USCEE REPORT #425

Semiannual Technical Report
Covering Research Activity During the Period
1 March 1972 to 31 August 1972

by

William K. Pratt

August 1972

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 1972 to 31 August 1972 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. Four 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 Detection and Measurement Projects: the recognition of objects within pictures and quantitative measurement of image features; (4) Image Processing Support Projects, a description of the USC image processing facilities accessible over the ARPANET.

PROJECTS PARTICIPANTS

Project Director

William K. Pratt

Research Faculty

Harry C. Andrews
Lee D. Davisson
Ali Habibi
Anil K. Jain
Richard P. Kruger
Nasser Nahi
Alexander A. Sawchuk
Lloyd R. Welch

Support Staff

Angus B. Cossey
Carolyn Matthews
James M. Pepin
Mark A. Sanders
Michael Reichik
Linda M. Webster

Students

Tooraj Assefi
Wen-Hsiung Chen
Faramarz Dauarian
Roy M. Glantz
Michael Huhns
Mohammed Jahanashahi
I. David Levy
Clanton Mancill
Nelson Mascarenhas
Manuher Naraghi
Sumeet Pasricha
Clifford Reader
Stuart Robinson
Dennis Smith
Edward Sutton
Fred Tydeman
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 1972 to 31 August 1972.

The image processing research study has been subdivided into five projects:

Image analysis projects
Image coding projects
Image restoration and enhancement projects
Image detection and measurement projects
Image processing support projects

The image analysis project comprises the background research effort into the basic structure of images in order to develop meaningful quantitative characterizations of an image. 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 detection and measurement projects are the registration of images, detection of objects within pictures and measurements of image features. 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 6 describe the research effort on the projects listed above during the reporting period. Section 7 contains a short description of new projects that are being initiated, and are not yet to the reporting stage. Section 8 is a list of publications by project members.

2. Research Project Activities

The following sections describe some of the significant project activities of the past six months.

Publications. July, 1972 has been a banner month for publications in the field of image processing. The Institute of Electrical and Electronic Engineers published a special issue of the Proceedings of the IEEE on digital image processing. Professor Harry C. Andrews of USC acted as co-editor of the special issue. The journal contained five papers written by the USC staff. The IEEE also devoted the July issue of the Transactions on Computers to the subject of two-dimensional digital signal processing. USC staff members contributed two papers to the issue. Finally, the IEEE carried a cover article on digital image processing featuring the USC research on pseudocolor image enhancement. A reprint of this paper is included in Section 4.7.

Image Coding Conference. The University of Southern California has been chosen as the host for the Fourth Picture Coding Conference to be held on 22-24 January 1973. Approximately 80 to 100 research workers in the field of image coding are expected to attend. Topics of the conference include: properties of the human observer, facsimile coding, intraframe picture coding, interframe television coding, color image coding, and multispectral image data coding. An image coding contest will be held at the conference to determine the best image coding algorithms and systems for monochrome and color images.

3. Image Coding Projects

The goal of the image processing research study - to develop digital image coding techniques capable of representing monochrome and color images with a minimal number of code bits - is not likely to be satisfied, at least in the near term, by a single coding system. Accordingly the research effort has been directed toward a wide variety of coding methods designed for various applications. The results of the research study during the past six months are summarized here and presented in greater detail in the subsequent subsections.

A study on quantization techniques for natural color images is the subject of the first report. Using a quantitative measure of color error, previously developed, it was found that the minimum error color representation using standard PCM coding techniques is the red, green, blue representation. Other coordinate systems such as the YIQ system employed for commercial television transmission lead to significantly greater quantization error unless a complicated quantizer is employed.

The study on deltamodulation and differential PCM coders for monochrome and color images has led to worthwhile results. In particular, simulation of a dual mode coder which uses deltamodulation for low detail image regions and differential PCM for high detail image areas has indicated that subjectively good results can be obtained with about 1.5 to 2.0 bits/pixel. Such a coder could be implemented in real time with a relatively modest system

The research work on transform coding systems has continued with the development of the slant transform coder. The slant transform possesses basis functions (expansion waveforms) that more closely approximate natural imagery than other more commonly employed transforms. Analysis and supporting experimentation has indicated that the Slant transform is capable of image coding with about 1.0 to 2.0 bits/pixel. Another recent development in the application of transforms to image coding is the transform differential coding method. In this system a unitary transform is used to obtain a low detail estimate of the image. The pixel differences between the original and estimate are then coded. With this system it is

possible to code images with as low as 3.5 bits/pixel with no error or with about 1.5 to 2.0 bits/pixel with a scarcely noticeable error. It now appears that these transform coding systems can be implemented in real time using acoustic surface wave delay line implementation.

A problem common to almost all types of image coding systems, is the efficient coding of variables, e.g. image samples, differential PCM samples, or image transform coefficients. A comprehensive comparison of several quantization techniques has been completed.

In the final research study reported upon in this section, some preliminary results are presented on a new coding process called Universal Coding. In universal coding systems, coding is accomplished without complete knowledge of the source statistics, e.g. statistically coding an image without complete knowledge of the pixel correlation. Results of this theoretical investigation should be useful in establishing performance bounds on image coders operating with incomplete statistical knowledge of the image data.

3.1 Color Image Quantization

Anil K. Jain
William K. Pratt

When color images are converted to digital form for computer processing or digital communication, consideration should be given to the effects of quantization errors. Quantization errors can cause luminance, hue, and saturation shifts of a color that result in a significant image degradation. In quantization of color images, the quantization error depends on: (a) the relative number of bits assigned to the different coordinates; and (b) the distribution of the image colors in the color space corresponding to the respective color coordinate system.

Figure 1 contains a general block diagram for a color image quantization system. A source image described by source tristimulus values R, G, B is converted to three components x_1, x_2, x_3 which are then quantized. The quantized components $\hat{x}_1, \hat{x}_2, \hat{x}_3$, are then converted back to the original color coordinate system producing the tristimulus values $\hat{R}, \hat{G}, \hat{B}$.

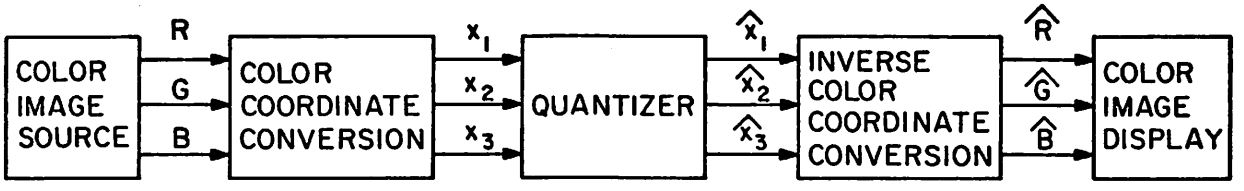


Figure 3.1-1 Color Image Quantization Model

The quantizer in Figure 1 effectively partitions the color space of the color coordinates x_1, x_2, x_3 into quantization cells and assigns a single color value to all colors with a cell. To be most efficient, the three color components x_1, x_2, x_3 should be quantized jointly. However, implementation problems often dictate separate quantization of the color components. In such a system x_1, x_2, x_3 are individually quantized over their maximum ranges. In effect, the physical color solid is enclosed in a rectangular solid which is then divided into rectangular quantization cells. Many of these quantization cells will lie outside the physical color solid, and will be wasted.

In the present analysis it will be assumed that each color component is linearly quantized over its maximum range into 2^{n_i} levels where n_i represents the number of bits assigned to the component x_i. The total number of bits allotted for the coding is fixed at

$$N = n_1 + n_2 + n_3 \tag{1}$$

Let b_i represent the upper bound of x_i and a_i the lower bound. Then each quantization cell has dimension

$$\epsilon_i = \frac{b_i - a_i}{2^{n_i}} \tag{2}$$

Thus the coordinates of the quantized color become

$$y_i = x_i \pm \epsilon_i \tag{3}$$

subject to the conditions

$$a_i \leq y_i \leq b_i \quad , \qquad i = 1, 2, 3 \tag{4}$$

Observe that the values of y always lie within the smallest rectangular solid enclosing the color solid for the given color coordinate system.

The total error due to quantization can be specified by the geodesic color distance [1] between x_i and y_i as given by

$$S = \int_{(x_i)}^{(y_i)} \left[\sum_{j,k=1}^{3} C_{jk} dx_j dx_k \right]^{\frac{1}{2}} \tag{5}$$

where the coefficients C_{jk} represent the sensitivity of average human perception of differences in the j^{th} and in the k^{th} coordinates of a color coordinate system. Using equation (5) it is possible to evaluate the total error due to a given bit assignment. Thus, for a given number of bits, N, it is possible to find the optimal bit assignment for each color. However, for the coding of color images, one would like to determine the optimal bit assignments for all the colors in a given coordinate system. In this study consideration is given to the problem of determining non-inferior bit assignments for the N.T.S.C. receiver primary major colors in the following coordinate systems: i) C.I.E. Uniform Chromaticity scale; UVW and uvY; ii) N.T.S.C. transmission system, YIQ; iii) N.T.S.C. receiver phosphor system, RGB and rgY. Choosing N = 12, equation (5) was used (via a color distance routine [1]) to determine the maximum value of S (over all possible y), for a fixed color and a fixed bit assignment n_1, n_2, n_3 such that

$$n_i \leq 6 , \quad i = 1, 2, 3 \tag{6}$$

$$N = 12 \tag{7}$$

This was repeated for the seven major colors and all possible bit assignments satisfying (6) and (7).

The bit assignments giving relatively large values of S were considered inferior and thus ignored. Table I for example, shows the non-inferior set of bit assignments for various colors and in the RGB system (see reference [2] for more results).

Quantization of a 256 by 256 color image of a girl [3], was done with bit assignments (4, 4, 4), (5, 4, 3) and (3, 5, 4) in the RGB, YIQ and UVW systems. The color shifts in the RGB system were found to be smallest

and were in relative agreement with the error values of Table I. In the UVW and YIQ systems, relatively large amount of color shifts were observed (compared to the values obtained computationally [2]) and were found to be very sensitive to the quantization range of each coordinate. These large color shifts occurred because the actual ranges of the I, Q and U, W components of the image were much smaller than the corresponding ranges of these coordinates in their color solids, and also a significant number of quantization cells in the rectangular solid enclosing the actual color solid lie outside that color solid. The conclusion from this study is that for a fixed number of quantization cells, the cells should be packed into the volume occupied by the color solid. Such a quantization procedure, however, is difficult to implement except for the RGB coordinate system.

TABLE I. Quantization errors in the RGB system for different non-inferior bit assignments.

Bit Assignment	Red	Green	Blue	Cyan	Magenta	Yellow	White
4,4,4	15.4	7.0	20.6	4.5	8.2	5.0	3.4
4,5,3	21.0	10.3	1.5	5.1	3.6	7.8	3.5
5,3,4	22.8	6.7	2.9	5.0	13.8	0.8	6.0
5,4,3	23.2	9.3	13.9	3.9	7.2	7.0	4.5
3,5,4	12.9	11.5	11.1	8.5	0.3	7.7	5.7
3,4,5	12.3	3.3	16.9	8.3	0.2	6.8	5.7

References

1. A. K. Jain, "Color Distance and Geodesics in Color 3-Space," (to appear in the J. Opt. Soc. Am.).

2. A. K. Jain and W. K. Pratt, "Color Image Quantization," Presented at the National Telecommunications Conference, Houston, December, 1972.

3. W. K. Pratt, "Spatial Transform Coding of Color Images," IEEE Trans. on Comm. Tech., Vol. COM-19, No. 6, December, 1971, pp. 980-992.

3.2 Deltamodulation and Differential PCM Coding of Monochrome and Color Signals

A. Habibi

Deltamodulation and Differential PCM image coding systems offer substantial promise for real time television coding applications because of their relatively modest implementation requirements. The following sections present an introductory discussion of these systems and descriptions of their application to the coding of monochrome and color images.

DPCM Systems. A differential PCM image coder, as shown in figure 1, consists of a linear predictor that estimates the incoming pixel value, a subtractor that subtracts the estimated pixel value from its actual value to form a differential signal, and a quantizer that maps the differential signal into a set of discrete levels. An efficient predictor for video signals is one that exploits the correlations of the signal in all spatial directions. Experimental results have indicated that employing a third order predictor results in a substantial improvement over the first order predictor [1]. Further increases in the prediction order are usually not warranted.

Deltamodulation Systems. In a deltamodulation system each image sample is compared to its estimate and a positive or negative signal is produced depending on the comparative amplitude of the difference. A block diagram of a delta modulator is shown in figure 2. The output of the comparator is multiplied by a constant in the feedback loop and is used as an input signal to an integrator whose output is the estimate of the incoming signal. The size of the constant (step size) is an important parameter in the encoder. A large step size degrades the encoded signal when the signal is changing smoothly by causing a large granular noise; a small step size limits the ability of the encoder to build up to large signal change. This is avoided by introducing some memory in the quantizer that considers the polarity of the previous binary digits at the output of the quantizer and changes the step size accordingly. A small step size for a slowly varying signal and a large step size for a fast varying signal enables the adaptive deltamodulator to track the input signal more accurately. Abate [2] has developed a workable system that uses step sizes of 1, 2, and 4.

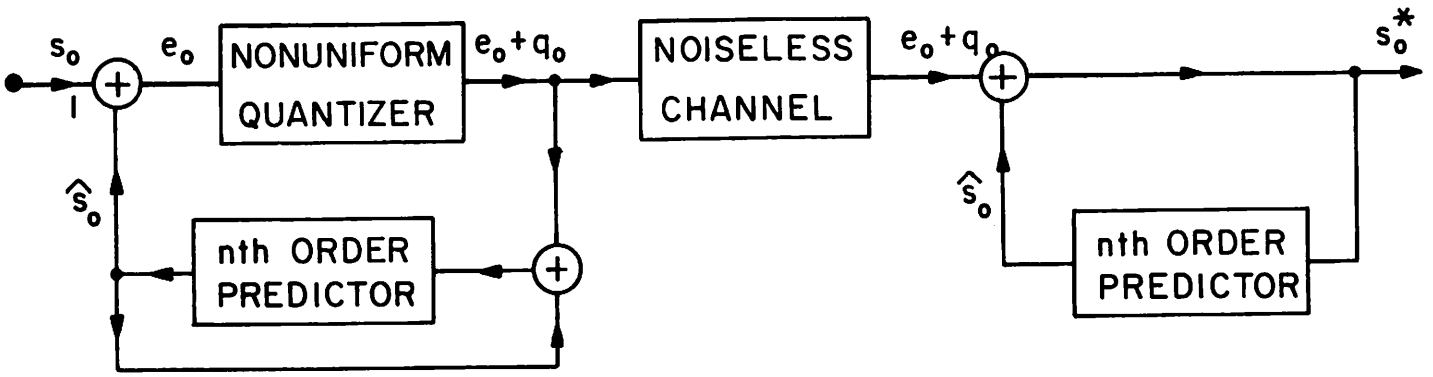


Figure 3.2-1 Block Diagram of an n^{th} Order DPCM Image Coder

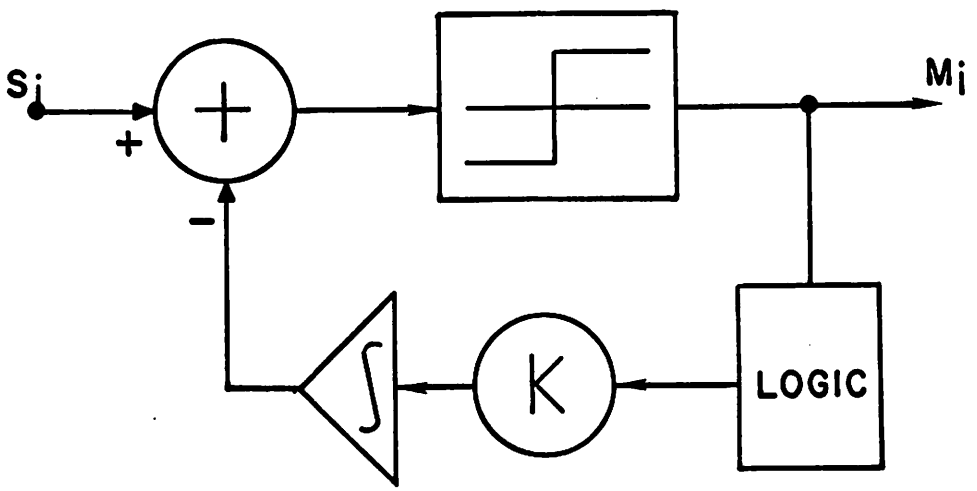


Figure 3.3-2 Delta Modulation Image Coder

Dual Mode Deltamodulation-DPCM Systems*. The combination of the deltamodulation and the DPCM coding techniques has been found to provide an improvement in information transfer rate for a given picture quality for monochrome images when compared to either system separately. Use of DPCM alone is inefficient because the codeword output symbols are not typically generated with equal probability. Use of deltamodulation alone is inefficient because of quickly changing amplitude. The dual mode coder combines the best of both modes (deltamodulation for slowly changing regions and DPCM for quickly changing regions) and potentially achieves a bit rate reduction while preserving picture quality. The simulation results confirm this hypothesis.

Figure 3 presents a pictorial view of the principal somponents of the dual mode coder system. The position of switches S1 and S2 determine the current coding mode. The switching is performed automatically by the system. This system is similar to one considered by Frei et al [3] with the difference that here a third order predictor is used for the DPCM system.

The quantizer for deltamodulation is uniform, with two levels. The deltamodulation reconstruction level Q determines the extent of mode switching in the system. The DPCM quantizer is nonuniform with 8 levels.

Prior to processing, the number of pixels per line is expanded by a factor of three. The new pixels are interpolated from the original data. The deltamodulation coder uses the expanded lines where the DPCM coder uses only the original data. During the deltamodulation mode, groups of three pixels are coded. The sequences of three bits produced by the quantizer are then compressed to one bit according to a majority decision. If all three bits are identical they are transmitted uncompressed; the coder also switches to DPCM. During the DPCM mode each of the original pixels is coded for transmission. When two successive transmissions are at the lowest quantization level, but alternate sign (idling condition), the coder switches back to the delta mode since such a condition indicates the resumption of a smooth picture region. One drawback of this dual mode coding is the generation of data at an uneven rate which requires buffering the data prior to its transmission.

* The results pertaining to the combination of DPCM and deltamodulation coder were obtained by Thomas Middleton a graduate student in Electrical Engineering Department of the University of Southern California.

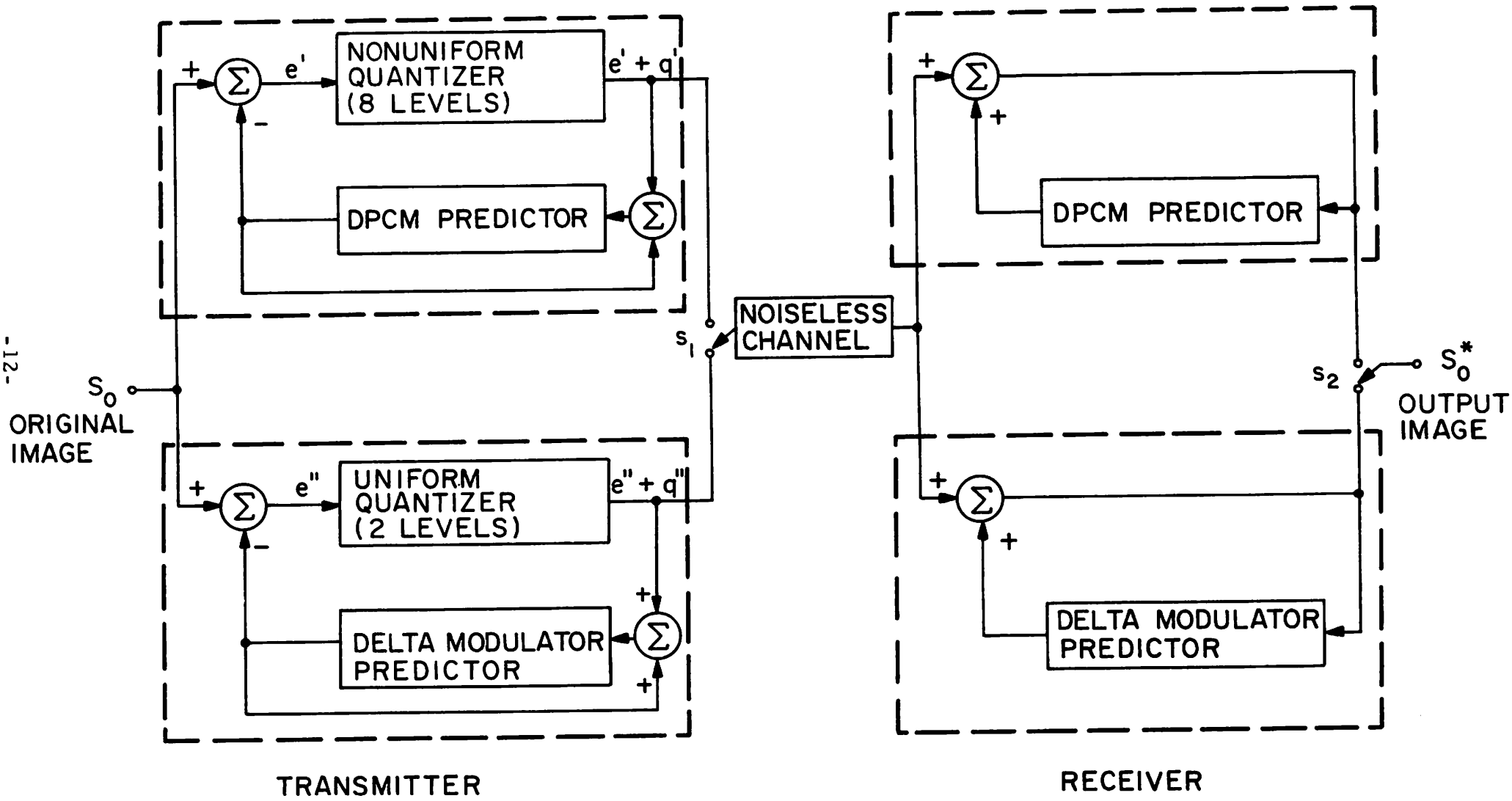


Figure 3.2-3 **Dual mode deltamodulation - DPCM image coder**

Photographs of the original and processed pictures appear in figures 4 and 5. Results are quite encouraging and support the results reported by Frei [3]. All of the simulation run parameters were trained on the tank picture. With the delta reconstruction threshold set at Q=4.0, the tank showed only slight degradation coded at 1.72 bits/pixel (Huffman DPCM Coding). Using these same parameters, the girl was coded in 1.15 bits/pixel, suffering slightly more degradation. With Q=5.0 and a bit rate of 1.61 (3 bit DPCM coding) the tank showed slightly higher contouring but still a good picture. At Q=10.0 and 1.16 bits/pixel, the contouring was more pronounced. Thus an acceptable picture can probably still be coded at about 1.5 bits/pixel, the rate achieved by Frei [3].

Based on these results the dual mode coder appears to be a workable scheme, particularly in a noiseless environment such as storing video data digitally. In a noisy transmission channel coder system it is not as attractive because redundancy must be added to keep performance at a higher level. Differential coding schemes are unstable in the presence of noise and do not recover until reinitialization such as at the beginning of a line.

Comparison of DPCM and Delt Modulation Systems for Color Images. Both the encoder and the decoder of a DPCM and the delta modulation systems were simulated on a digital computer to code the luminance and the chromaticity components of a color video signal. The luminance component of this video signal is coded by the DPCM and the delta modulation systems at various bit rates. The results are used to evaluate the performance of each coding system and compare their performances.

The mean square values of the coding errors (MSE) are plotted in figure 6. The performance of both the first and the third order DPCM systems improve at about the same rate by increasing the number of quantization levels. The first order DPCM system with a two-level quantizer is essentially the same as the simple delta modulator operating at one bit per picture element. The adaptive delta modulator, taking advantage of the feedback properties, outperforms both of these systems at coding rates of

(a) Original

(b) DPCM, 2.66 bits/pixel

(c) Dual

(d) Dual
Q = 5, 2.61 bits/pixel

(e) Dual
Q = 10, 1.16 bits/pixel

(f) Dual
Q = 4, 1.18 bits/pixel

Figure 3.3-4 Dual mode image coding using "TANK" picture

(a) Original

(b) Dual
Q = 4, 1.16 bits/pixel

(c) Dual
Q = 6, 1.07 bits/pixel

Figure 3.3-5 Simulation results using "GIRL" picture

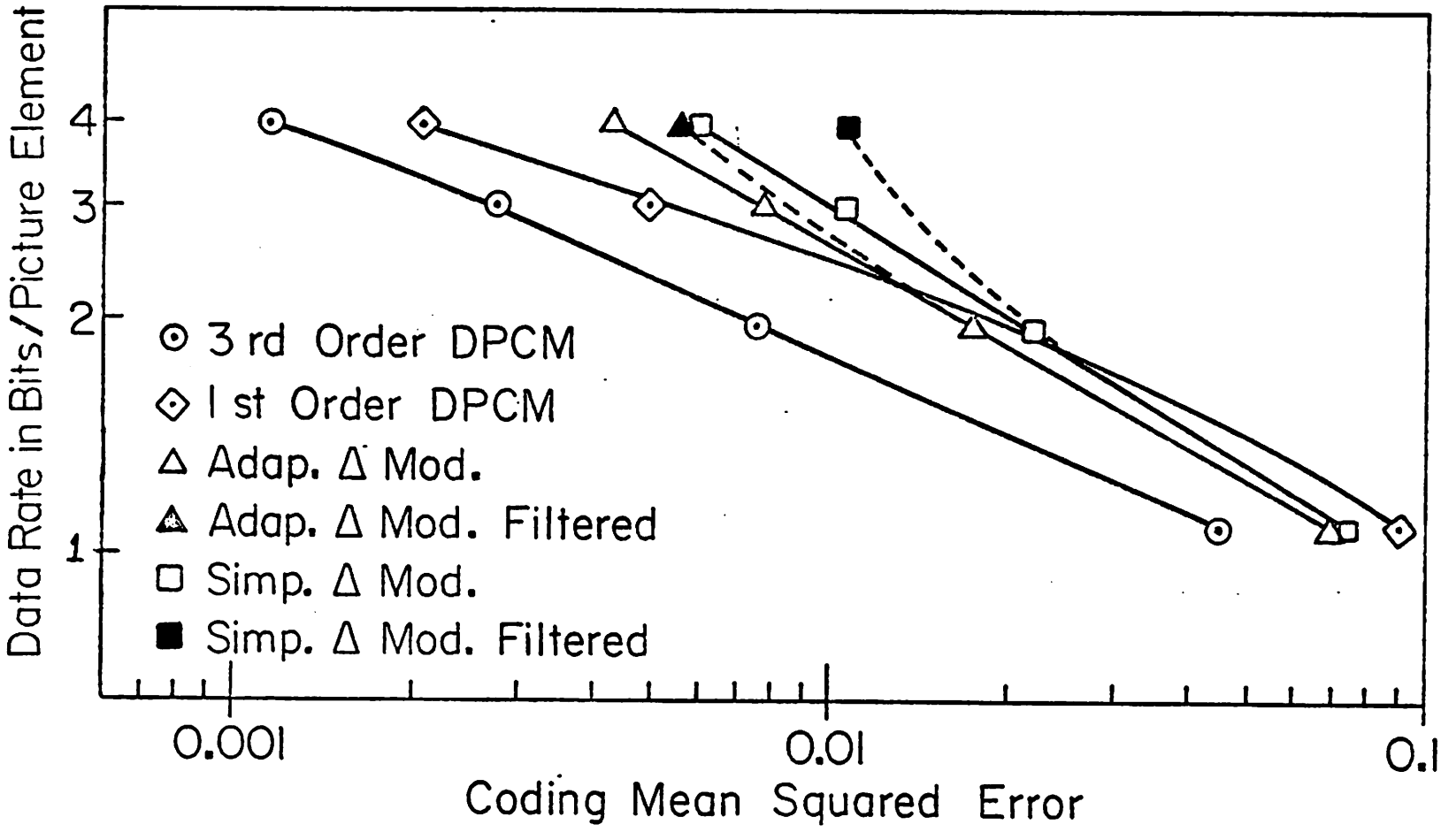


Figure 3.2-6 Coding Mean Squared Error at Various Data Rates for the Normalized Picture

one and two binary digits per picture element. However, the first order DPCM system achieves a better coding capability at higher bit rates. As shown in figure 6 the performance of the third order DPCM system is better than the performance of the delta modulators at all bit rates.

The results of the delta modulators at bit rates higher than one bit per picture element requires sampling the analog data at a higher sampling rate. The previously sampled data was interpolated to generate a larger number of samples. The linear interpolation used to double, triple, and quadruple the number of samples introduces higher frequency errors and other undesirable distortions. The higher frequency error is eliminated in practice by low pass filtering the coded signal. In calculating the coding error for the simulated systems the effect of the high frequency error was eliminated by two different techniques (see [4]). The solid lines on figure 6 corresponds to complete elimination of these errors, and the curves shown by dashed lines correspond to partial elimination of the high frequency errors.

The coded pictures corresponding to some of the points on figure 6 are shown on figure 7. The degradation in the picture coded by the third order DPCM system at a bit rate of 2 bits per picture element is not noticeable. Using the other schemes a rate of 3 bits per picture element is needed to obtain similar results.

The color video signal is very insensitive to degradations in the chromaticity components, thus these signals can be coded using only a fraction of binary digits per signal element. Figure 8 is a display of the I and Q signals coded by the adaptive delta modulator using 1/4 bit per pixel. To achieve this bit rate both the I and Q signals are subsampled to reduce the resolution in both spatial directions by a factor of two. This signal is then coded and the result is linearly interpolated to increase the number of picture elements to 256 by 256. Combining these signals with the coded luminance component results in coded video signal which does not show any noticeable degradation. The total bit rate using the luminance component coded by the third order DPCM system is 2.5 bits per picture element. The bit rate corresponding

(a) Original

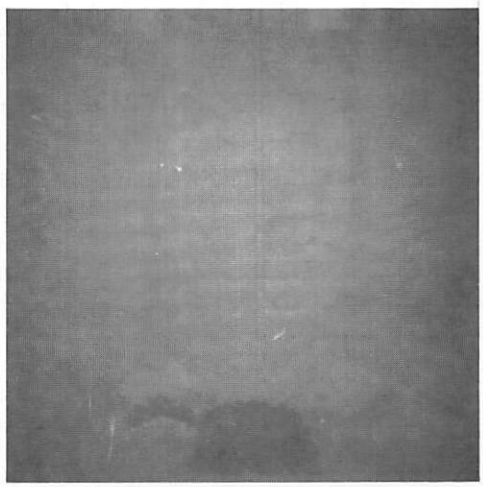
(b) 3rd Order DPCM, 2 bits/pixel

(c) Adaptive Delta, 3 bits/pixel

(d) Simple Delta, 3 bits/pixel

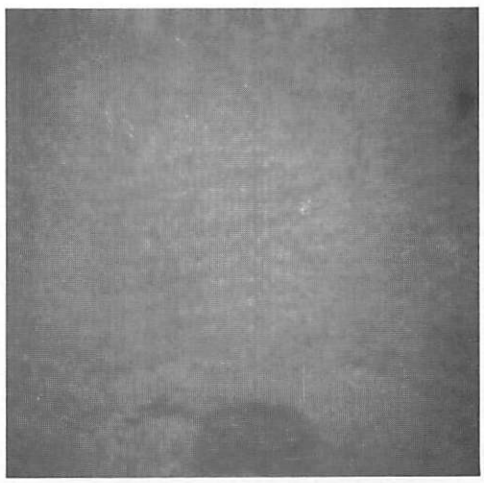
Figure 3.2-7 Lumination of a Color Signal Coded by DPCM, Adaptive Delta, and Simple Delta Modulators.

(a) Original I



(b) Original Q

(c) I, Adaptive Delta
0.25 bits/pixel



(d) Q, Adaptive Delta
0.25 bits/pixel

Figure 3.2-8. Original and Coded Chromaticity Components of the Color Video Signal

to the luminance signal coded by other schemes is 3.5 bits per picture element.

References

1. A. Habibi, "Comparison of nth-Order DPCM Encoder with Linear Transformations and Block Quantization Techniques", IEEE Trans. on Communication Technology, Vol. COM-19, No. 6, December, 1972, pp. 933-944.
2. J.E. Abate, "Linear and Adaptive Delta Modulation", Proc. of the IEEE, Vol. 55, No. 3, March, 1967, pp. 298-308.
3. A.H. Frei, H.R. Schindler, and P. Veitinger, "An Adaptive Dual-Mode Coder/Decoder for Television Signals", IEEE Trans. on Communication Technology, Vol. Com-19, No. 6, December, 1972, pp. 933-944.
4. A. Habibi, "Delta Modulation and DPCM Coding of Color Signals", Proc. of the 1972 International Telemetering Conference, October, 1972.

3.3 Slant Transform Image Coding of Monochrome Images

Wen-Hsiung Chen and William K. Pratt

In transform image coding using separable unitary transforms the two dimensional transform $F(u, v)$ of an image $f(x, y)$ may be written as

$$[F(u, v)] = [A][f(x, y)][A]^T$$

where $[A]$ is a matrix whose rows are basis functions of the transformation. An efficient transformation for image coding is one for which the image energy is redistributed to a relatively small number of transform coefficients $F(u, v)$. The remaining coefficients can then be quantized with a small number of quantization levels without significiant error. The energy compaction of a unitary transform will be "high" if the basis functions of the transform "resemble" typical lines of an image. In most natural images a large number of lines or line segments are of nearly constant brightness. Also many lines and line segments either increase or decrease in brightness over their length in a nearly linear fashion. Thus, one of the basis functions should be a constant and another a discrete sawtooth or slant waveform. With this thought in mind, a family of basis functions containing a flat and a sawtooth waveform member have been developed. The remaining basis functions have been chosen so that the sequency (number of zero cross-

ings) of each basis function is equal to its row number minus one. Also, by design, the resulting set of basis functions, called the Slant transform, possesses a fast computation algorithm. The algorithm is described in detail in ref. [1]. Figure 1 is a superimposed plot of the Slant and Hadamard transform basis functions for a vector length of 16.

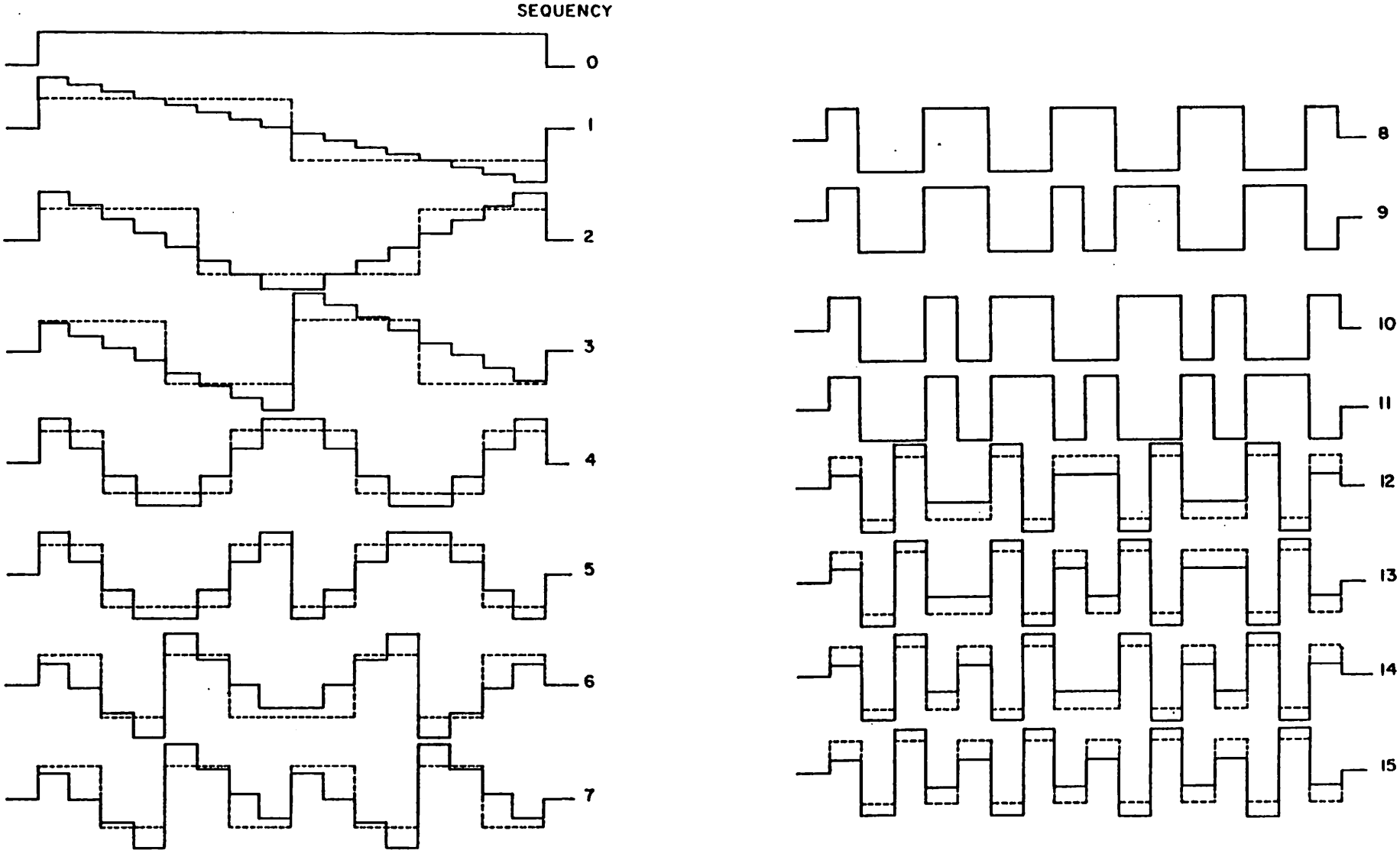
An intensive study has been performed on techniques for efficiently quantizing transform domain samples for the commonly employed image transforms. An optimum strategy from the standpoint of minimizing the mean square error between the original and the reconstructed images has been found. First, the number of bits assigned to each transform coefficient is made proportional to the logarithm of the variance of the coefficient. The quantizer levels are then selected according to the Max quantization rule. Figure 2 contains a plot of the mean square error obtained with several image transforms as a function of the image block size processed for image coding with an average of 1.5 bits/pixel. In this example, the image field was modelled as a first order Markov process with row and column correlation factors of 0.95. From this figure it is seen that the Slant transform is much superior to the Hadamard transform and closely approaches the optimal performance of the Karhunen-Loeve transform. Also, the curve indicates that the mean square error does not decrease much after an image block size of 16 by 16 is reached.

Figure 3 contains original eight bit, 256 by 256 pixel images and their reconstructions coded with an average of 1.5 bits/pixel using the quantization strategy described above. Subjectively, the reconstructions appear to be of good quality with little visible degradation.

Further research on slant transform image coding is underway. The effects of channel errors are being studied, and the Slant transform is also being applied to color images.

References

1. W.K. Pratt, L.R. Welch, and W. Chen, "Slant Transforms for Image Coding", Proceedings Symposium on Applications of Walsh Functions, March, 1972.



solid = Slant transform

dotted = Hadamard transform

Figure 3.3-1 Comparison of Hadamard and Slant Basis Vector of Length 16

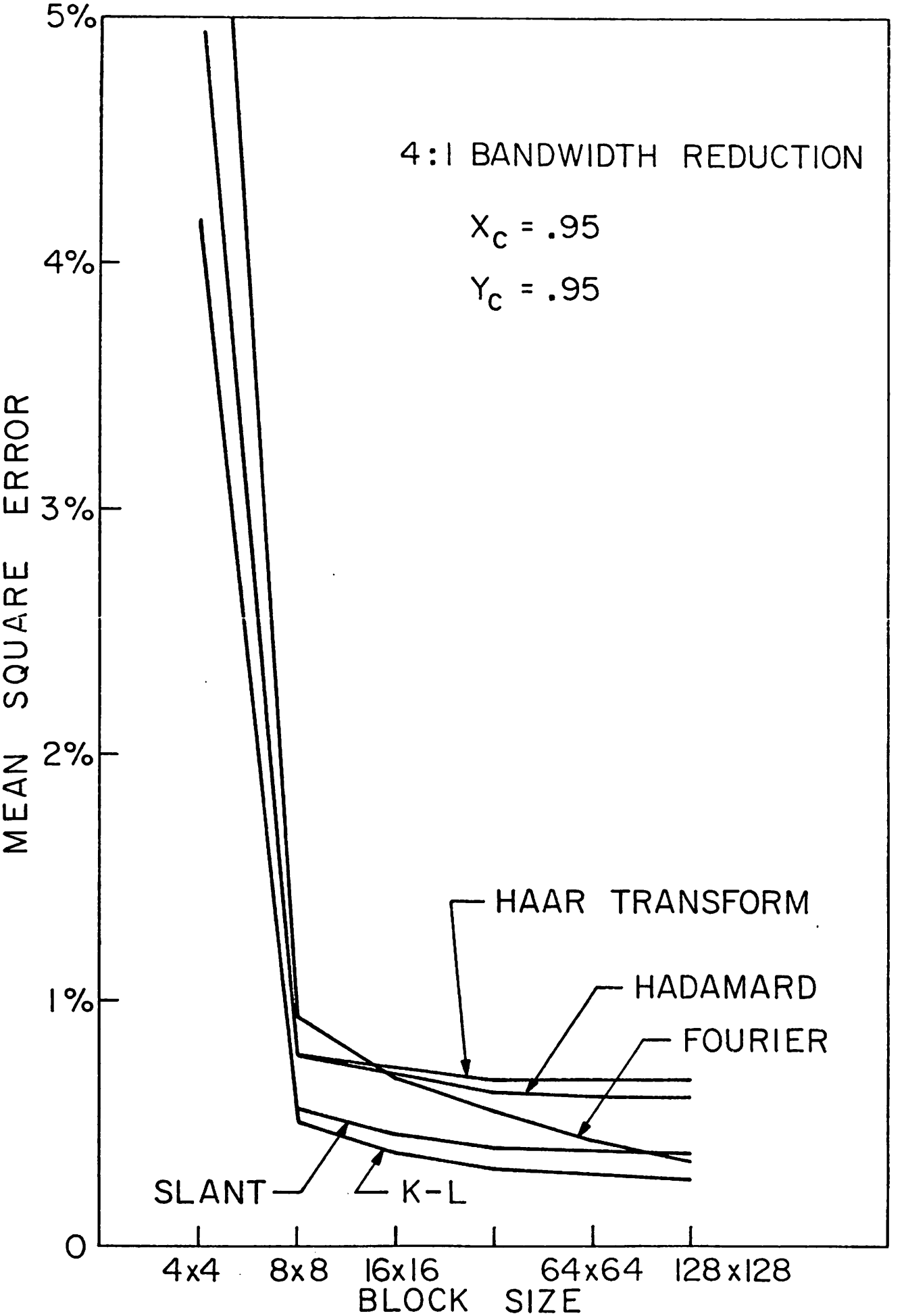


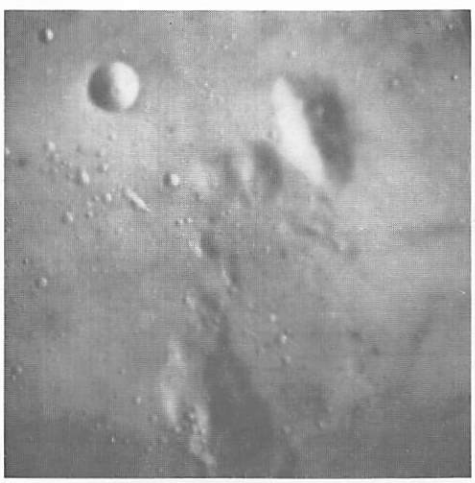
Figure 3.3-2 Mean Square Error Performance of Transform Image Coders

(a) Original

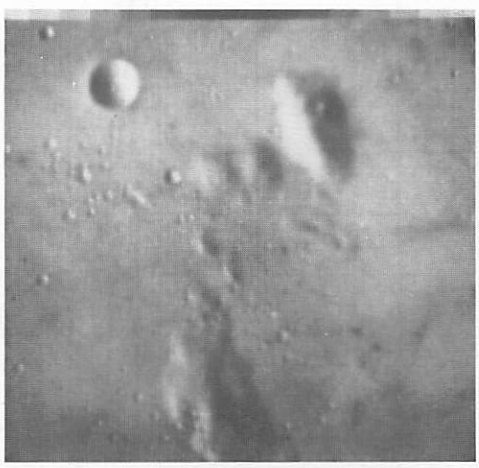
(b) Zonal Coding
1.5 bits/pixel

(c) Original

(d) Zonal Coding
1.5 bits/pixel



(e) Original



(f) Zonal Coding
1.5 bits/pixel

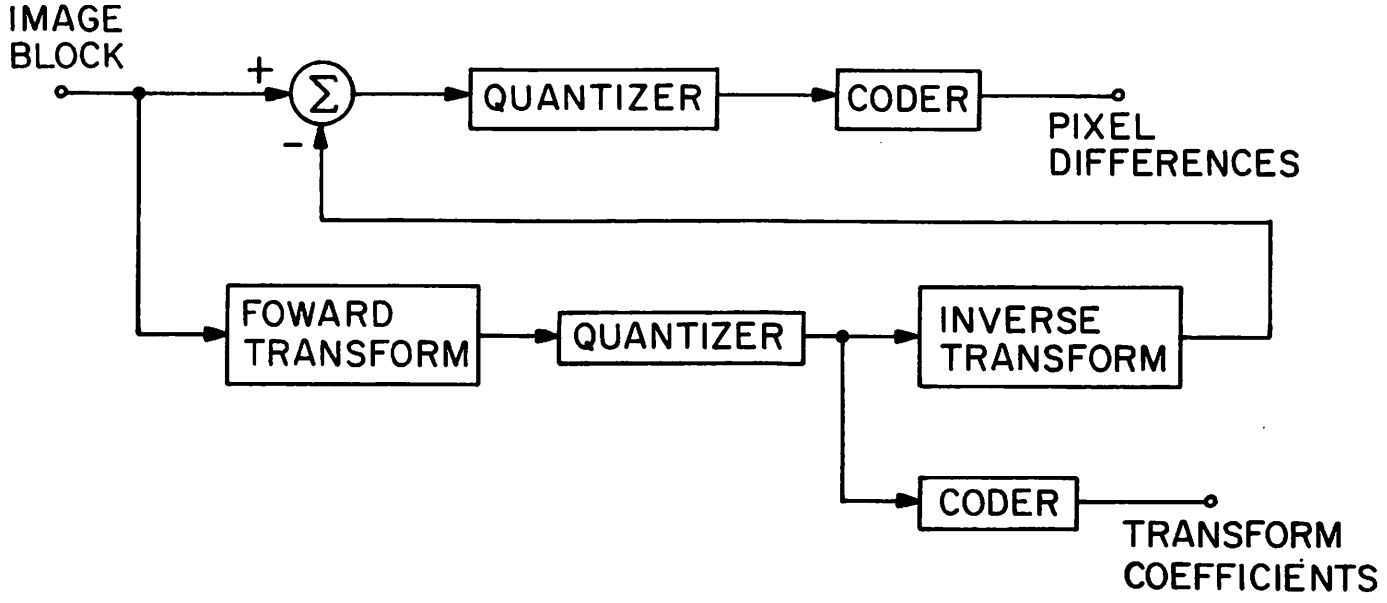
Figure 3.3-3. Slant Transform Image Coding Examples.

3.4 Transform Differential Coding of Monochrome Images

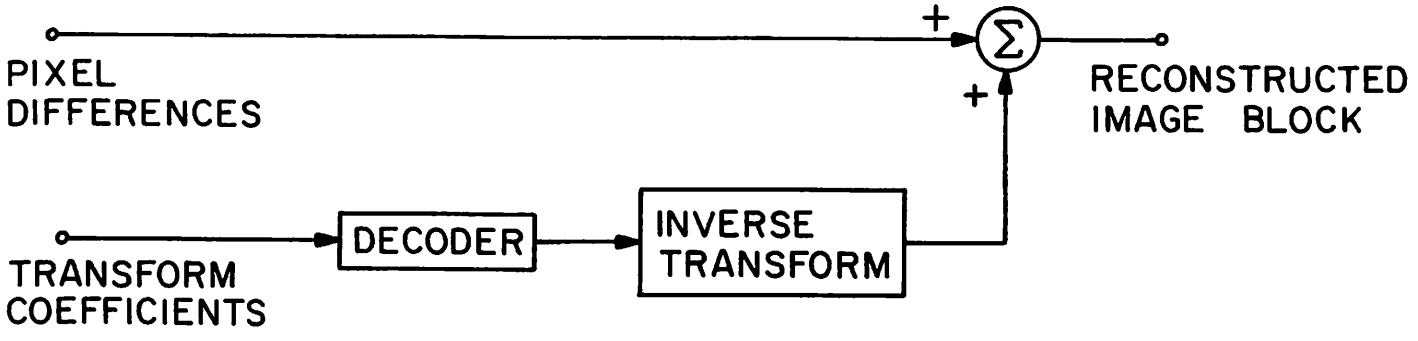
Faramarz Davarian and William K. Pratt

In conventional transform image coding a surprisingly recognizeable monochrome image can be reconstructed using about 5-10% of the transform coefficients with the largest magnitude. The reconstructed image generally possesses the correct overall grey scale shading, but lacks high frequency detail. One could consider this reconstructed image to be a first order estimate of the original image in the sense that the feedback signal in a differential PCM image coder is an estimate of the scanned pixel values. This idea has led to the development of a coding technique called transform differential image coding as shown in Figure 1. In this system an image block undergoes a two dimensional unitary transform. The transform coefficients are grossly quantized, coded, and transmitted. The quantized transform coefficients are then subjected to an inverse two dimensional transformation. The resultant image estimate is then subtracted pixel by pixel from the original image. The pixel difference signals are quantized, coded, and transmitted. At the receiver the pixel difference signals are added to the image estimate to obtain a reconstruction of the image block. A bandwidth reduction can be obtained with the transform differential image coding by two coding techniques. One of the techniques is information preserving and yields a moderate compression ratio, and the other technique permits a small reconstruction error in order to achieve a greater compression ratio.

Information Preserving Coder. In the information preserving coder version of the transform differential coder, the pixel difference signals are quantized to the same grey scale resolution as the pixels of the original image and then coded using a variable length Huffman coder. Typically, an average of about 2.0 to 2.5 bits/pixel is allotted to the pixel difference signals. It should be noted that the quantization process on the transform coefficients does not create any reconstruction error since the pixel difference signals are perfectly coded, but the transform coefficient quantization process does affect the coding performance of the overall system. Figure 2



(a) Coding system



(b) Decoding system

Figure 3.4-1. **Transform differential image coding system**

illustrates entropy of the pixel difference signal as a function of the number of transform coefficients transmitted for a typical picture using a 16 by 16 pixel block slant transform. Figure 2 also contains plots of the number of bits allotted to the transform coefficients and the total number of code bits required for the image using an optimum coder for the pixel differences. The maximum pixel length has been limited to a difference of five quantization levels in this example. If a pixel difference greater than five occurs the actual pixel difference is coded. The simulations indicate that for an optimized system, an image can be coded with an average of about 3.5 to 4.0 bits/pixel. This result is somewhat disappointing since a coding of about 4.0 bits/pixel can be achieved simply by coding the pixel differences along an image line.

Non-Information Preserving Coder. In the non-information preserving coder, the pixel difference signal is quantized using a nonlinear quantizer with a relatively small number of levels. The quantizer design follows the same general rules as for a differential PCM image coder. In an initial set of simulation experiments shown in Figure 3, the slant transform coefficients were coded with an average of 2.0 bits/pixel and the pixel difference signal was quantized to only three levels and coded with 1.0 bits/pixel. The resulting reconstruction of 3.0 bits/pixel is quite good. It appears possible to approach a coding of 1.5 bits/pixel with this technique using an adaptive quantizer. Further investigations along this line are underway.

3.5 Quantization Error and Entropy of a Single Random Variate

A. Habibi

Most practical coding systems that are used to convert analog signals to binary digits are composed of three serial processors. The first attempts to process the input data to form a set of uncorrelated samples with a continuous amplitude; the second, a memoryless quantizer, maps the continuous-valued signal to a set of a given number of discrete levels; and finally the third processor converts this data to a set of binary

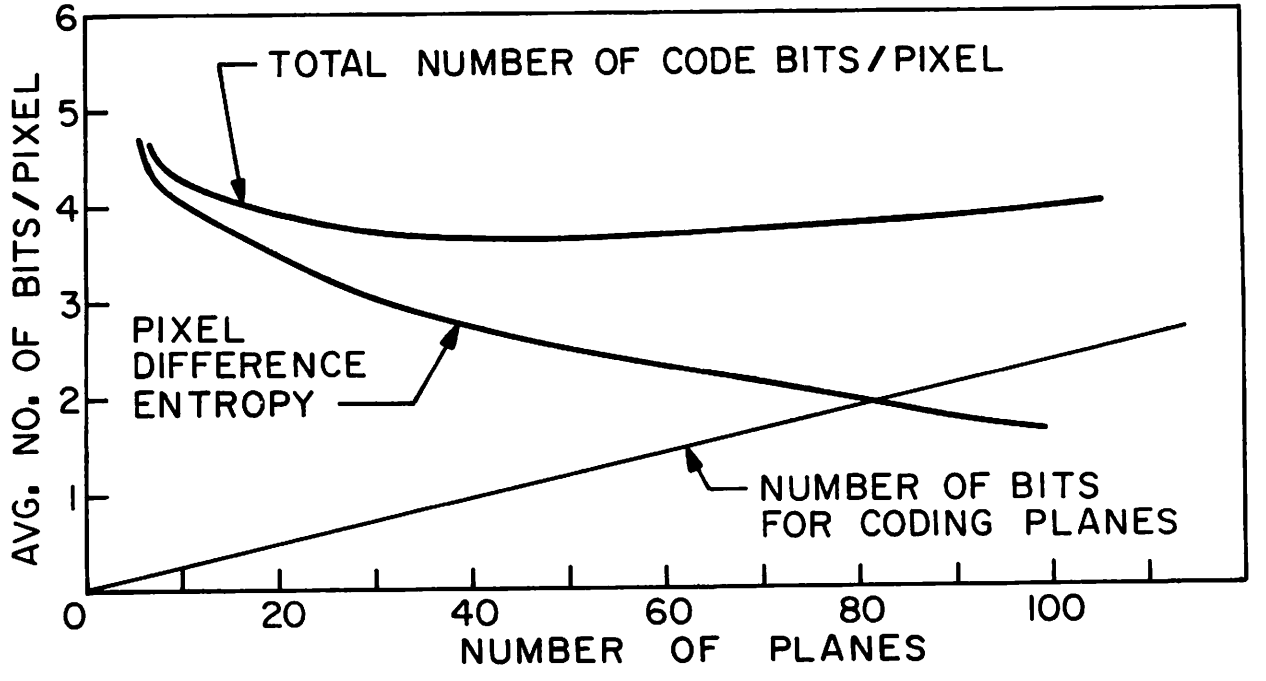
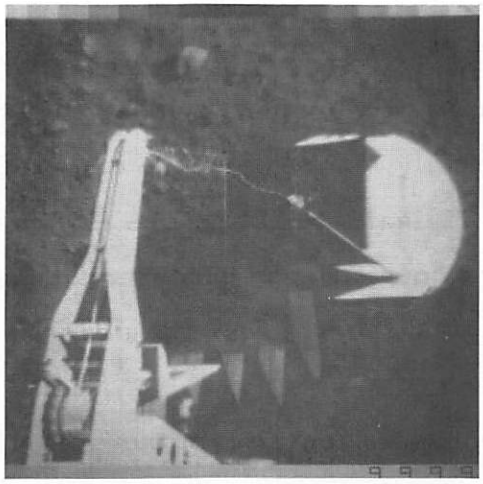
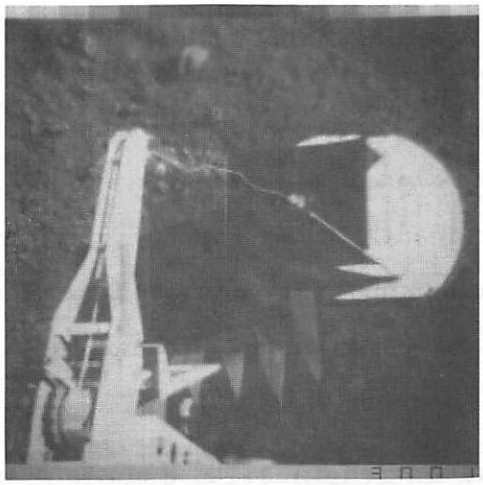


Figure 3.4-2. Coding performance example for differential transform image coder



(a) Original



(b) Transform differential coded
3.0 bits/pixel

Figure 3.4-3. Transform differential image coding examples

digits by a fixed-length or a variable length coder. The use of the memoryless quantizer is justified if the signal at the input of the quantizer is uncorrelated. Examples of decorrelating processors are the sampler and linear predictor in differential PCM and the Karhunen-Loeve transformation in the transform coding schemes [1,2]. The variable coding scheme in the output of the quantizer is employed to generate binary data at an average rate equal to the entropy of the discrete data. This rate is, in general, lower than the rate of data generated by a fixed-length coder which is equal to the base-two logarithm of the number of quantization levels. Naturally, increasing the data rate will reduce the degradation between the original signal and one that is reconstructed from the binary digits by inverse operations. This relationship is bounded by Shannons rate distortion function which is the absolute bound for any coding system [3,4].

In this report attention is focused on the operation of the memoryless quantizer. The criteria of optimality are both the mean square error and the entropy of the discrete signal. Thus the system which performs closest to the rate distortion function is the ideal system. Using this criterion the performance of some existing quantizers in coding uncorrelated signals of various distributions is studied. Of specific interest to image coding are the Max quantizer, the instaneous companding, quantizer and the uniform quantizer for input signals possessing Gaussian or double-sided exponential probability density functions. These probability densities are of interest because of their occurrance in DPCM and transform coding schemes for pictorial data.

Review of Quantization Schemes. A quantizer maps a sample x, a continuous variable with a continuous probability density $p_x(x)$ into one of a finite set $\{y_1, \ldots, y_N\}$ of rational numbers. This mapping follows the procedure:

$$\text{if } x_j < x \leq x_{j+1} \text{ then } x \text{ is mapped to } y_j (j=1, \ldots, N)$$

The $x_j (j=2, \ldots, N)$ are the transition levels, with x_1 and x_{N+1} the greatest lower bound and least upper bound, respectively, to the input signal, and

the $y_j (j=1, 2, \ldots, N)$ are the quantization values. Given the total number N, of discrete levels allowed, a quantizer is specified by the sets of transition levels x_j and the reconstruction values y_j.

In his search for a quantizer with a minimum distortion, Max [5] minimized the expected value of an arbitrary function of error between the input and the reconstructed signals, and obtained a set of sufficient conditions. For a square function of error the sufficient conditions are

$$y_j = 2x_j - y_{j-1} \qquad j=2, \ldots, N \qquad (1)$$

and

$$\int_{x_j}^{x_{j+1}} (x-y_j)\, p(x)dx=0 \qquad j=1, \ldots, N \qquad (2)$$

Note that this technique minimizes the mean square error without considering the entropy of the quantized numbers. Equations (1) and (2) can only be solved by iterated techniques. Various schemes of approximating the above quantizer have been suggested that use easier techniques of computing the transition and the reconstruction values [6-8]

In a uniform quantizer, as shown in figure 1, both the transition and the reconstruction levels are spaced at equal intervals, q. For a given number of quantization levels N, the maximum quantization levels is $X=(N-1)\, q/2$. The parameter q can be chosen to minimize the mean square error [5], or to minimize the entropy of the quantized signal or a combination of the two in order to code at a rate closest to the Shannon rate distortion function [4]. The Instantaneous companding quantizer is a third type of quantizer that, in a sense, combines attractive features of the other two types. This system uses a uniform quantizer that is pr - ceeded by companding of the input signal. The type of compression depends on the probability density function of the signal. The quantizer is followed by an expandor that compensates for the effect of the compression by expanding the quantized signal. The system is shown in figure 2 where f(.) and g(.) are the pre-and the post-quantization mappings. Algazi [6]

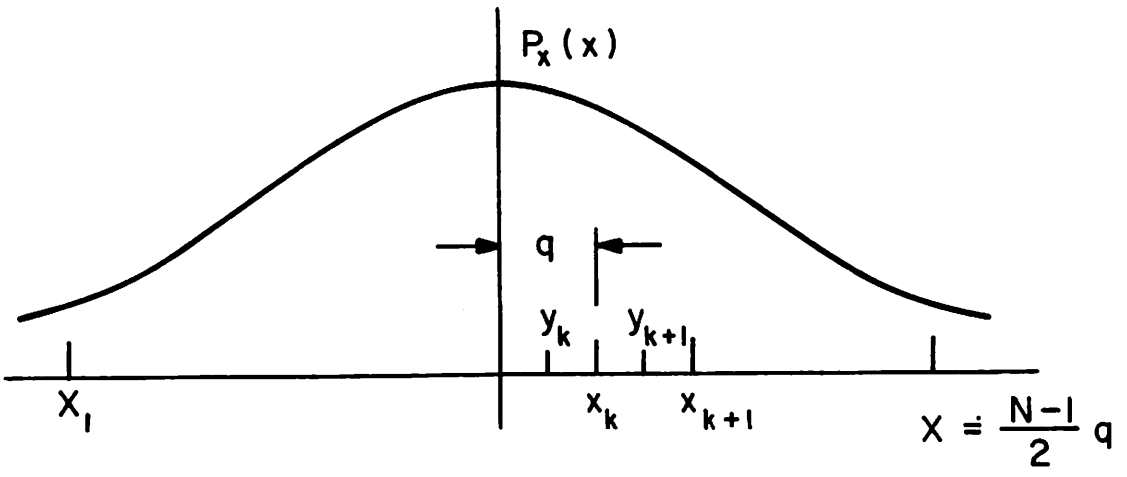


Figure 3.5-1. **Uniform quantizer**

Figure 3.5-2. **Block diagram of the instantaneous companding quantizer**

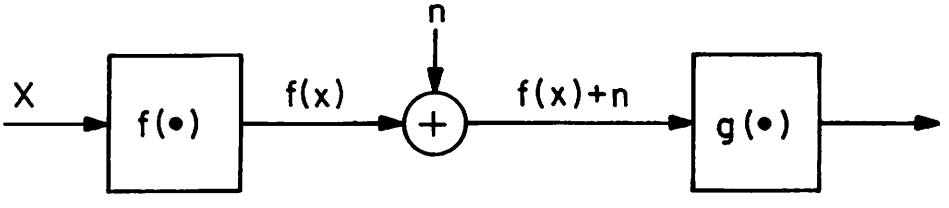


Figure 3.5-3. **The analog model of the instantaneous companding quantizer**

modeled this quantizer by the linear system of figure 3. Under the assumptions that the quantization noise is small, additive and is uncorrelated with the signal, Algazzi showed that the system will minimize the mean square error if mapping f(.) is given by

$$f(x) = K_1 \int [P_x(x)]^{\frac{1}{3}} dx + K_2 \tag{3}$$

and

$$g(.) = f^{-1}(.) \tag{4}$$

The same results have been obtained by Panter and Dite using a different approach [9, 10]. Mappings f(.) and g(.) are obtained from (3) and (4) for a given probability density $p_x(x)$. Then the uniform quantization parameter (for a given number of quantization levels) is chosen to minimize the mean square error or to operate at the closest point to the rate distortion function.

Comparison of Quantization Schemes. The Max, uniform, and the instaneous companding quantizers discussed in the previous sections have been designed for two-sided exponential and Gaussian probability density functions. These probability density functions have been chosen because of their importance in image coding. Equations (1) and (2) were then solved iteratively for the two-sided exponential and Gaussian probability density functions to obtain the quantization error and the entropy of the quantized signals for various number of quantization levels. The number of binary digits required to achieve this coding error is the base-two logarithm of the number of quantization levels for the system that uses fixed-length code words. When variable length coding is employed the number of binary digits is equal to the entropy of the quantized signal. Figures 4 and 5 show the bit rate versus the mean square error for the exponential and Gaussian probability densities for fixed-length codes while figures 6 and 7 show a similar relationship for the variable-length coder. In figures 4 and 5 the parameter q of the uniform quantizer is chosen to minimize the mean square error. The parameter q varies over a range of values for different numbers of quantization levels in

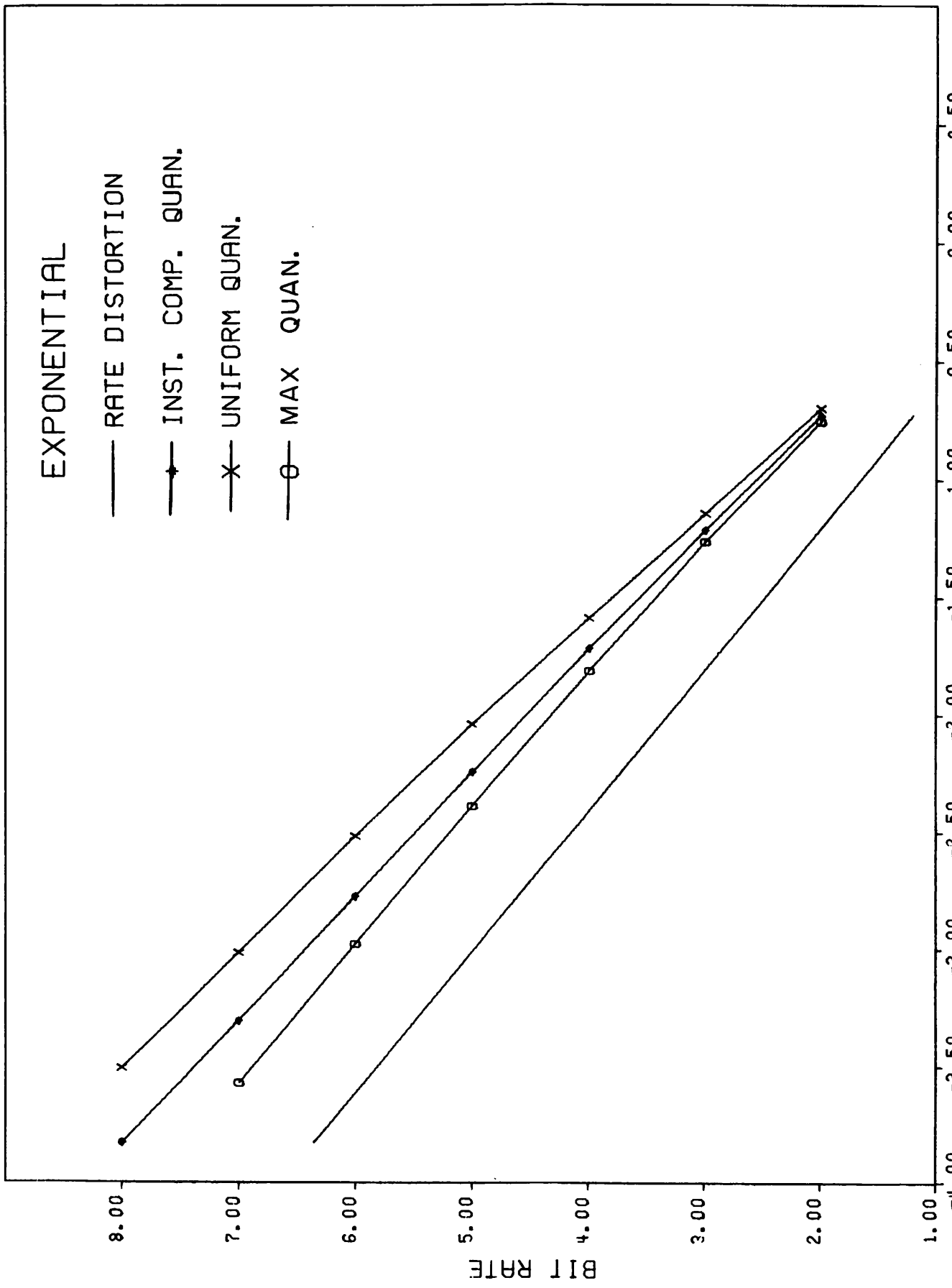


Figure 3.5-4. Bit rate versus the mean square error for fixed-length code words-exponential density

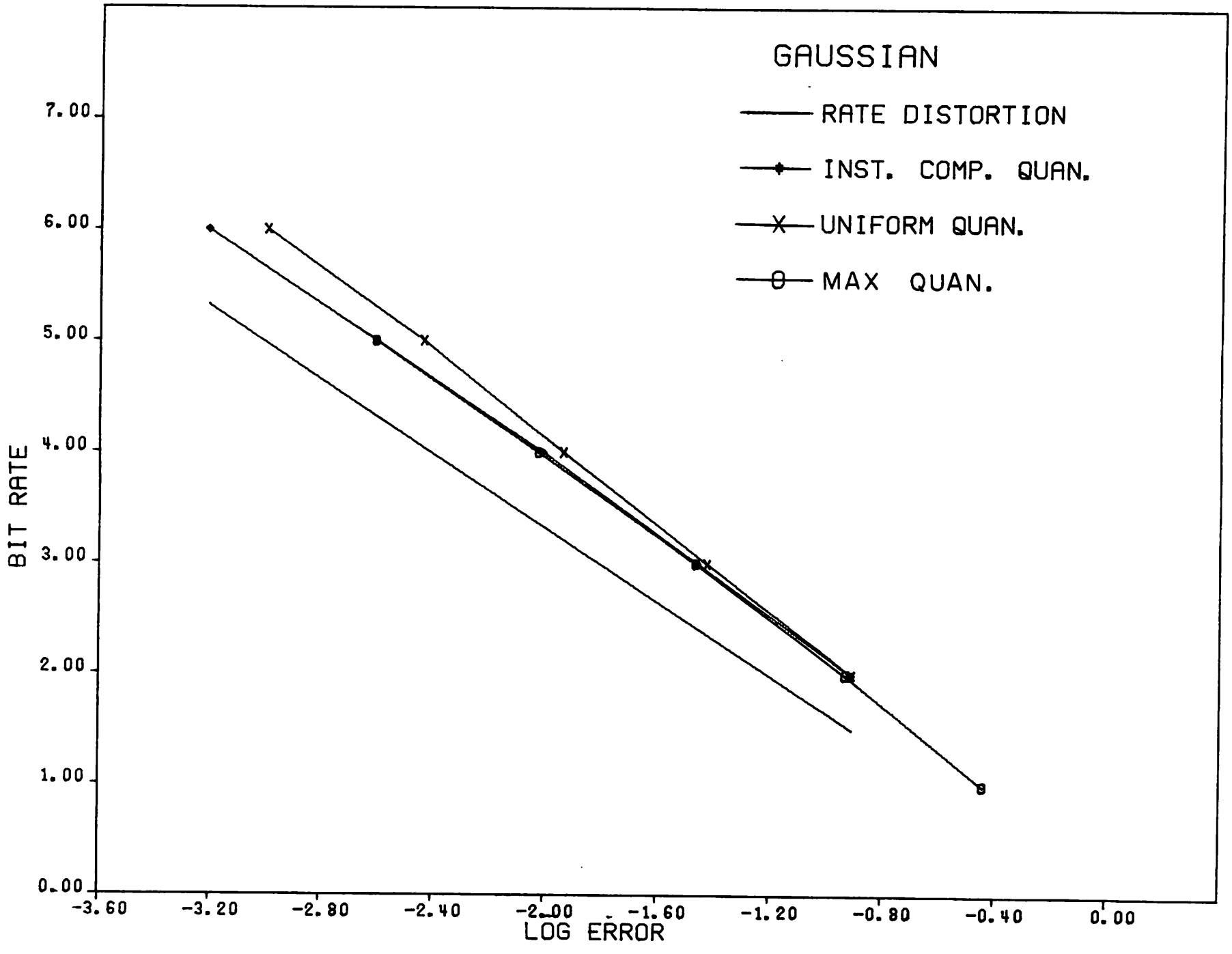


Figure 3.5-5. Bit rate versus the mean square error for fixed-length code words-Gaussian density

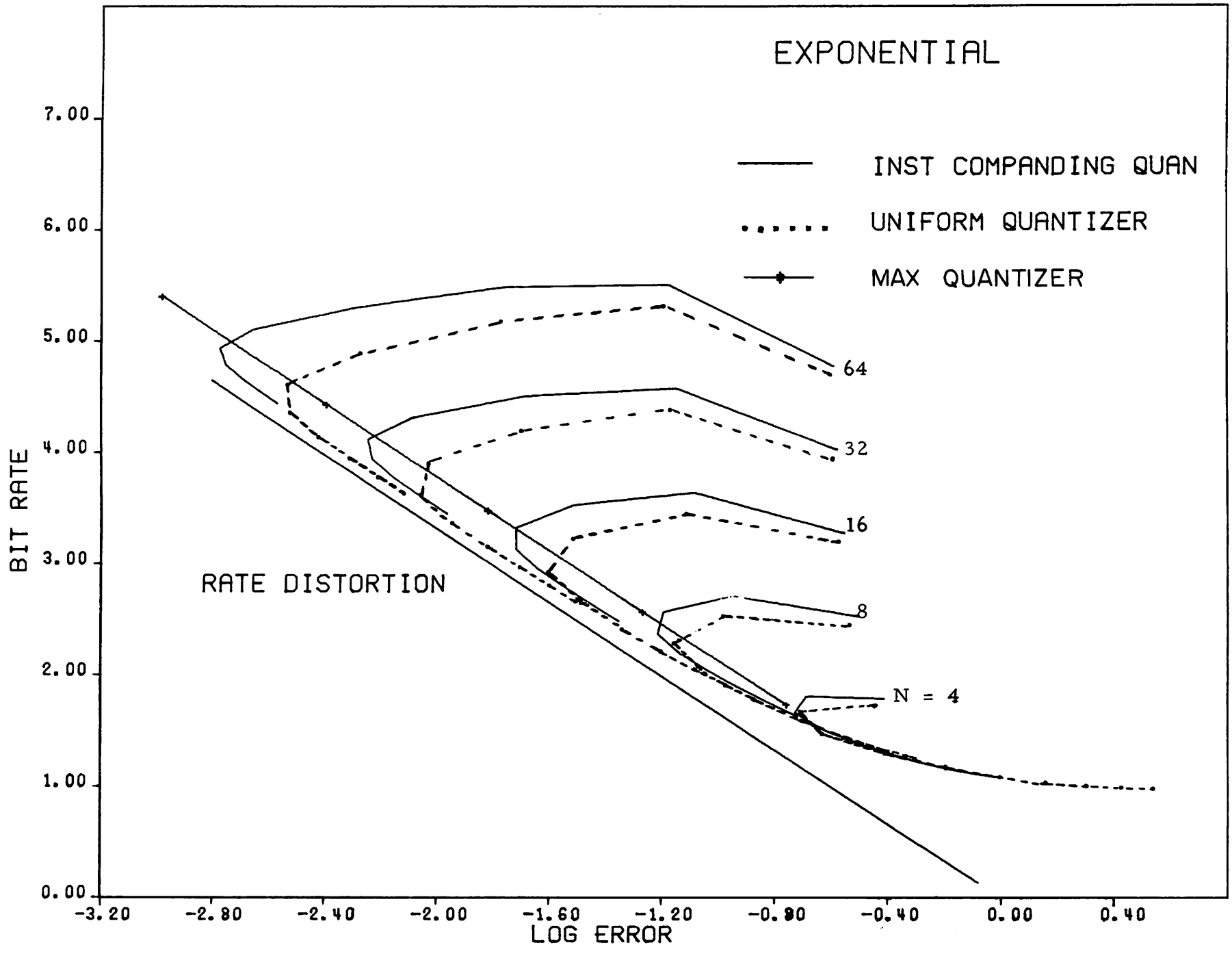


Figure 3.5-6. Bit rate versus the mean square error for variable-length coder-exponential density

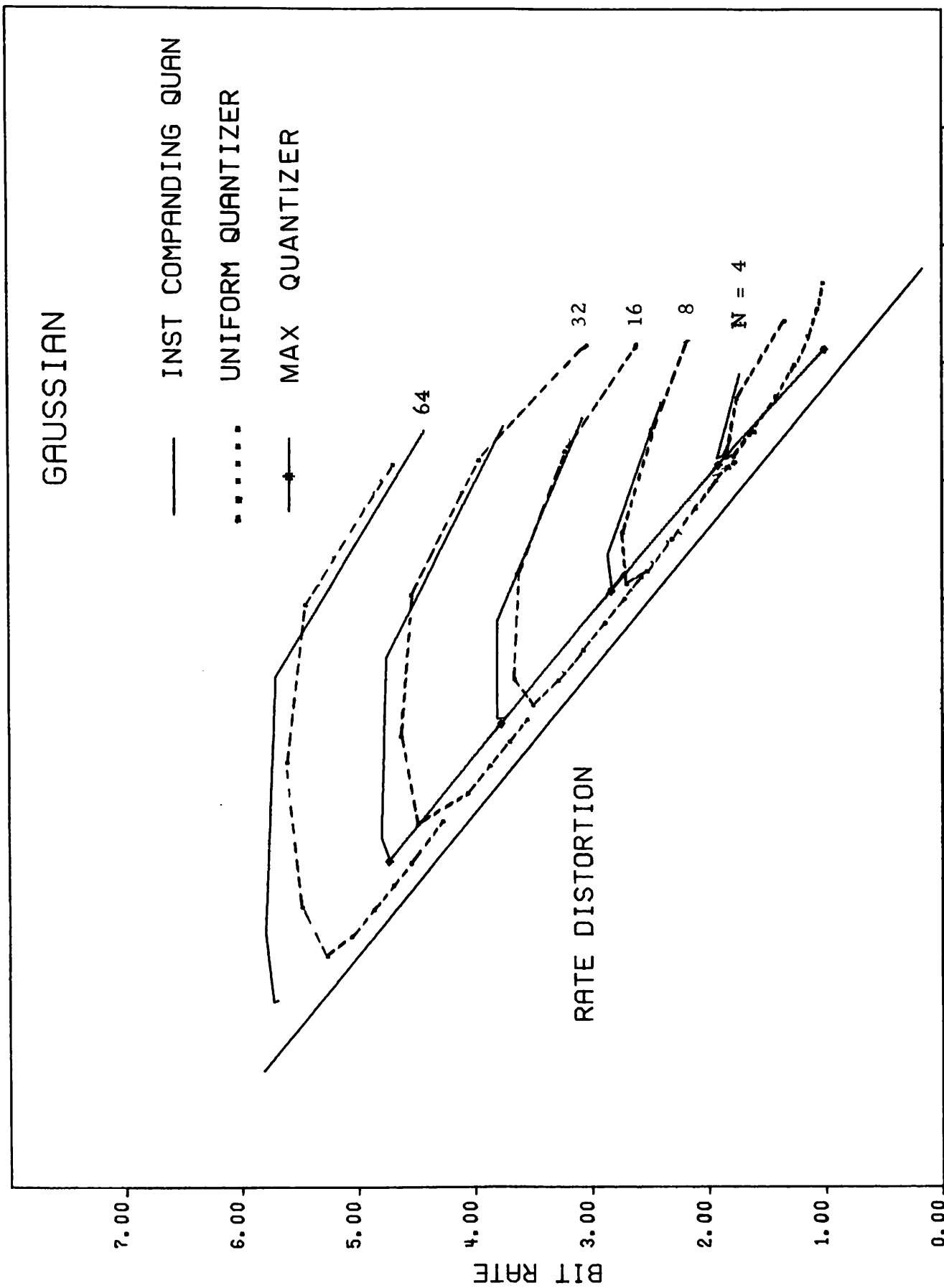


Figure 3.5-7. Bit rate versus the mean square error for variable-length coder-Gaussian density

figures 6 and 7. For the Gaussian probability density function the tangent to these curves is within 1/4 of a binary digit from the rate distortion function, and for the two-sided probability density it is within 1/8 of a binary digit. Figures 4 and 5 indicate that for a fixed-length coder the efficiency of the instantaneous companding quantizer is worse than the Max quantizer, but it is better than the uniform quantizer. Figures 6 and 7 show that the instaneous companding quantizer does not perform as well as the uniform quantizer using a variable coding scheme, but it outperforms the Max quantizer.

From the results above it is concluded that: (1) If the quantizer is followed by a variable length coding scheme (such as the Huffman Coder) then the best results will be obtained by using a uniform quantizer. The performance is not very sensitive to the least upper bound of the transition levels X. Good coding results are obtained with a value of X about three times of the standard deviation of the signal. The performance of the instantaneous companding quantizer is close to the uniform quantizer, but it remains inferior to it with the Max quantizer doing worse than both schemes. (2) Using fixed-length code words the performance of the Max quantizer is superior closely followed by the instantaneous companding and the uniform quantizers.

References

1. A. Habibi, "Comparison of nth-Order DPCM Encoder with Linear Transformations and Block Quantization Techniques," IEEE Trans. on Comm. Tech., Vol. COM-19, No. 6, December, 1971, pp. 948-956.
2. A. Habibi and P.A. Wintz, "Image Coding by Linear Transformations and Block Quantization", IEEE Trans. on Comm. Tech., Vo. COM-19, No. 1, February, 1971, pp. 50-63.
3. C.E. Shannon, "A Mathematical Theory of Communication", Bell Systems Tech. Journal, Vol. 17, October, 1948, pp. 623-656.
4. T.J. Goblick, Jr. and J.L. Holsinger, "Analog Source Digitization: A Comparison of Theory and Practice", IEEE Trans. Information Theory, Vol. IT-13, April, 1967, pp. 323-326.
5. J. Max, "Quantizing for Minimum Distortion", IRE Trans. Information Theory, Vol. IT-6, March, 1960, pp. 7-12.

6. V.R. Algazzi, "Useful Approximations to Optimum Quantization", IEEE Trans. Comm. Tech., Vol. COM-14, June, 1966, pp. 297-301.

7. R.C. Wood, "On Optimum Quantization", IEEE Trans. Information Theory, Vol. IT-15, March, 1969, pp. 248-252.

8. L.I. Bluestein, "Asymptotically Optimum Quantizers and Optimum Analog to Digital Converters," IEEE Trans. Information Theory, Vol. IT-10, July, 1964, pp. 242-246.

9. P.F. Panter and W. Dite, "Quantization Distortion in Pulse Code Modulation with Nonuniform Spaceing of Levels", Proc. IRE, Vol. 39, January, 1951, pp. 44-48.

10. B. Smith, "Instantaneous Companding of Quantized Signals", Bell Sys. Tech. J., Vol. 36, May, 1957, pp. 653-709.

3.6 Universal Coding

Lee D. Davisson

Suppose one is given a time and state discrete source (e.g., a sampled and quantized image) to block encode without error to minimize the average code word length, the code depending on some unknown parameters of the source message probabilities. The unknown parameter could be something as simple as, say, the probability of a one in a binary independent letter source or the transition probability matrix for a Markov model of an image, or as general as the complete set of message probabilities for the source. Such sources are called composite.

In some cases, although not usually, the unknown parameter may itself have a known prior distribution so that all ensemble message probabilities are known. There is no reason to suppose, however, that coding for the whole ensemble will be efficient for the actual parameter in effect. One of the important results of this study is to establish that coding for the whole ensemble does, under certain circumstances have a very important meaning in an asymptotic sense. For example, consider a composite binary source whereby the probability, θ, of a "one" is chosen by "nature" randomly according to the uniform probability law on $[0, 1]$ and then fixed for all time. The user is not told the θ, however. Subsequently, the source produces independent letter outputs with the chosen, but unknown, probability. The probability, given θ, that any message block of length N

contains n ones in any given pattern is then

$$\theta^n(1-\theta)^{N-n} \tag{1}$$

If θ were known, a per letter average code word length of

$$H(\theta) = -\theta \log_2 \theta - (1-\theta)\log_2(1-\theta) \tag{2}$$

bits would be required to block encode the source. Although θ is unknown, the composite source output probabilities are known. In fact, the probability of any given message containing ones in N outputs in simply

$$\int_0^1 \theta^n (1-\theta)^{N-n}\, d\theta = \frac{1}{N+1}\binom{N}{n}^{-1} \tag{3}$$

(Note that the composite source does not produce independent letters. This is due to the fact that past observations provide information about the value of θ in effect.)

There is no reason to suppose that the use of the probabilities of (3) to design a variable length code will do well for any θ that comes along in the sense of approaching (2) by using large blocks. One of the surprising results of this study is that the probabilities of (3) provide codes as good asymptotically as the probabilities of (1) with θ known exactly.

Consider the following encoding procedure. It is seen from (1) that all sequences with the same number of ones are equally probable. Send the message in two parts. Part one consists of a fixed length code for the number of ones using at most $\log(N+1)+1$ bits. Part two is a code indicating the location of the ones within the block, one of $\binom{N}{n}$ possibilities. At most the $\log\binom{N}{n}+1$ bits, a variable length depending on the number of ones, is required for this information. The per letter coding rate is then

$$\leq \frac{1}{N}\left[\log(N+1)+2+\log\binom{N}{n}\right] \tag{4}$$

The first part goes to zero as $N \to \infty$. The expected value of the second part must converge to the per letter entropy of equation (2) because the coding rate cannot be below the source entropy, and, at the same time, the second part of the code is optimal, all messages conditioned on the first part being equally likely, independent of θ. This will be demonstrated quan-

titatively subsequently. Note that a comparison of (4) and (3) shows that coding with respect to the whole ensemble with θ uniform on $[0, 1]$ is asymptotically optimum for this example.

Formulation of the Universal Coding Problem. Several measures of optimal universal coding effectiveness can be proposed in terms of a source block output random variable, $\underline{X}_N$ and an unknown parameter $\underline{\theta}$. As they all involve differences between the average code word length, denoted $\ell(\underline{X}_N \mid \underline{\theta})$ and the entropy per block symbol, $H(\underline{X}_N \mid \theta)$, the measure is a form of redundancy. In each case the infimum is taken over the class C_N of uniquely decipherable codes on block N. They are listed in order of ascending redundancy:

Definition 1. Weighted Redundancy

$$\mathcal{R}_w = \lim_{N \to \infty} \inf_{C_N} \int_\Lambda \frac{\ell(\underline{X}_N \mid \underline{\theta}) - H(\underline{X}_N \mid \underline{\theta})}{N} dw(\underline{\theta})$$

where $w(\underline{\theta})$ is a probability measure.

A code for which $\mathcal{R}_w$ approaches zero is called weighted universal.

Definition 2. Maximin Redundancy

$$\underline{\mathcal{R}} = \lim_{N \to \infty} \sup_{w \in W} \inf_{C_N} \int_\Lambda \frac{\ell(\underline{X}_N \mid \underline{\theta}) - H(\underline{X}_N \mid \underline{\theta})}{N} dw(\underline{\theta})$$

If $\underline{\mathcal{R}}$ is zero, the code sequence which approaches it is called maximin-universal.

Definition 3. Minimax Redundancy

$$\overline{\mathcal{R}} = \lim_{N \to \infty} \inf_{C_N} \sup_{\underline{\theta} \in \Lambda} \left[\frac{\ell(\underline{X}_N \mid \underline{\theta}) - H(\underline{X}_N \mid \underline{\theta})}{N} \right]$$

If the minimax redundancy $\overline{R}$ is zero, then the user can transmit the source output at a rate per symbol arbitrarily close to the source entropy rate uniformly over all values $\underline{\theta} \in \Lambda$ (unlike the other forms of universality) by taking N large enough. $\overline{\mathcal{R}} = 0$ is the strongest and most desirable condition. However, in some cases it is too much to hope for. If a sequence of codes for which

$\overline{\mathcal{R}} = 0$ does exist, it is called minimax-universal.

It can be shown that:

Theorem 1. The necessary and sufficient condition for the existence of weighted universal codes is that the per letter average mutual information between the message ensemble $\{x_i\}$ and $\underline{\theta}$ be zero.

Theorem 2. Maximin universal codes exist if and only if the per letter channel capacity between $\underline{\theta}$ and $\{x_i\}$ is zero.

Theorem 3. A necessary and sufficient condition for the existence of minimax universal codes is that there exist a mixture probability mass function $p(\underline{x}_N)$ for which the conditional mutual information be zero uniformly in $\underline{\theta}$ and N, i.e.

$$\lim_{N \to \infty} \frac{1}{N} \sup_{\underline{\theta} \in \Lambda} I(\underline{x}_N, \underline{\theta}) = 0$$

The determination of the existence or non-existence minimax universal codes is mot easily made by searching the class of weighted-universal codes. This is analogous to decision theory where Bayes rules are more easily formed than minimax rules so that minimax rules are constructed by searching the class of admissible Bayes rules by finding the Bayes rule with respect to a "best guess" least favorable prior.

For the binary example presented here:

$$\frac{1}{N} I(\underline{X}_N, \underline{\theta}) = \frac{1}{N} \sum_{n=0}^{N} \binom{N}{n} \theta^n (1-\theta)^{N-n} \log\left[(N+1)\binom{N}{n} \theta^n (1-\theta)^{N-n}\right]$$

$$= \frac{\log(N+1)}{N} - \frac{1}{N} \times \left[\begin{array}{l}\text{entropy of a binomial random} \\ \text{variable with parameter } \theta\end{array}\right]$$

$$\leq \frac{\log(N+1)}{N}$$

$$\to 0$$

$$N \to \infty$$

Hence, the code constructed by taking θ uniform results in a minimax universal code, as well as a maximum and weighted universal code.

Summary. The preceeding has charted the basic concepts of universal coding for arbitrary information sources. Studies are underway to

apply universal coding to theoretical image models and to digitized images.

3.7 Analog Real Time Implementation of Transform Image Coders*

Lloyd R. Welch

The invention of surface wave acoustic devices has made possible the design of tapped dely line filters of long lengths. (Filters with several hundred delays are practical.) Since an N point transform can be thought of as N, tapped delay line filters, the thought naturally arises as to the possibility of the real time mechanization of transforms using surface wave devices.

Since N distinct surface wave substrates with associate driving circuits would be expensive, various investigators have considered multiple use of a delay line by connecting the taps through semiconductor switches. A switch arrangement is shown in figure 1 when the transform coefficients are $\pm$ 1, 0. The number of such taps is 2N-1. To produce the inner product of the signal with the first row of the transform matrix at time 1, $S_{\pm}^{1}$ through $S_{\pm}^{N}$ are appropriately activated while $S_{\pm}^{k}$ for k >M are all off. To produce the inner product with the second row at time 2, $S_{\pm}^{2}$ to $S_{\pm}^{N+1}$ are appropriately activated while $S_{\pm}^{1}$ and $S_{\pm}^{k}$ for k >N+1 are all off. It is somewhat doubtful, however, that the signal amplitude out of the switches can be controlled well enough to produce an accurate summation. Furthermore, when the transform has coefficients with amplitudes other than unity, the switches cannot be controlled well enough to produce these amplitudes accurately.

Recent studies by material science investigators have shown that it is not difficult to lay several paths on a substrate driven by a single transducer. This offers an alternative configuration where each row of the transform matrix corresponds to a delay line channel on the substrate with its own tap configuration. Since the tap apertures can be precisely cut, summation coefficients can be controlled easily to less than one percent. Several problems remain, however. Multichannel devices result in parallel output. If several outputs are desired, the tap configuration on consecutive delay lines must be offset by one sample time, thus the total amount of delay

* supported in part by the Air Force Office of Scientific Research

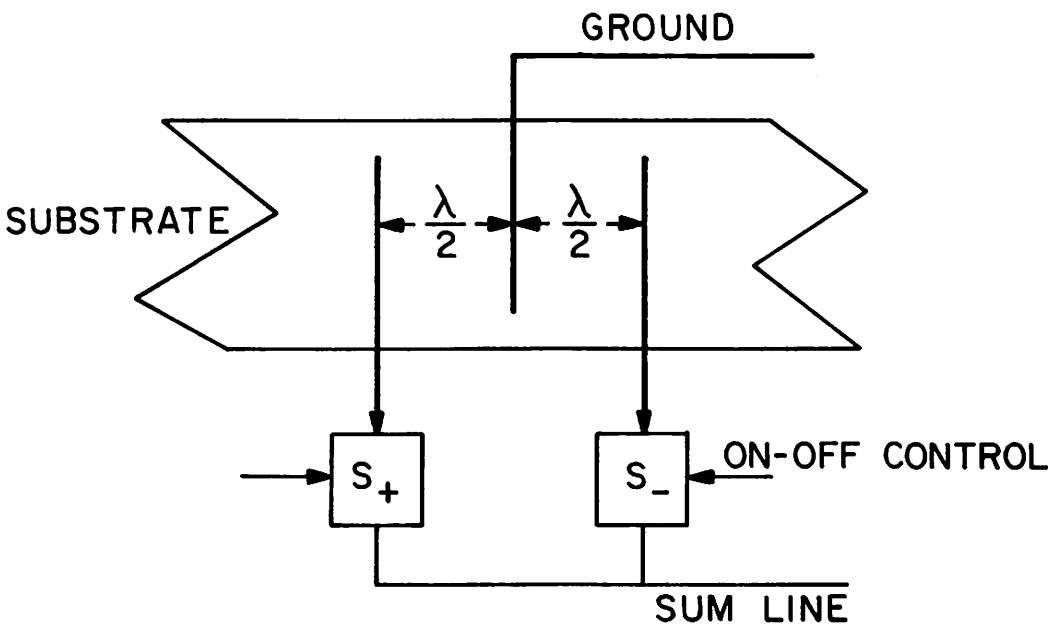


Figure 3.6-1. **Acoustic surface wave delay line taps**

needed on the substrate is still 2N-1 units of sample time. A second problem not so easily solved is as follows: The signal on the substrate is g(t) cos (2 πft) where f is the carrier frequency and g(t) is either the signal to be processed or the signal with a bias term added. In either case, the tap configuration on many of the channels will be orthogonal to the bias term and the output $\hat{g}$ cos (2πft) may have a negative value for $\hat{g}$. Therefore a simple amplitude detector will preserve the absolute value but lose the algebraic sign. The method used to overcome this problem will depend on what further analogue digital processing will be done. This problem is currently under investigation.

4. Image Enhancement and Restoration Projects

Image enhancement, as defined here, is the manipulation of pictoral data for improvement in presentation to a human viewer, or for simplification in format for subsequent machine processing. In image enhancement processes there is no overt attempt to preserve or improve fidelity. Image restoration processes attempt to reconstruct an image according to some pre-conceived notion of what the image should have been if there had been no distortion in its formation. The latter problem of image restoration is amenable to hard analysis, while the former problem of image enhancement is usually subject to heuristic approaches. Image enhancement and restoration research topics considered during the past six months are summarized below.

The first three research tasks deal with the application of generalized Wiener filtering techniques to the restoration of images that are statistically described in terms of covariance functions of the original image and the noise or interference. In the first report consideration is given to restoration of images subjected to a linear, but not necessarily space invariant, blurring and additive noise. A process has been developed for sequential row and column filtering of the image which yields substantial improvements in image quality. The second report is concerned with the computational aspects of generalized Wiener filtering for images subjected to additive noise. A procedure for restricting the Wiener filter to be causal, i.e., only operating on past data samples, is given. It is demonstrated that the increase in reconstruction error under a causality condition is small if Hadamard transform preprocessing of the data is employed. The third

report deals with the combination of the scalar Wiener filtering and the homomorphic filtering concepts. A logarithmic operation is performed on the image to separate multiplicative interference from an image. Linear Wiener filtering is then employed to minimize the interference, followed by an exponentiation operation to restore the image. Significant improvements in the detection of image detail are demonstrated.

The following two reports discuss the problem of image restoration for images subjected to linear space variant degradation. It is shown that, in a wide class of problems, the space variant operator can be cast into the cascade of an invertible geometric operator, a linear space invariant system, and another invertible geometric operator. Such systems can be "undone" and restored by conventional space invariant techniques such as scalar Wiener filtering. An example of space variant restoration is given for an image subjected to motion during film exposure.

The next report outlines some beginning attempts at image restoration by nonlinear filtering. It is pointed out that conventional linear filtering often results in non-physically realizable reconstructions, i.e., the representation of negative intensities. A nonlinear recursive technique is developed for image restoration, and examples of its performance are presented.

The last report is in the form of an inserted reprint from the July 1972 issue of the IEEE Spectrum. This paper outlines several image enhancement techniques and discusses the uses of pseudocolor for image enhancement.

4.1 Image Deblurring by Generalized Wiener Filtering

William K. Pratt

The concept of generalized Wiener filtering previously developed [1] for the restoration of images degraded by additive sensor noise can be extended to compensate for spatial blurring of the image as well as noise.

Image Degradation Model. Figure 1 contains a block diagram of a model for image degradation. The ideal continuous image is subjected to a spatial blurring mechanism which is assumed linear, but not necessarily spatially invariant. Next, the blurred image is combined with additive noise that is uncorrelated with the image, and the blurred and noisy image is then spatially sampled and quantized for digital processing. The spatially sampled image function can be expressed as

$$f_S(x, y) = [f_B(x, y) + n(x, y)]\, s(x, y) \tag{1}$$

where the blurred image is given by the superposition integral

$$f_B(x, y) = \int_{-\infty}^{\infty} f_I(\alpha, \beta) b(x, y; \alpha, \beta) d\alpha\, d\beta \tag{2}$$

with $b(x, y; \alpha, \beta)$ representing the impulse response of the physical blur mechanism. The sampling array

$$s(x, y) = \Delta x \Delta y \sum_{m=-M}^{M-1} \sum_{n=-N}^{N-1} j(x - m\Delta x, y - n\Delta y) \tag{3}$$

is assumed to be composed of $(M \cdot N)$ identical sample pulses $j(x, y)$, arranged in a grid of spacing $(\Delta x, \Delta y)$. The image at the sample points

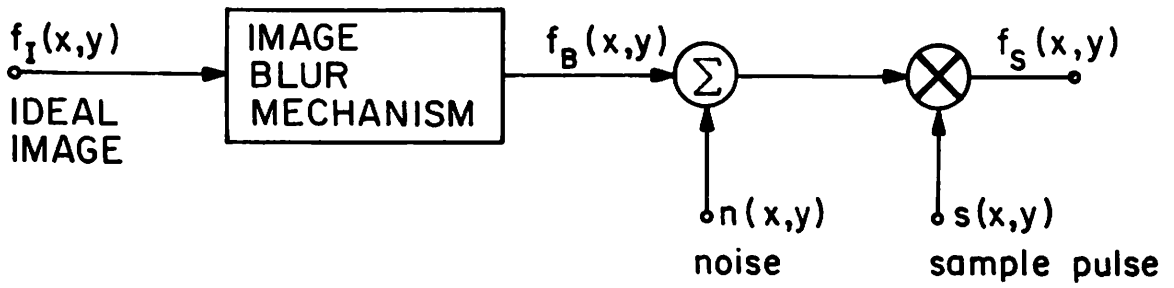


Figure 4.1-1 Model for image degradation

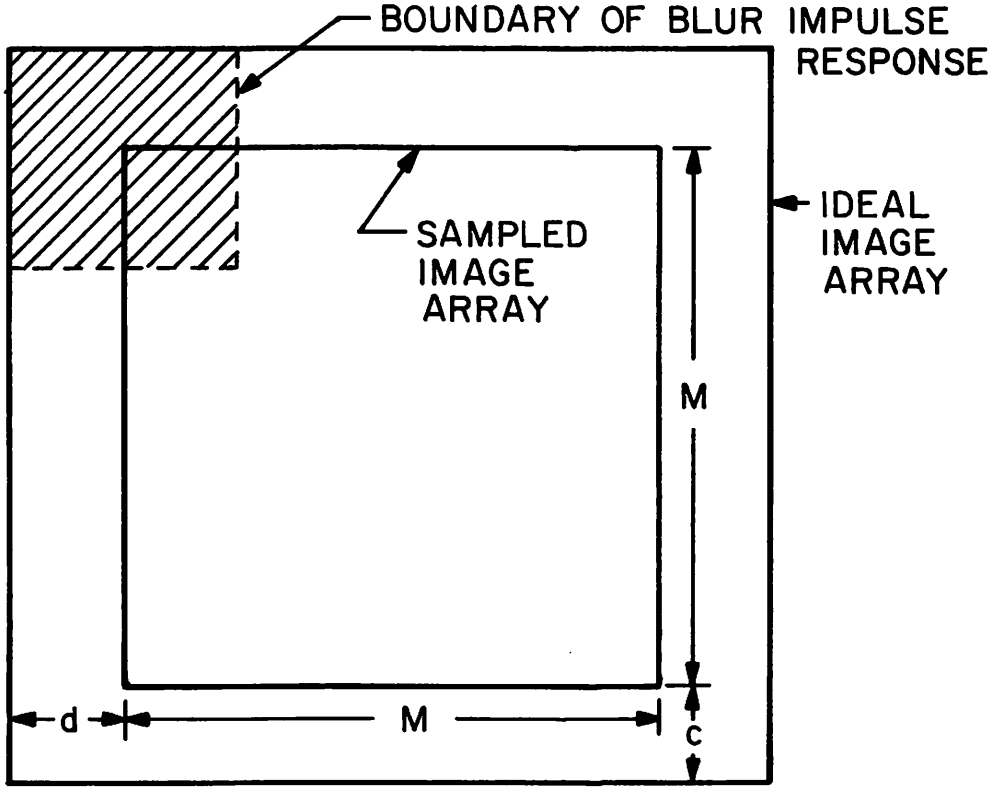


Figure 4.1-2. Geometric arrangement of image sample points

$(j = m\Delta x, k = n\Delta y)$ can be approximated as the double summation

$$f_S(j,k) = \sum_{p=j-c}^{j+c} \sum_{q=k-d}^{k-d} f_I(p,q)b(j,k;p,q) + n(j,k) \qquad (4)$$

where the constants (c, d) define an enclosing boundary for the blur mechanism impulse response. Figure 2 illustrates the geometric arrangement of sample points in the ideal image, blur impulse response, and the sampled blurry image. It should be noted that the sample points outside the physical sampling array contribute to samples of the blurred image.

In general the blur impulse response is a four dimensional function that is spatially convolved with the ideal image. If the blur function is separable in orthogonal coordinates, i. e.

$$b(j,k;p,q) = b_j(j;p)b_k(k;q) \qquad (5)$$

then the rows and columns of the image array can be processed sequentially. In this case eq. (4) can be expressed in matrix form as

$$[f_S(j,k)] = [b_j(j;p)][f_I(p,q)][b_k(k;q)]^T + [n(j,k)] \qquad (6)$$

The development of the generalized Wiener filter for image deblurring and noise smoothing presented here is limited to the sequential processing of rows and columns of an image. The more general case of two dimensional filtering is under investigation.

Sequential Row and Column Wiener Filtering. For notational simplicity let

$\underline{s}$ = P × 1 column vector of ideal image row or column

$\underline{f}$ = M × 1 column vector of observed image row or column

$\underline{n}$ = M × 1 column vector of noise along image row or column

$\underline{b}$ = M × P matrix of blur impulse response along image row or column

Then, the observed image vector can be expressed as

$$\underline{f} = \underline{b}\,\underline{s} + \underline{n} \tag{7}$$

In the generalized Wiener filtering solution to the image restoration problem, as illustrated in Figure 3, an estimate $\hat{\underline{s}}$ of the ideal image line $\underline{s}$ is obtained by performing a unitary transform on the observed image line $\underline{f}$, multiplying the transformed data vector by a filter matrix $\underline{G}_A$, and performing an inverse transform. The resulting output vector is then

$$\hat{\underline{s}} = \underline{A}^{-1}\underline{G}_A\underline{A}\,\underline{f} \tag{8}$$

The reason for introducing the unitary transform $\underline{A}$ is to decorrelate the observed data vector and permit subsequent computational simplifications in the filter function multiplication. Transforms that have been considered include the Fourier, Hadamard, and Karhunen-Loeve transforms. This subject has been covered in Ref. [1] and is extended in the following section for the case of additive sensor noise, but no image blur. In developing the optimum filter matrix $\underline{G}_A$ for a particular unitary transform matrix $\underline{A}$, it is possible to compute the optimum filter matrix for the identity transform, $\underline{G}_I$, and then perform the two dimensional transformation

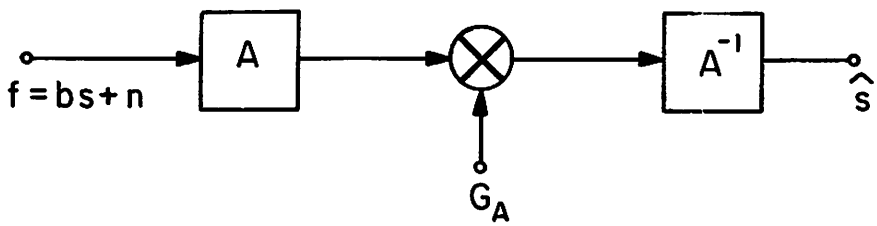


Figure 4.1-3. Generalized Wiener Filtering

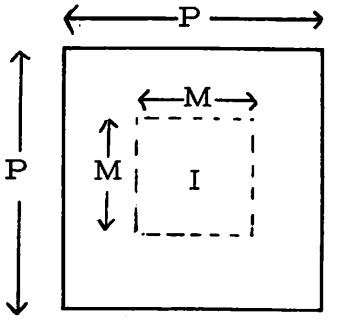
$$\underline{G}_A = \underline{A}\,\underline{G}_I\underline{A}^{-1} \tag{9}$$

The optimum Wiener filter matrix is, by definition, the filter matrix that minimizes the mean square error between $\underline{s}$ and its estimate $\hat{\underline{s}}$. However, the estimate will only be used over the vector length of the observed data. Thus, the mean square error may be written as the trace of the covariance matrix of the error over the data vector length.

$$e = \mathrm{Tr}\left\{\overline{[\underline{k}(\underline{s}-\hat{\underline{s}})][\underline{k}(\underline{s}-\hat{\underline{s}})]^T}\right\} \tag{10}$$

where k is an identity matrix of dimension M embedded in the center of a matrix of null elements of dimension P

$\underline{k}$ =



By the orthogonality condition the mean square error is minimum when

$$\overline{[\underline{k}(\underline{s}-\hat{\underline{s}})]\underline{f}} = 0 \tag{11a}$$

or

$$\overline{[\underline{k}(\underline{s}-\underline{G}_I\underline{b}\,\underline{s} - \underline{G}_I\underline{n})][\underline{b}\,\underline{s} + \underline{n}]^T} = 0 \tag{11b}$$

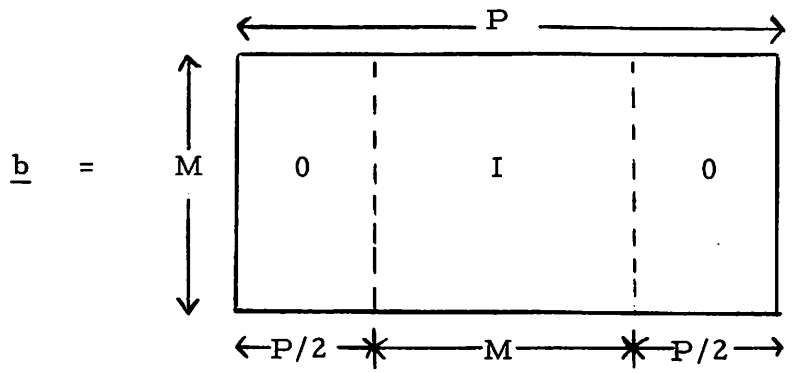
For the case in which the ideal image and noise are uncorrelated, the optimum filter matrix is found to be

$$\underline{G}_I = \underline{k}\underline{C}_s\underline{b}^T[\underline{b}\underline{C}_s\underline{b}^T + \underline{C}_n]^{-1} \tag{12}$$

and the resulting mean square error is

$$(e)_{min} = \underline{k}\underline{C}_s\underline{k} - \underline{G}_I\underline{b}\,\underline{C}_s\underline{k} \tag{13}$$

In the limiting case in which there is no blur, the blur matrix becomes



The optimum Wiener filter

$$\underline{G}_I = \underline{k}\underline{C}_s\underline{k}[\underline{k}\,\underline{C}_s\underline{k} + \underline{C}_n]^{-1} \tag{14}$$

then reduces to the form derived in Ref. [1] for no blur. If there is no image noise the optimum Wiener filter becomes

$$\underline{G}_I = \underline{k}\,\underline{C}_s\underline{b}^T[\underline{b}\,\underline{C}_s\underline{b}^T]^{-1} \tag{15}$$

Image Restoration Example. Consider the restoration process when the image lines are statistically described by a first order process with a covariance function

$$C_s\{i,j\} = \rho^{|i-j|} \tag{16}$$

and the noise is white with a covariance function

$$C_n\{i,j\} = \frac{1}{SNR}\,\delta(i,j) \tag{17}$$

where SNR is the image signal-to-noise ratio. The blur function is assumed to be Gaussian in form

$$b\{i,j\} = \exp\{-(i-j)^2/B_L\} \tag{18}$$

Figure 4 is a plot of the variance of each pixel of the image line estimate as a function of signal-to-noise ratio and blur for an image line section of 32 pixels and an observation of 16 pixels. The image correlation factor is $\rho = 0.8$. The reason for utilizing only the center 16 pixels of the 32 pixel estimate is readily apparent from Figure 4.

A series of simulations have been performed to evaluate the filtering process. An original image of 256 by 256 pixels and 6 grey levels was filtered over its entire size with the filter function of eq. (18) using Fourier domain technique. Numerically generated, zero mean Gaussian noise was added to each pixel value. The resultant blurry and noisy image was then requantized to 64 grey levels to serve as an input to the restoration filter. Figures 5a, 5c, and 5e are photographs of blurred and noisy originals. The corresponding restorations are shown in Figures 5b, 5d, and 5f. It is clear that the restoration filter has resulted in a substantial visual improvement. Also, there is no apparent checkering effect in the images even though the filtering is limited to 16 x 16 pixel blocks.

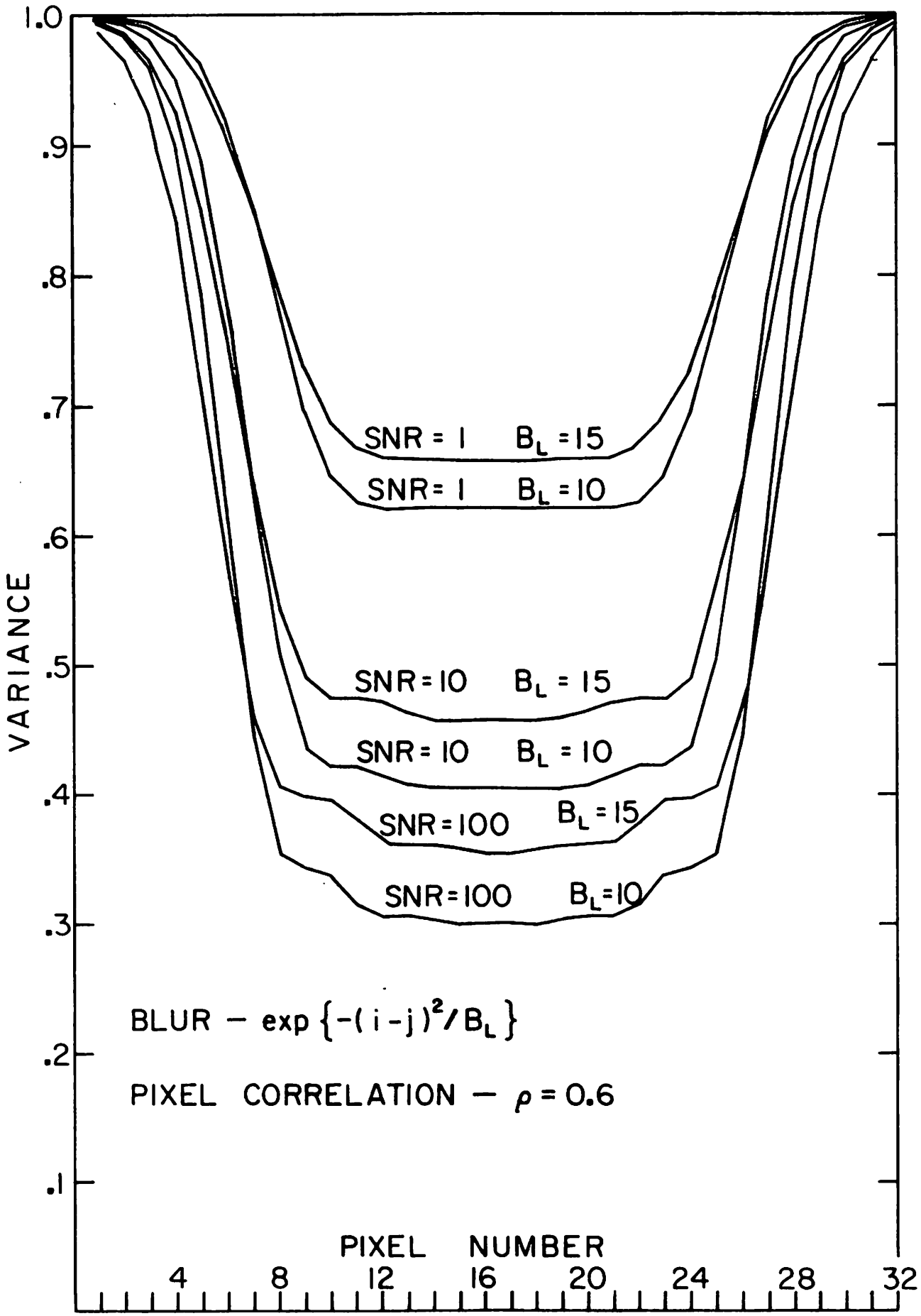


Figure 4.1-4. Variance of pixel estimates

(a) Original

$\frac{S}{N} = \infty$, B = 10

(b) Restoration

$\frac{S}{N}$ = 100, B = 10, R = .6

(c) Original

$\frac{S}{N}$ = 10, B = 5

(d) Restoration

$\frac{S}{N}$ = 10, B = 5, R = .8

(e) Original

$\frac{S}{N}$ = 10, B = 10

(f) Restoration

$\frac{S}{N}$ = 10, B = 10, R = .6

Figure 4.1-5. Generalized Wiener filtering image deblurring examples.

Reference

1. W. K. Pratt, "Generalized Wiener Filtering Computation Techniques", IEEE Transaction on Computers, Vol. C-21, No. 7, July, 1972, pp. 636-641.

4.2 Causal Generalized Wiener Filtering for Image Restoration

Nelson D. A. Mascarenhas

Generalized Wiener filtering computational techniques have been used with excellen results in one and, more recently, two dimensional processing. A typical example of the last case is the problem of the restoration of an image corrupted by additive noise [1]. It has also been shown that suboptimal filtering can be done, with reduced computational demands, without severe degradation in performance. This section presents the derivation of a suboptimal Wiener filter satisfying the additional constraint that the matrix multiplication that performs the filtering must be causal or with a preassigned delay. Major advantages result from this restriction. First, the number of operations on the filter matrix multiplication can be considerably decreased. Second, the delay in starting this operation is reduced. This also leads to a further increase on the speed of the overall filtering operation.

Derivation of the Suboptimal Filter. The discrete causal filter for a mean square error criterion was first derived by Friedland [2]. The concept is extended here by permitting unitary preprocessing on the data. Figure 3 of Section 4.1 shows the block diagram of a generalized Wiener filter in a system for sequential image row and column processing. In this system $\underline{f}$ is the input (Mx1) data vector, $\underline{A}$ is a unitary (MxM) matrix

and $\underline{G}_A$ is the (M×M) filter matrix. No assumption is made on the stationarity of the input process, but it is assumed, for simplicity, that the signal and noise are uncorrelated, although the analysis can easily be extended to the other case. The covariance matrices of the signal ($\underline{C}_s$) and noise ($\underline{C}_n$) are assumed known.

The design of the filter $\underline{G}_A$ is based on minimizing the mean square error

$$e = \overline{\mathrm{Tr}\{(\underline{s}-\hat{\underline{s}})(\underline{s}^*-\hat{\underline{s}}^*)^T\}}$$

However, since $\underline{A}$ is a unitary transformation, it is equivalent to minimize

$$e = \overline{\mathrm{Tr}\{(\underline{S}-\hat{\underline{S}})(\underline{S}^*-\hat{\underline{S}}^*)^T\}}$$

Observe that if $\underline{f}$ is a stationary process, the process $\underline{F}$ may be nonstationary. Also, if the matrix $\underline{A}$ represents a time invariant filter (i.e. $\underline{A}(i,j) = A(i-j)$) then $\underline{F}$ will be stationary.

The optimum noncausal filter has been derived as [1]

$$\underline{G}_A = \underline{C}_S[\underline{C}_S + \underline{C}_N]^{-1}$$

where

$$\underline{C}_S = \underline{A}\,\underline{C}_s\underline{A}^{-1} = \underline{A}\,\underline{C}_s\underline{A}^{*T}$$

and

$$\underline{C}_N = \underline{A}\,\underline{C}_n\underline{A}^{-1} = \underline{A}\,\underline{C}_n\underline{A}^{*T}$$

In order to obtain an expression for $\underline{G}_A$ in the suboptimal filter, first consider the case where no delay is allowed on the product of $\underline{G}_A$ by $\underline{F}$. This restricts

$\underline{G}_A$ to be a lower triangular matrix. A direct approach to obtain $\underline{G}_A$ in this case would consist in imposing the orthogonality between the error $(\underline{S}-\hat{\underline{S}})$ and the data $(\underline{F})$ vectors. In the general, non-stationary case, a set of $1 + 2 + \cdots + M = M(M+1)/2$ linear equations would result in the same number of elements of $\underline{G}_A$ below or on the main diagonal. In the stationary case this number would be reduced to M. However, another approach can be taken, giving more insight into the problem and allowing the use of a general procedure for any delay. Starting from the result obtained for the noncausal filter, observe that $(\underline{C}_S + \underline{C}_N)$ is a positive definite matrix and, consequently, it can be uniquely decomposed as [3].

$$\underline{C}_S + \underline{C}_N = \underline{L} \cdot \underline{U}$$

where $\underline{L}$ is a lower triangular matrix and $\underline{U}$ its conjugate transpose. Thus, the optimum noncausal filter $[\underline{G}_A]_{onc}$ assumes the form

$$[\underline{G}_A]_{onc} = \underline{C}_S \underline{U}^{-1} \underline{L}^{-1}$$

where $\underline{U}^{-1}$ is upper triangular and $\underline{L}^{-1}$ is its conjugate transpose or simply transpose if $\underline{F}$ is a real process.

Now, consider the process

$$\underline{W} = \underline{L}^{-1} \underline{F}$$

The covariance matrix of $\underline{W}$ is given by

$$\underline{C}_W = \underline{L}^{-1} \underline{C}_F (\underline{L}^{-1})^{*T} = \underline{L}^{-1} \underline{C}_F \underline{U}^{-1}$$

But

$$\underline{C}_F = \underline{C}_S + \underline{C}_N = \underline{L}\,\underline{U}$$

so that

$$\underline{C}_W = \underline{L}^{-1}\underline{L}\,\underline{U}\,\underline{U}^{-1} = \underline{I}$$

Thus it is concluded that $\underline{L}^{-1}$ performs a whitening filter operation on $\underline{F}$. Since $\underline{L}^{-1}$ is a lower triangular, this is equivalent to the statement that $\underline{L}^{-1}$ represents the Gram-Schmidt orthogonalization process on the data $\underline{F}$. The problem now has been transferred to the task of designing an optimum causal discrete filter operating on the white noise process $\underline{W}$. This can be done by first examining the optimum noncausal filter operating on $\underline{W}$, which is given by

$$[\underline{G}_A]_{onc}\,\underline{L} = \underline{C}_S\underline{U}^{-1}$$

The product $\underline{C}_S\underline{U}^{-1}$ can be decomposed as a sum of a lower triangular and an upper triangular matrices

$$\underline{C}_S\underline{U}^{-1} = [\underline{C}_S\underline{U}^{-1}]_L + [\underline{C}_S\underline{U}^{-1}]_U$$

The first term operates only on past and present values of W, which are uncorrelated with future values of $\underline{W}$. This means that the best that can be performed when no future values of $\underline{W}$ are to be operated upon, is to take the optimum causal filter $[\underline{G}_A]_{oc}$ as

$$[G]_{oc}\,\underline{L} = [\underline{C}_S\underline{U}^{-1}]_L$$

or

$$[\underline{G}]_{oc} = [\underline{C}_S\underline{U}^{-1}]_L\,\underline{L}^{-1}$$

Observe that the procedure of taking $[\underline{G}_A]_{oc} = [\underline{C}_S\underline{U}^{-1}\underline{L}^{-1}]_L$ is, in general, wrong, since $\underline{F}$ is not necessarily white (an exception occurs when $\underline{A}$ is the Karhunen-Loeve transformation) so that future values of $\underline{F}$ are correlated with past or present values of this process.

The Suboptimal Filter with Non-Zero Delay. When a delay of D units is allowed in the operation, it is possible to use values of the process $\underline{W}$, D units of time after the present time. Then the optimum filter under the constraint of delay D, $(\underline{G}_A)_{od}$ has the expression

$$(\underline{G}_A)_{od}\,\underline{L} = [\underline{C}_S\underline{U}^{-1}]_{LD}$$

or

$$(\underline{G}_A)_{od} = [\underline{C}_S\underline{U}^{-1}]_{LD}\,\underline{L}^{-1}$$

where $[\underline{C}_S\underline{U}^{-1}]_{LD}$ includes the elements of a lower triangular matrix and subdiagonals up to the delay D.

The number of multiply and add operations involved in the filter matrix multiplication $[\underline{G}_A]_{od}\ \underline{F}$ when a delay $D(0 \le D \le M-1)$ is allowed can now be calculated. The first (M-D-1) rows of $[\underline{G}_A]_{od}$ (from top to bottom) have $(M-1),\ldots,(D+1)$ nonzero entries. Each of the last (D+1) rows have M nonzero entries, so that the total number is given by

$$(D+1)M + \frac{(M-D-1)}{2}(D+M) = \tfrac{1}{2}(M(M+D) + (D+1)(M-D))$$

Calculation of the Additional Mean Square Error. When the optimum noncausal generalized Wiener filter is used, it can be shown that the mean square error is [1].

$$e_{min} = Tr\{\underline{C}_S\underline{C}_N(\underline{C}_S+\underline{C}_N)^{-1}\}$$

Now consider the error when a constraint of delay D is imposed on the filter. In this case the error is expressed by

$$e = Tr\{\overline{(\underline{s}-\hat{\underline{s}}_D)(\underline{s}^*-\hat{\underline{s}}_D^*)^T}\} = Tr\{\overline{(\underline{S}-\hat{\underline{S}}_D)(\underline{S}^*-\hat{\underline{S}}_D^*)^T}\}$$

The error can also be given by

$$e = Tr\{\overline{(\underline{S}-\hat{\underline{S}}+(\hat{\underline{S}}-\hat{\underline{S}}_D))(\underline{S}^*-\hat{\underline{S}}^*+(\underline{S}^*-\hat{\underline{S}}_D^*))^T}\}$$

where $\hat{\underline{S}}$ is the estimate for the optimal noncausal filter. Taking into consideration that $\underline{S}-\hat{\underline{S}}$ is uncorrelated with any linear function of the input data, in particular $\hat{\underline{S}}-\hat{\underline{S}}_D$, one obtains

$$e = e_{min} + Tr\{[\underline{G}_A]_{dif}\ (\underline{C}_S + \underline{C}_N)[\underline{G}_A]_{od}$$

where

$$[\underline{G}_A]_{dif} = [\underline{G}_A]_{onc} - [\underline{G}_A]_{od}$$

The fact that $(\underline{C}_S + \underline{C}_N)$ is a positive definite matrix guarantees that the error is strictly increased for any nonmaximum allowed delay.

The mean square error for different transformation varying the preassigned delay from 0 to 15 has been calculated, for a one-dimensional (16 x 1) input data vector. The input process consists of a Markov process signal in the presence of white noise as defined in Section 4.1. It can be observed from the curves of Figure 1 that only a small price, in terms

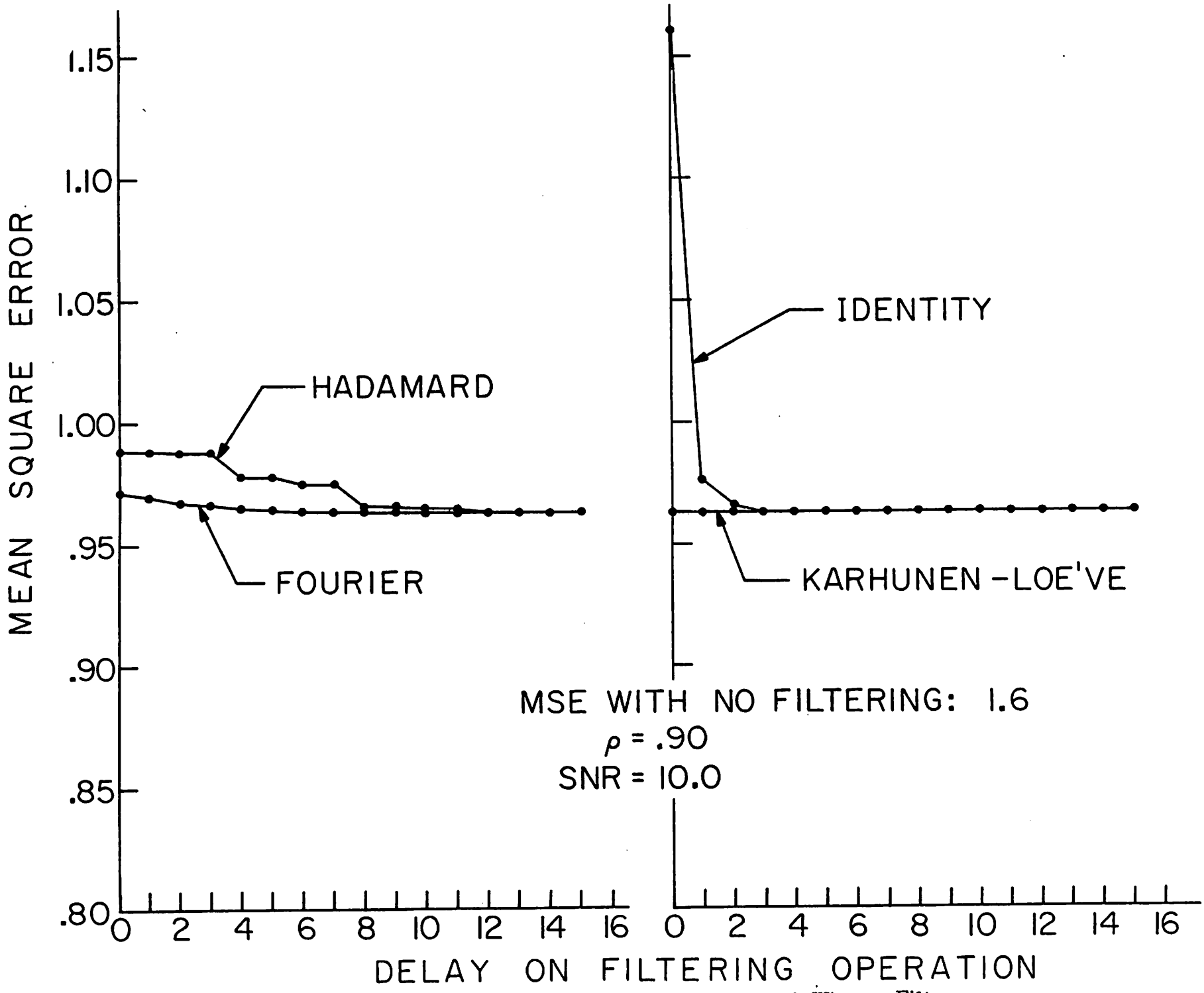


Figure 4.2-1. Mean Square Error Performance of Causal Wiener Filters

of mean square error, is paid by allowing a few units of delay or even no delay, for any of the considered transforms. This is due to the concentration of energy on and around the main diagonal of the filter matrix. The most severe increase in mean square error occurs for the identity operator, followed by the Hadamard transform. As should be expected, the Karhunen-Loeve filter, being diagonal, is not affected by the restriction of causality.

However, the final judgement for the selection of the best transformation requires a joint consideration of the curves of Figure 1 and the number of operations involved. Another possible element for consideration comes from a detailed analysis (beyond the scope of the present paper) of numerical errors in computing the inverse of $(\underline{C}_S + \underline{C}_N)$ in the general case, for different transformations.

It should be observed that, although the present method results in causal operation only on the matrix $\underline{G}_A$, in the case of the identity operator, the overall transformation is causal.

Finally, consider the fact that the $\underline{L}\,\underline{U}$ factorization method (also called square root factorization or Cholesky method) suggests an efficient way of inverting the covariance matrix $(\underline{C}_S + \underline{C}_N)$, namely, obtaining the lower triangular matrix $\underline{L}$, inverting it and multiplying $(\underline{L}^{-1})^{*T}$ by $\underline{L}^{-1}$.

Experimental Results. The suboptimal filter has been used to enhance a 256 by 256 pixel, 64 grey level picture of a toy tank shown in Figure 2a. In Figure 2b, white Gaussian noise has been added under the condition of unity signal to noise ratio. This was done by setting the variance of the

(a) Original Picture

(b) Original plus Gaussian Noise (SNR = 1.0)

(c) Optimal Hadamard Filter (delay = 15, SNR = 1.0)

(d) Causal Hadamard Filter (delay = 0, SNR = 1.0)

Figure 4.2-2. Examples of Causal Wiener Filtering

noise equal to the measured variance of the pixel values. Figures 2c and 2d show the comparison of the performance of the optimal (delay = 15) and suboptimal (delay = 0) filters, with Hadamard preprocessing. The correlation coefficient giving the best result was found to be equal to $\rho = 0.5$ for this particular picture. Hardly any difference between the two pictures was noticed.

Conclusion. The concept of causal filtering has been extended to generalized Wiener filter computations. The method is based on a factorization of the covariance matrix of the signal plus noise. It has been shown that only a small price in mean square error is paid even when small or no delay is allowed on the filter matrix multiplication. The comparison of the performance of the optimal and suboptimal filters indicate that even though the number of operators for the filter multiplication is reduced to almost half for the zero delay filter, very little, if any, visual difference is detected on the filtered pictures for different unitary transforms.

References

1. W. K. Pratt, "Generalized Wiener Filtering Computation Techniques," IEEE Transactions on Computers, Vol. C-21, No. 7, July, 1972, pp. 636-641.

2. B. Friedland, "Least Squares Filtering and Prediction of Nonstationary Sampled Data," Information and Control, Vol. 1, December, 1958, pp. 297-313.

3. E. Isaacson and H. B. Keller, Analysis of Numerical Methods, John Wiley and Sons, New York, 1966.

4. H. L. VanTrees, Detection, Estimation, and Modulation Theory, Part 1, John Wiley and Sons, New York, 1968.

4.3 Homomorphic Scalar Wiener Filtering for Image Restoration

Robert H. Wallis

Images degraded by the effects of non-uniform illumination can be dramatically restored by using a non-linear filtering technique which utilizes the response of the human visual system [1]. Experimental evidence indicates that the human eye behaves in accordance with the model shown in Figure 1. This model explains the eye's ability to accommodate an enormous dynamic range of intensities. Consider the input to the system as

$$I_{xy} = i_{xy} \cdot r_{xy} \tag{1}$$

where i_{xy} represents the illumination of the scene and r_{xy} its reflectivity. The logarithm transforms this product into the sum

$$\log(I_{xy}) = \log(i_{xy}) + \log(r_{xy}) \tag{2}$$

Since the illumination of natural scenes varies gradually across the image while the reflectivity contains much sharp detail (edges, etc.), the high pass filter tends to extract the scene from the shadows. By treating the $\log(i_{xy})$ term as additive inteference, a processor which imitates the human eye can be constructed, reducing the effects of nonuniform lighting of the scene. A block diagram of this processor is provided in Figure 2.

The homomorphic filter adds an exponential function to the visual model in order to neutralize the logarithm. The log function transforms intensities into densities. The filtered densities are then transformed back into intensities for viewing. Since

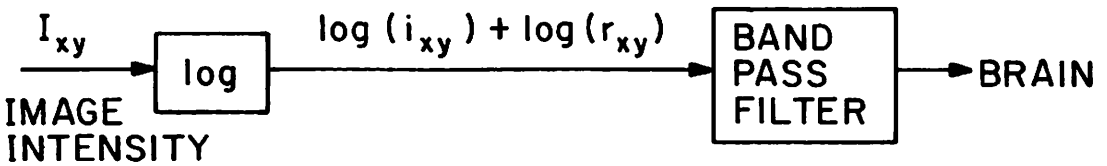


Figure 4.3-1. Model of Human Visual System

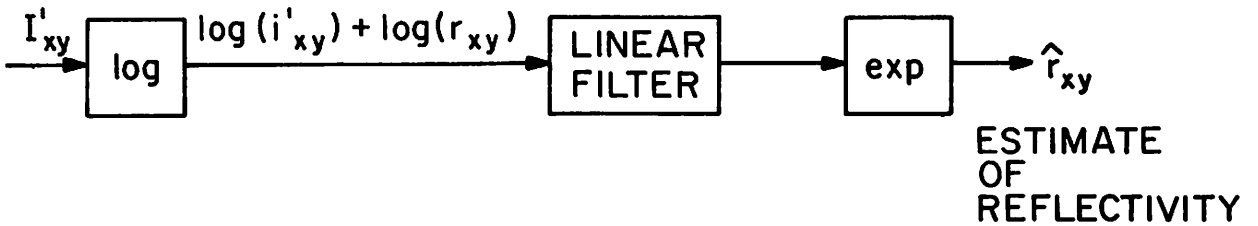


Figure 4.3-2 Homomorphic Filter

$$\log(I'_{xy}) = \log(i'_{xy}) + \log(r'_{xy}) \tag{3}$$

can be treated as signal plus additive interference, the linear filter which best estimates the signal in a mean square sense can be found using Wiener filtering theory.

Wiener filtering. Wiener filtering is a technique which estimates a signal in the presence of additive noise, or interference, with a minimum mean square error. Consider

$$\underline{y} = \underline{s} + \underline{n} \tag{4}$$

where

$\underline{y}$ = available data vector

$\underline{s}$ = signal vector to be estimated

$\underline{n}$ = interference vector

All vectors are of length N. For simplicity let the structure of the estimator be as follows

$$\underline{h}^T \underline{y} = \hat{\underline{s}}_k \tag{5}$$

That is, the estimate of the k^{th} element of the vector $\underline{s}$ is given by the vector dot product

$$\underline{h}^T \underline{y} = \sum_{i=1}^{N} \underline{h}_i y_i \tag{6}$$

The problem then, is to solve for the $\underline{h}$ that minimizes the squared error

$$E(e^2) = E[\underline{h}^T \underline{y} - \underline{s}_k]^2 \tag{7}$$

Where E denotes expected value. Defining $\nabla_{\underline{h}}$ to be the gradient with respect to $\underline{h}$, it is required that

$$\nabla_{\underline{h}}\{E[\underline{h}^T\underline{y} - \underline{s}_k]^2\} = 0 \tag{8}$$

Therefore

$$\underline{h}^T E(\underline{y}\,\underline{y}^T) = E(\underline{s}_k\underline{y}^T) \tag{9}$$

or

$$\underline{h} = [E(\underline{y}\,\underline{y}^T)]^{-1}E(\underline{s}_k\underline{y}) \tag{10}$$

Assuming the signal to be uncorrelated with the interference, and that both are zero mean, then

$$E(\underline{y}\,\underline{y}^T) = E(\underline{s}\,\underline{s}^T) + E(\underline{n}\,\underline{n}^T) \tag{11a}$$

or

$$E(\underline{y}\,\underline{y}^T) = C_s + C_n \tag{11b}$$

where C_s and C_n are the covariance matrices of the signal and interference respectively. Therefore

$$\underline{h} = [C_s + C_n]^{-1}E(\underline{s}_k\underline{s}) \tag{12}$$

An Example of Homomorphic Wiener Filtering. For simplicity assume that both the illumination densities (interference) and the reflectivity densities can be modeled by Markov processes, i.e.,

$$C_n(i,j) = E(n_i n_j) = \sigma_n^2 \rho_n^{|i-j|} \tag{13a}$$

$$C(i,j) = E(s_i s_j) = \sigma_s^2 \rho_s^{|i-j|} \tag{13b}$$

$$E(\underline{s}_k \underline{s}) = \sigma_s^2 \begin{bmatrix} \rho^{|k-1|} \\ \rho^{|k-2|} \\ \vdots \\ 1 \\ \rho \\ \rho^2 \\ \vdots \\ \rho^{|N-k|} \end{bmatrix} \tag{13c}$$

Since the interference is highly correlated (slowly varying) and the signal less correlated, ρ_n should be chosen larger than ρ_s. The following parameters were selected somewhat arbitrarily.

$$\sigma_n = \sigma_s = 1 \tag{14a}$$

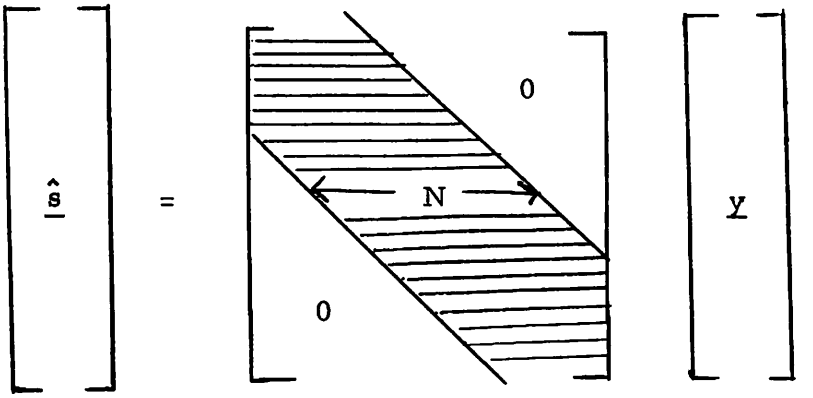
$$\rho_n = .9 \tag{14b}$$

$$\rho_s = .6 \tag{14c}$$

$$k = 8 \tag{14d}$$

$$N = 16 \tag{14e}$$

Thus, the resulting $\underline{h}$ vector is 16 elements long and estimates the 8th element of the $\underline{s}$-vector. This scalar filter can be used to filter large records of data by convolving $\underline{h}$ with a long data vector. This can be thought of as multiplication of the data array by a large circulant matrix.



(15)

Each row of the filter matrix is a shifted version of $\underline{h}$ centered at the diagonal. Of course, the first and last N/2 elements of the $\hat{\underline{s}}$ vector will be inaccurate estimates, however, for the large arrays used in image processing, this error is negligible.

The actual convolution is carried out most efficiently using the Fast Fourier Transform (FFT) algorithm [3]. The processing is extended to two dimensions by using a two dimensional FFT whose frequency domain filter function is the product of two one-dimensional filter functions, i.e.,

$$H(f_x, f_y) = H(f_x) \cdot H(f_y) \tag{16}$$

where f_x and f_y are spatial frequencies in the x and y directions.

Although the filter parameters have been chosen somewhat arbitrarily, the results are quite good. Figures 3a and 3b present original and restored images of a tank sampled on a 256 by 256 grid. Note that the restored version shows much detail not visible in the original, and seems to be more uniformly illuminated. It is expected that more careful modelling will yield even better results.

(a) Original

(b) Reconstruction

Figure 4.3-3. Homomorphic Filter

Conclusion. This is a preliminary study and neglected many important considerations such as the calibration of the intensities in the original image. Questions such as "what effect does the length of the filter's impulse response have on its performance?" remain unanswered. The results however are quite encouraging and future study on extensions of the technique are planned.

References

1. T. G. Stockham, Jr., "Image Processing in the Context of a Visual Model," Proc. IEEE, Vol. 60, No. 7, July, 1972, pp. 828-842.

2. W. K. Pratt, "Generalized Wiener Filtering Computation Techniques," IEEE Transactions on Computers, Vol. C-21, No. 7, July, 1972, pp. 636-641.

3. J. W. Cooley, P. A. W. Lewis and P. D. Welch, "The Finite Fourier Transform," IEEE Trans. Audio Electroacoustics, Vol. AU-17, June, 1969, pp. 77-85.

4.4 Image Restoration By Space-Variant Decomposition

Alexander A. Sawchuk

In the well-known linear superposition model for incoherent imaging, the degradation from object to image is given by

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

where $\mathcal{O}(\underline{u})$ is the original object intensity function, $\mathcal{I}(\underline{x})$ is the image intensity recorded by the system, and $h(\underline{x}, \underline{u})$ is the response at image coordinates $\underline{x} = (x_1, x_2)$ to a unit impulse at $\underline{u} = (u_1, u_2)$ in the object coordinates. In general, the form of $h(\underline{x}, \underline{u})$ varies with the position of the object impulse $\underline{u}$ and is called a space-variant point-spread function (SVPSF) in the context of image processing. If $h(\underline{x}, \underline{u})$ can be written as a function only

of the difference $\underline{x}-\underline{u}$, then $h(\underline{x}, \underline{u})$ is a space-invariant point-spread function (SIPSF) and (1) reduces to a convolution.

A large number of enhancement and restoration methods are available to process degraded images, but so far most of them are of practical use only when the system degradation (1) is space-invariant. Restoration in the general space-variant case is theoretically possible using brute force sampling and digitization of the functions in (1) followed by solution of the resulting tensor equation, but the enormous computational capacity required for practical realization makes such a technique useless in most cases. With these facts in mind, it is evident that the restoration of space-variant degradations might be easily accomplished if the SVPSF could be uniquely related to an equivalent space-invariant system. This approach to image restoration is referred to as space-variant decomposition.

Given a general linear SVPSF describing an optical system degradation, the basic idea behind decomposition is to express the SVPSF as an equivalent cascade of invertible geometrical distortion operations with space-invariant systems. Figure 1 illustrates the concept of a space-variant system $h(\underline{x}, \underline{u})$ and its equivalent. Because the distortions are one-to-one mappings of points between coordinate systems, they are invertible even though they may be nonlinear in the spatial coordinates. The center block is a space-invariant system $h_I(\underline{z}-\underline{v})$ which gives the transformed intermediate image $\mathcal{I}_z(\underline{z})$ by operating on the transformed object $\mathcal{O}_v(\underline{v})$. Once $h(\underline{x}, \underline{u})$ is reduced to this form, the restoration proceeds by inverting the distortions and applying computationally simple space-invariant estimation or inverse

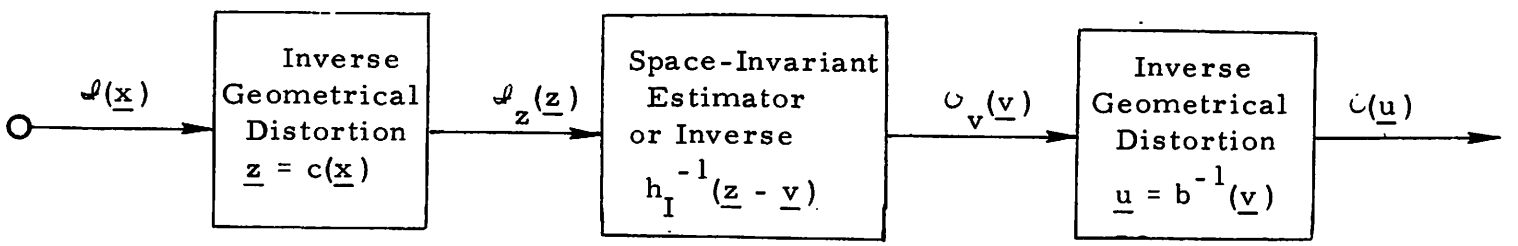


Figure 4.4-1. General Coordinate Transformation Restoration

filtering to recover the object function [3]. This idea is summarized in Figure 2 and is a generalization of techniques reported previously [2], [3].

The procedure for testing whether or not the decomposition is possible for a given SVPSF involves manipulating $h(\underline{x}, \underline{u})$ into a functional form in which the geometrical distortions and space-invariant $h_I(z-v)$ are more obvious. Simply stated, if $h(\underline{x}, \underline{u})$ is expressible in the form

$$h(\underline{x}, \underline{u}) = \alpha(\underline{x})\beta(\underline{u})h_I(c(\underline{x}) - b(\underline{u})) \tag{2}$$

where $\alpha(\underline{x})$, $\beta(\underline{u})$, $c(\underline{x})$ and $b(\underline{u})$ are one-to-one, continuous, and invertible functions, with $\alpha(\underline{x}) \neq 0$ and $\beta(\underline{u}) \neq 0$ over their domains, and $h_I(r)$ is a Fourier transformable function, then space-variant decomposition and restoration are possible. The SVPSF $h(\underline{x}, \underline{u})$ may also be put in alternative related forms from which restoration follows [1]. Writing $h(\underline{x}, \underline{u})$ in the form (2) immediately gives the geometrical transformations

$$\underline{z} = c(\underline{x}) \tag{3a}$$

$$\underline{v} = b(\underline{u}) \tag{3b}$$

and the SIPSF $h_I(\underline{r})$. Here the intermediate functions are given by

$$\mathcal{O}_v(\underline{v}) = \beta(b^{-1}(\underline{v}))J_{b^{-1}}(x)\mathcal{O}(b^{-1}(v)) \tag{4}$$

and

$$\mathcal{I}_z(\underline{z}) = \frac{\mathcal{I}(c^{-1}(\underline{z}))}{\alpha(c^{-1}(\underline{z}))} \tag{5}$$

where $J_{b^{-1}}(\underline{v})$ is a Jacobian function.

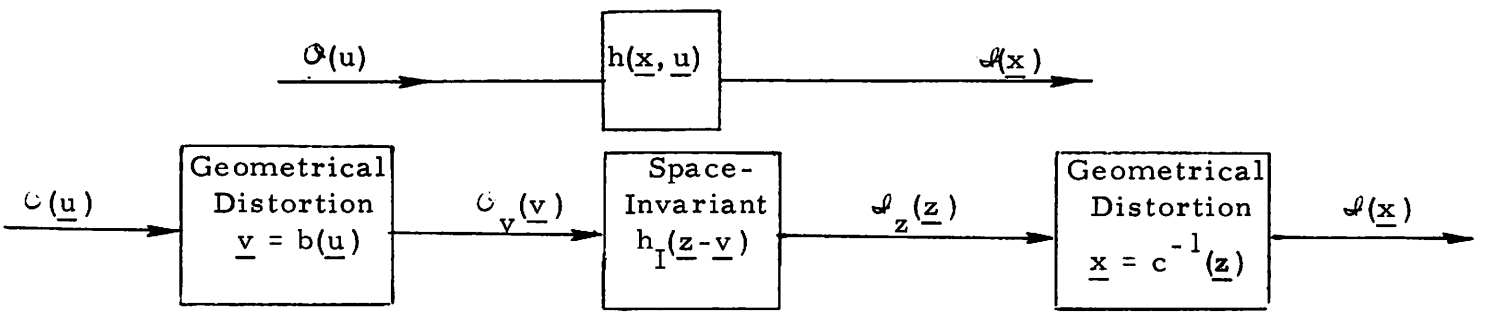


Figure 4.4-2. General Space-Variant Decomposition

By examining physical sources of optical image degradation from a system point of view, it has been found that the decomposition can be derived for a large class of space-variant imaging systems [1], [2], [5]. The application of this technique to incoherent motion blur is discussed in the following section, while the remainder of this section is concerned with the restoration of optical aberrations.

In optical imaging, degradations due to the system aberrations are usually of greater magnitude than those due to diffraction. An inherent difficulty in attempting restoration arises because these aberrations are usually quite space-variant. Robbins and Huang [3] have derived the form of the aberration point-spread function as

$$h(r_i, \theta_i; r_o, \theta_o) = \frac{1}{r_o^{2n}} h^* \left(\frac{r_i \cos(\theta_i - \theta_o) - r_o}{r_o^n} + 1, \frac{r_i \sin(\theta_i - \theta_o)}{r_o^n} \right) \tag{6}$$

where (r_o, θ_o) and (r_i, θ_i) are object and image respectively, expressed in polar coordinates for simplicity due to the circular symmetry inherent in many optical systems. In this formulation the function $h^*(x_i, y_i)$ is the response to an input impulse at standard position $r_o = 1$, $\theta_o = 0$, and the expression for the overall system operation becomes

$$\mathcal{I}(r_i, \theta_i) = \int_0^{2\pi} \int_0^{\infty} h(r_i, \theta_i, r_o, \theta_o) \mathcal{O}(r_o, \theta_o) r_o \, dr_o \, d\theta_o \tag{7}$$

with n = 0 representing spherical aberration, n = 1 coma, and n = 2 astigmatism and curvature of field.

With knowledge of the degrading expression (6), it is often possible to use the decomposition technique for space-variant restoration. Robbins and Huang [3] reported a method for the restoration of coma (n = 1) which is essentially a special case of the general technique. By making the changes of variable (the geometrical distortions in Figure 1)

$$z_1 = \ln(r_i) \tag{8a}$$

$$v_1 = \ln(r_o) \tag{8b}$$

$$z_2 = \theta_i \tag{8c}$$

$$v_2 = \theta_o \tag{8d}$$

one obtains

$$h^*\left(e^{z_1-v_1}\cos(z_2-v_2); e^{z_1-v_1}\sin(z_2-v_2)\right) = h_I(z_1-v_1, z_2-v_2) \tag{9}$$

where the $1/r_o^2$ factor from the right hand side of (6) is absorbed into $\mathcal{O}(r_o, \theta_o)$. It can be shown [1] that a Mellin transform method which accomplishes the same coma restoration is really the same decomposition technique.

The application of decomposition restoration to spherical aberration (n = 0) and other aberrations (n = 2) is not quite so straightforward but is possible for certain special cases of $h^*(x_i, y_i)$ and is a currently active research area.

The decomposition method is also useful when the degradation is a space-variant blur caused by tilt of the image plane. For this case, the

SVPSF can be represented approximately by [3]

$$h(x_1, x_2; u_1, u_2) = \frac{1}{u_1^2} h^*\left(\frac{x_1 - u_1}{u_1}, \frac{x_2 - u_2}{u_1}\right) . \qquad (10)$$

When $h^*(u_1, u_2)$ factors into functions $h_1^*(u_1)h_2^*(u_2)$ it is called separable, and for this large class of functions space-variant restoration of tilt can be accomplished by decomposition.

The restoration methods described here are very general and useful in both noisy and noise-free systems. Future work will be directed at determining other areas of application and implementing digital simulations of the technique. Several papers reporting on the details of this work are in preparation and portions will be presented at the Fall Convention of the Optical Society of America [5].

References

1. A. A. Sawchuk, "Space-Variant Image Motion Degradation and Restoration," Ph.D. Dissertation, Department of Electrical Engineering, Stanford University, Stanford, California, 1972.

2. A. A. Sawchuk, "Space-Variant Motion Degradation and Restoration," Proc. of the IEEE, Vol. 60, No. 7, July, 1972, pp. 854.

3. G. M. Robbins, and T. S. Huang, "Inverse Filtering for Linear Shift-Variant Imaging Systems," Proc. of the IEEE, Vol. 60, No. 7, July, 1972, pp. 862.

4. A. A. Sawchuk, "Linear Space-Variant Motion Degradation and Restoration," 1972 Spring Meeting, Optical Society of America, New York, April, 1972.

5. A. A. Sawchuk, "Coordinate Transformations in Space-Variant Image Enhancement and Restoration," 1972 Fall Meeting, Optical Society of America, San Francisco, October, 1972.

4.5 Restoration of Motion Degraded Images

Alexander A. Sawchuk

The coordinate transformation restoration technique described in the previous section is of particular value in the restoration of images degraded by motion. In a recent paper [1] an equivalent linear space-variant system containing all the motion effects has been derived given a mechanical description of the motion. The mechanical description has the general parametric form

$$u_1 = g_1(x_1, x_2; t) = g_1(\underline{x}; t) \tag{1a}$$

$$u_2 = g_2(x_1, x_2; t) = g_2(\underline{x}; t) \tag{1b}$$

which uniquely relates any object point $\underline{u} = (u_1, u_2)$ to the location $\underline{x} = (x_1, x_2)$ of its image in the fixed frame as a function of time for the exposure interval $[0, T]$.

Although the equivalent space-variant point-spread function (SVPSF) which can be derived from (1) is rather complicated [1], [2] it reduces to some simply expressions for common types of motion. For parallel object and image intensity function planes which translate during exposure, equation (1) takes the form

$$u_1 = g_1(\underline{x}; t) = x_1 - m_1(t) \tag{2a}$$

$$u_2 = g_2(\underline{x}; t) = x_2 - m_2(t) \tag{2b}$$

where only the coordinate difference $\underline{x}-\underline{u}$ is a time function, and the $m_i(t)$ describes the planar translation. This particular form of motion function

leads to an equivalent motion point-spread function which is space-invariant, yielding a simple solution to the restoration problem.

More complex space-variant motion blur occurs when object and image planes are not parallel. For a large class of aerial imaging systems, the motion functions (1) can be expressed as

$$b_1(u_1, u_2) = c_1(x_1, x_2) - m_1(t) \tag{3a}$$

$$b_2(u_1, u_2) = c_2(x_1, x_2) - m_2(t) \tag{3b}$$

where the $m_i(t)$ express a translation between transformed coordinate planes. Making a change of variable

$$v_1 = b_1(u_1, u_2) \tag{4a}$$

$$v_2 = b_2(u_1, u_2) \tag{4b}$$

$$z_1 = c_1(x_1, x_2) \tag{4c}$$

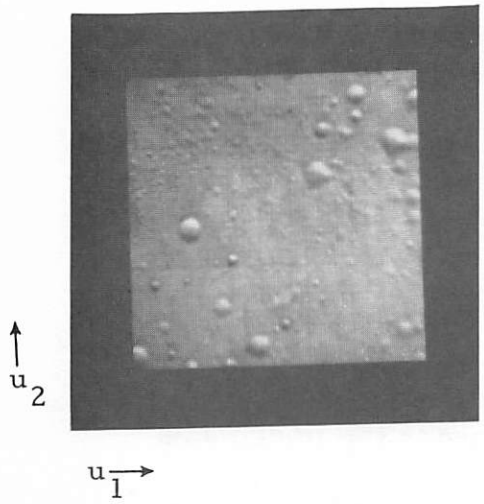
$$z_2 = c_2(x_1, x_2) \tag{4d}$$

immediately puts (3) into space-invariant form and simultaneously identifies the coordinate distortions for space-variant decomposition described in Figure 1 of Section 4.4. The general restoration method of Figure 2 of Section 4.4 is then applied directly with a space-invariant estimator or inverse filter. The space-invariant equivalent system $h_I(w_1, w_2)$ is given by

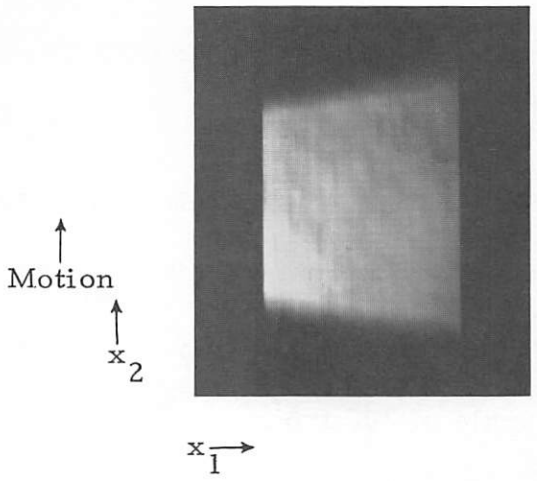
$$h_I(w_1, w_2) = \begin{cases} \dfrac{1}{\left[(m_1(t))^2 + (m_2(t))^2\right]^{\frac{1}{2}}} \Bigg|_{\substack{t=m_1^{-1}(w_1) \\ t=m_2^{-1}(w_2)}} & \begin{matrix} m_1(0) \le w_1 \le m_1(T) \\ m_2(0) \le w_2 \le m_2(T) \end{matrix} \\ 0 & \text{elsewhere} \end{cases} \tag{5}$$

Coordinate transformation restoration has been applied in a computer simulation of space-variant aerial imaging. Figure 1a is the ideal lunar object scene taken by a Ranger spacecraft. Figures 1b and 1c respectively show the motion degradation when the object is imaged by side and forward oblique tilted cameras with constant velocity translation at fixed altitude during exposure. In Figure 1b, the camera looks sideways to the left at an angle of 20^{o} from vertical. The movement smears the image as a function of position, verifying the space-variant nature of the system. Figure 1c shows the image recorded by a forward oblique system tilted forward in the direction of movement by an angle 20^{o} from vertical. As before, the blurring length and direction changes with position, and in contrast to the side oblique case there is blur in both the x_1 and x_2 directions, even though the movement is along the u_2 axis. For these two space-variant systems, it turns out that the first geometrical transformation $\underline{v} = b(\underline{u})$ Figure 1 of Section 4.4 is not present, so restoration proceeds by first using the appropriate inverse geometrical distortion $\underline{z} = c(\underline{x})$ to obtain the intermediate image $\mathcal{I}_z(\underline{z})$ which appears as Figure 1d. Using an inverse FFT filter then gives the final restored object in Figure 1e.

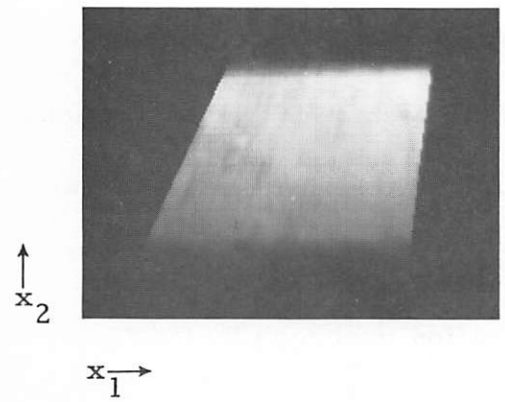
Another application of coordinate transformation restoration occurs when there is relative rotation between parallel object and image planes. The equivalent SVPSF may be derived in this case [1], where it is shown that the polar coordinate transformations



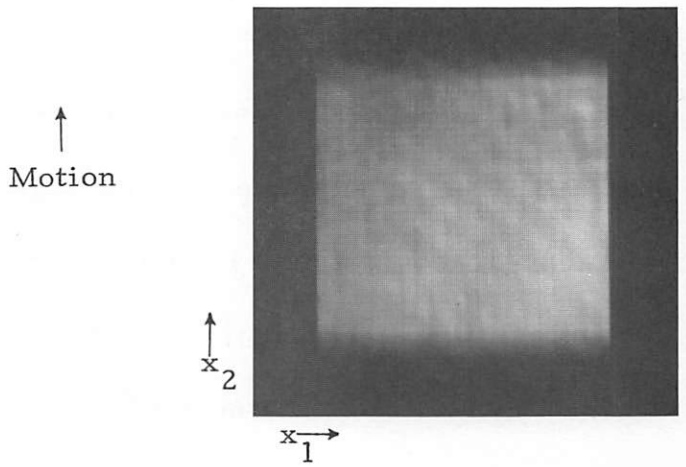
(a) Original Lunar Object



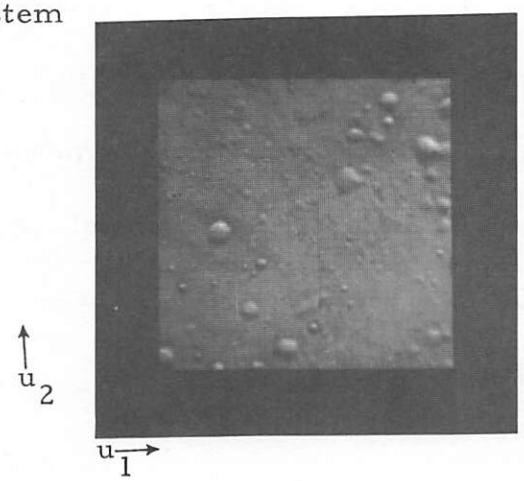
(b) Degradation by Moving Side Oblique System



(c) Degradation by Moving Forward Oblique System



(d) Intermediate Image



(e) Final Restored Object

Figure 4.5-1. Space-Variant Image Motion Restoration

$$x_1 = r_z \cos \theta_z \tag{6a}$$

$$x_2 = r_z \sin \theta_z \tag{6b}$$

$$u_1 = r_v \cos \theta_v \tag{6c}$$

$$u_2 = r_v \sin \theta_v \tag{6d}$$

used for the geometrical distortions in the model of Figure 1 of Section 4.4 convert the rotation into an equivalent space-invariant $h_I(\underline{z}-\underline{v})$ for restoration. Several computer simulations to verify the rotational restoration technique are now underway, as well as investigations of other types of decomposable space-variant motion.

References

1. A. A. Sawchuk, "Space-Variant Image Motion Degradation and Restoration," Proc. of the IEEE, Vol. 60, No. 7, July, 1972, pp. 854.

2. A. A. Sawchuk, "Space-Variant Image Motion Degradation and Restoration," Ph.D. Dissertation, Department of Electrical Engineering, Stanford University, Stanford, California, 1972.

4.6 Linear and Nonlinear Interpolation for Image Restoration

Anil K. Jain

In many parameter estimation, data interpolation and other similar situations, one has a knowledge of the lower and upper bounds of the estimates. For example, in the processing of image data, obtained by scanning a black and white picture, it is known the data should always be positive (since the intensity of reflected light at any point is ≥ 0). The estimates obtained via optimal filtering may violate these bounds due to inaccuracies in modelling, e.g., when a two dimensional image is modelled by a linear

Markov process, given the noisy observations and the bounds on the estimates themselves (or hard constraints). The problem is formulated by considering the maximum likelihood function, and the presence of these bounds (or hard constraints) leads to a nonlinear filter [1].

Consider the class of two-dimensional images represented by the following first order vector Markov process model [2],

$$\underline{x}_{k+1} = \underline{A}\,\underline{x}_k + \underline{B}\,\underline{u}_k \,, \tag{1}$$

$$\underline{y}_k = \underline{c}\,\underline{x}_k + \underline{\eta}_k \,, \quad k = 0, 1, \ldots, N \tag{2}$$

where, $\underline{x}_k$ are (N × 1) state vectors, $\underline{y}_k$ are (R × 1) observation vectors, $\underline{u}_k$ and $\underline{\eta}_k$ are (N × 1) are zero mean uncorrelated, white Gaussian noise vectors with,

$$E[\underline{u}_k \underline{u}_s^T] = K\delta(k-s)$$

$$E[\underline{\eta}_k \underline{\eta}_s^T] = L\delta(k-s)$$

Now, suppose there exists a priori knowledge of the lower bound on the observations, viz.,

$$\underline{y}_k \geq \underline{\ell} \qquad k = 0, \ldots, N \tag{3}$$

where $\underline{\ell}$ is an (R × 1) vector. The problem then is to determine the optimal estimates of $\underline{x}_k$, given the noisy observations $\underline{y}_k$ and the constraints so that the conditional probability,

$$\eta = \max_{\underline{x}_k,\, k=0,N} \{p(\underline{x}_k, \underline{c}\,\underline{x}_k + \underline{\eta}_k > \underline{\ell} \,|\, \underline{y}_k,\ k = 0, \ldots, N)\} \tag{4}$$

is maximized. The solution of this maximization problem, obtained via the calculus of variations after using the Bayes rule and the probability densities of $\underline{\eta}_k$ and $\underline{u}_k$, gives the following nonlinear equations of the interpolator [1] (or the Maximum Likelihood Filter).

$$\underline{x}_{i+1} = \underline{A}\,\underline{x}_i + \underline{B}\,\underline{K}\,\underline{B}^T \underline{q}_{i+1} \tag{5}$$

$$\underline{q}_i = \underline{A}^T \underline{q}_{i+1} + (\underline{c}^T/L)(\underline{y}_i - \underline{c}\,\underline{x}_i) + \underline{c}^T \underline{h}(x_i) \tag{6}$$

$$\underline{q}_0 = \underline{P}^{-1} \underline{x}_0 \tag{7}$$

$$\underline{q}_N = 0 \tag{8}$$

where

$$h(x_i) = \frac{\frac{1}{\sqrt{2\pi L}} \exp\left\{-\frac{(\underline{\ell} - \underline{c}\,\underline{x}_i)^2}{2L}\right\}}{1 - z(x_i)} \,. \tag{9}$$

and

$$z(x_i) = \frac{1}{\sqrt{2\pi L}} \int_{-\infty}^{\underline{\ell} - \underline{c}\,\underline{x}_i} \exp\left[\frac{-t^2}{2L}\right] dt \,. \tag{10}$$

It can be easily checked that for $\underline{\ell} = -\infty$ (i.e., in the absence of any constraints), equation (6) reduces to a linear equation and equations (5) and (6) represents the standard optimal linear interpolator equations. These linear set of equations subject to conditions (7) and (8) are easily solved by using a Riccati Transformation [1]. Figures 1 and 2 show the original and noisy image. Figure 1 shows the optimal linear interpolation on the two-dimensional 32 x 32 noisy image with S/N ratio of 7/3. The S/N is improved by roughly

20.5db. Let u_i represent the linear interpolation estimate. The nonlinear interpolator is determined by first linearizing equation (6) about the linear estimate u_i. The linearized equations are then solved via a Riccati Transformation as before. Figure 2 shows this interpolation estimate. The major improvements in the picture are on the left and right sides of the white square (about 1.1 db). The improvement in the picture depends on how often the constraint in violated. Better approximation to the nonlinear filter may be obtained by taking more iterations. For a picture of size M x M, the storage requirement for interpolation are nM(M+1) locations, where n is the order of the vector $\underline{x}[2]$. However, an estimator or one step predictor corresponding to the nonlinear filter would require only n storage locations, thereby, making on line filtering feasible.

References

1. A. K. Jain, "Linear and Nonlinear Interpolation for 2-Dimensional Image Enhancement," Conference on Decision and Control, New Orleans, December, 1972.

2. N. E. Nahi and T. Assefi, "Bayesian Recursive Image Enhancement," IEEE Trans. Computers, Col. C-21, No. 7, July, 1972, pp. 734-738.

3. L. E. Franks, "A Model for Random Video Process," Bell Sys. Tech. Journal, Vol. 45, April, 1966, pp. 609-630.

4.7 Pseudocolor Image Enhancement Techniques

Harry C. Andrews, A. G. Tescher, Richard P. Kruger

The article commencing on the following page appeared in the July, 1972 issue of the IEEE Spectrum as a tutorial paper describing digital image processing. It is included in this report because of the relatively new results on the use of pseudocolor which appeared on the latter part of

the article. The article lists a variety of techniques on the use of color as an aid to viewing monochrome imagery, and illustrates a few of the methods by color reproductions. Applications of the techniques might range from biomedical to photoreconnaisance imagery, the success being related to the applicability of the given pseudocolor technique to the viewing objective desired. Unfortunately, this is more of a user oriented criterion, and one in which theory has yet been quite limited. Future pursuits of the technique will be directed along the lines of maximizing the entropy of the pseudocolor displays and utilizing just noticeable difference (jnd) uniform chromaticity display scales.

5. Image Detection and Measurement Projects

The image detection and measurement projects comprise a mixture of tasks concerned with techniques and tools for formatting images and extracting information from pictures. The problems considered under this classification include the registration of pairs of images, the measurement of image features, and the detection of objects within pictures. In many instances the solutions to the problems are very much unique to the problem, but there are also universal problems. The latter have been given primary consideration in the study.

The first report describes the preliminary effort in the application of recursive estimation techniques to the detection of objects within pictures. During the restoration of images by recursive filtering, one of the by products of the estimation is a measure of the likelihood that the image contains an object of interest possessing a certain statistical description. This information can be utilized with a minimal amount of effort to determine the presence of objects in the presence of substantial amounts of image noise.

Another research topic of the past six months has been a study of edge detection techniques. Several methods have been surveyed, and one of them - called a bug tracer - has been implemented by software. This edge detector has been evaluated on a sequence of poor visual quality radiographic images. The performance of the edge detector is quite promising.

5.1 Detection of Objects in Noisy Pictures

Nassar E. Nahi

Simple recursive procedures have been developed to detect the presence of objects described by statistical measures in pictures containing additive noise. Let the output of an image scanner be denoted by the scalar discrete time function $z(k) = y(k) + v(k)$ observed for $k=1, \ldots, N$ where $y(k)$ and $v(k)$ are independent scalar stochastic processes. Statistical properties of $v(k)$ and $y(k)$ are assumed known. The function $y(k)$

is called the signal process and v(k) the noise process. Given z(k), k=1, . . . , N, the problem is to decide on the presence or absence of the signal.

In formulating the detection problem let H_1 be the hypothesis that the signal is present, and H_o be the hypothesis that the signal is absent.

$$H_1 : z(k)=y(k) + v(k) \qquad K=1, \ldots, N$$

$$H_o : z(k) = v(k)$$

Define f_j, j=0, 1 as the joint probability density of z(1), . . . , z(N) when H_j is true and define the corresponding likelihood ratio as $t(N)=\ln \frac{f_1}{f_o}$ If t(N) exceeds some appropriate threshold, H_1 is accepted; otherwise H_0 is accepted. Formulation of t(N), for the case where f_1 and f_o are Gaussian, leads to $t(N) = \sum_{k=1}^{N} \frac{z^2(k)}{\sigma^2} - \frac{[z(k) - \hat{y}(k)]^2}{L^2 + \sigma^2}$

where $\hat{y}(k)$ is the Bayes estimate of y(k) obtained recursively [1, 2], σ^2 is the noise variance, and L^2 is the ~~estimated signal~~ error variance. The required threshold can be obtained from specification of error probability (false alarm and miss) and $p(t|H_o)$ and $p(t|H_1)$, where p is the probability distribution of t given H_j, j=0, 1. The errors corresponding to wrong decisions will then be found from:

$$\text{false alarm} = \int_t^{\infty} p(t|H_o)\, dt \text{ and}$$

$$\text{error of second kind} = \int_{-\infty}^{t} p(t|H_1)\, dt.$$

The above procedure has been applied to a number of pictures under very low signal to noise conditions (peak to peak signal level/variance of noise of about $\frac{6}{20}$). This is far below the level where a human observer can detect the presence of signals. The preliminary results have been very encouraging. Attempts are being made to use the detection process in conjunction with estimation to improve the image enhancement process beyond what is achievable by <u>linear</u> estimation.

References

1. N. E. Nahi and T. Assefi, "Bayesian Recursive Image Estimation", IEEE Trans. on Computers, Vol. C-21, No. 7, July, 1972, pp. 734-738.
2. N. E. Nahi, "Role of Recursive Estimation in Statistical Image Enhancement", Proc. of IEEE, Vol. 60, No. 7, July, 1972, pp. 872-877.

5.2 Edge Detection Techniques

Charles B. McGregor and Richard P. Kruger

A fundamental process in pattern recognition and image processing is the detection of edges in a scene. An investigation of edge detection techniques applied to continuous motion images has been undertaken. The problem under study consists of identifying and measuring the area of a brightened region on an image frame and tracking this area as it changes from frame to frame. The work performed to date has consisted of an examination of differentiator and tracing types of edge detectors.

Differentiating Edge Detectors. The computation of edge elements by differentiation on a discrete image takes the form of simple additions and substractions of pixel intensities. The simplest such differencing operation takes the form:

$$d_{i,j} = a_{i,j} - a_{i-1,j}$$

Other forms which have been used include the gradient [1,2]

$$d_{i,j} = (a_{i,j+1} - a_{i,j})^2 + (a_{i-1,j} - a_{i,j})^2$$

and the Roberts gradient

$$d_{i,j} = |a_{i,j} - a_{i+1,j+1}| + |a_{i,j+1} - a_{i+1,j}|$$

Variants on these basic approaches, including the Laplacian, have been reported in the literature.

A fundamental problem occurring with simple differencing techniques is that in real images large differences may occur on a pixel to pixel basis at points both interior and exterior to a desired edge. This causes difficulty if a threshold criterion is to be used to collect the set of edge points from the differentiated scene. One approach to this problem is the use of heavy smoothing in the computation of the differences, eg:

$$d^4_{i,j} = (a_{i-4,j} + \ldots + a_{i-1,j}) - (a_{i,j} + \ldots + a_{i+3,j})$$

where the superscript indicates the number of elements included in each of the averaging sums. A further refinement involves the use of a product of terms:

$$d_{i,j} = d^1 \cdot d^2 \cdot d^4 \cdot \ldots$$

While this product edge detector does succeed in removing many undesired detections from the results, it does so at the price of lost resolution and the generation of a set of sequencies at which the detector is totally blind to periodic edges. The blind sequency effect can be seen from Figure 1. Consider the edge detector $d^N_{i,j}$. As this detector is translated along the x-axis, it can be observed that any waveform which causes the integral under the positive and negative cusps of the detector to vary identically will not contribute to the detector output. There exists a broad class of periodic waveforms having period equal to N/k (k an integer) which possess this property. Thus, for example, the detector d^8 is blind to waveforms with sequencies of 8, 16, 32,... While such an effect is acceptable in situations in which the data is known to have no spectral components at the blind sequencies, it is totally unacceptable for cases in which the spectrum of the data is unknown. While further proceedures might be developed, possibly using different values of N with suitable combining logic, the growing complexity of the detector becomes unattractive for an operation which must be repeated on each pixel of a scene.

In the presence of noise, the performance of the differentiating detector is severely degraded. Noise will increase the number of randomly distributed points which cross the derivative threshold, causing a rough, thickened edge region rather than the desired thin line. As the differentiating detector does not label interior points in the desired set, edge tracer techniques as described in the next section are not implemented as effectively as with threshold detection processing.

Threshold-Tracing Edge Detectors. The determination of the edge of a set defined by threshold criteria begins with the comparison of each

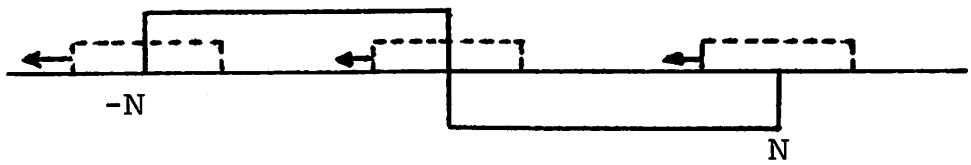


Figure 5.2-1. Blind sequency effect in edge detector

6	7	8
5	X	1
4	3	2

Figure 5.2-2. Unit radius "circle" used by bug

pixel with a threshold. Pixels below the threshold values are set to zero and those exceeding the threshold are set to a uniform value. Thus the data is initially compressed to a one-bit representation on the intensity axis before further processing. The threshold for this comparison may be either externally specified or computed by seeking an extremum of the histogram of the scene. The output from this process will contain isolated and discontinuous sets of points, similar to the output of a differentiating edge detector. At this point, continuity of the edge is invoked and an algorithm which seeks connected points on the frontier of the set is used. In topological terms, this algorithm seeks points which (1) are adjacent to previously determined perimeter points and (2) are the centers of unit radius neighborhoods containing points of both the circumscribed set and its complement.

The operation of this algorithm is best described by ascribing its attributes to an intelligent bug which walks about the image plane performing the tracing [3]. Initially, the bug is either placed on a known boundary point or it is caused to enter the scene along a suitable horizontal or vertical line and proceeds until it identifies an edge point. One typical procedure defines the initial edge point as the first interior point of the set which is followed by a continuous line of N points also belonging to the interior. When thus initialized, the bug looks toward the interior of the set and begins searching a unit radius path in a clockwise direction until it locates a point belonging to the complement of the set. Having found such a point, the bug continues its scan from that point until it encounters the first member of the desired set. The bug then moves to the new point and enters it into a list of perimeter points. Next, it looks back to the previous point and begins the search for an exterior point, repeating the above procedure. The bug repeats this procedure, moving in a clockwise direction around the desired set until it returns to the initial point. Having traced a closed path, the bug retires. It is noted that the bug lives in a square world and considers the eight elements shown in Fig. 2 as belonging to a neighborhood of unit "radius". Since the bug moves in fixed length steps, a complete record of its path consists of

the coordinates of the initial point and the series of directions it moves on each succeeding step. This highly compact representation of the perimeter is known as a chain code [4]. This encoded data contains all information necessary to compute the enclosed area and the moments of the traced centour [5].

In the presence of noise, the bug algorithm performs admirably since it only senses and reacts to noise elements which lie within a unit radius of a perimeter point. Further the probability that the bug will deviate from the desired perimeter by N units equals the probability that a chain of N adjacent noise elements will occur with one end adjacent to the perimeter. This probability varies with the Nth power of the probability of a single noise element. The only restriction on the tracing of closely adjacent set perimenters is the probability the noise will bridge between the perimenters, causing two sets to appear as one.

An Experimental Edge Detection Study. The performance of the threshold-tracing edge detector has been evaluated using radiographic films of the human heart. The cardiac catheterization procedure used in the evaluation of severe cases of heart disease involves injection of radio-opaque dye into the left ventricle (the output pumping chamber) of the heart while motion pictures are made of the flouroscopic image of the heart. From these films, the opaque image of the ventricle can be traced to show its contraction as a function of time while the heart beats. From the cross sectional area shown, the volume may be estimated as a function of time and the volume of blood expelled by the heart per beat can be computed for diagnostic use. At a motion picture rate of thirty frames per second, a manual processing of these fifty images requires ten to twelve hours of a skilled technician's time. Thus a need for computer data processing arises.

The experimental data consists of strips of 35 millimeter film which have been scanned in a flying spot scanner. A view of the heart with no dye injected is available to allow subtraction of background opacity from the frame to be processed. A frame of film showing a ball of known diameter in the position later occupied by the heart is available for calibration purposes.

Initial processing consists of subtraction of the background picture from the first image to be processed and computation of a gray level histogram for the resulting picture. A bimodal histogram is assumed and a curve is fitted to the valley of the histogram. The central minimum of this curve is used as the threshold for edge detection. After extensive investigation, a version of the bug which does not require prior thresholding of the data has evolved. This bug accepts threshold levels from the calling program and makes the necessary threshold comparisons as it traces its path through a field of unthresholded data. This procedure reduces the threshold comparisons from N^2 to approximately 16N operations and allows the data to be viewed with the threshold tracing superimposed. A typical tracing is accomplished in 0.8 seconds. Examples of raw data and the data with an edge tracing superimposed are shown in figures 3 and 4.

Points of Departure. The next task in applying the contour tracing to frame to frame tracking involves reducing the large amount of data present in a chain-encoded contour to a few descriptive paramenters. An initial approach will use a Fourier expansion on the closed boundary. The minimum number of Fourier coefficients needed to approximate the boundary within an area error limit will be established. A parallel investigation is underway which may result in the integration and chain code processing portions of this project being used with manual Graf-pen input of data to a time-sharing computer system for interim man-machine data processing.

References

1. A. Rosenfeld, R. Thomas, Y. Lee; "Edge and Curve Enhancement in Digital Pictures", Tech. Report., 69-93, AEC Contract AT(40-1)-3662.
2. A. Rosenfeld, M. Thurston, "Edge and Curve Detection for Visual Scene Analysis", IEEE Trans. Computers, Vol. C-20, May, 1971.
3. Denis Rutovitz, "Data Structures for Operations of Digital Images", Medical Research Council, London.
4. H. Freeman, "On the Encoding of Arbitrary Geometric Configurations", IEEE Trans. Computers, Vol. EC-10, No. 2, June, 1961.
5. H. Freeman, "Techniques for the Digital Computer Analysis of Chain-Encoded Arbitrary Plane Curves", Proc. NEC, October, 1961.

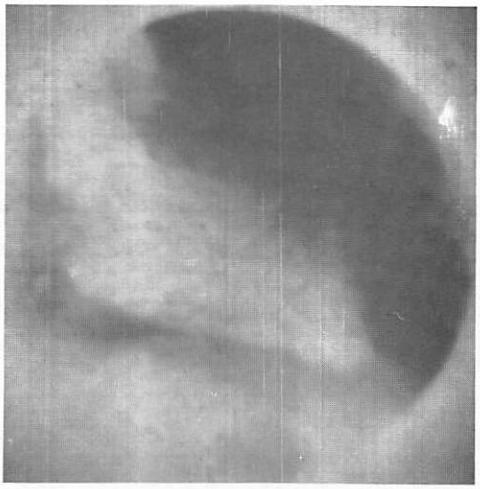


Figure 5.2-3. Diastolic (dilated) Ventricle, raw data

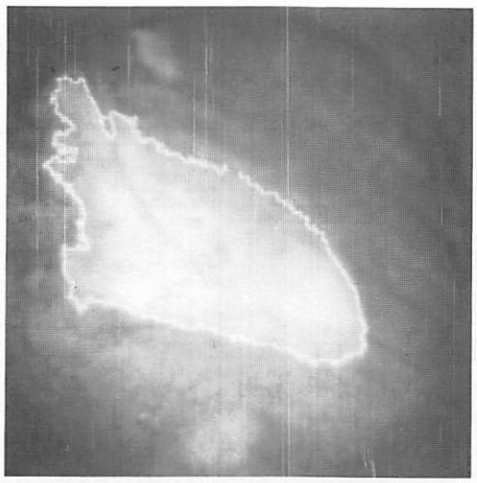


Figure 5.2-4. Data of Figure 3 with edge traced

6. Image Processing Support Projects

The following sections describe the status of the U. S. C. image processing research facilities with particular emphasis on the progress toward making the image processing hardware and software facilities available over the ARPANET. A report is also given on the initial steps towards the establishment of a standard set of original images for the image processing research community.

6.1 U. S. C. Image Processing Hardware and Software Facilities

Richard P. Kruger

Image Processing Hardware Facilities. Figure 1 is a general block diagram depicting the hardware facilities utilized by the U. S. C. Image Processing Laboratory (IPL) for various image acquisition, computation, and display operations. The IPL facility consists of a Hewlett-Packard 2100 which is used to command and control the image acquisition and display devices, and an Interdata IV which is used for simulation of image processing systems.

At present a conventional flying spot scanner is used for image digitization and hard copy display of color and monochrome images on photographic prints. The device can raster scan transparencies up to 70 millimeters square with resolutions up to 1024 by 1024 pixels. The scan time varies, but is usually less than 1 minute per frame. Normal analog to digital conversion is eight bits. Display is usually on Polaroid or 35 mm film. Color digitization and display is accomplished serially under computer control using a color filter wheel.

The Muirhead color facsimile device is a two unit drum transmitter and receiver. Its capabilities include 100 line per inch resolution with simultaneous ten bit per color quantization as well as display of the three primary colors on an 8 by 10 inch print. The scan time for this device is 12 minutes. While the maximum scan or display size is

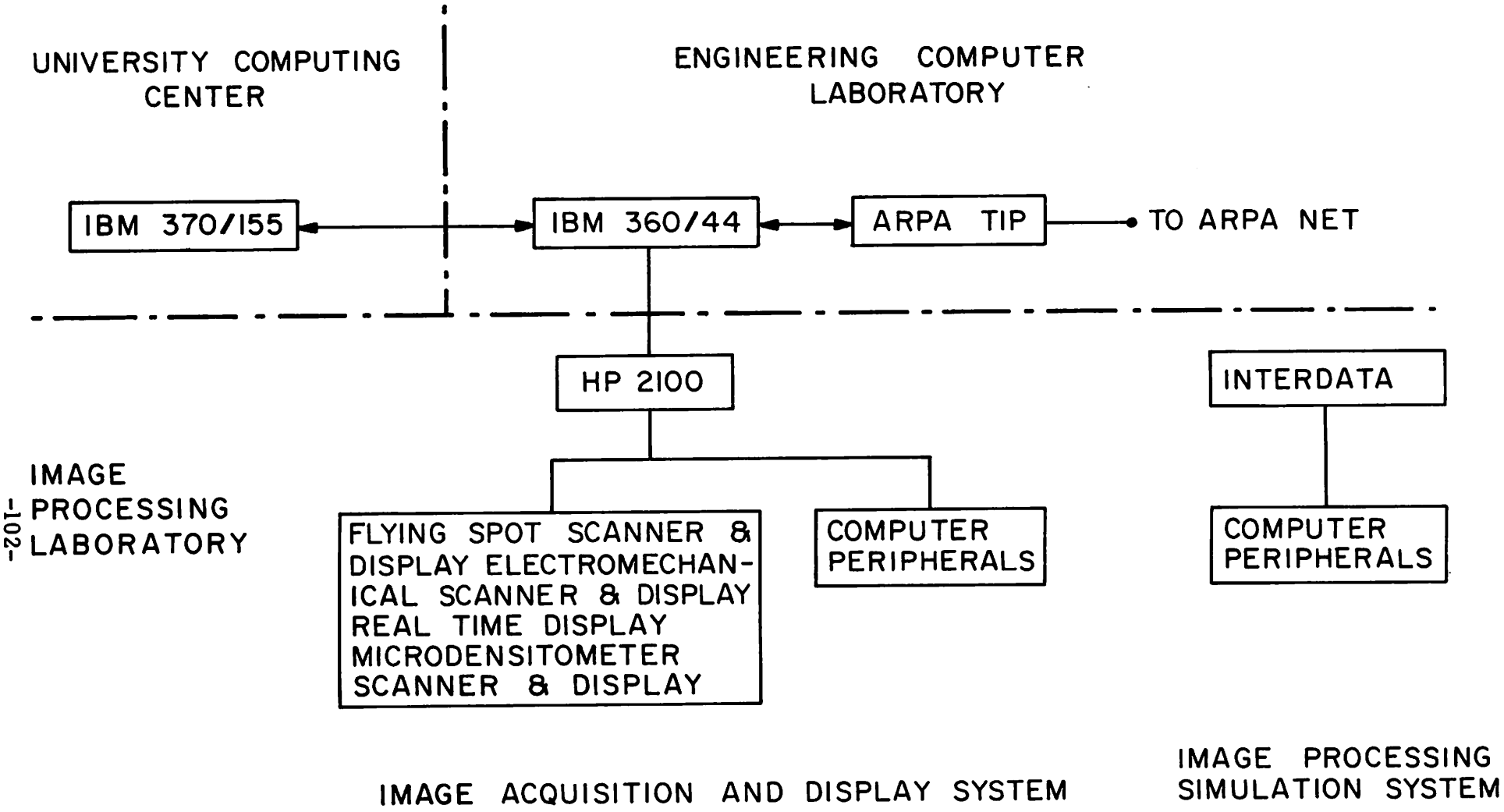


Figure6.1-1. USC Image Processing Laboratory Facilities

fixed on an 8 by 10 inch field, options exist for scan or display of any 2.5 or 5.0 inch square portion of a larger field at resolutions of 256 by 256 or 512 by 512 pixels.

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

An Optronics flat bed scanning microdensitometer is scheduled for delivery by 1 January 1973. A specification of the unit is enclosed as an appendix to this report. The Optronics unit will scan and display 16, 35, and 70 millimeter color or monochrome roll film on a registered transport under computer control as well as provide prints or transparencies up to 10 inches square. Specifications call for a minimum aperture size of 2 microns square, 12 bits or precision, a scanning velocity of 8 inches per second in the direction of travel and a 0 to 4 specular density range. Color digitization and display is sequential using a color wheel under program control.

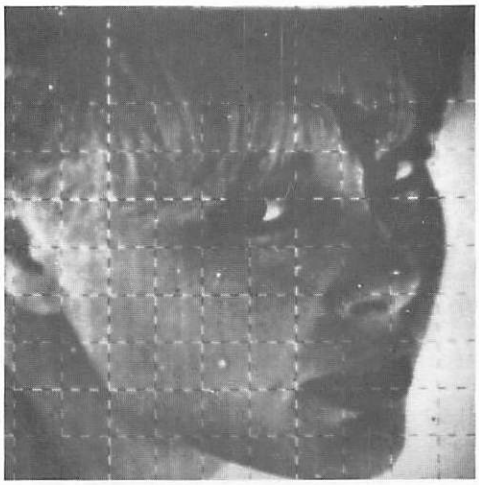
The IPL computers and peripherals are hardware interfaced to the Engineering Computer Laboratory (ECL) IBM 360/44. Images may be transferred from the disk on the HP 2100 to a disk on this machine, processed and then transferred back to the HP 2100. Work is continuing on an operating system that will provide remote job entry between the HP 2100 and the IBM 360/44 for image processing tasks. The IBM 360/44 is optically linked to the University Computer Center (UCC) IBM 370/155 for remote job entry between the machines. The IBM 360/44 also acts as the U.S.C. host computer for the ARPANET using a TIP interface.

Image Processing Software Capability. Image processing tasks on the IBM 370/155 are performed with conventional FORTRAN based software and a processing software package called VICAR. The VICAR sys-

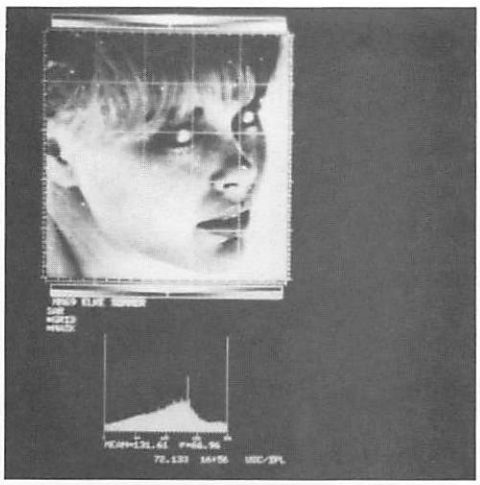
tem is also being implemented on the IBM 360/44. The VICAR (Video Image Communication and Retrievel) system was acquired from the Jet Propulsion Laboratory, and a modified form has been operational at U. S. C. for the past year.

The purpose of VICAR is to facilitate the acquisition, digital processing, and recording of image data on a production basis by scientific personnel not familiar with systems programming, applications programming, and problem programming. The system programmer requires knowledge of applications and problem programming only in a global sense. The applications programmer is capable only of using existing problem programs and requires no knowledge of either system or problem programming. The results of processing an image with problem programs, GRID, MAPGRID and CONCAT are shown in Figure 2a, b and c. The problem programmer must have knowledge of application program protocol but need not have system programming knowledge. From a system viewpoint, the operation of VICAR is very similar to that of a standard utility program. The application programmer submits a card deck containing a limited number of VICAR control statements to define the processing tasks. These statements include READ and WRITE tape instructions in 7 or 9 track, 6 or 8 bit formats, disk data set RESERVE cards and program EXECUTE cards. All tape and disk labels and job control statements are generated so as to be transparent to the user. Up to 14 simultaneous I/O requests can be made. User labels as well as history and system labels, are routinely placed at prescribed locations preceding the image data on a tape or disk file. This information can be retrieved and used for picture annotation purposes such as shown in Figure 2d. This is one display option of the problem program MASK. Labels can be deleted in total or in part at the user's option. The problem programmer unlike the applications programmer can add programs to the program library which can be executed by himself or by an applications programmer. A problem program can be written either in Fortran

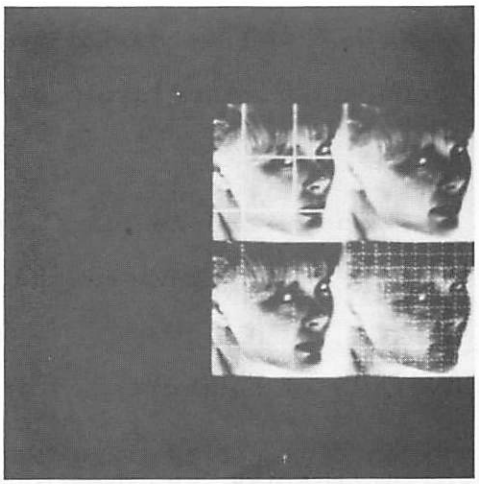
(a) Output of GRID



(b) Output of MAPGRID



(c) Output of CONCAT



(d) Output of MASK

(e) Original Image

(f) Image processed by Gaussian isotropic high frequency emphasis filter followed by histogram equalization.

Figure 6.1-2. Examples of Processed Images

or Assembly Language and differs from a normal subroutine only in its I/O structure. In addition to normal Fortran I/O which is optional, several VICAR I/O statements must be included. Most existing conventional subroutines can be easily converted to VICAR problem programs through a modification of its I/O structure. At present VICAR contains over 30 application programs which perform such varied tasks as Fourier transformations, picture annotation, line plotting and filter design. An updated version of VICAR with block I/O transfer has been installed at UCC. This modification will decrease I/O time and significantly reduce I/O costs involved in image processing on that machine.

APPENDIX

Specifications for Proposed Microdensitometer Film Scan and Playback System for University of Southern California Image Processing Laboratory

A. Digitizer/Display Specifications

1. The ability to scan and playback sprocketed roll film transparencies of 16, 35, and 70 millimeter sizes using an index registered automatic film transport which may be advanced or reversed manually and under program control an arbitrary number of frames. The transport device must provide frame to frame registration error of not greater than $\pm$ 12 microns perpendicular and parallel to direction of film travel.
2. The ability to directly digitize and playback sheet film photographic transparencies of opacities (prints) up to a maximum size of 25.6 cm square.
3. Color separation by means of a 3 primary color wheel under manual and program control to digitize or playback color images in a sequential mode.
4. The capacity to digitize and playback images in a range from 0 to

4 specular sensity units (0 - 100% transmittance) with at least 12 bit A/D and 10 bit D/A conversion with at least 8 bits of significance minimum in any application. (Of course film base fog and maximum density sensitivity will affect particular applications.)

5. Linearity of ± .02 specular density units (± .5 percent transmittance) over the above stated range and half the above variation over any 2 density unit window within the 0 - 3 density unit range.
6. Stability of less than ± .02 density unit drift (± .5 percent transmittance) after one hour of warm up to be sustained for at least 10 hours.
7. Repeatability of ± .01 density units or less after one hour of warm up.
8. Optical system depth of focus to exceed 250 microns at all aperture sizes or a means whereby the transported roll film or mounted sheet film curvature will not exceed the depth of focus at the film plane for any selected aperture.
9. Square apertures capable of 256, 512, and 1024 line resolutions for a square frame scan or replay on all types roll film as well as for sheet film sized of 4, 5, 8, and 10 inches square with no aperture overlap. (It is expected that standard apertures available with the device will also be included and that other apertures may be easily added in the future.)
10. X and Y axis spatial repeatability of less than or equal to ± 1 micron.
11. X and Y axis straightness of travel variation less than or equal to ± 5 microns.
12. X and Y axis orthogonality of ± .5 seconds of arc.
13. Position accuracy of 1 micron RMS in X and Y direction in 25.6 cm of travel. (6 micron peak to peak)

14. Digitization or replay speed variable from approximately 5 to 20 cm/second minimal.
15. Peak signal-to-noise ratio exceeding 50 db in large aperture clear field operation and to exceed 40 db under all operating conditions.
16. Photomultiplier surface material must be capable of extended red wavelength response to provide minimum signal-to-noise ratio requirement for color image scanning. It is suggested that S-20, S-25, or Gallium Arsenide photosurfaces with a thermal electric PMT cooler be investigated to meet this specification.

B. Other Specifications

1. A viewing screen for accurate manual film positioning as well as a means for viewing small portions of the film plane.
2. Switchable logarithmic converter.
3. Direct digital display of film density information.
4. Console display with convenient switches and indicators such as end of scan light, etc.
5. Means for accurate initial and interval calibration of instrument to ensure continuing densitometer accuracy.
6. The proposed system will be interfaced to a Hewlett-Packard 2100 with 16 bits in and out, flag in and out configuration. A typical I/O card for this purpose is the microcircuit interface card No. 12566. Cooperation with our computer staff and Hewlett-Packard will be available to aid in software development and consultation.
7. Switching circuitry to allow bit packing into a density or transmittance range existing within a particular transparency or print so that the potential exists for full register length usage.
8. A complete maintenance manual so that our laboratory staff can perform all repairs and periodic adjustments.
9. Scan and replay capability in a normal room light environment.

6.2 USC/ARPANET Image Processing System

Richard P. Kruger and William K. Pratt

The University of Southern California image processing hardware and software facilities previously discussed will soon be available for use over the ARPA computer network. A remote site user may:

(a) read an image from an image file and transfer the image to the remote site;

(b) store an image, transferred from the remote site, on an image file;

(c) display and make a hard copy of an image stored on an image file;

(d) digitize a hard copy image and store on an image file;

(e) execute a VICAR language image processing program.

The use of the USC/ARPANET image processing system by a remote site user requires, as a minimal configuration, a computer terminal connected to a TIP. With this configuration image processing jobs can be initiated from the remote site with operations performed on previously stored images and hard copy image results returned to the user by mail. The next level of capability entails the addition of an image recording device at the remote site for display of processed images. A complete remote site system would also include an image scanner and digitizer for supplying input images.

Figure 1 contains a block diagram describing the information flow in the USC/ARPANET image processing system. In operation a user at a remote site establishes a connection with the USC network control program and notifies the NCP that an image processing job is to be executed. A data transfer program receives a request for an image processing job from the NCP and then acts as a monitor for control of the job instructions and data. The remote user's job parameters, consisting of instructions defining the image processing operations to be performed and image data loca-

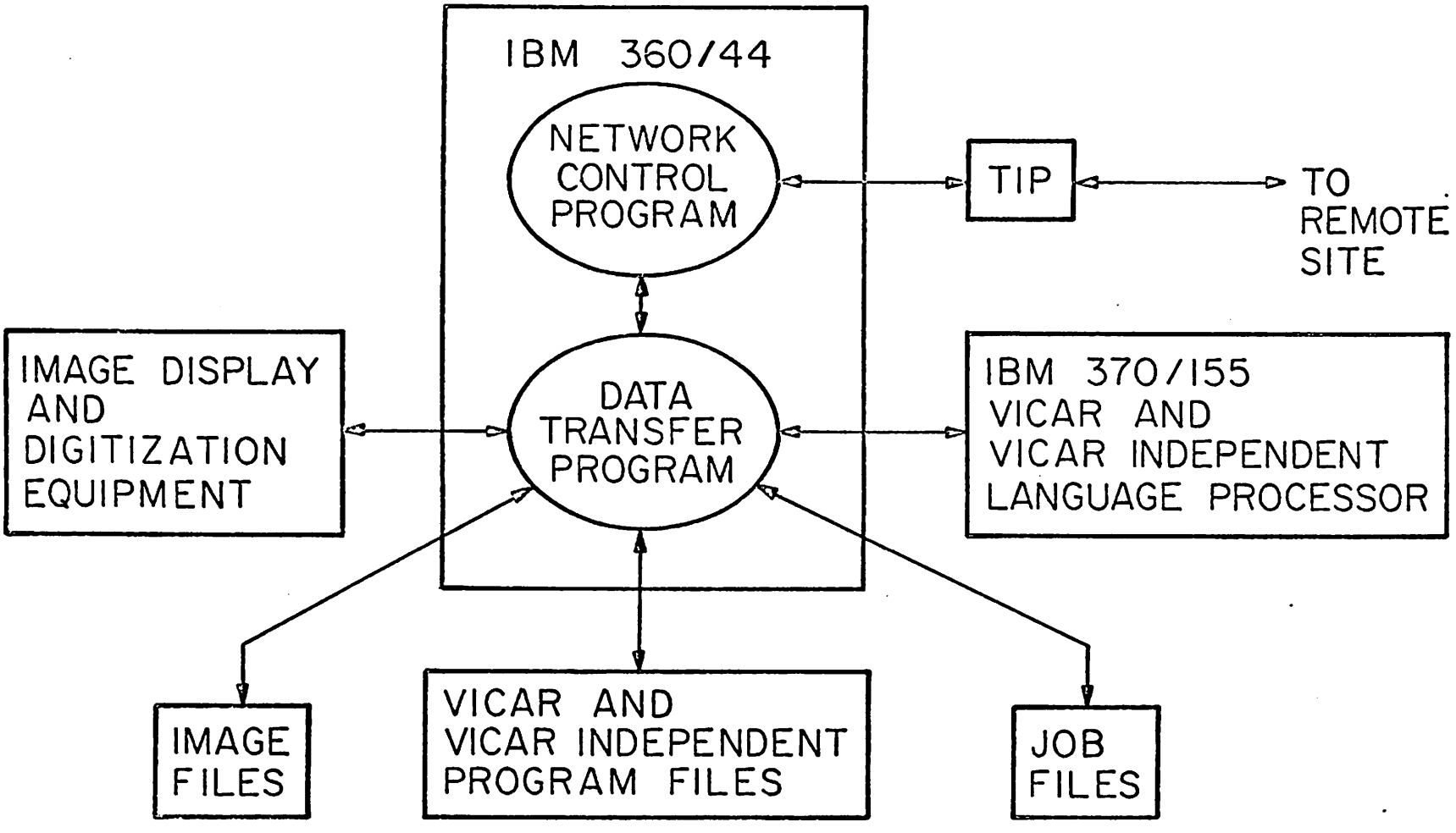


Figure 6.2-1 USC/ARPANET IMAGE PROCESSING SYSTEM

tions, are stored in a job file preparatory to execution. If the user submits a VICAR program, the source program is stored on a mass image file. When the remote user's job is ready for execution the data transfer program transfers the appropriate image data and instruction to the processors. For example, if a VICAR job is to be performed the data transfer program transfers the remote user's VICAR language source program and the input images to be processed from the image files to the VICAR language processor on the IBM 370/155. After the VICAR processing is completed, the output image is returned to the image file, and subsequently returned to the remote site through the TIP.

The majority of messages transmitted over the network are of short length and can be placed in a single packet. However, bulk files such as digital images usually exceed the packet length, and therefore, must be partitioned for transmission. The partitioning is automatically handled by the host and user network control programs, and is not of concern to the user. For digital image processing operations, images are usually coded with eight bits (one byte) per picture element. Furthermore, images are often restricted in size to a binary integer, e.g. (256 by 256, 512 by 512, etc.) in order to take advantage of fast computational algorithms which impose this restriction. In the interest of standardization, the USC/ARPANET image processing system has been designed to handle image files with a record length of 256N bytes per record where N is an integer. The number of records in a file can be specified by the user, and is not restricted to a binary integer. At present the fastest transmission speed over the ARPANET is 50kbs. Considering the transmission control data that automatically is added to each message packet and transmission delays encountered at each node, the average user data rate is likely to be about 25kbs. At this rate the time required to transmit a 256 by 256 byte image is about 21 sec. A 1024 by 1024 image requires about 5.6 min. for transmission. Plans are underway to provide

a higher speed transmission service to the network. Also, image bandwidth reduction techniques are being considered to decrease the transmission time.

6.3 Development of Original Images for Image Processing

William K. Pratt

A master set of three monochrome and two natural color images has been assembled as a provisional standard for image coding studies. The following sections describe the original master set. It is anticipated that future master sets will contain a wider variety of imagery, useful for image coding, restoration, enhancement, etc. studies, and that the subject pictures will be better calibrated, analyzed, and documented.

General Picture Description. All test pictures contain 256 by 256 pixel points. The luminance (intensity) at each pixel point has been linearly quantized into 256 levels and binary coded into the range (0, 255).

The test pictures are listed below:

File #	Name
1	girl - red tristimulus value
2	girl - green tristimulus value
3	girl - blue tristimulus value
4	couple - red tristimulus value
5	couple - green tristimulus value
6	couple - blue tristimulus value
7	girl - monochrome luminance
8	couple - monochrome luminance
9	moonscape - monochrome luminance

The monochrome luminance picture of the girl and couple have been obtained from the natural color representation through the coordinate conversion

$$Y_T = .299R_T + .587G_T + .114B_T$$

where Y_T = image luminance in integer form ($0 \leq Y_T \leq 255$)

R_T, G_T, B_T = red, green, blue image tristimulus value in integer form ($0 \leq R_T \leq 255$), ($0 \leq G_T \leq 255$), ($0 \leq B_T \leq 255$).

The moonscape image is a picture of the moon taken from a spacecraft and subsequently processed by the Jet Propulsion Laboratory to enhance its subjective qulaity. No information is readily available as to characteristics of the image sensor or the post processing algorithms.

The color images have been scanned by Kodak from color slides distributed by the Society of Motion Picture and Television Engineers (SMPTE). The digital tapes and supporting documentation were supplied to USC by Dr. Oleh Tretiak of M.I.T. In the scanning process, the scanner output (nominally photographic density) was linearly quantized to 256 levels. System calibration curves were then generated to relate a scanner digital number to true density and true transmittance. At USC the calibration curves were used to convert the scanner digital numbers to a set of digital numbers linearly proportional to the transmittance of the test slides for each tristimulus value. The resulting transmission scale between 0. and 0.6 has been linearly quantized to 256 levels. The SPMTE slides were scanned over an equi-space raster of 520 lines and 780 pixels per line. The couple picture of the master set was obtained from the scan of SMPTE set #2 by extracting the center 512 by 512 pixels and spatially averaging the corrected transmittance in groups of four pixels to obtain a 256 by 256 pixel image. The girl picture of the master set was obtained by extracting a 256 by 256 pixel area of the same digitized SMPTE slide #3.

Statistical measurements have been performed on the master set images. These measurements include:

mean, variance, standard deviation
row and column covariance
grey scale histogram (64 levels)
first order entropy
second order row and column entropy

For storage of images on magnetic tape, the tape format is:

2400 ft. reel
9 track
800 bpi
one pixel per byte
256 bytes per record
256 records per file
no tape label
record gap after each record (IBM compatable)
file mark after each file (IBM compatable).

7. New Research Projects

The following are summary descriptions of new research studies that will be initiated during the next six months in addition to the continuation of topics presented earlier.

7.1 Interframe Coding of Television Signals

Ali Habibi

The contour tracing algorithm has been employed as a technique of encoding two dimensional digital data by tracing the contours of equal gray levels and transmitting the addressing information that enables the receiver to reconstruct these contours. This technique relies upon the spatial correlation to reduce the data redundancy. It is proposed to use the contour tracing algorithm to encode interframe differential pictures. These pictures are created by taking the difference of two successive frames in the television signal and indicating the change along the temporal axis. The proposed technique reduces the redundancy in both sptaial and temporal domains, therefore achieving a further reduction in the redundancy than is possible with either one of the techniques. Its potential use includes coding both monochrome and color video signals.

Preliminary results indicate that good coding capabilities may be achieved by combining unitary transformations with a DPCM encoder. Using this technique for intraframe coding, first a one dimensional unitary transformation is used to obtain the transform of each line of a still picture. This gives a signal in the transform domain which has a strong correlation in the vertical direction, then this signal is coded by a DPCM coder taking advantage of this correlation.

It is planned to extend this technique to interframe coding by taking the two-dimensional unitary transform of each frame and then applying the DPCM encoder to code the transformed data in the temporal direction. Such a system will explore the correlation of moving pictures in spatial as well as temporal directions and allow a parallel operation that can be used to implement real-time interframe coders.

7.2 Space Variant Point Spread Functions

Harry C. Andrews

Due to the increase in computing technology, it is becoming feasible to consider the task of computationally implementing image restoration systems degraded by space variant point spread functions (SVPSF). Five specific areas of investigation under this subject are:

1. Object dependent SVPSF's
2. Eigenfunction Analysis
3. Degree of Freedom and SVPSF's
4. Gaussian Kernel SVPSF's
5. Atmospheric degradating SVPSF's

Each one of these areas is being investigated with respect to analysis and computational implementation. The object dependent SVPSF is modeled after the equation

$$g(x, y) = \int_{-\infty}^{\infty}\int f(\xi, \eta)\, h(x, y, \xi, \eta, f)\, d\xi d\eta$$

An approach to the solution of this equation is a Taylor Series expansion of h(.) about f(.). The Eigenfunction analysis utilizes the equation

$$\lambda_{nm}\psi_{nm}(x, y) = \int_{-\infty}^{\infty}\int \psi_{nm}(\xi, \eta)\, h(\xi, \eta, x, y)\, d\xi d\eta$$

to develop a system of Eigenfunctions (if they exist) for expansion of the image function.

The degrees of freedom analysis results in a four-dimensional singular valued decomposition to describe $h(x, y, \xi, \eta)$ more efficiently for computational inversion. The Gaussian kernel SVPSF analysis attempts to invert a point spread function given by

$$h(\underline{p}, \underline{q}) = A(\underline{p}) \exp - \{(\underline{p} - \underline{q})^t K^{-1}(\underline{p})(\underline{p} - \underline{q})\}$$

where

$$\underline{p} = (x, y)$$

$$\underline{q} = (\xi, \eta)$$

Various simplications are introduced for computational simplicity. The atmospheric degradation SVPSF is really a phase sensitive SVPSF

system in which the image and point spread function phases are particularly significant. Phase unwrapping and phase filtering will probably be required for such restoration processes.

7.3 Fast Recursive Restoration for Two Dimensional Images

Anil K. Jain

The standard methods used for recursive image restoration have the following difficulties:

a) The filters are not isotropic, i.e., the improvement in the S/N ratio depends considerably on whether the image is scanned horizontally or vertically;

b) The computation time for large images is considerable;

c) Vector scanning gives a better S/N ratio but increases dimensionality which leads to large storage requirements.

A second order vector Markov model for two-dimensional images has been developed and it is believed (based on some initial results) that an isotropic filter with scalar filter equations (for vector scanning) can be developed leading to a fast recursive restoration algorithm [1].

Reference

1. E. Angel and A.K. Jain, "A Nearest Neighbors Approach to Multidimensional Filtering," Conference on Decision and Control, New Orleans, December, 1972; also to appear in IEEE Transactions Aut. Control, April, 1973.

7.4 Pseudocolor Mapings Via Color Distance

Anil K. Jain

This project under consideration will deal with the determination of hue and saturation distribution as functions of illuminance (or intensity) of a black and white image so that a suitably chosen criterion in terms of color distance [1] is satisfied.

Reference

1. A.K. Jain, "Color Distance and Geodesics in Color-3 Space", to appear in the Journal of the Optical Society of America.

7.5 Image Registration

William K. Pratt and Alexander A. Sawchuk

A common image processing problem is that of spatially registering a pair of images from the same scene. For example, one image may be from a visible wavelength image sensor and the other from an infrared sensor. The general problem is to spatially register the images and compensate for translation, rotation, and magnification differences in the presence of sensor noise and sensor nonlinearities.

A preliminary approach taken to the problem has been the development of an N-dimensional correlation function expressing the N degrees of freedom of misregistration. This technique has been implemented by software for translation errors with good results. The unique feature of the program is its utilization of the spatial correlation of each image plane.

Further studies will expand the correlation technique for magnification and rotation. Also, consideration will be given to sequential search techniques.

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.

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

H. C. Andrews, "N Topics in Search of an Editorial: Heuristics, Superresolution, and Bibliography," Proceedings IEEE, Vol. 60, No. 7, July, 1972.

L. D. Davisson, "Rate-Distortion Theory and Application," Proceedings IEEE, Vol. 60, No. 7, July, 1972.

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

A. Habibi, "Coding Color Images by Differential Pulse Code Modulation," to be published in Proceedings International Telemetering Conference, October, 1972.

A. Habibi, "Quantization Errors and Entropy Considerations," Proc. Conf. on Computer Image Processing and Recognition, University of Missouri at Columbia, August, 1971.

A. K. Jain, "Color Distance and Geodesics in Color 3 Space," to be published in the J. Opt. Soc. Am.

A. K. Jain, "Linear and Nonlinear Interpolation for Two-Dimensional Image Enhancement," to be presented at Conference on Decision and Control, New Orleans, December, 1972.

A. K. Jain and W. K. Pratt, "Color Image Quantization," to be published in Proceedings of the National Telecommunication Conference, Houston, Texas, December, 1972.

N. D. Mascarenhas, "Suboptimal Generalized Wiener Filters," Proc. Conf. on Computer Image Processing and Recognition, University of Missouri at Columbia, August, 1971.

N. E. Nahi and T. Assefi, "Bayesian Recursive Image Enhancement," IEEE Transactions on Computers, July, 1972.

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

J. Pearl, H. C. Andrews and W. K. Pratt, "Performance Measures for Transform Data Coding," IEEE Transactions on Communication Technology, June, 1972.

W. K. Pratt, "Walsh Functions in Image Processing and Two-Dimensional Filtering," Symposium on Applications of Walsh Functions, March, 1972.

W. K. Pratt, L. R. Welch and W. Chen, "Slant Transforms for Image Coding," Symposium on Applications of Walsh Functions, March, 1972.

W. K. Pratt, "Generalized Wiener Filtering Computation Techniques," IEEE Transactions on Computers, July, 1972.

W. K. Pratt, "Binary Symmetric Channel Error Effects on PCM Color Image Transmission," IEEE Transactions on Information Theory, September, 1972.

W. K. Pratt and R. P. Kruger, "Image Processing Over the ARPA Computer Network," to be published in Proceedings International Telemetering Conference, October, 1972.

C. Reader, "Spatial Subsampling for the Transform Coding of Images," Symposium on Applications of Walsh Functions, Washington, D. C., March, 1972.

A. A. Sawchuk, "Linear Space-Variant Motion Degradation and Restoration," Spring Meeting of the Optical Society of America, New York, April, 1972.

A. A. Sawchuk, "Space-Variant Image Motion Degradation and Restoration," Proceedings IEEE, Vol. 60, No. 7, July, 1972.

A. A. Sawchuk, "Coordinate Transformations in Space-Variant Image Enhancement and Restoration," Fall Meeting of the Optical Society of America, San Francisco, October, 1972.

Security Classification

DOCUMENT CONTROL DATA - R & D

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

1. ORIGINATING ACTIVITY *(Corporate author)* Image Processing Laboratory, Electronics Sciences Laboratory, University Of Southern California University Park, Los Angeles, California 90007	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

Semiannual Technical Report to the Director, Advanced Research Projects Agency, for the period 1 March 1972 to 31 August 1972.

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

Technical Semiannual 1 March 1972 to 31 August 1972

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

William K. Pratt (Project Director)

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
FO8606-72-C-0008 b. PROJECT NO. c ARPA Order No. 1706 d.	USCEE Report 425 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)* None

10. DISTRIBUTION STATEMENT

Approved for release; distribution unlimited

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

13. ABSTRACT

This technical report summarizes the image processing research activities performed by the University of Southern California during the period 1 March 1972 to 31 August 1972 under Contract No. FO8606-72-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. Four 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 Detection and Measurement Projects, the recognition of objects within pictures and quantitative measurement of image features; (4) Image Processing Support Projects, a description of the USC image processing facilities accessible over the ARPANET.

****

14. Image Processing, Digital Image Processing, Image Coding, Image Enhancement, Image Restoration, Image Processing Support Projects, Color Image Transforms, Pseudocolor Processing

DD FORM 1 NOV 65 1473

Security Classification

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