. Reilly, P. D. Schnelle, Journal of the Optical Society of America, Vol 48, #10, 1958, p. 736.

4.9 Psycho-Visual Pseudo-Color Mapping of Imagery

Werner Frei and Mark Stein

Pseudo-color techniques generally map two-dimensional functions of space variables into a pre-selected gamut of colors, thus creating artificially or pseudo-colored images for easier and more efficient information extraction by human observers. These techniques can be broadly classified into two categories: (a) mappings of a scalar spatial function such as low-contrast achromatic images into some fixed arbitrary color scale so that minute differences of grey in the original image may then appear as easily recognizable color differences [1];(b) mapping of a vector-valued spatial function into color space. Since the human visual system is able to recognize three attributes of colors more or less independently of each other (attributes such as brightness, hue and saturation) one can theoretically map three components of a vector-valued spatial function into three suitably chosen attributes and instruct the observer that for example, brightness corresponds to function A, hue to function B and saturation to function C.

It is felt that many pseudo-color techniques often fail to reach their potential because they are designed with little understanding of the visual system. Arbitrary color assignments, indeed, tend to introduce artificial contours and destroy other valuable features creating more confusion than clarification. Careless pseudo-coloring may thus result in pictures with a mere artistic interest. In this report, these points will be illustrated by an application of pseudo-color to nuclear medicine imagery.

Comparison of Inhalation and Perfusion Lung Scintigrams Lung scintigrams or scans are images of the lung obtained by imaging the output of a γ- ray detector which is placed over a patient who has received an internally administered radio nuclide. In lung scanning, two images are obtained which reflect the lung blood perfusion and ventilation respectively. An area that is normally ventilated but not perfused most likely has a blood clot in the vessels (pulmonary embolus). This distinction is extremely important in clinical therapeutic decision-making. Since the γ- ray camera is an imaging system with a low spatial resolution and the images are extremely noisy, the correspondence determination between the two images is very difficult for human observers, especially when multiple defects are present on both types of scintigrams. Identification of this correspondence requires the ability to determine: (a) degree (severity) of defects; (b) polarity of defects (which image has the defect;(c) anatomic limits (lung outlines). Viewing the perfusion and inhalation scans adjacent to, or superimposed upon one another results in an inability to extract these three information elements accurately.

The perfusion and inhalation scans, denoted as two functions $f_p(x_1, x_2)$ and $f_I(x_1, x_2)$ of the same space variables x_1, x_2, contain combined anatomic (structural) and defect information. In order to combine the information of interest into one color image, a field

$$Y(x_1, x_2) = k_1 \max [(f_p(x_1, x_2), f_I(x_1, x_2))] + k_2 \qquad (1)$$

which preserves structural information, and a field

$$\delta(x_1, x_2) = k_3 [f_p(x_1, x_2) - f_I(x_1, x_2)] \qquad (2)$$

which measures differences in count rates are formed (k_1, k_2, k_3 are arbitrary scaling factors). Next, $Y(x_1, x_2)$ is mapped into luminance and $\delta(x_1, x_2)$ into an axis of complementary colors. In the general case, such an axis is defined in the CIE chromaticity diagram by a straight line

$$y = ax + b \qquad (3)$$

$$y_w = ax_w + b$$

Choosing a zero difference (normal area) to be represented by white,

$(x_w = y_w = 1/3)$, the chromaticity of the displayed pixel is

$$x(\delta) = 1/3 + \delta\,(1 + a^2)^{-\frac{1}{2}} \tag{4}$$

$$y(\delta) = 1/3 + \delta\,(1 + a^{-2})^{-\frac{1}{2}} \tag{5}$$

Since the chromaticity coordinates x, y are given by

$$x = \frac{X}{X + Y + Z} \qquad\qquad y = \frac{Y}{X + Y + Z}$$

where X, Y, Z are tristimuli referred to the CIE non-physical primaries (X), (Y), (Z), one obtains

$$Y = k_1 \max\,[f_p(x_1, x_2), f_I(x_1, x_2)] + k_2$$

$$X = Y\,\frac{x(\delta)}{y(\delta)} \tag{6}$$

$$Z = Y\,\frac{1 - x(\delta) - y(\delta)}{y(\delta)} \tag{7}$$

where x(δ), y(δ) are given by eqs. (4) and (5). Finally the signals to drive a color TV monitor are obtained by a change of coordinates in tristimulus space as given by

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} \begin{matrix} = \\ = \\ = \end{matrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{8}$$

A much more computationally economic solution is obtained by restricting the choice of complementary colors to an axis passing through one of the receiver primaries. Considering further the energy limitations of the display primaries, it is found that the most efficient axis is the green-magenta axis, because the luminous efficiencies of the NTSC primaries give $y_G = 0.585$ and $y_R + y_B = y_M = 0.413$. The pseudo-color mapping, such that $\delta = 0$

maps into white becomes

$$\frac{G}{G + M} = \delta + 0.5 \qquad (9)$$

and

$$Y = 0.585G' + 0.413 M \qquad (10)$$

where $G(x_1, x_2)$ and $M(x_1, x_2) = R(x_1, x_2) = B(x_1, x_2)$ are the green, red and blue signals for a TV display. Letting

$$K = \frac{0.5 - \delta}{0.5 + \delta} \qquad \text{for} \quad -0.5 < \delta \leq 0.5$$

gives

$$G(x_1, x_2) = \frac{Y(x_1, x_2)}{0.585 + 0.413} \qquad (11)$$

and

$$M(x_1, x_2) = k G = R(x_1, x_2) = B(x_1, x_2) \qquad (12)$$

With the mapping described, pseudo-color images are obtained where intensity naturally carries outline information, hue indicates the polarity of the defects, and saturation shows severity, while normal areas appear white. The constant k_2 insures that the intensity does not drop below a level that would make color discrimination difficult.

Initial studies have shown that the ability of the physician to determine accurately correspondences in perfusion and ventilation scans can be greatly enhanced by this single pseudo-colored images as compared to two separate black-and-white images. In addition this rational use of color may increase the detectability of defects. Both results have been recognized in the initial work, but more experience and rigorous psycho-physical testing will be needed to develop confidence of the medical community in this method.

References

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

5. Image Data Extraction Projects

Image data extraction is a name given to a collection of projects that are concerned with the detection of features within an image and methods of measuring these features.

A hybrid optical-digital feature extraction system for images is described in the first report. The system includes a two dimensional Fourier transformation stage implemented by optics followed by a diffraction pattern electronic sensor which supplies digital signals to a computer processor. The hybrid system can be programmed for edge detection, texture detection, or to detect a variety of other image features.

In the next report a description is given of an interactive edge tracing device for photographic transparencies. The device consists of an electro-acoustic sensor that provides coordinate points on an edge contour which are then formatted for remote computer entry.

The following two reports are concerned with techniques for obtaining two and three dimensional perspective views of an object from one or two dimensional density profiles. Methods based upon the Fourier transform, convolution, algebraic reconstruction, and Kalman filtering are explored.

In the last report a method of image boundary estimation is discussed. Experimental results are given for several test images.

5.1 An Optical-Digital System for Feature Extraction From Images

Richard P. Kruger and Ernest L. Hall

A hybrid approach to image feature extraction is described which combines the parallel processing capability of a coherent optical system with the logical processing ability of a computer controlled scanner. A novel feature of the system is provided through the use of an automated film transport which permits parallel combinations of optical and digital processing on local or global areas of an image rather than a cascade of sequential optical-digital or digital-optical operations. This system configuration has been motivated by the desire to combine both digital and optical processing capabilities in a single system, thus permitting a convenient switch from one system to the other when desired for a particular task. For example, the digital system may be used to locate objects and the optical system to measure properties of the objects.

A block diagram of the hybrid system is shown in Figure 1. At present a film transparency is loaded into a fixed frame which may be manually registered in both the optical and digital systems. A future modification will permit access to the film via the film transport shown in Figure 2a. The film scanner is an image dissector type as shown in Figure 2b. The principal components of the optical system are the sampling sensors shown in Figure 3a and the electronic amplifier unit shown in Figure 3b. The controlling computer is a PDP 11/10 which is interfaced via an HP 2100A to an IBM 360/44.

For a practical application of the hybrid system, the transformations and operations should use the particular device to advantage. For example, Fourier transformations are easily computed with the coherent optical system. Nonlinear operations such as threshold computations are easily performed with the computer controlled scanner. Furthermore, for a feature extraction system, a large computer storage is not required since the film is an excellent read only memory and the number of required measurements is usually small.

In the hybrid system a diffraction pattern sampler is used for the optical measurements. The Fourier transform of f(x, y) is defined as

$$F(u, v) = \iint_{-\infty}^{\infty} f(x, y) \exp\{-(uv+vy)\} dxdy$$

and the k^{th} measurement, m_k, is given by

$$m_k = \iint_{R_k(u, v)} |F(u, v)|^2 dudk$$

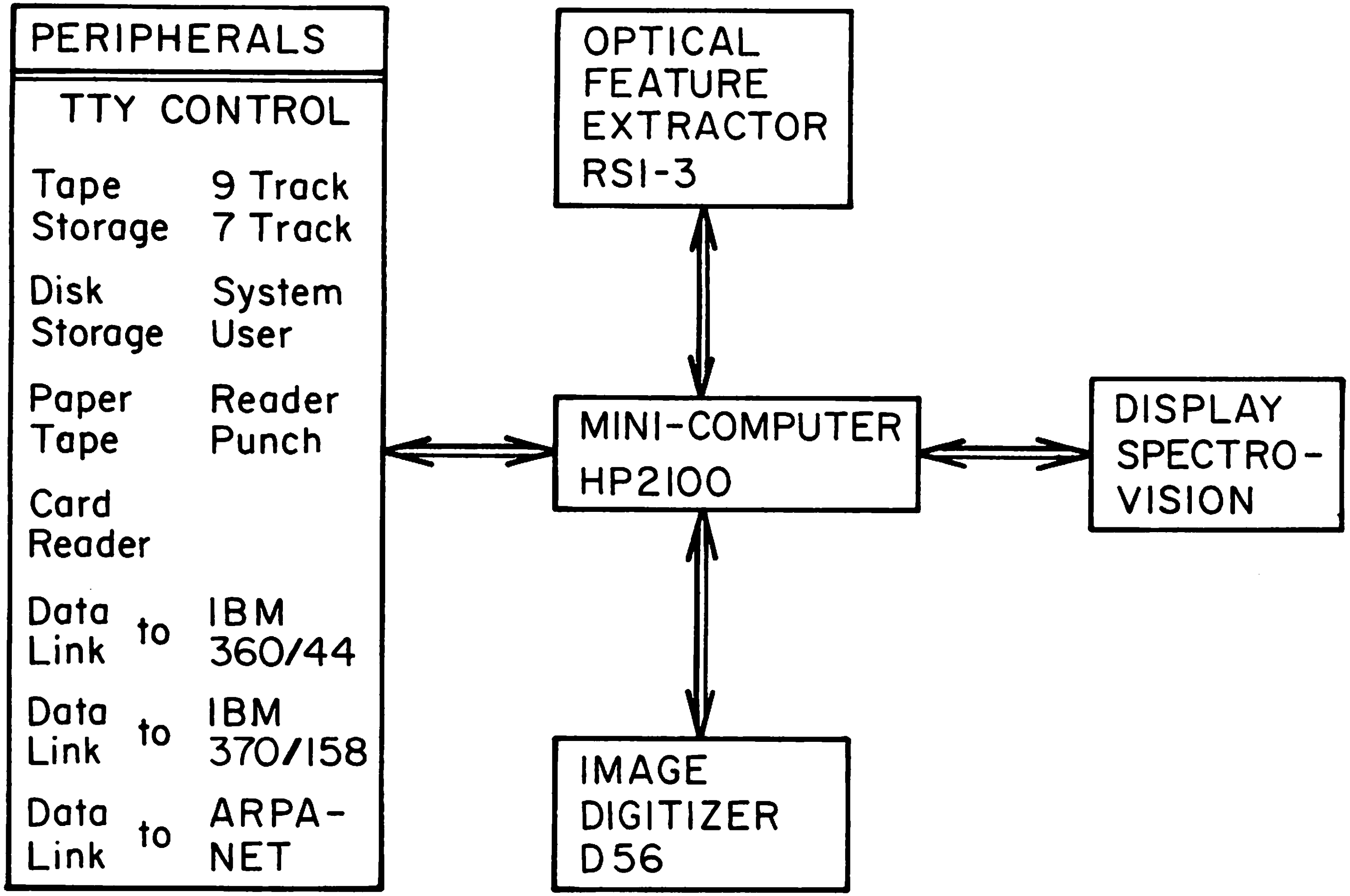
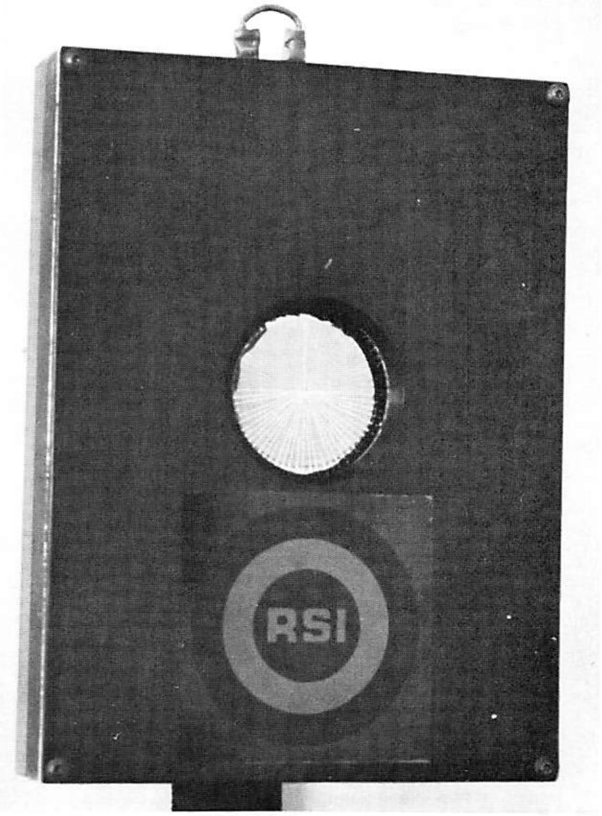


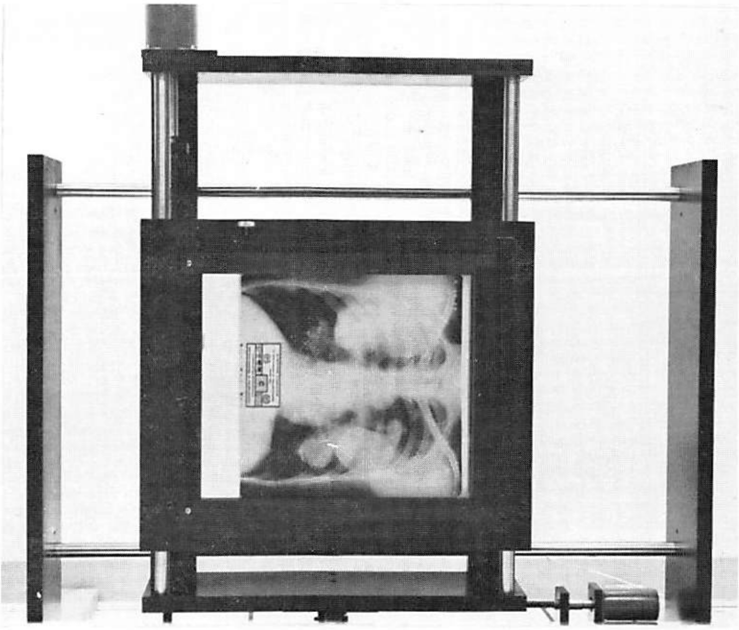
Figure 5.1-1. Block diagram of hybrid optical/digital image processing system.

(a) optical spectrum analyzer

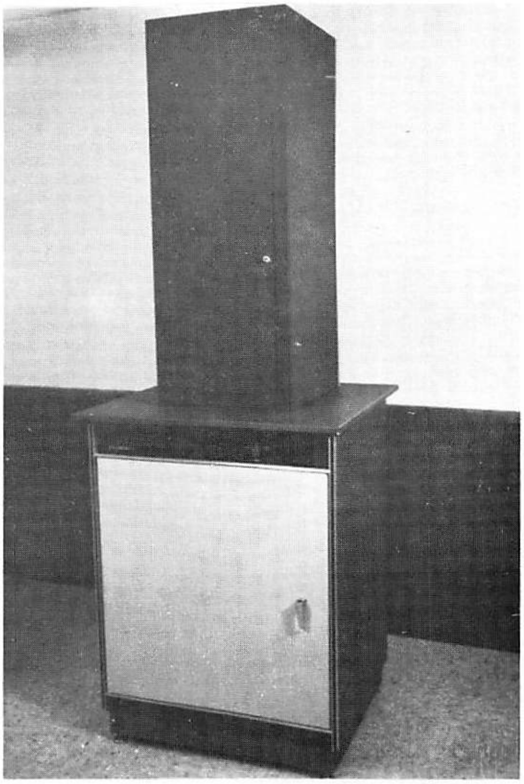


(b) ring and wedge detector assembly.

Figure 5.1-2. Optical spectrum analyzer and detector.



(a) optical X-Y transport assembly.



(b) image digitizer.

Figure 5.1-3. Optical transport and digitizer.

where the regions $R_k(u, v)$ are annular ring or wedge shaped areas in the spatial frequency domain [1]. These measurements have been found useful in many applications. By the use of the film transport, any desired region of the image may be sampled in the above manner.

The digital computation is completely programmable and limited only by the resolution of the scanning device and speed of the computer. Typical operations might include a low resolution scan to produce a computer image, $f(i, j)$, computation of the histogram, $h(f)$ and spatial profiles

$$\varphi(x) = \sum_j f(i, j)$$

$$\varphi(y) = \sum_i f(i, j)$$

which are used to locate a center point (x, y) for the optical measurements.

Another set of digital measurement which are presently used are spatial moments [2] $m_{pq}(f)$ defined by

$$m_{pq}(f) = \iint_{-\infty}^{\infty} x^p x^q f(x, y)dxdy$$

The moment computation may be normalized to obtain translation or size invariance. The moments may also be extended to reflect edge structure of texture patterns. For example, the picture function at any point (ξ, η) may be expressed as a Taylor's series about the point (x, y). The first three terms of this expansion are

$$f(x+\xi, y+\eta) = f(x, y) + \xi \frac{\partial f(x, y)}{\partial x} + \eta \frac{\partial f(x, y)}{\partial y}$$

Therefore, the picture function at the point (ξ, η) may be expressed in terms of the moments

$$m_{pq}[f(x+\xi, y+\eta)] = m_{pq}[f(x, y)] + \xi m_{pq}[f_x(x, y)] + \eta m_{pq}[f_y(x, y)]$$

The operations which combine the optical and digital measurements appear to be highly application dependent and will be studied for particular applications.

A current application of this hybrid system involves a hybrid system for automatic screening of chest radiographs to detect the presence or absence of coal workers black lung disease (pneumoconiosis). The proposed system incorporates coarse digitization of the entire chest x-ray film to either automatically reject a film with improper exposure or detect and measure gross anatomical features in the film. In addition aperture centers in all six lung zones are computed for transmission to the computer controlled film transport.

Figure 4 depicts the output of this first digital process. Shown are least squares polynomial estimates of the lung outlines, rectangular estimates of the six lung zones, and simulated aperture centers for each zone. The second stage of processing will involve the movement of the film transport to each of the aperture centers, measurement of their respective spectral content and conservative initial diagnostic classification. If the film is classified normal in all zones, it will be removed from the system. If this conservative initial classification indicates a possible abnormal situation, the random access capability of the digital scanner will be used to raster scan the suspicous rectangular zones at high resolution. This third stage of processing will involve the computation of textural measures obtained within the boundaries of the interpolated least square estimates of lung zones in question. These textural measures will be used to obtain a second diagnostic classification. This final decision will either place the film into the normal group or place it in one or several disease classes.

This flexible hybrid approach allows for processing economy at each stage of analysis. In general use of the film as a read only memory, different optical detectors, random access digital scanning and a computer controlled film transport will provide a flexible laboratory tool for image analysis and data extraction.

References

1. R. P. Kruger, W. B. Thompson and A. F. Turner, "Automated Computer Diagnosis of Pneumoconiosis from the Standard Posterior-Anterior Chest Radiograph," IEEE Transactions on Systems, Man and Cybernetics, Jan. 1974.

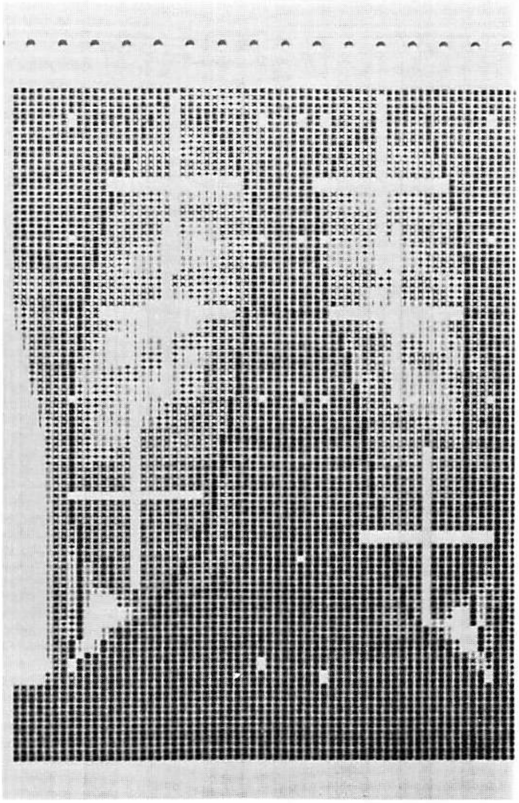


Figure 5.1-4. Lung outline and optical aperture centers.

2. E. L. Hall, W. O. Crawford and F. E. Roberts, "Moment Measurement for Computer Texture Descrimination in Chest X-Rays," Available from IEEE Computer Society Repository.

5.2 A Terminal/Timeshare Based Device for Interactive Boundary Tracing and Analysis

Richard P. Kruger

The ability to input graphic information such as image boundary information and strip chart data in digital form into a computer for analysis can be approached in many ways. The specific facility to be described here is particularly applicable in circumstances which preclude the use of a local minicomputer system, and where it is desired to share a common program library with several remote users.

A system with these general attributes has been installed at the USC-LA County Medical Center and connected to the USC main campus computer via standard dial up telephone lines. The library consists of Fortran language programs accessed under IBM time share option for analysis of various aspects of left ventricular heart function. The system configuration for data collection is shown in Figure 1. The digital input device consists of a graf pen digitizing tablet.

The Graf-Pen graphic tablet digitizer utilizes acoustic ranging to digitally measure the position of a stylus on the surface of the tablet. The measurement is performed by measuring the time required for sound to travel from a spark sound source on the tip of the stylus to a pair of linear microphones located along two axes of the tablet. This time measurement is translated into a number between zero and 1999 for each axis, effectively dividing the tablet into a matrix of 2000 x 2000 points.

The four digit binary-coded decimal numbers generated by the Graf-Pen for each data point are converted to a sequence of characters suitable for transmission to a computer terminal or teletype unit by a data converter. This data converter, in addition to performing the parallel to serial data conversion, inserts control characters into the data string to make it compatable with a computer.

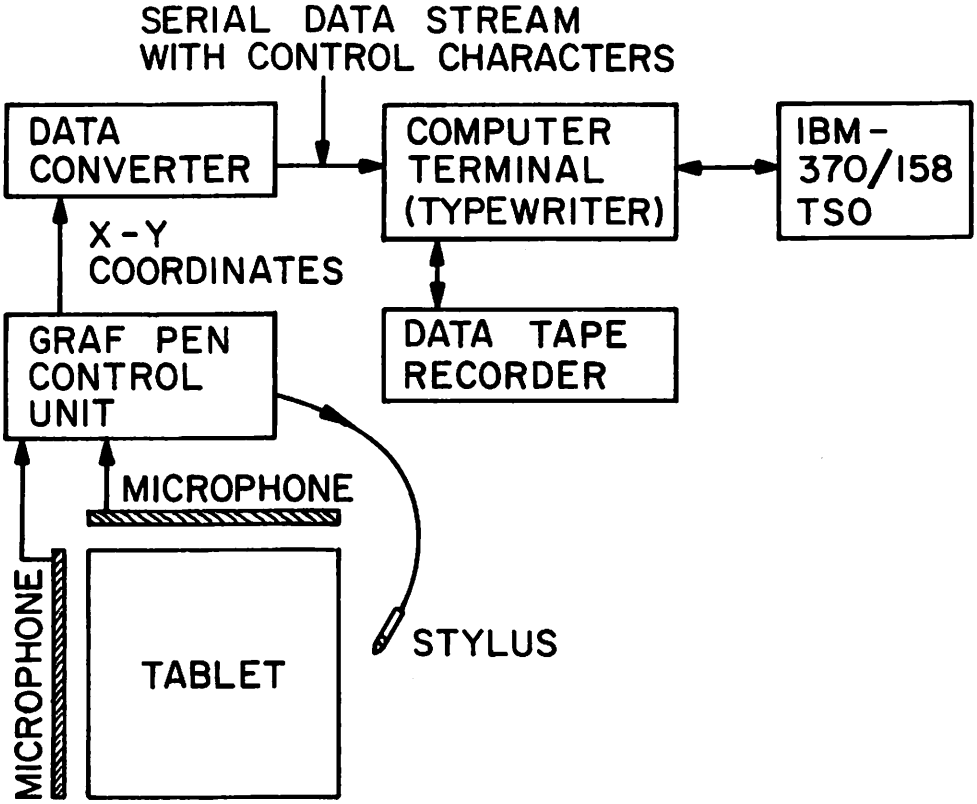


Figure 5.2-1. Data acquisition and processing system.

In the present installation, data generated by the Graf Pen is passed from the data converter to a Novar computer terminal which is equipped with a cartridge data tape recorder. This tape recorder is used as a data buffer to hold pen data until it is required by a computer program and to provide for maximally efficient transmission of data to the computer during program execution. The sporadic manner in which time-shared computers schedule program execution makes such a buffer a necessity. During the execution the operation of the tape recorder is controlled automatically by the computer making manual intervention unnecessary.

Data transfer rates for the graphic data terminal are determined by the rate limitations of the computer terminal and the computer's input equipment. Since the existing equipment is limited to no more than 30 characters per second, the data acquisition and transfer rate is limited to approximately two X-Y coordinate pairs per second. During digitization operations the digitizer can be set to acquire data only when the operator indicates a desired point by pressing on the data stylus.

In operation a stop motion movie projector image of the dye filled left ventricle is projected onto the graf pen surface where the operator digitizes 30 to 50 perimeter points obtained from 2 to 60 frames of cine-film per patient. This boundary information as well as digitized strip chart recording of a simultaneously recorded pressure versus time curve is stored on cassette tape for transmission.

At this point the IBM time share option is accessed with resident programs to accept, analyze the data and return a report to the operator. The programs initially interact with the operator in a question and answer mode followed by computation and return of 20 quantitative indices of ventricular function. In addition, a high resolution (30 characters per inch) plot routine returns a plot of the end diastole (full expansion and end systole (full contraction) boundary trace (see Figure 2) as well as volume vs time,

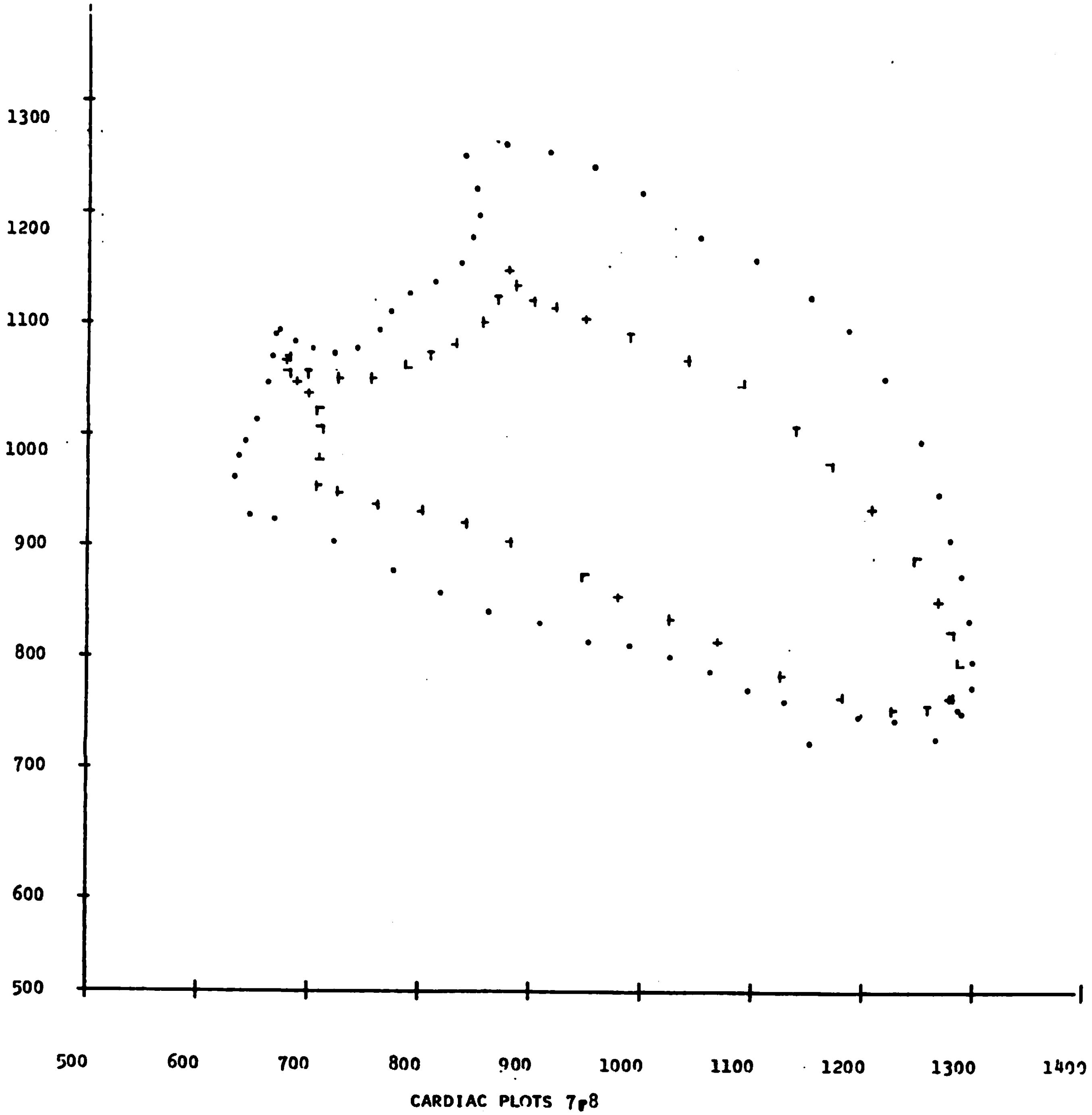


Figure 5.2-2. Ventricular outlines from high resolution, terminal based plotting routine.

pressure vs volume, and force vs length plots where applicable. This information then becomes part of the permanent patient record.

This interactive terminal based device has significantly reduced the time to complete such an analysis to the point of practical cost-effectiveness. The modest cost of such a system will allow many remote users to have access to a large general purpose computer where they can use as well as augment existing ventricular analysis programs for use by the user community. This concept has the potential for encouraging resource sharing within its limited scope.

5.3 Reconstruction of Three-Dimensional Object Arrays From Their Projections

Alexander A. Sawchuk and Ernest L. Hall

A problem of some interest in image processing concerns the reconstruction of arrays of data from projection measurements having a lower spatial dimensionality. Typically, the objective is to reconstruct three-dimensional object arrays from two-dimensional projections, or to reconstruct two-dimensional object arrays from one-dimensional projections. The technique has applications in a wide variety of areas, not all of which are traditionally connected with image processing. Some of these application areas include: radiography, or X-ray imaging [1,2]; nuclear scintillation imaging [3]; electron microscopy [4]; electromagnetic and pressure wave field mapping such as in radar, radioastronomy and sonar [5]; measurement of volume refractive index and temperature profiles [6]; and multidimensional picture representation and coding [7].

A large number of different techniques have been proposed for accomplishing image reconstruction. In its most general form, the M-dimensional projection vector $\underline{g}$ is obtained by a linear combination of the N-dimensional orginal object vector $\underline{f}$ as expressed by

$$\underline{g} = \underline{H}\,\underline{f} \qquad (1)$$

where $\underline{H}$ is an $M \times N$ matrix. All reconstruction algorithms attempt to produce an estimate $\hat{\underline{f}}$ of the true $\underline{f}$. There are many obvious similarities to the image restoration problem, and the special nature of reconstruction has led to several general approaches: Fourier transform [1,4], convolution [5,8], and algebraic [2]. No one approach has been found superior to the others, and extensive analytical comparison is difficult. Other problems, such as measurement noise in taking the data according to eq. (1), and imperfect knowledge of $\underline{H}$ also led to different tradeoffs between techniques. Noise is utlimately the most significant problem in image processing, and very few reconstruction methods really consider it adequately.

The Fourier transform method of reconstructions is one of the simplest techniques although it requires that projection data be taken in a full 180° around the object to be imaged. The method relies on the fact that the Fourier transform of a two-dimensional projection of a three-dimensional transform of the object. If $\underline{f}(x_1, x_2, x_3)$ represents the object, the three-dimensional Fourier transform is given by

$$\underline{F}(f_1, f_2, f_3) = \iiint_{-\infty}^{\infty} \underline{f}(x_1, x_2, x_3)\exp[-2\pi_j(f_1x_1 + f_2x_2 + f_3x_3)]dx_1dx_2dx_3 \quad (2)$$

and the central section of the transform is

$$\underline{F}(f_1, f_2, 0) = \iiint_{-\infty}^{\infty} \underline{f}(x_1, x_2, x_3)\exp[-2\pi_j(f_1x_1 + f_2x_2)]\,dx_1dx_2dx_3 \quad (3)$$

After performing the integration with respect to x_3 in eq. (3) and identifying

$$\underline{f}_3(x_1, x_2) = \int_{-\infty}^{\infty} \underline{f}(x_1, x_2, x_3)dx_3 \quad (4)$$

as the projection on the x_1-x_2 axes and $\underline{F}(f_1, f_2, 0)$ as its transform, the full three-dimensional transform can be built up plane by plane using the transforms of different projection views of the object. It can be shown that a coordinate system (s_1, s_2) rotated, as shown in figure 1a, by θ with respect to (x_1, x_2) moves in exactly the same manner as the rotated Fourier transform space (μ_1, μ_2). Thus, projections are taken at different θ orientations, transformed, and inserted into the two-dimensional transform plane shown in figure 1b. The entire transform of the object can be built up plane-by-plane by this method. Since only a finite number of projections are available, some interpolation in the transform space is required. Several methods including pulse approximation, linear interpolation, Fourier series interpolation, and sinc function expansion for band and space-limited objects are being considered, and tradeoffs and relative accuracies are under study. Related to this are problems of appropriate sampling rate and projection density for adequate reconstruction as a function of object parameters. Although the Fourier transform method needs 180° of data, it is relatively easy to implement and gives some of the best quality results with relatively low required computation when the projection data is not too noisy.

The second major reconstruction technique is known as the convolution method [5, 8] and operates entirely in the spatial domain without the use of any transforms. Denoting the desired object to be reconstructed in polar coordinates by $f(r, \theta)$, then a forward Fourier transform relationship to transform $F(R, \Phi)$ is

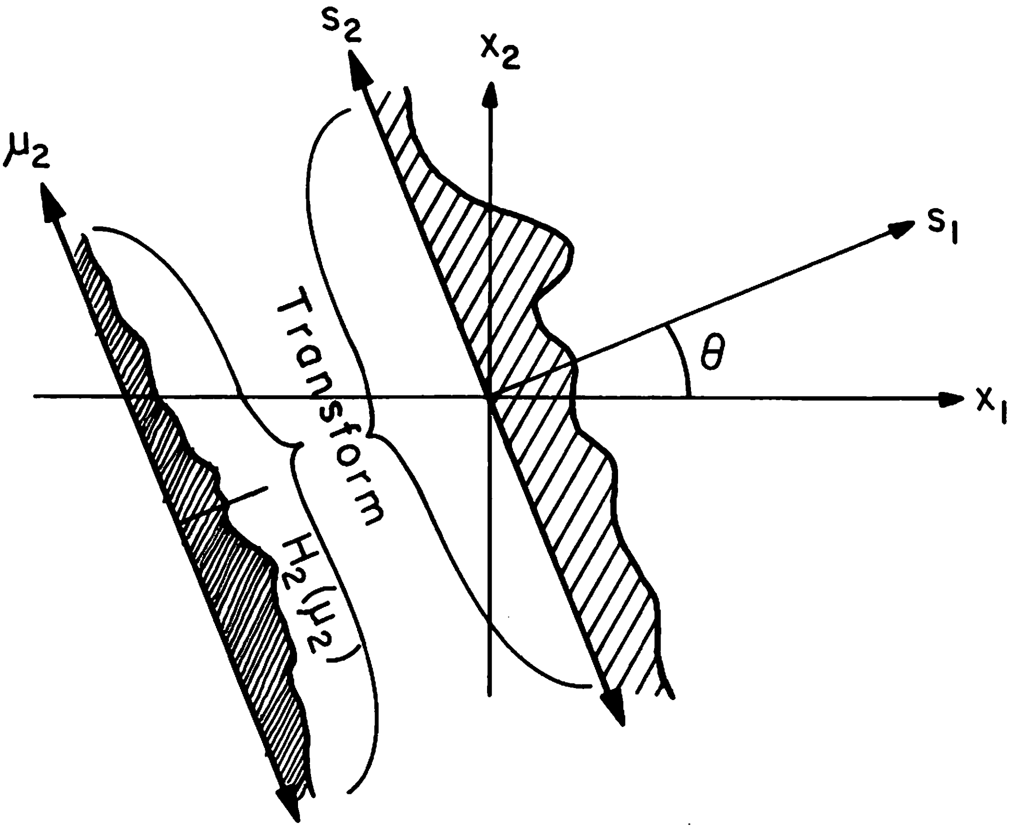
$$F(R, \Phi) = \int_0^{2\pi} \int_0^{\infty} f(r, \theta)\, e^{-j2\pi R \cos(\theta - \Phi)} r\,dr\,d\theta \tag{5}$$

with reverse transform

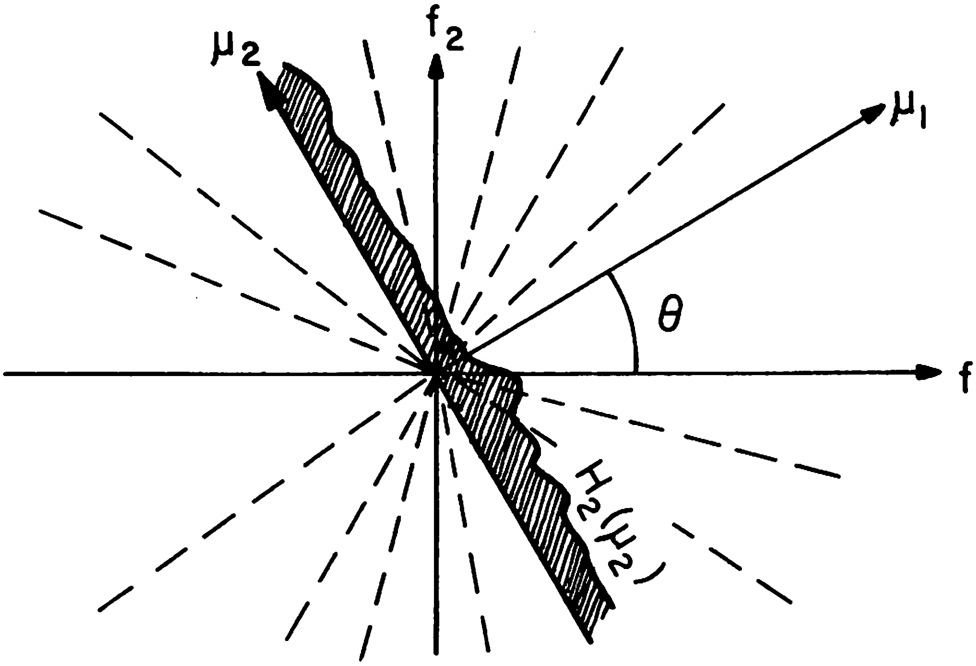
$$f(r, \theta) = \int_0^{2\pi} \int_0^{\infty} F(R, \Phi) e^{+j2\pi r R \cos(\theta - \Phi)} R\,dR\,d\Phi\,. \tag{6}$$

Equation (6) can also be expressed as

$$f(r, \theta) = \int_0^{\pi} \int_{-\frac{A}{2}}^{\frac{A}{2}} |R| F(R, \Phi) e^{+j2\pi r R \cos(\theta - \Phi)} dR\,d\Phi \tag{7}$$



a) Projection data – Fourier transform reconstruction



b) Assembly of Fourier transform

Figure 5.3-1. Geometry for Fourier transform technique.

where the finite limits $-\frac{A}{2}, \frac{A}{2}$ on R represent the bandlimited nature of $F(R, \Phi)$. Equation (7) is the basis for the technique.

The method begins by first measuring projections $g(\ell, \Phi)$ as a function of the variable ℓ at angle Φ as shown in figure 2a. This can be expressed as an inverse transform

$$g(\ell, \Phi) = \int_{-\frac{A}{2}}^{\frac{A}{2}} F(R, \Phi) e^{+j2\pi R\ell}\, dR \tag{8}$$

in a frequency space identical to that shown in figure 1b. The second step is to linearly filter $g(\ell, \Phi)$ with a space-invariant impulse response $h_\Phi(\ell)$ given by

$$h_\Phi(\ell) = \int_{-\frac{A}{2}}^{\frac{A}{2}} |R| e^{j2\pi R\ell}\, dR \tag{9}$$

to produce $g'(\ell, \Phi)$ according to the convolution relationship

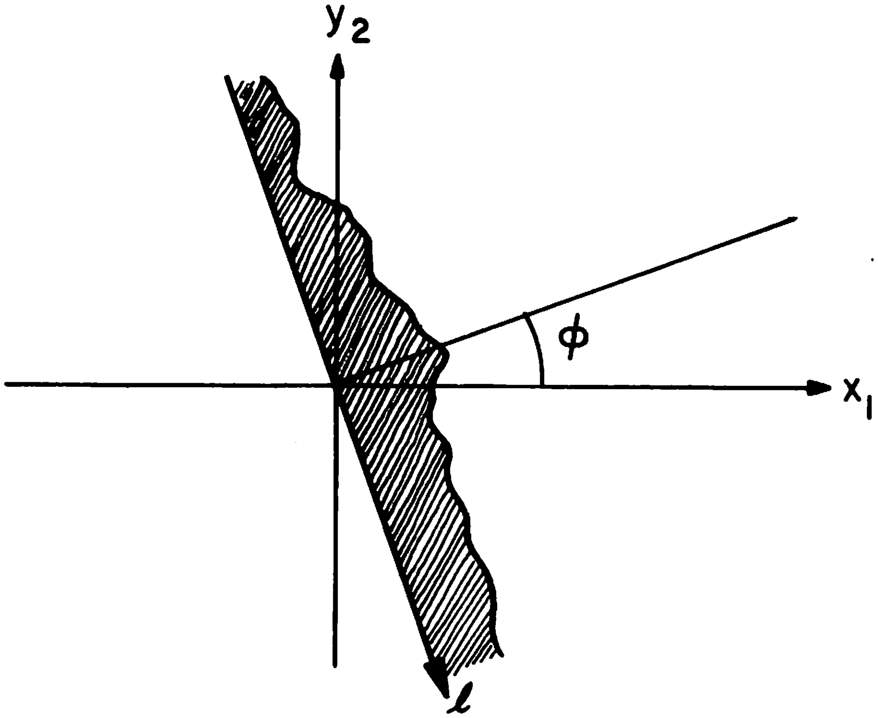
$$g'(\ell, \Phi) = g(\ell, \Phi) * h_\Phi(\ell) \tag{10}$$

This operation is known as "rho-filtering" due to the radial spatial frequency amplification of eq. (9) shown in figure 2b. By the convolution theory of transform theory, eq. (10) is the same as

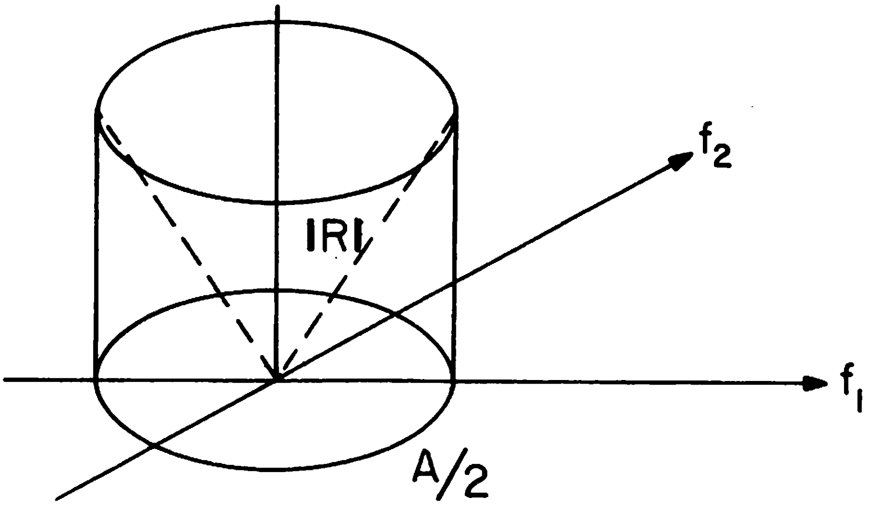
$$g'(\ell, \Phi) = \int_{-\frac{A}{2}}^{\frac{A}{2}} |R|\, F(R, \Phi) e^{j2\pi R\ell}\, dR\ . \tag{11}$$

The final step is to use $g'(\ell, \Phi)$ and compute

$$f(r, \theta) = \int_0^\pi g'(r \cos(\theta - \Phi), \Phi)\, d\Phi \tag{12}$$



a) Convolution projection data



b) Rho filter for convolution restoration

Figure 5.3-2. Geometry for convolution techniques.

to give the object density $f(r, \theta)$. Substituting eq. (11) into eq. (12) gives the identity of eq. (7) and proves the method. In a discrete sampled-data format the operations of eq. (10) and eq. (12) becomes

$$g'(na, \theta) = a \sum_{m=-\infty}^{\infty} g(ma, \theta)h((m-n)a) \qquad (13)$$

and

$$f(r, \Phi) = f(jr_0, k\Phi_0) = \sum_{t=1}^{N} g'[jr_0 \cos(k\Phi_0 - t\theta_0), t\theta_0] \qquad (14)$$

respectively. Other alternative techniques may involve reversing the order of the operations of eqs. (10) and (12), or combining them with part of the data taking operation of eq. (8). As with the transform method, the discrete nature of the computer implementation requires interpolation, sampling, and numerical integration; work is underway to study the best method of performing these tasks. Some study of the effects of noise is also being considered for this system, and improvements to the inverse filtering operation in the rho filter are possible.

The third reconstruction technique may be called the algebraic approach. The problem is illustrated in figure 3 and may be formulated in the following manner. Define a linear set of projection equations

$$\underline{H}\,\underline{f} = \underline{g}$$

where

$$\underline{f} = (f_1, f_2, \ldots, f_N)^T \qquad N = n \times n$$

$$\underline{g} = (g_1, g_2, \ldots, g_N, g_{N+2}, \ldots, g_{2N}, \ldots, g_{2N+2}, g_{2N}, \ldots g_K)^T$$

with $K = M - (m_a - 1)$ linearly independent projection equations, and $\underline{H}$ is a $K \times N$ matrix depending only upon the geometry. These equations are exactly similar in form to the restoration equation and only differ by construction. Some of the possible solution constraints include:

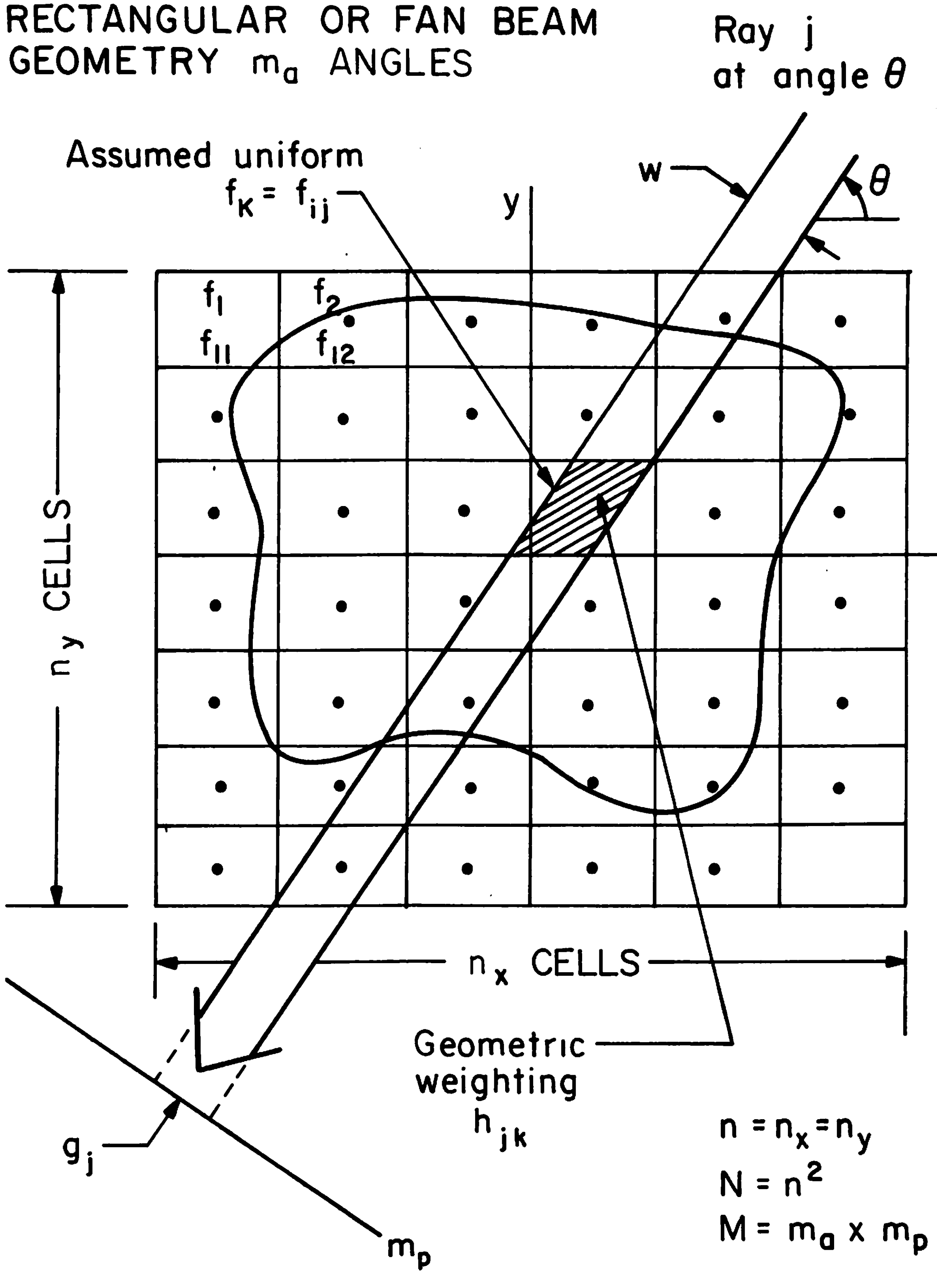


Figure 5.3-3. Algebraic reconstruction geometry.

a) Minimum Energy

$$J_1(\underline{f}) = \underline{f}^T\underline{f}$$

b) Smoothness

$$J_z(\underline{f}) = \underline{f}^T C\underline{f}$$

c) Minimum mean square error. Given a measurement with noise

$$\underline{z} = \underline{g} + \underline{\zeta}$$

$$J_3(\underline{f}) = \|\underline{z}-\underline{H}\underline{f}\|$$

Solutions may easily be developed for these minimization problems. For example, one may minimize

$$J(\underline{f}) = \underline{f}^T\underline{f} + \alpha\underline{f}^T\underline{C}\,\underline{f}$$

subject to the constraint

$$\underline{H}\,\underline{f} = \underline{g}$$

which gives

$$\hat{\underline{f}} = -2[\underline{H}(\underline{I} + \alpha\underline{C})^{-1}\underline{H}^T]^{-1}\,\underline{g}$$

Similarly, one may minimize

$$J(\underline{f}) = \|\underline{z} - \underline{H}\,\underline{f}\|^2 + \alpha\underline{f}^T C\underline{f} + \underline{f}^T\underline{f}$$

subject to the noise measurement constraint which gives

$$\hat{\underline{i}} = [\underline{I} + \underline{H}^T\underline{H} + \alpha\underline{C}]^{-1}\,\underline{H}^T\underline{Z}$$

The inverses in these equations may be shown to exist by construction of $\underline{F}$ and $\underline{C}$.

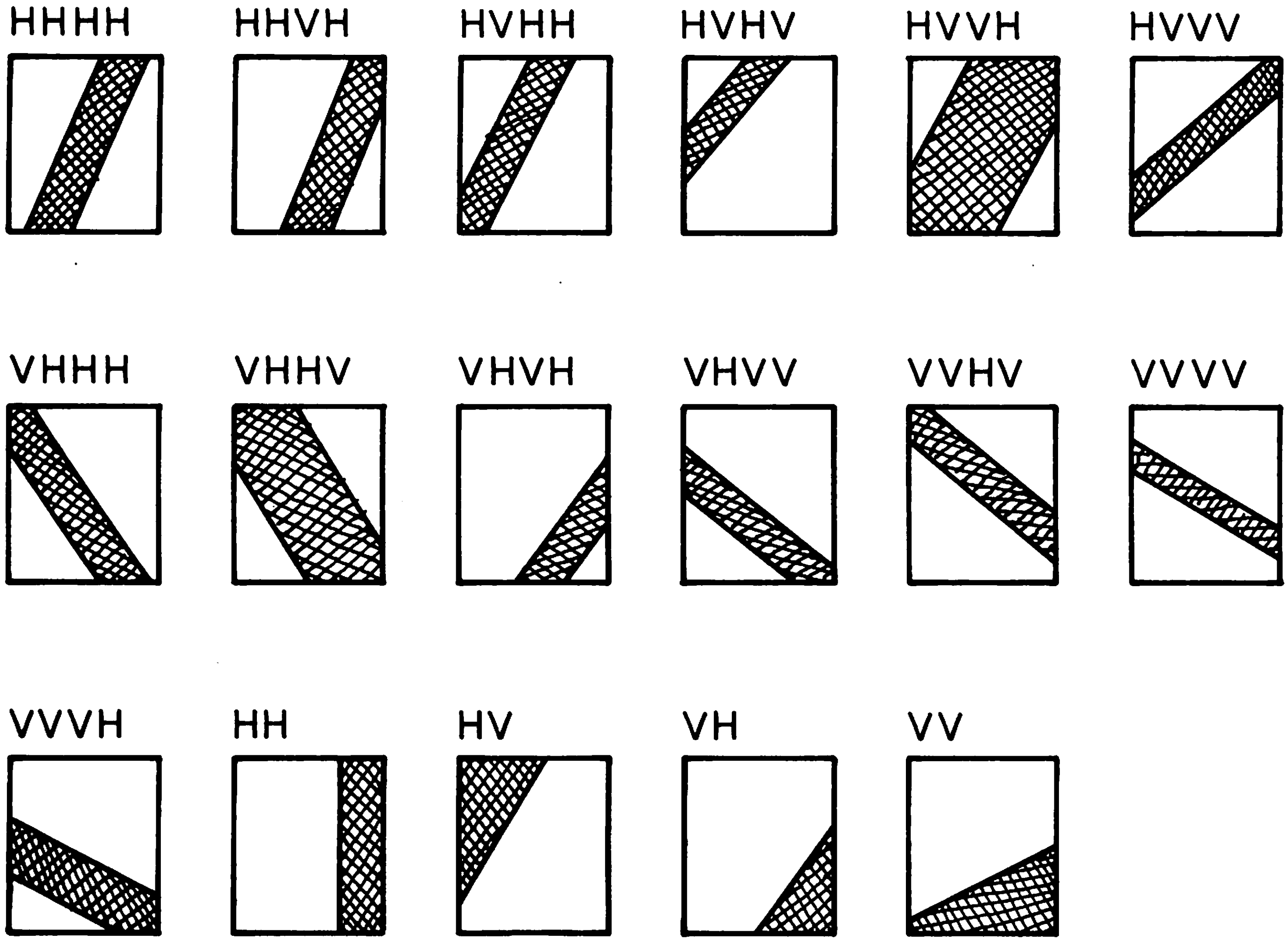
Thus, the algebraic approach is characterized by the solution of a large set of linear equations. For example in the EMI scanner [9] the unknown image is reconstructed on an 80 X 80 grid which gives a 6400 equivalent vector dimension. Also $m_a = 180$ projection angles are used with $m_p = 160$ points per projection angle. Thus, K is approximately 28,000. The previous solutions for $\underline{\hat{\mu}}$ if directly solved would require inverting 6400 X 6400 matrices.

The large amount of computations involved motivate two avenues of continued study. The first involves the use of iterative solutions of the large set of equations. The second involves studying the structure of the H matrix so that efficient computational algorithms may be developed.

Certain structures in the H matrix are obvious. For example, suppose the coefficient values are selected proportional to the intersection area of a given ray and the elemental unknown area as shown in figure 4. Then assuming either one or two rays intersect each elemental area there are only 18 possible intersection areas as shown in figure 1a. Thus, although the intersection areas changes for each projection angle, the formulas for the computation are identical.

References

1. R. H. T. Bates and T. M. Peters, "Towards Improvements in Tomography," New Zealand J. Science, vol. 14, pp. 883-896, 1971.
2. G. T. Herman and S. Rowland, "Three Methods for Reconstructing Objects from X-rays: A Comparative Study," Proc. San Diego Biomedical Symposium, pp. 285-297, 1972.
3. M. Goitein, "Three-Dimensional Density Reconstruction From a Series of Two-Dimensional Projections," Nucl. Inst. and Meth., vol. 101, pp. 509-518, 1972.
4. R. A. Crowther, D.J. DeRosier, and A. Klug, "The Reconstruction of a Three-Dimensional Structure from Projections and its Application to Electron Microscopy," Proc. Roy. Soc. London A., vol. 317, pp. 319-340, 1970.
5. R. N. Brucewell and A. C. Riddle, "The Inversion of Fan-Beam Scans in Radioastronomy," Astrophysical J., Vol. 150, pp. 427-434, 1967.
6. D. W. Sweeney and C. M. Vest, "Reconstruction of Three-Dimensional



NOTE: Cases HHHV, HHVV, VVHH cannot occur with parallel rays.

Figure 5.3-4. Intersection areas for parallel rays and elemental areas.

Refractive Index Fields from Multidirectional Interferometric Data," Appl. Opt., Vol. 12, pp. 2649-2664, 1973.

7. R. L. Kashyap and M. C. Mittal, "Picture Reconstruction from Projections," First International Conference on Pattern Recognition, Washington, D. C., pp. 286-292, 1973.

8. G. N. Ramachandran and A. V. Lakshminarayanan, "Three-Dimensional Reconstruction from Radiographs and Electron Micrographs: Application of Convolutions Instead of Fourier Transforms," Proc. Nat. Acad. Sci., Vol. 68, pp. 2236-2249, 1971.

9. EMI, Ltd, "Method of and Apparatus for Examining a Body by Radiation Such as X- or Gamma Radiation," Canadian Patent No. 887891, Dec., 1971.

5.4 Application of Kalman Filtering to the Reconstruction of Images from Noisy Projections

Firouz Naderi, Ali Habibi

A digital monochromatic image can be modeled by a two dimensional array $x(i,j)$ $i,j = 1,\ldots,N$ where $x(i,j)$ denotes the magnitude of the gray level at point (i,j). Then a one dimensional projection of the $N \times N$ image in direction θ is defined as an N vector $\underline{Y}_\theta$ where each component of $\underline{Y}_\theta$ is the sum of the elements of the image along a line making an angle θ with a fixed reference axis. Figure 1 shows examples of two such projections for $\theta = 90$ and 45 degrees. Note that for the $\theta = 90^o$ projection

$$\underline{Y}_{90^o} = [y^1_{90^o}, y^2_{90^o}, \ldots, y^N_{90^o}]^T \tag{1}$$

where

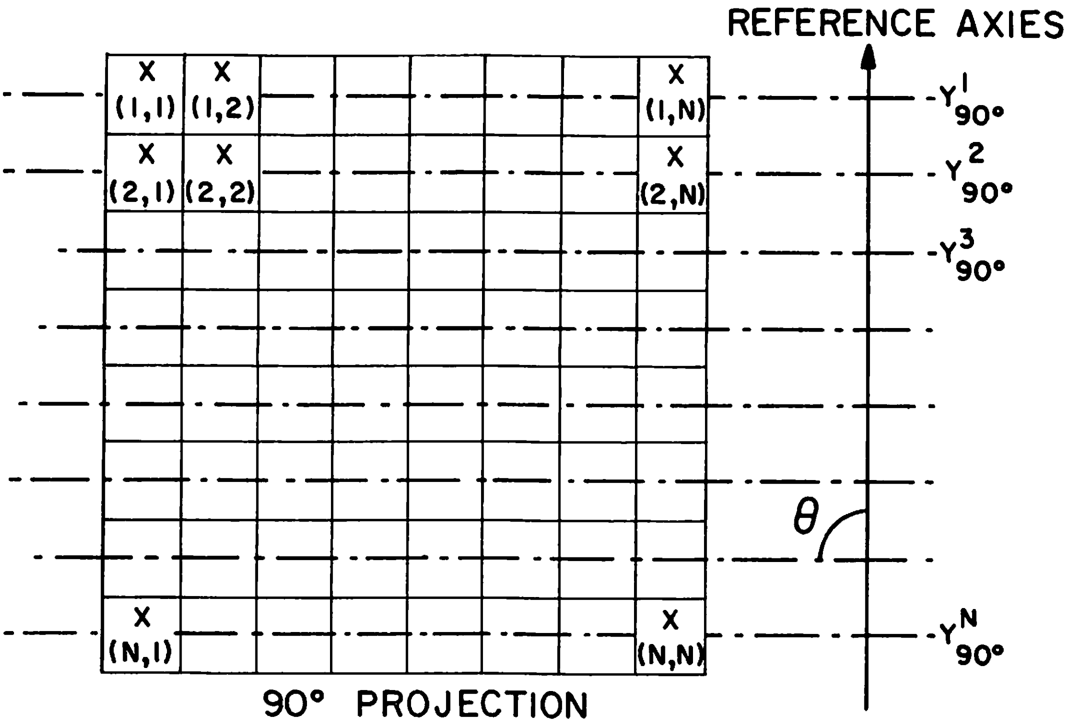
$$y^i_{90^o} = \sum_{j=1}^{N} x(i,j) \tag{2}$$

Now, let $\underline{x}$ be an N^2 vector defined as

$$\underline{x} = [x(1,1), x(1,2), \ldots, x(1,N), x(2,1), \ldots x(N,N)]^T \tag{3}$$

A one dimensional projection of an image can now be written as

$$\underline{Y}_\theta = \underline{C}_\theta \underline{x} \tag{4}$$



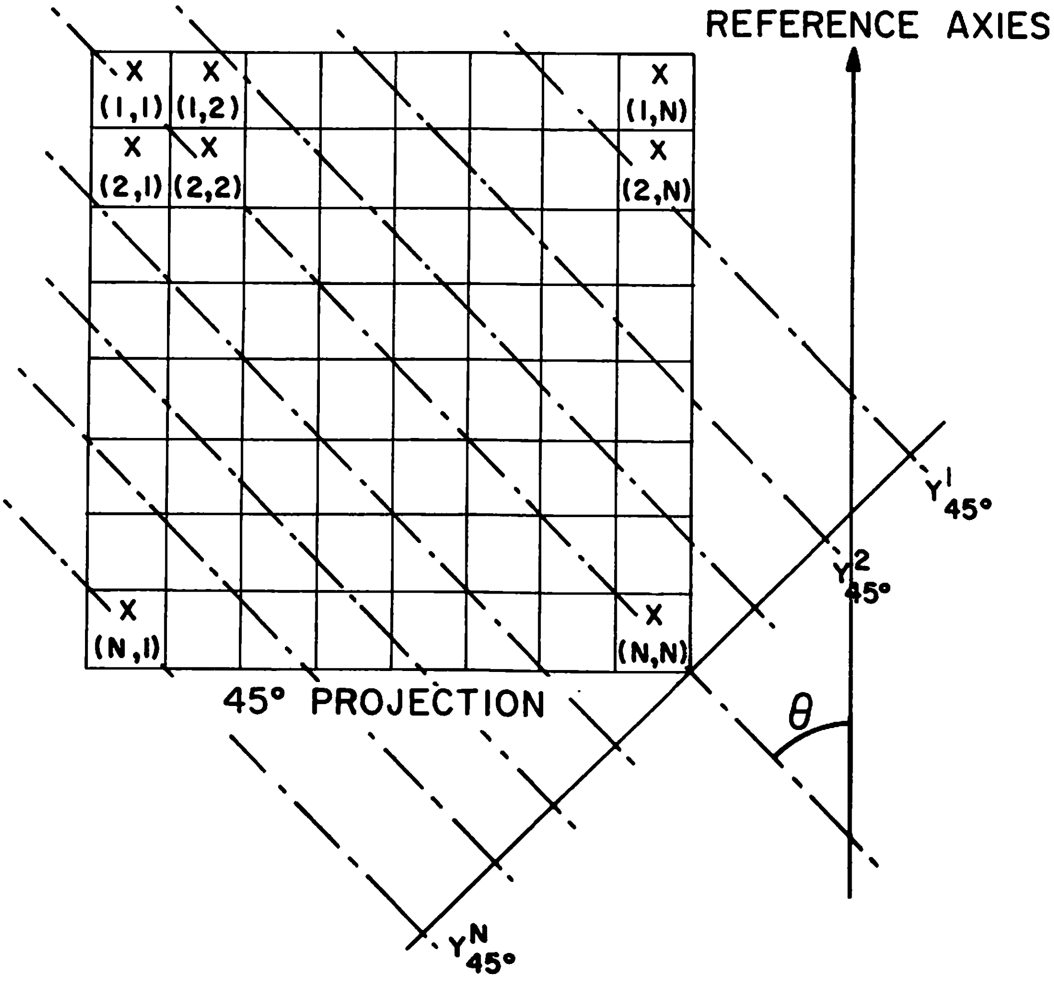


Figure 5.4-1. Examples of one dimensional projections of an image.

Where the N by N^2 matrix $\underline{C}_0$ is called the projection matrix at angle θ. For the example of figure 1 when $\theta = 90^\circ$ matrix $\underline{C}_\theta$ is

$$\underline{C}_{90^\circ} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 & \cdots & & & & & & & 0 \\ 0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 & & \cdots & & 0 \\ 0 & 0 & \cdots & 0 & & \cdots & & 0 & 1 & 1 & \cdots & 1 & 0 & & \cdots & 0 \\ \vdots & & & \vdots & & & & \vdots & & & & & & & & \vdots \\ & & & & & & & & & & & & & & & \\ 0 & 0 & \cdots & & & & & & & & & 0 & 1 & 1 & \cdots & 1 \end{bmatrix}$$

Given a finite number of projections $\underline{Y}_{\theta_1}, \ldots, \underline{Y}_{\theta_L}$ and the corresponding projection matrices $\underline{C}_{\theta_1}, \ldots, \underline{C}_{\theta_L}$ it is desired to reconstruct the image, i.e. to determine the N^2 elements $x(i, j)$, $i, j = 1, \ldots, N$. This problem can be viewed as an attempt to solve a set of $N \times L$ algebraic equations with N^2 unknowns. In most cases the number of linearly independent equations available is not equal to N^2; therefore obtaining a unique solution requires forming a meaningful criterion and solving the equations based on this additional constraint.

Another approach discussed below is to model the image as a two dimensional random process, and use some a priori statistical knowledge of the process to obtain an estimate of the image.

Reconstruction of Images Using Kalman Filtering The approach here is to derive a linear dynamic model to represent the imaging system, and then use linear estimation theory to obtain an estimate of the image which is optimal using mean square error criterion. Let Y(1), . . . , Y(L) be L one dimensional projections of the image at different angles. As a result of faulty measurement, $Y(1), \ldots, Y(L)$ are not the true projections, but contain some noise. Let this noise be an additive white noise denoted by the N element vector $\underline{n}(k)$ where $\underline{n}(k)$ is the noise corresponding to projection $\underline{Y}(k)$. Then

$$\underline{Y}(k) = \underline{C}(k)\underline{x} + \underline{n}(k) \tag{6}$$

where $E[\underline{n}(k)^T \underline{n}(j)] = L\delta(k-j)$.

Clearly since all the projections correspond to the same fixed image, the dynamic equations representing the system are

$$\underline{x}(k+1) = \underline{x}(k) \tag{7a}$$

$$\underline{Y}(k) = \underline{C}(k)\underline{x}(k) + \underline{n}(k) \tag{7b}$$

Given projections $\underline{Y}(1), \ldots, \underline{Y}(L)$ it is desired to find a linear estimate of $\underline{x}$ such that the mean square error is minimum.

The recursive solution to the above is the well known Kalman estimator. Denoting by $\hat{\underline{x}}(k+1)$ the estimate of the image obtained after observing the projections $\underline{Y}(1), \ldots, \underline{Y}(k)$ one obtains

$$\hat{\underline{x}}(k+1) = [\underline{I} - \underline{F}(k)\underline{C}(k)]\hat{\underline{x}}(k) + \underline{F}(k)\underline{Y}(k) \tag{8}$$

where

$$\underline{F}(k) = \underline{P}(k)^T\underline{C}(k)[\underline{C}(k)\underline{P}(k)^T\underline{C}(k) + L\underline{I}]^{-1} \tag{9}$$

and

$$\underline{P}(k) = E\{[\underline{x}(k) - \hat{\underline{x}}(k)][\underline{x}(k) - \hat{\underline{x}}(k)]^T\} \tag{10}$$

The term $\underline{P}(k)$ is the covariance of the error after k iterations and has the recursive form

$$\underline{P}(k+1) = [\underline{I} - \underline{F}(k)\underline{C}(k)]\underline{P}(k) \tag{11}$$

Equations (8), (9) and (11) provide a recursive method of finding the estimate of the image. However along with this equation initial conditions for $\hat{\underline{x}}(k)$ and for $\underline{P}(k)$ are needed.

A reasonable value for the initial conditions are [2]

$$\hat{\underline{x}}(0) = E(\underline{x})$$

$$\underline{P}(0) = E[\underline{x} - E(\underline{x})][\underline{x} - E(\underline{x})]^T$$

Some aspects of this approach are intriguing enough to warrant more investigation of this method. For example:

1) The error covariance equation (eq. 11) is asymptotically stable. Given that projection matrices, $\underline{C}(k)$, are known for various angles, the error covariance can be calculated for all k before any projection data is actually available. This suggests that it might be possible to decide how many projections are necessary to achieve a certain fidelity.

2) The error covariance depends on the projection matrices which in turn depend on the angle of projection. It might be possible then to decide the angles of projections more judiciously by proper evaluation of the error covariance.

At the present the biggest disadvantage of the method is the excessive computations which are required because of using large matrices. However this method has distinct advantages over other methods that are used for reconstructing images from noisy projections. Namely, it is iterative and provides a measure for the error and is based on an established estimation theory procedure.

References

1. R. L. Kashyap, "Picture Reconstruction from Projections," Proceedings of the First International Joint Conference on Pattern Recognition, Oct. 1973, Washington, D. C., pp. 286-292.

2. N. E. Nahi, Estimation Theory and Application, John Wiley and Sons, 1969.

5.5 Boundary Estimation of Statistical Objects in Noise

Nasser E. Nahi and Mohammad Jahanshahi

The problem considered is determination of the boundary of an object in noise. Standard boundary detection techniques, in general, utilize some form of differencing or differentiation, [1]. These methods, however, are sensitive to noise. Various extensions of these techniques appear in [2], [3]. They are mainly effective for high signal to noise ratio images.

The images considered in this work can be partitioned into two regions: background and foreground. The foreground is assumed to form a horizontally convex set in the plane.

Definition: A set $E \subset R^2$ is said to be horizontally convex if given $\underline{x} = (x_1, x_2) \epsilon E$, $\underline{y} = (y_1, y_2) \epsilon E$ with $x_1 \neq y_1$, and $x_2 = y_2$, then $w\underline{x} + (1-w)\underline{y} \epsilon E$, where $0 < w < 1$.

Example: Sets E_1 and E_2, in figure 1, are horizontally convex. Set E_3 is not.

An object of interest whose intensity levels dominate those of the background will always be assumed to exist in the image. This object forms the foreground.

Problem Statement Assuming

1. An image consisting of a background and a horizontally convex object is given;
2. The image is represented by $b(z, n)\Gamma(z, n)$, where $b(z, n)$ is a random process representing grey levels within the object, and

$$\Gamma(z, n) = \begin{cases} 1 & \text{when } (z, n) \text{ is a point in the object,} \\ 0 & \text{when } (z, n) \text{ is a point in the background;} \end{cases}$$

3. The process $b(z, n)$ is characterized in a statistical sense, in terms of its first and second order moments;
4. The operator $\Gamma(z, n)$ defined the object boundary;
5. The observables

$$y(z, n) = b(z, n)\Gamma(z, n) + v(z, n) ,$$

where $v(z, n)$ represents the observation noise with known first and second order statistics, are given.

The goal is to find an estimate of the object boundary.

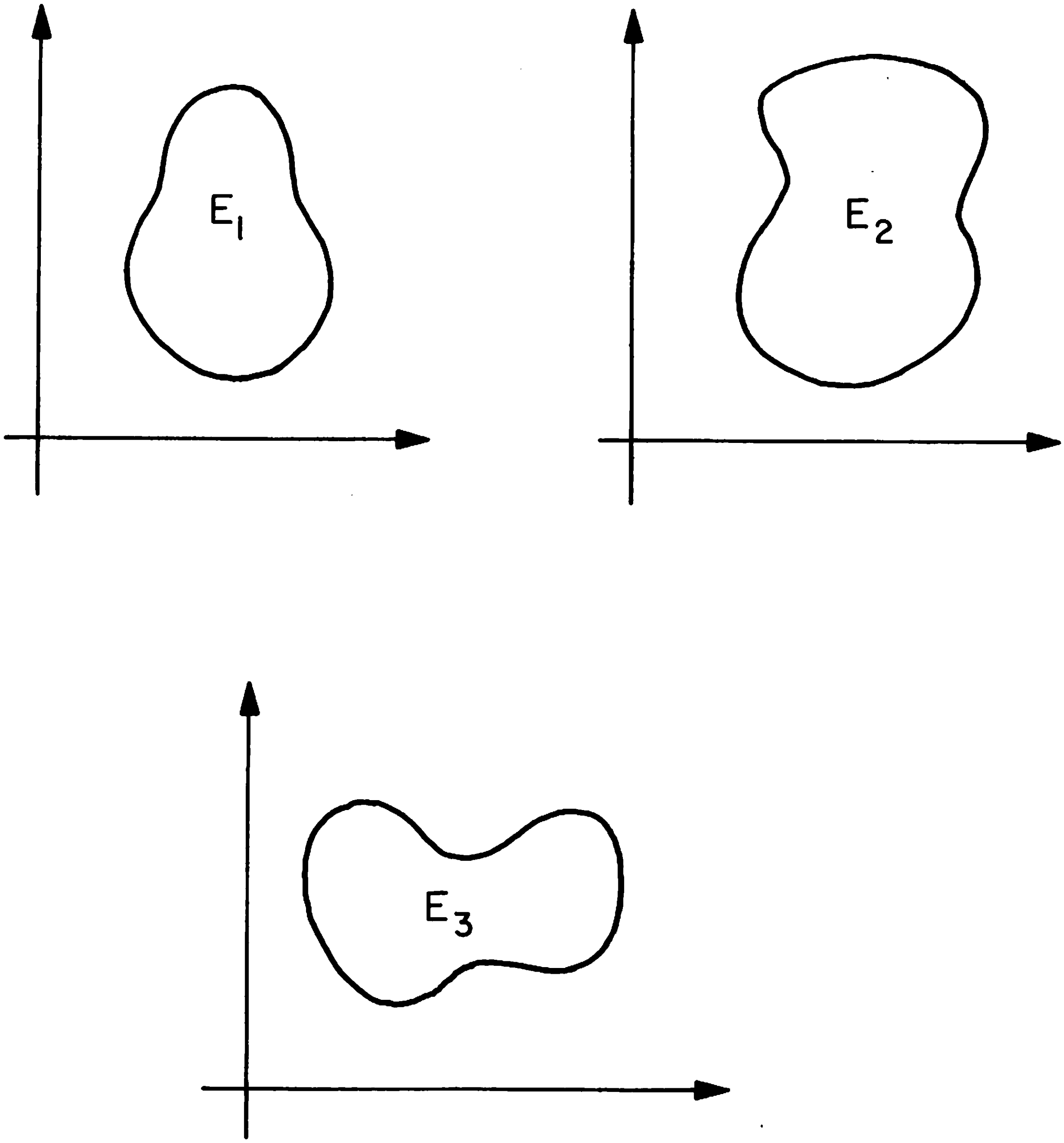


Figure 5.5-1. Examples of convex sets.
Sets E_1 and E_2 are horizontally convex.
Set E_3 is not.

A Solution Method The image, as stated, is characterized by a two dimensional random process. Since the solution of the proposed method will utilize conventional digital computers with litter or no parallel-computation capabilities, an image scanner is employed. Therefore, the problem is restated as the following: Assuming

1. The observation

$$y(t) = s(t)\lambda(t) + v(t) \qquad \text{for } 0 \leq t \leq T'$$

is given, where s(t) denotes the output of a line scanner, representing the grey level of the image at time t, and where

$$\lambda(t) = \sum_{\ell=m_1}^{m_2} u(t-a_\ell) - u(t-c_\ell)$$

with

m_1 = first line of the image containing the object;

m_2 = last line of the image containing the object;

a_ℓ = start of the object in line ℓ;

c_ℓ = end of the object in line ℓ;

$u(\cdot)$ = step function;

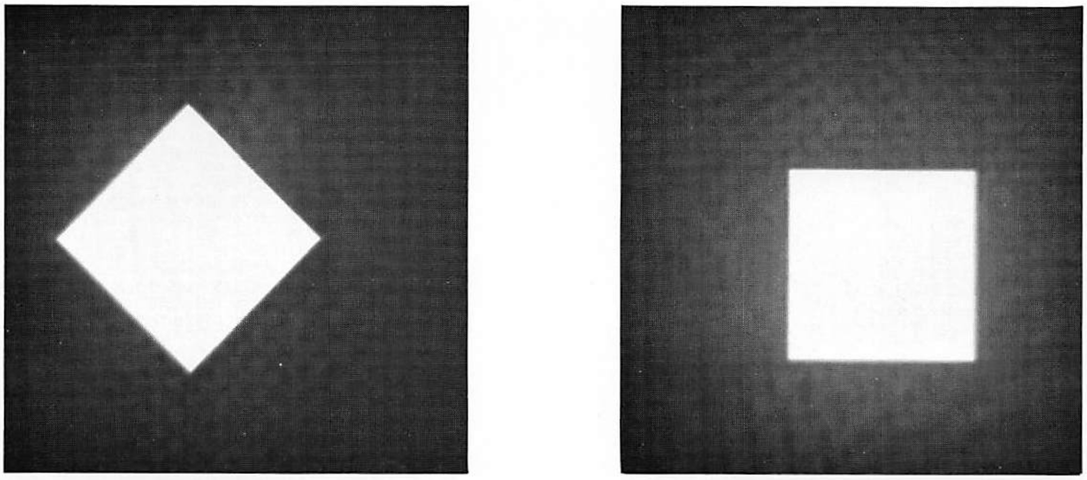
and

$v(t)$ = process representing the inaccuracies introduced by the scanner or any other phenomenon distorting the image.

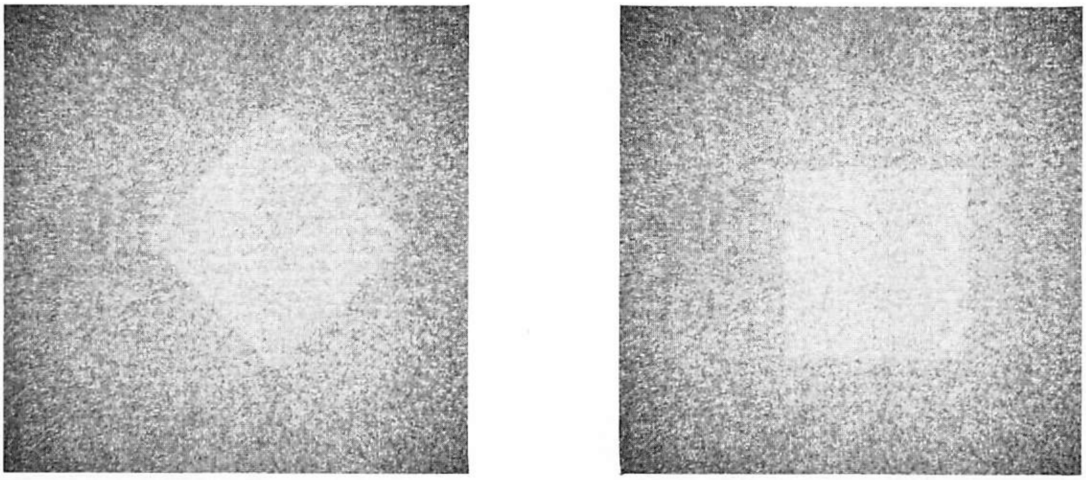
2. The sequence $x = (a_\ell, c_\ell)$, $\ell = m_1, m_1+1, \ldots, m_2$, is Markov.
3. The density function $f(x_\ell | x_{\ell-1})$ is Gaussian.

Find the M. A. P. estimates of a_ℓ and c_ℓ, $m_1 \leq \ell \leq m_2$, denoted by $\hat{a}_\ell$ and $\hat{c}_\ell$, respectively, and determine the values of m_1 and m_2.

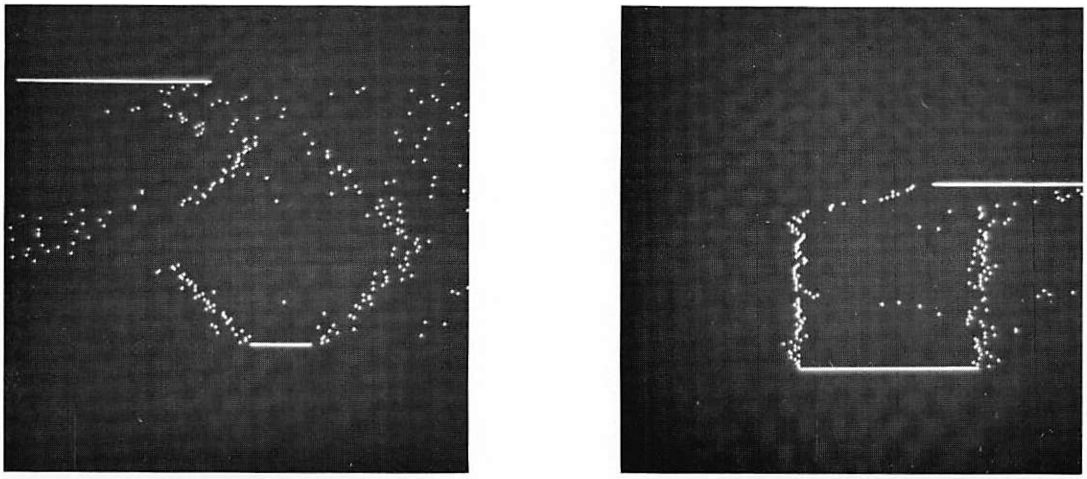
A recursive solution to the above problem has been obtained. Figure 2 contains two examples showing the feasibility of the boundary estimation technique. Further investigations are in progress as to the refinement of some of the theoretical aspects of the problem.



(a) original diamond (b) original square



(c) noisy diamond (d) noisy square



(e) boundary estimate (f) boundary estimate

Figure 5.5-2. Boundary estimation examples.

References

1. A. Rosenfeld, Picture Processing by Computer, New York, Academic Press, 1969.

2. A. Rosenfeld, "A Nonlinear Edge Detection Technique," Proceedings of the IEEE, Vol. 58, No. 5, May 1970.

3. E. Argyle, "Techniques for Edge Detection," Proceedings of the IEEE, Vol. 59, No. 2, Feb. 1971.

4. N. E. Nahi, Estimation Theory and Applications, New York, Wiley, 1969.

5. N. E. Nahi, T. Assefi, "Bayesian Recursive Image Estimation," IEEE Transactions on Computers, Vol. C-21, No. 7, July 1972.

6. Image Analysis Projects

The image analysis projects are concerned with the background technology necessary to effectively design image coding, restortion, 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.

The following project traces the development of a model of human color vision. This model has been found to accurately predict known visual phenomena such as spatial color contrast and color blindness.

6.1 A New Model of Color Vision and Some Practical Implications

Werner Frei

Color image coding and processing techniques usually operate on three scalar functions of some space variables such as red, green and blue color separations. Under the laws of colorimetry, these scalars called tristimulus values represent light energy quantities and the operations defined on them by color matching experiments conveniently satisfy the requirements of linear mappings: for example, the additive mixture of two arbitrary colors is mapped into the sum of their respective tristimulus values.

In terms of the human observer, however, colorimetric color representations have serious disadvantages, because the tristimuli are not representative of the physiological response evoqued in the human visual system. It has been shown that the visual response is a highly nonlinear function of the energetic tristimulus quantities. Tristimulus spaces are therefore not euclidean, and unfortunately, ordinary superposition of tristimuli does not entail superposition of responses. For example, if the tristimulus values of a color are doubled, one does not obtain a color that appears twice as bright, twice as saturated, etc. This remark applies of course to spatial arrangements of colored stimuli so that visual modulation transfer functions are, strictly speaking, not defined in tristimulus space.

It is therefore no surprise that distortion measures such as mean square error, or visual modulation transfer functions, etc. are of little relevance to the observer of a picture processed under that kind of criterion.

In view of these difficulties, the simple visual model developed here has a variety of practical implications in image processing and possibly in vision research. It is shown that its structure enables one to define an algebraic system in tristimulus space with a generalized superposition consistent with known perceptual phenomena.

As a result, one can define error measures in a euclidian domain that is a linear vector mapping of tristimulus space, given a particular set of basis vectors. This gives analytical convenience in the design and evaluation of coding and processing techniques.

On the other hand, with superposition defined, visual MTF can be defined and it can be shown that spatial color contrast phenomena such as simultaneous color contrast, color shadows, color mach bands, etc., can be modeled by linear spatial filters.

Superposition is also expected to enhance the effectiveness of pseudo-color techniques. Since the visual system is able to recognize more or less independently three attributes of colors (such as brightness, hue and saturation), pseudo-color techniques attempt to display multiple functions (at most three) of the space variables as a single composite color picture. If the observer is expected to be able to recognize the individual functions in the pseudo-color image, it is quite evident that the functions of interest must be mapped into color space in accordance with subjective superposition.

The visual model discussed here provides a tool to design precisely that kind of mapping, and, in addition, it predicts the spatial resolution to be expected for each function mapped into color space.

Visual Model Figure 1 shows a block diagram of the model. The basic ideas and the physiological evidence justifying its structure have been discussed in a previous report [1]. In brief, the first stage converts the input spectral energy distirbution of light $C(\lambda)$ (a function of the space variables x and y) into a tristimulus vector $t = t(x, y)$ whose three components represent the respective amounts of light energy absorbed by three types of photo-receptors (Young-Helmholm theory):

$$t = [t_1, t_2, t_3]^T \qquad t_i = t_i(x, y)$$

$$t_i = \int_{\lambda_L}^{\lambda_U} C(\lambda) t_i(\lambda) d\lambda \qquad i = 1, 2, 3$$

where $t_i(\lambda)$ is the spectral absorbtion function of the i-th type of receptor. The second stage of the model represents the conversion of absorbed energies to neural signals, according to an approximate logarithmic relationship (Weber-Fechner type of response). The components of the resulting vector $\underline{t}^*$ are: $\log t_1$, $\log t_2$, $\log t_3$ (the superscript star denotes the logarithmic domain). The last stage of the model contains a matrix of spatial filters with transfer function

$$\underline{H}^*(ju^*, jv^*) = \begin{bmatrix} H_1^* & 0 & 0 \\ -H_2^* & H_2^* & 0 \\ -H_3^* & 0 & H_3^* \end{bmatrix} \qquad H_{ij}^* = H_{ij}^*(ju^*, jv^*)$$

where u^*, v^* are spatial frequencies in the logarithmic domain. This stage simultaneously represents two phenomena:

(a) linear differences (inhibitions) between the outputs of the two types of receptors (demonstrated in the retinas of monkeys by deValois [2]). These differences are represented by equal terms of opposite sign within a row of $\underline{H}^*$. They can be seen to generate two "perceptual" chromatic signals by

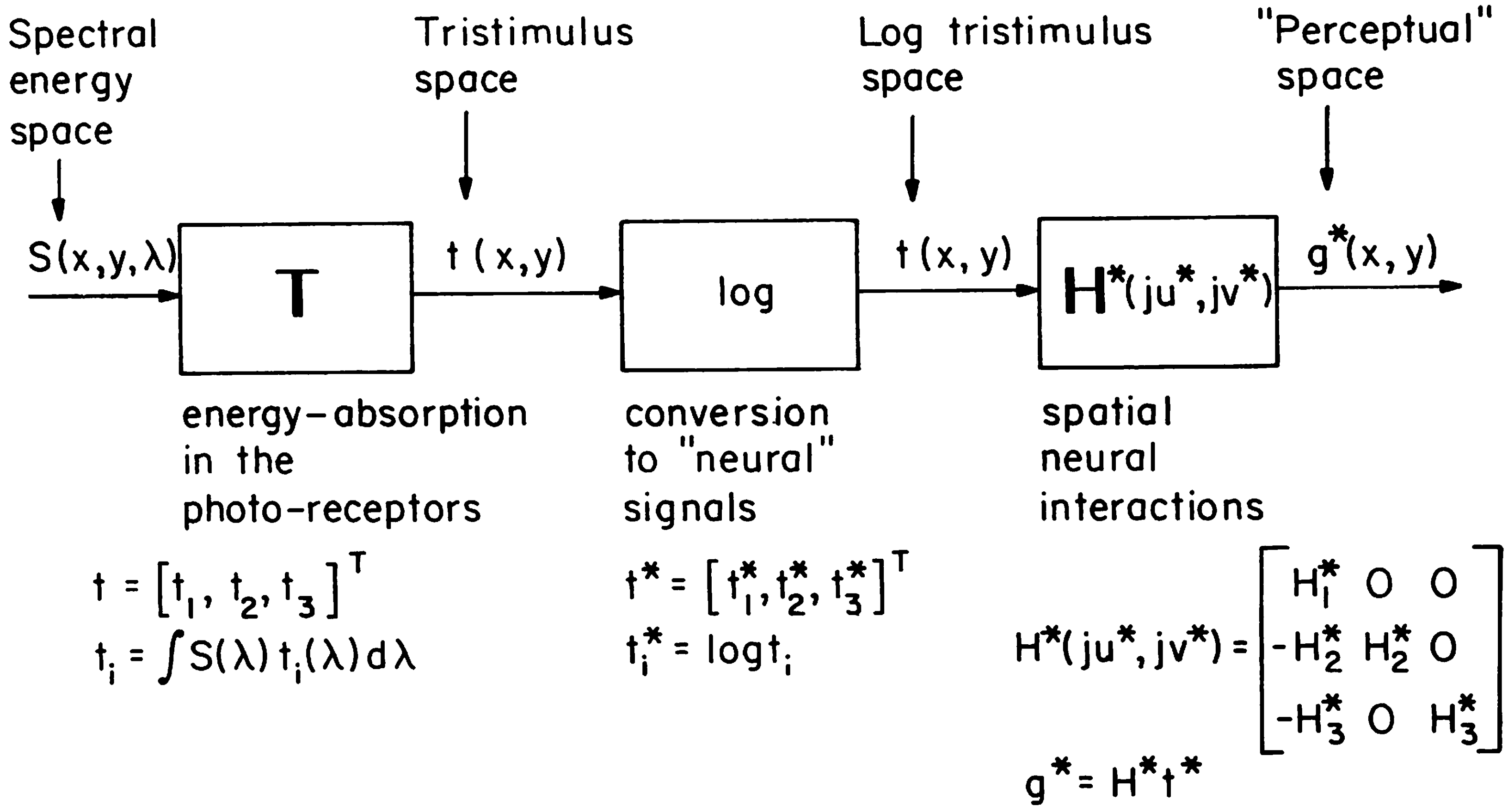


Figure 6.1-1. Model of color vision.

observing that

$$g_2^* = H_2^*(t_2^* - t_1^*) = \log\left(\frac{t_2}{t_1}\right)^{H_2^*}$$

$$g_3^* = H_3^*(t_3^* - t_1^*) = \log\left(\frac{t_3}{t_1}\right)^{H_3^*}$$

Under the assumption of exact logarithmic receptor responses, g_2^* and g_3^* are functions of the ratios of pairs of tristimulus values, e.g. chromatic quantities independent of the absolute light intensity. (This approximation is not valid at very low or very high intensities, but quite acceptable within most of the range of interest. The intensity is represented by

$$g_1^* = \log(t_1)^{H_1^*}$$

(b) The matrix H* also represents weighted linear summations between the outputs of spatially distributed like receptors. Assuming shift-invariance, suppose that each summation output $g^*(x_i, y_j)$ is related to the outputs of the receptor array $t^*(x_k, y_\ell)$ by the sum

$$g^*(x_i, y_j) = \sum_k \sum_\ell t^*(x_k, y_\ell) h^*(x_i - x_k, y_j - y_\ell)$$

where $h^*(x, y)$ denotes the impulse response matrix in the logarithmic domain. Taking the discrete two-dimensional Fourier transform, gives

$$g_i^*(x, y) = \mathcal{F}^{-1}\left(\sum_{j=1}^{3} \mathcal{F}[t_j^*(x, y)] H_{ij}^*(ju^*, jv^*)\right)$$

(Note: the shorthand notation $g_i^* = \sum_j H_{ij}^* t_j^*$ is used instead of the above expression.) Letting $t_1(\lambda)$ be equal to the luminous efficiency function V_λ, and considering the scalar expression for g_1^*, a familiar model of achromatic vision

$$g_1^*(x, y) = \mathcal{F}^{-1}(\mathcal{F}[\log I(x, y)] H^*(ju^*, jv^*))$$

is obtained [3]. This model has been successfully used to predict various visual phenomena such as Mach bands, simultaneous contrast, etc., as well as for image coding and processing [4].

Spatial Color Contrast Phenomena Examination will now be directed toward the mechanisms by which chromatic contrast effects such as color shadows, etc., can be represented by the model. Suppose that a small test spot with tristimulus values t_1, t_2, t_3 is viewed against a colored background with values t_{10}, t_{20}, t_{30}. Assuming total lateral inhibition, e.g. mutual subtractions of the outputs from neighboring receptors, the apparent chromatic quantities $g_2^{*'}$, $g_3^{*'}$ and $g_{02}^{*'}$, $g_{03}^{*'}$ of the test color and background respectively, become

$$g_2^{*'} = \log\left(\frac{t_2}{t_1}\right) - \log\left(\frac{t_{20}}{t_{10}}\right) = \log\left(\frac{t_2}{t_1}\,\frac{t_{10}}{t_{20}}\right)$$

$$g_3^{*'} = \log\left(\frac{t_3}{t_1}\,\frac{t_{10}}{t_{30}}\right)$$

$$\left.\begin{aligned} g_{20}^{*'} &\approx \log\left(\frac{t_{20}}{t_{10}}\,\frac{t_{10}}{t_{20}}\right) = 0 \\ g_{30}^{*'} &\approx \log\left(\frac{t_{30}}{t_{10}}\,\frac{t_{10}}{t_{30}}\right) = 0 \end{aligned}\right\} \triangleq \text{White}$$

In other words, the background appears to be white, while the test color is shifted towards a color complementary to that of the background. The same apparent chromaticities are predicted by the model if one assumes that the responses of the photo-receptors are decreased in proportion to the average energies absorbed (bleaching). The new tristimuli are

$$t_i' = \frac{t_i}{t_{i0}}$$

And the apparent chromatic quantities become

$$g_2^{*\prime} = \log\left(\frac{t_2}{t_1}\,\frac{t_{10}}{t_{20}}\right)$$

$$g_3^{*\prime} = \log\left(\frac{t_3}{t_1}\,\frac{t_{10}}{t_{30}}\right)$$

The apparent chromatic shifts predicted correspond to the vonKries coefficient law regardless of the physiological assumption we make (inhibition or bleaching).

In practice however, the visual system does not ignore completely a non-white average sceme chromaticity, particularly at saturated average chromaticities. Letting lateral inhibition be weighted by a factor $k^* < 1$ the following apparent chromaticities are obtained

$$g_2^{*\prime} = \log\left(\frac{t_2}{t_1}\right) - k_1^* \log\left(\frac{t_{20}}{t_{10}}\right)$$

$$= \log\left[\frac{t_2}{t_1}\left(\frac{t_{10}}{t_{20}}\right)^{k_1^*}\right] \qquad k_1^* < 1$$

$$g_3^{*\prime} = \log\left[\frac{t_3}{t_1}\left(\frac{t_{10}}{t_{30}}\right)^{k_2^*}\right] \qquad k_2^* < 1$$

Since the influence of various colored areas within the visual field on each other's corresponding perceptual quantities is a function of their geometric distance, let k_1^*, k_2^* be functions of the geometric distance between the excited receptors and write for each summation output

$$g_2^*(x_i, y_j) = \sum_k \sum_\ell \frac{t_2^*(x_k, y_\ell)}{t_1^*(x_k, y_\ell)} \qquad k_1^*(x_i - x_k, y_j - y_\ell)$$

$$g_3^*(x_i, y_j) = \sum_k \sum_\ell \frac{t_3^*(x_k, y_\ell)}{t_1^*(x_k, y_\ell)} \qquad k_2^*(x_i - x_k, y_j - y_\ell)$$

It is then observed that $k^*(x, y)$ is an impulse response, and the color contrast phenomena can be represented by the spatial filters $H_2^*(ju^*, jv^*)$, $H_3^*(ju^*, jv^*)$

$$g_2^* = \mathcal{J}^{-1}\left[\mathcal{J}\left(\log \frac{t_2(x,y)}{t_1(x,y)}\right) H_2^*(ju^*, jv^*)\right]$$

$$g_3^* = \mathcal{J}^{-1}\left[\mathcal{J}\left(\log \frac{t_3(x,y)}{t_1(x,y)}\right) H_3^*(ju^*, jv^*)\right]$$

Figure 2 shows a comparison of measured visual modulation transfer functions. The question whether the chromatic response is attenuated at very low spatial frequencies (spatial band-pass response) is subject to controversy. Note that the measurements of chromatic responses were not made in a logarithmic domain.

Spectral Sensitivities of the Receptors According to Konig's theory, color blindness known as dichromatism occurs whenever one of the three types of photo-receptors is inoperative in some fashion. The theory therefore predicts the existence of the three types of dichromats whose color matches are uniquely determined by two of the three energy absorption quantities

$$t_i = \int S(\lambda)t_i(\lambda)d(\lambda) \qquad i = 1, 2$$

Under Konig's assumption, colors with two fixed values t_i, t_j and arbitrary value t_k should all look alike to the corresponding dichromat. It can be shown that the chromaticity of such equivalent colors are sets of converging straight lines in the CIE x-y chromaticity diagram, and that the centers of convergence are the chromaticities of the respective missing primaries.

Experiments show that there exist indeed three types of dichromats, called protanopes, deuteranopes and tritanopes and that their color matches correspond well with the predicted behavior (fig. 3).

Let A be a 3 x 3 matrix which maps the CIE XYZ tristimulus space into the model's tristimulus space

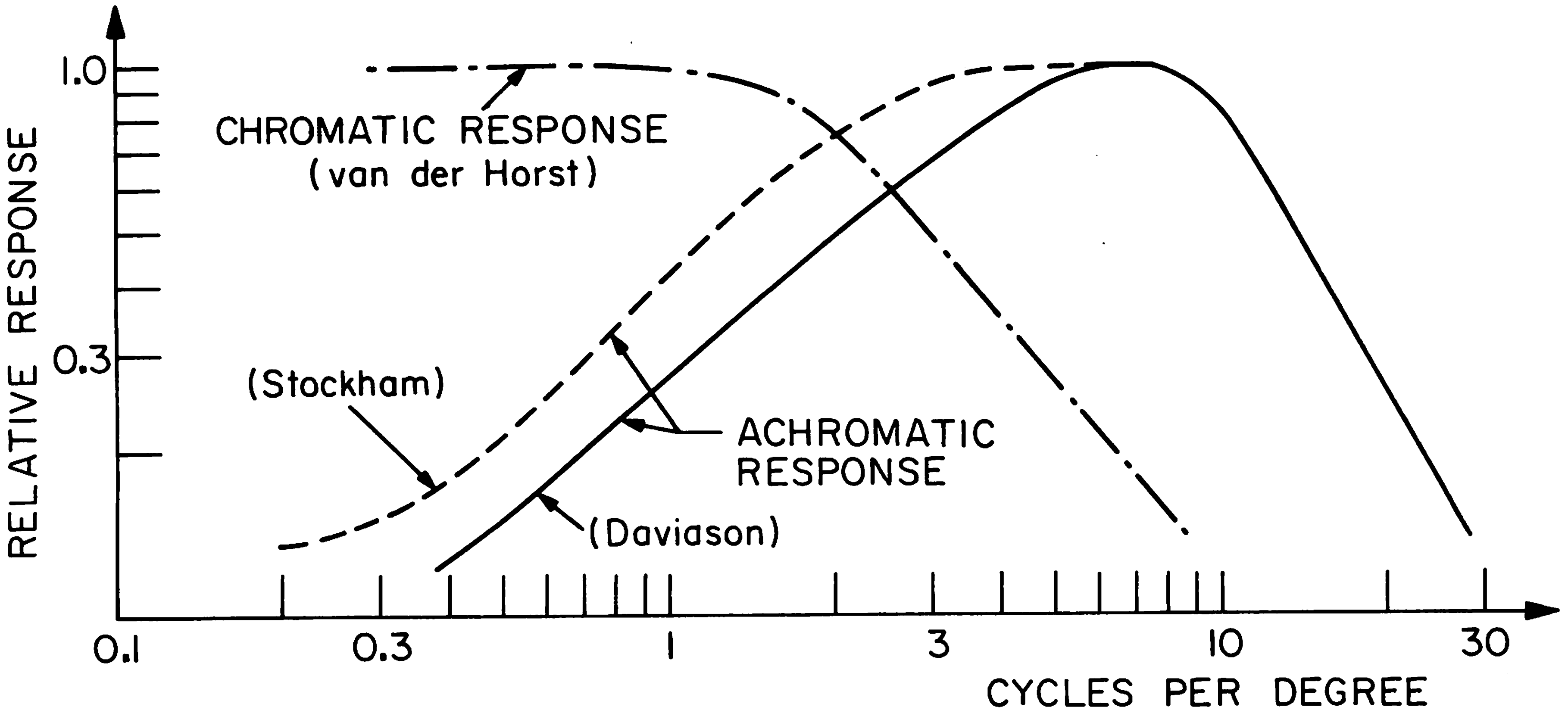


Figure 6.1-2. Comparison of modulation transfer functions.

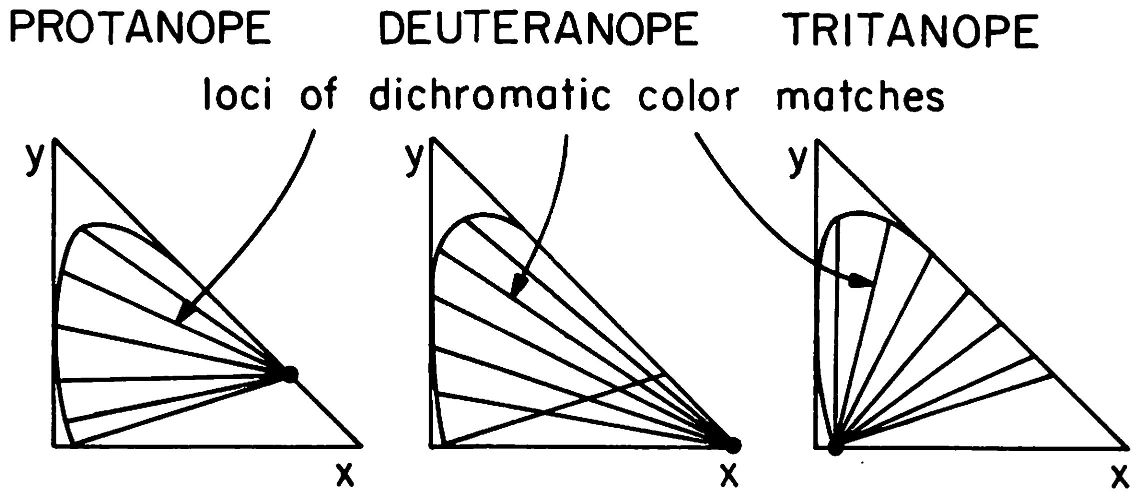


Figure 6.1-3. Confusion loci for the three types of dichromats in the CIE x-y chromaticity diagram.

$$\begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Also let $x_k, y_k, z_k, k = 1, \ldots 3$ be the chromaticity coordinates of the convergence for the protanope, deuteranope and tritanope respectively. Since x_k, y_k, z_k are the chromaticities of the unknown primarties, the following equations can be written to determine $A = [a_{ij}]$

$$\begin{bmatrix} t_1 \\ 0 \\ 0 \end{bmatrix} = [A] \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ t_2 \\ 0 \end{bmatrix} = [A] \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ t_3 \end{bmatrix} = [A] \begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix}$$

Three more equations can then be written by defining an arbitrary reference white for the model, for which $t_1 = t_2 = t_3 = 1$. Finally, the spectral sensitivity functions of the receptors are given by

$$\begin{bmatrix} t_1(\lambda) \\ t_2(\lambda) \\ t_3(\lambda) \end{bmatrix} = [A] \begin{bmatrix} \bar{x}_\lambda \\ \bar{y}_\lambda \\ \bar{z}_\lambda \end{bmatrix}$$

where $\bar{x}_\lambda, \bar{y}_\lambda, \bar{z}_\lambda$ are the CIE color matching curves for the standard observer. Using essentially the method outlined, and rounding Pitts coordinates for the deuteranopic convergence point from $x = 1.08$, $y = -0.08$ to $x = 1.00$ and $y = 0.0$, Judd obtained the following transformation matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -0.460 & 1.359 & 0.101 \\ 0 & 0 & 1 \end{bmatrix}$$

The spectral sensitivities of the receptors thus obtained are unimodal positive functions of the wavelength, as one would except from a physiological standpoint (fig. 4). Furthermore, one of the receptors exhibits a V_λ, or luminance response, as is required for our model to be considered with Abney's law of luminance addition. Figure 4 also shows the remarkable agreement between the measured luminous efficiency functions of the normal and dichromatic observers, and the luminous efficiencies predicted with the model.

Uniform Perception Space g^* While the visual model has been based upon physiological evidence, quite a few approximations have been made. It remains to verify whether the model predicts measured color differences.

First note that g_1^* = log luminance is in agreement with Weber's law of perception for intensities. As far as the chromatic quantities g_2^* and g_3^* are concerned, consider concentric circles centered at the origin and straight lines radiating from the origin in the g_2^*, g_3^* plane. If the model were perfect, the circles would be loci of constant saturation colors, and the lines loci of constant hue colors. Figure 5 shows a set of such equi-distant circles and lines, mapped into the CIE x-y chromaticity diagram. The agreement with the Munsell equi-distant set is remarkable if scaling constants c_1, c_2 (to be included in the filter function H_2^* and H_3^*) are introduced such that

$$g_2^* = c_1 \log \frac{t_2}{t_1}$$

$$g_3^* = c_1 \log \frac{t_3}{t_1}$$

Assuming that the sensation of brightness is evoked by the total activity of the three perceptual channels, brightness is then equal to the norm of g^*, $|g^*| = (g_1^{*2} + g_2^{*2} + g_3^{*2})^{\frac{1}{2}}$. The mapped set of circles and lines now represents equi-brightness geodesics. A remarkable agreement with the geodesics obtained with Stile's line element is noted [5]. Improvements may be expected by replacing the logarithmic approximation by a more realistic function of the type $\Delta I/K+I$. However, the original approximation shall be retained for the next discussion.

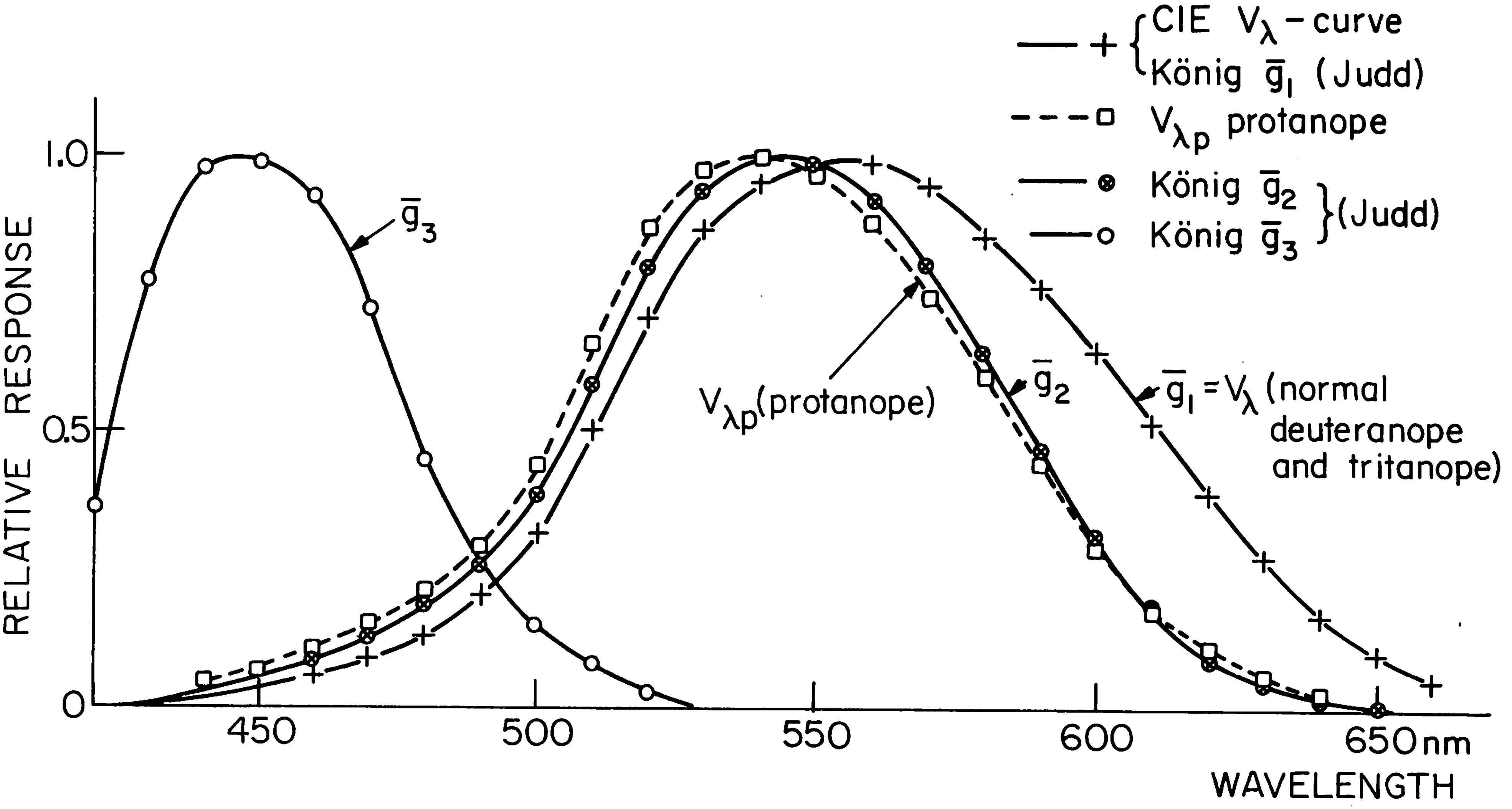


Figure 6.1-4. Comparison of Konig-type receptor sensitivities and luminous efficiency functions of normal and dichromatic observers.

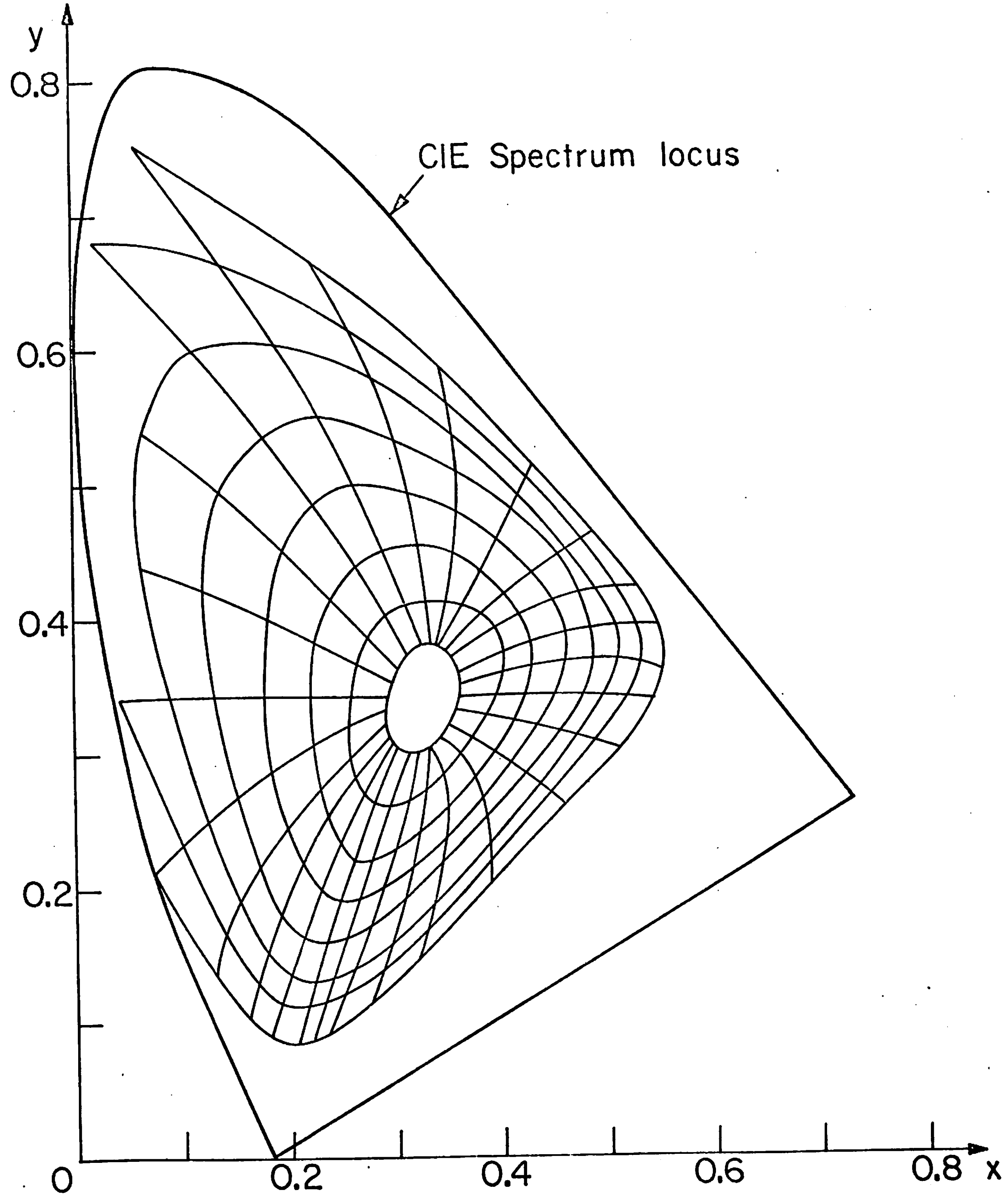


Figure 6.1-5. Equi-distant loci of constant hue and constant saturation predicted by the model.

Generalized Superposition in Tristimulus Space Define two vector spaces V and W over the field of real numbers $r \epsilon R$, with $t \epsilon V$, $t = (t_1, t_2, t_3)$ and $g^* \epsilon W$, $g^* = (g_1^*, g_2^*, g_3^*)$. Let the components t_1, t_2, t_3 of t be tristimulus values with respect to the spectral receptor sensitivities of the model and some arbitrary reference white, and define the following laws of composition on V

$$t_A, t_B \epsilon V$$

$$t_A + t_B = (t_{1A}t_{1B}, t_{2A}t_{2B}, t_{3A}t_{3B})$$

$$rt = (t_1^r, t_2^r, t_3^r)$$

Next, let the components g_1^*, g_2^*, g_3^* of g^* be the perceptual quantities defined in the model, and let ordinary addition and multiplication by scalars be the laws of composition on W. Since the three receptor sensitivity curves selected are positive between λ_L, λ_U (the limits of the visible spectrum), t_1, t_2, t_3 are always larger or equal to zero. Excluding zero light intensities (which is quite acceptable for practical images) the vector mapping

$$T: t \longrightarrow g^*$$

defined by the model as

$$T: (t_1, t_2, t_3) \longrightarrow \left(H_1^* \log t_1, H_2^* \log \frac{t_2}{t_1}, H_3^* \log \frac{t_3}{t_1}\right)$$

can easily be shown to be linear. The inverse mapping T^{-1} is

$$T^{-1}: (g_1^*, g_2^*, g_3^*) \longrightarrow \left(\exp(g_1^*)^{1/H_1^*}, \exp(g_1^*, g_2^*)^{1/H_2^*}, \exp(g_1^*, g_3^*)^{1/H_3^*}\right)$$

It is seen that the operations defined on t (which is an element of tristimulus space, relative to a particular basis determined by the model's receptor sensitivities) correspond to ordinary addition and multiplication by scalars of perceptual quantities in the euclidean space W. For example, multiplying the tristimuli of a color by a scalar

$$r(t_1, t_2, t_3) = (t_1^r, t_2^r, t_3^r) \triangleq (rg_1^*, rg_2^*, rg_3^*)$$

results in a color which appears r-times as light, r-times as saturated and, letting brightness be approximately proportional to the norm of g^*, a color r-times as bright as the original color is obtained.

Conclusions Generalized superposition thus defined opens a broad variety of image coding and processing related problems to analytical treatment. In particular, homomorphic filtering can be readily applied to color images, if tristimuli are referred to the basis indicated. The simple mapping $T: t \longrightarrow g^*$ can be used in any real-time environment to solve the old problem of optimal color quantization. Finally, it is hoped that the reflections on superposition will help to clarify some of the controversy on chromatic modulation transfer functions in the visual system.

References

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

2. de Valois, "Analysis of Response Patterns of LGN Cells," Journal Optical Society of America, Vol. 56, July 1966, pp. 966-977.

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

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

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

7. Image Processing Hardware and Software Projects

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

The first project describes the progress toward the development of a real time digital image display system for the ARPANET. This device which is connected to a port on the TIP provides a flicker free display of digital images transmitted over the network.

In the second report plans are outlined for the construction of a digital image recording and display system. This system will be capable of recording real time color television for playback into a digital computer; and will record computer generated images for real time playback.

The last report is concerned with progress on software development for image processing and networking.

7.1 Real Time ARPANET Digital Image Display

Toyone Mayeda

The digital color image display for use on the ARPANET is presently ready for software checkout. The refresh memory integrated circuits were finally received during February 1974.

A block diagram of the system is shown in Figure 1, and photographs of the system are shown in Figure 2. The display system specifications are listed below:

1. Receive asynchronous digital picture information from the ARPANET TIP with brightness resolution up to 64 levels and at input rates up to 19.2K band;
2. Store the received data in an array of up to 256 x 256 six bit picture values;
3. Display a true 6 bit monochrome image on the monitor;
4. Display a pseudo color image by use of a random access memory which is addressed by the refresh memory output. The RAM can be remotely programmed from the TIP or by local switch con-

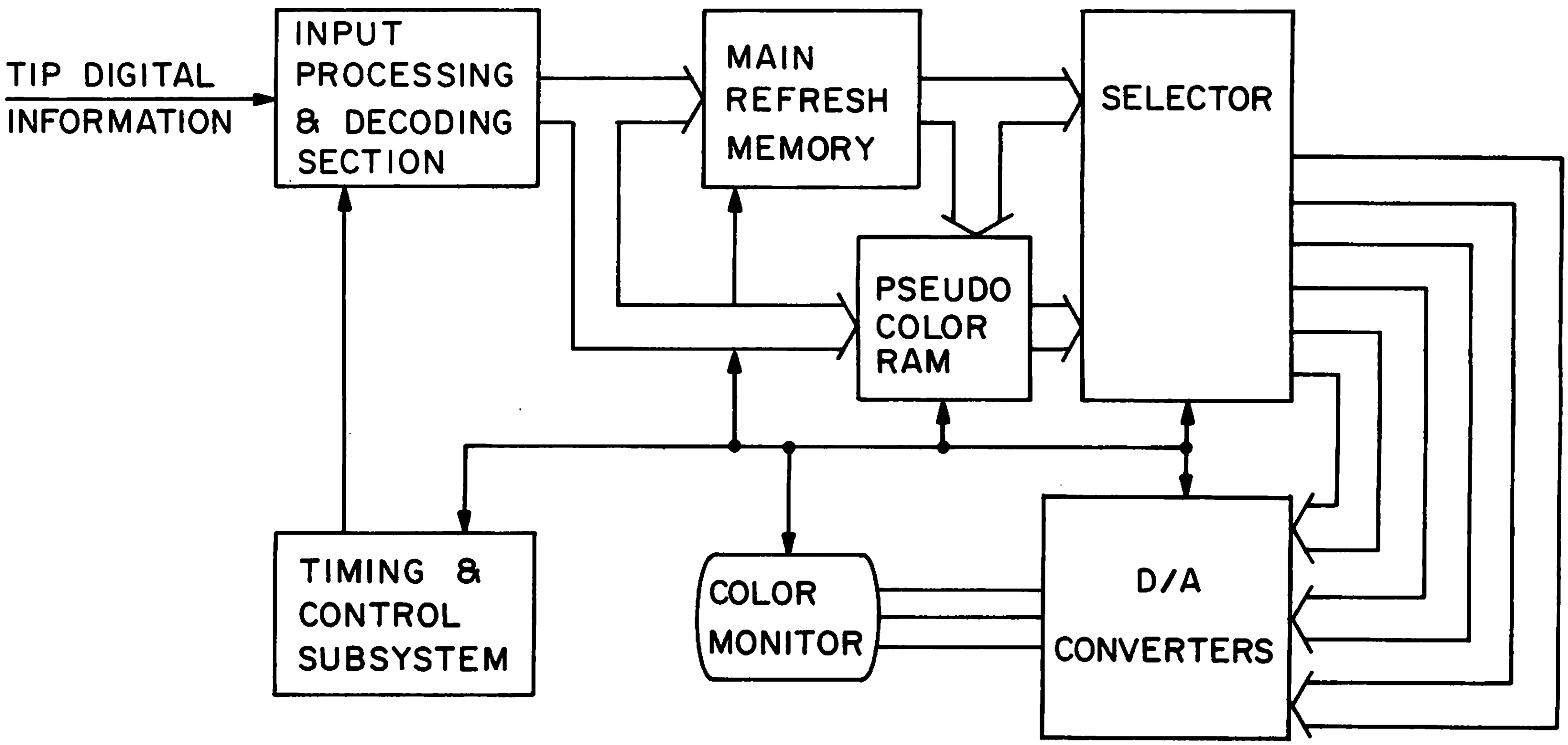
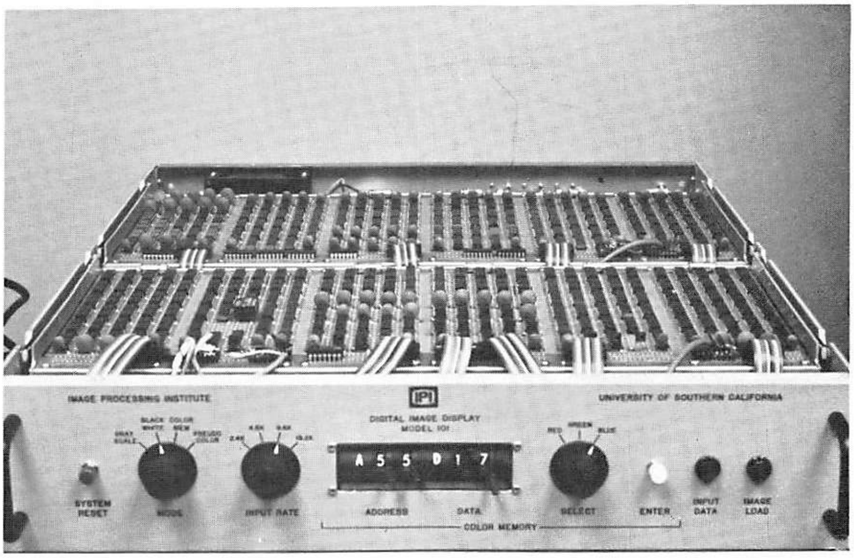
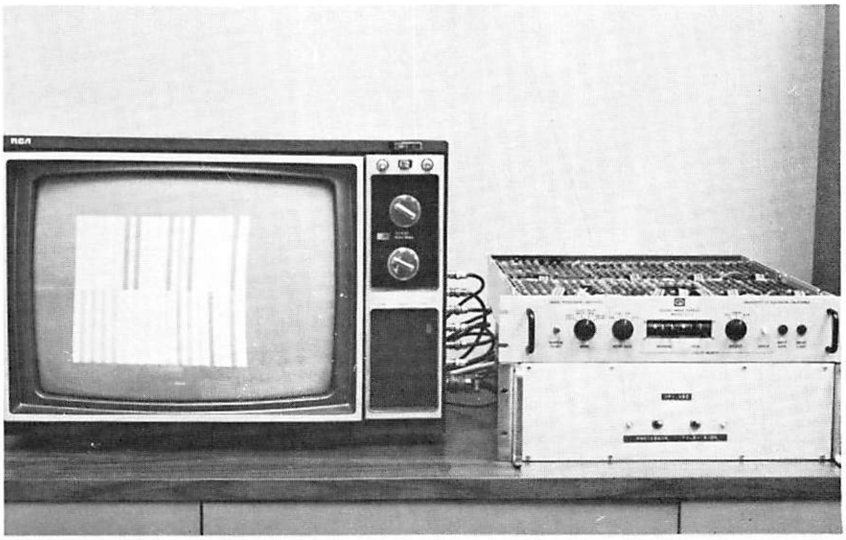


Figure 7.1-1. ARPANET digital image display.



(a) processor electronics



(b) total system - RCA color TV, processor electronics and power supply.

Figure 7.1-2. ARPANET digital image display.

trol. Over 4096 different color combinations of hue, saturation and luminance are available for pseudo-coloring.

Design of an 8 bit monochrome digital image display has been completed and is presently in the wire listing phase (prior to fabrication). The specifications of this system are:

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 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.

7.2 Real Time Color Image Digital Recorder and Display

Toyone Mayeda

Design has been started on a real time image digital recorder and display system as shown in Figure 1. The system will initially be designed to process an eight bit monochrome image, but eventually, will have the capability to process a color image using additional tracks of the recorder. The digital recorder and analog to digital converter have been ordered. With reference to the digital recorder in Figure 1, the monochrome system operates in the following four modes:

1. In the 600 ips record mode, video from the monochrome TV camera is digitized at approximately 10 MHz by the A/D conver-

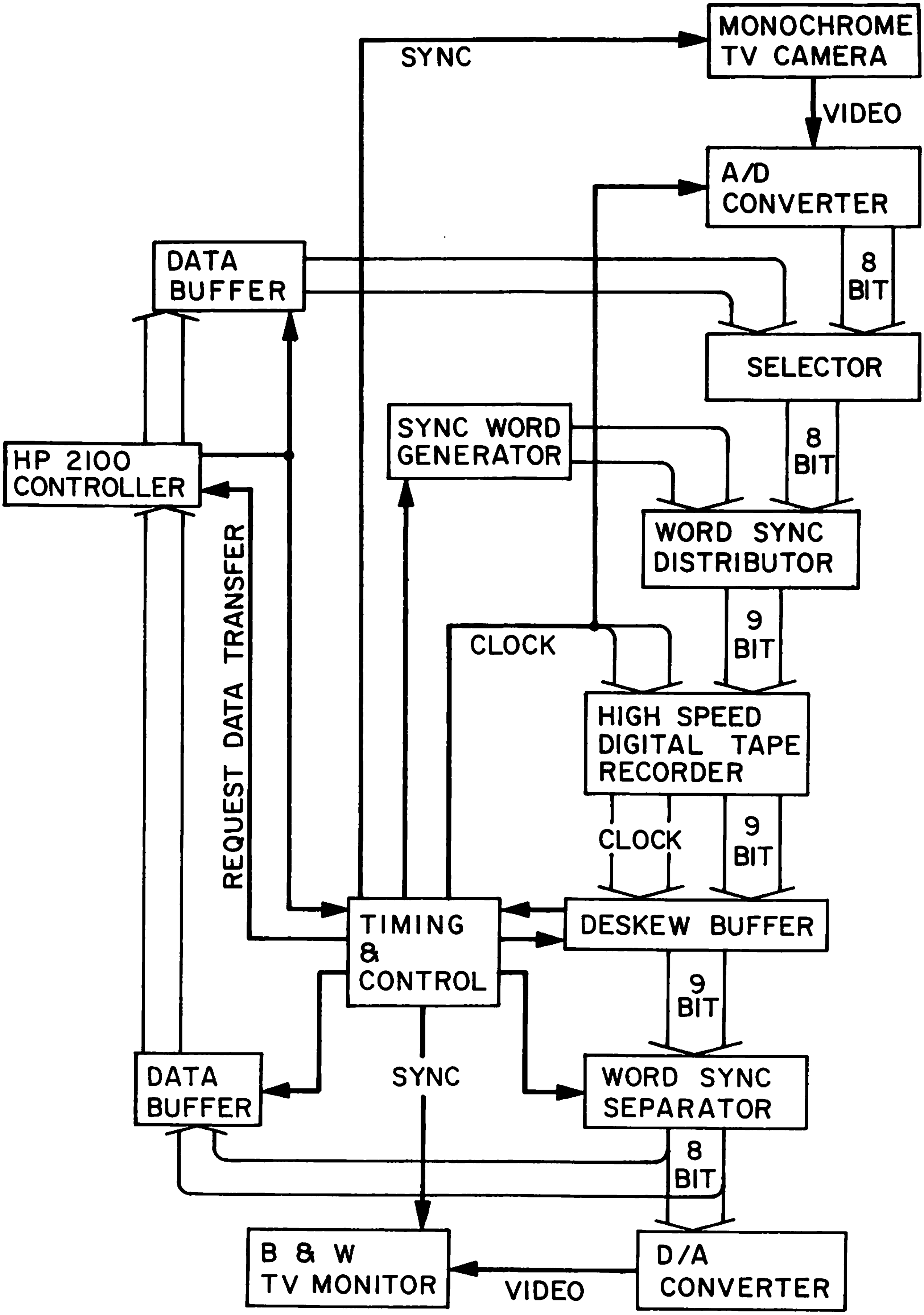


Figure 7.2-1. Real time image digital recorder and display.

ter and the eight digitized bits are selected. A ninth channel, which consists of a continuous 16 bit sync word pattern, is distributed with the eight data channels and all nine channels are recorded.

2. In the 1 7/8 ips playback mode, the nine channels are aligned in the deskew buffer. After the sync words are separated, the data in the eight remaining channels are transferred to the controller to be processed.
3. In the 1 7/8 ips record mode, the processed data from the controller is recorded with the sync word as in mode 1.
4. In the 600 ips playback mode, the nine channels are aligned and the sync words are separated as in mode 2. The eight remaining data channels are applied to a D/A converter and the output video data is displayed on a TV monitor.

7.3 Software Progress Report

James Pepin

In the last six months several important projects have been initiated by the programming staff. The most visible is a new network timesharing monitor. The second area of work is the interfacing of the Optronics scanner and digitizer. A third area of effort has been in the front-end image processing system design. Finally a PDP-11/10 has been added to the Image Processing Instite (IPI) hardware. This machine well act as a front-end to the HP2100, relieving it of some device control functions.

The new network support has many invisible improvements as well as some that are at once apparant. A password capability has been added to allow users data integrity. This also will allow more awareness of who is using the system and what they are doing. Another important feature added to the monitor is the capability to perform multi-tasking within the timesharing system itself. This will allow implementation of some complicated network protocals without complex code in each program. The system also has been made more efficient in its input/output and user swapping areas. This improvement should be very noticeable in heavy loading situations. Other improvements in this area include new and better help files. A file structure scheme has been implemented that will allow for faster and easier searches of HELP and other similar data bases.

In December IPI received an Optronics microdensitomiter. This device has required a large software effort for interfacing to the HP2100. This microdensitometer has many options that must be implemented by the computer so that a user can scan or display an image without detailed knowledge of the control logic.

At the end of January the Institute received a PDP 11/10. This computer will be used as a device controller to relieve the HP2100 of some of its controlling duties. As it now stands the 2100 has to handle all the 'bit watching' for the devices in the lab; with the addition of the 11/10 some of this task can be moved to the 11/10, allowing the 2100 to be free to handle other tasks. This technique will enable operation of two or more devices simultaneously.

During the last six months the staff has been participating in several committees that have as their responsibility the formulation of the needs of the image processing community on the ARPA net. There have been two areas of prime interest: software requirements and hardware requirements. The software area investigation has lead to the conclusion that some sort of insulation is necessary between the average image processing user and the large machines on the net. The approach taken is that a software package should be implemented that will present to the image processing user one command language. This will allow the user to use any large machine while only learning the command for image processing. Following this line of thought it follows that it would be desirable to have an image processing 'front-end'. This would be a computer system that would run this common language, convert that into the Job Control of the proper machine, and then send the JCL along with all pictures and other data needed to complete the user's task. This system could also act as a conferencing tool for image processors and could be a common collection point for literature in the field. Various hardware and software approaches have been investigated.

8. Publications

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

Harry C. Andrews, "Digital Fourier Transforms as a Means of Scanner Evaluation," Applied Optics, Vol. 13, No. 1, January, 1974, pp. 146-149.

Harry C. Andrews and C. K. Chen, "Nonlinear Intrinsic Dimensionality Computations," IEEE Transactions on Computer, Vol. C-23, No. 2, February, 1974.

Harry C. Andrews, "Digital Image Restoration-Survey," IEEE Computer, Vol. 7, No. 5, May, 1974.

Harry C. Andrews and Andrew 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.

William O. Crawford and Ernest L. Hall, "Computer Classification of Radiographs of Coal Workers Pneumoconiosis," Radiological Society of North America, November,1973.(also supported by NIOSH grant HSM-99-72-30.)

L. D. Davisson, "Universal Noiseless Coding," IEEE Transactions on Information Theory, November, 1973.

Ernest L. Hall, "Almost Uniform Distributions for Computer Image Enhancement, IEEE Transactions on Computers,February,1974.

Ernest L. Hall and William O. Crawford, "Automated Classification of Profusion of Opacities in Radiographs of Coal Workers," 26th annual Conference on Engineering in Medicine and Biology, September, 1973. (also supported by NIOSH grant HSM-99-72-30.)

Ernest L. Hall, Kendell Preston, and F. E. Roberts, "Classification of Profusion of Black Lung Disease from Chest X-rays," First International Conference on Pattern Recognition, October 1973. (also supported by NIOSH grant HSM-99-72-30.)

E. L. Hall, R. P. Kruger and A. F. Turner, "An Optical Digital System for Automatic Processing of Chest X-rays," Accepted for publication, Optical Engineering.

Ali Habibi, "Application of Lower-Triangular Transformations in Coding and Restoration of Two-Dimensional Sources," Proceedings of National Telecommunications Conference, November 1973, Atlanta, Georgia, pp. 12D1-12D6.

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

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

Ali Habibi, Guner S. Robinson, "A Survey of Digital Picture Coding," IEEE Computer(special issue on Digital Image Processing).

Ali Habibi, William K. Pratt, Guner Robinson, et.al., "Real Time Image Redundancy Reduction Using Transform Coding Techniques," 1974 International Communications Conference, June, 1974.

Anil Jain, "Image Modelling for Unification of Transform and DPCM Coding of Two-Dimensional Images," National Electronics Conference, Chicago, October, 1973.

Anil K. Jain, "Computational Considerations in Generation of Geodesics in Color Space," Optical Society of America Annual Meeting, 1973.

Anil K. Jain, and Edward Angel, "Image Restoration, Modelling and Reduction of Dimensionality," accepted for publication in IEEE Transactions on Computers, May, 1974.

Richard P. Kruger, William B. Thompson, and A. Franklin Turner, "Computer Diagnosis of Pneumoconiosis," IEEE Transactions on Systems Man and Cybernetics, January, 1974. (also supported by NIOSH grant HSM-99-72-30.)

Richard P. Kruger, Ernest L. Hall, and A. Franklin Turner, "A Prototype Radographic System for Pneumoconiosis Chest X-Ray Screening," Proc. IEEE Conference on Systems Man and Cybernetics, Boston, Mass., November, 1973. (also supported by NIOSH grant HSM-99-72-30.

Nasser E. Nahi, Ali Habibi, "Nonlinear Adaptive Recursive Image Enhancement," Proceedings of First International Joint Conference on Pattern Recognition, Washington, D.C., Oct.1973, pp. 485-491.

William K. Pratt, "Digital Transforms," 1973 IEEE NEREM Signal Processing Seminar, Boston, Mass., November, 1973.

William K. Pratt, "Transform Image Coding Spectrum Extrapolation," Seventh Hawaii International Conference on System Sciences, January, 1974.

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

William K. Pratt, Wen H. Chen, Lloyd R. Welch,"Slant Transform Image Codng," IEEE Transactions on Communications,(to appear 1974).

Alexander A. Sawchuk, "Space-Variant Image Restoration by Coordinate Transformations", Journal of the Optical Society of America, Vol. 64, February, 1974, pp. 124-130.

James Winter, Mark A. Stein, and Alexander A. Sawchuk, "Transverse Axial Tomography Using Coherent Optical Image Reconstruction," Proceedings of Conference on Computerized Transverse Axial Tomography, UCLA, Los Angeles, November, 1973.