USCEE REPORT #459

Semiannual Technical Report
Covering Research Activity During the Period
1 March 1973 to 31 August 1973

by

William K. Pratt

August 1973

Signal and Image Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
Department of Electrical Engineering-Systems
3740 McClintock Avenue, Room 404
Los Angeles, CA 90089-2564 U.S.A.

ABSTRACT

This technical report summarizes the image processing research activities performed by the University of Southern California during the period of 1 March 1973 to 31 August 1973 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 Enhancement and Restoration 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
Wen-Hsiung Chen
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
James M. Pepin
Easter D. Russell
Mark A. Sanders
George Soen
Wai Szeto
John Tahl
Florence B. Tebbets
Linda M. Webster

Students

Ben Britt
Steve Dashiell
Faramarz Davarian
Roy M. Glantz
Michael Huhns
Steve Hui
Mohammad Jahanashahi
Yik Kwoh
Alan Larson
Robert Liles
Eduardo Lopez
Simon Lopez-Mora
Clanton Mancill
Nelson Mascarenhas
Firouz Naderi
Michael Patton
Mohammad Peyrovian
Stuart Robinson
Dennis Smith
Robert Wallis
Carl Wedberg
Pamela Welch

TABLE OF CONTENTS

1. Research Project Overview

This report describes the progress and results of the University of Southern California image processing research study for the period of 1 March 1973 to 31 August 1973. 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:

Powell Hall. The U. S. C. Image Processing Institute has moved to the recently completed Powell Hall of Information Science and Systems Engineering on the U. S. C. main campus. Powell Hall is a six story building of over 120 offices with the ground floor devoted to laboratories. These laboratories include the Engineering Computer Laboratory (the IBM 360/44 computer host to the ARPANET), the Communication Systems Laboratory, and the Image Processing Laboratory. The Image Processing Laboratory consists of separate rooms for digital image processing, image digitization and display, optical processing and holography, photographic processing, and biomedical image processing.

Image Processing Software Systems Meeting. On July 17 and 18 a meeting was held at U. S. C. on image processing software systems. Over sixty members of the image processing research and development community attended the meeting. The major objectives of the meeting were to assess the existing hardware and software systems for image processing and to discuss possible future plans for an ARPANET based universal image processing system. Presentations were made by representatives of some of the large scale computational centers on the network summarizing their capabilities. Then, talks were given on the characteristics of existing image processing software systems and languages. Finally, discussions were held on future plans. As a result of the meeting, the following committees were established to formulate preliminary plans for an ARPANET image processing system.

Systems Committee

William K. Pratt, U. S. C. - Chairman

* system management
* hardware and software coordination

Hardware Committee

Thomas Stockham, Utah - Chairman

* resource site selection
* front end hardware
* networking considerations
* image digitizer and displays, standards and systems

Software Committee

Harry C. Andrews, U.S.C. - Chairman

* command languages
* functional algorithms

3. Image Coding Projects

The research effort in image coding is directed towards bandwidth reduction systems for monochrome and color television systems both for real-time and slow scan operation, considering the various fidelity specifications required in each case. The results of the image coding research activity during the past six months are summarized here and presented more extensively in the following subsections.

In transform image coding, one of the means for achieving compression is to discard transform coefficients according to a given strategy. The reconstruction is usually obtained by setting the discarded coefficients to zero. Observing that the transform coefficients are usually partially correlated, a more sophisticated reconstruction scheme is studied, in which the missing coefficients are extrapolated instead of set to zero. Computer simulations show significant quality improvement of reconstructed images.

In the next report, a class of interpolative models for one and two dimensional signal representation is introduced for coding of image signals. It is shown that these models give lower mean square error (and entropy) for a given bit rate compared to standard DPCM models, and the reconstruction algorithm gives the optimum quantizer. This representation is found to be relatively insensitive to variations in image statistics.

A class of lower triangular transformations is studied, which are not orthogonal, but result in a desired uncorrelated transformed signal. It is shown that this transformation can be performed recursively for Markov sources, in a fashion much similar to DPCM. Due to this similarity, DPCM systems may be regarded as transform coders and readily compared with that class of coders.

The next report describes an adaptive interframe transform coder in which the amplitudes and phases of the transform domain are selected and quantized according to a predicted statistical model. The prediction is derived from the previously coded block. Computer simulations demonstrate the feasibility of very high compression factors, good quality

reconstructions having been obtained with as few as 0.5 bits/pixel.

Past reports have summarized results for the encoding of sources with unknown probabilities and no distortion applied to image coding. Preliminary results are presented here on coding with distortion, and there is an indication that it is possible to approach the theoretic distortion bound for a video source without apriori knowledge of its statistics.

Finally, the application of singular value decomposition techniques to image matrices is studied. Significant dimensionality reduction can be achieved as shown by pictures reconstructed from a small set of singular values. A scheme is presented for dimensionality reduction under the constraint of a given mean square error, and the significance of small singular values is discussed.

3.1 Transform Domain Spectrum Extrapolation

William K. Pratt

In transform image coding one of the coding techniques is to transmit only those transform coefficients that lie in some geometrical zone in the transform domain; these are usually the low frequency components. At the receiver it has been common practice to insert zeros for the missing components prior to the inverse transformation operation. However, it is known that for all transformations, other than the Karhunen-Loeve transform, the coefficients are partially correlated. Hence, it is possible to utilize knowledge of this correlation to estimate the missing coefficients rather than arbitrarily setting their values to zero.

Transform Spectra For purposes of analysis it is convenient to regard the pixels of an image or image block as an N element vector, $\underline{f}$, which is a sample of a random process with zero mean and known covariance matrix $\underline{K}_f$. This vector undergoes a linear transformation defined by the N×N element operator matrix $\underline{A}$. Typical transformations of interest include the Fourier, Hadamard, Haar, Slant, and Karhunen-Loeve transforms. The N element vector $\underline{f}$ denotes the one dimensional transformation of $\underline{f}$

as given by

$$\underline{\mathcal{f}} = \underline{A}\underline{f} \tag{1}$$

The elements of $\underline{\mathcal{f}}$ represent the spectral components of an expansion of $\underline{f}$ in terms of a set of basis vectors – the rows of $\underline{A}$. Since $\underline{f}$ is zero mean, $\underline{\mathcal{f}}$ also possesses a zero mean. For a unitary transformation it is easily seen that the covariance matrix of $\underline{\mathcal{f}}$ is

$$\underline{K}_{\mathcal{f}} = \underline{A}\,\underline{K}_f\,\underline{A}^{-1} \tag{2}$$

Spectrum Extrapolation-Noise Free Condition Figure 1a contains a diagram illustrating the generation of the spectral components of a signal, the selection or truncation of these components, and the estimation of the discarded components. The complete vector of spectral components $\underline{\mathcal{f}}$ is operated upon by an M x N (M< N) selection matrix S, that extracts certain elements of $\underline{\mathcal{f}}$ and records them in an M element vector $\underline{\mathcal{f}}_T$. The selection matrix contains a unit element at S(i, j) to transfer the jth element of $\underline{\mathcal{f}}$ to the ith element of $\underline{\mathcal{f}}_T$.

In the diagram of fig. 1a, the vector $\underline{\mathcal{f}}_T$ is multiplied by an N x M estimation matrix $\underline{W}$ which provides an estimate, $\hat{\underline{\mathcal{f}}}$, of the complete set of spectral components of $\underline{\mathcal{f}}$. An inverse unitary transformation $\underline{A}^{-1}$ then reconstructs an estimate, $\hat{\underline{f}}$, of the signal vector, $\underline{f}$. The estimator matrix $\underline{W}$ is chosen to minimize the mean square error, $\mathcal{E}$, between the signal vector and its estimate as defined by*

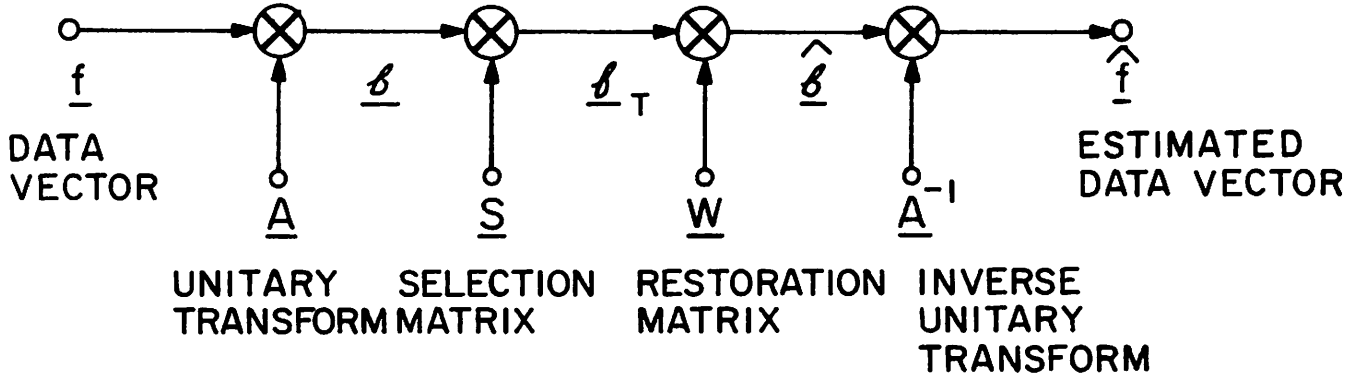
$$\mathcal{E} = \mathrm{tr}\,[E\{(\underline{f}-\hat{\underline{f}})(\underline{f}-\hat{\underline{f}})^{*T}\}] \tag{3}$$

The optimum choice of $\underline{W}$ can be found by forcing the dynamic error $(\underline{f}-\hat{\underline{f}})$ to be orthogonal to the observation $\underline{\mathcal{f}}_T$. Thus, by setting

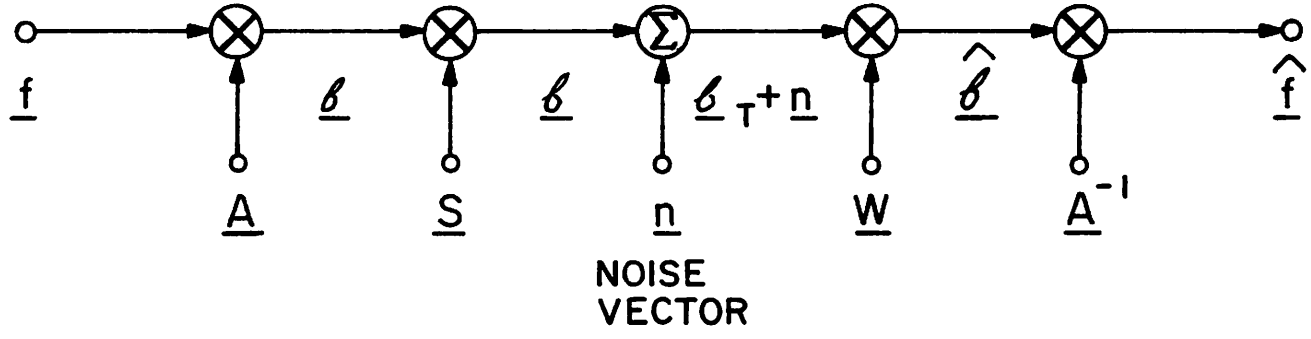
$$E\{(\underline{f}-\hat{\underline{f}})\underline{\mathcal{f}}_T^{*}\} = 0 \tag{4}$$

---

* Notation: $E\{\cdot\}$ = expected value; tr-trace of a matrix; * = complex conjugate; T = transpose.



(a) Noise-free estimation



(b) Additive noise estimation

Figure 3.1-1. Transform spectrum extrapolation operations.

one obtains

$$E\{(\underline{f} - \underline{A}^{-1}\underline{W}\underline{S}\underline{A}\underline{f})\,(\underline{S}\underline{A}\underline{f})^{*T}\} = \underline{0} \tag{5}$$

This leads directly to the optimum solution

$$\underline{W} = \underline{A}\underline{K}_f\underline{A}^{-1}\underline{S}^T[\underline{S}\underline{A}\underline{K}_f\underline{A}^{-1}\underline{S}^T]^{-1} \tag{6}$$

Premultiplying $\underline{W}$ by the selection matrix $\underline{S}$ yields

$$\underline{S}\underline{W} = \underline{I}_M \tag{7}$$

where $\underline{I}_M$ denotes an MxM identity matrix. Since

$$\hat{\underline{f}} = \underline{W}\underline{f}_T \tag{8}$$

It is clear from eq.(6) that

$$\underline{S}\underline{f} = \underline{f}_T \tag{9}$$

Thus, the estimator simply copies the elements of $\underline{f}_T$ into the appropriate elements of $\hat{\underline{f}}$. The remaining elements of $\hat{\underline{f}}$ are, in general, obtained by a linear combination of the elements of $\underline{f}_T$. In essence, the estimator linearly extrapolates the values of the observation of known spectral components to determine the unknown spectral values.

Spectrum Extrapolation - Additive Noise In physical situations an observation usually is contaminated by noise. If the noise is additive, as in fig. 1b, the optimum estimator that minimizes the mean square error, will, in effect, perform an additional task of noise filtering. Applying the orthogonality principle, one obtains

$$E\{(\underline{f}-\hat{\underline{f}})\,\underline{v}^{*T}\} = \underline{0} \tag{10a}$$

or

$$E\{[\underline{f}-\underline{A}^{-1}\underline{W}\,(\underline{n}+\underline{S}\underline{A}\underline{f})]\,[\underline{n}^{*T}+\underline{f}^{*T}\underline{A}^{-1}\underline{S}^T]\} = \underline{0} \tag{10b}$$

For the usual case in which the signal and noise are uncorrelated, the

optimum estimator is found to be

$$\underline{W} = \underline{A}\underline{K}_f\underline{A}^{-1}\underline{S}^T[\underline{S}\underline{A}\underline{K}_f\underline{A}^{-1}\underline{S}^T + \underline{K}_n]^{-1} \tag{11}$$

Performance Evaluation The usual alternative to spectrum extrapolation is simply to set the values of the discarded spectral components to zero and perform an inverse transformation. For such a strategy the restoration matrix is simply $\underline{W} = \underline{S}^T$ and the estimated vector becomes

$$\hat{\underline{f}} = \underline{A}^{-1}\underline{S}^T\underline{S}\underline{A}\underline{f} \tag{12}$$

Substitution into eq.(3) yields the sub-optimal mean square error

$$\mathcal{E}_{So} = \mathrm{tr}[\underline{K}_f(\underline{I}_N - \underline{A}^{-1}\underline{S}^T\underline{S}\underline{A})] \tag{13a}$$

or

$$\mathcal{E}_{So} = \mathrm{tr}[\underline{K}_f(\underline{I}_N - \underline{S}^T\underline{S})] \tag{13b}$$

If the optimal restoration matrix of eq.(6) or eq.(11) is employed the resulting mean square error is found to be

$$\mathcal{E}_o = \mathrm{tr}[\underline{K}_f(\underline{I}_N - \underline{W}\underline{S})] \tag{14}$$

Therefore, a reduction in mean square error by the amount

$$\mathcal{E}_D = \mathcal{E}_{So} - \mathcal{E}_o = \mathrm{tr}[\underline{K}_f(\underline{W} - \underline{S}^T)\underline{S}] \tag{15}$$

can be realized by spectrum extrapolation.

Figure 2 contains a plot of the mean square error of estimated spectral components for the Hadamard, and the Slant transforms. In this example, the data was assumed generated by a first order Markov process with a correlation factor of $\rho = 0.95$.

Experimental Results Computer simulation experiments have been performed to evaluate the spectrum extrapolation process for image coding. In the experiments illustrated in fig. 3, an image was Hadamard transformed

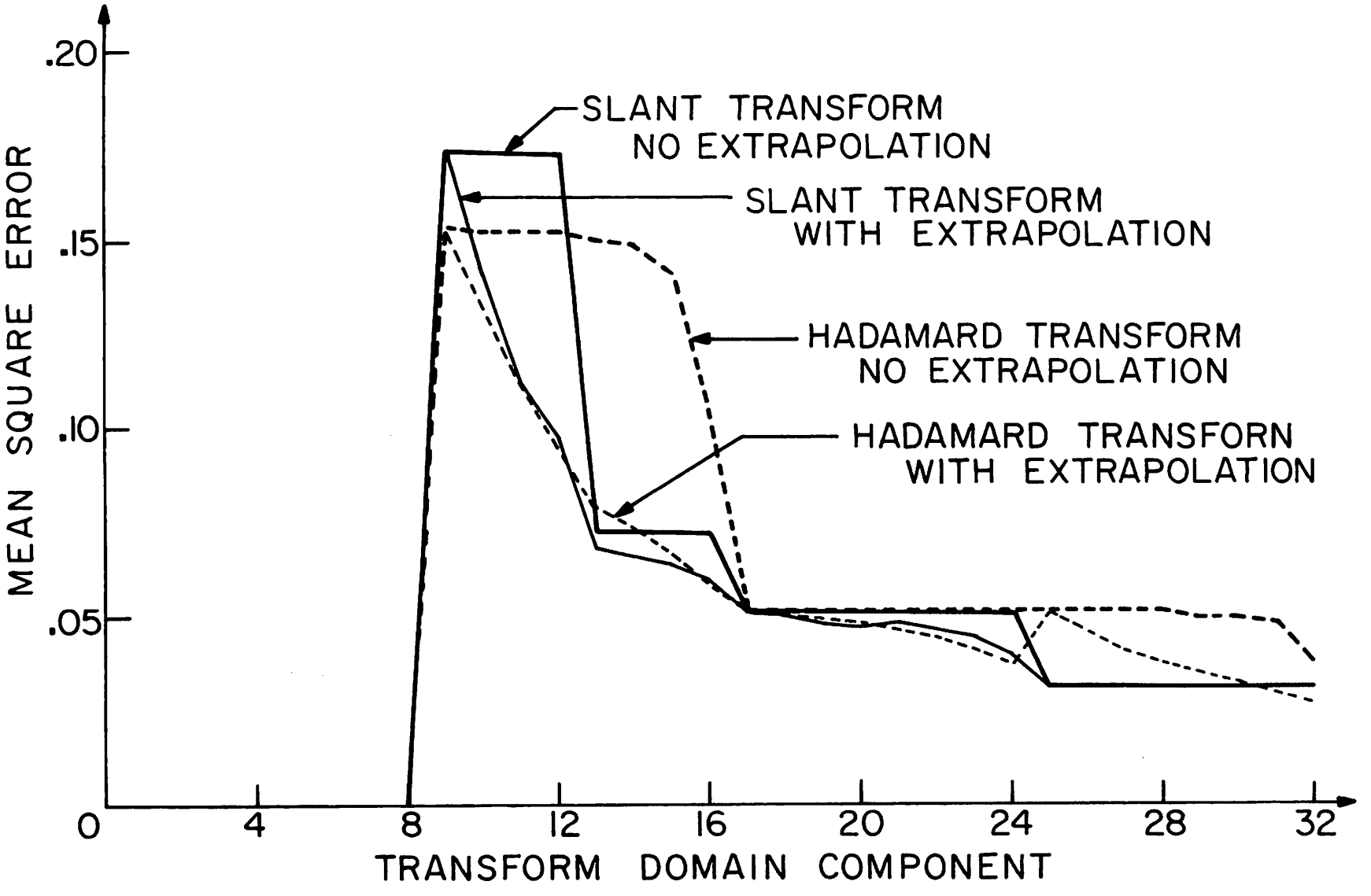


Figure 3.1-2. Mean square error of spectral components.

(a) original

(b) 4:1 zonal selection

(c) 4:1 zonal selection spectrum extrapolation

(d) 10:1 zonal selection

(e) 10:1 zonal selection spectrum extrapolation

Figure 3.1-3. Hadamard transform spectrum extrapolation examples.

in 16x16 blocks, and the coefficients lying outside of 8x8 and 5x5 low sequency zones, respectively, were set to zero before the inverse transformation. In figs. 3c and 3e the coefficients outside the zones were extrapolated assuming the image was a sample of a first order Markov process. The improvement in subjective quality is readily apparent from the photographs.

Conclusions The spectrum extrapolation technique, when applied to zonal transform image sampling, has been shown to offer significant improvement in image quality. It should be noted that the extrapolation processing is performed only at the receiver. Thus, the technique offers promise where a simple coder is essential, but a more complex decoder is permitted. Effort is continuing to apply the spectrum extrapolation to variable bit assignment transform algorithms to compensate for quantization error.

3.2 Role of Interpolative Models in Coding of One and Two Dimensional Signals

Anil K. Jain

Consider a stationary first order one dimensional Markov signal, u_i, $i = 1, \ldots, N$ represented by the correlation function.

$$E[u_i u_j] = \rho^{|i-j|}$$
$$0 \leq \rho \leq 1, \quad i, j = 1, \ldots, N \tag{1}$$

and let

$$E[u_i] = 0 \tag{2}$$

In DPCM coding of such a signal, one forms the differential signal

$$\epsilon_i \triangleq u_i - \rho u_{i-1} \tag{3}$$

and the sequence ϵ_i, $i=1, \ldots, N$, is quantized and coded. It is easy to show that this sequence has zero mean and the elements ϵ_i are uncorrelated and that

$$\sigma_1^2 \triangleq E[\epsilon_1^2] = 1-\rho^2 \tag{4}$$

Eq. (3) may be interpreted by saying that the differential signal is formed by scanning the signal in the direction of increasing i, and forming the difference between its scanned value u_i and its estimated value (ρu_{i-1}) from the previously scanned element u_{i-1}. In this sense eq. (3) is an online model, and can be rewritten for the signal representation as

$$u_i = \rho u_{i-1} + \epsilon_i \tag{5}$$

Suppose one wishes to represent u_i using two previous samples, i.e.,

$$u_i = \rho_1 u_{i-1} + \rho_2 u_{i-2} + \epsilon_i \tag{6}$$

If the model for the correlation function in eq. (1) is to hold, then it can be shown [1] that in eq. (6),

$$\rho_1 \equiv \rho \tag{7}$$

$$\rho_2 \equiv 0 \tag{8}$$

Thus it is seen that the sequence u_i with correlation matrix of eq. (1) cannot be represented by a second (or higher) order predictive Markov process. Now, suppose u_i is represented by

$$u_i = \rho_1 u_{i-1} + \rho_2 u_{i+1} + \nu_i \tag{9}$$

This means the estimate of u_i is obtained by using one prescanned and one postscanned element, thereby requiring a one step delay. It can be shown that in eq. (9)

$$\rho_1 \equiv \rho_2 \equiv \frac{\rho}{1+\rho^2} \tag{10}$$

and

$$E[\nu_i \nu_j] = \frac{1-\rho^2}{1+\rho^2}\,\delta_{ij}, \text{ for } \rho \simeq 1, \tag{11}$$

where δ_{ij} is the kronecker delta function. Comparison between eqs. (4) and (11) immediately shows that

$$E[\nu_i^2] \leq E[\epsilon_i^2] \qquad 0 \leq \rho \leq 1 \tag{12}$$

Therefore, if ν_i is used as the differential signal for quantization and coding, lower quantization error can be achieved for the same bit rate. This sequence $\{\nu_i\}$ will be called the "Interpolative Differential Signal".

Reconstruction Algorithm It has been shown [1] that the received signal ν_i^* can be used to reconstruct u_i by solving the following equations,

$$u_{i+1} = \gamma_i u_i + s_i \tag{13}$$

$$\gamma_{i-1} = \frac{1}{\left(\frac{1}{\rho_1} - \gamma_i\right)} \qquad \gamma_N = 0, \tag{14}$$

$$s_{i-1} = \left(s_i + \frac{\nu_i^*}{\rho_1}\right) \gamma_{i-1} \qquad s_N = 0 \tag{15}$$

Alternatively one can solve

$$\hat{u}_i = -\frac{\hat{\nu}_i^*}{\rho_1 \lambda_i} \tag{16}$$

where

$$\hat{u}_i = \sqrt{\frac{2}{N+1}} \sum_{j=1}^{N} u_i \sin\left(\frac{ij\pi}{N+1}\right) \qquad \hat{\nu}_i^* = \sqrt{\frac{2}{N+1}} \sum_{j=1}^{N} \nu_i^* \sin\left(\frac{ij\pi}{N+1}\right) \tag{17}$$

and

$$\lambda_i = \frac{1}{\rho_1} - 2\cos\left(\frac{i\pi}{N+1}\right) \tag{18}$$

The sequence u_i is then given by

$$u_i = \sqrt{\frac{2}{N+1}} \sum_{j=1}^{N} \hat{u}_i \sin\left(\frac{ij\pi}{N+1}\right). \tag{19}$$

Optimum Quantizer From eq. (16) the number of bits n_i, for minimum quantization error, can be shown to be given by

$$n_i = p + \frac{2}{N} \sum_{k=1}^{N} \log_2(\lambda_k) - 2\log_2(\lambda_i) \tag{20}$$

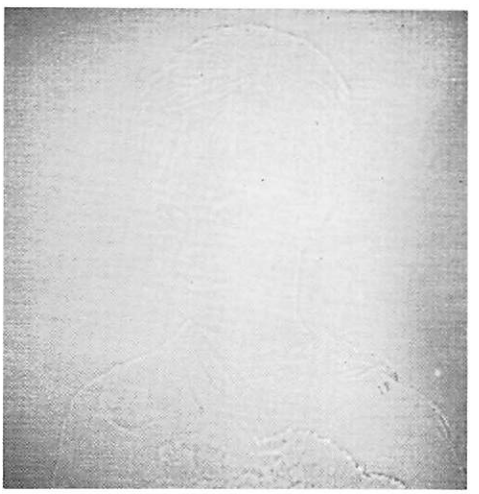
where p is the chosen bit rate in terms of number of bits per element. The optimum quantizer corresponding to eqs. (13), (14) and (15) can also be derived.

Results and Discussion The algorithm described above for one dimensional signals can be generalized to two (or more) dimensions. The quantization strategy (eq. 20) depends only on the eigenvalue distribution and for $\rho \simeq 1$ is very insensitive to variations in image statistics. The eigenvalues λ_i, for all such models are known in closed form as in eq. (18). Figures 1a to 1d show the implementation of the coding scheme mentioned here. The bit assignment in fig. 1d was chosen as $n_i \alpha - \log_2 i$. Experiments are currently under way using eq. (20) for the bit assignment strategy. Observe that $\hat{u}_i$ and $\hat{v}_i$ are related to the Fourier transforms of u_i and v_i (eq. (17)) and if one sets $\rho = 0$, then $\rho_1 \lambda_i \to 1$ and $\hat{u}_i^* = \hat{v}_i^*$. Thus for $\rho = 0$, the model reduces to the straight transform coding. Experiments performed for different values of ρ have indicated that for the values of ρ near unity, (0.9 to 1.0) the value of ρ may be set equal to unity without any appreciable effect on coder efficiency. The coding in figs. 1b to 1d actually corresponds to setting $\rho = 1$. For greater details and computational comparisons see reference [1].

References

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

(a) original image



(b) interpolative differential signal ϵ_{ij}

(c) encoded image with quantization levels proportional to variance of the error signal

(d) encoded image with quantization levels depending on eigenvalue distribution

Figure 3.2-1. Examples of interpolative image coding.

3.3 A Unified Representation of DPCM and Transform Coding Systems

Ali Habibi and Ronald S. Hershel

In recent years orthogonal unitary transformations have been used with considerable success in coding and restoration of discrete signals. One orthogonal transformation that results in the least mean square error is the Karhunen-Loeve transformation that generates an uncorrelated transformed signal. Other orthogonal transformations such as discrete Fourier, Hadamard, Haar, slant, cosine and others have been used to approximate the performance of the Karhunen-Loeve transformation because of their efficient computational properties. The Karhunen-Loeve transformation is indeed the only orthogonal transformation that results in an uncorrelated transformed signal. However, if the orthogonality constraint is removed, a second transformation of a lower triangular type can be found that also uncorrelates the signal. This transformation is computationally more efficient than the Karhunen-Loeve transformation, but it does not possess many of its desirable properties. An interesting property of this transformation is that the transformation can be accomplished recursively for Markov data. Indeed, it will be shown that coding Markov processes using this transformation results in a modified form of the differential pulse code modulator (DPCM), thus making DPCM a member of the class of transform coding systems.

Transformation by Lower-Triangular Operators Consider an N-dimensional data vector $X = (x_1, x_2, \ldots, x_N)^T$ and let X represent a sample vector from an ensemble of N-dimensional zero mean random variables. A vector $Y = (y_1, y_2, \ldots, y_N)^T$ can always be generated from a linear combination of x_i's as

$$y_1 = x_1 \tag{1a}$$

$$y_j = x_j - \sum_{k=1}^{j-1} \ell_{jk} x_k \qquad \text{for } j=2, \ldots, N \tag{1b}$$

or in vector form

$$Y = LX \tag{2}$$

where L is a unit lower-triangular matrix; i. e.,

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ -\ell_{21} & 1 & 0 & 0 & \cdots & 0 & 0 \\ -\ell_{31} & -\ell_{32} & 1 & 0 & \cdots & 0 & 0 \\ . & . & & & & . & . \\ . & . & & & & . & . \\ . & . & & & & . & . \\ -\ell_{N,1} & -\ell_{N,2} & & & & \ell_{N,N-1} & 1 \end{bmatrix} \tag{3}$$

Denoting the covariance matrices of X and Y by C_X and C_Y respectively, eq. (2) implies that

$$C_Y = L\, C_X\, L^T. \tag{4}$$

Choelsky has shown that for every symmetric positive definite matrix C_X there exists a real non-singular lower-triangular matrix L such that matrix $L\, C_X\, L^T$ is diagonal [1]. Martin and Wilkinson have considered numerical algorithms for finding L and C_Y and have developed efficient techniques requiring only $N^3/6$ multiplications [2, 3].

Some of the significant properties of the transformation by the unit lower-triangular operator are:

1) The unit lower-triangular operator L is not unitary; thus the transformation does not preserve the length of a vector. Although the determinant of the covariance matrix of X is invariant under this transformation, the trace of the covariance matrix is not invariant.

2) This transformation does not share the optimum concentration of energy in the first $M \leq N$ components of Y exhibited by the method of principal components. Indeed, for an n^{th} order Markov process, the variances of the Y components, except the first n components, are all equal.

3) Transformation with the lower triangular operator L does not

require a transformation delay.

Note that if the components of X are samples from an n^{th} order Markov process, the stochastic linear model of eq. (1) will be

$$y_j = x_j - \sum_{k=1}^{n} \alpha_k x_{j-k} \qquad j=1,2,\ldots,N \tag{5}$$

where $x_i = 0$ for $i=0,-1,-2,\ldots$. Then the operator L_n will be a banded matrix of n+1 bands

$$L_n = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & & & 0 \\ -\alpha_1 & 1 & 0 & 0 & \cdots & & & 0 \\ -\alpha_2 & -\alpha_1 & 1 & 0 & \cdots & & & 0 \\ \cdot & \cdot & & & & & & \cdot \\ \cdot & \cdot & & & & & & \cdot \\ \cdot & \cdot & & & & & & \cdot \\ -\alpha_n & -\alpha_{n-1} & \cdots & -\alpha_1 & 1 & 0 & \cdots & 0 \\ 0 & -\alpha_n & \cdots & -\alpha_2 & -\alpha_1 & 1 & 0 \cdots & 0 \\ \cdot & \cdot & & \cdot & & & & \cdot \\ \cdot & \cdot & & \cdot & & & & \cdot \\ \cdot & \cdot & & \cdot & & & & \cdot \\ 0 & 0 & \cdots & 0 & -\alpha_n & \cdots & -\alpha_1 & 1 \end{bmatrix} \tag{6}$$

Transformation of an N vector with operator L_n requires less than nN multiplication as compared to $N^2/2-N$ multiplications needed for transformation with the unit lower-triangular operator L in its general form. Furthermore, since transforming with the L_n operator requires only the n most recently encountered components of X, this transformation can be performed by using a feedback loop identical to one in an n^{th} order DPCM system to perform the transformation recursively. In this case, the complexity of the transformation is independent of the dimensionality of X and depends only on the order of the Markov process n. The block diagram of the system using the feedback loop for operation L_nX is shown in fig. 1.

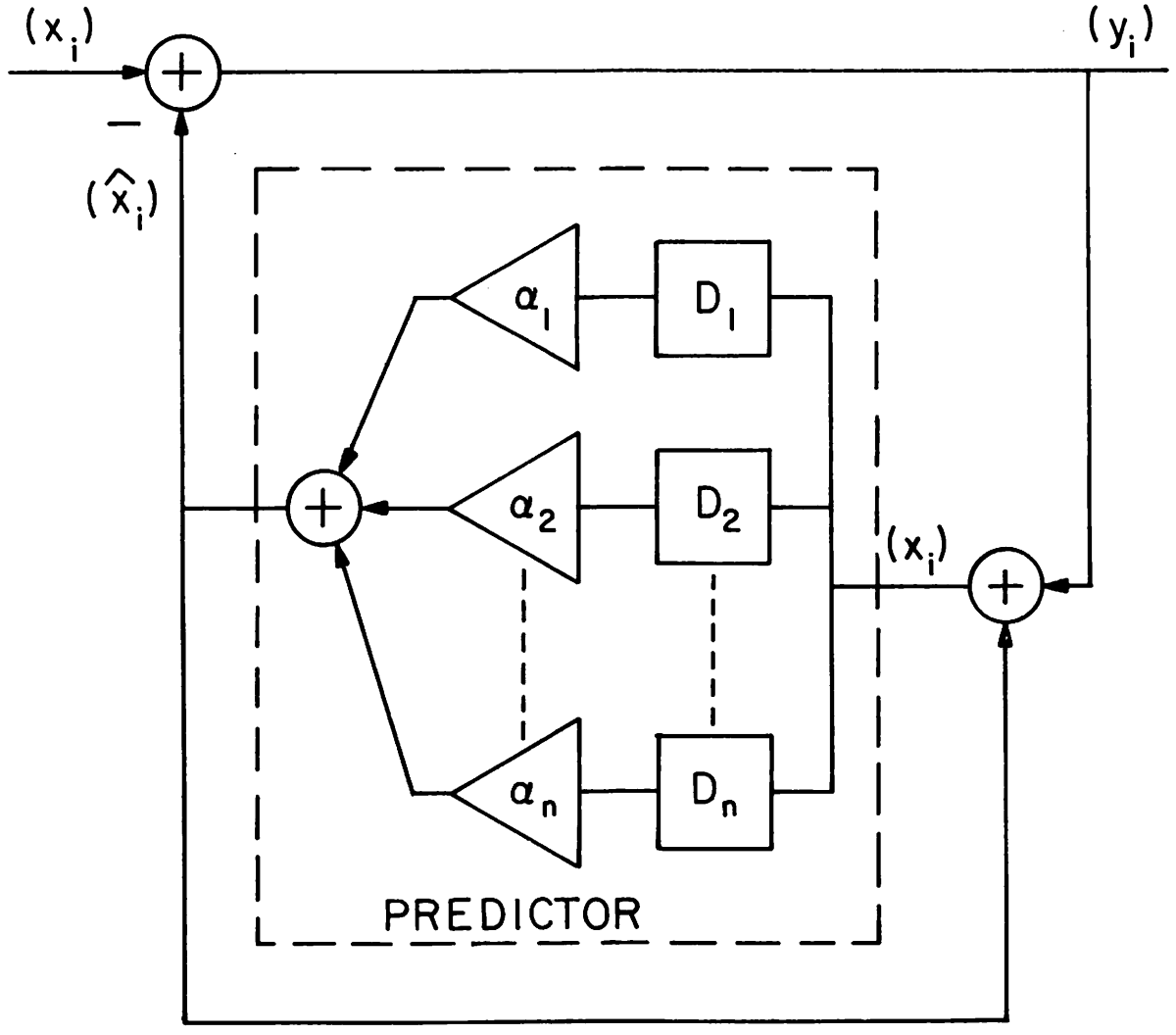


Figure 3.3-1. Block diagram of a system performing the transformation $Y = L_n X$.

Coding by Lower-Triangular Transformation In a transform coding system that uses a Karhunen-Loeve transformation and block quantization, the average coding error is minimum if the vector X is transformed to an uncorrelated vector Y by the matrix of eigenvectors of C_X. Block quantization of Y involves assigning a given number of bits to the N components of Y such that the quantization error is the same for all components of Y [4]. Denoting this error by Δ and observing the fact that the quantization error is additive and uncorrelated for various components of Y, it can be shown that the average coding error is bounded by

$$\epsilon_b^2 = \frac{1}{N} \operatorname{tr}\{\Delta I\} = \Delta. \tag{7}$$

Transformation of the vector X by the lower-triangular operator L results in a vector Y where the components of Y are uncorrelated and in general have unequal variances. Components of Y are quantized using the block quantization technique and are transmitted. The vector X can be reconstructed at the receiver, within some level of degradation, by operating on the coded vector $Y+Q$ with operator L^{-1}. The average coding error is bounded by

$$\epsilon_L^2 = \frac{\Delta}{N} \operatorname{tr}[L^{-1}(L^{-1})^T]. \tag{8}$$

Comparison of eqs. (7) and (8) shows that using the lower-triangular operator, a non-unitary matrix, rather than the Karhunen-Loeve transformation prior to block quantization, results in an inferior coding system. This is evident from the fact that L^{-1} is a unit lower-triangular matrix, thus

$$\operatorname{tr}[L^{-1}(L^{-1})^T] \geq \operatorname{tr} I. \tag{9}$$

When components of X belong to a first-order Markov process, the transformation is accomplished recursively using the block diagram on fig. 1 with a single feedback loop. Then, using linear prediction theory, it can be shown that the variance of y_1 is the same as the variance of x_1 where

the variances of y_2 through y_N are all equal to $(1-\alpha_1^2)$. Thus, with the block quantization coding system shown in fig. 2, only two quantizers are required. One quantizer encodes y_1 and the other y_2 through y_N; since y_2 through y_N have identical variances. From published literature, the performance of this encoder is improved by including the quantizer in the predictor loop as shown in fig. 3. This combines the operation of the quantizer with the transformation, and is identical to a DPCM system with the stipulation that a separate quantizer is used to encode the first component of the differential signal. Naturally the effect of using only one quantizer is negligible for large N; thus the two systems are identical. This is easily generalized to conclude that transforming an n^{th} order Markov process by a lower triangular operator L_n results in an n^{th} order DPCM system.

Since L is a triangular matrix, the operation of the quantizers can always be combined with the transformation to give a generalized DPCM system. In reference 5, the authors propose a combination of transformation and block quantization, similar to one in the DPCM system, and show that the performance of the generalized DPCM system at high bit rates approaches the performance of the encoder using the method of principal components and block quantization. This system uses $N^2/2-N$ multiplications to transform the data and requires no coding delays. It simplifies to an n^{th} order DPCM encoder if components of data X belong to an n^{th} order Markov process.

References

1. L. Fox, "Practical Solution of Linear Equations and Inversion of Matrices," Appl. Math. Ser. Nat. Bur. Stand., Vol. 39, 1954, pp. 1-54.

2. R. S. Martin and J. H. Wilkinson, "Symmetric Decomposition of a Positive Definite Matrix," Numerische Mathematik, Vol. 7, 1965, pp. 362-383.

3. R. S. Martin and J. H. Wilkinson, "Symmetric Decomposition of Positive Definite Band Matrices," Numerische Mathematik, Vol. 7, 1965, pp. 355-361.

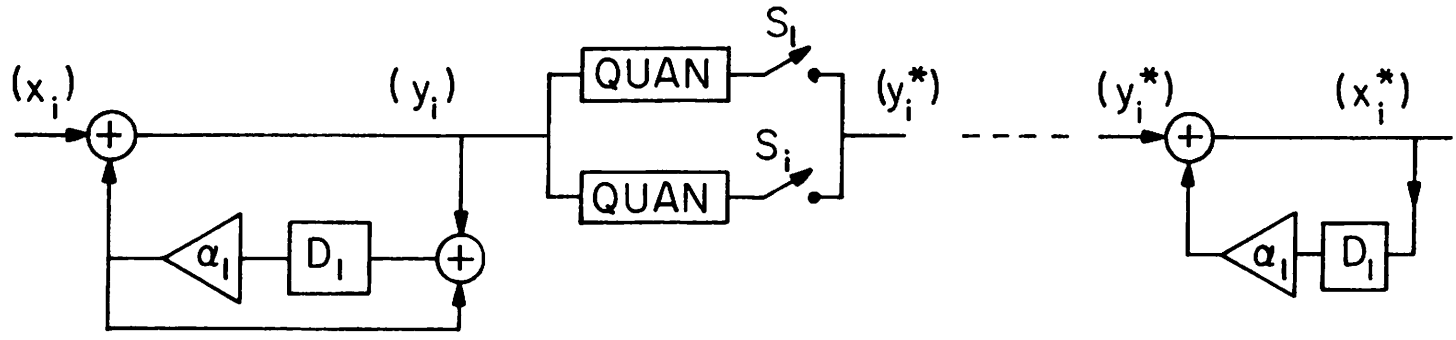


Figure 3.3-2. Block diagram of an encoder using the L_1 transformation and block quantization.

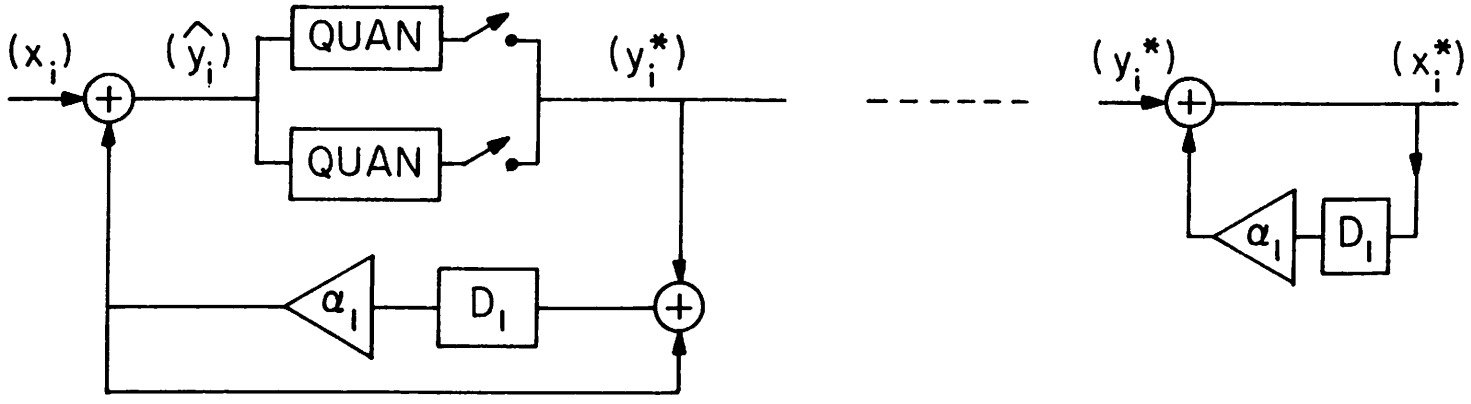


Figure 3.3-3. Block diagram of a modified DPCM encoder.

4. J. H. Y. Huang and P. M. Schultheiss, "Block Quantization of Correlated Gaussian Random Variables," IEEE Trans. Commun. Syst., Vol. CS-11, September, 1963, pp. 289-296.

5. A. Habibi, R. S. Hershel, "A Unified Representation of DPCM and Transform Coding Systems," to be published.

3.4 Interframe Transform Coding of Images

Andrew G. Tescher

A new transform image coding technique which is characterized by a high degree of adaptivity and is capable of very significant bandwidth reduction has been developed. The various implementations include the coding of two dimensional (monochrome and interframe) as well as three dimensional data (interframe and color). In this section, the indicated coding technique is discussed as related to interframe images.

An interframe image coder removes the high degree of redundancy which is present in a sequence of images representing the temporal variation in visual scenes. The particular implementation operates on a sequence of four images. The basic steps of this interframe coding procedure are shown in fig. 1.

The sequence of four frames, $I = \{I_1, I_2, I_3, I_4\}$ is processed by a three dimensional transform (Fourier and Walsh transforms are utilized in the specific implementations). The three dimensional transform consists of four 256 x 256 two dimensional transforms followed by the four element one dimensional transform along the temporal axis. In the final image transform, $\tilde{I} = \{\tilde{I}_1, \tilde{I}_2, \tilde{I}_3, \tilde{I}_4\}$. The image elements are considered to be uncorrelated within each transform plane as well as between different transform planes.

The coder "dynamically" determines the model of the transform domain on the basis of the previously coded transform elements and specifies the number and location of the required quantum levels for amplitude and phase quantization. The amplitude and phase are assumed to have Rayleigh and uniform probability density functions, respectively. The

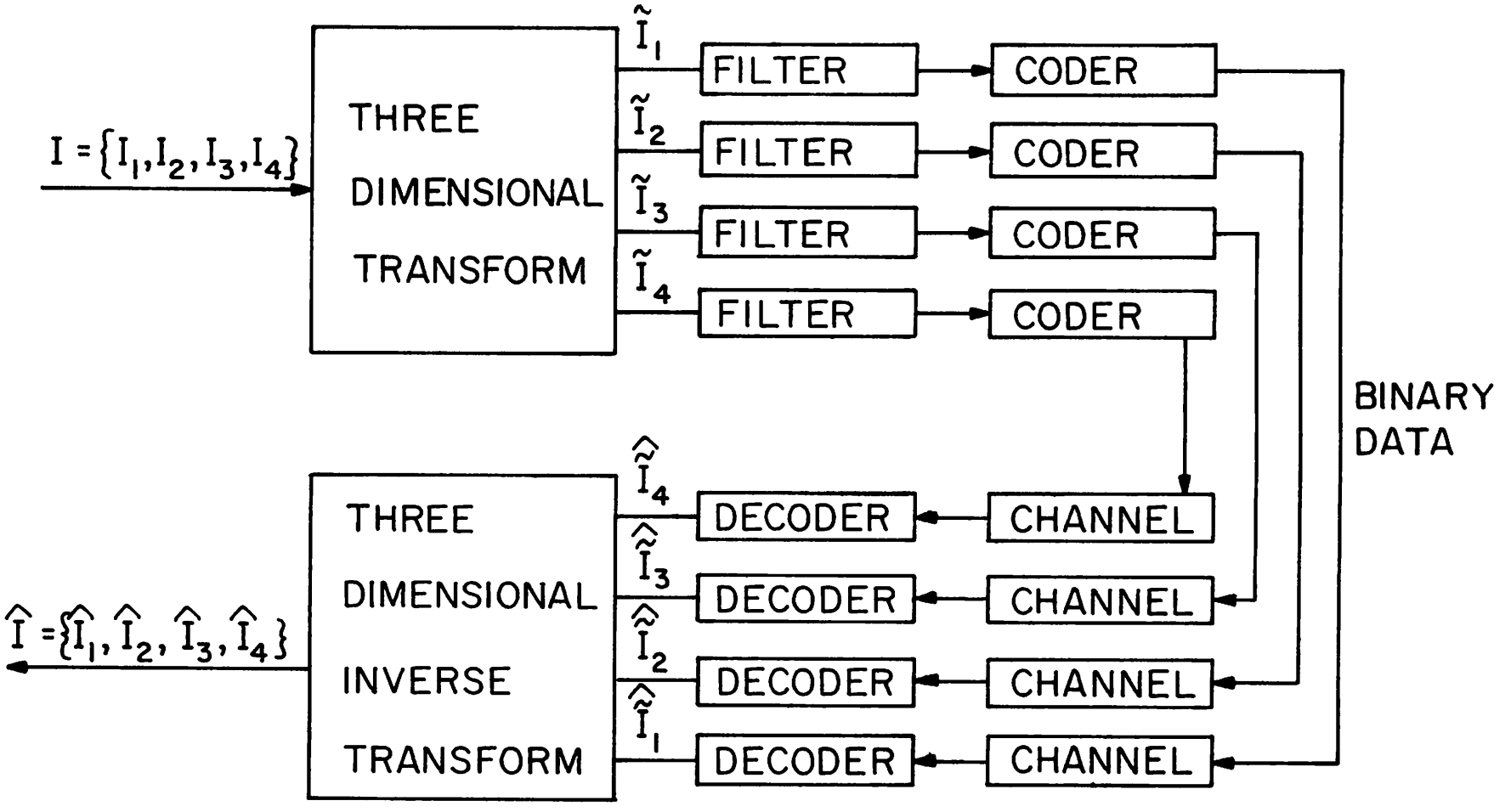


Figure 3.4-1. Interframe image coding system.

amplitude variance is estimated locally and is made proportional to the number of quantum levels for the companded amplitude. The number of bits for the phase is normally one higher except when the amplitude code is one bit in which case three bit quantization is utilized for the phase. The coder does not generate code words if the estimated amplitude variance drops below a specific threshold.

The coder shown in fig. 2, in effect, adapts to the three dimensional image power spectral density. The latter quantity may be modified by the multiplicative (linear or nonlinear) filters prior to the coding process. For example, the data rate may be reduced by low frequency filtering.

The decoder of fig. 3 simply repeats the estimating steps of the coder and determines the variance of the next amplitude and then reconstructs the quantified image amplitude and phase. The subsequent utilization of the three dimensional inverse transform completes the decoding process yielding the estimate of the four frames, $\hat{I} = \{\hat{I}_1, \hat{I}_2, \hat{I}_3, \hat{I}_4\}$.

The novel features of this image coding process are: a) a very high degree of adaptivity without any bookkeeping; b) the lack of any predetermined transform model (such as for example, the exponential Markov image correlation model); c) the capability for varying the amount of bandwidth reduction via the application of the filters preceeding the coder.

An example for the interframe coding procedure is shown in fig. 4. The eight bit images have been reconstructed using only 0.3 bit per picture element. The three dimensional Fourier transform was utilized for this case. An example for the adaptively determined bit assignment is shown in fig. 5. Here, the numbers indicate the adaptively determined amplitude code word lengths in the left half of a Fourier transform plane. Similar application of the above described transform coding technique to monochrome and color images demonstrates good quality image reconstruction with as low as 0.4 and 0.5 bits respectively.

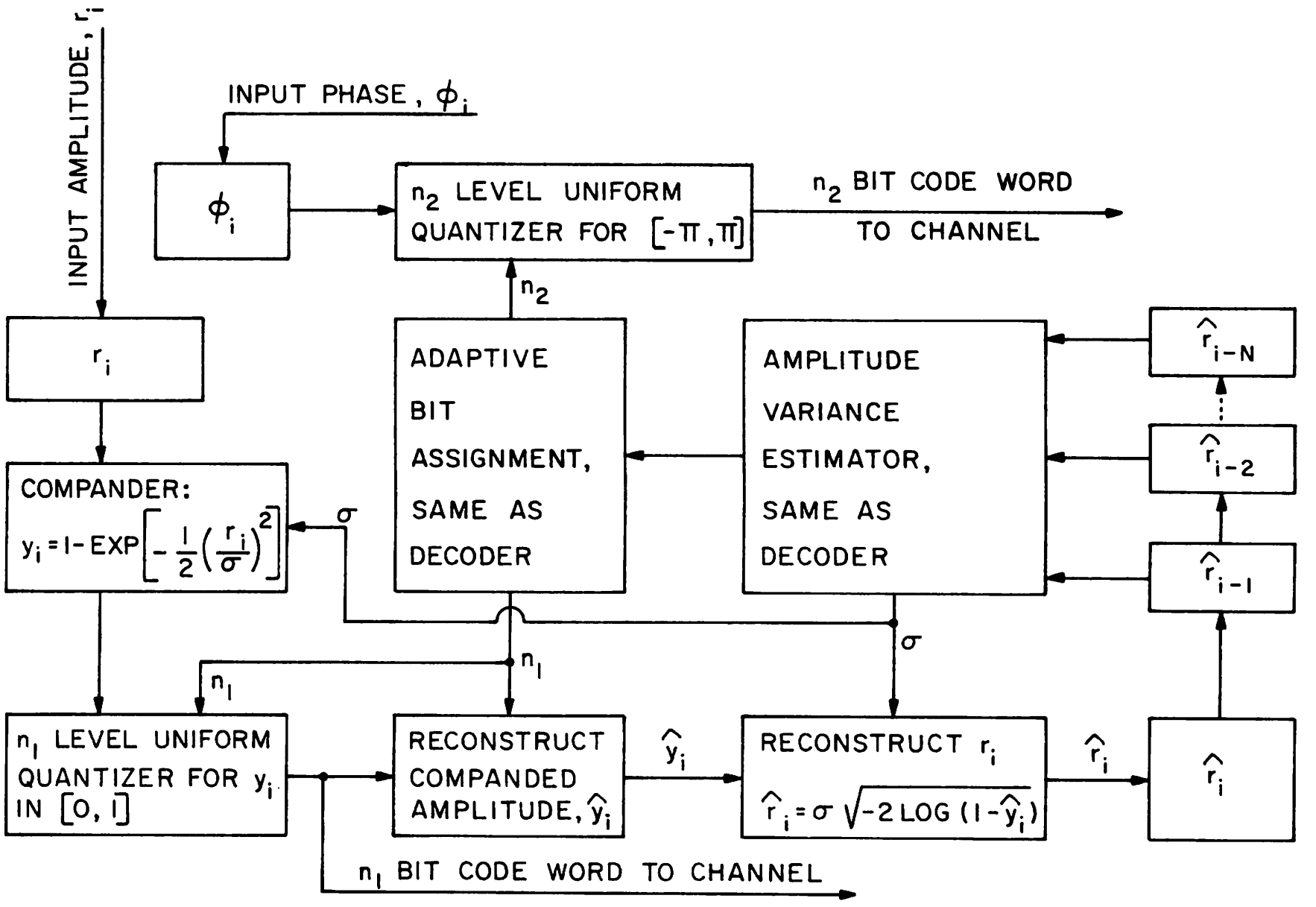


Figure 3.4-2. Adaptive coding algorithm.

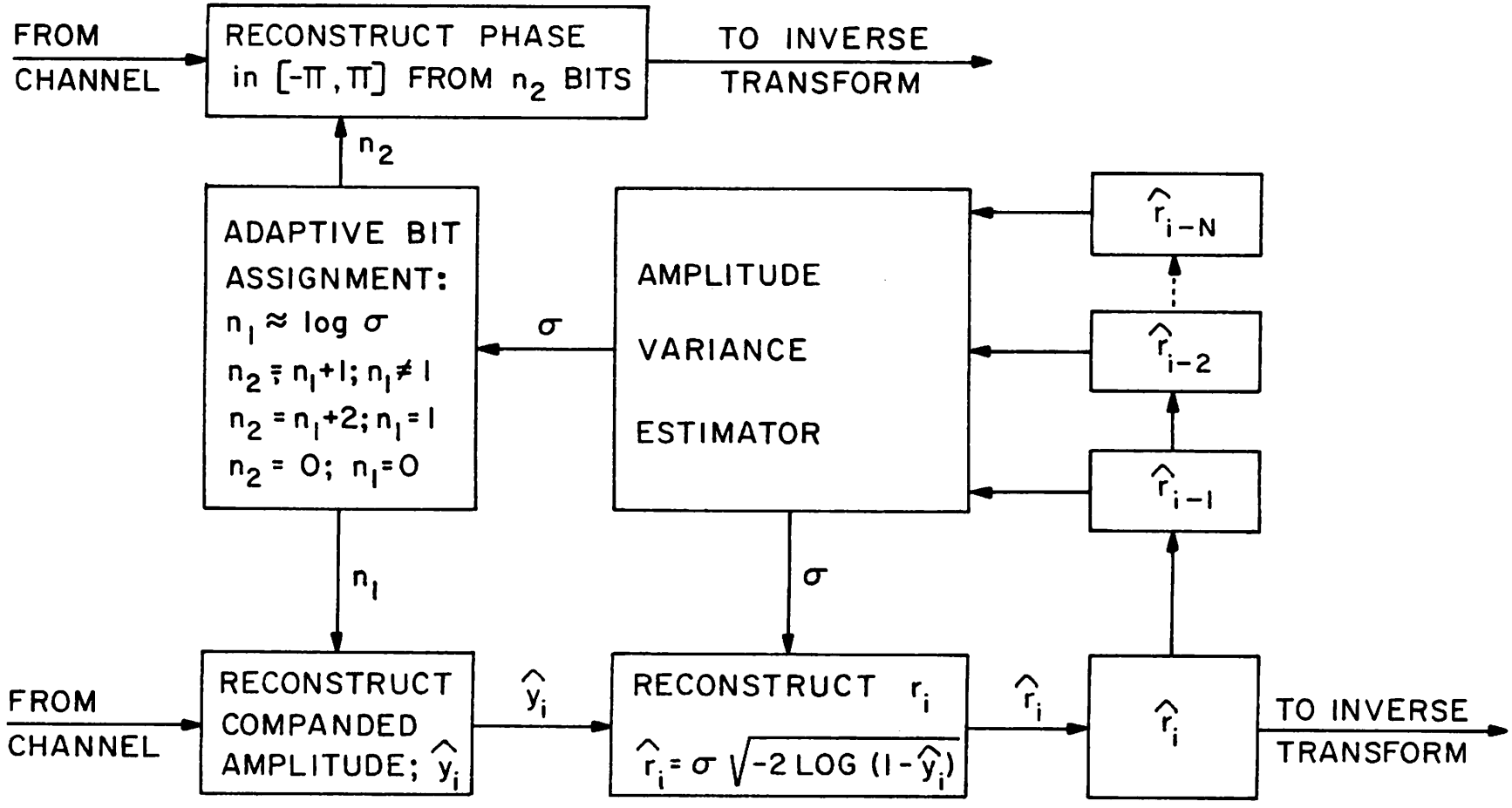
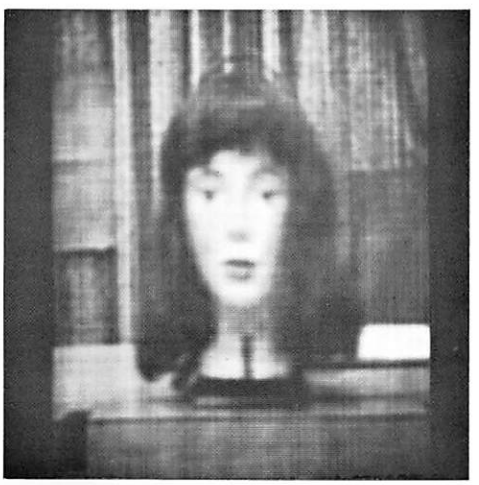


Figure 3.4-3. Adaptive decoding algorithm.

(a) original
8 bit/pixel

(b) original,
8 bit/pixel



(c) reconstructed image,
0.3 bit/pixel

(d) reconstructed image,
0.3 bit/pixel

Figure 3.4-4. Interframe coding of a four-image sequence
(the first two images are shown only)

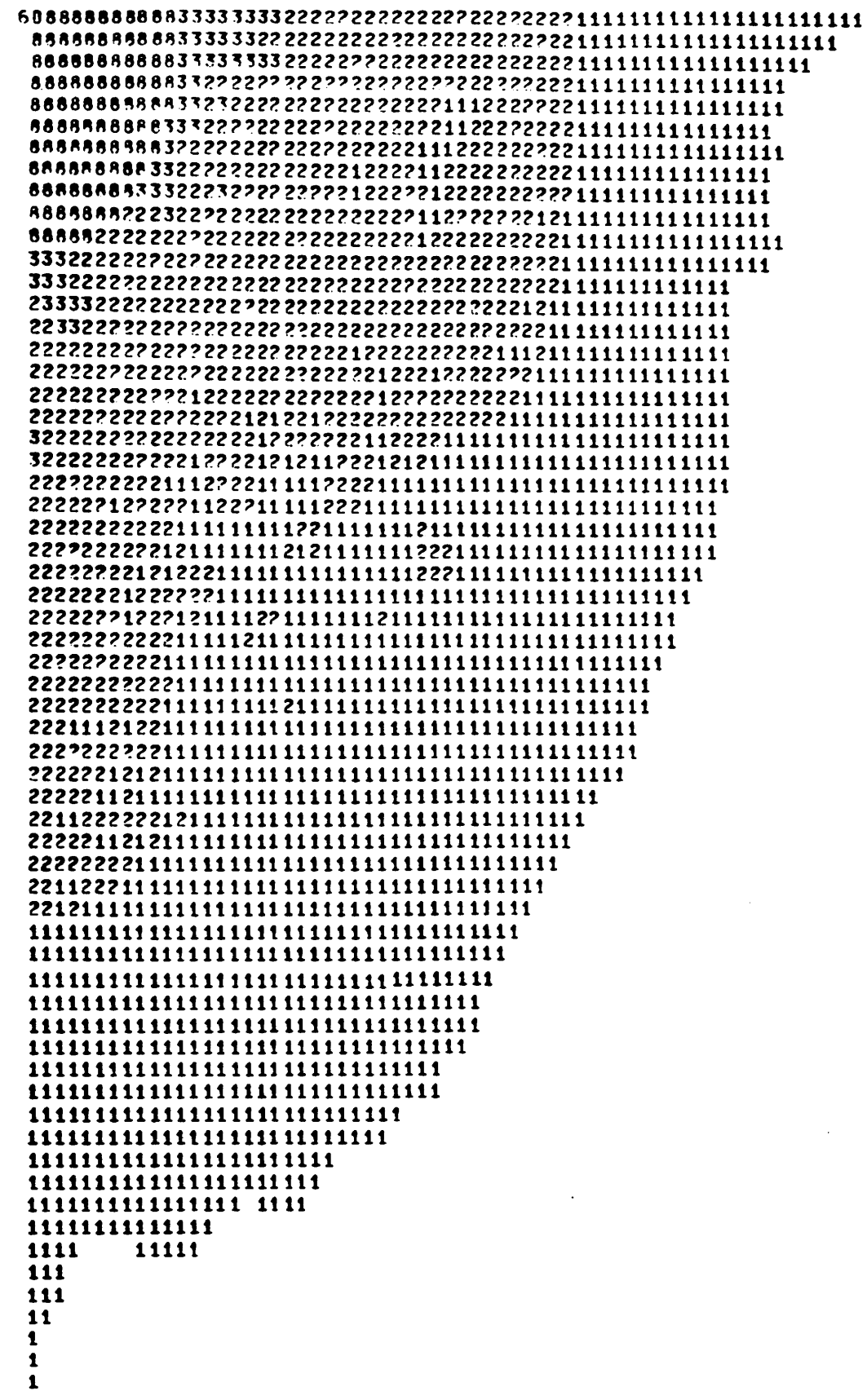


Figure 3.4-5. Adaptively determined bit assignment for amplitudes.

3.5 Coding Subject to a Fidelity Criterion for Sources with Unknown Probabilities

Lee D. Davisson

Past reports have summarized results for the encoding of image sources with unknown output probabilities without distortion. This report presents some preliminary results on coding when distortion is allowed. Of particular interest is the minimum possible average distortion when the bit rate of the coding procedure is fixed.

Assume a video source with individual output samples, called "letters", drawn from some "alphabet" A_0 of quantized values and a reproducing alphabet $\hat{A}_0$. A blocklength N code C_N is simply a collection of, say, $M = M(N)$ N-tuples $y_i^N = (y_{i_0}, \ldots, y_{i_{N-1}})$, $i=0, \ldots, M-1$, where each letter of each N-tuple is drawn from $\hat{A}_0$ so that $y_i^N \in \hat{A}^N$, the N-fold cartesian product of $\hat{A}_0$. A nonnegative single-letter distortion measure ρ_N defined on $A^N \times \hat{A}^N$ is given by

$$\rho_N(x^N, y^N) = N^{-1} \sum_{i=0}^{N-1} \rho_1(x_i, y_i) \tag{1}$$

where ρ_1 is called the per-letter distortion measure. A source is encoded by looking at a source N-tuple x^N and mapping it into the codeword $y^N \in C_N$ such that $\rho_N(x^N, y^N)$ is minimized. The resulting codeword is $\hat{x}^N(x^N)$, and then

$$\begin{aligned} \rho_N(x^N | C_N) &= \rho_N(x^N, \hat{x}^N(x^N)) \\ &= \min_{y^N \in C_N} \rho_N(x^N, y^N) \end{aligned} \tag{2}$$

The rate R of a code C_N is defined by $R = N^{-1} \ln M$ and the average distortion of a code $\rho(C_N)$ is given by

$$\rho(C_N) = E\{\rho_N(X^N | C_N)\} \tag{3}$$

where X^N denotes a random source vector and the expectation is over the source ensemble.

There are two approaches to studying the "optimum" performance obtainable in source coding systems. The first and by far the most common is the rate-distortion approach wherein the desired average distortion is fixed and an attempt is made to determine the minimum rate for which there exists a code meeting the fidelity constraint, i.e., one studies the function r(D) where

$$r(D) = \inf_N \inf_{C_N : \rho(C_N) \leq D} N^{-1} \ln M(C_N) \tag{4}$$

It is well known that for stationary ergodic sources r(D) can be given in terms of the rate-distortion function R(D), which is defined as an information theoretic minimization. Specifically,

$$r(D) \hat{=} R(D) \tag{5}$$

and for arbitrary $\epsilon > 0$, there exists a sufficiently large N and a code C_N having

$$\begin{aligned} N^{-1} \ln M(C_N) &\leqq R(D) + \epsilon \\ \rho(C_N) &\leqq D + \epsilon \end{aligned} \tag{6}$$

An alternative approach is to specify a fixed rate, and find the minimum possible average distortion attainable. This is the distortion-rate approach and is the more natural when one begins with a capacity constraint or a fixed number of quantization levels as opposed to an average fidelity constraint. In this case one is interested in the function $\delta(R)$ defined by

$$\delta(R) = \inf_N \inf_{C_N : N^{-1} \ln M(C_N) \leqq R} \rho(C_N) \tag{7}$$

and the answer for stationary-ergodic sources is given in terms of the distortion-rate function (inverse rate distortion function) D(R)

$$\delta(R) \hat{=} D(R)$$

and for arbitrary $\epsilon > 0$, there exists an N and a code C_N such that

$$\begin{aligned} N^{-1} \ln M(C_N) &\leqq R + \epsilon \\ \rho(C_N) &\leqq D(R) + \epsilon \end{aligned} \tag{8}$$

For the usual stationary, ergodic case the two views are entirely equivalent (although one may be more natural for a particular problem) and the functions are evaluated via identical parametric equations [1, Ch. 2]. This equivalence no longer holds, however, for fixed rate source coding stationary nonergodic sources as is typical for video sources where pictures of nonhomogeneous scenes are to be encoded.

Ergodicity is crucial in proving coding theorems since the law of large numbers or more general ergodic theorems provide a crucial step in random coding arguments. Roughly speaking, an ergodic source is a stationary source for which all doubly infinite output sequences are "typical" so that relative frequencies of functions of a finite number of letters of the process converge with probability one to their expectation. As an example, if one considers a Bernoulli process consisting of a sequence of independent, identically distributed binary random variables with $Pr(1) = p$, $Pr(0) = 1-p$, then the relative frequency of ones in a block of length N converges to p with probability one as $N \to \infty$. Now suppose that at time $-\infty$, Nature flips a fair coin and chooses either $p = p_1$ or $p = p_2$, $p_1 \neq p_2$, and then sends a Bernoulli random process with the chosen parameter. The resulting source is a mixture source and the probability of k ones in a block of length N is given by

$$\tfrac{1}{2} p_1^k (1-p_1)^{N-k} + \tfrac{1}{2} p_2^k (1-p_2)^{N-k}$$

so that the resulting mixture source is stationary, but it is no longer independent, identically distributed and no longer ergodic. It is not ergodic since any doubly infinite sequence (string) is not "typical" of the

mixture source, it is "typical" of one or the other of the subsources comprising the mixture, i.e., the relative frequency of ones approaches either p_1 or p_2, but not $\frac{1}{2}(p_1+p_2)$. Similarly, if Nature spins a fair wheel to determine p so that p is a uniform [0,1] random variable, one would have a stationary, nonergodic source consisting of a mixture of an uncountable number of stationary-ergodic subsources [9]. The ergodic decomposition theorem states that all stationary sources having an alphabet with a separable σ-algebra can be considered as a unique mixture of a possibly uncountable number of stationary-ergodic subsources so that the Bernoulli example above is in a sense typical.

Consider now the case of a source that is a mixture of a finite number, say K, of stationary ergodic subsources with a prior probability w_k, $k=0,\ldots,K-1$. This can also be viewed as a composite source with memory where the particular subsource is chosen at $-\infty$ and fixed for all time [1, Section 6.1]. A codebook for the mixture is constructed as follows. For each of the subsources a rate $R+\epsilon/2$ code $C_N(k)$ is developed such that for N sufficiently large,

$$\rho(C_N(k)|k) \leqq D_k(R) + \epsilon \tag{9}$$

where $\rho(C_N(k)|k)$ is the average distortion of a code C_N if the k^{th} subsource is the "true" source chosen by Nature. Then N is chosen large enough so that a $C_N(k)$ satisfying eq. (9) can be found for each $k=0,1,\ldots,K-1$. A super codebook

$$C_N = \bigcup_{k=0}^{K-1} C_N(k) \tag{10}$$

consisting of all of the words in each of the subcodes is then formed. The source encoder still finds the best word in C_N for whatever source N-tuple it sees. Thus C_N has average distortion

$$\rho(C_N) = \sum_{k=0}^{K-1} \rho(C_N|k)\, w_k$$

$$\leqq \sum_{k=0}^{K-1} \rho(C_N(k)|k)\, w_k$$

$$\leqq \sum_{k=0}^{K-1} D_k(R) w_k + \epsilon$$

and the number of codewords $M(C_N)$ contained in C_N is

$$M(C_N) = \sum_{k=0}^{K-1} M(C_N(k)) \leqq K e^{N(R+\epsilon/2)} \tag{11}$$

$$= \exp\{N(R+\epsilon/2 + N^{-1} \ln K)\}$$

so that by choosing N large enough, C_N has rate $\leqq R+\epsilon$. If the k^{th} subsource is present, then $\rho(C_N|k) \geqq D_k(R)$. Thus for arbitrarily small ϵ

$$\sum_{k=0}^{K-1} D_k(R) w_k \leqq \delta(R) \leqq \sum_{k=0}^{K-1} D_k(R) w_k + \epsilon \tag{12}$$

It is the weighted average of the distortion-rate functions of the individual ergodic components that yields $\delta(R)$, and not the usual distortion-rate function of the mixture source.

In fact, more than eq. (12) is true. Given a code C_N satisfying eq. (12), one might ask how well this code would perform for a specific subsource as opposed to its performance for the mixture. Specifically, what is the average distortion $\rho(C_N|k)$ if the k^{th} subsource is in effect? Since

$$\rho(C_N|k) \leqq \rho(C_N(k)|k) \leqq D_k(R) + \epsilon \tag{13}$$

one obtains the surprising result that when the ergodic decomposition is finite, there exists a single code C_N of rate approximately R such that the average distortion obtained under any subsource is the best possible, ignoring second order terms, i.e., is just as small as if we knew in

advance which subsource would be chosen by Nature.

The preceeding results can be extended to a nonfinite number of ergodic subsources as is more typical of video sources. The result is essentially the same although the methods of establishing the more general result are much more complicated. Thus it is known that it is possible to approach the information theoretic rate distortion bound for a video source without a priori knowing the source output probabilities. The implication of this theory to practical video coding is presently under investigation.

3.6 Image Data Compression Using Singular Value Decomposition

Monty Adler and Harry C. Andrews

The primary area of this research involves the use of singular value decomposition (SVD) as a tool for dealing with the inversion (or pseudo inversion) of large, nearly singular, matrices. As a by-product of this research, there have been attempts to use the SVD technique on images for the purpose of data compression and perhaps image enhancement.

SVD Definition and Properties In the SVD technique a digitized and sampled portion of an image (not necessarily square) is expanded as the sum of matrices of rank 1. If H is an NxN image, the first step in the SVD technique entails finding the eigenvectors and eigenvalues of the matrices HH^T and H^TH. Let HH^T have the eigenvector decomposition (λ_1, U_1), $(\lambda_2, U_2) \cdots (\lambda_N, U_N)$, i.e., $HH^TU_i = \lambda_i U_i$. Similarly, let H^TH have the eigenvector decomposition $(\eta_1, V_1)(\eta_2, V_2) \cdots (\eta_N, V_N)$, i.e., $H^THV_i = \eta_i V_i$. The following properties can then be shown:

1) $\lambda_i \geq 0 \quad i = 1, \ldots, N$

2) $\lambda_i = \eta_i \quad i = 1, \ldots, N$

3) $H = \sum_{i=1}^{N} \lambda_i^{1/2} U_i V_i^T$ where $\lambda_i^{1/2}$ is positive.

The terms $\lambda_i^{1/2}$ are the "singular values" of H and property 3 is the

"singular value decomposition" of H. If the λ_i are ordered in decreasing magnitude, it is possible to approximate H by

$$H_k = \sum_{i=1}^{k} \lambda_i^{1/2} U_i V_i^T$$

It can be shown that the error in approximating H by H_K is

$$||H - H_k||^2 = \sum_{i=k+1}^{N} \lambda_i.$$

Use of SVD for Data Compression A direct representation of the image matrix H requires N^2 data elements, while H can be approximated by H_K using only 2KN elements. The following elementary algorithm for data compression will permit reconstruction to within a required mean squared error, ε.

Step 1 - Compute $||H||^2$, set k = 1, S = 0.

Step 2 - Compute and save λ_k, U_k, V_k.

Step 3 - $S = S + \lambda_k$.

Step 4 - If $||H||^2 - S < \varepsilon$, done. Otherwise step 5.

Step 5 - k = k+1; go to step 2.

This technique has been used on three 120x120 images. All three cases exhibited basically the same results which are summarized here:

(a) When the singular values were plotted, the curve was convex and the singular values dropped off very quickly. For one of the images, the first 10 singular values were 1907, 442, 338, 183, 149, 115, 93, 80, 67, 59. Singular values with indices 20, 30 and 40 were 25, 16 and 12, respectively.

(b) The percent mean squared error in truncation as defined by

$$\text{MSE}(k) = \frac{\sum_{i=k+1}^{120} \lambda_i}{\sum_{i=1}^{120} \lambda_i} \times 100\%$$ dropped off to below 1% at approximately k=15.

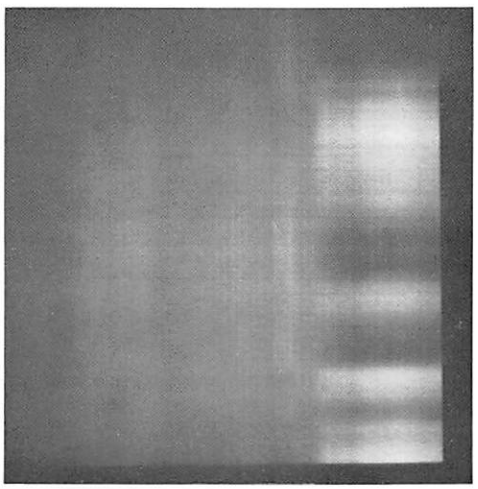
(c) Successive values of H_k were displayed on a high resolution display in sequence. Viewers could barely differentiate H_{15} from H_{20} and the original.

(d) The number of data values using SVD is $2 \times k \times 120 + k$ which for k=15 yields a reduction of 75% and for k=20, a reduction of 67%. It should, however, be noted that the original pixel elements are integers of 6-10 bits. However, the vectors U_i, V_i are made up of real signed numbers less than 1 in absolute value which could possibly require more than 8 to 10 bits for their representation.

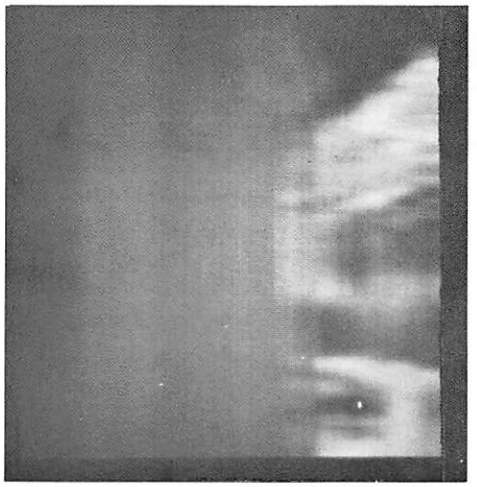
(e) The algorithm to compute the SVD of a matrix will in general compute the λ_i in decreasing order. A "real time" algorithm should have the capability of "handing over" the λ_i, U_i, V_i values one at a time and letting the user decide whether any more terms are required.

(f) Figures 1, 2 and 3 show images and their reconstruction for different values of k. In each case the images have been treated as a 120x120 array. In figs. 1 and 2 note that for k=20 and k=22, respectively, the comparison with the original is quite good. However in fig. 3 it is necessary that k=26 before the reconstruction is good. As expected the plot of $\lambda_k^{\frac{1}{2}}$ vs k drops off more sharply for figs. 1 and 2 than for fig. 3.

For the sample 120x120 images, the singular values had the property of decreasing sharply at first and then after a while decreasing linearly. This can be seen in the table below which gives every 10^{th}



(a) K = 2



(b) K = 6

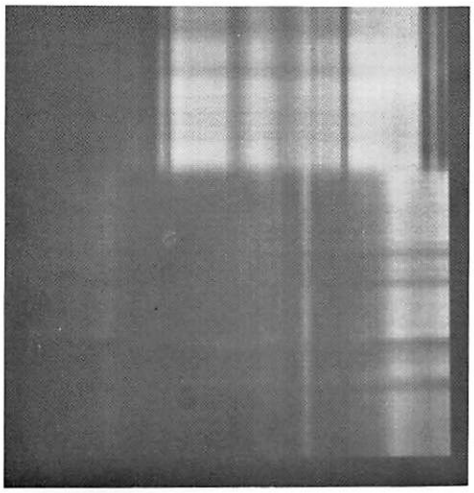
(c) K = 12

(d) K = 16

(e) K = 20

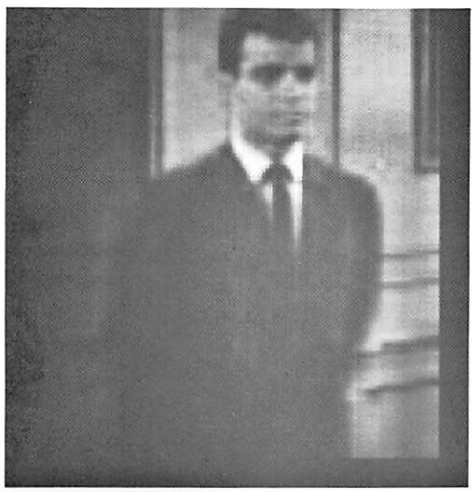
(f) original

Figure 3.6-1. Expansion of an image by singular values - image no. 1.

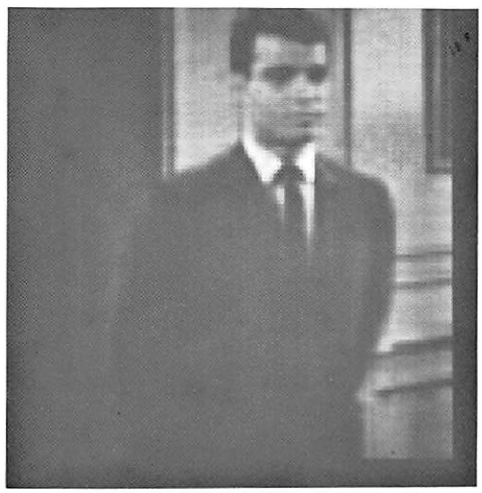


(a) K = 2

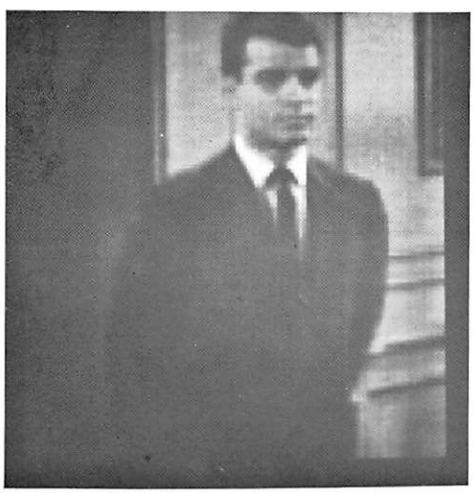
(b) K = 6



(c) K = 12



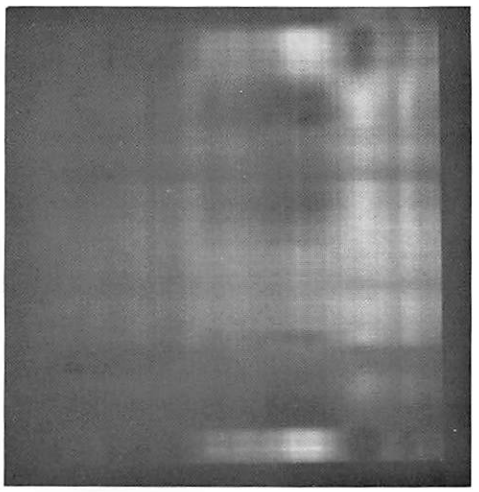
(d) K = 18



(e) K = 22

(f) original

Figure 3.6-2. Expansion of an image by singular values - image no. 2.



(a) K = 2

(b) K = 6

(c) K = 14

(d) K = 20

(e) K = 26

(f) original

Figure 3.6-3. Expansion of an image by its singular values - image no. 3.

singular value for two different pictures.

k	Singular Value Picture 1	Picture 2
1	1272	1625
11	109	96
21	60	52
31	38	34
41	27	21
51	18	16
61	13	11
71	9.3	7.8
81	6.7	5.6
91	4.8	3.6
101	3.0	2.4
111	1.3	1.0

Considering the amount of redundancy in the pictures, one might have expected that singular values would exist which were either zero or small numbers resulting from computer round-off. A tentative conclusion is that the nonzero but small singular values are the result of low level noise in the original images. In these cases, H_k might be a better representation of the "true" object than H. In the image shown in fig. 1, low level noise showed up as "patches" in the relatively flat area on the left side. For certain values of k this noise was not present. Although the above discussion is tenuous, it does lead to the question of the significance of individual terms $\lambda_k^{\frac{1}{2}} U_k V_k^T$ and how they could be treated.

Areas of Further Investigation

(a) Methods of coding the components of U_i and V_i to achieve better data compression.

(b) A method for image enhancement in which the expression

$H_k = \sum_{i=1}^{k} W_i \lambda_i^{\frac{1}{2}} U_i V_i^T$ where the W_i are chosen to bring out certain features.

(c) Examination of the significance of the individual terms $\lambda_i^{\frac{1}{2}} U_i V_i^T$.

4. Image Enhancement and Restoration Projects

Image enhancement techniques have a variety of purposes such as to improve the subjective aspect of images, improve the visibility of particularly interesting features, or to translate the image information into a form more suitable for subsequent machine data extraction. Image restoration attempts to reconstruct images as they would have been formed by ideal imaging systems, using available degraded image values and some apriori knowledge of the physical imaging system and/or image statistics.

The first report analyzes distribution transformations which are used in image enhancement to equalize the gray level histogram of an image. This technique is often applied to increase the perceptibility of features with a very low contrast range.

Many imaging systems can be modeled by a set of linear equations. In that case, image restoration entails the solution of the set of linear equations. One classical mechanism for solving a set of linear equations is the matrix pseudoinverse, whose properties and performance are discussed here. Computational difficulties and possible instabilities are pointed out.

In the following report, a restoration algorithm is studied, which combines matrix and gradient methods to attain numerical efficiency and accuracy. It is hoped that this algorithm will provide a practical numerical tool for a wide variety of nonlinear image restoration problems involving large two dimensional arrays.

Constrained image restoration takes advantage of some apriori knowledge of the image. The use of equality and inequality constrained estimators in the underdetermined and overdetermined case is studied and compared with unconstrained methods.

The resultant quality of image enhancement is limited by the restriction imposed by linearity. A nonlinear statistical image model is introduced, upon which a new recursive image enhancement procedure is developed.

This technique utilizes simultaneous statistical detection and estimation theories and attempts to treat "signal" and "background" in images separately. Simulations show improved enhancement particularly at edge boundaries.

In the following report, the concept of a nonlinear statistical image model is explored further. The problem is to estimate the parameters of the model which determine the location of an object in the image to be restored. The presence of the object in the noisy image data is postulated as apriori knowledge.

A method of image restoration for incoherent optical systems with third order aberrations is presented. The technique uses coordinate transformations to initially operate on the degraded image with a geometrical distortion. Following this, space invariant inverse filtering or estimation and another transformation are used to complete the process. The technique makes restoration practical by effectively reducing the system dimensionality.

The last two reports of this section are concerned with digital color image processing. The reproduction errors in color imaging systems are partly due to the inherent dimensionality reduction process by which the colors of the scene are reduced to three quantities or tristimuli. Errors occur when the spectral sensitivities of the color sensors do not correspond to the human color matching curves. The problem considered is to estimate the original color tristimuli, given the tristimuli of a photographic color image and the characteristics of the film.

When photographing a color CRT display like a TV monitor, the imperfection of color films can be partly compensated for because each color displayed has a unique spectral distribution. When the spectral emissivity of the CRT and the spectral sensitivities of the film are known, a pre-distortion may be designed to obtain the desired colors in the photo. Methods to cope with physical gamut restrictions are also discussed.

4.1 Almost Uniform Distributions for Computer Image Enhancement

Ernest L. Hall

The desirable information in an image often occupies only a small portion of the available contrast range. Thus, certain objects may not be distinguishable because their gray level values are too nearly the same. Computer contrast enhancement techniques are often used to increase the perceptibility of this low detail information. Certain useful contrast enhancement techniques may be based on the histogram or empirical cumulative distribution function, CDF, (integral of the histogram) of the gray level values in the image. For general image enhancement, one may design a contrast transformation which produces an image with a uniform or other desirable gray level distribution. Transformations of this type have been used for image enhancement [1, 2] as well as for an image normalization [3, 4].

The conventional solution to the problem of producing a uniform distribution is to form the empirical CDF $F_{EX}^{(x)}$ and use the $F_{EX}^{(x)}$ transformation. This transformation produces a discrete variable whose empirical CDF, $F_{EY}^{(y)}$, might be expected to be approximately uniform because of the well known distribution transformation. The purpose of this note is to investigate how closely such a discontinuous function can approximate the uniform. The conclusion is stated in the following lemma.

Lemma Let X be a discrete random variable with values

$$x_i = \frac{(b-a)}{N} i + a, \quad i = 0, \ldots N,$$

i.e., $a \leq x_i \leq b$ and distribution function, $F_X(x) = Pr\{X \leq x\}$, which is a jump function with maximum jump equal to ϵ. That is:

$$\max_i \{F_x(x_{i+1}) - F_x(x_i)\} = \epsilon \geq 0, \quad i = 1, 2, \ldots, N.$$

Then the discrete random variable, Y, defined by: $Y = F_X(X)$, has an approximately uniform distribution in the sense that:

$$F_Y(y) = 1, \quad 1 \leq y$$

$$F_Y(y) = 0, \quad y \leq 0$$

and $0 \leq y - F_Y(y) \leq \epsilon, \quad 0 < y < 1.$

Obviously if an absolutely continuous function, F(x), defined on [a, b] is approximated by its sampled values, $F(x_i)$, at N equi-spaced points, the approximation improves, i.e., ϵ approaches zero, as N increases. The point of the lemma is to develop a bound for the approximation for a finite N and ϵ. The finite distribution transform thus changes the range and the spacing of the variable and provides an approximation to the uniform distribution.

To obtain a good approximation, ϵ must be small, which may not be the case for some empirical distributions. However, certain modifications of the discrete transform which are easily implemented on a computer may be used to improve the approximation.

The simplest method for improving the approximation is simply to rescale the transformed variable. The discrete distribution transformation produces a variable

$$y_i = F_{EX}(x_i), \quad i = 0, 1, \ldots, N$$

with range

$$F_{EX}(a) \leq y_i \leq 1.$$

In the previous lemma, the distribution of Y is compared to a uniform [0, 1] distribution. If Y is rescaled and compared to a uniform $[F_{EX}(a,), 1]$ distribution then it is easily shown that

$$|z - F_Y(y)| < \epsilon_2 < \epsilon$$

where z represents the uniform $[F_{EX}(a), 1]$ variable. Quantizing the rescaled variable to equispaced values may or may not improve the approximation but may reduce the number of levels and produce a "false contouring"

effect in pictures.

It may also appear that the number of non-zero histogram values of the original variable, X, determines the effectiveness of the transformation. However, according to the lemma, it is the values of the non-zero histogram values which determine the accuracy of the approximation. A greater number of quantization levels generally tends to produce a smaller ϵ and thus a better approximation. Of course, if the original variable is a constant then increasing the quantization levels does not reduce ϵ.

Another variation which may be considered is to compress two or more x_i values into a single y_i value. This compression would increase ϵ but possibly also increase $F_{EX}(a)$, thus the approximation would not be improved but the range enhancement may be improved.

For many classes of images, the "ideal" distribution of gray levels is a uniform distribution and the above theorem indicates the method for obtaining an approximate uniform distribution with a simple gray level transformation. A uniform distribution of gray levels tends to make equal use of each quantization level and to enhance low detail information because of the range compression. To use this transformation, one must:

(a) compute the histogram of the image gray level values;

(b) compute the empirical distribution function;

(c) use this distribution curve for the gray level transformation;

(d) rescale and quantize the resulting values.

Finally, although this transformation is very effective for enhancing low contrast detail, it does not discriminate between low contrast information and noise.

References

1. E. L. Hall, et. al., "A Survey of Preprocessing and Feature Extraction Techniques for Radiographic Images," IEEE Trans. on Computers, Vol. C-20, No. 9, September, 1971.

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

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

4. R. Haralick and D. Anderson, "Textural-Tone Study with Applications to Digitized Imagery," TR 182-2, Univ. of Kansas, November, 1971.

4.2 Application of the Matrix Pseudoinverse to Image Restoration

William K. Pratt and Nelson D. A. Mascarenhas

In many imaging systems it is possible to model the imaging process by a set of linear equations. Image restoration then entails the solution by the set of linear equations. One classical mechanism for the solution of a set of linear equations is the matrix pseudoinverse. Its properties and performance for image restoration are discussed here.

Sampled Imagery Model A large class of imaging systems may be mathematically described by a linear systems model. Consider an ideal, continuous, infinite extent image function $F(x, y)$ which is subject to some linear spatial degradation (blur) resulting in an observed continuous, infinite extent image function $\tilde{F}(x, y)$. The linear spatial degradation is assumed to be well modelled by the convolution integral

$$\tilde{F}(x, y) = \int_{-\infty}^{\infty}\int F(\alpha,\beta)\, G(x-\alpha, y-\beta)\, d\alpha\, d\beta \qquad (1)$$

where $G(x, y)$ represents the impulse response of a linear spatially invariant system degradation. If the observed image $\tilde{F}(x, y)$ is physically sampled over an $M \times M$ grid of resolution Δ by an ideal* array of Dirac delta functions, the observed image samples are given by

$$\tilde{F}(m_1\Delta, m_2\Delta) = \int_{-\infty}^{\infty}\int F(\alpha,\beta)\, G(m_1\Delta - \alpha,\ m_2\Delta - \beta)\, d\alpha\, d\beta + N(m_1\Delta, m_2\Delta) \qquad (2)$$

$$1 \le m_1,\ m_2 \le M$$

* The effect of finite width sample pulses can be lumped together with the spatial image degradation.

where $N(m_1\Delta, m_2\Delta)$ denotes an additive noise component in the image sensor and sampler. For numerical purposes the continuous integration can be replaced by a summation by numerically sampling the ideal image $F(x, y)$ with an $N \times N$ grid of Dirac samples at a resolution of δ and then invoking quadrature integration. Also, in the discrete representation, it is necessary to truncate the impulse response operator to some spatial limit, say $(L\delta, L\delta)$. Then, the samples image may be described by

$$\tilde{F}(m_1, m_2) = \sum_{n_1 = m_1}^{L+m_1-1} \sum_{n_2 = m_2}^{L+m_2-1} F(n_1\delta, n_2\delta)\, H(m_1\Delta - n_1\delta + L\delta, m_2\Delta - n_2\delta + L\delta) + N(m_1\Delta, m_2\Delta) \tag{3}$$

where the array H, assumed to be zero outside its range of indicies, represents the sampled impulse response and incorporates all quadrature integration factors. In order to prevent serious approximation errors at the boundaries of $\tilde{F}$, N should be chosen such that

$$N\delta \geq M\Delta + (L-1)\delta \tag{4}$$

Usually the numerical sampling resolution δ is kept quite small to limit numerical approximation errors of the continuous integration of eq. (2) by the discrete summation of eq. (4). In such cases N is usually greater than M, that is, there are more representational points in the ideal image, than physical sample points of the observed image.

It is possible to represent the linear eq. (3) in vector space form by defining vectors $\underline{f}$, $\underline{\tilde{f}}$, and $\underline{n}$ which are obtained by column scanning $\underline{F}$, $\underline{\tilde{F}}$, and $\underline{N}$, respectively. Then,

$$\underline{\tilde{f}} = \underline{B}\,\underline{f} + \underline{n} \tag{5}$$

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

$$B = \begin{bmatrix} \underline{B}_{1,1} & \underline{B}_{1,2} & \cdots & \underline{B}_{1,L} & 0 & 0 & & 0 \\ 0 & \underline{B}_{2,2} & \cdots & \underline{B}_{2,L} & \underline{B}_{2,L+1} & \cdots & & 0 \\ \vdots & & & & & & & \vdots \\ 0 & & & & & & & \underline{B}_{M,N} \end{bmatrix} \tag{6}$$

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

$$B^{(m_2, n_2)}_{m_1, n_1} = H(m_1\Delta - n_1\delta + L\delta,\ m_2\Delta - n_2\delta + L\delta) \tag{7}$$

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 Consider a set of linear equations represented by the vector equation

$$\underline{y} = \underline{A}\underline{x} + \underline{e} \tag{8}$$

where $\underline{A}$ is a PxQ matrix and $\underline{e}$ is a Px1 vector representing error or uncertainty in a measurement. A common problem is to determine some vector $\underline{x}_0$ that satisfies the vector equation. In general, eq. (8) may either have: (1) a unique solution; (2) many solutions, or (3) no solution at all.

The pseudoinverse of a PxQ matrix $\underline{A}$ is a QxP matrix $\underline{A}^+$ that satisfies the relation [1]

$$\underline{A}\,\underline{A}^+\underline{A} = \underline{A} \tag{9a}$$

$$\underline{A}^+\underline{A}\,\underline{A}^+ = \underline{A}^+ \tag{9b}$$

Also $\underline{A}\,\underline{A}^{+}$ and $\underline{A}^{+}\underline{A}$ must be symmetric. If $\underline{A}$ is of rank P, then the pseudoinverse can be computed by

$$A^{+} = A^{T}(A\,A^{T})^{-1} \tag{10a}$$

and if A is of rank Q, then

$$A^{+} = (A^{T}A)^{-1}\,A^{T} \tag{10b}$$

The pseudoinverse can be utilized to compute solutions to eq. (8) if such solutions exist, and to compute "best" approximate solutions in the case of nonexistent exact solutions.

For e = 0, it is known [1] that a solution exists for eq. (8) if and only if

$$\underline{A}\,\underline{A}^{+}\underline{y} = \underline{y} \tag{11}$$

Also, if a solution to eq. (8) exists, then the solution is of the form

$$\underline{x}_0 = \underline{A}^{+}\underline{y} + (\underline{I}_P - \underline{A}^{+}\underline{A})v \tag{12}$$

where $\underline{v}$ is an arbitrary P x 1 vector and $\underline{I}_P$ is a P x P identity matrix. If the solution is unique

$$\underline{A}^{+}\underline{A} = \underline{I}_P \tag{13}$$

For an arbitrary error vector $\underline{e}$, the pseudoinverse solution

$$\underline{x}_0 = \underline{A}^{+}y \tag{14}$$

is a best approximate solution in the sense that

$$(Ax-y)^{T}(Ax-y) \geq (Ax_0-y)^{T}(Ax_0-y) \tag{15a}$$

and

$$\underline{x}^{T}\underline{x} > \underline{x}_0^{T}\underline{x}_0 \tag{15b}$$

That is, the pseudoinverse solution is the best solution in the mean square sense and also the lowest energy solution.

Image Restoration Application The pseudoinverse solution may be employed to determine an estimated solution $\hat{\underline{f}}$ to the vector space representation of a sampled image as given by eq. (6). As an example consider a blur matrix which is spatially separable such that the direct product relation

$$\underline{B} = \underline{B}_C \quad \underline{B}_R \tag{17}$$

holds where $\underline{B}_C$ and $\underline{B}_R$ are MxN matrices describing the image blur in horizontal and vertical directions. Furthermore, the blur matrix $\underline{B}_C$ is assumed to be of the form

$$\underline{B}_C = \begin{bmatrix} h_C(L) & h_C(L-1) & \cdots & h_C(1) & 0 & 0 \cdots & 0 \\ 0 & h_C(L) & \cdots & h_C(2) & h_C(1) & \cdots & 0 \\ \vdots & & & & & & \\ 0 & \cdots & \cdots & 0 & h_C(1) & \cdots & h_C(1) \end{bmatrix} \tag{18}$$

where

$$h_C(j) = K \exp\{-[j-(L-1)/2]^2/S_c\}$$

with K and S_c constants. The blur matrix $\underline{B}_R$ is of similar form. The noise is assumed to be independent of the image and to be white with covariance matrix $\underline{K}_n = \sigma^2\underline{I}$. It should be observed that with this direct product structure of the blur matrix, the computational task of performing the pseudoinverse can be considerably reduced since

$$\underline{B}^+ = \underline{B}_C^+ \quad \underline{B}_R^+ \tag{19}$$

Several computer simulation experiments have been performed to evaluate the pseudoinverse restoration technique for the imaging model

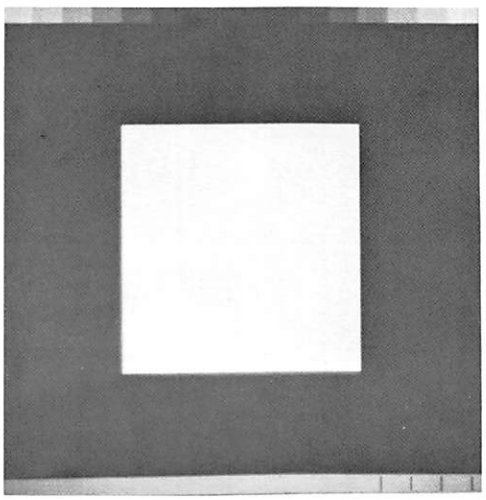
described above. In these experiments the original 8 x 8 pixel image of fig. 1 consisting of a bright square of value 245 on a constant background of value 10 over a 0 to 255 scale has been subjected to linear blur and additive noise by computer simulation. The noise free restoration for various degrees of blur is illustrated in fig. 2. The results indicate that the restoration is subjectively better than the blurred image but the original picture is not necessarily obtained. This is due to the fact the pseudoinverse has produced one of an infinite number of possible restorations, namely, the minimum norm solution.

The experiments have been repeated with a small amount of noise added to the observation, as shown in fig. 3. In this figure restoration values outside the range 0-255 have been clipped. It can be seen that the small noise perturbation can cause a severe change in the result, as is the case with the blur coefficient equal to 5. This can be explained in terms of the ratio of the largest to the smallest nonzero eigenvalue of the matrix ($B^T B$). Since the non zero eigenvalues of $B^T B$ are equal to the nonzero eigenvalues of BB^T and since B^T is assumed to have full column rank, BB^T is nonsingular and that ratio is the condition number of BB^T. In fig. 4 this number is plotted against the blur coefficient. This result implies, quite interestingly, that the worst results in pseudoinverse restoration should be expected with moderate amounts of blur, as compared with small or large values.

Conclusions The pseudoinverse technique has been studied for image restoration in the hope that this method for solving linear equations would provide good quality image restoration. The results of the study indicate several shortcomings of the method. First, computational problems occur when attempting to compute the pseudoinverse, especially for large size blocks. Secondly, observation noise can lead to numerical instability of the solution at moderate values of blur.

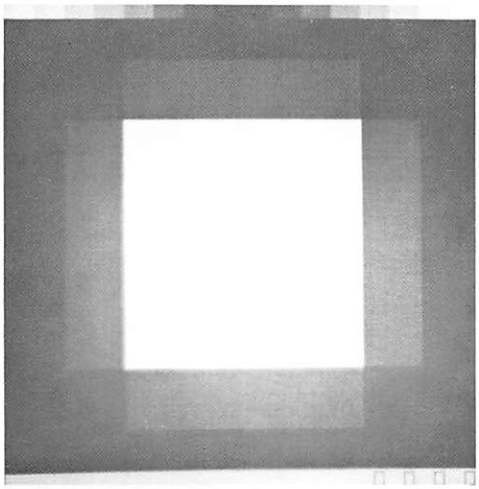
References

1. F. A. Graybell, Introduction to Matrices with Applications in Statistics, Wadsworth Publishing Co., Belmont, Calif., 1969.



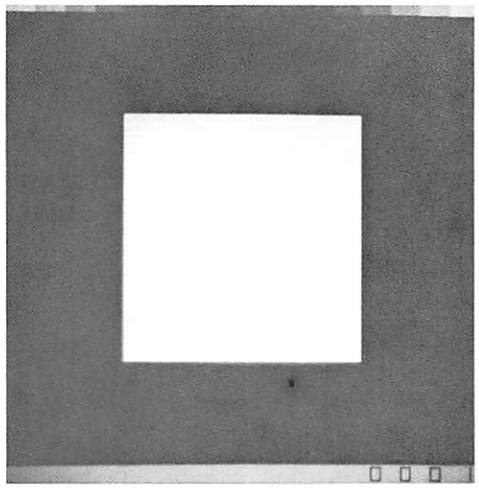
original

Figure 4.2-1. Original image for pseudoinverse restoration examples.

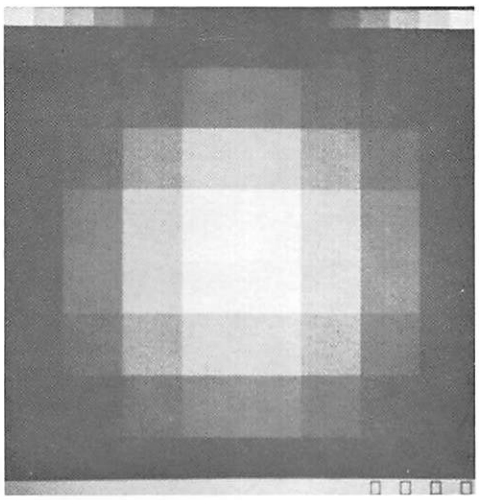


(a) observation

$S_c = S_R = 0.5$ $\sigma^2 = 0$

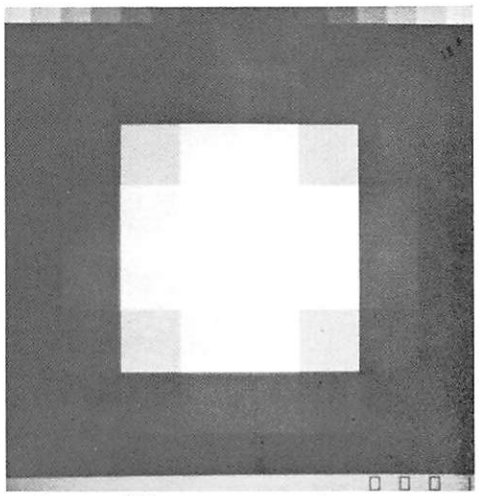


(b) restoration

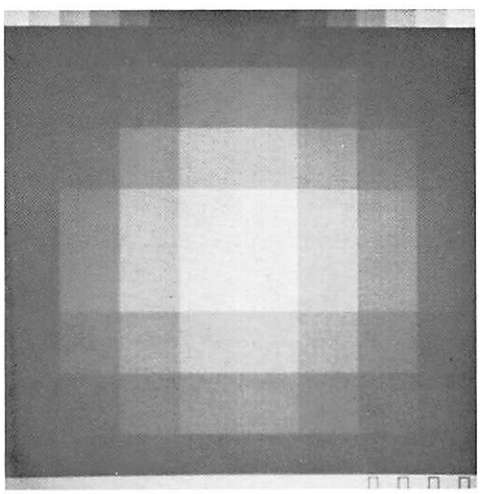


(c) observation

$S_c = S_R = 5$ $\sigma^2 = 0$

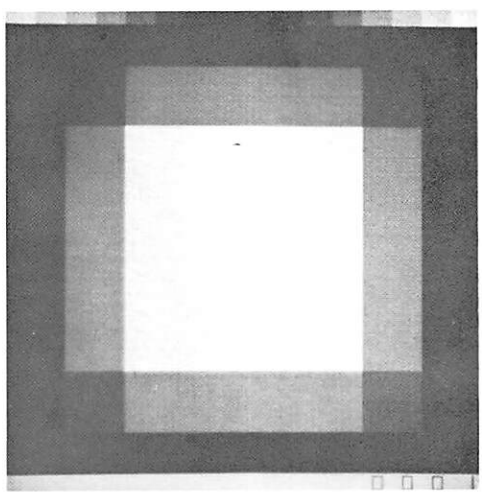


(d) restoration



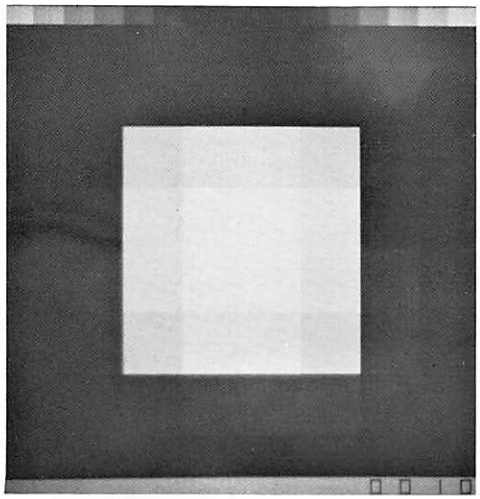
(e) observation

$S_c = S_R = 500$ $\sigma^2 = 0$



(f) restoration

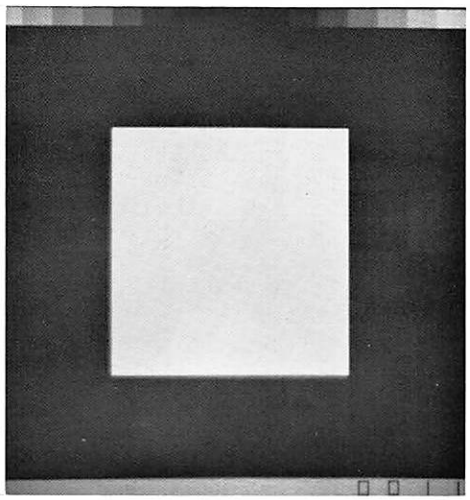
Figure 4.2-2. Pseudoinverse restoration or noise free observation.



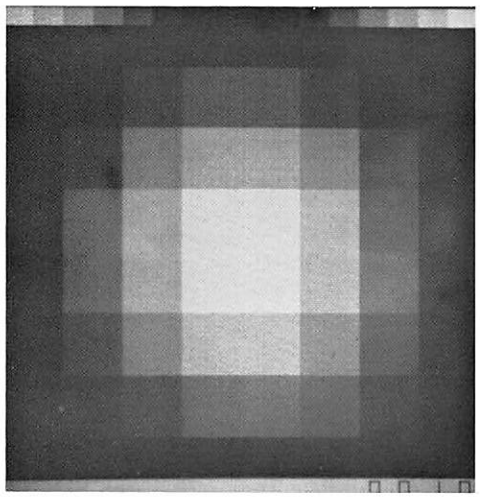
(a) observation

$S_c = S_R = 0.5$

$\sigma^2 = 10$



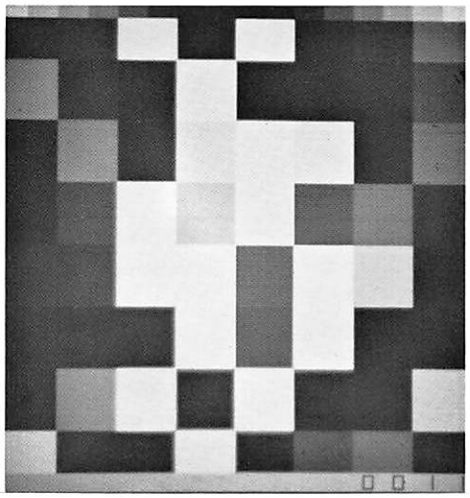
(b) restoration



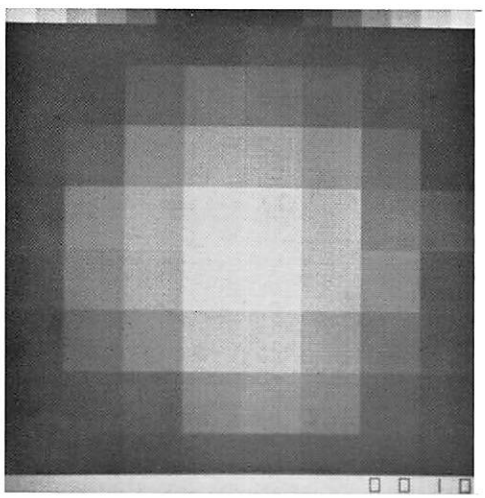
(c) observation

$S_c = S_R = 5$

$\sigma^2 = 10$



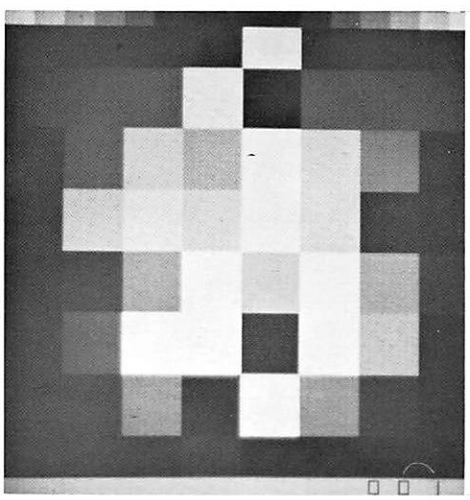
(d) restoration



(e) observation

$S_c = S_R = 500$

$\sigma^2 = 10$



(f) restoration

Figure 4.2-3. Pseudoinverse restoration for noisy observation.

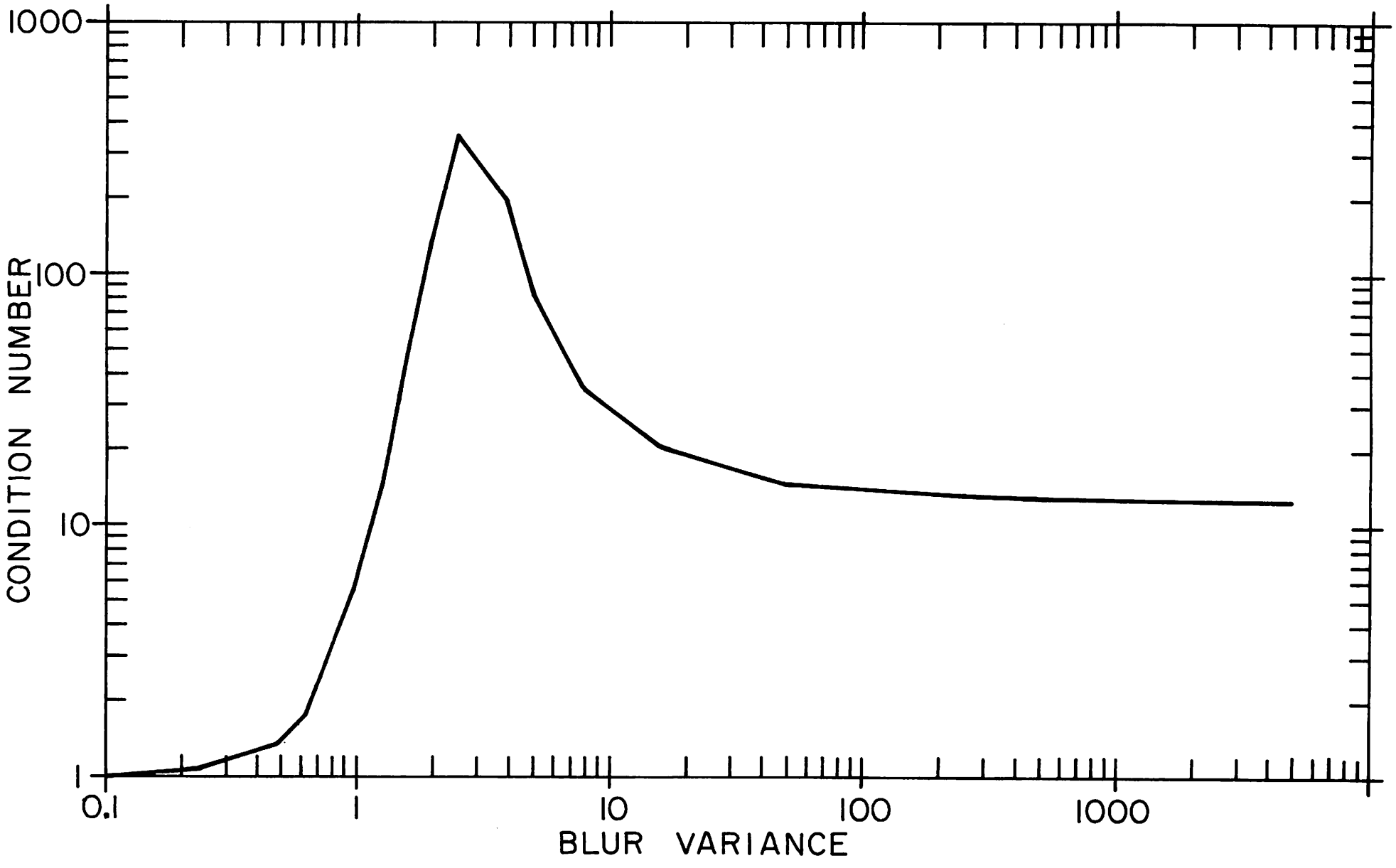


Figure 4.2-4. Condition number as a function of Gaussian blur variance.

4.3 Positive Restoration by Deconvolution

Ronald L. Hershel

Whenever an object scene $\vec{o}$ is degraded by an imaging system (characterized by a linear transform S) and additive noise $\vec{n}$, the image $\vec{r}$ may be expressed as

$$r_i = \Sigma S_{ij} o_j + n_i, \qquad i = 1, 2, \ldots, N. \tag{1}$$

With certain statistical assumptions which include a positive constraint on $\vec{o}$, one may estimate both the object and noise by the following restoration formulas [1]

$$o_i^* = \exp\left(\sum_j S_{ij} a_j^*\right) \tag{2}$$

$$n_i^* = \beta a_i^* \qquad \beta = \sigma_n^2/\sigma_0^2 \tag{3}$$

where the unknown parameters a_i^* $(i = 1, 2, \ldots, N)$ are determined through satisfying eq. (1). The expressions in eq. (2) ensure a positive estimate o^* while no such constraint appears on the noise estimate (assumed uncorrelated with zero mean).

The main effort in this research has been to develop a suitable numerical algorithm for solving the simultaneous nonlinear equations associated with eqs. (1) and (2). Due to an explicit use of an NxN matrix, standard procedures such as Newton-Raphson [2], secant [3], and variable metric [4] suffer from a limitation on the number of unknowns that can be effectively handled ($N < 100$). Other methods which do not require matrix handling (such as the steepest descent [5], conjugate gradients [6], and the generate Partan [7]) either converge too slowly (lacking quadratic convergence) or break down due to accumulated round off error.

Mathematical Considerations Consider the scalar function

$$f(\vec{a}) = \sum_i^M \exp\left(\sum_j^N S_{ij} a_j\right) + \frac{B}{2} \sum_i^N a_i^2 - \sum_i^N a_i r_i$$

where $M > N$ to accommodate the higher resolution associated with positive restorations [1]. This equation may be expressed in vector notation as

$$f(\vec{a}) = \vec{u}^T \exp(S\vec{a}) + \beta/2\, \vec{a}^T\vec{a} - \vec{a}^T\vec{r}; \qquad u_i = 1,\ i = 1, 2, \ldots, M \tag{4a}$$

For small deviations $\vec{\Delta}$, $f(\vec{a})$ may be expanded to second order

$$f(\vec{a}+\vec{\Delta}) = f(\vec{a}) + \vec{\Delta}^T\vec{g} + \tfrac{1}{2}\vec{\Delta}^T J\vec{\Delta} \tag{4b}$$

where the gradient $\vec{g}$ and Jacobian J are defined as

$$g_i = \frac{\partial f}{\partial a_i} \qquad J_{ij} = \frac{\partial^2 f}{\partial a_i \partial a_j} \qquad i, j = 1, 2, \ldots, N$$

By eq. (4), one may write

$$\vec{g} = S\vec{o} + \vec{n} - \vec{r} \tag{5}$$

$$J = S\theta S^T + \beta \tag{6}$$

where the matrix θ is defined as

$$\theta_{ij} = \left\{ \begin{array}{ll} 0 & i \neq j \\ o_i & i = j \end{array} \right| \quad i = 1, 2, \ldots, M$$

with the object values o_i determined by eq. (2). Since $o_i(\vec{a}) \geq 0$, it is apparent from eq. (6) that J is positive-definite symmetric for all real values of $\vec{a}$. Hence $f(\vec{a})$ displays global convexity and possesses a unique minimum $f(\vec{a}^*)$. This minimum coincides with a zero gradient ($\vec{g}^* = 0$) which by eq. (5) is written as

$$S\vec{o}^* + \vec{n}^* - \vec{r} = 0$$

Hence finding the minimum of $f(\vec{a})$ is equivalent to solving the positive restoration problem expressed by eqs. (1)-(3). With these considerations an algorithm can be developed which incorporates the salient features of

both the conjugate gradient and matrix methods in an attempt to minimize the overall computational requirements.

Mixed Algorithm for Rapid Convergence The initial estimates $\vec{o}^{(o)} = \vec{u}$ and $\vec{n}^{(o)} = 0$ are used to begin the iterative process. These estimates correspond to $\vec{a}^{(o)} = 0$ as seen by eqs.(2)-(3). Hence, from eq.(3), the initial value of $f(\vec{a})$ is

$$f(o) = \vec{u}^T\vec{u} = M$$

which is simply the object sample size. Denoting $\vec{v} = S\vec{u}$, eq.(5) gives

$$\vec{g}_1 = \vec{r} - \vec{v}$$

from which the following quantities are defined

$$c_1 = |\vec{g}_1|^2 \tag{7a}$$

$$d_1 = \vec{r}^T\vec{g}_1 \tag{7b}$$

$$\vec{h}_1 = S^T\vec{g}_1 \tag{7c}$$

The next estimate for $\vec{a}$ is along the direction of steepest descent

$$\vec{a} = b_1\vec{g}_1$$

whereby one may write $f(\vec{a})$ as a function of the unknown scalar b_1,

$$f(b_1) = \vec{u}^T \exp(b_1\vec{h}_1) + \beta/2\, c_1 b_1^2 - d_1 b_1$$

and proceed to find the $b_1^{(1)}$ that minimizes $f(b_1)$. Denoting $\vec{a}^{(1)} = b_1^{(1)}\vec{g}_1$ and $\vec{g}_2 = \vec{g}(\vec{a}^{(1)})$ the process is continued as follows. For the k^{th} iteration:

(1) Find $c_k = |\vec{g}_k|^2$, $d_k = \vec{r}^T\vec{g}_k$ and $h_k = S^T\vec{g}_k$.

(2) Let $\vec{a} = \sum_{i=1}^{k} b_i\vec{g}_i$ and find those values $b_i = b_i^{(k)}$ $(i = 1, 2, \ldots, k)$ which minimize $f(\vec{a})$.

(3) Set $$\vec{a}^{(k)} = \sum_{i=1}^{k} b_i^{(k)} \vec{g}_i$$

$$\vec{o}^{(k)} = \exp\left(\sum_{i=1}^{k} b_i^{(k)} \vec{h}_i\right)$$

$$\vec{n}^{(k)} = \beta \vec{a}^{(k)}$$

(4) Find $$\vec{g}_{k+1} = S\vec{o}^{(k)} + \vec{n}^{(k)} - \vec{r}$$

(5) Set k = k+1 and return to step 1.

One important feature of the above algorithm is that only two S transforms are required for each iteration. Step (2) is performed efficiently using the Fletcher-Powell algorithm [8] which is particularly well suited to the form of $f(\vec{a})$. Furthermore, Step (2) ensures orthogonality of the $\vec{g}$ vectors in succession

$$\vec{g}^T \vec{g}_{k+1} = 0 \qquad i = 1, 2, \ldots, k$$

and in effect reduces the hyperspace of the error $\vec{a}^* - \vec{a}^{(k)}$ by one dimension. Hence complete convergence is assured after k = N iterations where the $\vec{g}_i$, i = 1, 2, . . . , N, completely span the space of $\vec{a}^*$. As to the behavior of the incomplete solutions $\vec{a}^{(k)}$ (k < N), preliminary tests indicate rapid convergence to the near quadratic region of $f(\vec{a})$. Once in this region, the algorithm is entirely equivalent to the conjugate gradient method [6].

Modification for Restricted Storage The k^{th} iteration normally requires k linearly independent vectors to minimize $f(\vec{a})$ over a k dimensioned subspace. However in the vicinity of $f(\vec{a}^*)$, the quadratic approximation of eq. (4a) becomes valid and only a single conjugate direction is required for each iteration [6]. This property suggests a procedure which retains less than k conjugate directions. Assume, for example, that there is a restriction to five direction vectors for use in Step (2) due to storage limitations. Then as each new gradient $\vec{g}_{k+1}$ is found, the $\vec{g}_{k-5}$ is eliminated

after J-orthogonalization with $\vec{g}_{k-4}, \vec{g}_{k-3}, \ldots, \vec{g}_{k+1}$. Each new vector becomes

$$\vec{g}_i^* = \vec{g}_i + \alpha_i \vec{g}_{k-5}$$

where α_i is found by requiring

$$\vec{g}_{k-5}^T J^{(k)} \vec{g}_i^* = 0, \qquad i = k-4, k-3, \ldots, k+1 \tag{8}$$

Since $\vec{g}_i^T J \vec{g}_k$ is equivalent to $\vec{h}_i^T \theta \vec{h}_j$, the calculations envolved in eq. (8) are relatively fast.

Conclusions A restoration algorithm has been presented which combines matrix and gradient methods to attain numerical efficiency and accuracy. Though for any given problem the convergence rate is dependent on the imaging transform, object structure and noise level, it is hoped that this algorithm will provide a practical numerical tool for a wide variety of nonlinear image restoration problems involving large two dimensional arrays.

References

1. R. Hershel, "Unified Approach to Restoring Degraded Images in the Presence of Noise," Tech. Report 72, Optical Sciences Center, Tucson, Arizona, December, 1971.

2. A. S. Householder, Principles of Numerical Analysis, McGraw-Hill, New York, 1953.

3. J. G. P. Barnes, "An Algorithm for Solving Nonlinear Equations Based on the Secant Method," Computer Journal, Vol. 8, 1964, pp. 66-72.

4. R. Fletcher, "A New Approach to Variable Metric Algorithms," Computer Journal, Vol. 13, No. 3, 1970, pp. 317-322.

5. H. D. Curry, "The Method of Steepest Descent for Nonlinear Minimization Problems," Qu. App. Math., Vol. 2, 1944, p. 258.

6. M. R. Hestenes and E. Stiefel, "Methods of Conjugate Gradients for Solving Linear Systems," J. Res. N. B. S., Vol. 49, 1952, pp. 409-438.

7. B. V. Shah, R. J. Buehler, and O. Kempthorne, "Some Algorithms for Minimizing a Function of Several Variables," J. Soc. Indust. Appl. Math., Vol. 12, No. 1, 1964, pp. 74-92.

8. R. Fletcher and M. J. D. Powell, "A Rapidly Convergent Descent Method for Minimization," Computer Journal, Vol. 8, 1964, pp. 163-168.

4.4 Constrained Restoration of Digital Pictures

Nelson D. A. Mascarenhas

The problem of restoring digital pictures degraded by blur and corrupted by noise can be put into the framework of a linear regression model

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

where $\underline{y}$ is a (mxl) vector of observed values, $\underline{B}$ is the (mxn) blur matrix, $\underline{x}$ is a (nxl) vector of original pixel values, $\underline{z}$ is the random (mxl) vector representing a noise process. A considerable improvement in the restoration can be obtained if the experimenter incorporates any prior knowledge about the possible values for the parameter vector $\underline{x}$. In this section two kinds of constraints will be considered: equality and inequality constraints.

Equality Constraints A priori knowledge of linear relations involving the original pixel values can be expressed by

$$\underline{G}\,\underline{x} = \underline{u}$$

where $\underline{G}$ is an kxn matrix and $\underline{u}$ is an kxl vector in the range of $\underline{G}$.

Assuming white noise, under the least squares criterion, one searches for the minimization of $\|\underline{y} - \underline{B}\underline{x}\|^2$ over the set $\{\underline{x} : \underline{G}\,\underline{x} = \underline{u}\}$. The solution of this problem may not be unique if $\underline{B}$ does not possess full column rank, as is the case in an undetermined system, i.e. when there are more parameters to be estimated than observed values. The minimum norm solution, however, is unique and it is given by [1]:

$$\hat{\underline{x}} = \underline{G}^{+}\underline{u} + \overline{\underline{B}}^{+}\overline{\underline{y}}$$

where + denotes the pseudoinverse and

$$\overline{\underline{B}} = \underline{B}(\underline{I} - \underline{G}^{+}\underline{G})$$

$$\overline{\underline{y}} = \underline{y} - \underline{B}\,\underline{G}^{+}\underline{u}.$$

The equality constrained method has been applied to the restoration of digital pictures. Figure 1a shows the center part (8x8) of the original (12x12) picture. In this case the system is underdetermined. Figure 1b represents the (8x8) observed values blurred by a Gaussian shaped transfer function and corrupted by white noise. When the constraint that the sum of the (8x8) pixel values of the center part of the restored picture be equal to the correspondent sum in the input picture, the result is shown in fig. 1c. It can be seen that an improvement over the unconstrained case as shown in fig. 1d occurs.

This improvement can be made explicit by the reduction on the covariance matrix of the estimated values. In the case of the overdetermined system, the covariance matrix of the equality constrained estimator is given by [2]

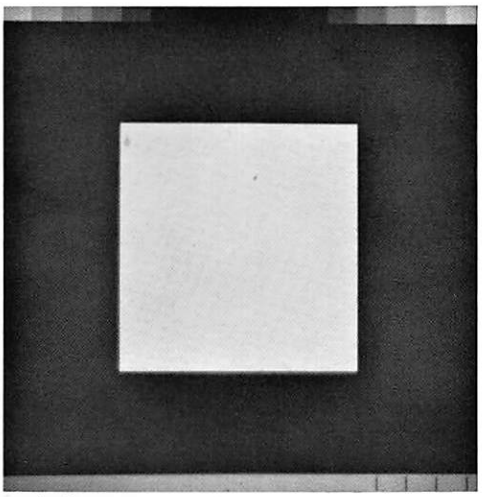
$$\mathrm{Cov}[\hat{\underline{x}}] = (\underline{B}^T\underline{B})^{-1} - (\underline{B}^T\underline{B})^{-1}\underline{G}^T[\underline{G}(\underline{B}^T\underline{B})\underline{G}^T]^{-1}\underline{G}(\underline{B}^T\underline{B})^{-1}$$

Since $\underline{B}^T\underline{B}$ is the covariance matrix of the unconstrained estimator, the reduction is clear.

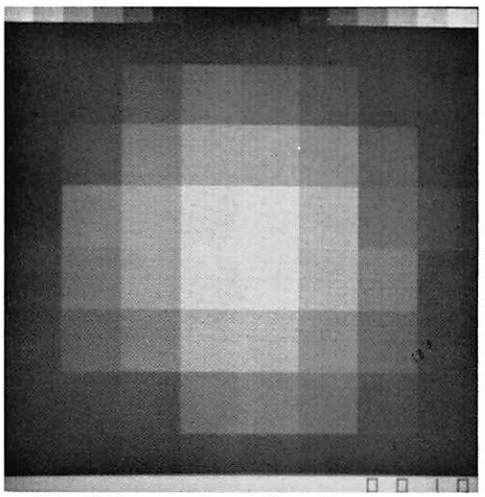
The imposition of linear constraints may imply bias, however, when the restrictions are not true [3]. Even in this case, there may be an improvement in the mean square sense, taking into account both bias and variance.

Linear constraints allow also for the calculation of confidence intervals for pixel values. Figure 2 shows the curves of the residual norm $\|\underline{y} - \underline{B}\hat{\underline{x}}\|$ in the one-dimensional overdetermined system when the fourth estimated pixel value is constrained to vary over a certain range. The same figure shows the confidence interval for the estimation of the fourth pixel value under 95% confidence level.

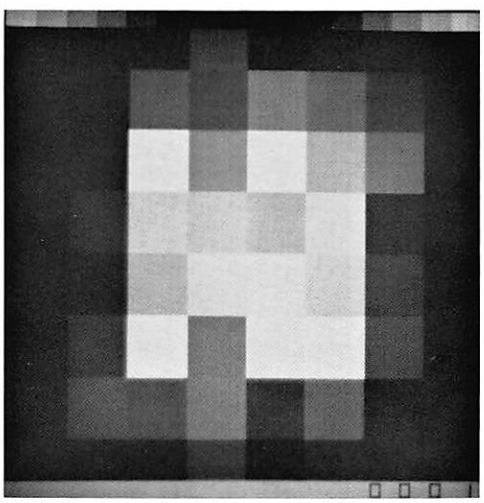
Inequality Constraints The fact that the original pixel values are nonnegative quantities can be incorporated in the model, leading to improved restoration procedures. A prior knowledge of upper bounds or any other inequality restrictions can also be used.



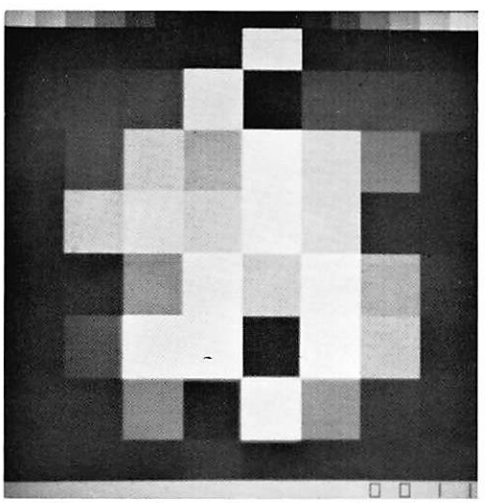
(a) original



(b) blurred and noisy
B_L = 500
VAR = 10
Gaussian blur



(c) restored, equality constrained



(d) restored, unconstrained

Figure 4.4-1. Comparison of unconstrained and equality constrained restorations.

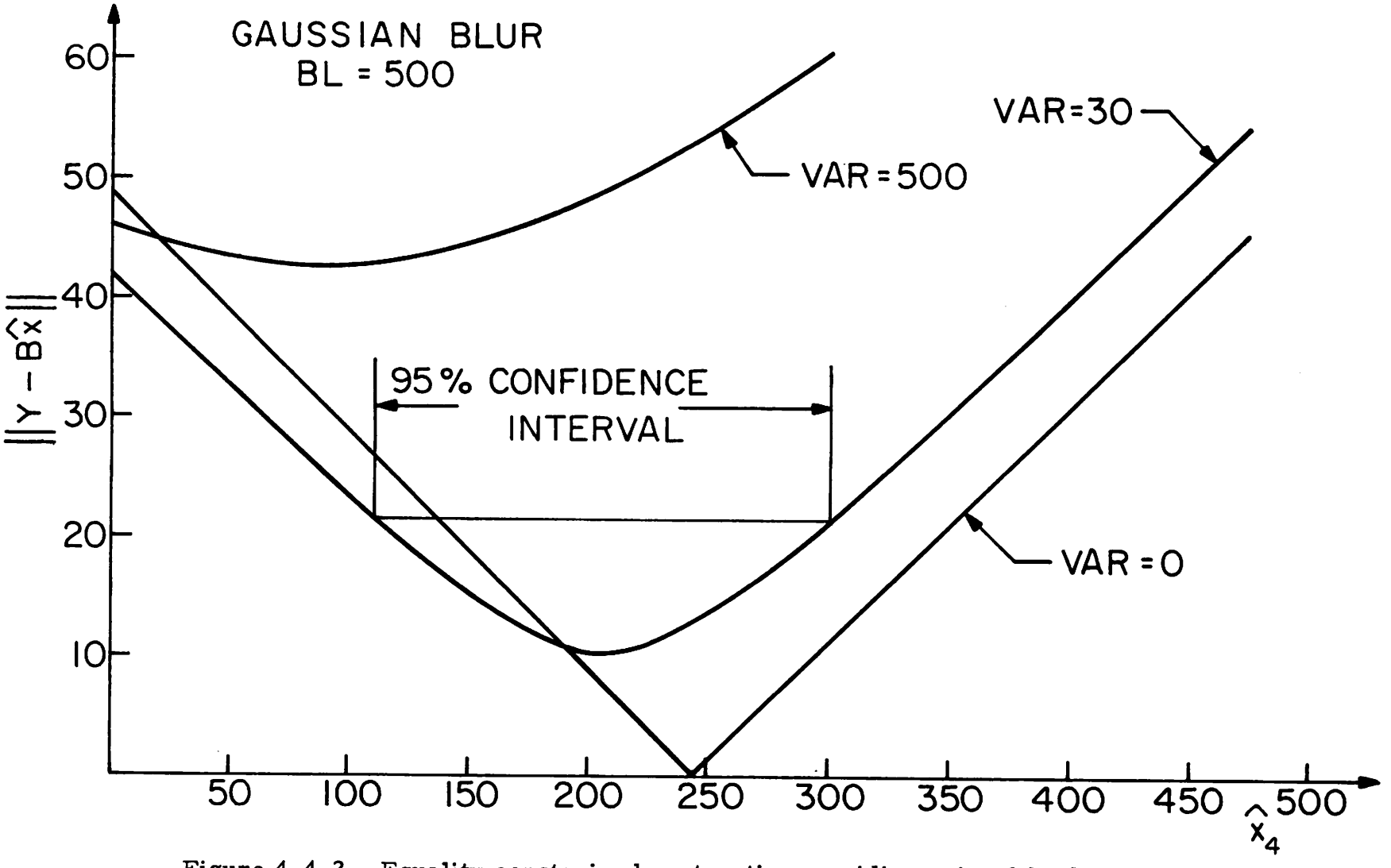


Figure 4.4-2. Equality constrained restoration - unidimensional M=8, N=12.

Under the least squares criterion a quadratic programming problem results. In the overdetermined case the matrix of the quadratic expression ($\underline{B}^T\underline{B}$) is positive definite while in the underdetermined case it is positive semidefinite.

The main advantages in the use of inequality constrained restoration are the following: (1) In the underdetermined case there are parametric functions of the pixel values that are inestimable, i.e., their confidence interval is infinite, in the unconstrained method. This drawback can be removed by bounding the range of estimated values in the n-dimensional space. (2) There is a reduction in confidence interval as the result of the use of additional prior statistical information. Figure 3 shows the improvement in the 95% confidence interval by the use of the constraints, $0 \leq x \leq 255$ on the one-dimensional overdetermined model. The quadratic programming problem was solved by the use of Dantzig's version of the quadratic programming algorithm originally developed by Wolfe [4]. (3) The very important effect of ill conditioning can be improved. This occurs by the reduction on the size of the elongated confidence ellipsoids of the ill conditioned problem, due to the inequality restriction. Even if an unconstrained solution is attempted followed by clipping the values to their bounds, this implies the nonoptimal and sometimes disastrous procedure of simply projecting the unconstrained solution on the boundaries. The constrained restoration, however, obtains the optimal solution within the restrictions. A disadvantage of the procedure is that it imposes bias on the solution. This can be minimized, however, by imposing both lower and upper bounds.

Computational problems arise in image processing due to the large dimensionality that may be involved. In an attempt to minimize this problem a sequential restoration of rows and columns was performed in the two dimensional case. The result tends to be better when there is little blur. This occurs because on one hand a moderate blur implies large condition number and the input to the column restoration will be clipped. This tends to invalidate the model of additive white noise that is the basis of the quadratic programming algorithm. On the other hand, for large blur, the

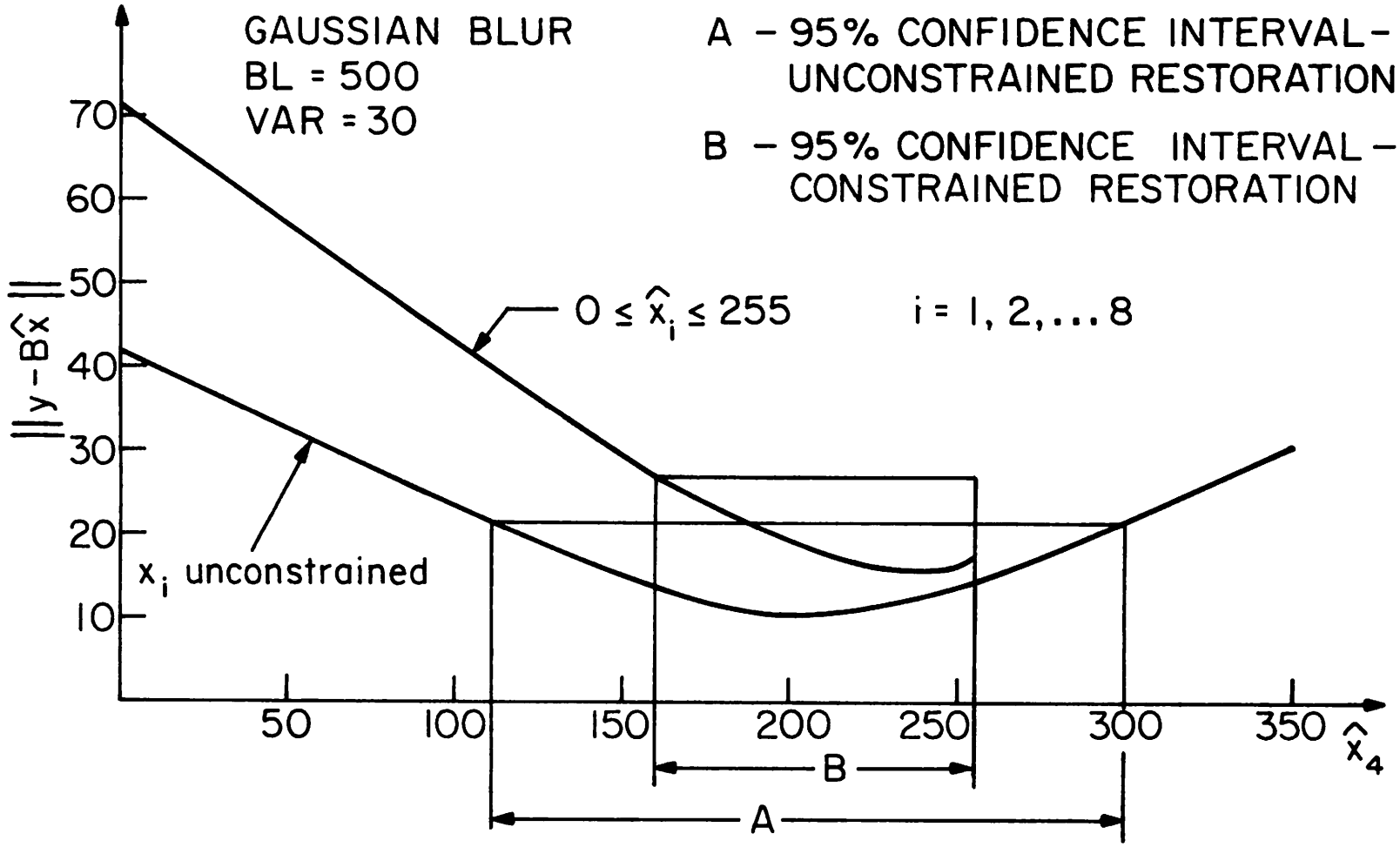


Figure 4.4-3. Comparison between unconstrained and inequality constrained restorations.

condition number is moderate, but now the random variables in the same row will tend to be correlated and this is not taken into account in the column restoration. Figures 4a to 4d show the comparison of unconstrained and constrained restoration under large condition number. In this case the blur function is $\left(\frac{\sin x}{x}\right)^2$, simulating the effect of diffraction. The unconstrained solution was obtained by clipping at 0 and 255. It can be observed that a substantial improvement occurs with the use of the constrained method, transforming an unfeasible restoration into a feasible one.

References

1. A. Albert, "Regression and the Moore-Penrose Pseudoinverse," Academic Press, New York, 1972.

2. T. O. Lewis and P. L. Odell, "Estimation in Linear Models," Prentice Hall, New Jersey, 1971.

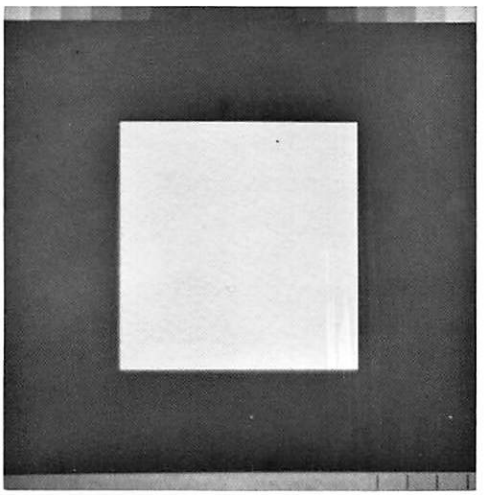
3. D. S. Huang, "Regression and Econometric Methods," John Wiley and Sons, New York, 1970.

4. L. Cutler and D. S. Pass, "A Computer Program for Quadratic Mathematical Models to be used for Aircraft Design and other Applications Involving Linear Constraints," Rand Corp., Santa Monica, 1971.

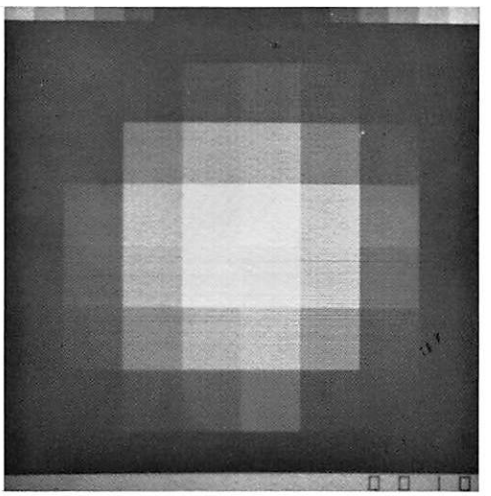
4.5 Nonlinear Adaptive Recursive Image Enhancement

Nasser E. Nahi and Ali Habibi

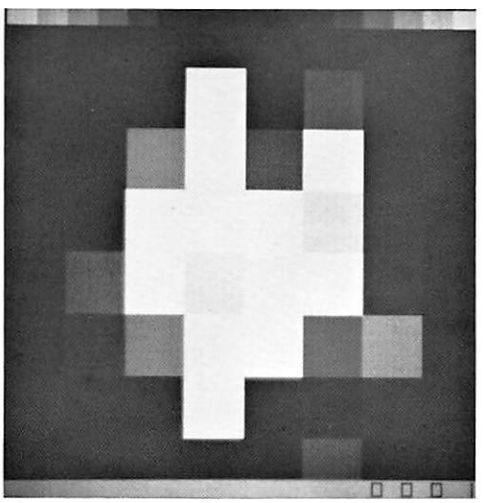
Various linear recursive algorithms for digital image enhancement have been introduced in the literature [1-3]. While these algorithms represent efficient methods for performing the task of linear image enhancement, the resultant quality of enhancement is limited due to the restriction imposed by linearity. A nonlinear statistical model of an image is introduced which characterizes more faithfully the discontinuities existing in typical pictures due to the boundaries of various texture segments. Based on this model a new recursive image enhancement procedure is developed utilizing simultaneously the concepts of statistical detection and estimation theories. Hence, the algorithm is called detection-directed image enhancement procedure.



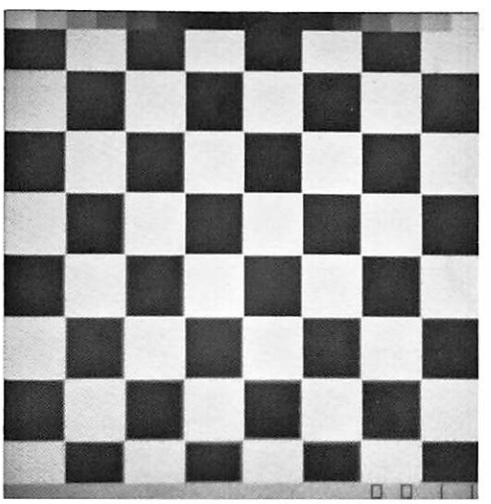
(a) original



(b) blurred and noisy
$B_L = 10$
VAR = 40
$sinc^2$ blur



(c) restored, inequality constrained



(d) restored, unconstrained

Figure 4.4-4. Comparison of unconstrained and inequality constrained restorations.

A Nonlinear Statistical Model of Image The linear enhancement filters (estimators) described [1-3] reflect a continuity characteristic in that the brightness level at any point on the enhanced picture is a function of brightness levels of all points of the original picture. This in effect will introduce a global smoothing over the entire picture causing elimination of sharp boundaries which usually signifies various textured segments.

Consider a picture composed of an object with one texture in a background; extension to multi-textured pictures is straightforward. Let $b_0(m, n)$ and $b_b(m, n)$ denote random processes describing the object and background brightness levels, respectively. Consequently, $b_0(m, n)$ and $b_b(m, n)$ will assume the statistics of object and background textures. A picture can now be thought of as a random process $b(m, n)$ where

$$b(m, n) = \begin{cases} b_b(m, n) & \text{within the background segment} \\ b_0(m, n) & \text{within the object segment} \end{cases}$$

Equivalently

$$b(m, n) = \alpha(m, n)\, b_b(m, n) + [1 - \alpha(m, n)]\, b_0(m, n)$$

where

$$\alpha(m, n) = \begin{cases} 1 & \text{within the background segment} \\ 0 & \text{within the object segment} \end{cases}$$

In other words, the picture (in the absence of noise) represents a replacement process of a section of random process $b_b(m, n)$ by a section of the random process $b_0(m, n)$ where generally the binary replacement variable is unknown since the location of the object within the picture is not known a priori. This model for $b(m, n)$ is a nonlinear statistical model where its sample functions possess discontinuities reflecting the boundaries of textured segments.

Edge Detection in Presence of Noise In a typical problem $\alpha(m, n)$ is not known apriori and must be estimated using the noisy observation

y(m, n). Two methods of estimating $\alpha(m, n)$ are suggested. The first method is one dimensional and operates on the scanned lines of the observed data y(m, n), while the second method uses a two-dimensional estimator directly.

Let the first and the second order statistics of the image and the noise be denoted as follows:

$$E\{v(m, n)\} = 0$$

$$E\{v(m, n)\, v(m+i, n+j)\} = \sigma_n^2 \delta(i, j)$$

$$E\{b_b(m, n)\} = b_0$$

$$E\{(b_b(m, n)-b_0)(b_b(m+i, n+j)-b_0)\} = \sigma_b^2 R_b(i, j)$$

$$E\{b_0(m, n)\} = S_0$$

$$E\{(b_0(m, n)-S_0)(b_0(m+i, n+j)-S_0\} = \sigma_s^2 R_s(i, j)$$

Now let H_0 and H_1 indicate the hypotheses that no edge has occurred at the (m, n)th pixel and an edge has occurred on the (m, n)th pixel, respectively. Appropriate likelihood ratios are then formed to test hypothesis H_0 versus H_1 by considering the joint probability density functions of y(m, n) and a number of adjacent pixels. Naturally due to the spatial nature of the correlation of the picture the adjacent pixels should be those surrounding the (m, n)th pixel in various spatial directions. This gives a two dimensional edge detection technique which is more appropriate to be used with the two dimensional Bayesian estimator [3]. For a one dimensional estimator [1, 2], a similar approach is used to develop a one dimensional edge detection technique. For details on both techniques see [4].

Decision-Directed Image Enhancement The block diagram of this decision directed estimator is shown on fig. 1. Each pixel is first examined to establish if it belongs to the signal or the background, i. e. $\alpha(m, n) = 0$ or $\alpha(m, n) = 1$. Then it is directed to one of two Kalman filters F_b or F_s

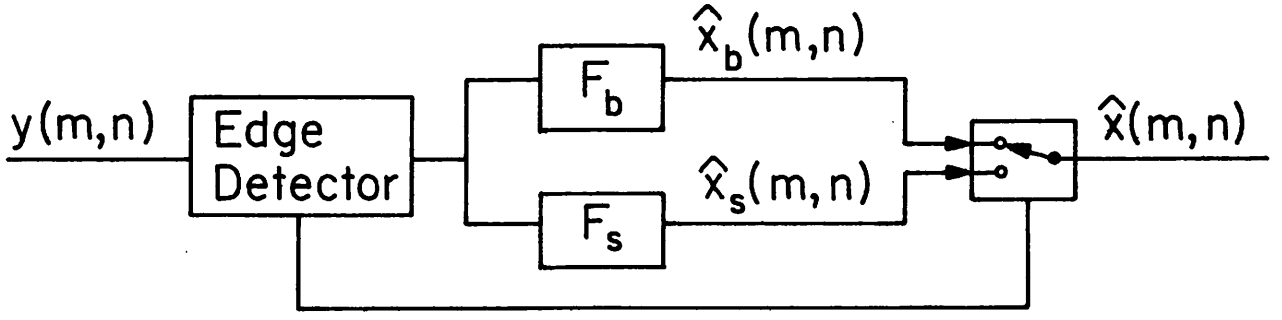
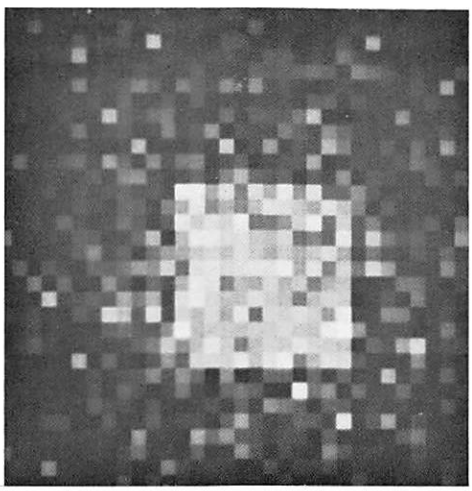


Figure 4.5-1. Block diagram of the nonlinear estimator.

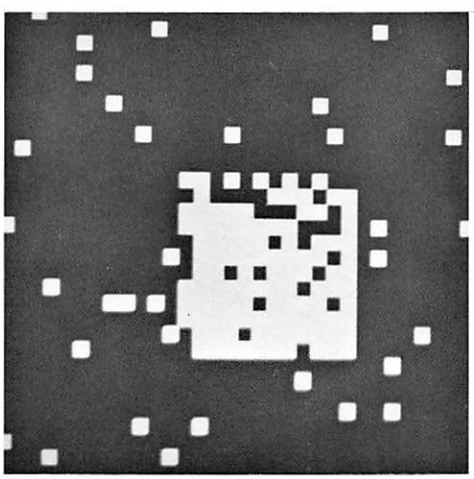
respectively. These filters are designed based upon the statistics of the signal or the background. Naturally for any (m, n)th pixel only one of these filters is active. The idle filter retains the values of the parameters most recently processed and uses them as the initial values for proceeding with the estimation of the next pixel when it would switch to an active mode. In the two dimensional estimator in addition to the presence or absence of an edge at (m, n)th pixel the angle of the edge is also indicated to the filters so that a proper updating of the initial values is performed. Design and operation of the one and two dimensional Kalman filters are discussed in references [1-3].

Experimental Results To evaluate the performance of the edge detection technique a grid size of 32 by 32 was used and two signals of constant brightness levels were generated against a zero background. These signals are corrupted by gaussian white noise to generate the pictures shown on figs. 2a and 3a. These pictures correspond to a signal and a background that both have zero variances and unit correlations, the only factor distinguishing the signal in each picture from its background is a higher mean values for the signals. Figures 2b and 2c and 3b and 3c show the signals as detected for two different thresholds. A negative threshold gives less overall error while a threshold of zero gives a better signal in expense of more error in the background.

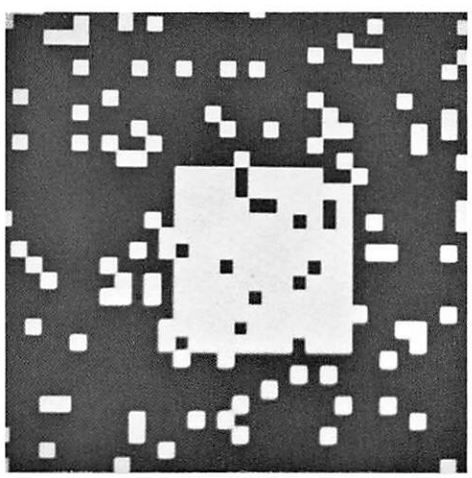
The second example considered is a girl's picture on a grid size of 256 by 256 where each pixel is represented by 8 bits. The original and the noisy picture, corrupted by a white gaussian noise are shown on figs. 4a and 4b. The nonlinear Kalman filter composed of the two dimensional edge detector and two two-dimensional Kalman filters as shown in fig. 3 was simulated on a digital computer. The filtered picture is shown on fig. 4c. The noisy picture was separated into a "signal" and a "background" by using the mid-gray value as a threshold. The statistics of the "signal" and the "background" were estimated separately. For this particular picture the statistics of the "signal" and the "background" were identical except for the mean values, thus two Kalman filters differed only on their



(a) noisy signal

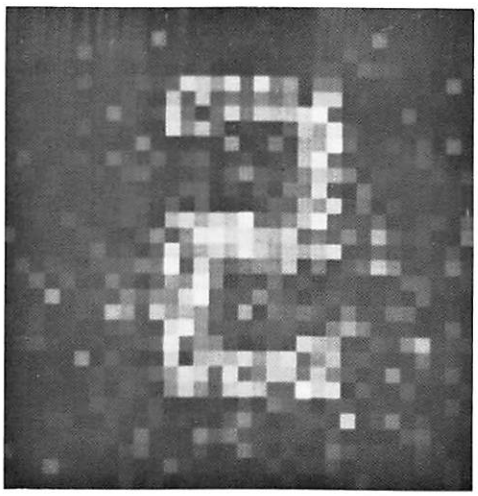


(b) detected signal
threshold = -3

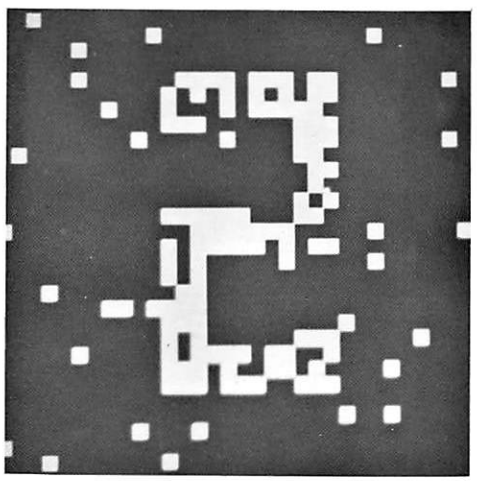


(c) detected signal
threshold = 0.0

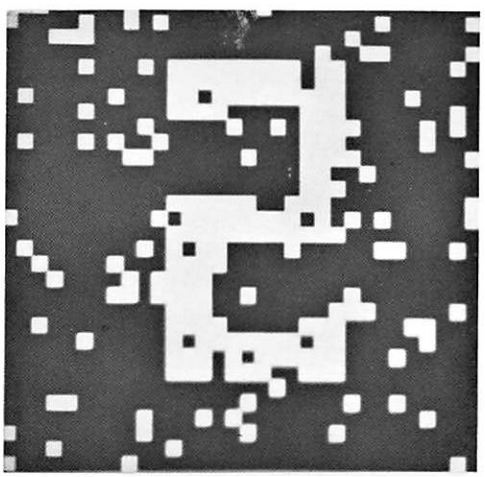
Figure 4.5-2. Input and outputs of the edge detector for threshold values of zero and -3 for square image.



(a) noisy signal



(b) detected signal
threshold = -3



(c) detected signal
threshold = 0.0

Figure 4.5-3. Input and outputs of the edge detector for threshold values of zero and -3 for numeral 2 image.

(a) original

(b) noisy signal

(c) nonlinearly filtered

(d) linearly filtered

Figure 4.5-4. Original, noisy and enhanced images using two-dimensional linear and nonlinear Kalaman filtering.

mean values. Figure 4d is the filtered signal using one Kalman filter as discussed in [3]. Comparison of figs. 4c and 4d indicates that nonlinear filtering eliminates, to some extent, edge smearing which is characteristic of standard Kalman filtering.

References

1. N. E. Nahi, "Role of Recursive Estimation in Statistical Image Enhancement," Proc. of the IEEE, Vol. 60, No. 7, July, 1972, pp. 872-877.

2. N. E. Nahi and C. A. Franco, "Recursive Image Enhancement Vector Scanning," IEEE Transactions on Communications, Vol. COM-21, No. 4, April, 1973, pp. 305-311.

3. A. Habibi, "Two Dimensional Bayesian Estimate of Images," Proc. of the IEEE, Vol. 60, No. 7, July, 1972, pp. 878-883.

4. N. E. Nahi and A. Habibi, "Nonlinear Adaptive Recursive Image Enhancement," Proceedings of 1973 International Joint Conference on Pattern Recognition, October, 1973.

4.6 Nonlinear Recursive Image Enhancement

Nasser E. Nahi and Mohammad H. Jahanashahi

A typical image in the absence of any distortion or noise consists of an object of interest within a background. The object, such as a human face, usually contains detailed information which is essential in the quality of the final enhanced image. In addition to this detailed information content, the object represents a shape defined by its boundary. Generally, both the detailed information and boundary information may be characterized by statistical measures such as the mean and autocorrelation. This is the only a priori knowledge available to the image restoration system.

If one now attempts to represent the entire image (the object within a background) by first and second order moments, the statistics of the boundary will usually overwhelm the statistical information on the details of the object. It is then expected that while the minimum mean square estimate of the object given a noisy observation provides good restoration of the object as an area within the image, it does result in undesirable

blurring of the detail within the object. This is clearly apparent in references [1]-[4] especially in fig. 8 of the latter reference.

Consequently, the estimate of the object should be composed of two estimates: 1) estimates of the object boundary, and, 2) estimate of object detail. Let the output of the scanner be represented by s(t). Furthermore, let $s(t) = a(t)\,\lambda(t)$ where

$$\lambda(t) = \sum_{j=m_1}^{m_2} [u(t-\alpha_j) - u(t-\beta_j)]$$

$$(m_1 - 1)T < t \leq m_2 T$$

$$m_1 = 1, 2, \ldots, M$$

$$m_2 = m_1, m_1 + 1, \ldots, M$$

with

$a(t)$ = gray level of the image at scanning time t

m_1 = starting line of the object

m_2 = terminal line of the object

M = number of lines in the image

$u(\cdot)$ = step function

α_j = start of the object in line j

β_j = end of the object in line j

T = time to scan each line of the image

J = number of pixels in each line of the image

The justification for choosing s(t) as above is due to the a priori knowledge that there exists an object of interest in the image. The problem is to estimate the values of m_1, m_2, α_j, and β_j which would determine the location of the object. Moreover, it is desirable to obtain an estimate of the unknown process a(t) which would result in a high detail restoration of the object.

Sampling s(t) at equal intervals of $t = \frac{T}{J}$ will result in $s(k) = a(k)\,\lambda(k)$,

$$k=(m_1-1)T+\frac{T}{J},\ (m_1-1)T+\frac{2T}{J},\ \ldots,\ m_2T.$$

It will be assumed, without any loss of generality, $\frac{T}{J}=1$. Now given the observation sequence $y(k)=s(k)+v(k)$, with $v(k)$ as the observation noise, one may estimate the unknown process $a(k)$, $k=(m_1-1)J+1,\ldots,m_2J$, and α_j, β_j, $j=m_1, m_1+1, \ldots, m_2$, based on the maximum aposteriori probability criterion. The values of m_1 and m_2 are determined on the basis of statistical inference theory.

Let $a(k)$ be independent of α_j and β_j. Assume $a(k)$ is a zero mean Gaussian random process $E[a(i)\,a(j)]=P(i,j)$, and $v(k)$ is a zero mean Gaussian white noise with $E[v(i)\,v(j)]=\sigma^2\Delta(i-j)$, where $\Delta(\cdot)$ denotes the Kronecker delta. A MAP estimate of $a(k), \alpha_j, \beta_j$ is found to satisfy the following relationships:

$$\hat{a}(k)=L(k,k)\,y(k)+\sum_{\ell=(\hat{m}_1-1)J}^{k-1}L(k,\ell)\,[y(\ell)-\hat{a}(\ell)] \tag{1}$$

where

$$L(k,\ell)=\frac{P(k,\ell)\,\hat{\lambda}(\ell)}{\sigma^2+P(k,k)\,\hat{\lambda}(k)} \tag{2}$$

$$k=(\hat{m}_1-1)J+1,(\hat{m}_1-1)J+2,\ldots,\hat{m}_2J$$

and

$$\hat{\lambda}(\ell)=\sum_{j=\hat{m}_1}^{\hat{m}_2}[u(\ell-\hat{\alpha}_j)-u(\ell-\hat{\beta}_j)] \tag{3}$$

minimum

$$\alpha_{m_1},\alpha_{m_1+1},\ldots,\alpha_{m_2}\ \left\{-2\mathrm{Log}_e f(\alpha_{m_1},\alpha_{m_1+1},\ldots,\alpha_{m_2},\beta_{m_1},\beta_{m_1+1},\ldots,\beta_{m_2})\right.$$
$$\beta_{m_1},\beta_{m_1+1},\ldots,\beta_{m_2} \tag{4}$$
$$\left.+\frac{1}{\sigma^2}\sum_{\ell=(\hat{m}_1-1)J+1}^{\hat{m}_2J}[\hat{a}^2(\ell)-2\hat{a}(\ell)\,y(\ell)]\,\lambda(\ell)\right\}$$

where f(x) is the joint density function. Furthermore, $\hat{m}_1$ and $\hat{m}_2$ are values of m_1 and m_2 which maximize $F(m_1, m_2)$ subject to $m_2 - m_1 > 0$; where

$$F(m_1, m_2) = \Lambda(m_2) - \Lambda(m_1 - 1)$$

$$m_1 = 1, 2, \ldots, M$$

$$m_2 = m_1, m_1+1, \ldots, M$$

and

$$\Lambda(\ell) = \Lambda(\ell-1) + \sum_{i=(\ell-1)J+1}^{\ell J} [2\hat{a}(i)\, y(i) - \hat{a}^2(i)]$$

with

$$\Lambda(0) = 0$$

$$\ell = 1, 2, \ldots, M.$$

It can be shown that $\hat{a}(k)$ for $\hat{\alpha}_j \leq k \leq \hat{\beta}_j$, $j = m_1, m_{1+1}, \ldots, m_2$, takes the same values as estimates of s(k) obtained by a Kalman filter as in [1], [2].

A recursive algorithm for computing $\hat{\alpha}_j$ and $\hat{\beta}_j$ from relation (4) has been developed which results in suboptimal values for $\hat{\alpha}_j, \hat{\beta}_j$. The degree of optimality of this algorithm is under investigation. Experimental results, however are favorable as to the suitability of this algorithm. Once $\hat{m}_1, \hat{m}_2, \hat{\alpha}_j$, and $\hat{\beta}_j$, $j = \hat{m}_1, \ldots, \hat{m}_2$, are determined one can obtain an estimate of the object boundary. Furthermore, values of $\hat{a}(k)$ should result in a high detail restoration.

References

1. N. E. Nahi, "Role of Recursive Estimation in Statistical Image Enhancement," Proceedings of IEEE, Vol. 60, No. 7, July, 1972, pp. 872-877.

2. N. E. Nahi and T. Assefi, "Bayesian Recursive Image Estimation," IEEE Transactions on Computers, Vol. C-21, No. 6, July, 1972, pp. 734-738.

3. S. R. Powell and L. M. Silverman, "Modeling of Two-Dimensional Covariance Functions with Application to Image Restoration," presented at the 1972 Conference on Decision and Control.

4. A. Habibi, "Two-Dimensional Bayesian Estimate of Images," Proceedings of IEEE, Vol. 60, No. 7, July, 1972, pp. 878-884.

4.7 Restoration of Space Variant Aberrations by Coordinate Transformations

Alexander A. Sawchuk

Aberrations in linear incoherent optical systems are most generally described by a space variant point spread function (SVPSF) $h(\underline{x}, \underline{u})$ in a superposition integral

$$\mathcal{I}(\underline{x}) = \int_{-\infty}^{\infty} h(\underline{x}, \underline{u})\, \mathcal{O}(\underline{u})\, du \qquad (1)$$

where $\mathcal{O}(\underline{u})$ is the object radiance and $\mathcal{I}(\underline{x})$ is the image irradiance. Here the shorthand notation $\underline{x} = (x_1, x_2)$ and $\underline{u} = (u_1, u_2)$ denotes the image and object coordinates, respectively, so that $\underline{x} - \underline{u}$ denotes $(x_1 - u_1, x_2 - u_2)$. In general, restoration of space variant degradations is exceedingly difficult [1]; however, by taking advantage of certain degrading system properties, the dimensionality of the problem may be reduced so that practical solutions can be obtained by coordinate transformation restoration (CTR) [2]-[4].

The technique for space variant image restoration involves the transformation of object, image, and SVPSF to a related space invariant system. After the transformation, space invariant processing is used for restoration. Many kinds of aberrations can be exactly described by the cascaded system shown in fig. 1, where $h_I(\underline{z} - \underline{v})$ represents a space invariant operation. The expressions

$$\underline{v} = b(\underline{u}) \qquad (2)$$

and

$$\underline{x} = c^{-1}(\underline{z}) \tag{3}$$

are shorthand notations for geometrical distortions and have been previously discussed [5]. The functions $\mathcal{O}_v(v)$ and $\mathcal{I}_z(\underline{z})$ represent intermediate quantities that appear in the decomposition.

Whenever any space variant degradation can be decomposed to the form shown in fig. 1, the general CTR procedure outlined in the block diagram of fig. 2 can be applied. The first step is to invert the distortion of eq. (3) to produce an estimate of $\mathcal{I}_z(\underline{z})$. Then, the blur-removing part of the restoration is accomplished by using space invariant inverse filtering or statistical estimation to find the function $\mathcal{O}_u(\underline{u})$. The estimator in fig. 2 is followed by another inverse distortion point mapping given by eqs. (4) and (5), and the result is the estimate $\hat{\mathcal{O}}(\underline{u})$ of the object. Considerations on performing distortions by digital and optical techniques are given in Ref. [1]. Simple modifications for filtering with system noise are also possible [1].

To apply CTR to the restoration of space variant aberrations, object and image coordinates are assumed normalized to unit magnification for simplicity. In the geometrical optics description of incoherent imaging systems, the displacement of an image point from its ideal (Gaussian) location in the image plane is specified by a pair of ray aberration functions of the form

$$x_1 = u_1 + \Delta_1(\varepsilon_1, \varepsilon_2, u_1, u_2) \tag{4a}$$

$$x_2 = u_2 + \Delta_2(\varepsilon_1, \varepsilon_2, u_1, u_2) \tag{4b}$$

where $\underline{u} = (u_1, u_2)$ is the object point, $\underline{x} = (x_1, x_2)$ is the image point, and $\underline{\varepsilon} = (\varepsilon_1, \varepsilon_2)$ are coordinates in the exit pupil σ of the optical system [6]. Equation (4) specifies that a ray leaving (u_1, u_2) and passing through exit pupil point $\underline{\varepsilon}$ intersects the image plane at $\underline{x}$. If series expansions of Δ_1 and Δ_2 in eq. (4) are limited to the lowest order terms, these terms appear

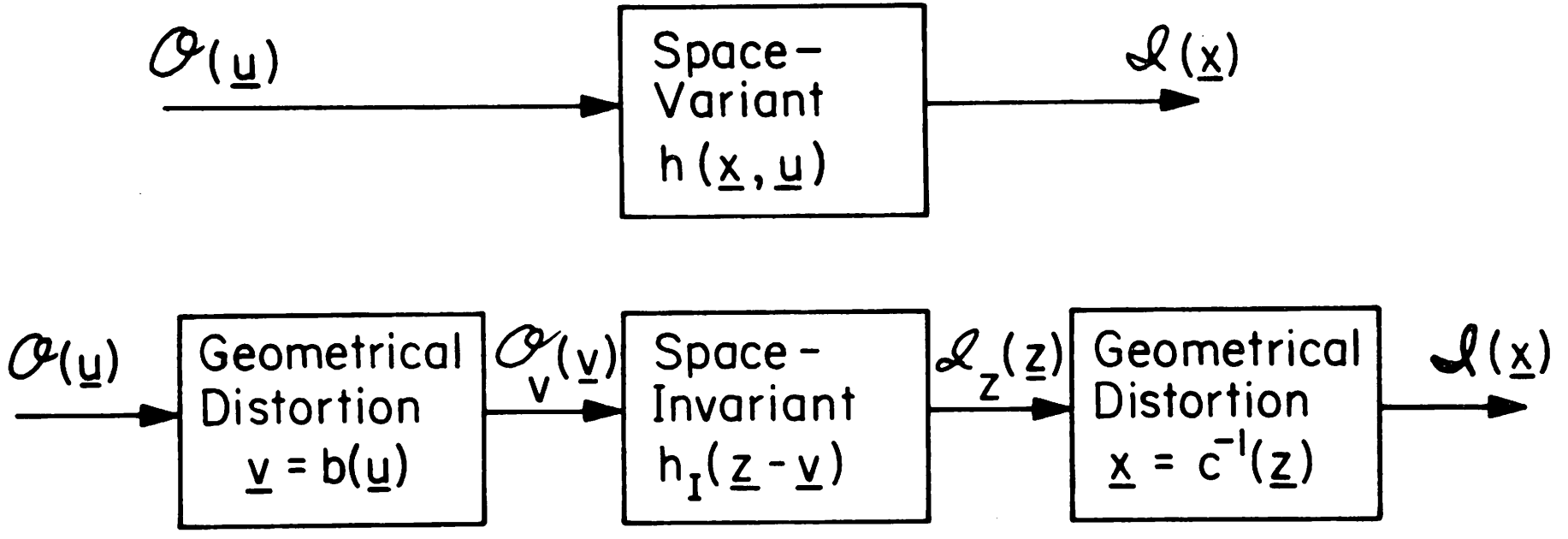


Figure 4.7-1. Model for decomposition of space variant impulse response.

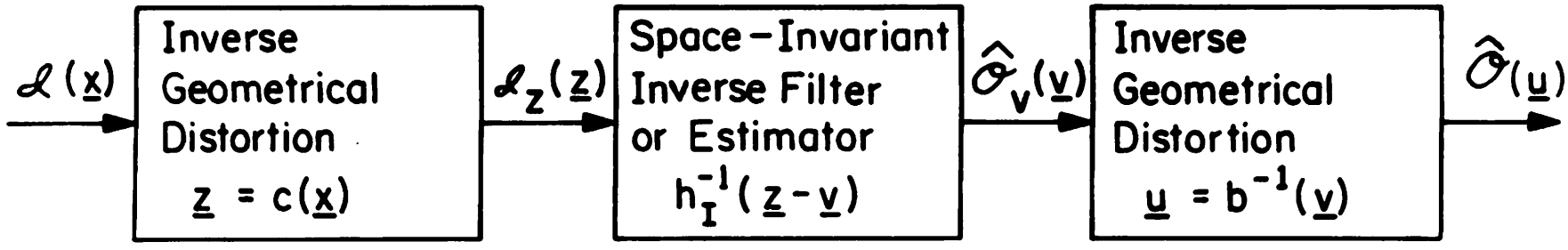


Figure 4.7-2. Model for restoration with a space variant impulse response.

in five combinations that describe the well known five primary Seidel aberrations [6]. Of these, the term for spherical aberration is independent of x_1 and x_2, implying a space invariant degradation because the blur does not change over the image plane. The term for distortion aberration depends only on x_1 and x_2, implying no blurring at all. The remaining terms for coma, astigmatism and curvature of field depend on both $\underline{x}$ and $\underline{\epsilon}$ and imply space variant imaging. If all aberration coefficients are zero, the Δ_1 and Δ_2 terms of eq. (4) are zero and ideal imaging occurs.

The form of the SVPSF for these aberrations can be derived by a procedure similar to that used for finding space variant motion degradation [1], [2], [8]. The major difference is that the summation is over the exit pupil σ instead of over the exposure time interval. The object point $\underline{u}$ is assumed near the optical axis so that the irradiance at the exit pupil is a constant independent of position $\underline{\epsilon}$. Assuming unity irradiance at and conservation of energy in an infinitesemal area in the mapping from $\underline{\epsilon}$ to $\underline{x}$, the SVPSF response $h(\underline{x}, \underline{u})$ satisfies

$$h(\underline{x}, \underline{u})\, d\underline{x} = d\underline{\epsilon} \tag{5}$$

where $dx = dx_1 dx_2$ and $d\epsilon = d\epsilon_1 d\epsilon_2$ are area elements in the image and exit coordinates, respectively. Using the aberration functions, these area elements are related by

$$J_\Delta(\underline{\epsilon}, \underline{u})\, d\underline{\epsilon} = d\underline{x} \tag{6}$$

where

$$J_\Delta(\underline{\epsilon}, \underline{u}) = \begin{vmatrix} \dfrac{\partial x_1}{\partial \epsilon_1} & \dfrac{\partial x_1}{\partial \epsilon_2} \\ \dfrac{\partial x_2}{\partial \epsilon_1} & \dfrac{\partial x_2}{\partial \epsilon_2} \end{vmatrix} \tag{7}$$

is a Jacobian determinant. In general, there are N exit pupil points $\left(\epsilon_1^1, \epsilon_2^1\right)$, $\left(\epsilon_1^2, \epsilon_2^2\right) \cdots \left(\epsilon_1^N, \epsilon_2^N\right)$ that map into $\underline{x}$, and combining eq. (5) with eq. (6) and summing over the N points gives the SVPSF for aberrations as

$$h(\underline{x}, \underline{u}) = \sum_{i=1}^{N} \left| J_{\Delta}(\underline{\epsilon}, \underline{u}) \right|^{-1}_{(\epsilon_1, \epsilon_2)\epsilon\sigma} \tag{8}$$

for $\underline{x}$ and $\underline{u}$ satisfying eq. (4) with $\underline{\epsilon} \in \sigma$, and zero elsewhere. In evaluating eq. (8), reverse substitution to eliminate $\underline{\epsilon}$ must be made in $J_{\Delta}(\underline{\epsilon}, \underline{u})$ using eq. (4).

The CTR procedure for space variant aberrations simply involves solving eq. (5) and rewriting them in the form

$$b_1(u_1, u_2) = c_1(x_1, x_2) - m_1(\epsilon_1, \epsilon_2) \tag{9a}$$

$$b_2(u_1, u_2) = c_2(x_1, x_2) - m_2(\epsilon_1, \epsilon_2). \tag{9b}$$

Whenever this is possible, the aberration is represented as a space invariant blur between distorted planes and CTR can be used. As before, $b(\underline{u})$ and $c(\underline{x})$ are the distortions, and evaluating eq. (8) gives

$$h_I(\underline{z}-\underline{v}) = \sum_{i=1}^{N} \left| J_{\Delta}(\underline{\epsilon}) \right|^{-1}_{\substack{\epsilon_1 = m_1^{-1}(z_1-v_1, z_2-v_2) \\ \epsilon_2 = m_2^{-1}(z_1-v_1, z_2-v_2)}} \tag{10}$$

for $\underline{x}$ and $\underline{u}$ related by eq. (18) with $\epsilon \in \sigma$, and zero elsewhere.

This procedure can be used for the restoration of coma-like aberrations and tilt in a cylindrical lens system [1], [7], [8]. Because of the circular symmetry in many imaging systems it is desirable to express coordinates in the $\underline{u}$, $\underline{x}$, and $\underline{\epsilon}$ planes as polar variables (u_r, u_θ), (x_r, x_θ) and $(\epsilon_r, \epsilon_\theta)$. For a system with coma aberrations alone, the aberration functions may be written

$$x_r = u_r(1 - F\epsilon_r^2(2+\cos 2\epsilon_\theta)) \tag{11a}$$

$$x_\theta = u_\theta - \pi/2 + \arctan[(1 - F\epsilon_r^2(2+\cos 2\epsilon_\theta))/-F\epsilon_r^2 \sin 2\epsilon_\theta] \tag{11b}$$

where F is the coma aberration coefficient. In this polar representation

the system is invariant in the θ variables because $x_\theta - u_\theta$ depends only on $\underline{\varepsilon}$ variables. Taking the natural logarithm of eq. (11a) and identifying

$$z_r = \ell n\, x_r \tag{12a}$$

$$v_r = \ell n\, u_r \tag{12b}$$

$$z_\theta = x_\theta \tag{12c}$$

$$v_\theta = u_\theta \tag{12d}$$

$$m_r(\varepsilon_r, \varepsilon_\theta) - \ell n(1 - F\varepsilon_r^2(2+\cos 2\varepsilon_\theta)) \tag{12e}$$

$$m_\theta(\varepsilon_r, \varepsilon_\theta) = \pi/2 - \arctan[(1 - F\varepsilon_r^2(2+\cos 2\varepsilon_\theta))/-F\varepsilon_r^2 \sin 2\varepsilon_\theta] \tag{12f}$$

puts the system in the form of eq. (4) for CTR. Robbins and Huang [8] have obtained very good computer restoration of coma and cylindrical lens tilt by this method using a fast-Fourier-transform (FFT) inverse filter and Wiener filter for $h_I^{-1}(\underline{z}-\underline{v})$ in fig. 2.

For astigmatism and curvature of field the aberration functions are

$$x_r = u_r(1+(2C+D)\,\varepsilon_r\, u_r \cos \varepsilon_\theta \tag{13a}$$

$$x_\theta = u_\theta - \pi/2 + \arctan[(2C+D)/D \tan \varepsilon_\theta)] \tag{13b}$$

where C and D are the aberration coefficients. Such a degradation may be inverted by first performing a geometrical distortion to an (r, θ) system, making the degradation invariant in θ. A transform in the θ variable is then performed and a different one-dimensional space variant estimator for each θ frequency is used to restore in the r direction. The θ frequencies are then filtered, and an inverse transform is θ and reverse polar coordinate mapping produces the estimate $\hat{\mathcal{O}}(\underline{u})$. The CTR method does not apply as directly as it does for coma, but still reduces the system from four dimensions to three. Simulations of this procedure are to be carried out.

References

1. A. A. Sawchuk, "Space-Variant Image Restoration by Coordinate Transformations," submitted to Journal of the Optical Society of America.

2. A. A. Sawchuk, "Image Restoration by Space-Variant Decomposition," USC Semiannual Technical Report, USCEE 425, 1 March 1972 - 31 August 1972, pp. 75-82.

3. A. A. Sawchuk, "Coordinate Transformations in Space-Variant Image Enhancement and Restoration," J. Optical Society of America, Vol. 62, November, 1972, p. 1337. Annual Meeting, Optical Society of America, San Francisco, 1972.

4. A. A. Sawchuk, "Space-Variant System Analysis of Image Motion," Journal of the Optical Society of America, Vol. 63, September, 1973, p. 1052.

5. A. A. Sawchuk, "Space-Variant Image Motion with Different Geometrical Distortion Models," USC Semiannual Technical Report USCEE 444, 1 September 1972 - 28 February 1973, pp. 81-87.

6. M. Born and E. Wolf, Principles of Optics, Fourth Ed., Pergamon Press, London, 1970.

7. G. M. Robbins, Mass. Inst. Tech. Res. Lab. Electronics Quart. Prog. Rpt. No. 93, 1969, p. 243.

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

4.8 Color Image Restoration

Clanton Mancill

Color imaging systems such as color television and color photography attempt to provide a spectral output at each image point which appears to the human visual system to match as nearly as possible the color of the corresponding point in the original scene. The errors in color reproduction which are introduced by the imaging system can be put in three categories:

1. Imperfect sensors. Ideal color sensors would have spectral sensitivities equivalent to some set of human visual color matching curves. When this condition is not met, some color information is lost, and exact

color restoration is impossible. Statistical knowledge about the input scene can provide partial compensation, however, for the errors due to imperfect sensors.

2. Signal contamination. Color signals can be contaminated by noise or crosstalk from other color signals. Photographic dyes, for example, absorb light in undesired spectral regions causing the absorptions of the other dye layers to appear to be excessive. Color crosstalk can, in theory, be corrected fully if the system is properly modelled.

3. Nonlinearities. Some system nonlinearities may be fully corrected, e.g., gamma correction. Others, such as dynamic range limiting, may result in uncorrectable color errors.

Digital Color Restoration The reduction of color errors in a photographic recording of a scene may be done digitally with significant advantages. The flexibility and speed of computation with a digital computer allows the restoration of color in a digitally stored image in ways which are not possible with photographic restoration techniques.

The image which is to be restored is first recorded on color photographic film and converted to electrical signals by scanning the film transparency with a color separation scanner. Analog-to-digital conversion then is performed and the resulting digital array (three color values per pixel) is stored on tape.

The properties of a color reversal photographic film and a flying spot scanner have been modelled in a computer simulation program. The color errors due to imperfect sensitivities, crosstalk and nonlinearities are then present in the scanner signals which are produced as output from the simulation. These scanner signals are then used as input to various color estimation algorithms with the purpose of estimating as accurately as possible the color of each image point of the original scene as it would be perceived by a human observer. The human observer's response to input and output colors are described by computing the tristimulus values of the colors in a standard system such as the 1960 C. I. E. uniform

chromaticity scale (U. C. S.).

The estimation of the tristimulus values of the original image points using as observables the digital scanner signals may be accomplished in two independent steps. First the three film layer exposures are estimated. Then using the estimated exposures as inputs the original tristimulus values are estimated.

The estimation of film exposures from scanner signals requires the solution of three simultaneous integral equations. This may be done with arbitrary accuracy if sufficient computation is performed and if the system is modelled accurately. Exposure estimation has been described in a previous report [3].

The estimation of tristimulus values poses a more difficult problem for two reasons. First, information is lost when the film sensitivities differ from color matching curves. Secondly, the performance of any tristimulus estimator depends on the spectral intensity distribution of the light which exposed the film. Spectral intensity distributions are continuous curves with an infinity of degrees of freedom. In order to investigate tristimulus estimation techniques, a generator of spectral intensity curves possessing desired tristimulus values and properties of colors existing in nature has been developed.

Color Generators Three types of color generating algorithms have been investigated:

1) Variational method [2]
2) Pseudoinverse method
3) Gaussian curve method.

The first two methods generate intensity curves which have any desired tristimulus values, but which are not guaranteed to have the desirable property of being everywhere nonnegative. The third method generates nonnegative curves, but the resulting tristimulus values only approximate the desired values. Taken together, these methods can generate large

numbers of satisfactory colors for testing tristimulus estimators. A measured spectral intensity is shown in fig. 1 along with visually equivalent distributions which have been generated by methods 1) and 2).

Tristimulus Estimation The tristimulus estimators which have been tested thus far are linear estimators. The general form is the simple matrix equation:

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} X_R \\ X_G \\ X_B \end{bmatrix} \tag{1}$$

where X_R, X_G, X_B are the estimated film exposures and U, V, W are the resulting estimated UCS tristimulus values of the color. The constants m_{ij} are to be chosen in order to minimize the expected perceived color error in some sense.

Some simple adhoc choices of constants have been tested with good results. The nine values of m_{ij} can be chosen so that three test colors, widely separated in chromaticity space, will be estimated exactly, with some error to be expected for all other input spectral distributions. The constants can be chosen to minimize the mean squared tristimulus error over a large class of test colors. Since all colors are weighted sums of spectral (single frequency) colors, the class of spectral colors suggests itself as a test color. The results, however, indicate that a class of test colors containing desaturated as well as saturated colors is better when tested against measured natural colors (see fig. 2).

Under investigation at this time is a method of tristimulus estimation which utilizes stochastic linear regression [2]. The color spectral distribution $c(\lambda)$ is estimated, then the resulting tristimulus values are obtained. The estimated $c(\lambda)$ is obtained as the solution to the set of Fredholm integral equations

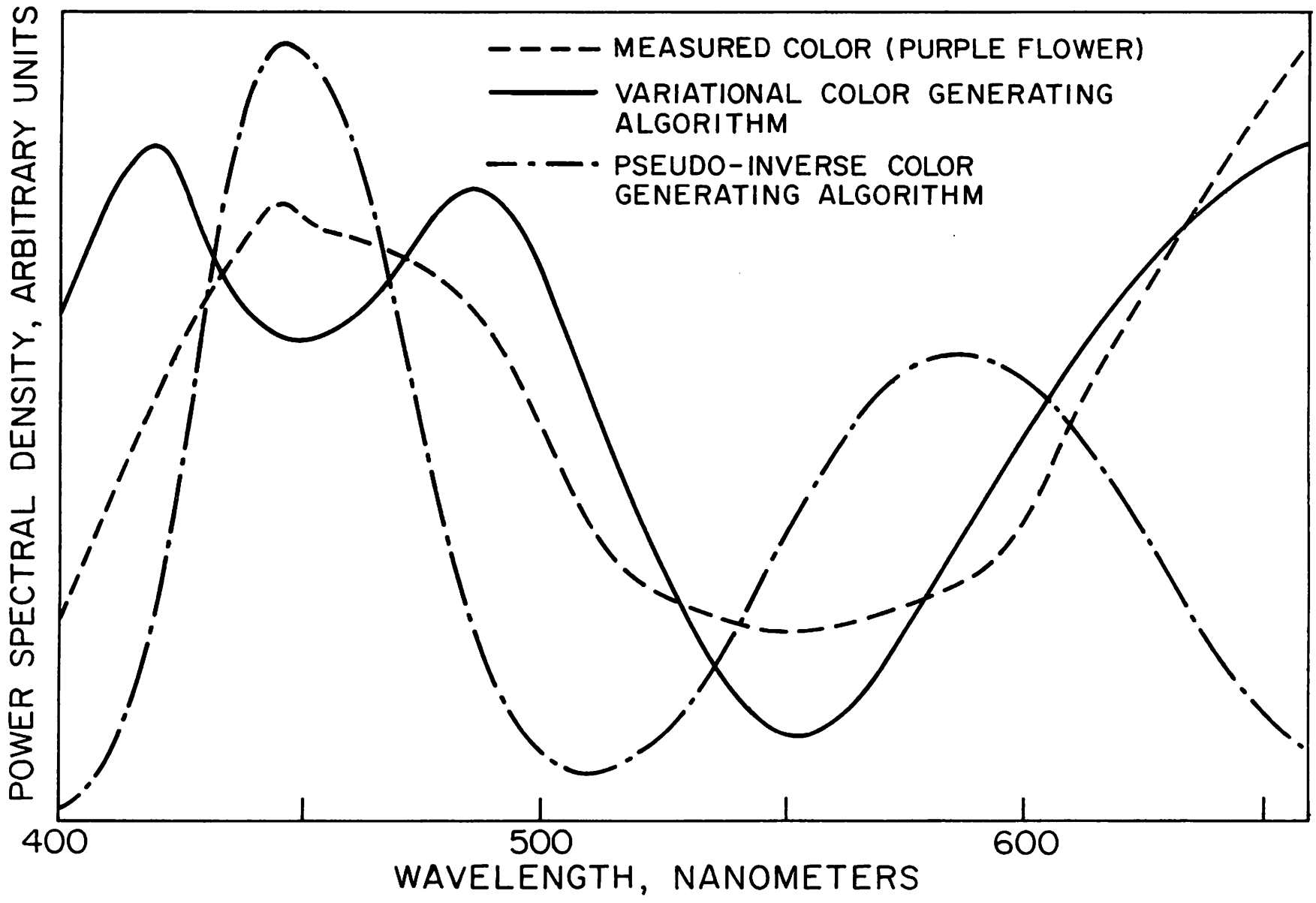


Figure 4.8-1. Comparison of three visually equivalent (metameric) colors, one measured and two computer generated

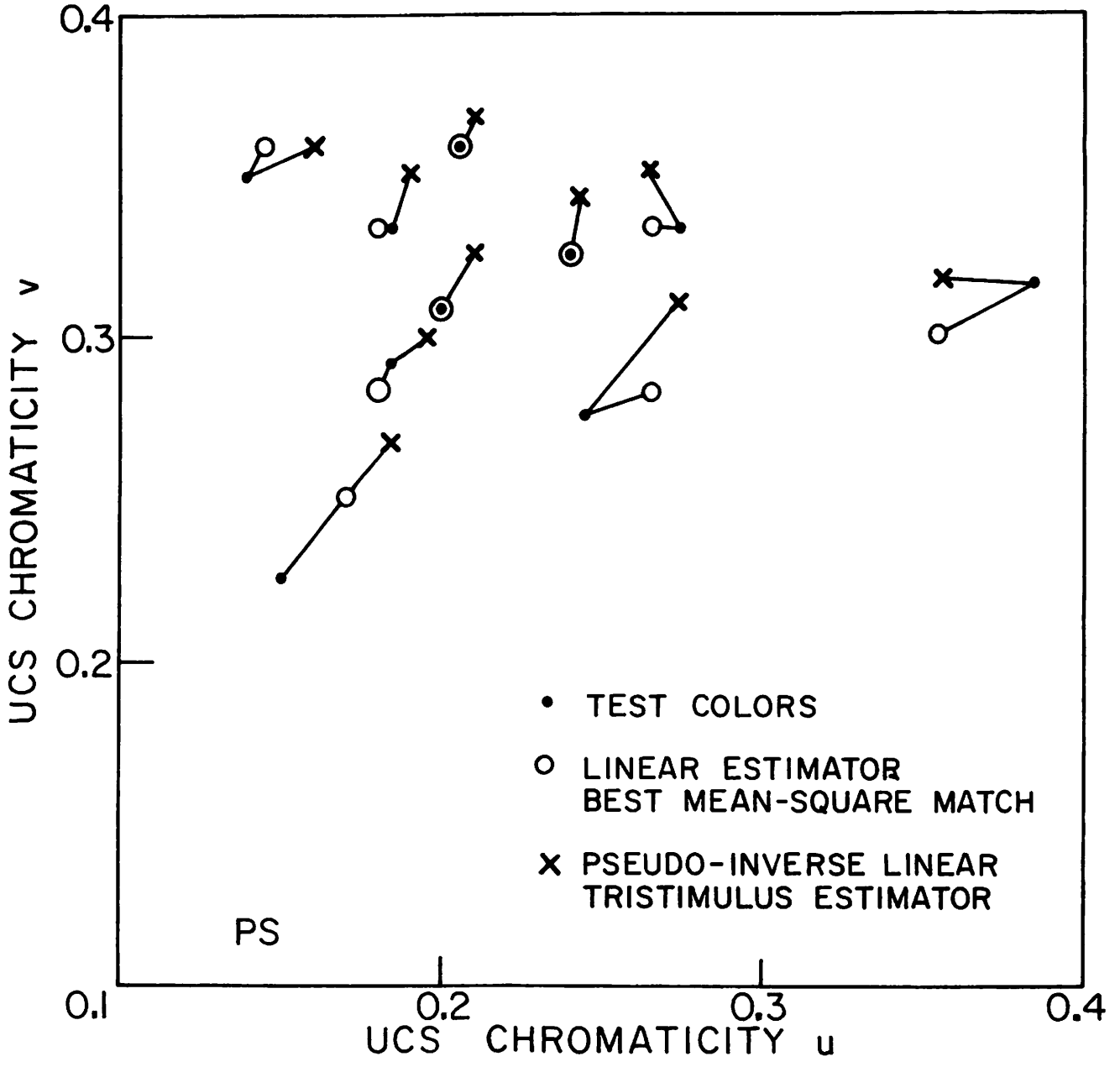


Figure 4.8-2. Chromaticity errors of two types of tristimulus estimators against ten test colors.

$$x_i = \int_{\lambda_1}^{\lambda_2} s_i(\lambda)\, c(\lambda)\, d\lambda + \Sigma_i, \qquad i=1, 2, 3 \tag{2}$$

where λ is wavelength, the x_i are the previously estimated exposures, the $s_i(\lambda)$ are the known film sensitivities, and Σ_i is an additive term representing noise and measurement uncertainty. These equations are discretized and the resulting matrix equations can be solved for $c(\lambda)$ or for scalar functions of $c(\lambda)$, such as the tristimulus values of $c(\lambda)$.

References

1. K. Takahama and Y. Nayatani, "New Method for Generating Metameric Stimuli of Object Colors," J. Opt. Soc. America, Vol. 62, 1972, pp. 1516-1520.

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

3. W. K. Pratt, Semiannual Technical Report, University of Southern California, USCEE Report No. 411, March, 1972.

4.9 Perfect Photographic Reproduction from an Additive Color Display

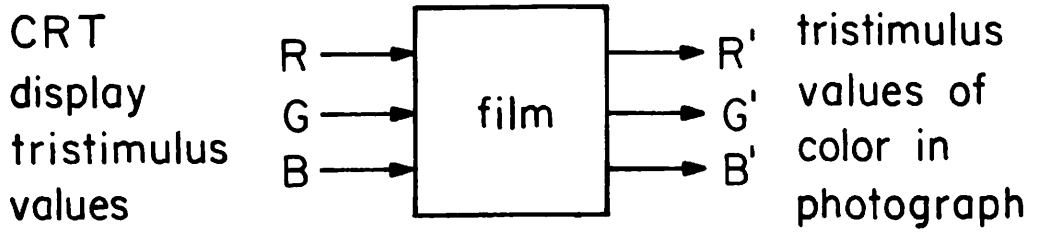
Robert Wallis

It is well known that a given color can be matched by a weighted sum of three primaries. When these weights are scaled such that unity amounts of all three yield a reference white, they are known as tristimulus values [1]. Unfortunately, the spectral sensitivities of the three layers in present day color films are such that the tristimulus values of the colors being photographed are not preserved. Thus, it is possible that colors which appear the same to the human eye (metameric colors) will photograph differently and vice versa. If however, the goal is not to photograph a natural scene, but to photograph a color image on a CRT, the problem becomes tractable. This is because, unlike natural scenes, each color which the CRT can produce is a unique weighted sum of its red, green, and blue primaries, and thus has a unique spectral distribution associated with it. The same uniqueness property holds for color film in spite of the fact

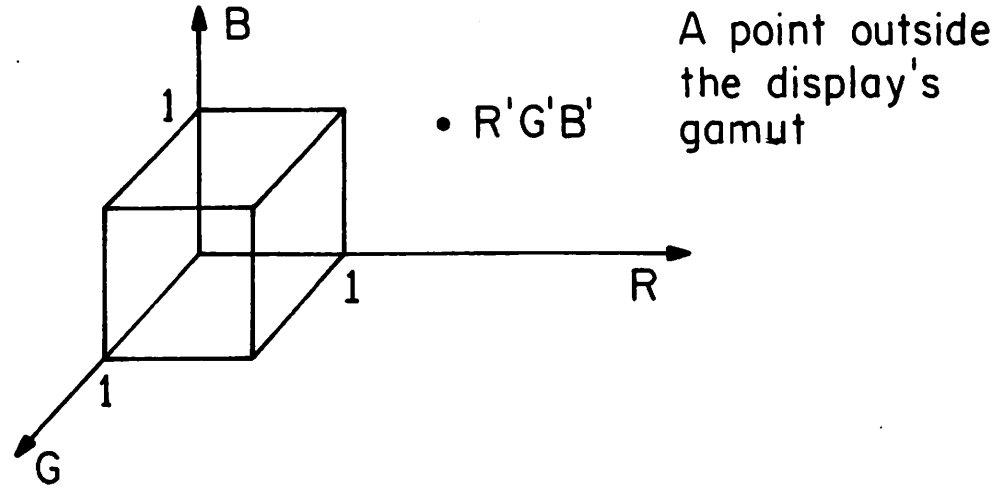
that it is a subtractive instead of an additive system. This means that the CRT-film system can be modelled as a mapping which converts each set of tristimulus values which the CRT can generate into a (usually) different set on the photograph (see fig. 1a).

Thus, it is theoretically possible to force the film to yield the desired tristimulus values. This can be done by using the inverse of the mapping and photographing a color which will be distorted into the desired color. This should always be possible provided that the gamut of colors which the film can produce is a subset of the CRT's gamut. If this is not true, a number of different strategies are possible to cope with the problem. Some possible techniques are discussed below.

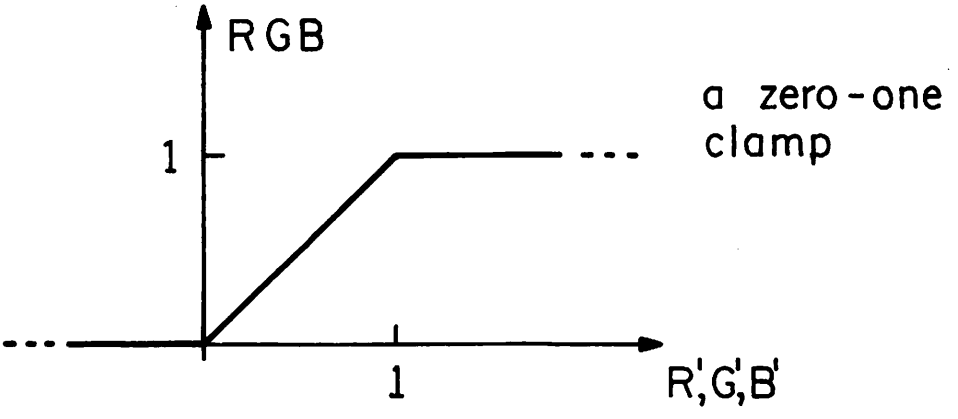
Gamut Restrictions Any color reproduction system is only capable of producing a certain range, or gamut of colors. This gamut can be represented as a solid in tristimulus value space. Consider for instance a color CRT display with red, green and blue primaries. If the amount of each primary is constrained to be between 0 and 1, then the display's gamut is a unit length cube in R, G, B space (see fig. 1b). The problem considered here is how should one represent a color lying outside the possible reproduction gamut, such as one requiring a negative amount of red for instance? A reasonable approach is to use the color on the display solids' surface which is geometrically closest to the color R'G'B'. For the special case in which the solid is a unit cube, this strategy is equivalent to merely "clipping" the R'G'B' coordinates between zero and one. This transfer function is shown in fig. 1c. Simulations on real color images have shown that such a trivial scheme works surprisingly well. However, there is much room for improvement since the RGB space is typically not a uniform color space [3]. In other words, equally separated colors in the space do not seem equally separated perceptually. For example, errors in saturation seem less noticeable than errors in hue, for most colors. This suggests another simple scheme, that of desaturating (adding white) the color towards grey until it lies within the possible reproduction gamut. A convenient way to represent these approaches is the use of a chromaticity diagram in which



(a)



(b)



(c)

Figure 4. 9-1. Color film gamut correction.

luminance information is ignored (fig. 2).

A more relevant approach would be to use one of the uniform color scales. That is, a space in which metric distances more nearly correspond to perceptual distance. Good examples are the $U^*V^*W^*$ and L, a, b (cube root) coordinate systems [3]. Such nonlinear systems would of course lead to a distorted display solid, and render the analytical problems of finding a point within that solid closest to a given point outside the solid much more difficult.

Computer simulations of the above mentioned techniques along with the development of a mathematical model of color film are currently underway. The goal of this research is the capability of producing high fidelity color photographic reproductions of color CRT images.

References

1. W. K. Pratt, Digital Color Image Coding and Transmission, USCEE Report 403, 1971.

2. R. M. Evans, W. T. Hanson, and W. L. Brewer, Principles of Color Photography, John Wiley and Sons, New York, 1953.

3. G. Wyszecki and W. S. Stiles, Color Science, John Wiley and Sons, New York, 1957.

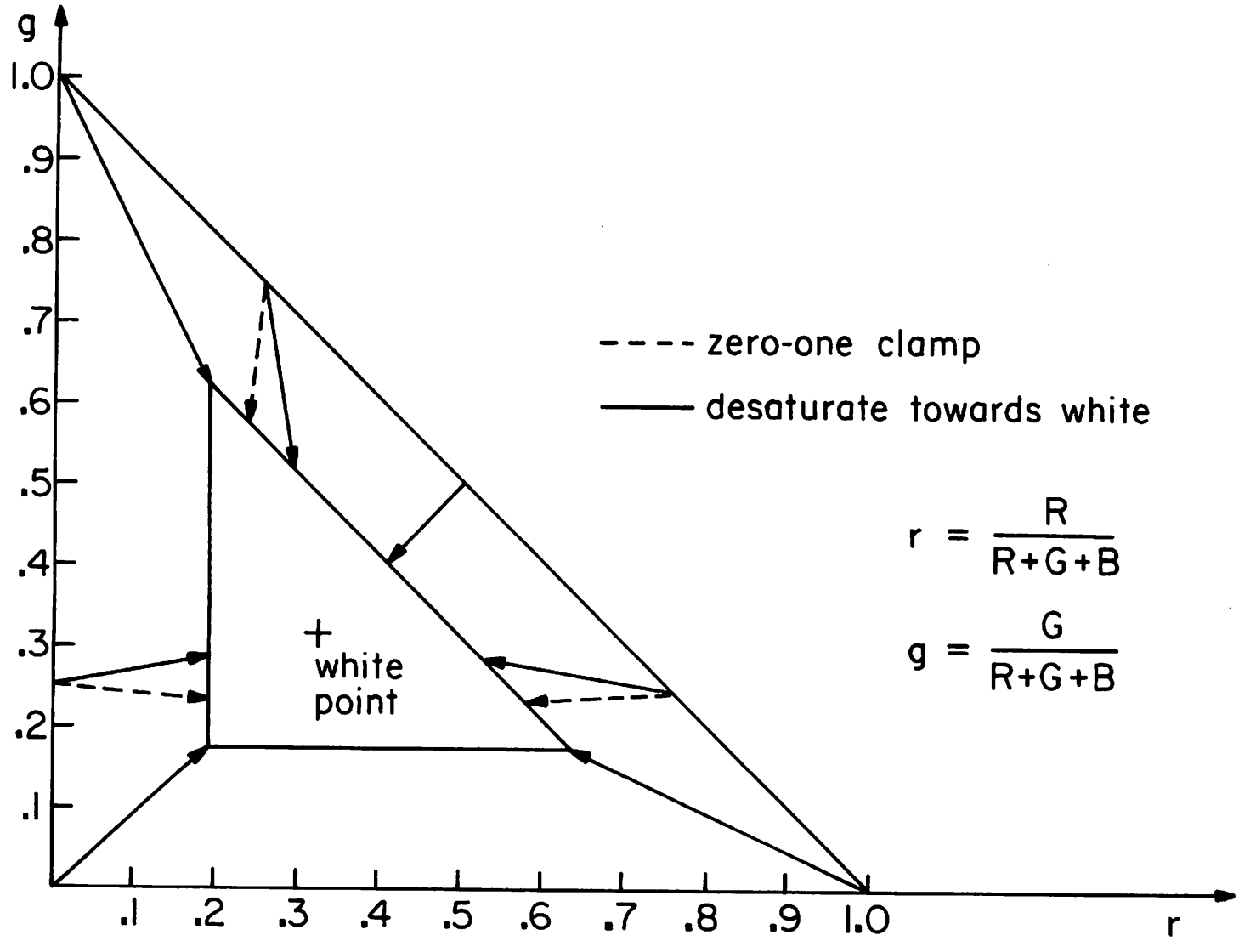


Figure 4.9-2. Chromaticity diagram showing two strategies of dealing with gamut restrictions. The smaller triangle represents the restricted gamut. ie. points from the large triangle are mapped into the smaller, more restricted triangle.

5. Image Data Extraction Projects.

Image data extraction is oriented towards the automated or interactive detection of specific features and the measurement of parameters in images. Various image enhancement techniques such as histogram equalization, spatial registration etc. are typically performed prior to the data extraction process for purposes of normalization and for more efficient extraction.

The following report is a study of the potentials of image data extraction for the automated detection of female breast carcinoma. One diagnostic method consists in determining anomalies in the body surface temperature patterns. Images can be obtained with an infra-red scanner where a suitable portion of the temperature scale is mapped onto a grey scale. In a mass-screening situation, these images can be analyzed by computer, and further tests such as biopsies would be indicated upon a positive computer diagnostic.

5.1 Quantitative Measures of Asymmetry in Contiguous Image Regions

Richard P. Kruger and Mark A. Stein

The human visual system is excellent at detecting local variations in picture brightness. However, it is poorly adapted for quantitative image brightness estimation and contiguous area comparisons [1]. Digital processing techniques have proven useful in the detection of objects within images and performing quantitative image interpretation.

A project recently undertaken by members of the biomedical image processing group of the Image Processing Institute entails detection of female breast carcinoma from thermogram scans. Basically, the techniques involve the location of continguous regions of constant temperature within each breast area and determination of possible asymmetry of these areas. The technical problems associated with the asymmetry measurements appear to be of general interest, and are described below.

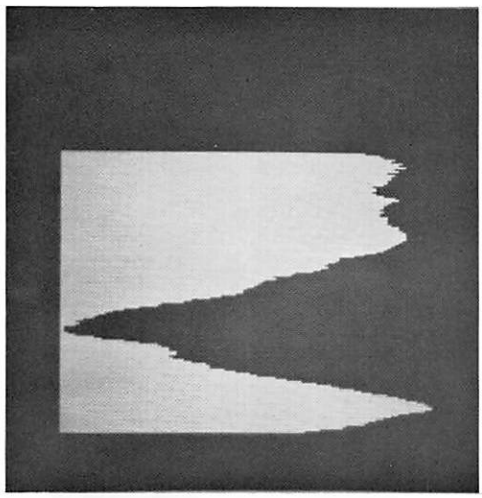
Quantitative Thermographic Measurements. A quantitative computer approach to asymmetry measurements will of necessity be experimental. The goal is to extract quantitative measures indicative of this asymmetry. The most salient common factor in these criterion is that of asymmetry between the left and right breast regions caused by temperature differences. The hypothesis therefore is that asymmetry between the breast regions in a thermogram are indicative of po-

tential abnormality with elevation of temperature as a primary component cause. It has been reported that temperature elevation of at least 1 degree centigrade in an involved breast region over that of an uninvolved region of the same breast or a similar region of the contralateral breast was present in 95 percent of subsequently proven malignancies in a screening program involving 3,500 women [2].

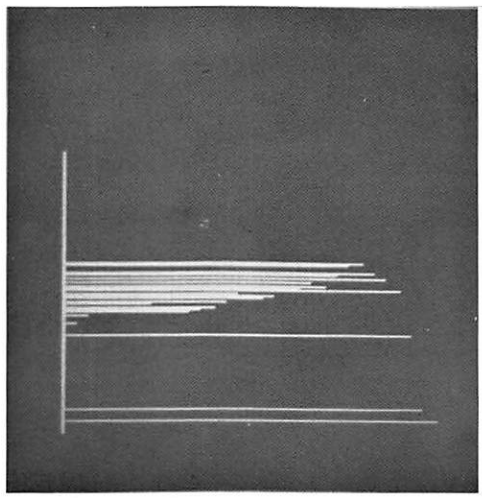
A feasibility study is being undertaken consisting of 15 benign and 16 malignant thermographic images. These images were recorded on polaroid film and later digitized to 256 by 256 pixels. The malignant cases were confirmed by biopsy.

Spatial Signature Analysis. Spatial signature analysis [3, 4, 5] is a technique which has aided in the localization of anatomical features in thyroid scintigram, bone lesions, and chest radiographs. These signatures are obtained by either summing the rows or columns of a digital array within a prescribed region. In the present application, the spatial signatures are being studied for two potential uses. First it is hoped that they will be useful in automatic location of the breast region in the thermogram. Second, it is hoped that measures which characterize one or both of the signatures will be useful asymmetry features for the computer diagnostic screening. Preliminary evidence indicates that these signatures may be of use for both purposes. The case illustrated in Figure 1 is an example of their potential.

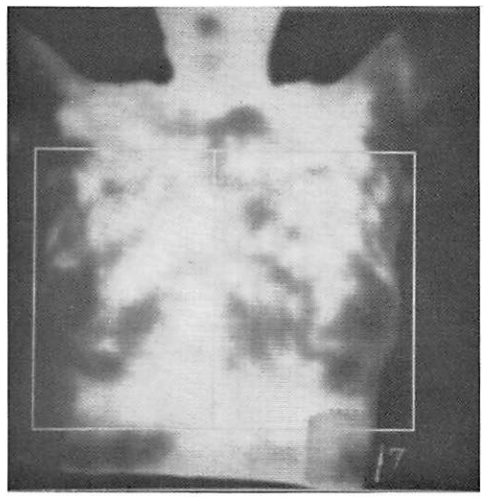
The horizontal signature obtained by summation of the indicative subframe columns indicates a characteristic pattern shown in figure 1. This signature indicates that non anatomical background can easily be distinguished from either the left or right lateral chest wall. In addition, the chest middle is detected by the relatively "hot" mediastinal region. The vertical signature is characterized by a lateral summation of the fraction of picture elements in rows between the lateral chest walls. This signature is characterized by two "hot" peaks between a cool region associated with the breast. The upper peak is due to the relatively uniform heat pattern detected superior to the breast region. The inferior peak is caused primarily by a lateral summation across the intra mammary fold. By these signatures, isolation of both the breast region and individual breast areas appears to possible. This preprocessing is however not the objective. However the horizontal signature has also provided interesting insights. The thermogram horizontal signature in figure 1 also demonstrates diffuse elevation of the left breast temperature. What remains is to further reduce this signature information to a manageable set of features.



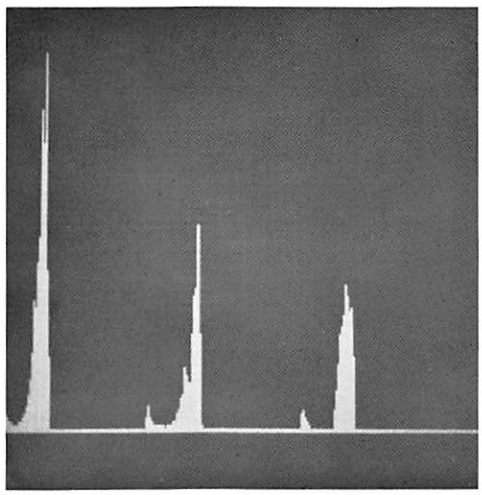
Vertical signature



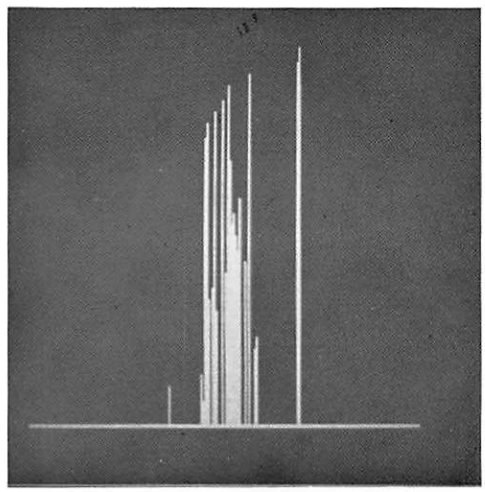
Vertical moment



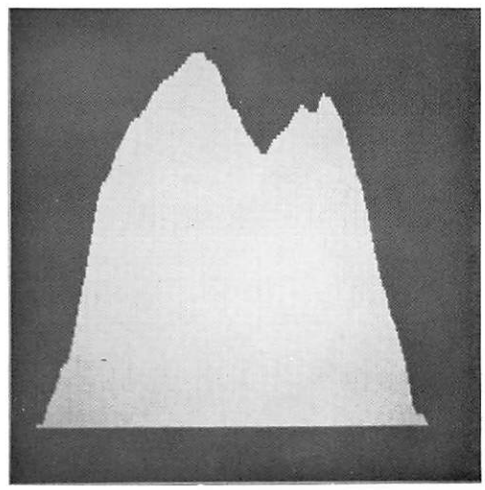
Digitized normal thermogram



Total, right and left breast histograms respectively



Horizontal moment



Horizontal signature

Figure 5.1-3. Normal thermogram.

This example illustrates the use of spatial signature for detecting and localizing relevant anatomy and shows a potential for lesion detection. However, other measures will also be presented.

Histogram Analysis. After the breast regions have been isolated it is possible to compute measures of symmetry based upon the histogram or first order statistics (thermal density distribution) of rectangular regions which enclose each breast. Temperature distributions computed for one breast may be compared with those for the contralateral breast in order to provide a measure of thermal asymmetry. The 30 level histograms shown in Figure 1 represents the total right and left breast histograms respectively. These histograms indicate the left breast has a greater proportion of its area at a higher temperature when compared with that of the right breast. Quantitative measurement of this asymmetry is being undertaken using several measurements often used to characterize probability density functions.

The histogram h(t) of respective breast areas will be denoted as:

$$h(t) \quad t=1,\ldots,30 \tag{1}$$

In addition to the mean (μ) and variance (V) which are well known measures of average gray level (temperature) and dispersion respectively, two other measures are being extracted. Skewness (s) measures the departure from histogram symmetry and kurtosis (K) expresses the tendency of a distribution to either cluster about μ or disperse towards the tails of the distribution.

These measures are expressed as follows.

The r-th moment about zero M_r is given by

$$M_r = \sum_{t=1}^{30} t^r h(t) \tag{2}$$

The mean $\mu=m_1$. The r-th central moment is given by

$$C_r = \sum_{t=1}^{30} (t-\mu)^r h(t) = \sum_{j=0}^{v} \binom{v}{j} M_{r-j} (-M_1)^j \tag{3}$$

The variance V, is given by

$$V = C_2 = M_2 - M_1^2$$

The skewness S is given by

$$S = \frac{\sqrt{B_1}\,(B_2+3)}{2\,(5B_2-6B_1-9)} \qquad (4)$$

where B_1 is referred to as the coefficient of skewness and B_2 is the coefficient of Excess

$$B_1 = \frac{C_3^{\,2}}{C_2^{\,3}} \qquad B_2 = \frac{C_4}{C_2^{\,2}} \qquad (5)$$

The kurtosis K is given by

$$K = \frac{C\mu}{C_2^{\,2}} - 3 \qquad (6)$$

These measures are computed for each breast region. Initial automated diagnostic classification will be attempted with the following features

$$f_1 = |\mu_r - \mu_L| \qquad f_3 = |S_r - S_L|$$

$$f_2 = |V_r - V_L| \qquad f_4 = |K_r - K_L| \qquad (7)$$

This is an attempt to detect asymmetry via absolute difference of the studied measures under the hypothesis that complete symmetry would yield a zero value for each f_i of course other measures may evolve but these will initially be tried.

Spatial moments. Spatial moments are commonly used to describe a mass distribution. The total mass, centroid, and radius of gyration are often used in mechanics to describe the properties of a single object of a collection of objects. The generality of the method of moments and the use of moment in varients in pattern recognition was described by Deutsch [6]. He proposed that the information contained in the relation of a point to the weights of its neighbors can be hypothesized to demonstrate the invariances to size, location, orientation, mirror image, and symmetry which have been observed in human perception experiments. Furthermore, he proposed that the process of human perception was based on the assumption that the human measures moment type information in an image. He proposed the theory mainly as a basis for developing other, more explanatory, theories.

The mathematical generality of the set of moment invariants has been pointed out by Hu [7], Duda and Hart [8], and others. The set of infinite moments of any bounded density function, exactly, completely, and uniquely describe the density function. All possible measureable features of an image are represented in the set of moments. Various normalization procedures may be performed on the computed moments to obtain metrics which are invariant to certain transformations such as translation, rotation, size change, or symmetry. The moments may be easily calculated for any finite image function, f(s, y), the (p+q)th order moments are defined by:

$$m_{pq} = \sum_x \sum_y x^p \, y^q \, f(x, y) \tag{8}$$

$p, q = 0, 1, \ldots$

$$m_{00} = \sum_x \sum_y f(x, y) \tag{9}$$

represents the total mass.

$$m_{10} = \sum_x \sum_y x \, f(x, y) \tag{10a}$$

$$m_{01} = \sum_{x}\sum_{y} y\, f(x, y) \tag{10b}$$

are used in determining the centroid, $(\bar{x}, \bar{y})$, or center of mass, since

$$x = \frac{m_{10}}{m_{00}}, \quad y = \frac{m_{01}}{m_{00}} \tag{11}$$

The central moments, m_{pq}, are defined by:

$$\mu_{pq} = \sum_{x}\sum_{y} (x - \bar{x})^{p} (y - \bar{y})^{q} f(x, y)$$

where $p, q = 0, 1, \ldots$

It is easily shown that the central moments are translation invariant. It is also possible to normalize the central moments to obtain invariance to a change in size.

For this study, only raw moments were used, since it was determined that translation invariance was not desirable, but size invariance was necessary. The moment computation has been initially limited to the set of first order moments. The use of measurements based only on the histogram moments negate to a great extent contextual structural anatomical difference between the breasts. The initial use of x and y centroids for each of the 30 gray levels in the horizontal (X) and vertical (Y) planes have been calculated for all the films in the test group. The X and Y centroids were computed in the following manner

$$\bar{X}(t) = \frac{\sum_{y=1}^{M} \sum_{x=1}^{N} x\, f_t(x, y)}{h(t)} \tag{12}$$

$t=1,\ldots,30$

$$\bar{Y}(t) = \frac{\sum_{y=1}^{M}\sum_{x=1}^{N} y\, f_t(x,y)}{h(t)} \qquad (13)$$

where h(t) is the histogram computed previously and M and N represent the dimension of a rectangular region containing the breasts.

Figure 2 indicates the pattern exhibited by these centroids for an ideal image consisting of a diagonal pattern of temperatures with decreasing value from upper left to lower right. The centroid pairs $(\bar{X}(t), Y(t))$ corresponding to each of t=10 temperature "points" to the area of the field where the particular temperature (t) as a centroid. Within the context of these measurements one would expect that asymmetry of either breast would be detected by a noted asymmetry in $\bar{X}(t)$. In a similar manner the $\bar{Y}(t)$ might locate the lesion vertically. Figure 1 illustrates the use of the centroid measures on the test case. Note that the $\bar{X}(t)$ centroids of higher temperatures are heavily biased to the right indicating an abnormality of the left breast. The $\bar{Y}(t)$ centroids while usually not as specific is also often useful in locating the lesion. Several possibilities exist to further reduce this centroid information to a useable form for computer classification. One approach may be to view the $(\bar{X}(t), \bar{Y}(t))$ centroids within the context of the three dimensional (x, y, t) plane. Within this context, completely temperative symmetric breasts would present a linear contour of minimum length orthogonal to the X, Y plane. Any deviation from this ideal would both lengthen the contour and increase the variance from the ideal. In either case if random fluctuations are small with respect to deterministic changes caused by focal or diffuse temperature asymmetry either parameter should indicate it.

Figure 3 illustrates a benign case with the previously described features extracted. It should be noted that asymmetry of the horizontal signature and $\bar{Y}(t)$ is not present. There also appears little difference in the temperature histograms of the breast.

References

1. Stockham, T. G., Image Processing in the context of a visual model, Proc. IEEE 60:828-842, 1972.

2. Cohen-Gershon, J., "Medical Thermography", Scientific American, Vol. 216, No. 2, Feb. 1967, pg. 94.

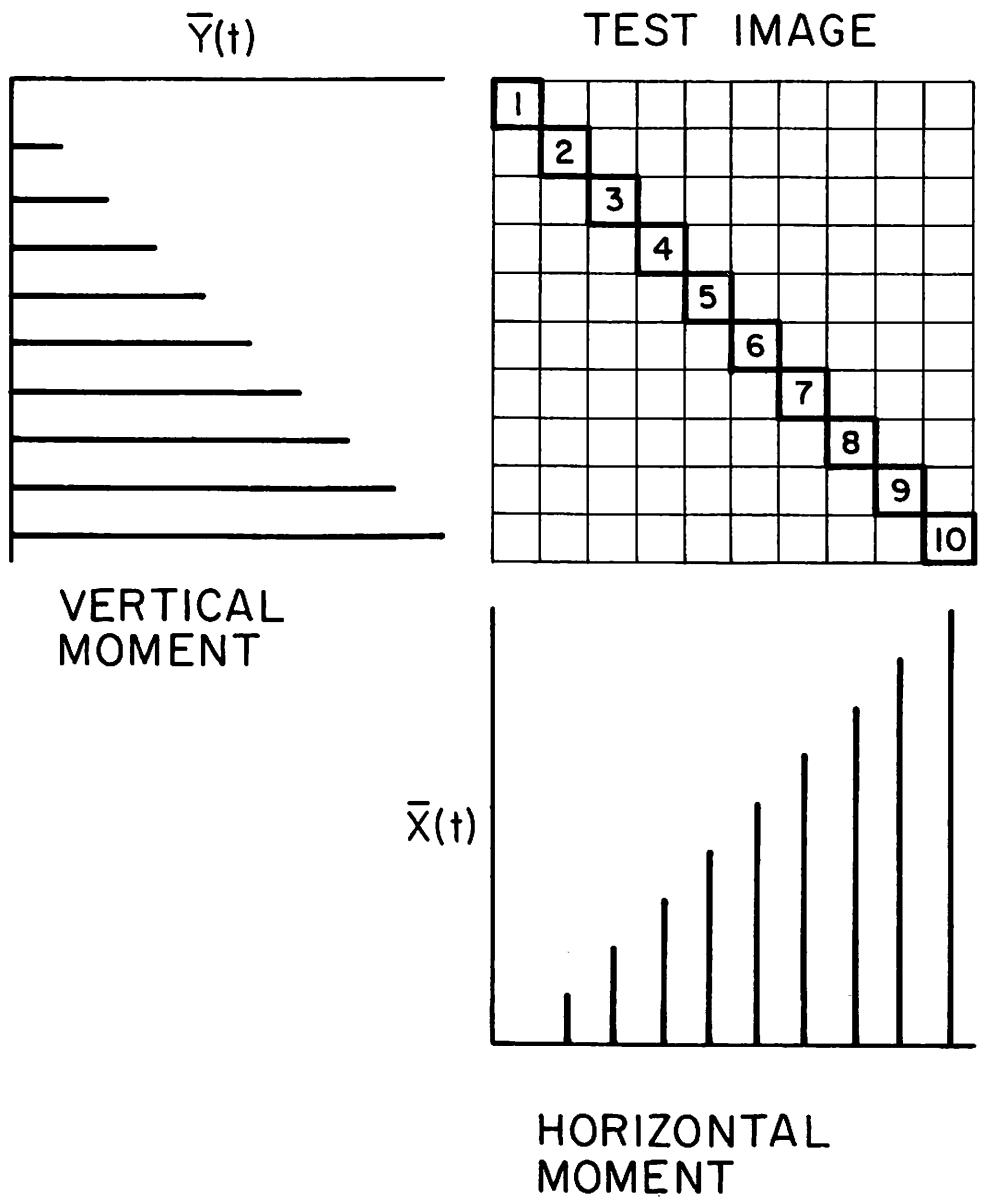
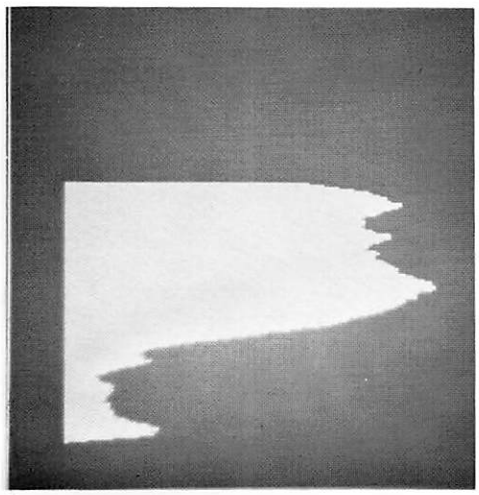
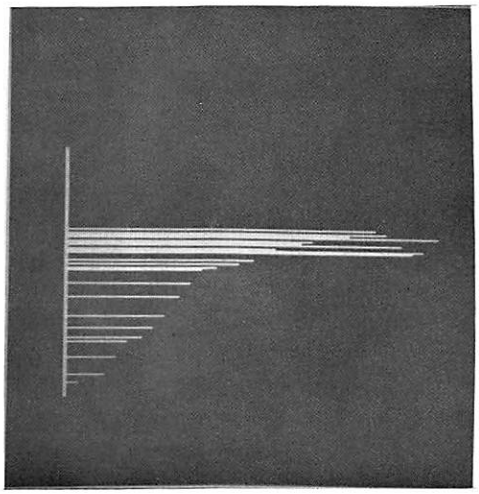


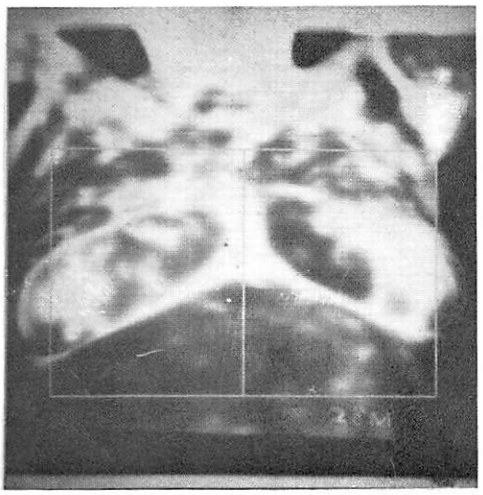
Figure 5.1-2. Simulated image.



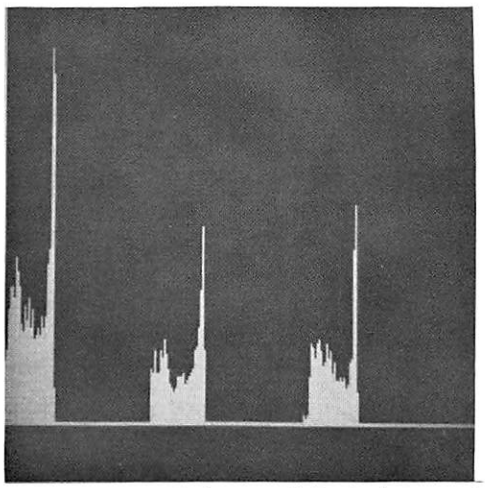
Vertical signature



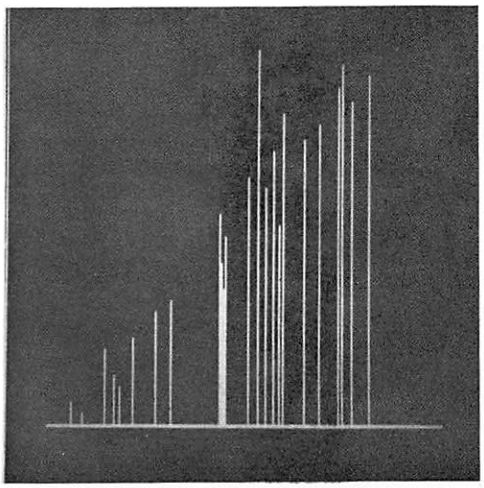
Vertical moment



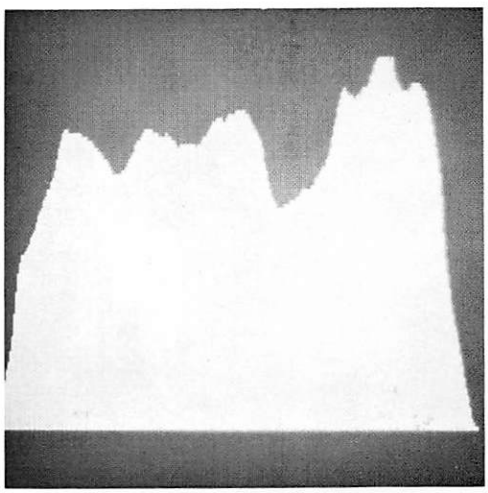
Digitized abnormal thermogram



Total, right and left breast histograms respectively



Horizontal moment



Horizonal signature

Figure 5.1-1. Abnormal thermogram.

3. Sprawls, P., "Digital Computer Interpretation of Radioisotope Distribution Patterns," Journal of Nuclear Medicine, Vol. 10, pg. 618, 1969.

4. Kruger, R. P., et al, "Automated Radiographic Diagnosis via Feature Extraction and Classification of Cardiac Size and Shape Descriptors", IEEE Transactions on Bio-Medical Engr. Vol. BME-19, No. 3, May, 1972, page 174.

5. Ausherman, D. A., "Extraction of Corrected Edges from Knee Radiograph," IEEE Transactions on Computers, Vol. C-21, No. 7, July 1972, pg. 753.

6. Deutsch, J. A., "A Theory of Shape Recognition," Brit. J. Psychol. XLVI, 1955, pg. 30.

7. Hu, M., "Visual Pattern Recognition by Moment Invariants," IRE Transactions on Info. Thy. Feb., 1962, pg. 179.

8. Duda, R. O., and Hart, P.E., "Pattern Classification and Scene Analysis," Wiley, New York, New York, pg. 364-366, 1973.

6. Image Analysis Projects

The image analysis projects focus on distinctive aspects of image information that are essential in the design of image coding and processing systems. In particular, quantitative image error measures are investigated considering some properties of the human visual system.

In the first report, a new model of color vision is presented, in which the color representation is more similar to the actual perceptual quantities than in conventional colorimetric representations. The effectiveness of color image processing may be improved if the operations or fidelity measurements are performed in the "perceptual" space defined by the model.

In the next section, the problem of generating constant brightness surfaces in color perception space is considered. A recursive algorithm coupled with a transversality condition is developed to improve computational accuracy.

6.1 Modelling Color Vision for Psychovisual Image Processing

Werner Frei

Color image coding and processing can take advantage of some properties of human vision which are not described by the conventional colorimetric color representations. After reviewing the Young-Helmholtz theory of vision, a simple non-linear model is presented, based upon physiological evidence and postulates currently used in vision research. This new model is intended to predict the colored sensation produced by picture elements as it is influenced by the surrounding image features.

Conventional color representations: the Young-Helmholtz theory of color vision. The fact that any colored light can be matched by superimposing three appropriately chosen colored lights forms the basis of the Young-Helmholtz theory of color vision. That theory postulates the existence of

three types of photoreceptors in the retina, whose spectral energy absorbtions are linearly independent. Accordingly, the perception of a colored light C characterized by its spectral energy distribution $C(\lambda)$ depends on three quantities x_i, $i = 1...3$, namely the respective amounts of energy absorbed by the receptors

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

where $t_i(\lambda)$ is the spectral energy absorbtion of the i-th receptor and λ_L and λ_U are the lower and upper limits of the visible spectrum.

Consequently, two colored lights C_1 and C_2 seen under the same viewing conditions (such as visual angle, background, previous adaptation) will match if the amounts of energy absorbed x_i are respectively equal, regardless of the particular spectral distributions $C(\lambda)$.

Colors can therefore be represented by vectors in the positive octant of a three dimensional space defined by the spectral absorbtion functions of the receptors, and Grassmann's experimental laws of color mixture can be expressed by vector addition in that space or any nonsingular linear transformation thereof. The basis vectors of such a space are called primaries, and the vector components tristimulus values, normalized with respect to an arbitrary "reference white". Tristimulus spaces are currently used for "objective" measurements of colors (referred to a well defined "standard observor") as well as for color representation in conventional color television [1] and color image processing [2],

Uniform color spaces. Unfortunately, the perception of colors is not a linear function of the colorimetric tristimuli, as is shown by the variances of color matching experiments[3] orWeber's law of luminous sensation [4]. Error criteria like the mean square error defined in tristimulus spaces are therefore of limited relevance to the human observor.

Various nonlinear transforms have been proposed which map the tristimulus spaces into approximately "uniform perception" representation spaces like the $U^* V^* W^*$ space, the cube root coordinate system, etc. Such spaces should be used for quantizing color information to obtain uniform perceptual increments in the discrete color representation space.

Spatial effects. The Young-Helmholtz theory predicts colored sensations under well defined viewing conditions only. In actual images however, the perceived color of picture elements is likely to be altered more or less by the surrounding elements due to a number of perceptual phenomena. For example, very small colored areas seen against a neutral background look colorless; a small colorless area seen against a colored background seems to have a hue complementary to that of the background; colored objects seen under a chromatic illumination appear to retain their original color, to a certain extent. Such effects are believed to be due to lateral neural interactions between the outputs of the photoreceptors in the retina, or to an unequal loss of sensitivity (bleaching) of the three types of receptors, or both [5]. Our present knowledge of the structure of the retina suggests models of color vision which predict these effects in terms of spatial frequencies.

First order model of spatial color vision. Figure 1 contains a psychophysical model for color vision. In the model it is assumed that:

(1) The retina contains three types of photo-receptors whose spectral energy absorbtion functions are linearly independent [6, 7];

(2) The neural outputs of the receptors are approximately logarithmic functions of the respective amounts of absorbed light energy [8];

(3) The spatial neural interactions between the outputs of the receptors are approximately linear summations; the interactions can therefore be represented at that stage by linear spatial filters [9].

The first stage of the model describes the respective energy absorbtions in the three types of photoreceptors according to the Young-Helmholtz theory of color vision. Let $\vec{c}$ be the spectral energy distribution of light evaluated at n wavelengths λ_i, i = i... n. The components of $\vec{t}$ are the

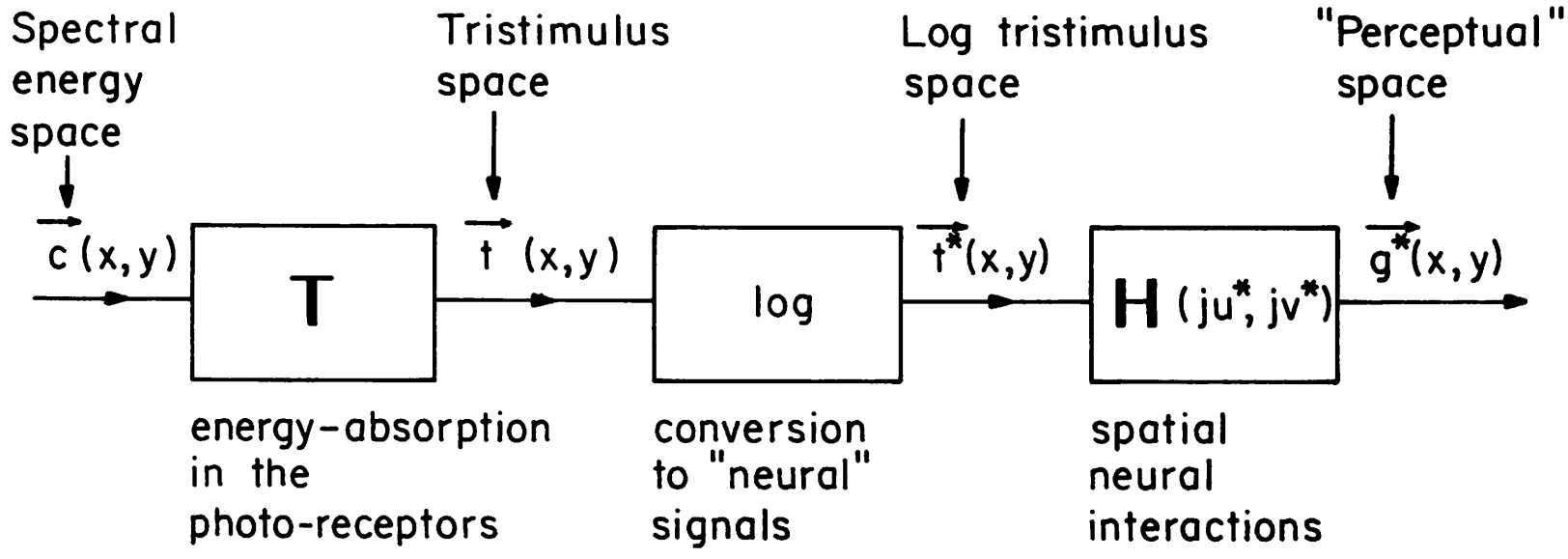


Figure 6.1-1. Model of color vision for psycho-visual image processing.

tristimulus values of $\vec{c}$ referred to a yet unknown set of primaries and reference white defined by the spectral absorbtions t_{ij}, $i = 1...3$, $j = 1...n$ of the receptors. In matrix notation

$$\vec{t} = T\,\vec{c} \qquad (2)$$

$$\vec{c} = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_u \end{bmatrix} \quad ; c_i \geq 0 \quad ; \quad \vec{t} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \quad ; \quad t_i \geq 0$$

$$T = \begin{bmatrix} t_{11}, & t_{12}, & \cdots & t_{1n} \\ t_{21}, & t_{22}, & \cdots & t_{2n} \\ t_{31}, & t_{32}, & \cdots & t_{3n} \end{bmatrix} \quad ; \; t_{ij} \geq 0$$

where $\vec{c}$ and $\vec{t}$ are functions of the spatial image coordinates x and y; assuming shift-invariance, T is constant. For the model to be consistent with Grassmann's laws of color mixture, T must be a linear transformation of the CIE - $\bar{r}_\lambda$, $\bar{g}_\lambda$, $\bar{b}_\lambda$ or $\bar{x}_\lambda$, $\bar{y}_\lambda$, $\bar{z}_\lambda$ color mixture data for spectral colors. That is,

$$T = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \bar{r}_{\lambda 1}, & \bar{r}_{\lambda 2}, & \cdots, & \bar{r}_{\lambda n} \\ \bar{g}_{\lambda 1}, & \bar{g}_{\lambda 2}, & \cdots, & \bar{g}_{\lambda n} \\ \bar{b}_{\lambda 1}, & \bar{b}_{\lambda 2}, & \cdots & \bar{b}_{\lambda n} \end{bmatrix} \qquad (3)$$

The second stage of the model represents the logarithmic conversion that is assumed to occur in the receptors between absorbed light energies t_i and output neural signals t_i^*. (The superscript star denotes the logarithmic domain.) This stage imposes a condition on the spectral absorbtion function of one type of receptor, because Abney's law states that the luminance

of a color mixture is equal to the linear sum of the luminances of the mixture's components.

Let $\vec{v}_\lambda^T$ represent the relative luminous efficiency of equal energy spectral lights evaluated a wavelengths λ_i , i = 1...n. The luminance L_c of $\vec{c}$ is given by

$$\begin{aligned} L_c &= k \vec{v}_\lambda^T \vec{c} \\ \vec{v}_\lambda^T &= [v_{\lambda 1}, v_{\lambda 2}, \cdots, v_{\lambda n}] \quad ; \quad 0 \leq v_{\lambda i} \leq 1 \end{aligned} \tag{4}$$

where k is a constant. The linear additivity of luminances can only be satisfied here if one of the receptor types exhibits a spectral absorbtion proportional to the relative luminous efficiency function, for example

$$t_{li} \sim v_{\lambda i} \tag{5}$$

The third stage of the model represents the spatial neural interactions between the receptor outputs. It has been shown elsewhere[5] that linear interactions can be modeled by spatial filters. Let the components of $\vec{g}^*$ be the "perceptual" variables of interest, and H^* denote a 3x3 matrix of spatial filters h_{ij}^* (ju^*, jv^*) where u^* and v^* are the spatial frequencies in the logarithmic domain . Then

$$\vec{g}^* = H^* \vec{t}^* \tag{6}$$

Abney's law (see condition (4) above) imposes that one of the perceptual variables g_i^* has to be a function of the "luminance" receptors only. With eq.(5) therefore

$$h_{12}^* = h_{13}^* = 0; \quad h_{11}^* = h_1^* \tag{7}$$

The perceptual luminance variable $g_1^* = h_1^* t_1^* = h_1^* \log t_1$ satisfies Weber's law of luminous sensation.

The perception of chromaticity is a function of the ratios of the energies absorbed in the different types of receptors. Therefore we can define two "perceptual" chromaticity variables g_2^* and g_3^* by simply forming differences in the logarithmic domain.

$$H^* = \begin{bmatrix} h_1^* & 0 & 0 \\ -h_{21}^* & +h_{22}^* & 0 \\ -h_{31}^* & 0 & +h_{33}^* \end{bmatrix} \tag{8}$$

The perceptual quantities g_i^* are now

$$g_1^* = h_1^* \log (t_1) \tag{9a}$$

$$g_2^* = h_{21}^* \log \frac{(t_2)^{h_{22}^*/h_{21}^*}}{t_1} \tag{9b}$$

$$g_3^* = h_{31}^* \log \frac{(t_3)^{h_{33}^*/h_{31}^*}}{t_1} \tag{9c}$$

Note that if t_1 corresponds to a "green" receptor, t_2 to a "red" and t_3 to a "blue" one, g_2^* turns out to be a protanopic chrominance quantity and g_3^* a tritanopic one, which brings this model close to explaining defective color vision. In fact, one may use available data on color blindness to determine t_{2i} and t_{3i} in eq.(2), and hence the model's tristimulus space t [10]. Assuming that the hues of tritanopic and deuteranopic color pairs do not shift as a function of spatial frequency (no such shifts have been reported), one can write

$$h_{22}^{*} = k_2 h_{21}^{*} = k_2 h_2^{*}$$

$$h_{33}^{*} = k_3 h_{31}^{*} = k_3 h_3^{*} \qquad (10)$$

where k_2 and k_3 are constants and

$$g_2^{*} = h_2^{*} \log \frac{(t_2)^{k_2}}{t_1}$$

$$g_3^{*} = h_3^{*} \log \frac{(t_3)^{k_3}}{t_1}$$

It is interesting to compare the perceptual variables g_2^{*} and g_3^{*} defined above with the chrominance signals used in European color television systems. Figure 2 shows a linear quantization of the normalized European chrominance signals plotted in MacAdams geodesic chromaticity diagram[11]. These signals are defined as follows

$$c_1 = \frac{R - Y}{Y}$$

$$c_2 = \frac{B - Y}{Y}$$

where R, B, Y are the FFC red, blue, and luminance primaries respectively. The logarithmic like distribution of the quantum steps is clearly visible in figure 2. Slightly different primaries are of course required for the model because the television primaries are negative outside the FFC-RGB triangle.

The spatial filter functions h_1^{*}, h_2^{*} and h_3^{*}. Models essentially similar to the one presented here have been proposed for achromatic vision [5] and used for black and white image coding and processing[12]. Since this model reduces to Stockham's visual model in the case of achromatic vision, his results can

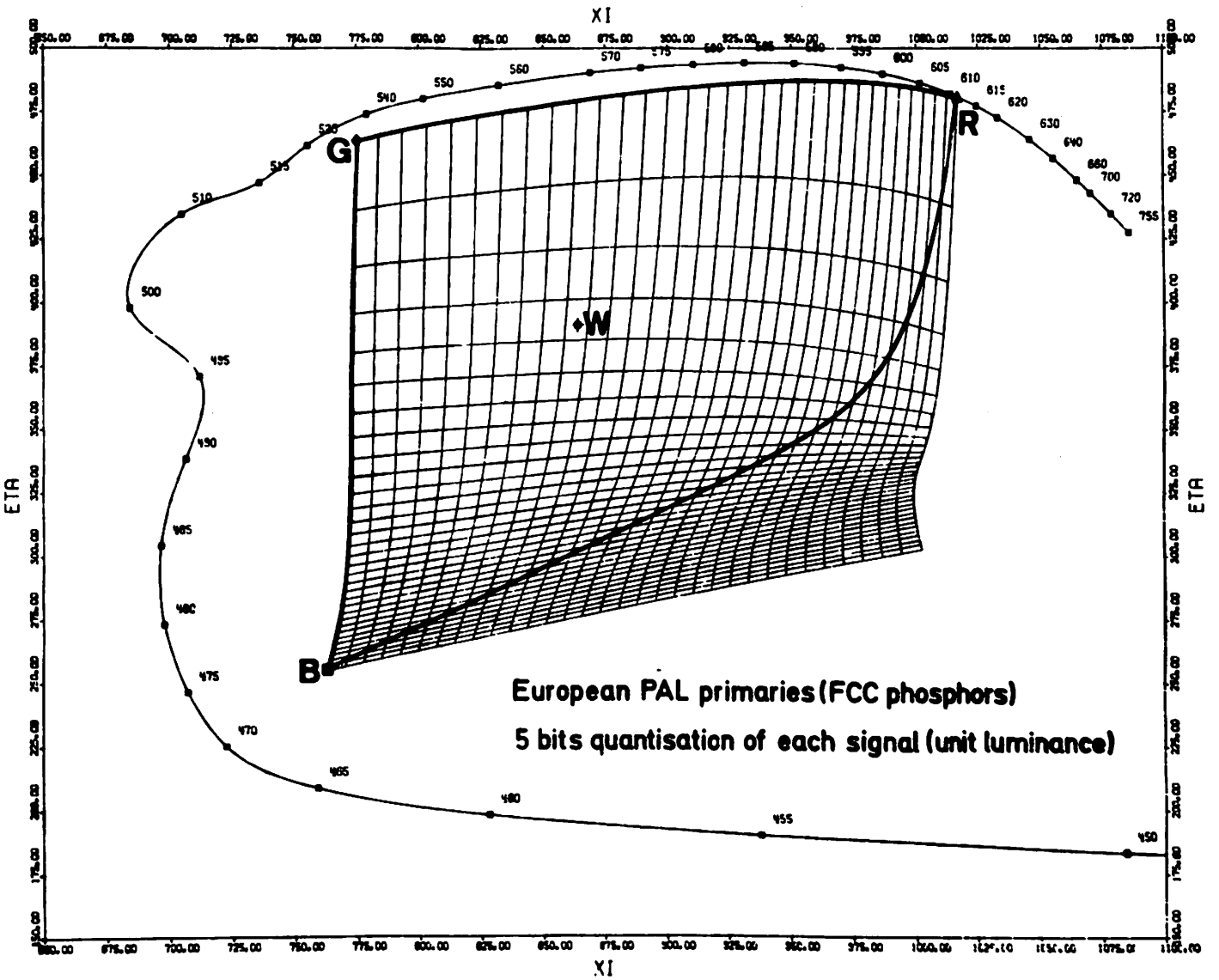


Figure 6.1-2. Linear quantization of television chrominance signals represented in geodesic chromaticity diagram.

be used to evaluate $h_1^*(ju^*, jv^*)$.

Since the resolution of the visual system at high spatial frequencies is maximum for luminance differences [13], the bandwidths of h_2^* and h_3^* are necessarily smaller than the bandwidth of h_1^*. Measurements of threshold noticeable chromaticity contrast using square-wave gratings [14] suggest that h_2^* and h_3^* should exhibit an increasing attenuation for spatial frequencies exceeding 1-2 cycles per degree of visual angle. On the other hand, the question whether the visual system actually attenuates low chromaticity spatial frequencies is still controversial [15, 16]. A low frequency attenuation of h_2^* and h_3^* would enable the model to predict simultaneous color contrast as well as chromatic adaptation according to the VonKries theory. Studies are under way to determine the filter functions to be used in this model.

Conclusion. A model of color vision for the primary purpose of image processing is presented. It maps the spatial spectral energy distribution of light into an approximately uniform perceptual representation space, taking spatial interactions into account. The nonlinear transformation of tristimulus space defined by the two latter stages of the model may be used for establishing error criteria more relevant to the human observer, as well as for problems involving limited dynamic ranges of display devices. The multiplicative structure of the model further suggests that homomorphic filtering [17] may be most appropriate for coding and enhancing color pictures. Another possible application lies in the design of color display devices for optimum human recognition.

References

1. W. T. Wrintingham, "Color Television and Colorimetry," Proc. IRE, Vol. 39, October 1957, pp. 1135 - 1172.
2. W. K. Pratt, "Digital Color Image Coding and Transmission, "USCEE, Rept. 403, June, 1971.
3. D. L. MacAdam, "Visual Sensitivities to Color Differences in Daylight," Journal Optical Society of America, Vol. 32, May 1942, pp. 247 - 274.
4. G. Wyszecky & W. S. Styles, Color Science, John Wiley and Sons, New York, 1967.
5. T. N. Cornsweet, Visual Perception, Academic Press, New York, 1970.

6. E. F. MacNichol Jr., "Three Pigment Color Vision", Scientific American, Vol. 211, 1964, pp. 48 - 56.

7. G. Wald, "The Receptors of Human Color Vision", Science, Vol. 145, 1964, pp. 1007 - 1016.

8. H. K. Harline & C. H. Graham, "Nerve Impulses from Single Receptors in the Eye", Journal Cell. Comp. Physiology, Vol. 1, 1932, pp. 277 - 295.

9. F. Ratliff & al., "Spatial and Temporal Aspects of Retinal Inhibitory Interactions," Journal Optical Society of America, Vol. 53, 1959, pp. 110 - 120.

10. Y. LeGrand, Light, Color, and Vision, Chapman and Hall, London, 1968.

11. W. Frei, P. A. Jaeger, P. A. Probst, "Quantisierung der Farbinformation bei der Bildcodierung, Nachrichtechnische Zeitung, Vol. 25, 1972, pp. 401 - 404.

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

13. P. K. Kaiser, "Brightness Difference and its Relation to the Distinctness of a Border," Journal Optical Society of America, Vol. 61, July, 1971, pp. 962 - 965.

14. G. J. Van der Horst & al., "Transfer of Spatial Chromaticity-Contrast at Threshold in the Human Eye," Journal Optical Society of America, Vol. 57, October 1967, pp. 1260 - 1266.

15. O. H. Schade, Journal Society of Motion Picture and Television Engineers, Vol. 67, 1958, p. 801.

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

17. A. V. Oppenheim & al, "Non-linear Filtering of Multiplied and Convolved Signals", Proc. IEEE, Vol. 56, August 1968, pp. 1264 - 1291.

6.2 Constant Brightness Surfaces in Color Perception Space

Anil K. Jain

In the theory of color vision a constant brightness surface is such that any two colors lying on this surface (in the color perception space) appear equally bright (not necessarily indistinguishable). Any curve that lies on this surface is called a constant brightness geodesic. If C_1 and C_2 are any two colors on this geodesic with coordinates (x_1, y_1, Y_1) and (x_2, y_2, Y_2)

in the CIE color system; then for fixed (x_1, y_1, Y_1, x_2, y_2), C_2 is such that the color distance between C_1 and any other color (x_2, y_2, Y) is minimum when Y equals Y_2. Thus, among all geodesics originating from C_1 and terminating at the straight line through (x_2, y_2), parallel to the Y axis, the color distance along the constant brightness geodesic is minimum. This minimum color distance property could be utilized in color gamut correction applications and in production of equally bright pseudocolor displays of image features.

For example, suppose the triangle R-G-B shown in figure 1 represents the color gamut obtainable with a set of primary sources P_1, P_2, P_3. If the colors in this triangle are to be reproduced with a different set of primary sources P_1', P_2', P_3' having a color gamut R'-G'-B', then all the colors $\{C\}$ lying outside R-G-B should be reproduced by colors such as $\{C'\}$ so that the color shift (color distance) between C and C' is minimum. Clearly, this can be achieved by locating C' on the intersection of the constant brightness geodesic through C and the color gamut triangle R'-G'-B'.

The problem of generating these curves is considered here. It is known (and can be proven) that these curves are horizontal at any gray. Theoretically, this information and the minimum property of these geodesics is sufficient to compute them. However, the accuracy of the computation can be improved (or checked) if one could verify some of their properties. It is shown [in details in ref [1]] that a transversality condition can be derived and appended to the algorithm to achieve higher computational accuracy.

Assuming the color space metric is given in the form [2]

$$ds = \left(\sum_{i,j} C_{ij} dx_i dx_j \right)^{\frac{1}{2}} = F(x_i, dx_i), \qquad (1)$$

it can be shown [1] that the constant brightness curves are solutions of

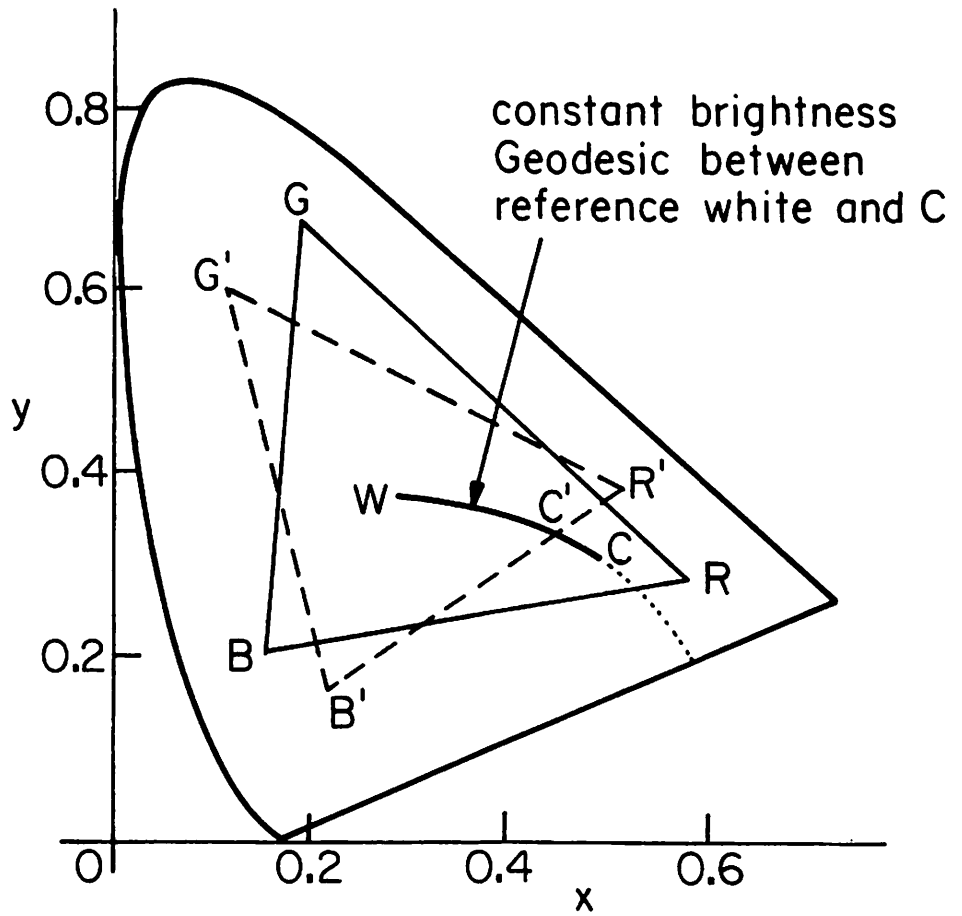


Figure 6.2-1. Color gamut corrections along constant brightness geodesics.

$$J = \min_{x_i(t),\, u_i(t),\, t_f} \int_{t_v}^{t_f} F(x_i, U_i, t)\, dt \qquad (2)$$

where $u_i(t) \equiv \dfrac{dx_i}{dt}$

t = parameter along the constant brightness curves

t_f = terminal value of the parameter

x_i = tristimulus values $i=1, 2, 3$ in the chosen color coordinate system.

Furthermore, the quality t_f must satisfy the condition

$$\left[F + \sum_{i=1}^{2} F_{u_i}(a_i - u_i)\right]_{t=t_f} = 0 \qquad (3)$$

where $F_{u_i} = \dfrac{\partial F}{\partial u_i}$

and $a_i = \dfrac{dx_i}{dt_f}$

Equation (3) gives the desired condition. Details of this derivation may be found in reference [1].

References

1. A. K. Jain, "Computational Considerations in Generation of Geodesics in Color Space," Optical Society of America, Annual Meeting, Rochester, New York, October, 1973.
2. A. K. Jain, "Color Distance and Geodesics in Color 3 Space," Journal Optical Society of America, Vol. 62, November, 1972, pp. 1287 - 1291.
3. A. K. Jain, "Role of Geodesics in Schrodinger's Theory of Color Vision," Journal Optical Society of America, Vol. 63, August, 1973, pp. 934 - 939.
4. K. D. Chickering, Journal Optical Society of America, Vol. 57, 1967, p. 537.

7. Image Processing Support Projects

The scope of the image processing support projects encompasses hardware and software developments for the needs of the image processint research effort.

The first report discusses the developments in image processing support software, where the effort has been directed towards improving the time sharing and text editing capabilities, as well as simplifying the use of the VICAR image processing language.

The next section reports the progress of the construction of a real time digital color image display system for use on the ARPANET. That unit is near completion, awaiting the delivery of memory components. Extended capabilities are specified for a more sophisticated version of the display system, whose initial design has been started.

7.1 USC/ARPANET Image Processing System

James Pepin

The past six months have been spent in refining support programs for network users, starting to implement the new FTP (File Transfer Protocol) and telnet protocols and investigating the facilities and requirements for a front-end system for image processing.

The first area of concern has been to improve and increase support of users on the ARPA net. This increased support has included such items as improvement in text editing capabilities and general improvements in the time sharing system. In the last few months IPI has begun to transfer large volumes of image data to Carnegie Mellon. This has resulted in the need for an online FTP program, which is now being implemented.

Work on the VICAR image processing software has been continued. Changes have been made to VICAR to allow users to run with less knowledge of the I/O system of the executing machine. New applications programs to perform filtering and image transformation have been implemented.

Also, the method which VICAR uses to create jobs has been more closely aligned with the method employed by O. S. (system on 370 /158). This will make the maintenance of VICAR on the two systems much easier.

The staff has been participating in the definition of the new FTP protocol by attending the FTP meeting in May and working with members of the FTP working group. After the protocol was defined, work began on converting existing FTP routines to run using the new, slightly changed syntax. During conversion to this protocol the FTP will be integrated into the system. This will allow FTP users to access the machine all the time, instead of by request.

The telnet protocol was recently redefined. This has required programming changes and these are proceeding now. Also the NCP has been modified to handle, in a sensible way, some of the error conditions that seem to be present in the net.

In July a meeting was held at USC to discuss processing requirements by the image processing users of the ARPA net. As a result of this meeting, committies were formed to investigate the possible methods of achieving resource sharing on the ARPA net. These committees will investigate the hardware and software requirements for the user community.

7.2 Development of Real Time ARPANET Image Display

John E. Tahl and Toyone Mayeda

An inexpensive digital image display for use on the ARPANET is under development and is presently in the check out phase. The development of the high resolution printer/scanner has been delayed, but the basic design and specification goals of the digital image display have not been changed and are listed as follows:

1. Receive, from the ARPANET TIP, digital picture information with brightness resolution up to 64 levels (6 bits) and at input rates up to 19.2 K baud.

2. Store the received data in an array of up to 256 x 256 six bit

picture values.

3. Present the received and stored information at standard television rates, to high speed output digital to analog converters, for application to the red, green and blue inputs of a color monitor. In this mode, the image presentation will be shades of gray. At a 19.2 K band input rate, a complete 256 x 256 image can be received and displayed within approximately 18 seconds.

4. In addition to the black and white display, the input will have full pseudo-coloring capability, using a high speed random access memory inserted between the output of the refresh memory and the digital to analog converters. The random access memory will be remotely programmed from the TIP or by local switch control. Over 4096 different color combinations of hue, saturation and luminance will be available for pseudo-coloring.

5. The displayed image will occupy only the center portion of the monitor, to provide a pleasing presentation without a blocked appearance.

A block diagram of the digital image display is shown in figure 1. All major components for this unit have been received -- with the exception of the refresh memory integrated circuits. All logic panels have been wire-listed and wire wrapped. Checkout of the unit has been accomplished as far as possible without the refresh memory section. No major checkout problems are anticipated after the refresh memory integrated circuits are received.

Initial design of a more sophisticated digital image display has been started. This second unit will have the following capabilities in addition to the ones of the first unit:

1. Accept an 8 bit image, or a 7 bit image and a 1 bit graphic overlay.

2. Include a function memory which can be used to translate the 8 bit data (from the refresh memory) with any desired transfer curve to produce 6 bit data.

3. An alphanumeric keyboard will be used to communicate with the ARPANET to receive image data, and will also be used to generate alphanumeric characters on the display.

4. The output video data and alphanumeric characters will be in composite RF format so that it can be displayed on any TV receiver using its antenna input.

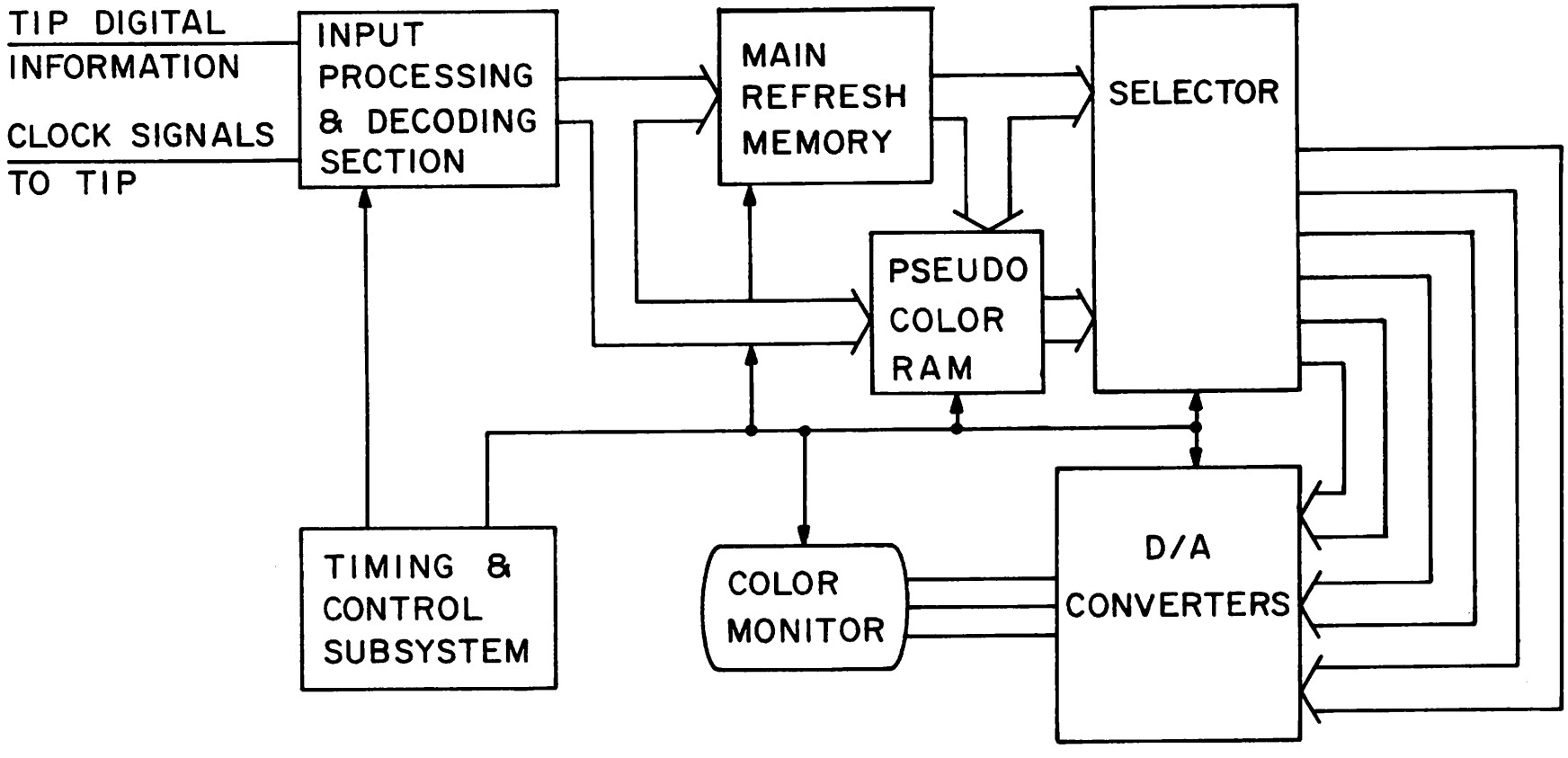


Figure 7.2-1. ARPANET digital image display.

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.

Andrews, H. C., "Digital Fourier Transforms as a Means for Scanner Evaluation," accepted for publication in Applied Optics, January 1974.

Chen, C. K. and Andrews, H. C., "Nonlinear Intrinsic Dimensionality Computations," accepted for publication in IEEE Transactions on Computers, December 1973.

Chen, W. and Pratt, W. K., "Color Image Coding with the Slant Transform," Symposium on Applications of Walsh Functions, Washington, D. C., April, 1973.

Crawford, W. O. and Hall, E. L., "Computer Classification of Radiographs of Coal workers Pneumoconiosis," Radiological Society of North America, November 1973.

Davisson, Lee D., "Universal Noiseless Coding," IEEE Transactions on Information Theory, November 1973.

Habibi, A., "A Cascade of Unitary Transformations and DPCM Systems for Coding Pictorial Data," Proceedings of Applications of Walsh Functions, April 1973.

Habibi, A., "A Cascade of Unitary Transformations and DPCM Systems for Coding Pictorial Data," to appear in IEEE Transactions on Communication, publication date unscheduled.

Habibi, A., "Application of Lower-Triangular Transformations in Coding and Restoration of Two-Dimensional Sources," Proceedings of 1973 National Telecommunication Conference, November 1973.

Hall, E. L., "Almost Uniform Distributions for Computer Image Enhancement," Accepted for publication in IEEETC.

Hall, E. L. and Crawford, W. O., "Automated Classification of Profusion of Opacities in Radiographs of Coal Workers," 26th Annual Conference on Engineering in Medicine and Biology, September 1973.

Hall, E. L., Crawford, W. O., Preston, K., and Roberts, F. E., "Classification of Profusion of Black Lung Disease from Chest X-rays," First International Conference on Pattern Recognition, October 1973.

Jain, Anil K., "Role of Geodesics in Schrodinger's Theory of Color Vision", JOSA, August 1973, Volume 63, Number 8, pages 934-939.

Jain, Anil K., "Image Modelling for Unification of Transform and DPCM Coding of Two-Dimensional Images", National Electronics Conference, Chicago, October 1973.

Jain, Anil K., "Computational Considerations in Generation of Geodesics in Color Space", Optical Society of America Annual Meeting, Rochester, New York, October 1973.

Kruger, R. P., Hall, E. L., and Turner, A. F., "A Prototype Radiographic System for Pneumoconiosis Chest X-Ray Screening" IEEE Conference on Systems Man and Cybernetics Proceeding, Boston, Massachusetts, November 1973.

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

Kruger, R. P., Winsor, T., Winsor, D., and Birsner, J. W., "A Quantitative Computer Approach to Thermography," American Thermographic Society Conference Proceedings, New York City, New York, June 1973.

Nahi, N. E. and Habibi, A., "Nonlinear Adaptive Recursive Image Enhancement," Proceedings of 1973 Joint International Conference on Pattern Recognition, October 1973.

Nahi, N. and Franco, C. A., "Recursive Image Enhancement - Vector Processing," IEEE Transactions on Communications, April 1973.

Pratt, W. K., "Digital Transforms," 1973 IEEE NEREM Signal Processing Seminar," Boston, Massachusetts, November 1973.

Sawchuck, A. A., "Digital Image Restoration and Enhancement: Space Variant and Invariant Processing," Proceedings of the Society of Photographic Scientists and Engineers 1973 Annual Conference, Rochester, New York, May, 1973.

Sawchuck, A. A., "Space - Variant System Analysis of Image Motion," Journal of the Optical Society of America, 63, page 1052, 1973.

Winter, J., Stein, M. A., and Kruger, R. P., "Automated Breast Thermography Diagnosis by Computer," American Thermographic Society Conference Proceedings, New York City, New York, June 1973.