OPTIMUM RESTORATION OF QUANTIZED CORRELATED SIGNALS

by

Michael N. Huhns

August 1975

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

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

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

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

1. ORIGINATING ACTIVITY (Corporate author)
Image Processing Institute
University of Southern California, University Park
Los Angeles, California 90007

2a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

2b. GROUP

3. REPORT TITLE
OPTIMUM RESTORATION OF QUANTIZED CORRELATED SIGNALS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Technical Report, August 1975

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

6. REPORT DATE
August 1975

7a. TOTAL NO. OF PAGES
211

7b. NO. OF REFS
74

8a. CONTRACT OR GRANT NO.
F08060-72-C-0008

b. PROJECT NO.
ARPA Order No. 1706

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)
USCIPI Report 600

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT
Approved for release: distribution unlimited

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY
Advanced Research Project Agency
1400 Wilson Boulevard
Arlington, Virginia 22209

13. ABSTRACT

An analysis of the optimum statistical restoration of quantized signals is presented. The restoration is based upon minimizing the mean square error between the input to a quantizer and its estimate. Since a quantizer is a nonlinear device, the estimation equation which is derived achieves an optimum nonlinear restoration. Its solution requires complete statistical knowledge of the quantizer input. Available statistical information usually includes the marginal distribution of each of the input variables and the correlation between them. Hence a technique is developed for generating correlated multidimensional probability density functions based on this available information. The technique is applied to gaussian, laplacian, and Rayleigh density functions. These multidimensional density functions characterize the outputs of transform coders, DPCM coders, and PCM coders, respectively. The quantized outputs of these coders are then restored by utilizing the multidimensional densities in the estimation equation. Examples of images which have been coded and restored by these techniques are presented. The results reveal a mean square error reduction. To achieve a visually subjective improvement also, a weighted mean square error criterion is employed, where the weighting corresponds to characteristics of the human visual system.

* *

14. Key Words: Digital image coding; Digital image restoration; Quantization; Multidimensional densities; Probability theory.

DD FORM 1 NOV 65 1473

UNCLASSIFIED

Security Classification

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to Professor William K. Pratt, whose assistance and invaluable suggestions made this dissertation possible. I am indebted to him for introducing me to the study of image processing and for his advice and guidance throughout my graduate career. I would like to thank Professor Harry C. Andrews for providing countless insights into image processing and for his critical review of this thesis. I am deeply grateful for the support, discussions, and friendship of Dr. Faramarz Davarian and Dr. Robert H. Wallis which formed an integral part of my education.

Thanks also go to my wife Mary for her love and encouragement throughout the course of this research.

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

ABSTRACT

An analysis of the optimum statistical restoration of quantized signals is presented. The restoration is based upon minimizing the mean square error between the input to a quantizer and its estimate. Since a quantizer is a nonlinear device, the estimation equation which is derived achieves an optimum nonlinear restoration. Its solution requires complete statistical knowledge of the quantizer input. Available statistical information usually includes the marginal distribution of each of the input variables and the correlation between them. Hence a technique is developed for generating correlated multidimensional probability density functions based on this available information. The technique is applied to gaussian, laplacian, and Rayleigh density functions. These multidimensional density functions characterize the outputs of transform coders, DPCM coders, and PCM coders, respectively. The quantized outputs of these coders are then restored by utilizing the multidimensional densities in the estimation equation. Examples of images which have been coded and restored by these techniques are presented. The results reveal a mean square error reduction. To achieve a visually subjective improvement also, a weighted mean square error criterion is employed, where the weighting corresponds to characteristics of the human visual system.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOTATION

$\underline{a}$	vector
$\underline{A}$	matrix
$\underline{A}^T$	transpose of the matrix $\underline{A}$
$\underline{A}^{-1}$	inverse of the matrix $\underline{A}$
$\underline{A}^*$	conjugate transpose of the matrix $\underline{A}$
$\underline{A} \otimes \underline{B}$	left direct product of $\underline{A}$ and $\underline{B}$
λ	eigenvalue
$\underline{e}_i$	eigenvector
$p(\cdot)$	probability density function
$Pr\{A\}$	probability of occurrence of the event A
$Tr\{\underline{A}\}$	trace of the matrix $\underline{A}$
$E\{\cdot\}$	expected value
$\underline{C}$	correlation matrix
R	quantization region in N-space
R_m	mth quantization region
d_k	kth decision level of quantizer
r_k	kth restoration level of quantizer
r	correlation coefficient
$U(\cdot)$	unit step function
$\delta(\cdot)$	Dirac delta function
$\mathcal{E}$	error

CHAPTER 1

INTRODUCTION

Quantization is a process inherent in all digital systems. Basically, quantization occurs whenever continuous physical properties are represented numerically. When this representation takes place in a digital computer, the quantization effect is called round-off, or truncation. As a mathematical operation, quantization is the processing of continuous functions to give a stepwise output, or the processing of sampled functions to give a sampled output. Even the value obtained in measuring a continuous quantity is the consequence of quantization. But wherever the occurrence, a fundamental aspect of quantization is that it results in an indeterminacy and a lack of complete information about the particular property under consideration.

In this report quantization is assumed to be a nonlinear operation which occurs within a quantizer--a zero-memory device that assigns an input to one of a countable number of possible output regions. This defines a broad class of devices that includes coders, digital transducers, and analog-to-digital converters. For many of these quantizing devices the input is a continuous variable; restricting a continuous input to a particular

region destroys some of the information about that input. For discrete inputs, combining their input regions into larger, and hence fewer allowable, output regions also decreases the amount of information available. These are both irreversible operations and the lost information cannot be recovered. It is thus important to optimize the quantizing process so that this lost information is held to a minimum.

The lost information can be minimized by decreasing the size of the output regions while simultaneously increasing their number. Unfortunately, this is not always possible or practical. It would also oppose one of the benefits of quantization: a smaller number of output regions requires less processing and less storage. A balance thus must be attained between accuracy and economy. This balance can be determined by an analysis of the quantization process.

Most analyses of quantization to date have focused on just one aspect of the problem, i.e., finding the best quantizing device to minimize the information lost. However, the eventual use of any quantized output is to accurately represent a continuous signal input. The output regions are ultimately utilized to estimate and restore the original quantizer input. It is this quantization restoration problem that has heretofore been neglected and

is the subject of this dissertation.

The simplest restoration procedure is to choose the midpoint of each quantization interval as the estimated value of the original input. However, this estimate can be improved, since it is based only on the output regions of the quantizer. The restoration to be described herein is based also on a priori knowledge of the statistics of the quantizer input. The input is assumed to consist of samples from a continuous random process. (The sampling presents few, if any, restrictions because digital systems require sampled and quantized signals, and the operations of sampling and quantizing are commutable.) The necessary statistics are the amplitude probability density function of the input samples. Where the complete statistics are unknown, a functional form for them is developed from known correlation functions and one-dimensional distributions. This statistical information is then combined with knowledge of the quantizer output to provide an optimum restoration. The restoration is optimum with respect to a desired error criterion.

A second problem requiring a similar solution occurs when quantized signals are transmitted through a noisy channel. Because of the errors that accrue during transmission, what is received does not exactly correspond to the quantizer output regions. The channel output

instead equals the quantizer output only within a specified probability. To achieve an optimum restoration the following available information must be utilized:

1. the a priori distribution of the quantizer input
2. the structure of the quantizer
3. the transmitted quantizer output region (which may be in error)
4. the channel error structure

The existence of the last constituent induces a modification in the solution discussed previously. The modification is also considered in this dissertation.

In this dissertation, solutions are presented for the restoration of quantized samples based on a priori knowledge of the multivariate probability density function of the quantizer input. The two cases considered are: (1), the quantizer output region is known exactly; and (2), the quantizer output is transmitted over a noisy channel and hence not known exactly.

These two situations arise in the coding and transmission of images. Quantizers are an integral part of all image coding systems. The goal of these systems is to make a coded image as similar as possible to an original image. Unfortunately cost, complexity, and hardware constraints often force a suboptimal coding scheme which results in a degraded image. The application of the

quantization restoration techniques discussed above can improve these degraded images. Experimental verification of this improvement is obtained by restoring images which have been coded and quantized.

CHAPTER 2

HISTORY OF QUANTIZATION AND SIGNAL RECONSTRUCTION

Developments in quantization have closely paralleled advances in digital systems. Although research had been conducted into areas such as uniform statistical grouping (which may be considered quantization) as early as 1898 [1], it is only since 1947 that deliberate attempts have been made to understand the process of quantization. By 1947 vacuum-tube technology had reached the stage for which digital systems were both possible and practical. At that time the concepts and the value of PCM (pulse code modulation--the first major application of both qunatization and digital hardware) were just being discovered and made known [2,3,4]. Bennett [5] then undertook an intensive investigation of the spectra of quantized signals. Bennett analyzed uniform quantizers, such as those uitlized in PCM systems, and found the characteristics of their output spectra for a white noise input spectrum. Since Bennett's initial work, developments in quantization have proceeded along three basic lines:

1. analyzing the results and the process of quantization
2. optimizing the quantization process with respect to various criteria and goals
3. reconstructing quantized signals to minimize the

degradation incurred through quantization

Each of these categories will be considered separately in the following paragraphs, and significant developments will be discussed in roughly chronological order.

2.1 Analysis of Quantization

One of Bennett's conclusions was that quantization uncertainty or noise, for a quantizer with many levels, is uniformly distributed throughout the signal band. This result was supported by Widrow in 1956 [6] in studies of the probability density functions of quantized signals and quantization noise. Widrow concluded that a quantizer could be modelled as a source of uniform, independent noise. In a later paper [7], Widrow attempted to define the limits of the region over which his additive noise conclusion would be valid, and then extended the statistical results to two dimensions (i.e., the quantization of two correlated samples). Myers [8] extended Bennett's analysis to the case of the uniform quantization of a signal corrupted by gaussian noise and derived the resultant noise distribution, which is no longer uniform. Velichkin [9], calculated the correlation function and output spectrum of a quantizer. In addition, Velichkin considered the more general cases of nonuniform quantization levels and arbitrary input signal spectrums. Velichkin's results, for the gaussian case, were in the

form of an infinite summation of Hermite polynomials of increasing order which, unfortunately, cannot be evaluated without simplifying assumptions.

Robertson [10] surmounted this difficulty by evaluating combinations of the terms of the summation such that the combinations tend to zero and hence obtained output spectrums for nonlinear and nonsymmetrical quantizers and for arbitrary input spectrum shapes. Chan and Donaldson [11] obtained a further generalization by finding the correlation function and spectrum of a quantized gaussian signal transmitted over a discrete memoryless channel. Their results reduce to those obtained by Velichkin when the channel is noiseless. For very coarse quantization, Curry and Vander Velde [12] suggested modelling the quantizer as a gain element, whose value is equal to the random-input describing function, plus an additive noise source. The inclusion of a gain element causes the noise source to appear more nearly white. The quantization process can then be analyzed more easily.

2.2 Optimum Quantizers

The analyses of quantization described above were all based on the concept that quantization introduces a noise or distortion. A number of researchers have attacked the fundamental problem of minimizing a measure of this

distortion by varying the location of N quantizing levels, given the characteristics of the quantizer input. They attempted to do this in an optimal fashion under different assumptions and conditions, and according to various criteria.

In 1951 Panter and Dite [13] tried to minimize the mean-square quantization error by utilizing statistical properties of the signal. They developed an optimum nonuniform quantizer based on the following assumptions:

1. the quantizer is symmetrical about zero
2. the probability density function, p(x), is an even function and is constant over each quantization interval
3. the signal is limited to the range [-V,V]
4. a signal quantized to a particular interval is restored to the midpoint of that interval

The resultant quantizer restoration levels, r_k, can then be calculated from

$$r_k = \frac{V\int_{\emptyset}^{2kV/N} [p(x)]^{-1/3}\,dx}{\int_{\emptyset}^{V} [p(x)]^{-1/3}\,dx} \qquad (2.1)$$

for k=1,2,...,n and where N=2n+1 is the total number of levels. The total distortion power for this choice of

levels is

$$\mathcal{E}=\frac{2}{3N^2}\left\{\int_{\emptyset}^{V}[p(x)]^{1/3}dx\right\}^3 \qquad (2.2)$$

Panter and Dite also suggested that this nonuniform level spacing could be realized by "companding"--compressing the original signal by a nonlinear function such as a logarithm, performing a uniform quantization, and then expanding the result by means of the inverse of the nonlinear function. This type of system was later analyzed by Smith [14] who provided a method for choosing the parameters of the nonlinearity with respect to a mean-square error criterion.

In a fundamental paper in 1960, Max [15] derived the necessary equations for finding the parameters of a quantizer having minimum distortion with respect to a convex error criterion. For a fixed number of quantization intervals, N, the decision levels, d_k, and the restoration levels, r_k, are obtained by a recursive solution of

$$d_k=(r_k+r_{k-1})/2 \qquad k=2,3,\ldots,N \qquad (2.3a)$$

$$\int_{d_k}^{d_{k+1}}\frac{\partial e(x-r_k)}{\partial x}p(x)dx=\emptyset \qquad k=1,2,\ldots,N \qquad (2.3b)$$

where p(x) is the distribution function of the quantizer input and e(·) is a convex error function. (Note that these equations do not require p(x) to be constant over the range of the quantizer.) In this notation, a signal quantized to the interval (d_k, d_{k+1}) would be restored to the point r_k. The set of simultaneous equations contained in eq. 2.3, except for trivial cases, cannot be solved explicitly and so must be evaluated numerically. Max also derived the equation which provides the optimum quantizer parameters for a uniform spacing of levels, and tabulated both the optimum uniform and optimum nonuniform quantization levels for a gaussian probability density function and a mean-square error criterion.

The difficulty in solving eq. 2.3 explicitly has led a number of researchers to consider various approximations. Garmash [16] simplified the integrals in eq. 2.3b by the trapezoidal rule and reduced the number of equations by choosing the size of the smallest interval. His results are valid only for a finite signal range. Roe [17] approximated the probability density function, p(x), by the first two terms of its Taylor series expansion about the midpoint of each quatization interval to obtain

$$\int_{0}^{d_k} [p(x)]^{1/(t+1)} dx \approx 2ak+b \qquad (2.4)$$

for k=1,2,...,N-1 where t is the order of the error criterion (t=2 for a mean-square error) and a and b are constants. This relation approximately provides the decision levels (and restoration levels for $k'=k+1/2$) for differentiable probability density functions, but the resultant quantizer is identical to that obtained by Panter and Dite [13]. Algazi [18], in attempting to find some simpler suboptimal algorithms, also rederived eq. 2.2. Williams [19] published a closed-form solution to eq. 2.3 for the special case of a laplacian distribution, but erred in assuming the restoration point to be the median, rather than the centroid, of the quantization interval. The correct quantization levels for a laplacian (and also a gamma) distribution were later calculated numerically [20].

No matter which method is utilized to calculate the quantization parameters, however, it will fail if the probability distribution is such that a unique minimum distortion point does not exist. For a mean-square distortion measure, Fleischer [21] derived the sufficient conditions under which a unique optimum quantizer can be found. To this end also, Bruce [22] used dynamic programming to find optimum quantizers for a variety of convex error criteria, and checked their uniqueness by locating and comparing all other extrema of the error surfaces.

Since quantizers usually operate on sampled signals, Velichkin [23] and Goodman [24] considered the joint optimization of both sampling and quantizing. Velichkin calculated the optimum parameters to achieve a minimum mean-square error, and Goodman compared this result to a lower bound obtained from rate distortion theory.

It has been found [25] that uniform quantizers approach this lower bound, i.e., uniform quantizers asymptotically have the lowest output entropy. Wood [26] and O'Neal [27] have taken advantage of this fact to derive (approximately) minimum entropy quantizers. Wood has shown that for a fixed output entropy, uniform quantizers have lower mean-square error than nonuniform (Max) quantizers. However, to achieve this error reduction, the quantizer output must be optimally coded (i.e., with a Huffman variable-length code) and this causes buffering problems. This difficulty may be partially overcome by permutation encoding [28], but this coding technique has the limitation of requiring very long block-lengths. For a non-buffered coding scheme, a quantizer having maximum output entropy could be considered to be optimum. It has been shown that the quantizers with minimum average error (such as Max's) are the same as those with maximum entropy, within a multiplicative constant [29].

Maximum entropy, or minimum error, quantizers must be

modified when their output is transmitted over a noisy channel [30]. For a mean-square error criterion, the decision and reconstruction levels for a noisy-channel quantizer can be found from

$$d_k = \frac{\sum_{i=1}^{N} r_i^2 (P_{ki} - P_{k-1,i})}{2 \sum_{i=1}^{N} r_i (P_{ki} - P_{k-1,i})} \qquad k=2,3,\ldots,N \qquad (2.5a)$$

$$r_k = \frac{\sum_{i=1}^{N} P_{ik} \int_{d_i}^{d_{i+1}} xp(x)dx}{\sum_{i=1}^{N} P_{ik} \int_{d_i}^{d_{i+1}} p(x)dx} \qquad k=1,2,\ldots,N \qquad (2.5b)$$

where P_{ki} is an element of the channel matrix, $\underline{P}$. For a noiseless channel ($\underline{P}=\underline{I}$), eq. 2.5 reduces to eq. 2.3. Minimum error quantizers also must be modified for nonstationary inputs or for correlated input samples. Golding and Schultheiss [31] and Stroh and Boorstyn [32] presented ad hoc adaptive quantizers designed to handle this situation. An optimum adaptive quantizer has not yet been successfully derived, nor has an optimum quantizer for correlated signals.

2.3 Quantized Signal Reconstruction

An important area of research in recent years has been the reconstruction of quantized signals. Some researchers have attempted to restore the input signal samples which were degraded by a quantizer, while others, realizing that many quantizers operate on sampled versions of analog signals, have tried to restore directly the analog waveform. In this latter category are the techniques of Ruchkin [33], Katzenelson [34], Steiglitz [35], Goblick [36], Kellogg [37], Hayes [38], and Chan and Donaldson [39]. Each of their approaches differed in the assumptions they made about the effect of the quantizer. Ruchkin found the best (with respect to mean-square error) linear filter to restore a quantized and sampled gaussian signal, under the assumption that a quantizer adds white gaussian noise. Katzenelson also found the best linear filter, but assumed the signal is a sample from a Markov process and the quantizer noise is colored. The same assumption about the quantizer was made by Steiglitz who analyzed specific, nonoptimum, reconstruction filters for special input power spectra and compared the results to a rate distortion bound. Steiglitz provided a trade-off between sampling rate and number of uniform quantization levels for the case of a minimum mean-square error and fixed capacity. Goblick, on the other hand, considered entropy coding and provided a trade-off between mean-square error and output

entropy. A prefilter was derived by Kellogg for the same case that Ruchkin considered, under the premise that a designer can tailor input and output signals to a fixed quantizer by means of prefilters and postfilters. Kellogg used an optimum, nonuniform quantizer for his simulations, and presented a numerical solution for the linear filters. Hayes added a constraint on the signal-to-noise ratio for his solutions, but otherwise numerically analyzed the same case as Kellogg. An exact analytical method for jointly optimizing the prefilter, postfilter, quantizer, and sampling rate was presented by Chan and Donaldson. Their optimum prefilter "whitens" the signal and removes the redundancy from the resultant signal samples. After quantization and transmission occur, a postfilter restores the continuous signal and removes the distortion added by both the quantizer and the channel. This method was general in that it considered the correlation between the signal and noise, included the effect of channel errors, and placed no restrictions on either the signal and noise spectra or the filter passbands <*>.

Curry [41] considered the problem of finding linear estimates based on quantized samples. He showed that these

---

<*> It should be mentioned that an excellent general treatment of quantizers, receivers, and noisy channels is available in Fine [40].

estimates can be computed in two steps: (1) find the conditional mean of the quantizer output, and (2) pass this mean through the linear filter that would have been used had the quantizer not existed. For a minimum mean-square error criterion this linear filter was found to be the Kalman or Wiener filter. The conditional mean for a gaussian process was found to be approximately

$$\underline{r}=(\underline{I}-\underline{D}\underline{C}^{-1})(\underline{b}+\underline{a})/2 \tag{2.6}$$

where $\underline{I}$ is the identity matrix, $\underline{C}$ is the covariance matrix of the quantizer input $\underline{x}$, $\underline{x}$ is quantized to the interval $[\underline{a},\underline{b}]$ by an arbitrary quantizer, and $\underline{D}$ is the diagonal matrix having elements

$$\underline{D}=\left\{\frac{(b_i-a_i)^2}{12}\delta_{ij}\right\} \tag{2.7}$$

Equation 2.6 provides a minimum mean-square error estimate of a quantizer's input, based only on its output, that is valid whenever the quantization is very fine. Clements and Haddad [42] derived a recursive techinque for finding this same result that is also applicable to the problem of nonlinear estimation using quantized data.

REFERENCES

1. W. F. Sheppard, "On the calculation of the most Probable Values of Frequency-Constants, for Data arranged according to Equidistant Divisions of a Scale," Proceedings of the London Mathematical Society, vol. 29, March 1898, pp. 353-380.

2. W. M. Goodall, "Telephony by Pulse Code Modulation," Bell System Technical Journal, vol. 26, July 1947, pp. 395-409.

3. A. G. Clavier, P. F. Panter, and D. D. Grieg, "PCM Distortion Analysis," Electrical Engineering, vol. 66, November 1947, pp. 1110-1122.

4. B. N. Oliver, J. R. Pierce, and C. E. Shannon, "The Philosophy of PCM," Proceedings of the IRE, vol. 36, November 1948, pp. 1324-1331.

5. W. R. Bennett, "Spectra of Quantized Signals," Bell System Technical Journal, vol. 27, July 1948, pp. 446-472.

6. B. Widrow, "A Study of Rough Amplitude Quantization by Means of Nyquist Sampling Theory," IRE Transactions on Circuit Theory, vol. CT-3, December 1956, pp. 266-276.

7. B. Widrow, "Statistical Analysis of Amplitude-Quantized Sampled-Data Systems," AIEE Transactions (Applications and Industry), vol. 80, January 1961, pp. 555-568.

8. G. H. Myers, "Quantization of a Signal Plus Random Noise," IRE Transactions on Instrumentation, vol. I-5, June 1956, pp. 181-186.

9. A. I. Velichkin, "Correlation Function and Spectral Density of a Quantized Process," Telecommunications and Radio Engineering, Part II: Radio Engineering, July 1962, pp. 70-77.

10. G. H. Robertson, "Computer Study of Quantizer Output Spectra," Bell System Technical Journal, vol. 48, September 1969, pp.2391-2403.

11. D. Chan and R. W. Donaldson, "Correlation Functions and Reconstruction Error for Quantized Gaussian Signals Transmitted over Discrete Memoryless Channels," IEEE Transactions on Information Theory, vol. IT-18, July 1972, pp. 519-523.

12. R. E. Curry and W. E. Vander Velde, "An Extended Criterion for Statistical Linearization," IEEE Transactions on Automatic Control, vol. AC-15, February 1970, pp. 106-108.

13. P. F. Panter and W. Dite, "Quantization Distortion in Pulse-Count Modulation with Nonuniform Spacing of Levels," Proceedings of the IRE, vol. 39, January 1951, pp. 44-48.

14. B. Smith, "Instantaneous Companding of Quantized Signals," Bell System Technical Journal, vol. 46, May 1967, pp. 653-709.

15. J. Max, "Quantizing for Minimum Distortion," IEEE Transactions on Information Theory, vol. IT-6, March 1960, pp. 7-12.

16. V. A. Garmash, "Quantization of Signals with Nonuniform Steps," Telecommunications, no. 10, 1957, pp. 10-12.

17. G. M. Roe, "Quantizing for Minimum Distortion," IEEE Transactions on Information Theory, vol. IT-10, October 1964, pp. 384-385.

18. V. R. Algazi, "Useful Approximations to Optimum Quantization," IEEE Transactions on Communications Technology, vol. COM-14, June 1966, pp. 297-301.

19. G. Williams, "Quantizing for Minimum Error with Particular Reference to Speech," Electronics Letters, vol. 3, April 1967, pp. 134-135.

20. M. D. Paez and T. H. Glisson, "Minimum Mean-Squared-Error Quantization in Speech PCM and DPCM Systems," IEEE Transactions on Communications, vol. COM-20, April 1972, pp. 225-230.

21. P. E. Fleischer, "Sufficient Conditions for Achieving Minimum Distortion in a Quantizer," 1964 IEEE International Convention Record, part 1, pp. 104-111.

22. J. D. Bruce, "Optimum Quantization," Technical Report 429, MIT Research Lab of Electronics, Cambridge, Massachusetts, March 1965.

23. A. I. Velichkin, "Quantization of Continuous Messages with Minimum Mean-Square Error," Telecommunications and Radio Engineering, vol. 18, February 1963, pp. 114-122.

24. L. M. Goodman, "Optimum Sampling and Quantization Rates," Proceedings of the IEEE, vol. 54, January 1966, pp. 90-92.

25. H. Gish and J. N. Pierce, "Asymptotically Efficient Quantizing," IEEE Transactions on Information Theory, vol. IT-14, September 1968, pp. 676-683.

26. R. C. Wood, "On Optimum Quantization," IEEE Transactions on Information Theory, vol. IT-15, March 1969, pp. 248-252.

27. J. B. O'Neal, Jr., "Entropy Coding in Speech and Television Differential PCM Systems," IEEE Transactions on Information Theory, vol. IT-17, November 1971, pp. 758-761.

28. T. Berger, "Optimum Quantizers and Permutation Codes," IEEE Transactions on Information Theory, vol. IT-18, November 1972, pp. 759-765.

29. D. G. Messerschmitt, "Quantizing for Maximum Output Entropy," IEEE Transactions on Information Theory, vol. IT-17, September 1971, p. 612.

30. A. J. Kurtenbach and P. A. Wintz, "Quantizing for Noisy Channels," IEEE Transactions on Communications Technology, vol. COM-17, April 1969, pp. 291-302.

31. L. S. Golding and P. M. Schultheiss, "Study of an Adaptive Quantizer," Proceedings of the IEEE, vol. 55, March 1967, pp. 293-297.

32. R. W. Stroh and R. R. Boorstyn, "A Technique for Adaptive Quantization," Proceedings of the UMR-Mervin J. Kelly Communications Conference, Rolla, Missouri, October 1970, pp. 17-5-1 to 17-5-6.

33. D. S. Ruchkin, "Linear Reconstruction of Quantized and Sampled Random Signals," IRE Transactions on Communications Systems, vol. COM-9, December 1961, pp. 350-355.

34. J. Katzenelson, "On Errors Introduced by Combined Sampling and Quantization," IRE Transactions on Automatic Control, vol. AC-7, April 1962, pp. 58-68.

35. K. Steiglitz, "Transmission of an Analog Signal Over a Fixed Bit-Rate Channel," IEEE Transactions on Information Theory, vol. IT-12, October 1966, pp. 469-474.

36. T. J. Goblick, Jr. and J. L. Holsinger, "Analog Source Digitization: A Comparison of Theory and Practicc," IEEE Transactions on Information Theory, vol. IT-13, April 1967, pp. 323-326.

37. W. C. Kellogg, "Information Rates in Sampling and Quantization," IEEE Transactions on Information Theory, vol. IT-13, July 1967, pp. 506-511.

38. J. F. Hayes, "Optimum Prefiltering and Postfiltering of Sampled Waveforms," Proceedings of the 1970 SWIEEECO Conference, pp. 24-28.

39. D. Chan and R. W. Donaldson, "Optimum Pre- and Postfiltering of Sampled Signals with Application to Pulse Modulation and Data Compression Systems," IEEE Transactions on Communications Technology, vol. COM-19, April 1971, pp. 141-156.

40. T. Fine, "Properties of an Optimum Digital System and Applications," IEEE Transactions on Information Theory, vol. IT-10, October 1964, pp. 287-296.

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

42. K. A. Clements and R. A. Haddad, "Approximate Estimation for Systems with Quantized Data," IEEE Transactions on Automatic Control, vol. AC-17, April 1972, pp. 235-239.

CHAPTER 3

NONLINEAR ESTIMATION WITH QUANTIZED MEASUREMENTS

As the previous chapter has shown, most analyses of quantization have focused on the quantizing process itself and on the determination of an optimum quantizer for each input sample. Unfortunately, little effort has been expended towards undoing the effects of quantization and recovering the original input. The core of the problem is that quantizers treat each signal individually and neglect the random process which models the signals. As the analysis below shows, an optimum restoration of quantized signals must utilize the joint probability density of the original input signals.

3.1 Vector Quantization

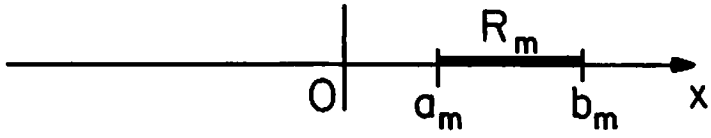
The input to a quantizer may be either a continuous function or discrete samples. A quantizer processes a continuous function to give a stepwise continuous output or processes a sampled function to give a sampled output. All digital systems require signals which are both sampled and quantized. Sampling and quantizing are mathematically commutable operations, i.e., the result is the same whether a signal is first quantized and the resultant stepfunction is sampled, or if the signal is first sampled and then the

samples are quantized. Hence in this analysis the input to a quantizer is assumed to be a sampled random process.

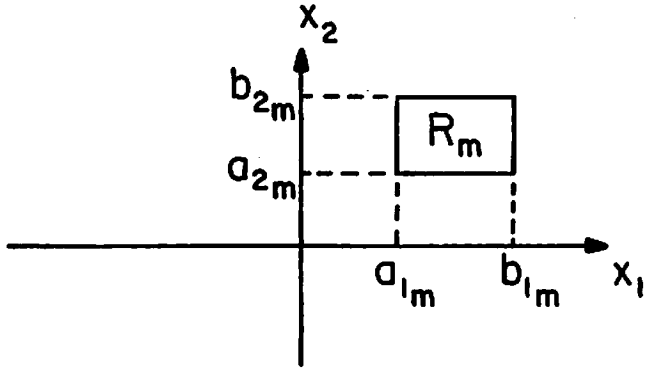
A quantizer assigns each sample to a particular interval according to its amplitude (quantization is sometimes considered to be sampling in amplitude). For an individual sample, the interval is a portion of the real line (see fig. 3-1a). However if two samples at a time are considered, then a quantizer assigns pairs of samples to regions in two-space, as shown in fig. 3-1b. Similarly, a quantizer assigns vectors of N samples to regions in N-space, R^N. If the vector components are quantized independently, the resultant region in R^N is rectangular. An estimate of the quantizer input based on its corresponding output region would be one of the points within that region. The goal in this chapter is to find the estimation point which is most similar to the quantizer input, with respect to a given error criterion. It should be mentioned that no attempt is made to find optimum quantization regions. Rather, a method is derived for finding the optimum restoration point within an arbitrary, given region. The restoration concepts can best be understood in the context of the following simple example.

3.2 Vector Restoration Example

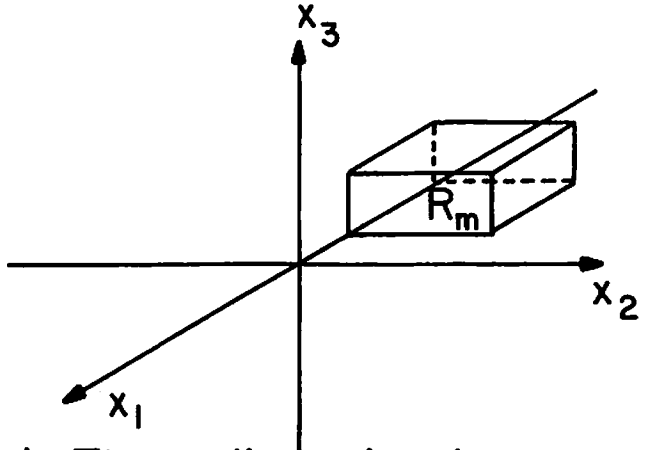
Consider a two level (one bit) quantization of a



(a) One-dimensional



(b) Two-dimensional



(c) Three-dimensional

Figure 3-1. Vector quantization regions.

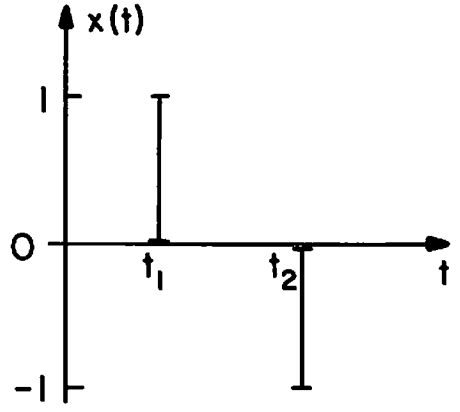
sampled signal limited in amplitude to a finite range. Specifically, for a signal limited to the range [-1,1], let the output of a quantizer be the information that the signal is either in [0,1] or in [-1,0). Next assume that two successive outputs at times t_1 and t_2 are as shown in fig. 3-2a.

$$x(t_1) \in [0,1]$$
$$x(t_2) \in [-1,0)$$

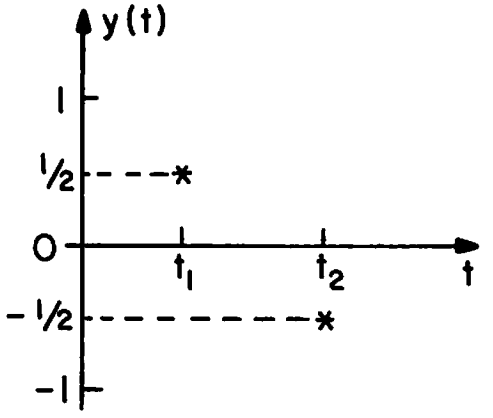
If nothing else is known about these signals, then the best restoration, as fig. 3-2b shows, would be

$$y(t_1)=1/2$$
$$y(t_2)=-1/2$$

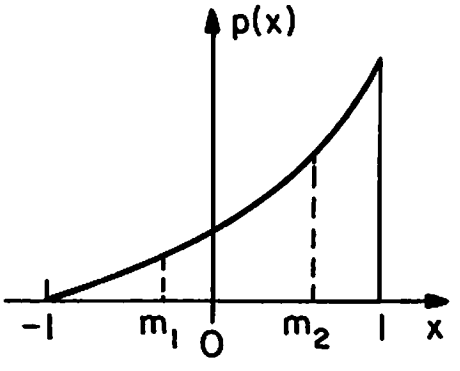
However, if the signals are samples from a random process with a known probability distribution, p(x), such as the one in fig. 3-2c, then a better restoration would be the mean values of each interval according to this distribution (see fig. 3-2d). (The restoration in fig. 3-2b would correspond to a uniform distribution.) Finally, if the signals at t_1 and t_2 are known to be correlated, then the restoration can be further improved by utilizing this correlation. Figures 3-2e and 3-2f show the restoration points for positive correlation and negative correlation,



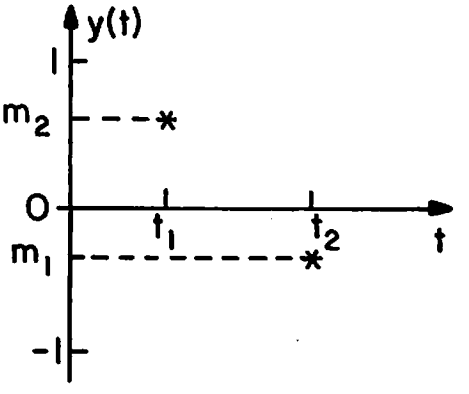
a. Quantization intervals for x_1 and x_2



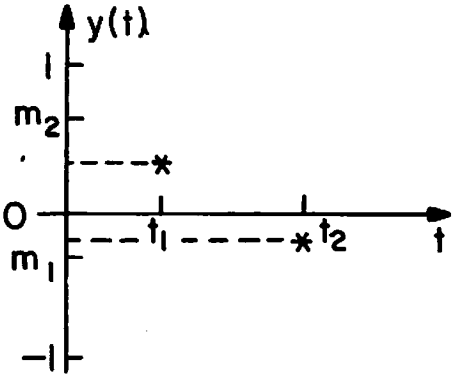
b. Restoration of x_1 and x_2 without prior information



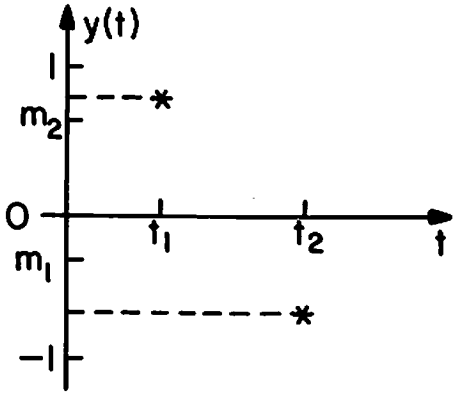
c. Probability distribution of x



d. Restoration of x_1 and x_2 according to p(x)



e. Restoration for strong positive correlation between x_1 and x_2



f. Restoration for negative correlation between x_1 and x_2

Figure 3-2. Vector restoration example.

respectively.

3.3 Nonlinear Estimator

To characterize the above concepts mathematically, an expression for the quantization error must first be derived. Let a vector of N samples of a random process be denoted as $\underline{x}$, where

$$\underline{x}=(x_1,x_2,\ldots,x_N) \tag{3.1}$$

and assume that this vector is statistically described by its probability density function, $p(\underline{x})$. If this vector is quantized, according to the techniques of Sec. 3.1, then it is assigned to one of M regions in R^N, denoted as R_m, for m=1,2,...,M. Next assume that the quantizer's desired output is its input signal. This is a reasonable assumption, for in most communication systems the quantized signal is intended to be an instantaneous replica of the input signal. Thus, it can also be assumed that $\underline{x} \in R_m$, i.e., that the input vector $\underline{x}$ is assigned to the region in which it is contained. Note that these assumptions are made for convenience only and place no restrictions whatsoever on the subsequent solution.

The point within R_m that is chosen as the restored value of the quantizer input is denoted as $\underline{y}_m$, where $\underline{y}_m$ is

a vector since it represents a point in N-space. The instantaneous error caused by restoring a quantized signal to the point $\underline{y}_m$ is then $\underline{x}-\underline{y}_m$. The restoration performance is to be determined by evaluating a function of this error, $e(\cdot)$. The goal is to minimize the mean value of this error function, for all possible quantizer inputs, by the proper choice of restoration points, $\underline{y}_m$. The expression for the total error which must be minimized is then

$$\mathcal{E} = \sum_{m=1}^{M} \int_{R_m} e(\underline{x}-\underline{y}_m)\, p(\underline{x})\, d\underline{x} \qquad (3.2)$$

The error weighting function, e, is arbitrary, but it is usually a nonnegative function because instantaneous errors of opposite sign should not cancel each other.

Many researchers have proposed solutions which minimize eq. 3.2 for different choices of the error weighting function, but all of their results are predicated on $\underline{x}$ being a one-dimensional or scalar input. Of these results, the most significant ones have been obtained by Bruce [1] who used dynamic programing techniques to find the optimum regions, R_m, and restoration points, $\underline{y}_m$, for arbitrary error weighting functions. His results generalized those obtained earlier by Max [2] for a squared error weighting.

A mean-square error criterion is also utilized here. The functional form of this error weighting is

$$e(\underline{x}-\underline{y}_m)=\mathrm{Tr}\{(\underline{x}-\underline{y}_m)(\underline{x}-\underline{y}_m)^T\} \tag{3.3}$$

where $\mathrm{Tr}\{\cdot\}$ denotes the trace of a matrix. The error function $e(\underline{x}-\underline{y}_m)$ is now a monotonic function. Then, under a simple assumption on $p(\underline{x})$ (that $p(\underline{x})$ is not entirely discrete), the error expression in eq. 3.2 can be minimized by the techniques of differential calculus. By making this assumption about $p(\underline{x})$, the relative extrema of the error surface $\mathscr{E}$ can be found from

$$\frac{\partial \mathscr{E}}{\partial \underline{y}_m}=\underline{\emptyset} \tag{3.4}$$

for m=1,2,...,M. Substituting equations 3.2 and 3.3 into eq. 3.4 yields

$$\frac{\partial \mathscr{E}}{\partial \underline{y}_m}=\int_{R_m} -2(\underline{x}-\underline{y}_m)p(\underline{x})d\underline{x} = \underline{\emptyset} \tag{3.5}$$

for m=1,2,...,M. This equation can be rearranged and the restoration point, $\underline{y}_m$, solved for to obtain

$$\underline{y}_m = \frac{\int_{R_m} \underline{x} p(\underline{x}) d\underline{x}}{\int_{R_m} p(\underline{x}) d\underline{x}} \tag{3.6}$$

for m=1,2,...,M. Equation 3.6 is thus a minimum mean-square error estimate of the input to a quantizer, based on the quantizer output and the a priori statistics of the input. Closer inspection of this equation reveals that

$$\underline{y}_m = E\{\underline{x} \mid \underline{x} \in R_m\} \tag{3.7}$$

for m=1,2,...,M. That is, $\underline{y}_m$ is the conditional mean estimate of $\underline{x}$ conditioned on the nonlinear information that $\underline{x}$ has been quantized to the region R_m. Because of this conditioning, $\underline{y}_m$ represents a nonlinear restoration of the quantized signal. The restoration is optimum only with respect to a mean-square error criterion, but it is completely independent of the specific form of the Quantizer.

Note that eq. 3.6 requires that a multidimensional density function, $p(\underline{x})$, be provided in order to solve for the optimal restoration. Since a gaussian density is the only known multidimensional density for correlated signals, only quantized gaussian random processes might be optimally

restored. If other multidimensional densities can be derived, then systems which are based on these densities might also be optimized. The subsequent chapter considers this derivation possibility and presents a technique for generating correlated multidimensional densities. The results are then utilized in succeeding chapters to analyze several common communication systems.

3.4 Estimation Covariance

The performance of the estimator derived in the previous section can be determined by computing the estimation error covariance matrix. The covariance of the estimator, based on the quantized measurement information, can be found from

$$\mathrm{cov}(\underline{x} \mid \underline{x} \in R_m) = E\{(\underline{x}-\underline{y}_m)(\underline{x}-\underline{y}_m)^T \mid \underline{x} \in R_m\} \qquad (3.8)$$

By comparison with equations 3.6 and 3.7 this can be expressed as

$$\mathrm{cov}(\underline{x} \mid \underline{x} \in R_m) = \frac{\int_{R_m} (\underline{x}-\underline{y}_m)(\underline{x}-\underline{y}_m)^T p(\underline{x})\, d\underline{x}}{\int_{R_m} p(\underline{x})\, d\underline{x}} \qquad (3.9)$$

The numerator can then be expanded and the resulting equation simplified to obtain the final form for the

estimation covariance

$$\text{cov}(\underline{x} \mid \underline{x} \in R_m) = \frac{\int_{R_m} \underline{x}\underline{x}^T p(\underline{x})\, d\underline{x}}{\int_{R_m} p(\underline{x})\, d\underline{x}} - \underline{y}_m \underline{y}_m^T \qquad (3.10)$$

Using this result, the average quantization error remaining after the minimum mean-square error restoration of Sec. 3.3 is

$$\mathcal{E} = \sum_{m=1}^{M} \text{Tr}\{\text{cov}(\underline{x} \mid \underline{x} \in R_m)\} \Pr(\underline{x} \in R_m) \qquad (3.11)$$

where Pr(A) denotes the probability of occurrence of the event A and

$$\Pr(\underline{x} \in R_m) = \int_{R_m} p(\underline{x})\, d\underline{x} \qquad (3.12)$$

Then using eq. 3.10, the error becomes

$$\mathcal{E} = \text{Tr}\left\{ \sum_{m=1}^{M} \left[\int_{R_m} \underline{x}\underline{x}^T p(\underline{x})\, d\underline{x} - \underline{y}_m \underline{y}_m^T \Pr(\underline{x} \in R_m) \right] \right\} \qquad (3.13)$$

which simplifies to

$$\mathcal{E} = \text{Tr}\{\underline{C} - \sum_{m=1}^{M} \underline{y}_m \underline{y}_m^T \Pr(\underline{x} \in R_m)\} \qquad (3.14)$$

where the matrix $\underline{C}$ is the covariance of the input vector $\underline{x}$. Alternatively, this same result could have been derived by substituting equations 3.6 and 3.3 into eq. 3.2 and then simplifying the result. The error expression in eq. 3.13 is used later to evaluate the performances of various optimal restoration techniques that are based on quantized measurements.

REFERENCES

1. J. D. Bruce, "Optimum Quantization," Technical Report 429, MIT Research Lab of Electronics, Cambridge, Massachusetts, March 1965.

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

CHAPTER 4

MULTIVARIATE PROBABILITY DENSITY FUNCTIONS

4.1 Introduction

Multivariate distributions of dependent random variables are often required in communications problems for the analysis of stationary random processes. The distributions arise whenever a continuous random process is sampled--a fundamental characteristic of all digital systems. The probability density function of N samples of a random process X(t) is usually denoted as

$$p(x_1, x_2, \ldots, x_N)$$

This expression is very general and represents an infinite number of possible density functions, only a few of which have ever been found. In fact, for N>2 and correlated variables, there is only the familiar multivariate gaussian distribution. Fortunately the gaussian is the most useful distribution (by virtue of the Central Limit Theorem making it the limiting distribution for many additive processes). There are, however, many processes which it does not model well. For these, the lack of other known distributions often means that a multivariate gaussian is used by default, or that the dependence between the variables is ignored and the density is written as the product of

independent marginal densities. In either case inaccuracies result. Thus, this chapter addresses the problem of finding the multivariate densities of non-gaussian variables whose marginal densities and correlation function are known.

4.2 Characteristics of Multivariate Distributions

The probability density function of a single random variable is uniquely determined by specifying all M+1 of its moments, $1 \leq M \leq \infty$, denoted by [1]

$$E\{x^m\}$$

for m=0,1,2,...,M. Uniqueness holds under the following easily satisfied conditions: when all of the moments are finite and when a power series with the moments as its coefficients converges. To uniquely specify the multivariate density of N random variables requires $(M+1)^N$ joint moments, similarly denoted as

$$E\{x_1^{m_1} x_2^{m_2} \ldots x_N^{m_N}\}$$

for $m_1, m_2, \ldots, m_N = 0, 1, \ldots, M$. For a given random process it is usually possible to determine the distribution, $p(x_i)$, of a single random variable (a marginal density function of the multivariate density), and to measure its correlation

with the other similarly distributed variables. If there are N variables, then the N marginal densities provide N(M+1) moments and the correlations provide N^2-N more. The remaining mixed moments are undetermined. If these are chosen arbitrarily, then there are an infinite number of possible multivariate densities which have the same $NM+N^2$ moments specified above. However, these densities differ from each other only in their higher moments, and thus are very similar.

A valid probability density function must satisfy the following six conditions [2]:

Property 1. $$p(x_1,x_2,\ldots,x_N)\geq 0 \tag{4.1}$$

Property 2. $$\int\ldots\int p(x_1,x_2,\ldots,x_N)dx_1dx_2\ldots dx_N=1 \tag{4.2}$$

Property 3. $$\int p(x_1,x_2,\ldots,x_N)dx_i=p(x_1,\ldots,x_{i-1},x_{i+1}\ldots,x_N) \tag{4.3}$$

for i=1,2,...,N.

Property 4. To accurately represent a given random process with correlation matrix $\underline{C}$, the density function must be

able to generate the elements of $\underline{C}$ from the equation

$$C_{ij}=\int\cdots\int x_i x_j p(x_1,x_2,\ldots,x_N)dx_1dx_2\ldots dx_N \qquad (4.4)$$

Property 5. When a stationary random process X(t) is sampled, the samples, in general, are correlated with a correlation matrix $\underline{C}$. If the sample spacing becomes so large, however, that the samples are uncorrelated and $\underline{C}$ is a diagonal matrix (assuming X(t) has no strictly periodical components), then the samples are also independent. Then the density function should be factorable into a product of independent density functions, as

$$p(x_1,x_2,\ldots,x_N)=p(x_1)p(x_2)\ldots p(x_N) \qquad (4.5)$$

Property 6. If two of the random variables are identical, e.g. $x_i=x_j$, then the density must become

$$p(x_1,x_2,\ldots,x_N)=\delta(x_i-x_j)p(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_N) \qquad (4.6)$$

4.3 Prior Multivariate Density Research

A few specialized multivariate densities have been found in the past. In 1945 S. O. Rice [3] derived

two-dimensional Rayleigh and sine-cosine densities. Subsequent researchers, such as Gumbel [4] and Parzen [5], found multivariate densities based on known marginals, but their densities failed to satisfy Properties 4 and 6 and so could not represent a given random process. The most significant results to date were obtained by Beckmann [6] who developed a general technique for constructing a two-dimensional density from its marginals and correlation coefficient. Beckmann's method required each marginal density to be the weighting function of a system of orthogonal polynomials. Unfortunately this resulted in a joint density containing an infinite series that often did not have a closed form solution. Also, by the inherent nature of orthogonal polynomials, the results were not extendable to more than two dimensions.

4.4 Derivation of Multidimensional Densities

To derive densities in higher dimensions, consider first an N-dimensional gaussian density. Assume, without loss of generality, that its mean is zero and its covariance is $\underline{C}$. Then this density can be written

$$p(\underline{x}) = (2\pi)^{-N/2} |\underline{C}|^{-1/2} \exp(-\tfrac{1}{2}\underline{x}^T \underline{C}^{-1} \underline{x}) \qquad (4.7)$$

It is possible, by a suitable linear transformation, to express this density in a new coordinate system in which

the components are statistically independent. The density transformation is effected through

$$p_Y(\underline{y}) = p_X(\underline{x}) \left| \frac{d\underline{x}}{d\underline{y}} \right| \tag{4.8}$$

Let

$$\underline{y} = \underline{E}^T \underline{x} \tag{4.9}$$

where $\underline{E}$ is the solution to the eigenvalue equation

$$\underline{E}\underline{\Lambda} = \underline{C}\underline{E} \tag{4.10}$$

with $\underline{\Lambda}$ a diagonal matrix of eigenvalues. Since $\underline{C}$ is symmetric and a covariance matrix, the eigenvalues are real and nonnegative. Also the eigenvectors which comprise the columns of $\underline{E}$ can be chosen to be orthonormal, i.e.

$$\underline{E}^T \underline{E} = \underline{I} \tag{4.11}$$

Now substituting eq. 4.9 into equations 4.8 and 4.7 yields

$$p_Y(\underline{y}) = p_X(\underline{E}\underline{y}) \, |\underline{E}^{-1}| = (2\pi)^{-N/2} \, |\underline{E}|^{-1} \, |\underline{C}|^{-1/2} \exp\left(-\tfrac{1}{2} \underline{y}^T \underline{E}^T \underline{C}^{-1} \underline{E}\underline{y}\right)$$

or

$$p_{\underline{Y}}(\underline{y}) = (2\pi)^{-N/2} |\underline{\Lambda}|^{-1/2} \exp(-\frac{1}{2}\underline{y}^T \underline{\Lambda}^{-1} \underline{y}) \qquad (4.12)$$

Notice that if $\underline{\Lambda}$ in this last equation is an identity matrix, then the original covariance matrix $\underline{C}$ must also be an identity matrix, for

$$\underline{C} = \underline{E\Lambda E}^T = \underline{EE}^T = \underline{I} \qquad (4.13)$$

The diagonal elements of $\underline{\Lambda}$ are just the variances of the components in the new coordinate system. If the components are scaled by the square root of these variances, i.e. $z_i = y_i\sqrt{\lambda_{ii}}$, then each new component has unit variance. Hence rotating these scaled coordinates by the inverse of the original transformation

$$\underline{w} = \underline{Ez} \qquad (4.14)$$

yields an uncorrelated unit variance probability density function

$$p(\underline{w}) = (2\pi)^{-N/2} \exp(-\frac{1}{2}\underline{w}^T \underline{E\Lambda}^{1/2} \underline{\Lambda}^{-1} \underline{\Lambda}^{1/2} \underline{E}^{-1} \underline{w})$$

$$= (2\pi)^{-N/2} \exp(-\frac{1}{2}\underline{w}^T \underline{w}) \qquad (4.15)$$

Reversing the above procedure thus leads to a method for generating a correlated multidimensional density function

from an uncorrelated one.

Reiterating the steps above in proper order yields the following general technique:

(1) Write an uncorrelated N-dimensional density as a product of its known marginals

$$p(\underline{x})=p(x_1)p(x_2)\ldots p(x_N) \tag{4.16}$$

(2) Define $\underline{E}$ such that $\underline{E}$ is orthonormal and

$$\underline{E}^T\underline{C}\underline{E}=\underline{\Lambda} \tag{4.17}$$

where $\underline{\Lambda}$ is diagonal and $\underline{C}$ is the desired covariance matrix. The matrices $\underline{E}$ and $\underline{\Lambda}$ are respectively the matrix of eigenvectors and the matrix of eigenvalues of $\underline{C}$.

(3) Transform $\underline{x}$ by $\underline{E}^{-1}$, i.e., $\underline{y}=\underline{E}^{-1}\underline{x}=\underline{E}^T\underline{x}$, so that

$$p_Y(\underline{y})=|\underline{E}|p_X(\underline{E}\underline{y})=p_X(\underline{E}\underline{y}) \tag{4.18}$$

(4) Scale the components of $\underline{y}$ by the square roots of the eigenvalues; then

$$\underline{z}=\underline{\Lambda}^{1/2}\underline{y} \tag{4.19}$$

and

$$p_Z(\underline{z}) = |\underline{\Lambda}|^{-1/2} p_Y(\underline{\Lambda}^{1/2}\underline{z}) = |\underline{\Lambda}|^{-1/2} p_X(\underline{E\Lambda}^{-1/2}\underline{z}) \qquad (4.20)$$

(5) Inverse transform $\underline{z}$ by $\underline{w}=\underline{E}\underline{z}$ to obtain

$$p_W(\underline{w}) = |\underline{E}|^T p_Z(\underline{E}^T\underline{w}) = |\underline{\Lambda}|^{-1/2} p_X(\underline{E\Lambda}^{-1/2}\underline{E}^T\underline{w})$$

or

$$p_W(\underline{w}) = |\underline{A}| p_X(\underline{A}\underline{w}) \qquad (4.21)$$

where

$$\underline{A} = \underline{E\Lambda}^{-1/2}\underline{E}^T \qquad (4.22)$$

To illustrate the effect of each step of the procedure, fig. 4-1 shows contour plots of two-dimensional gaussian and laplacian densities.

4.5 Proof of Density Properties

The function $p_W(\underline{w})$ in eq. 4.21 has correlation matrix $\underline{C}$ and satisfies the conditions for a valid probability density function as is next shown.

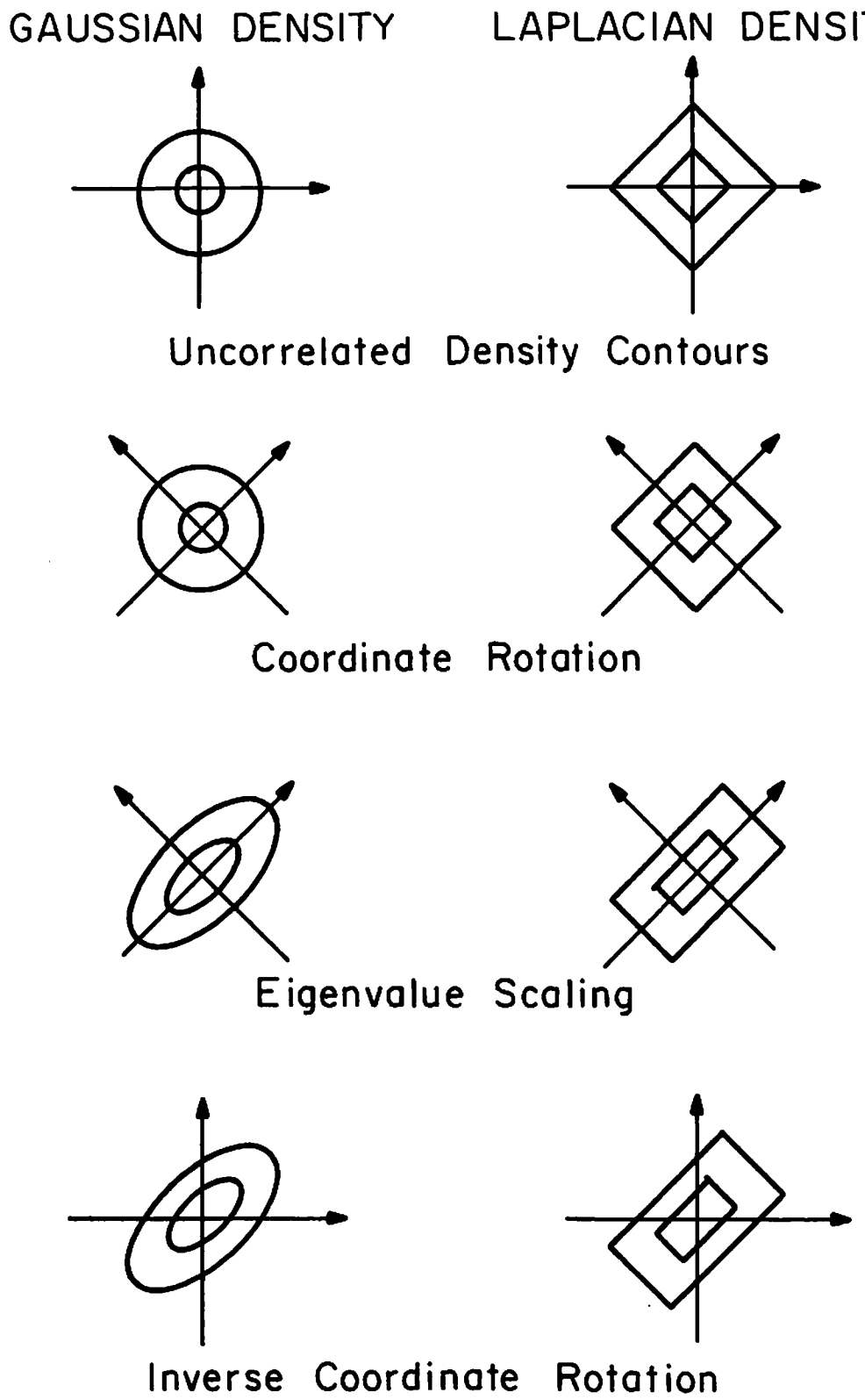


Figure 4-1. Contour plots of the two-dimensional density transformations used to generate correlated densities.

Property 1. $p_W(\underline{w})$ is clearly nonnegative because $p_X(\underline{x}) \geq 0$ for all $\underline{x} \in R^N$ and $\underline{Aw} \in R^N$.

Property 2. $$\int_{R^N} p_W(\underline{w})\,d\underline{w} = \int_{R^N} |\underline{A}|\,p_W(\underline{Aw})\,d\underline{w}$$

$$= \int_{R^N} p_X(\underline{x})\,d\underline{x} = 1 \tag{4.23}$$

Property 3. $$\int_R p_W(\underline{w})\,dw_i \neq p(w_1,\ldots,w_{i-1},w_{i+1}\ldots,w_N) \tag{4.24}$$

Property 4. $$E\{w_i w_j\} = E\{(\underline{A}^{-1}\underline{x})_i\,(\underline{A}^{-1}\underline{x})_j\}$$

$$= \underline{A}_i^{-1} E\{\underline{x}\underline{x}^T\}\underline{A}_j^{-1}$$

$$= (\underline{E}\underline{\Lambda}^{1/2}\underline{E}^T\underline{E}\underline{\Lambda}^{1/2}\underline{E}^T)_{ij}$$

$$= (\underline{C})_{ij} \tag{4.25}$$

Property 5. $\underline{C}=\underline{I}$ implies that $\underline{A}=\underline{I}$; then

$$p_W(\underline{w}) = |\underline{A}|\,p_X(\underline{Aw}) = p_X(\underline{w}) = p(w_1)\,p(w_2)\ldots p(w_N) \tag{4.26}$$

Property 6. If $x_i = x_j$, then $a_{ii} = -a_{ij} = a_{jj} = -a_{ji}$. Now $\lim_{x_i \to x_j} a_{ii} = a_{jj} = \infty$. Hence $\lim_{x_i \to x_j} |\underline{A}| = \infty$ and

$$p_{X_i}(\underline{A}_i\underline{w}) = p_{X_i}\left(\sum_{k=1}^{N} a_{ik}x_k\right) = \begin{cases} p_{X_i}(\infty)\ ; & x_i > x_j \\ p_{X_i}\left(\sum_{k=1}^{N} a_{ik}x_k\right); & x_i = x_j \\ p_{X_i}(-\infty)\ ; & x_i < x_j \end{cases} \tag{4.27}$$

Thus

$$\lim_{x_i \to x_j} p_W(\underline{w}) = \begin{cases} 0; & x_i > x_j \\ \infty\ ; & x_i = x_j \\ 0; & x < x \end{cases}$$

$$= \delta(x_i - x_j) \tag{4.28}$$

Unfortunately, Property 3 is not satisfied, i.e., the correlated density function no longer has the desired marginal distributions. In fact, its marginal distributions are a function of the correlation matrix $\underline{C}$. However, as is shown in more detail later, the correlated density function found above remains a good representation for a given random process, and is a good approximation to the exact (but unknown) probability density function.

4.6 Examples of Multidimensional Densities

To demonstrate the utility of the derivation procedure, some examples are now presented.

A. Gaussian probability density

$$p_X(\underline{x}) = (2\pi)^{-N/2} \exp(-\tfrac{1}{2}\underline{x}^T \underline{x}) \tag{4.29}$$

$$p_W(\underline{w}) = |\underline{\Lambda}|^{-1/2} (2\pi)^{-N/2} \exp(-\tfrac{1}{2}\underline{w}^T \underline{A}^T \underline{A}\underline{w})$$

$$= |\underline{C}|^{-1/2} (2\pi)^{-N/2} \exp(-\tfrac{1}{2}\underline{w}^T \underline{E}\underline{\Lambda}^{-1/2} \underline{E}^T \underline{E}\underline{\Lambda}^{-1/2} \underline{E}^T \underline{w})$$

$$= |\underline{C}|^{-1/2} (2\pi)^{-N/2} \exp(-\tfrac{1}{2}\underline{w}^T \underline{C}^{-1} \underline{w}) \tag{4.30}$$

B. Laplacian probability density

$$p_X(\underline{x}) = \frac{1}{2^{N/2}} \exp(-\sqrt{2} \sum_{i=1}^{N} |x_i|) \tag{4.31}$$

$$p_W(\underline{w}) = |\underline{C}|^{-1/2} \frac{1}{2^{N/2}} \exp(-\sqrt{2} \sum_{i=1}^{N} \left| \sum_{k=1}^{N} a_{ik} w_k \right|) \tag{4.32}$$

C. Rayleigh probability density

$$p_X(\underline{x}) = \left[\prod_{i=1}^{N} x_i U(x_i) \right] \exp(-\tfrac{1}{2}\underline{x}^T \underline{x}) \tag{4.33}$$

$$p_W(\underline{w}) = |\underline{C}|^{-1/2} \left[\prod_{i=1}^{N} \left(\sum_{k=1}^{N} a_{ik} w_k \right) U\left(\sum_{k=1}^{N} a_{ik} w_k \right) \right] \exp(-\tfrac{1}{2}\underline{w}^T \underline{C}^{-1} \underline{w}) \tag{4.34}$$

D. Maxwell probability density

$$p_{\underline{X}}(\underline{x}) = (2/\pi)^{N/2}\left[\prod_{i=1}^{N} x_i^{\,2}\, U(x_i)\right] \exp(-\tfrac{1}{2}\underline{x}^T\underline{x}) \qquad (4.35)$$

$$p_{\underline{W}}(\underline{w}) = \frac{(2/\pi)^{N/2}}{|\underline{C}|^{1/2}}\left[\prod_{i=1}^{N} \left(\sum_{k=1}^{N} a_{ik} w_k\right)^2 U\left(\sum_{k=1}^{N} a_{ik} w_k\right)\right] \exp(-\tfrac{1}{2}\underline{w}^T\underline{C}^{-1}\underline{w}) \qquad (4.36)$$

Figures 4-2 through 4-9 contain plots of two-dimensional gaussian, laplacian, and Rayleigh densities for both uncorrelated and correlated variables.

4.7 Marginal Densities and Random Process Simulations

As stated previously (see Section 4.5), the correlated multidimensional density function in eq. 4.21 and the uncorrelated density in eq. 4.16 have different marginal density functions. In generating the correlated density by steps (1) through (5), the marginal density becomes a function of the correlation matrix $\underline{C}$. For example, a two-dimensional correlated laplacian density can be written

$$p(x,y) = \frac{1}{2s_x s_y\sqrt{1-r^2}} \exp\left\{-\frac{1}{\sqrt{2(1-r^2)}}\left(\left|\frac{ax}{s_x} - \frac{by}{s_y}\right| + \left|\frac{ay}{s_y} - \frac{bx}{s_x}\right|\right)\right\} \qquad (4.37)$$

where

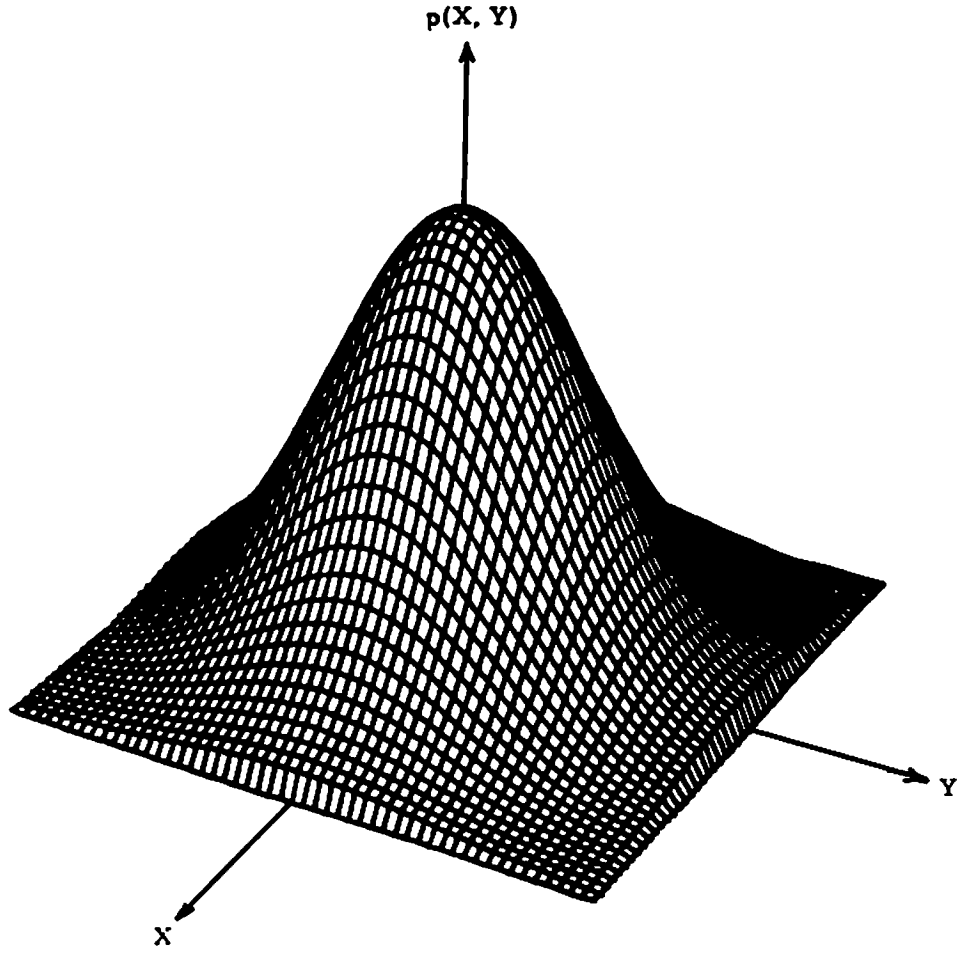


FIGURE 4-2. TWO-DIMENSIONAL GAUSSIAN DENSITY
Correlation = 0.0

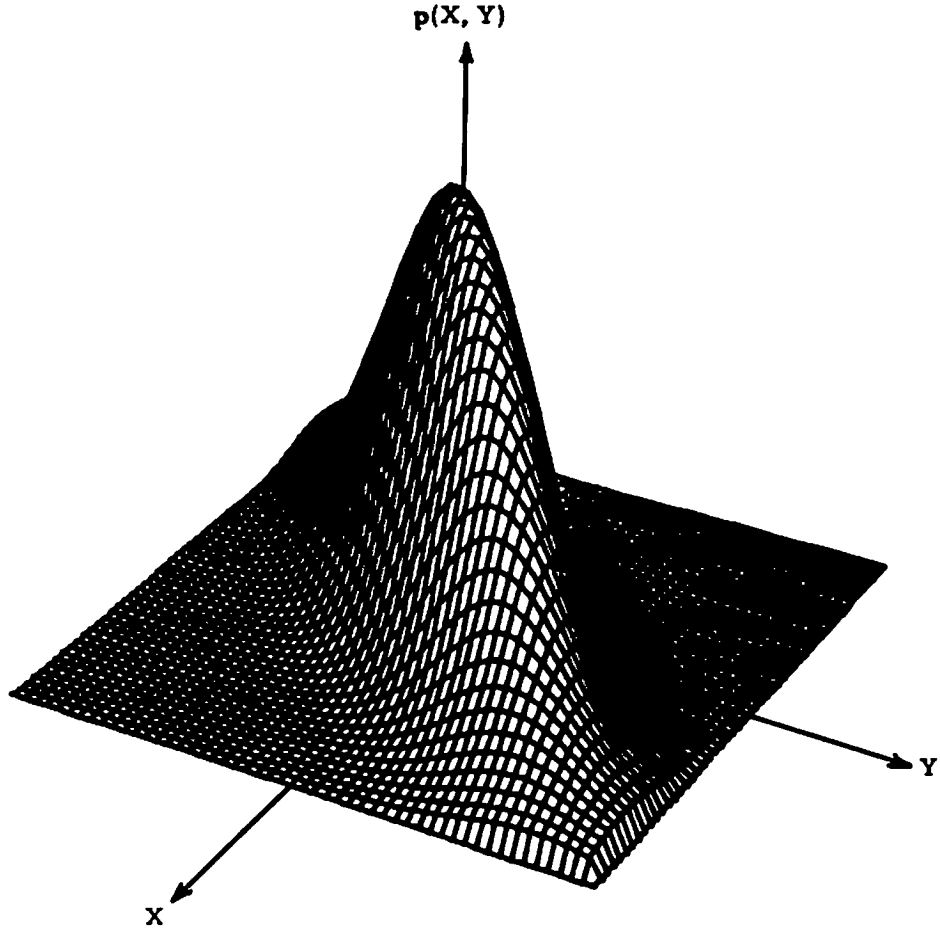


FIGURE 4-3. TWO-DIMENSIONAL GAUSSIAN DENSITY
Correlation = 0.8

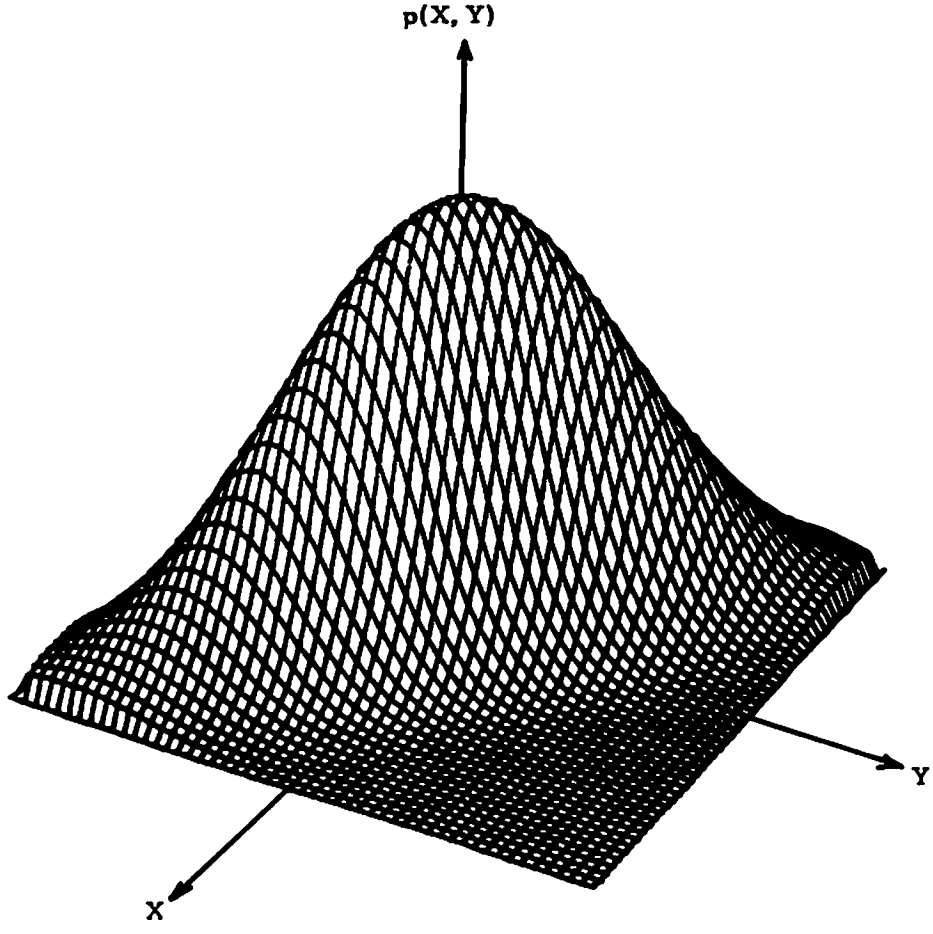


FIGURE 4-4. TWO-DIMENSIONAL GAUSSIAN DENSITY
Correlation = -0.8

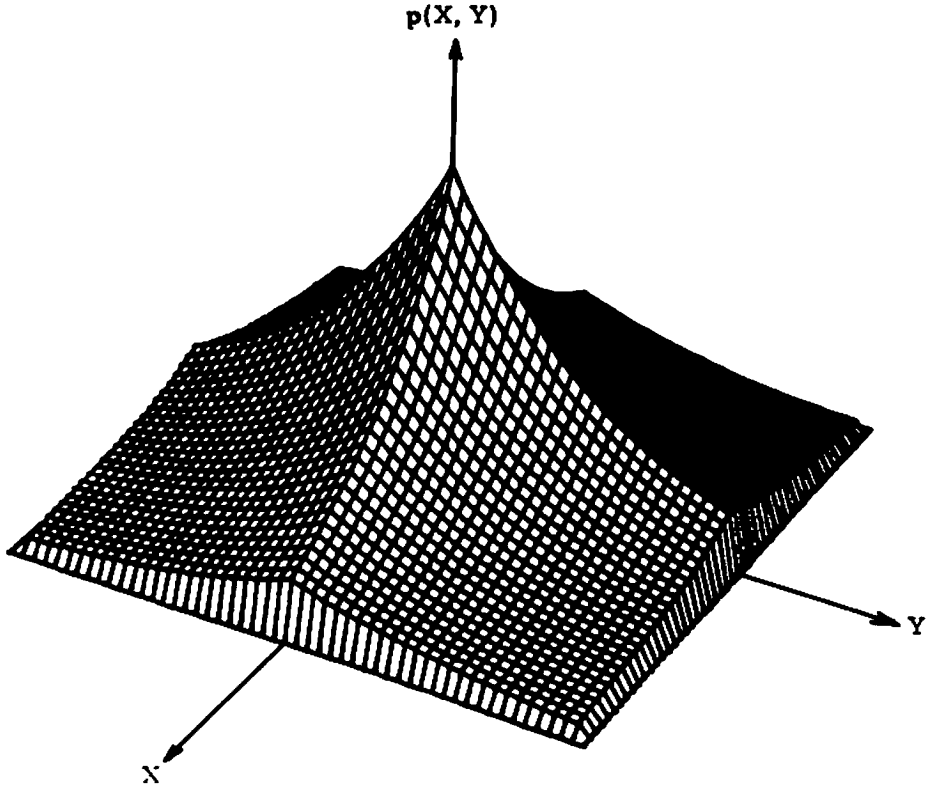


FIGURE 4-5. TWO-DIMENSIONAL LAPLACIAN DENSITY
Correlation = 0.0

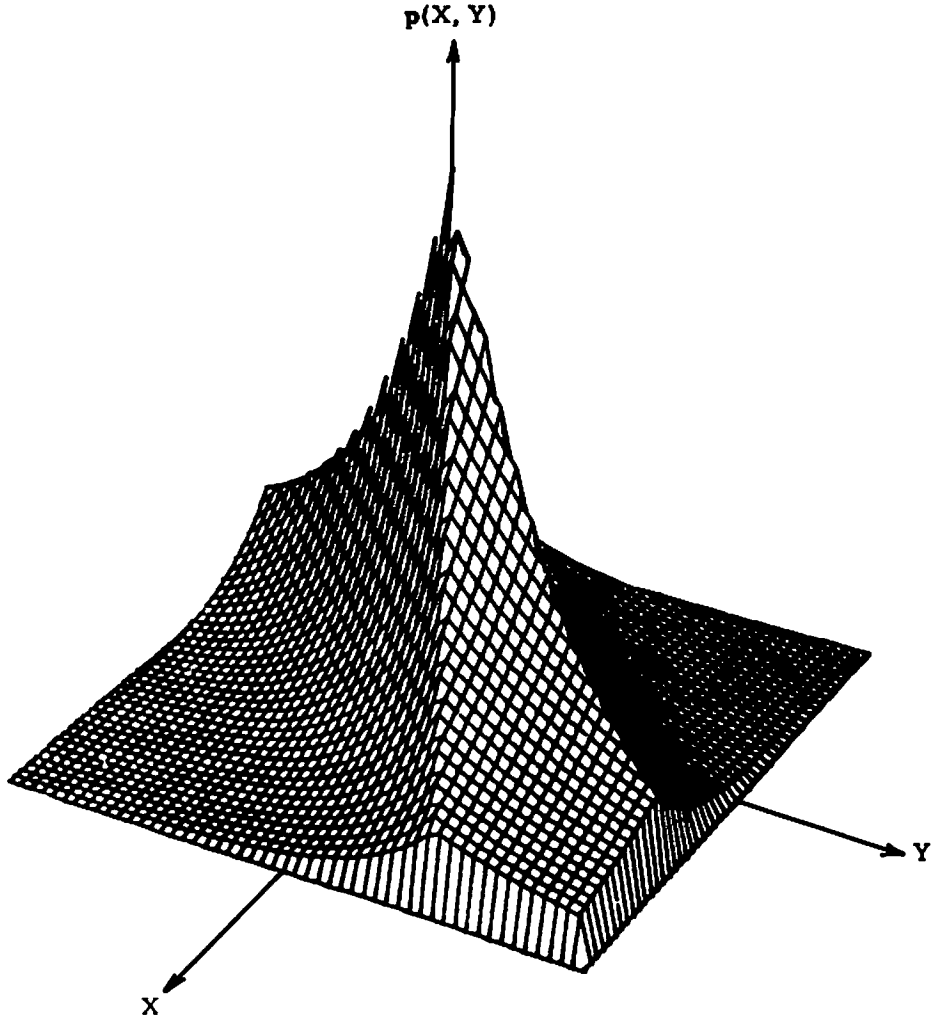


FIGURE 4-6. TWO-DIMENSIONAL LAPLACIAN DENSITY
Correlation = 0.8

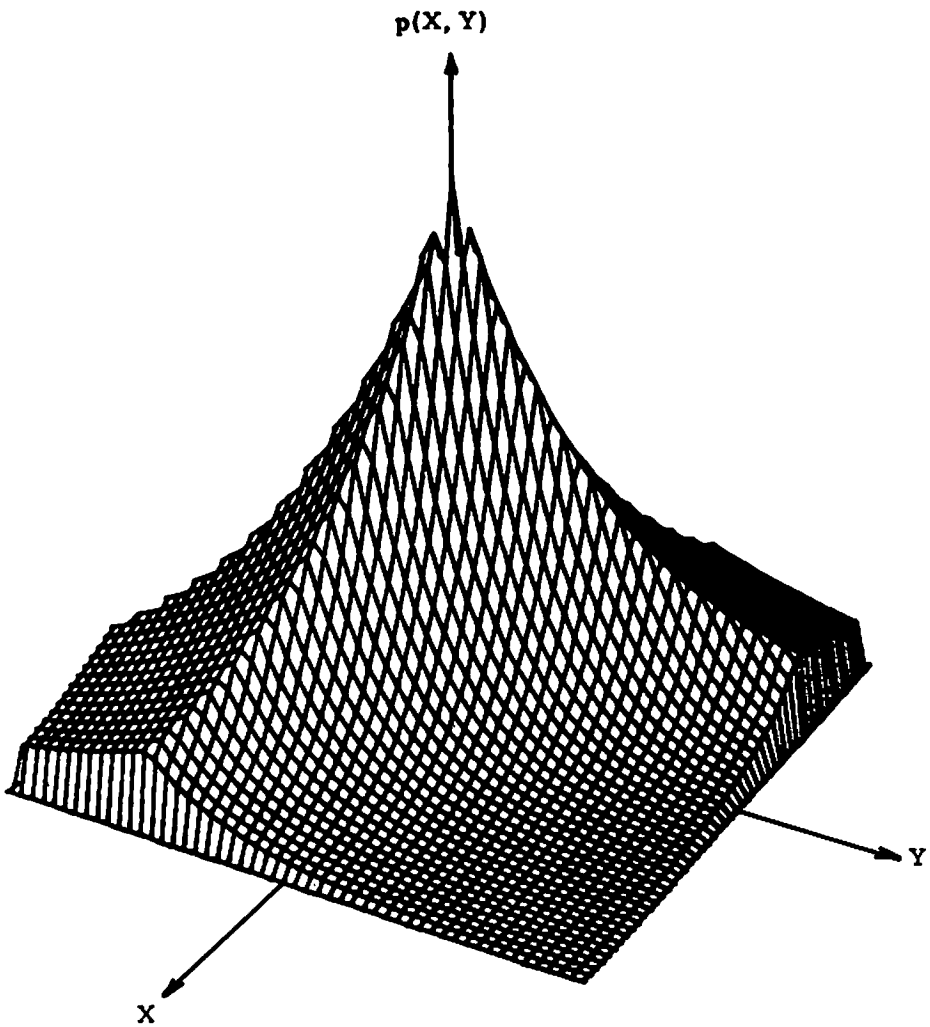


FIGURE 4-7. TWO-DIMENSIONAL LAPLACIAN DENSITY
Correlation = -0.8

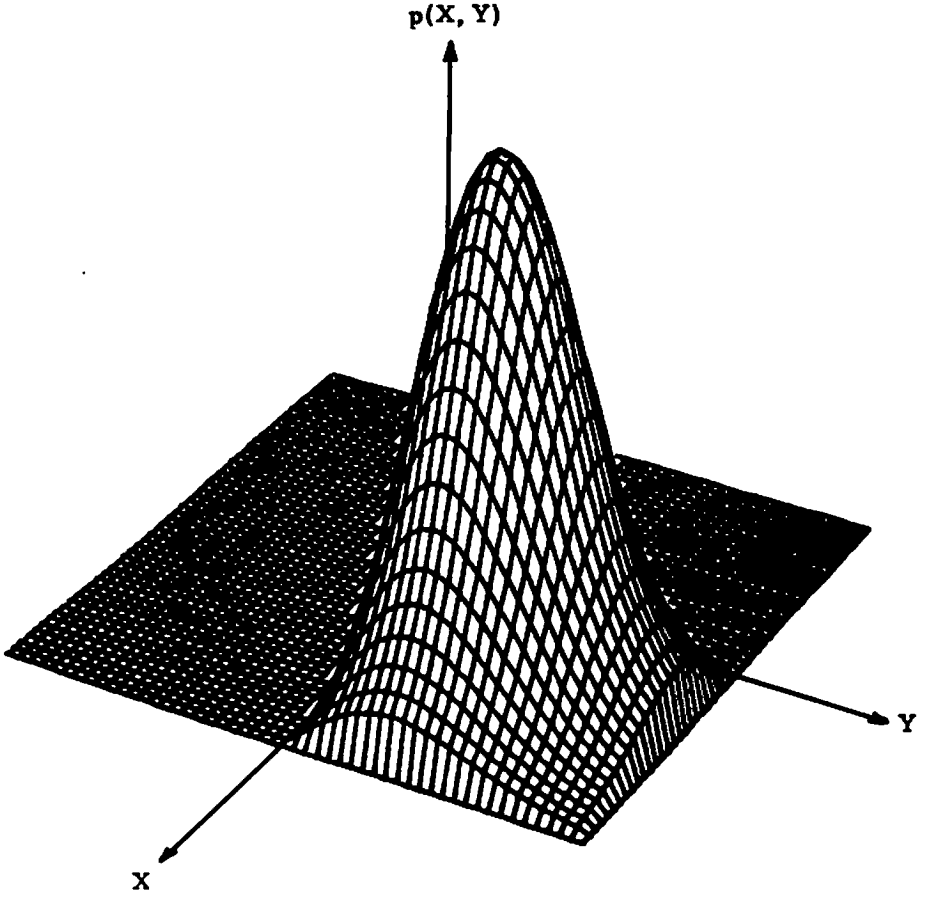


FIGURE 4-8. TWO-DIMENSIONAL RAYLEIGH DENSITY
Correlation = 0.0

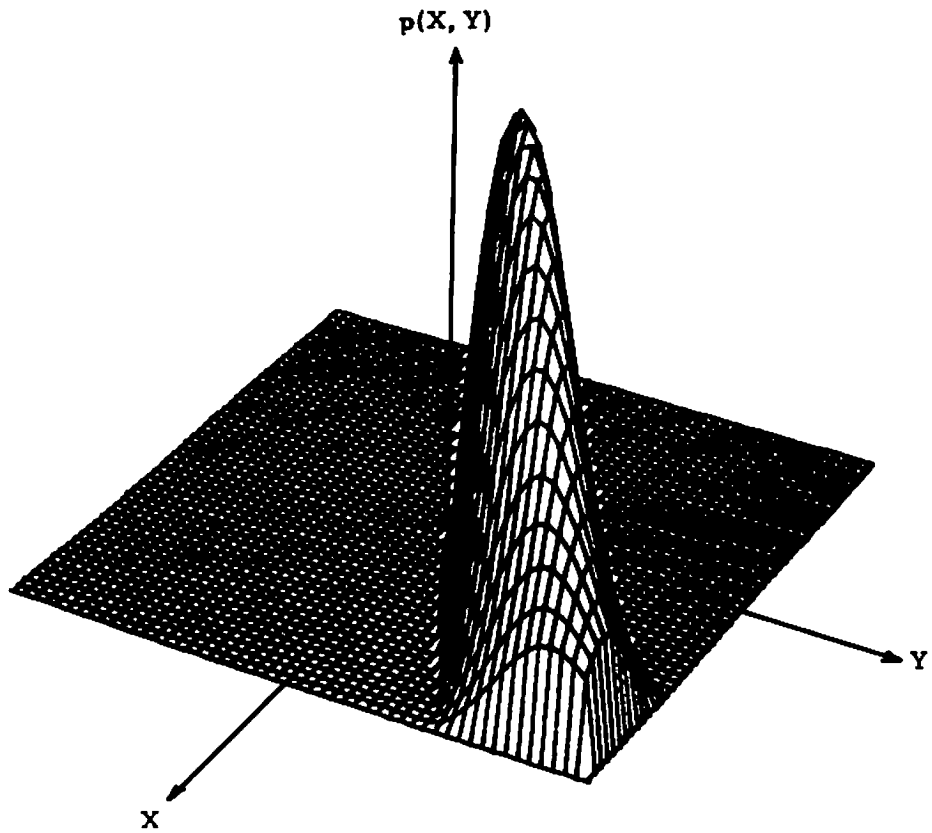


FIGURE 4-9. TWO-DIMENSIONAL RAYLEIGH DENSITY
Correlation = 0.8

$$s_x^2 = E\{x^2\} \tag{4.38}$$

$$s_y^2 = E\{y^2\} \tag{4.39}$$

$$r = \frac{E\{xy\}}{s_x s_y} \tag{4.40}$$

$$a = \sqrt{1+r} + \sqrt{1-r} \tag{4.41}$$

$$b = \sqrt{1+r} - \sqrt{1-r} \tag{4.42}$$

Its marginal distribution is

$$p(x) = \frac{1}{2s_x\sqrt{2(1-r^2)}}\left[|a|\exp\left(-\frac{2\sqrt{2}}{s_x}\left|\frac{x}{a}\right|\right) - |b|\exp\left(-\frac{2\sqrt{2}}{s_x}\left|\frac{x}{b}\right|\right)\right] \tag{4.43}$$

which is a function of the correlation, r. It becomes the classical laplacian distribution only when the correlation is zero, i.e.

$$p(x)\Big|_{r=0} = \frac{1}{\sqrt{2}s_x}\exp\left(-\frac{\sqrt{2}|x|}{s_x}\right) \tag{4.44}$$

However, non-zero correlations in eq. 4.43 result in marginal distributions that are very similar to the one

described by eq. 4.44. Fig. 4-10 shows these marginal distributions for various correlation coefficients--the differences are minimal for even a large change in the correlation. Thus either function could adequately model a random process such as the one characterized by fig. 4-11. This figure represents the measured distribution at the output of a DPCM coder having a sampled image as an input. This particular image has a measured average correlation of 0.4 between successive DPCM samples, resulting in the experimental two-dimensional distribution shown in fig. 4-12. Using r=0.4 in eq. 4.37, this distribution can be modeled as a two-dimensional laplacian density. This density is plotted in fig. 4-13.

Similarly, fig. 4-14 shows the marginal distributions of a two-dimensional Rayleigh density for several different correlations. The curves are plotted from

$$p(x)=\frac{1}{2s_x}\exp\left\{-\frac{x^2}{2s_x^2}\right\}\left\{\frac{x}{s_x}\exp\left[-\frac{x^2(\frac{a}{b}-r)^2}{2s_x^2(1-r^2)}\right](1-\sqrt{1-r^2})\right.$$

$$+\frac{x}{s_x}\exp\left[-\frac{x^2(\frac{b}{a}-r)^2}{2s_x^2(1-r^2)}\right](1+\sqrt{1-r^2})$$

$$\left.+r\sqrt{\frac{\pi}{2}}(\frac{x^2}{s_x^2}-1)\left[\mathrm{erf}\left[\frac{x(\frac{a}{b}-r)}{s_x\sqrt{2(1-r^2)}}\right]-\mathrm{erf}\left[\frac{x(\frac{b}{a}-r)}{s_x\sqrt{2(1-r^2)}}\right]\right]\right\}U(x) \quad (4.45)$$

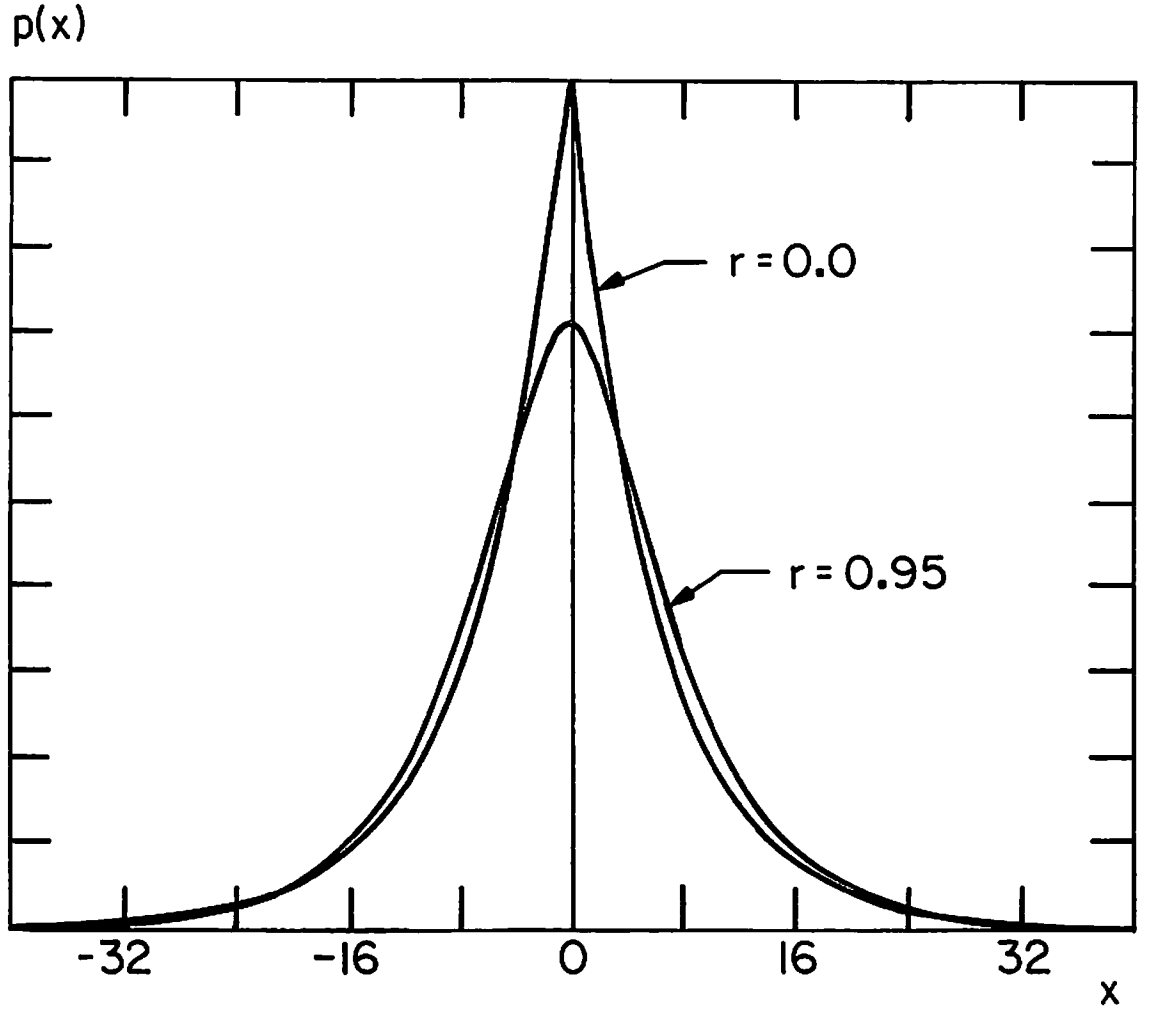


Figure 4-10. Marginal distributions of a correlated two-dimensional laplacian density function for various correlation coefficients (r).

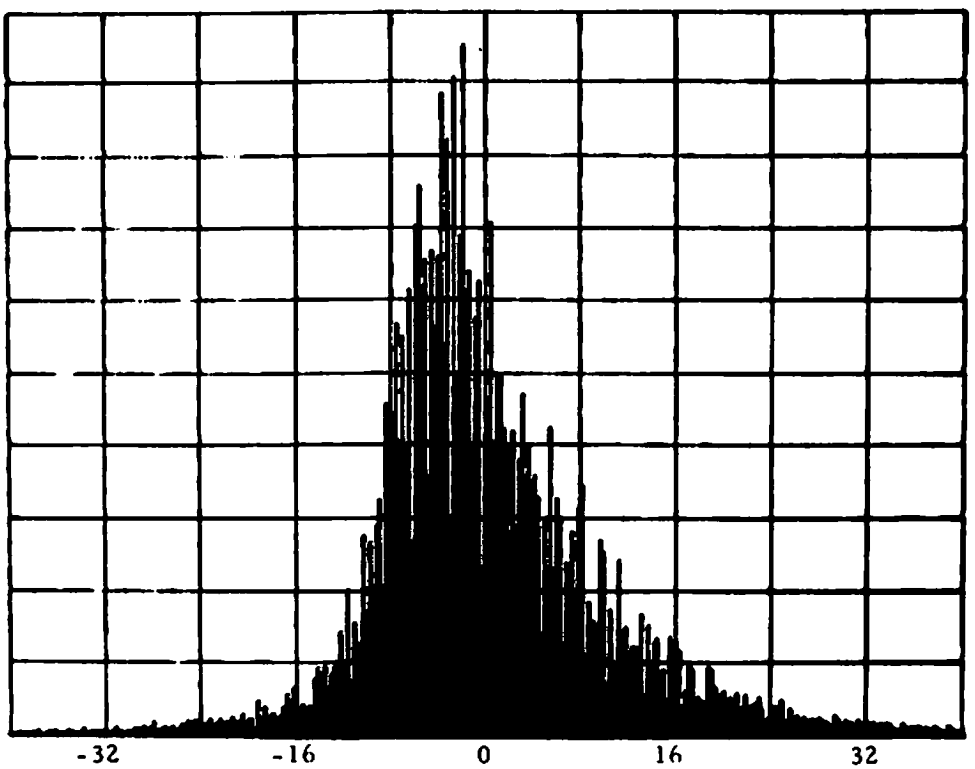


Figure 4-11. Histogram of the DPCM signal for the "girl" picture

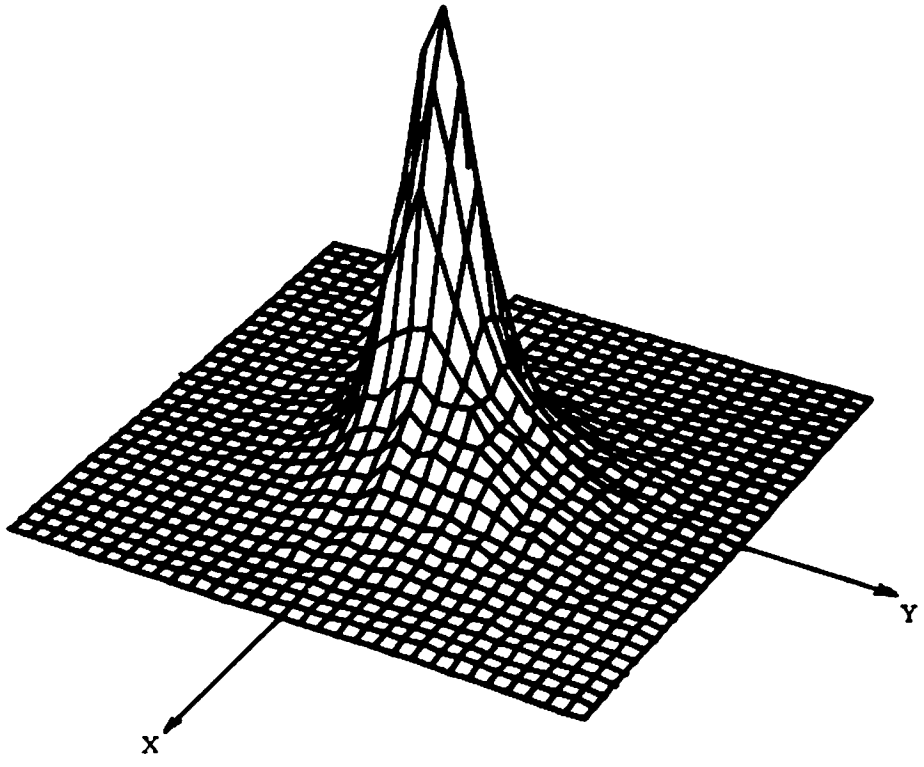


FIGURE 4-12. TWO-DIMENSIONAL HISTOGRAM OF THE DPCM CODED "GIRL" PICTURE

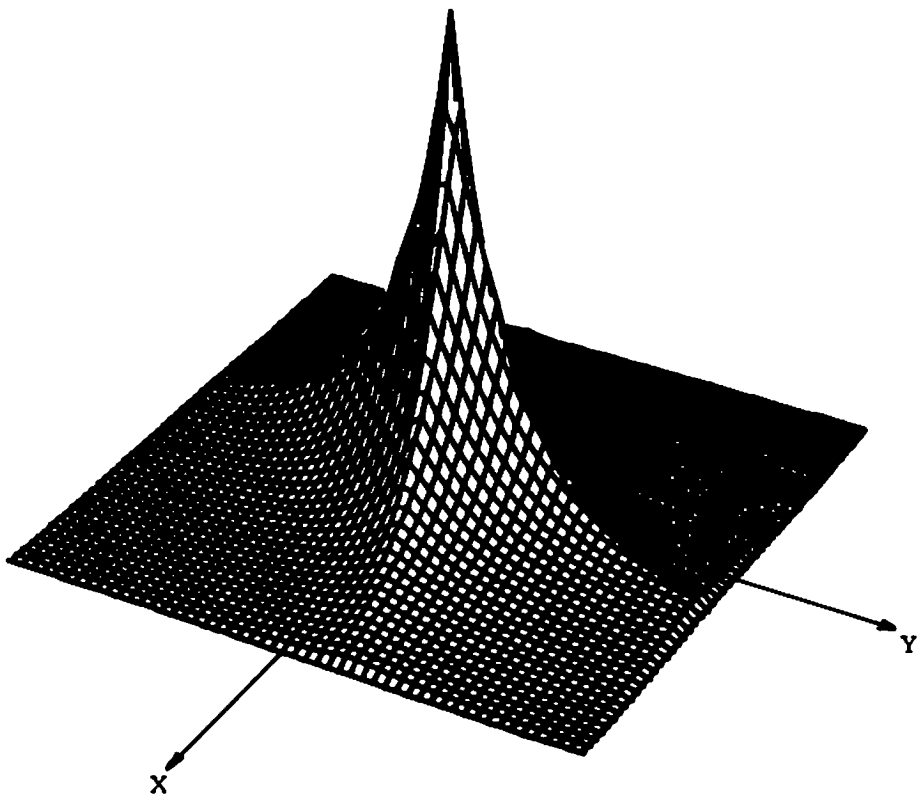


FIGURE 4-13. TWO-DIMENSIONAL LAPLACIAN DENSITY FUNCTION USED TO MODEL THE TWO-DIMENSIONAL DPCM SIGNAL SHOWN IN FIG. 4-12; CORRELATION =0.4

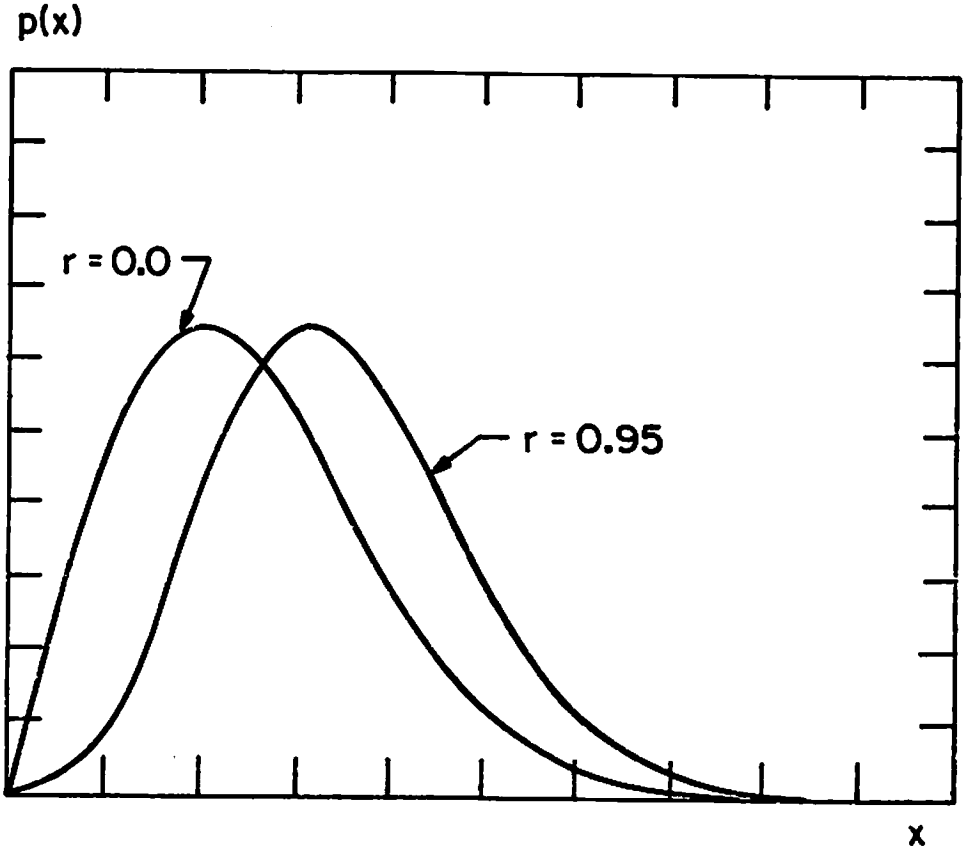


Figure 4-14. Marginal distributions of a correlated two-dimensional Rayleigh density function with correlation coefficient r.

which is found by integrating the two-dimensional Rayleigh density

$$p(x,y)=\frac{1}{4s_x s_y (1-r^2)^{3/2}}\left(\frac{ax}{s_x}-\frac{by}{s_y}\right)\left(\frac{ay}{s_y}-\frac{bx}{s_x}\right)U\left(\frac{ax}{s_x}-\frac{by}{s_y}\right)U\left(\frac{ay}{s_y}-\frac{bx}{s_x}\right)$$

$$\cdot\exp\left\{-\frac{1}{2(1-r^2)}\left(\frac{x^2}{s_x^2}-\frac{2rxy}{s_x s_y}+\frac{y^2}{s_y^2}\right)\right\} \qquad (4.46)$$

The parameters s_x, s_y, r, a, and b are the same as previously defined in equations 4.38 to 4.42. As expected, when the correlation between x and y becomes zero, the marginal density reduces to

$$p(x)\Big|_{r=0}=\frac{x}{s_x^2}\exp\left(-\frac{x^2}{2s_x^2}\right)U(x) \qquad (4.47)$$

the familiar one-dimensional Rayleigh density. However the figure shows that the shape of the marginal density function is insensitive to changes in the correlation. Hence a random variable approximately described by eq. 4.47 could be characterized equally well (though not as concisely) by eq. 4.45. Similar results are obtained for marginal distributions of the higher dimensional densities described by eq. 4.34. This equation therefore provides an effective model for a multidimensional correlated Rayleigh random process.

Thus, the technique developed above generates, from a given one-dimensional density, a multidimensional density that possesses a desired correlation function. Because the one-dimensional density function and correlation properties of a random process are usually measurable and known, the multidimensional density can be derived from this information and utilized to model the random process. This model can then be employed in the analysis and simulation of digital systems which operate on the random process. The potential applications for this modeling are widespread and several are analyzed in subsequent chapters.

REFERENCES

1. A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, Inc., New York, 1965, pp. 157-158.

2. Ibid., chap. 8.

3. S. O. Rice, "Mathematical Analysis of Random Noise," Bell System Technical Journal, vol. 24, 1945.

4. E. J. Gumbel, "Distributions à plusieurs variables, dont les marges sont données," Comptes Rendues de l'Academie des Sciences, Paris, vol. 246, 1958, pp. 2717-2720.

5. E. Parzen, Modern Probability Theory and Its Applications, Wiley, New York, 1960, p. 279.

6. P. Beckmann, Orthogonal Polynomials for Engineers and Physicists, Golem Press, Colorado, 1973, pp. 163-187.

CHAPTER 5

QUANTIZATION AND RESTORATION OF GAUSSIAN SAMPLES

In Chapter 3, an estimation equation was derived for the minimum mean-square error restoration of quantized samples from a random process. Solution of the estimation equation is dependent upon knowledge of both the structure of the quantizer and the multidimensional probability density function of the sampled random process. For many communication systems, the underlying random process can be described by a gaussian probability density function. A gaussian density arises naturally in many applications because it is the limiting distribution for all additive random processes, and the addition of random variables occurs often in communication systems. A gaussian random process is thus assumed to be the input to a quantizer and an optimal restoration is then derived based on this assumption.

5.1 Estimation of Quantized Gaussian Samples

Repeating one of the major results of Chapter 3, the restoration of a vector of samples $\underline{x}$, which has been quantized to one of M regions, R_m, can be found from

$$\underline{y}_m = \frac{\int_{R_m} \underline{x} p(\underline{x}) d\underline{x}}{\int_{R_m} p(\underline{x}) d\underline{x}} \tag{3.6}$$

for m=1,2,...,M, where $p(\underline{x})$ is the probability density function of the sampled random process. For a gaussian random process, $p(\underline{x})$ can be written

$$p(\underline{x}) = (2\pi)^{-N/2} |\underline{C}|^{-1/2} \exp(-\tfrac{1}{2}\underline{x}^T \underline{C}^{-1} \underline{x}) \tag{5.1}$$

where $\underline{C}$ is the covariance matrix of $\underline{x}$ and the mean is assumed to be zero. If the region R_m is rectangular, i.e., if each component of $\underline{x}$ is quantized individually, then the region can be expressed as

$$R_m = \{\underline{x} \mid x_i \in [a_{mi}, b_{mi}],\ i=1,2,\ldots,N\} \tag{5.2}$$

for m=1,2,...,M. For notational simplicity only a single region is henceforth considered and the subscript m is dropped, leaving

$$R = \{\underline{x} \mid \underline{x} \in [\underline{a}, \underline{b}]\} \tag{5.3}$$

Substituting eq. 5.1 into eq. 3.6 yields

$$\underline{y}=\frac{\int_{\underline{a}}^{\underline{b}}\underline{x}(2\pi)^{-N/2}|\underline{C}|^{-1/2}\exp(-\frac{1}{2}\underline{x}^T\underline{C}^{-1}\underline{x})d\underline{x}}{\int_{\underline{a}}^{\underline{b}}(2\pi)^{-N/2}|\underline{C}|^{-1/2}\exp(-\frac{1}{2}\underline{x}^T\underline{C}^{-1}\underline{x})d\underline{x}} \tag{5.4}$$

Unfortunately, no known analytical solution exists for either of the integrals in this equation. Curry [1] has obtained an approximate solution for finely quantized values of x_i, i.e., for

$$b_i-a_i < s_i \tag{5.5}$$

for i=1,2,...,N, where s_i is the standard deviation of the (i)th component of $\underline{x}$. His approach is to approximate the gaussian density by the first three terms of its Taylor series expansion about the midpoint of the region R. The integrations can then be performed, with the result that

$$E\{\underline{x} \mid \underline{x} \in R\}=(\underline{I}-\underline{D}\underline{C}^{-1})(\underline{b}+\underline{a})/2 \tag{5.6}$$

where the diagonal matrix

$$\underline{D}=\left\{\frac{(b_i-a_i)^2}{12}\delta_{ij}\right\} \tag{5.7}$$

for i,j=1,2,...,N.

An exact solution to eq. 5.4 can be obtained when the components of $\underline{x}$ are uncorrelated. In this case the covariance matrix of $\underline{x}$ can be expressed as

$$\underline{C}=\{s_i^2 \delta_{ij}\} \tag{5.8}$$

for i,j=1,2,...,N, so that the restoration equation can be rewritten as

$$\underline{y}=\frac{\int_{\underline{a}}^{\underline{b}} \underline{x} \exp\left(-\frac{1}{2}\sum_{i=1}^{N} x_i^2/s_i^2\right) d\underline{x}}{\int_{\underline{a}}^{\underline{b}} \exp\left(-\frac{1}{2}\sum_{i=1}^{N} x_i^2/s_i^2\right) d\underline{x}} \tag{5.9}$$

Each multidimensional integral can then be separated into a product of one-dimensional integrals, so that common factors can be cancelled, leaving

$$\underline{y}=\begin{bmatrix} \dfrac{\int_{a_1}^{b_1} x_1 e^{-x_1^2/2s_1^2} dx_1}{\int_{a_1}^{b_1} e^{-x_1^2/2s_1^2} dx_1} \\ \vdots \\ \dfrac{\int_{a_N}^{b_N} x_N e^{-x_N^2/2s_N^2} dx_N}{\int_{a_N}^{b_N} e^{-x_N^2/2s_N^2} dx_N} \end{bmatrix} \qquad (5.10)$$

Considering the (i)th component of this vector, since all of the components are identical in form, the integrals can be evaluated, as <*>

<*> In accordance with the usual convention,

$$\mathrm{erf}(x)=\frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2} dt$$

$$y_i = \sqrt{\frac{2}{\pi}}\, s_i \frac{\exp\left(-\frac{a_i^2}{2s_i^2}\right) - \exp\left(-\frac{b_i^2}{2s_i^2}\right)}{\operatorname{erf}\left(\frac{b_i}{\sqrt{2}s_i}\right) - \operatorname{erf}\left(\frac{a_i}{\sqrt{2}s_i}\right)} \tag{5.11}$$

A similar computation without the zero-mean assumption shows that, if $E\{\underline{x}\}=\underline{u}$, then

$$\underline{y}=\underline{u}+\sqrt{\frac{2}{\pi}}\begin{bmatrix} \dfrac{s_1\left\{\exp\left[-\dfrac{(a_1-u_1)^2}{2s_1^2}\right]-\exp\left[-\dfrac{(b_1-u_1)^2}{2s_1^2}\right]\right\}}{\operatorname{erf}\left(\dfrac{b_1-u_1}{\sqrt{2}s_1}\right)-\operatorname{erf}\left(\dfrac{a_1-u_1}{\sqrt{2}s_1}\right)} \\ \vdots \\ \dfrac{s_N\left\{\exp\left[-\dfrac{(a_N-u_N)^2}{2s_N^2}\right]-\exp\left[-\dfrac{(b_N-u_N)^2}{2s_N^2}\right]\right\}}{\operatorname{erf}\left(\dfrac{b_N-u_N}{\sqrt{2}s_N}\right)-\operatorname{erf}\left(\dfrac{a_N-u_N}{\sqrt{2}s_N}\right)} \end{bmatrix} \tag{5.12}$$

Gaussian variables which have been decorrelated by means of a Karhunen-Loeve transformation [2] and then quantized can be restored, according to a minimum mean-square error criterion, by utilizing these last two equations.

An exact analytical solution to eq. 5.4 also exists when an estimate of a single vector component, x_N, is desired based upon two types of information:

1. the other components, $x_1, x_2, \ldots, x_{N-1}$, which are known completely (quantized with an infinite number of bits)
2. the quantizer output, which nonlinearly specifies the interval containing x_N

To derive this solution, consider

$$\underline{y}=E\{\underline{x} \mid x_1=a_1, x_2=a_2, \ldots, x_{N-1}=a_{N-1}; a_N \leq x_N < b_N\} \qquad (5.13)$$

The vector $\underline{t}$ is defined such that

$$\underline{t}=\begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_{N-1} \\ x_N \end{bmatrix} \qquad (5.14)$$

Then, using eq. 5.4,

$$\underline{y}=\frac{\int_R \underline{t}\ \exp(-\frac{1}{2}\underline{t}^T\underline{C}^{-1}\underline{t})\,d\underline{x}}{\int_R \exp(-\frac{1}{2}\underline{t}^T\underline{C}^{-1}\underline{t})\,d\underline{x}}$$

or

$$\underline{y}=\frac{\int_{a_N}^{b_N} \underline{t}\ \exp(-\frac{1}{2}\underline{t}^T\underline{C}^{-1}\underline{t})\,dx_N}{\int_{a_N}^{b_N} \exp(-\frac{1}{2}\underline{t}^T\underline{C}^{-1}\underline{t})\,dx_N}$$

or

$$\underline{y}=\begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_{N-1} \\ \dfrac{\int_{a_N}^{b_N} x_N \exp(-\frac{1}{2}\underline{t}^T\underline{C}^{-1}\underline{t})\,dx_N}{\int_{a_N}^{b_N} \exp(-\frac{1}{2}\underline{t}^T\underline{C}^{-1}\underline{t})\,dx_N} \end{bmatrix} \qquad (5.15)$$

Denoting the elements of $\underline{C}^{-1}$ by

$$\underline{C}^{-1}=\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1N} \\ r_{21} & r_{22} & \cdots & r_{2N} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ r_{N1} & r_{N2} & \cdots & r_{NN} \end{pmatrix} \tag{5.16}$$

the exponential term in the Nth component of eq. 5.15 becomes

$$\exp\left\{-\frac{1}{2}\underline{t}\begin{bmatrix} \sum_{j=1}^{N-1} r_{1j}a_j+r_{1N}x_N \\ \sum_{j=1}^{N-1} r_{2j}a_j+r_{2N}x_N \\ \cdot \\ \cdot \\ \cdot \\ \sum_{j=1}^{N-1} r_{Nj}a_j+r_{NN}x_N \end{bmatrix}\right\}$$

$$=\exp\left\{-\frac{1}{2}\left[a_1\left(\sum_{j=1}^{N-1} r_{1j}a_j+r_{1N}x_N\right)+a_2\left(\sum_{j=1}^{N-1} r_{2j}a_j+r_{2N}x_N\right)\right.\right.$$

$$\left.\left.+\ldots+a_{N-1}\left(\sum_{j=1}^{N-1} r_{N-1,j}a_j+r_{N-1,N}x_N\right)+x_N\left(\sum_{j=1}^{N-1} r_{Nj}a_j+r_{NN}x_N\right)\right]\right\}$$

$$=\exp\left\{-\frac{1}{2}\left[a_1\sum_{j=1}^{N-1}r_{1j}a_j+a_2\sum_{j=1}^{N-1}r_{2j}a_j+\ldots+a_{N-1}\sum_{j=1}^{N-1}r_{N-1,j}a_j\right]\right\}$$

$$\cdot\exp\left\{-\frac{1}{2}\left[x_N\sum_{j=1}^{N-1}a_j(r_{jN}+r_{Nj})+r_{NN}x_N^2\right]\right\} \qquad (5.17)$$

The first exponential in eq. 5.17 is a constant and is common to both the numerator and denominator of the Nth component of $\underline{y}$ in eq. 5.15. Hence this factor can be cancelled, leaving

$$y_N=\frac{\int_{a_N}^{b_N}x_N\exp\left\{-\frac{1}{2}\left[x_N\sum_{j=1}^{N-1}a_j(r_{jN}+r_{Nj})+r_{NN}x_N^2\right]\right\}dx_N}{\int_{a_N}^{b_N}\exp\left\{-\frac{1}{2}\left[x_N\sum_{j=1}^{N-1}a_j(r_{jN}+r_{Nj})+r_{NN}x_N^2\right]\right\}dx_N} \qquad (5.18)$$

Completing the square within the exponential, each exponential becomes

$$\exp\left\{-\frac{r_{NN}}{2}\left[x_N^2+\frac{x_N}{r_{NN}}\sum_{j=1}^{N-1}a_j(r_{jN}+r_{Nj})+\left(\frac{1}{2r_{NN}}\sum_{j=1}^{N-1}a_j(r_{jN}+r_{Nj})\right)^2\right.\right.$$
$$\left.\left.-\left(\frac{1}{2r_{NN}}\sum_{j=1}^{N-1}a_j(r_{jN}+r_{Nj})\right)^2\right]\right\}$$

$$=\exp\left\{-\frac{r_{NN}}{2}\left[x_N+\frac{1}{2r_{NN}}\sum_{j=1}^{N-1}a_j(r_{jN}+r_{Nj})\right]^2\right\}$$
$$\cdot\exp\left\{+\frac{r_{NN}}{2}\left[\frac{1}{2r_{NN}}\sum_{j=1}^{N-1}a_j(r_{jN}+r_{Nj})\right]^2\right\} \qquad (5.19)$$

Again, the last exponential is constant and is common to both the numerator and denominator. Hence it can be cancelled. Next, performing the substitutions

$$u=x_N+\frac{1}{2r_{NN}}\sum_{j=1}^{N-1}a_j(r_{jN}+r_{Nj}) = x_N+v \qquad (5.20a)$$

$$du=dx_N \qquad (5.20b)$$

then

$$y_N = \frac{\int_{a_N+v}^{b_N+v} (u-v)\exp(-\frac{r_{NN}u^2}{2})\,du}{\int_{a_N+v}^{b_N+v} \exp(-\frac{r_{NN}u^2}{2})\,du}$$

$$= \frac{\int_{a_N+v}^{b_N+v} u\exp(-\frac{r_{NN}u^2}{2})\,du}{\int_{a_N+v}^{b_N+v} \exp(-\frac{r_{NN}u^2}{2})\,du} - v$$

$$= \frac{-\frac{1}{r_{NN}}\left\{\exp\left[-\frac{r_{NN}}{2}(b_N+v)^2\right]-\exp\left[-\frac{r_{NN}}{2}(a_N+v)^2\right]\right\}}{\sqrt{\frac{\pi}{2r_{NN}}}\left\{\mathrm{erf}\left[\sqrt{\frac{r_{NN}}{2}}(b_N+v)\right]-\mathrm{erf}\left[\sqrt{\frac{r_{NN}}{2}}(a_N+v)\right]\right\}} - v \qquad (5.21)$$

Since the covariance matrix $\underline{C}$ is symmetric, then $\underline{C}^{-1}$ is also symmetric. Hence

$$r_{jN} = r_{Nj} \qquad (5.22)$$

and therefore

$$\underline{y}=\begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_{N-1} \\ \sqrt{\frac{2}{\pi r_{NN}}}\,\frac{\exp\left[-\frac{r_{NN}}{2}(b_N+w)^2\right]-\exp\left[-\frac{r_{NN}}{2}(a_N+w)^2\right]}{\operatorname{erf}\left[\sqrt{\frac{r_{NN}}{2}}(b_N+w)\right]-\operatorname{erf}\left[\sqrt{\frac{r_{NN}}{2}}(a_N+w)\right]}-w \end{bmatrix} \tag{5.23}$$

where

$$w=\frac{1}{r_{NN}}\sum_{j=1}^{N-1}a_j r_{Nj} \tag{5.24}$$

If x_N is quantized to an infinite number of bits ($a_N=b_N$), then

$$y_N=a_N \tag{5.25}$$

as is expected. If x_N is quantized to zero bits, its quantization interval is the real line ($-a_N=b_N=\infty$), and then its estimate, y_N, is

$$y=\frac{1}{r_{NN}}\sum_{j=1}^{N-1}a_j r_{Nj} \tag{5.26}$$

This result is identical to one obtained by Pratt [3] in estimating an unknown spectral value based on known spectral components. (The equivalence between these results is shown in the next section.) However, eq. 5.23 is a more general result in that it can be utilized to estimate components that have been quantized to any number of bits by an arbitrary quantization scheme.

By comparing the Nth component, y_N, of the zero-mean version of eq. 5.12 (the solution for uncorrelated variables) with y_N found from eq. 5.23 (the solution for correlated variables), it can be seen that the correlation effects are embodied in the variable, w. In fact, y_N from eq. 5.23 reduces to y_N from eq. 5.12 when w=0, which in turn occurs when the correlation is zero. The variable w is simply the negative of the estimate of x_N based on both the correlation between all of the variables and values of the previous N-1 variables. In other words, -w is the estimate of x_N that is obtained when no direct knowledge of x_N is available.

5.2 Linear Spectrum Extrapolation

In some transform coding systems, transmission bandwidth is saved by truncating certain spectral components, notably the high frequency ones, to zero. (This may be alternately expressed by saying that these

components are transmitted with zero bits.) The low frequency components are retained because they contain more energy than the high frequency components. If the low and high frequency components are correlated, then an estimate of the high frequency components can be obtained by utilizing this correlation. Since higher frequencies are being estimated, the estimation technique is equivalent to a spectrum extrapolation. If the transform components are also samples from a gaussian random process, they can be estimated by a Wiener extrapolation technique developed by Pratt [3] who has derived the following linear extrapolation operator

$$\underline{W}=\underline{C}\underline{S}^{T}(\underline{S}\underline{C}\underline{S}^{T})^{-1} \tag{5.27}$$

based on a minimum mean-square error criterion, in which $\underline{C}$ is the NxN covariance matrix of the samples and the sampling matrix $\underline{S}$ has the form

$$\underline{S}=\left(\underline{I}_{N-k} \;\vdots\; \underline{\emptyset}_{k}\right) \tag{5.28}$$

where $\underline{I}_{N-k}$ is the (N-k)x(N-k) identity matrix, and $\underline{\emptyset}_{k}$ represents k columns of zeros. The operator provides an estimate of k truncated samples based on N-k known samples.

Equation 5.26, developed in the last section, provides an estimate, y_N, of a single vector component based on the known samples $a_1, a_2, \ldots, a_{N-1}$. Note that this also is a linear estimation of a sample from a gaussian random process. The estimate obtained by this method is identical to that obtained from eq. 5.27 for the special case k=1.

To demonstrate this equivalence, consider the following partition of the symmetric covariance matrix

$$\underline{C}=\left(\begin{array}{c|c} \underline{D} & \underline{E} \\ \hline \underline{E}^T & \underline{F} \end{array}\right) \tag{5.29}$$

Note that $\underline{D}$ and $\underline{F}$ are also symmetric matrices. For the present case, let $\underline{D}$ be (N-1)x(N-1) so $\underline{E}$ is Nx1 and $\underline{F}$ is 1x1. Now the inverse of $\underline{C}$ can be found in terms of these sub-matrices from [4]

$$\underline{C}^{-1}=\left(\begin{array}{c|c} \underline{D}^{-1}+\underline{D}^{-1}\underline{E}(\underline{F}-\underline{E}^T\underline{D}^{-1}\underline{E})^{-1}\underline{E}^T\underline{D}^{-1} & -\underline{D}^{-1}\underline{E}(\underline{F}-\underline{E}^T\underline{D}^{-1}\underline{E})^{-1} \\ \hline -(\underline{F}-\underline{E}^T\underline{D}^{-1}\underline{E})^{1}\underline{E}^T\underline{D}^{-1} & (\underline{F}-\underline{E}^T\underline{D}^{-1}\underline{E})^{-1} \end{array}\right) \tag{5.30}$$

In the notation of the previous section,

$$r_{NN}=(\underline{F}-\underline{E}^T\underline{D}^{-1}\underline{E})^{-1} \tag{5.31}$$

Next define the (N-1) vector $\underline{A}$ such that

$$\underline{A}=(a_1\ a_2\ \cdot\ \cdot\ \cdot\ a_{N-1})^T \tag{5.32}$$

Equation 5.26 can now be rewritten as

$$y_N = \frac{-1}{(\underline{F}-\underline{E}^T\underline{D}^{-1}\underline{E})^{-1}}[-(\underline{F}-\underline{E}^T\underline{D}^{-1}\underline{E})^{-1}\underline{E}^T\underline{D}^{-1}]\underline{A} = \underline{E}^T\underline{D}^{-1}\underline{A} \tag{5.33}$$

Using eq. 5.28 with k=1, it can be seen that

$$\underline{D}=\underline{S}\underline{C}\underline{S}^T \tag{5.34}$$

Next let $\underline{W}_N^T$ be an (N-1) vector representing the Nth row of $\underline{W}$, which was previously defined in eq. 5.27. Then the estimate, y_N, using this vector is

$$y_N = \underline{W}_N^T\underline{A} = (\underline{C}\underline{S}^T)_N(\underline{S}\underline{C}\underline{S}^T)^{-1}\underline{A} = \underline{E}^T\underline{D}^{-1}\underline{A} \tag{5.35}$$

which is identical to eq. 5.33. Therefore eq. 5.26 can provide a linear, minimum mean-square error, spectrum extrapolation.

5.3 Covariance of Gaussian Estimator

The conditional covariance of the estimate found in eq. 5.23 can be used as a measure of the performance of the

estimate. This covariance can be found by utilizing eq. 3.10 (without the subscript m) as follows:

$$\mathrm{cov}(\underline{x} \mid \underline{x} \in R) = \frac{\int_{a_N}^{b_N} \underline{t}\underline{t}^T p(\underline{t})\, dx_N}{\int_{a_N}^{b_N} p(\underline{t})\, dx_N} - \begin{bmatrix} a_1 \\ \vdots \\ a_{N-1} \\ y_N \end{bmatrix} [a_1 \cdots a_{N-1} \; y_N]$$

$$= \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \dfrac{\int_{a_N}^{b_N} x_N p(\underline{t})\, dx_N}{\int_{a_N}^{b_N} p(\underline{t})\, dx_N} - y_N^2 \\ 0 & 0 & \cdots & \end{pmatrix} \qquad (5.36)$$

Next, denoting the (N,N)th component of this matrix as e_{NN} and performing the integrations reveals

$$e_{NN} = -\sqrt{\frac{2}{r_{NN}}} \; \frac{(b_N+w)\exp\left[-\frac{r_{NN}}{2}(b_N+w)^2\right] - (a_N+w)\exp\left[-\frac{r_{NN}}{2}(a_N+w)^2\right]}{\mathrm{erf}\left[\sqrt{\frac{r_{NN}}{2}}(b_N+w)\right] - \mathrm{erf}\left[\sqrt{\frac{r_{NN}}{2}}(a_N+w)\right]}$$

$$+ \frac{1}{r_{NN}} - (y_N+w)^2 \qquad (5.37)$$

with r_{NN} and w as defined in equations 5.16 and 5.24, respectively. The quantity e_{NN} represents the average squared error that results when x_N is quantized to the region $[a_N,b_N]$ and an estimate of x_N is obtained based on:

1. the quantization region $[a_N,b_N]$
2. the correlation matrix, $\underline{C}$, of the variables
3. the previous N-1 variables

The information from these last two factors is contained in w, which is itself a random variable. To compare e_{NN} to the error that would have resulted had the correlation information been unavailable, it is necessary to remove the conditioning on w. This can be done by averaging e_{NN} over all values of w. Because w is a weighted sum of gaussian random variables ($a_1=x_1$ to $a_{N-1}=x_{N-1}$), w has a gaussian distribution. The parameters of this distribution are determined next.

The mean value of w can be found from

$$E\{w\} = E\{\frac{1}{r_{NN}} \sum_{j=1}^{N} r_{Nj} a_j\} = \frac{1}{r_{NN}} \sum_{j=1}^{N} r_{Nj} E\{a_j\} = 0 \qquad (5.38)$$

The variance of w is then

$$E\{w^2\} = E\{(\frac{1}{r_{NN}} \sum_{j=1}^{N} r_{Nj} a_j)^2\} = \frac{1}{r_{NN}^2} E\{(\underline{r}^T \underline{A})^2\} \qquad (5.39)$$

where

$$\underline{r}=\begin{bmatrix} r_{N1} \\ r_{N2} \\ \vdots \\ r_{N,N-1} \end{bmatrix} \tag{5.40}$$

and

$$\underline{A}=\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{N-1} \end{bmatrix} \tag{5.41}$$

Hence

$$E\{w^2\} = \frac{1}{r_{NN}^2}\underline{r}^T E\{\underline{A}\underline{A}^T\}\underline{r} = \frac{1}{r_{NN}^2}\underline{r}^T \underline{C}_{NN}\underline{r} \tag{5.42}$$

where $\underline{C}_{NN}$ is the submatrix of $\underline{C}$ formed by deleting the Nth row and the Nth column of $\underline{C}$. In the notation of Sec. 5.2, however, $\underline{C}_{NN}=\underline{D}$. Also,

$$\underline{r}=-\underline{D}^{-1}\underline{E}(\underline{F}-\underline{E}^T\underline{D}^{-1}\underline{E})^{-1} \tag{5.43}$$

and

$$r_{NN}=(\underline{F}-\underline{E}^T\underline{D}^{-1}\underline{E})^{-1} \tag{5.44}$$

Then

$$E\{w^2\}=\underline{E}^T\underline{D}^{-1}\underline{E} \tag{5.45}$$

so that

$$p(w)=\frac{1}{\sqrt{2\pi(\underline{E}^T\underline{D}^{-1}\underline{E})}}\exp\{-\tfrac{1}{2}w^2/(\underline{E}^T\underline{D}^{-1}\underline{E})\} \tag{5.46}$$

The average error that results from restoring to y_N a sample that has been quantized to the interval $[a_N, b_N]$ is thus

$$e_y=\int_{-\infty}^{\infty} e_{NN}\,p(w)\,dw \tag{5.47}$$

Unfortunately, a closed form solution for this integral does not exist when the expression for e_{NN} has the form given by eq. 5.37. Equation 5.47 therefore must be evaluated numerically.

The results from one such numerical evaluation are shown in table 5-1. The integrations, needed to calculate the errors listed in this table, have been performed by means of a 16-point gaussian-Hermite quadrature formula [5]. The quantization regions are the same as those calculated by Max for a unit variance, gaussian

TABLE 5-1

NORMALIZED MEAN-SQUARE RESTORATION ERROR FOR
A MAX-QUANTIZED, GAUSSIAN-MARKOV PROCESS
CORRELATION = 0.9
FILTERING CASE

Bits	Quantization Interval	Probability of Occurrence	Normalized Mean-Square Restoration Error	
			Not Using Correlation	Using Correlation
0	($-\infty$,+ ∞)	1.00000	1.000000	0.190000
1	(0.000,+ ∞)	0.50000	0.363344	0.085441
2	(0.000,0.982)	0.33685	0.076896	0.042723
	(0.982,+ ∞)	0.16315	0.201243	0.034435
3	(0.000,0.501)	0.19165	0.020687	0.016412
	(0.501,1.050)	0.16148	0.024683	0.017257
	(1.050,1.748)	0.10662	0.038160	0.018193
	(1.748,+ ∞)	0.04023	0.130557	0.014966
4	(0.000,0.258)	0.10188	0.005543	0.005140
	(0.258,0.522)	0.09742	0.005799	0.005309
	(0.522,0.800)	0.08871	0.006373	0.005667
	(0.800,1.099)	0.07616	0.007434	0.006256
	(1.099,1.437)	0.06047	0.009386	0.007122
	(1.437,1.844)	0.04271	0.013391	0.008260
	(1.844,2.401)	0.02445	0.023947	0.009205
	(2.401,+ ∞)	0.00818	0.093356	0.007747
5	(0.000,0.132)	0.05251	0.001451	0.001420
	(0.132,0.265)	0.05187	0.001466	0.001434
	(0.265,0.399)	0.05067	0.001502	0.001466
	(0.399,0.536)	0.04890	0.001558	0.001515
	(0.536,0.676)	0.04653	0.001636	0.001584
	(0.676,0.821)	0.04365	0.001746	0.001678
	(0.821,0.972)	0.04026	0.001893	0.001795
	(0.972,1.130)	0.03642	0.002091	0.001958
	(1.130,1.299)	0.03222	0.002367	0.002175
	(1.299,1.481)	0.02769	0.002754	0.002460
	(1.481,1.682)	0.02295	0.003325	0.002841
	(1.682,1.908)	0.01812	0.004225	0.003328
	(1.908,2.173)	0.01332	0.005764	0.004036
	(2.173,2.504)	0.00875	0.008842	0.004847
	(2.504,2.976)	0.00467	0.016968	0.005638
	(2.976,+ ∞)	0.00146	0.071316	0.004866

distribution [6]. For a given number of regions, or, equivalently, for a given number of bits, this choice of regions results in the smallest mean-square quantization error. This error, in fact, is the one that occurs when the restoration corresponding to eq. 5.23 is performed and the correlations are either zero or simply ignored. However, when successive samples of the distribution are correlated and this correlation is utilized, the restoration can be improved and the mean-square error decreased, as the last column of the table indicates. The decrease in mean-square error is seen to be most significant for coarse quantization at a low number of bits, but an improvement is evident for every quantizing region.

The specific correlation in this example is assumed to be due to a first-order Markov process with a correlation coefficient of Ø.9. The choice of a Markov process is a convenient one because the restoration of a random sample from this process depends only on the sample immediately preceding <*>. The error results are thus independent of the number of samples, N, used for the restoration, and independent of the size of the correlation matrix, $\underline{C}$. When only prior samples are utilized in the restoration of a

---

<*> By the definition of a Markov process,

$$p(x_N \mid x_1, x_2, \ldots, x_{N-1}) = p(x_N \mid x_{N-1})$$

random sample, the method is denoted as "filtering."

Table 5-2 contains similar error results to table 5-1, but for the technique of "smoothing." Smoothing utilizes information about all past and all future samples to restore a current sample. However, in the case of a Markov random process, only the sample immediately preceding and the sample immediately following the current sample are needed <**>. The use of this added information from future samples significantly reduces the resultant mean-square restoration error for any choice of quantizing region.

The results of tables 5-1 and 5-2 are presented in a different format in figures 5-1 and 5-2 to illustrate the total mean-square quantization errors that occur at various quantizing bit assignments. The errors are calculated by using the information from the tables in the following equation

$$\mathcal{E}=\sum_{m=1}^{M} e_{ym} \Pr\{x_N \in [a_N,b_N]_m\} \qquad (5.48)$$

where $[a_N,b_N]_m$ is one of the M regions to which x_N is quantized, and e_{ym} is the error, found from eq. 5.47, for

---

<**> For a Markov process it is also true that

$$p(x_N \mid x_1,\ldots,x_{N-1},x_{N+1},\ldots)=p(x_N \mid x_{N-1},x_{N+1})$$

TABLE 5-2

NORMALIZED MEAN-SQUARE RESTORATION ERROR FOR
A MAX-QUANTIZED, GAUSSIAN-MARKOV PROCESS
CORRELATION = 0.9
SMOOTHING CASE

Bits	Quantization Interval	Probability of Occurrence	Normalized Mean-Square Restoration Error	
			Not Using Correlation	Using Correlation
0	($-\infty$,+ ∞)	1.00000	1.000000	0.105000
1	(0.000,+ ∞)	0.50000	0.363344	0.049106
2	(0.000,0.982)	0.33685	0.076896	0.028041
	(0.982,+ ∞)	0.16315	0.201243	0.018555
3	(0.000,0.501)	0.19165	0.020687	0.012428
	(0.501,1.050)	0.16148	0.024683	0.012115
	(1.050,1.748)	0.10662	0.038160	0.010920
	(1.748,+ ∞)	0.04023	0.130557	0.007810
4	(0.000,0.258)	0.10188	0.005543	0.004501
	(0.258,0.522)	0.09742	0.005799	0.004577
	(0.522,0.800)	0.08871	0.006373	0.004720
	(0.800,1.099)	0.07616	0.007434	0.004908
	(1.099,1.437)	0.06047	0.009386	0.005144
	(1.437,1.844)	0.04271	0.013391	0.005160
	(1.844,2.401)	0.02445	0.023947	0.005431
	(2.401,+ ∞)	0.00818	0.093356	0.003924
5	(0.000,0.132)	0.05251	0.001451	0.001356
	(0.132,0.265)	0.05187	0.001466	0.001299
	(0.265,0.399)	0.05067	0.001502	0.001333
	(0.399,0.536)	0.04890	0.001558	0.001383
	(0.536,0.676)	0.04653	0.001636	0.001444
	(0.676,0.821)	0.04365	0.001746	0.001523
	(0.821,0.972)	0.04026	0.001893	0.001638
	(0.972,1.130)	0.03642	0.002091	0.001436
	(1.130,1.299)	0.03222	0.002367	0.001719
	(1.299,1.481)	0.02769	0.002754	0.001983
	(1.481,1.682)	0.02295	0.003325	0.002267
	(1.682,1.908)	0.01812	0.004225	0.002996
	(1.908,2.173)	0.01332	0.005764	0.003345
	(2.173,2.504)	0.00875	0.008842	0.003428
	(2.504,2.976)	0.00467	0.016968	0.004066
	(2.976,+ ∞)	0.00146	0.071316	0.003735

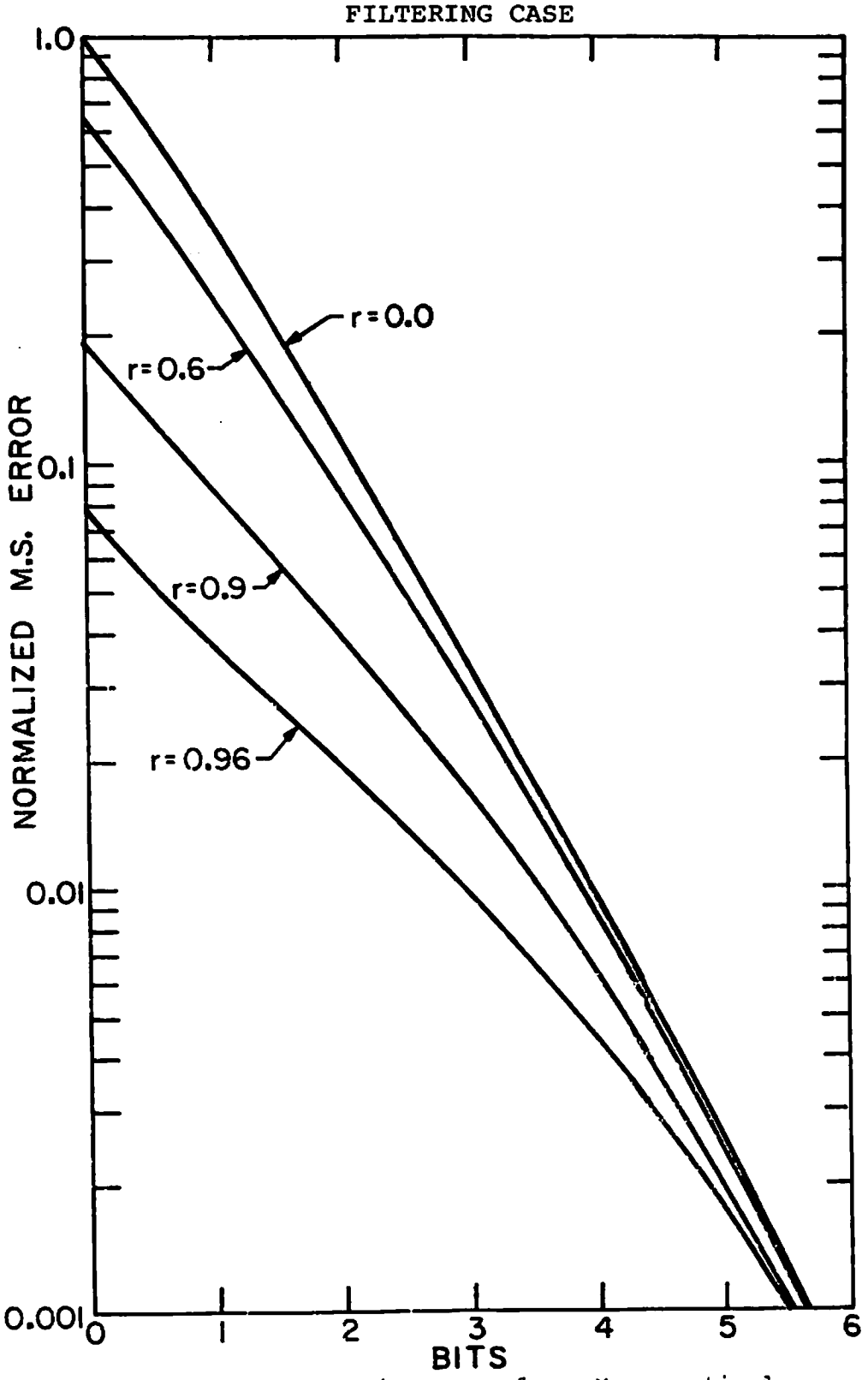


Figure 5-1. Mean-square restoration error for a Max quantized, gaussian-Markov process with correlation factor r.

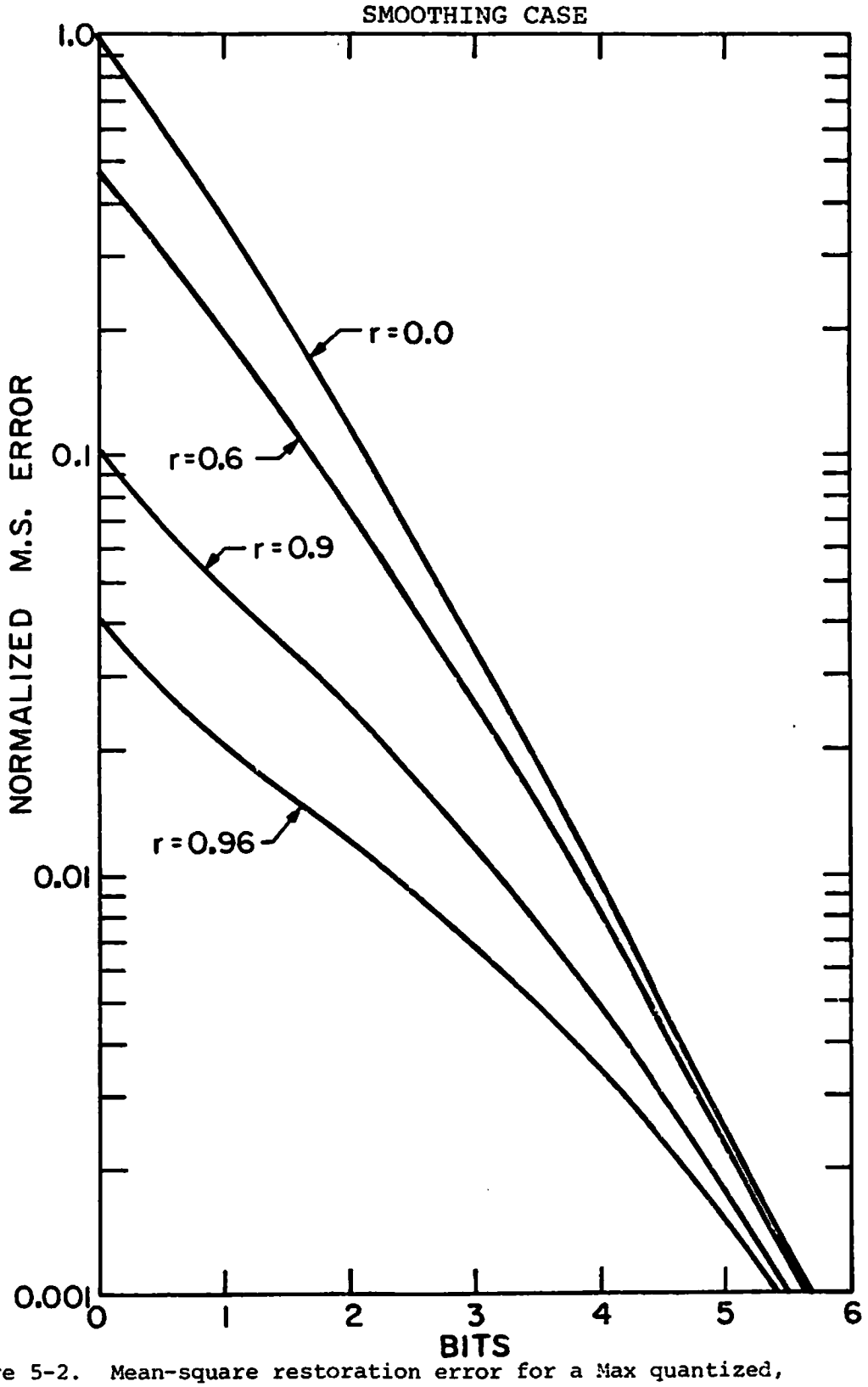


Figure 5-2. Mean-square restoration error for a Max quantized, gaussian-Markov process with correlation factor r.

this quantization region. The errors reveal the significant error reductions that can be obtained by utilizing available correlation information in an optimal fashion. For example, when samples of a random process have an average correlation of 0.96 and are quantized to one bit, utilizing this correlation in an optimum restoration yields a mean-square error equivalent to that from three bit Max quantization which does not utilize the correlation. Figure 5-1 shows the filtering case and fig. 5-2 the smoothing case.

5.4 Simulation Results for Gaussian Processes

In Sec. 5.1 a solution is described for achieving an optimal nonlinear restoration of quantized gaussian data. This solution has now been applied to the restoration of quantized one-dimensional random signals and two-dimensional transform domain zonal-quantized images. The results reveal a decrease in mean-square error in all cases. However, in spite of the error reduction, some images exhibit a degradation in subjective quality after being restored. Hence a nonlinear error criterion based on the human visual system has been used in place of the mean-square error function. Under this criterion a subjective image improvement, as well as a numerical error reduction, are obtained.

5.4.1 One-dimensional Markov Random Process Simulation

To demonstrate the utility of the restoration procedure, a randomly generated gaussian Markov signal has been quantized and restored. A signal having unit variance and a Markov correlation of Ø.95 is first generated. A two bit per sample Max quantization scheme is then employed to obtain the quantized approximation to this original signal. The quantized signal and the statistical knowledge about the original signal are used as inputs to the nonlinear estimator in eq. 5.23. The restoration is performed in block-lengths of 16, i.e., N=16 in the notation of this equation. However, the inputs to the equation are 16 quantization intervals rather than 15 known samples and one quantization interval as the estimator requires. Hence, to satisfy the conditions of the estimator, a point is chosen within each of the 15 intervals as an initial guess, and an estimate of the sixteenth sample is obtained based on this guess. This procedure is then repeated for each quantized sample in turn, with the guesses successively replaced by their calculated estimates. The method is thus recursive in nature. To facilitate the convergence of these estimates to their optimum values, the initial guesses are chosen to be the Max restoration points (when the sample correlation is ignored, these are the minimum mean-square error restoration points). It has been experimentally determined that one iteration of this procedure is

sufficient for convergence, and that more than one iteration provides negligible improvement.

A typical portion of the results from the above simulation is shown in fig. 5-3. The restoration decreases the mean-square quantization error by 33%, and is seen to provide a better approximation to the original signal than the quantized signal. The average improvement in mean-square error as a function of quantizing bit assignment for different correlation coefficients is shown in fig. 5-4. It can be seen from this graph that, as the amount of correlation in the Markov process approaches zero, the restoration provides no error improvement. There is no improvement as the number of quantizing bits becomes large and the differences between the original signal and the quantized signal vanish. There is also no improvement at zero bits when there is no information remaining in the quantized signal upon which to base a restoration. Thus, the above procedure represents a viable restoration technique only when the number of quantizing bits are small and the input samples to the quantizer are correlated.

5.4.2 Block Transform Zonal Image Coding

The conditions which were placed on the estimator derived in Sec. 5.1, and which were modified experimentally in the preceding paragraphs, are satisfied by the zonal

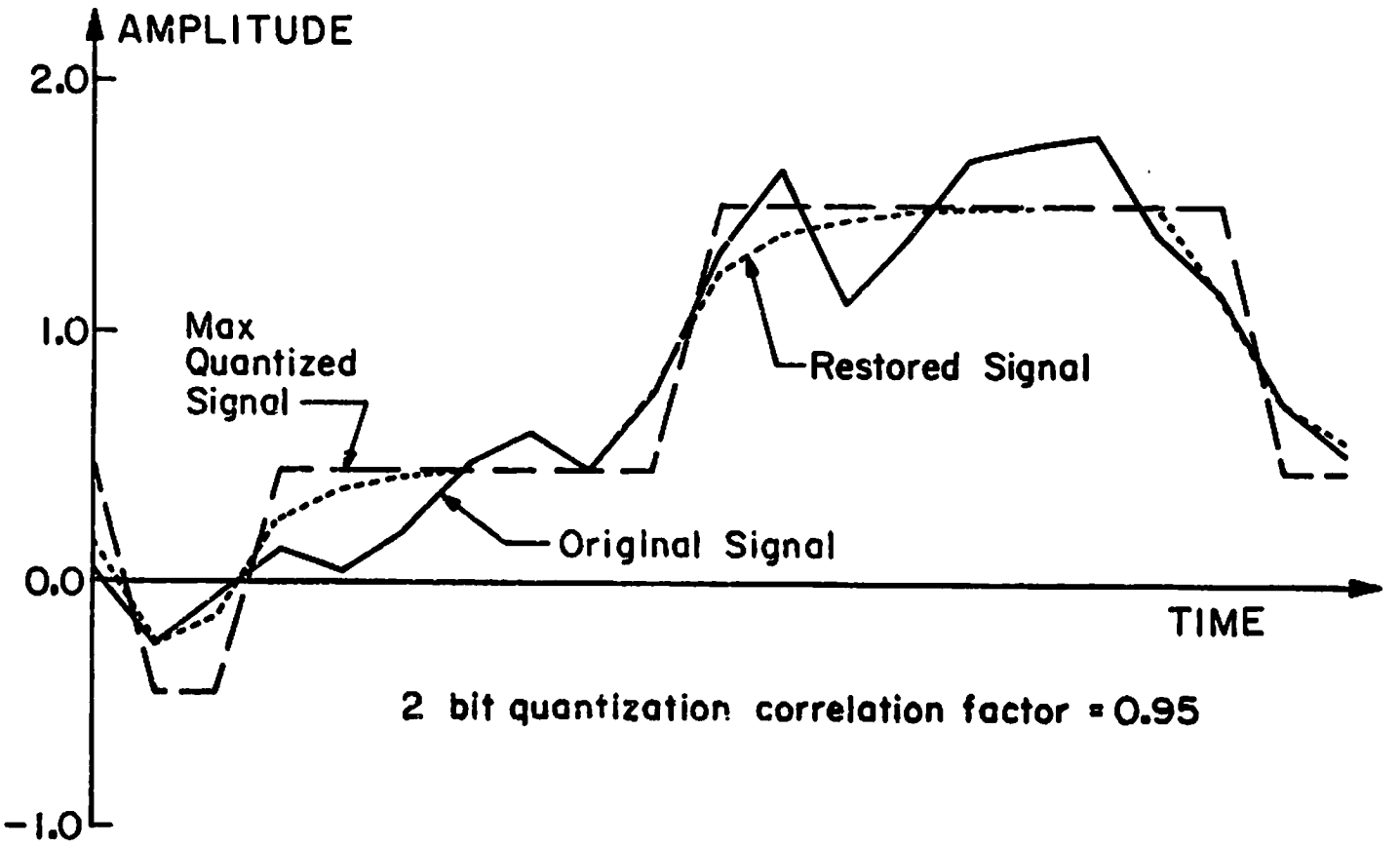


Figure 5-3. Restoration of a quantized gaussian-Markov signal.

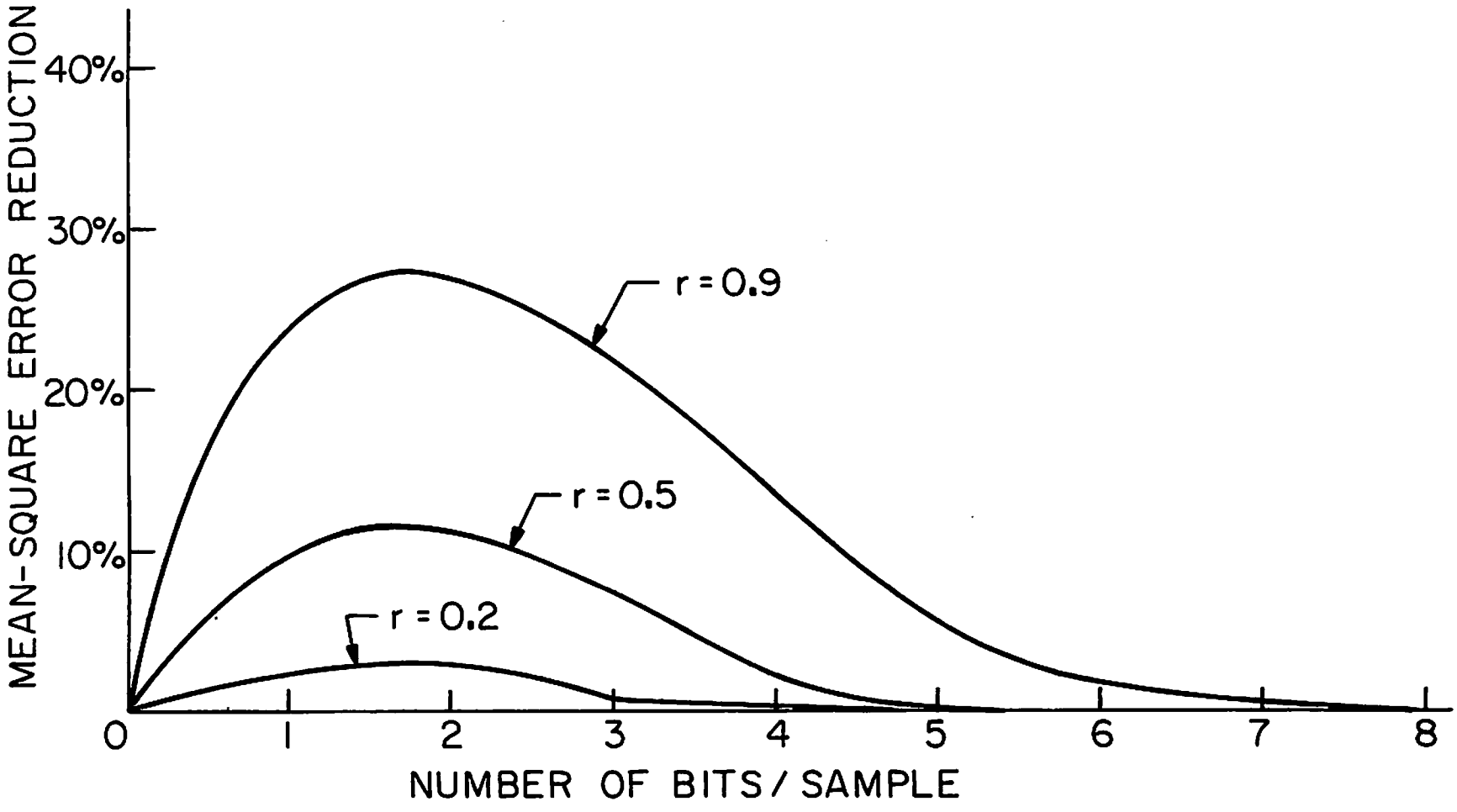


Figure 5-4. Mean-square error improvement for quantized gaussian-Markov signal with correlation coefficient r; samples restored in blocks of 16.

transform coding technique for images. First, transform samples typically have a gaussian distribution: each transform sample is the sum of a large number of random variables so that the central limit theorem can be invoked. Next, for all transforms except the Karhunen-Loeve, the transform samples are correlated. Finally, to achieve a bandwidth compression or a bit-rate reduction, some of the transform samples are quantized to a small number of bits. Hence, because all of the necessary conditions for a reconstruction are satisfied, it is possible and practical to restore zonal transform coded images.

The image that is presented to a transform coder is assumed to be in the form of a two-dimensional array of light intensities. (Each point of this array is known as a picture element or "pixel".) If the image is presented as a continuous field of intensities, however, it must first be sampled to obtain the image pixels. A two-dimensional discrete mathematical transform is then taken of these pixels. The transform is performed over the entire image or over subsections of the image known as blocks. The transform domain samples are next quantized and coded, either for storage or for transmission over a channel as fig. 5-5 shows. At the receiver the samples are decoded, restored to reduce the quantization effects, and inverse transformed to reconstruct the original image. The subsequent paragraphs discuss these operations in further

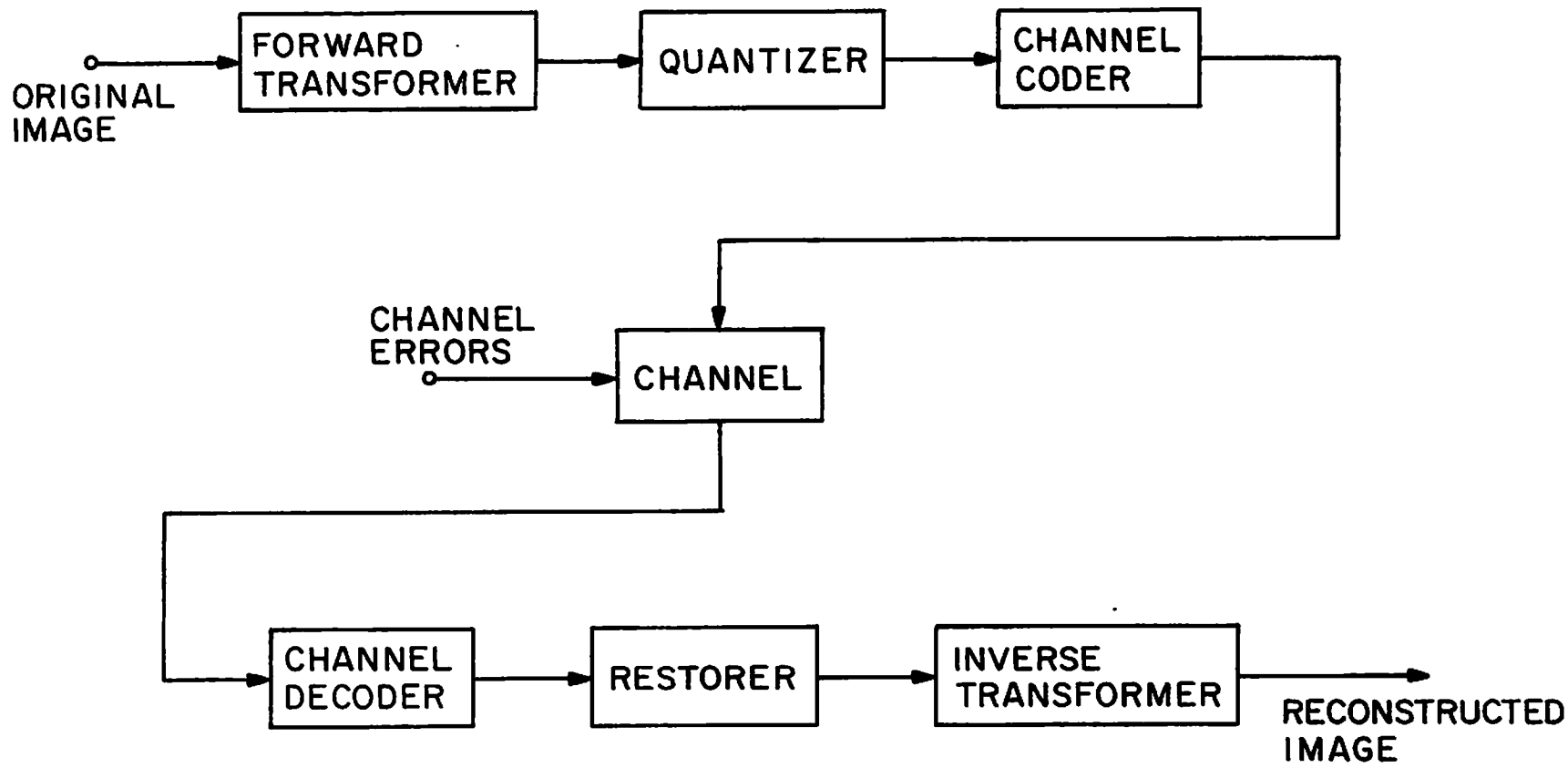


Figure 5-5. Diagram of a generalized transform image coding system.

detail.

5.4.2.1 Unitary Transformations

The transform operation is usually performed by one of a class of linear unitary operators <*> which have been recently applied to image coding [7]. A unitary transform is the most useful because it preserves the "length" or the energy of a vector component during the transformation, and because it transforms every orthonormal basis into another orthonormal basis. To interpret this concept for image coding purposes reveals: each unitary operator can be represented by a matrix composed of orthonormal basis functions (generalized spectral functions); each component in the transform domain then corresponds to the amount of energy of one of the basis (spectral) functions in the original (or image) domain. Unitary operators transform sums of squares into sums of squares, so that mean-square error calculations are equivalent in either domain. This also means that the total energy in the original and transform domains is the same. The unitary transforms which are most commonly used are

1. Karhunen-Loeve transform
2. Cosine transform

<*> A linear operator $\underline{H}$ such that $\underline{H}\underline{H}^*=\underline{H}^*\underline{H}=\underline{I}$ is called a unitary operator, where * denotes the conjugate transpose of the complex operator. In the real case, $\underline{H}$ is called orthogonal.

3. Fourier transform
4. Slant transform
5. Hadamard transform
6. Haar transform

These are listed in order of efficiency, i.e., from the most efficient transform to the least. They are, however, listed in reverse order of simplicity and speed of operation.

The Karhunen-Loeve (K-L) transform results in uncorrelated transform domain components. It has been found that quantizing these uncorrelated components produces a minimum possible mean-square quantization error [8]. Since the components are uncorrelated, they can be quantized and restored individually without loss of performance, so that the restoration method of Sec. 5.1 provides no advantage over simple Max quantization and restoration. Less efficient transforms result in greater quantization error than the K-L transform at the same bit rate, but they can be made to operate much faster. Their transform domain components have some statistical correlation remaining (if the components in the image domain were originally correlated), so that the restoration method of Sec. 5.1 can be applied. This restoration reduces the mean-square quantization error. The remaining error approaches the minimum that is obtainable by means of the K-L transform. Thus, almost the same performance as a

K-L transform can be obtained by utilizing a simpler transform together with a more complicated reconstruction technique. This procedure would be most useful, in practice, when an image coder is required to have a simpler design than the decoder. For the image coding simulations and reconstructions in this chapter, the Hadamard and the Haar transforms are utilized. These have fast and easily implementable algorithms, and also provide partially correlated transform domain samples which can then be restored.

5.4.2.2 Zonal Coding

There is, in general, a uniform distribution of information or energy throughout an image. Unitary transforms, although preserving the same total amount of energy as in the original image, rearrange this energy and concentrate it in a few of the transform domain components [9]. This concentration is achieved because the transform makes use of some of the correlation in the image. Zonal coding, which takes advantage of the energy concentration and rearrangement, entails the establishment of zones of constant bit assignments in the transform domain corresponding to zones of approximately constant energy. The bit assignments are chosen according to the assumed variances--which are a measure of the energy--of the transform domain components. A solution for this

assignment, first derived by Huang and Schultheiss [10], is

$$B_{ij}=b+\frac{1}{2\ \ln 2}\ln\left(\frac{V_{ij}}{|\underline{V}|^{1/N}}\right) \qquad (5.49)$$

where $\underline{V}$ is the matrix of transform domain variances with components V_{ij}, b is the desired average number of bits for each component, and B_{ij} is the resultant number of bits for the (i,j)th component. This equation provides the bit assignment needed to quantize gaussian variables based on a minimum mean-square error criterion. By definition, the number of bits assigned to each component must be a nonnegative integer. Since eq. 5.49 can produce non-integer and even negative values, the bit assignment it produces must be adjusted by trial-and-error techniques to obtain the final bit assignment for each component. A typical bit assignment for a 16x16, Haar transformed, image block that is zonal coded with an average of one bit is shown in fig. 5-6.

5.4.2.3 Spatial and Transform Domain Correlation Matrices

To calculate an optimum bit assignment using eq. 5.49, it is necessary to know the variance matrix of the transform domain samples. This can be derived if the correlation tensor of either the transform domain samples or the image domain samples is known. The exact

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

Figure 5-6. Typical transform domain quantizing bit assignment for one bit per pixel transform coding.

correlation information is usually unavailable, but the image domain correlation can often be accurately modeled as a Markov process. In fact, experimental evidence [11,12] indicates that a reasonable autocorrelation function for a large variety of pictorial data is given by

$$C(x,x',y,y')=\exp(-h|x-x'|-v|y-y'|) \tag{5.50}$$

(For convenience it is assumed that the continuous image, $F(x,y)$, has zero mean.) This function can be used to model images with different amounts of horizontal and vertical correlation by choosing different values for h and v, respectively. An image which has an autocorrelation function that is invariant to translation is said to be wide-sense stationary. The autocorrelation can then be rewritten as

$$C(\Delta x,\Delta y)=\exp(-h|\Delta x|-v|\Delta y|) \tag{5.51}$$

This autocorrelation function also possesses the property of horizontal and vertical separability, i.e., the correlation between any two points of the image is separable into the product of horizontal and vertical correlation functions.

Now, in discrete notation for an NxN sampled image $\underline{F}$, the correlation between any two pixels can be represented

by a four-dimensional tensor as

$$C(i,j,k,l)=E\{F(i,k)F(j,l)\} \tag{5.52}$$

for $1\leq i,j,k,l\leq N$. It is sometimes convenient to column-scan an image array into a data vector, $\underline{f}$, of resultant length N^2 [13]. In this case its correlation matrix, $\underline{C}_f$, is $N^2 x N^2$, and can be represented in partitioned form as

$$\underline{C}_f=\begin{pmatrix} \underline{C}_{1,1} & \underline{C}_{1,2} & \cdots & \underline{C}_{1,N} \\ \underline{C}_{2,1} & \underline{C}_{2,2} & \cdots & \underline{C}_{2,N} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \underline{C}_{N,1} & \underline{C}_{N,2} & \cdots & \underline{C}_{N,N} \end{pmatrix} \tag{5.53}$$

where $\underline{C}_{i,j}$ is the correlation matrix of the (i)th and (j)th columns of $\underline{F}$. Under the assumption of wide-sense stationarity, the correlation matrix has the block Toeplitz form

$$\underline{C}_f = \begin{pmatrix} \underline{C}_1 & \underline{C}_2 & \cdots & \underline{C}_N \\ \underline{C}_2 & \underline{C}_1 & \cdots & \underline{C}_{N-1} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \underline{C}_N & \underline{C}_{N-1} & \cdots & \underline{C}_1 \end{pmatrix} \tag{5.54}$$

Finally, when the correlation is horizontally and vertically separable, the correlation matrix can be written in direct product <*> form as

$$\underline{C}_f = \underline{C}_v \otimes \underline{C}_h \tag{5.55}$$

where $\underline{C}_v$ and $\underline{C}_h$ denote the NxN correlation matrices of the rows and columns of $\underline{F}$, respectively. If the image is now considered to be a sample of a Markov process with a correlation coefficient of v ($0 \leq v \leq 1$) between vertically adjacent pixels and coefficient h between horizontally adjacent pixels, then

---

<*> The symbol ⊗ denotes the matrix direct product. The definition employed here is the left direct product [15].

$$\underline{C}_v = s_v^2 \begin{pmatrix} 1 & v & v^2 & \cdots & v^{N-1} \\ v & 1 & v & \cdots & v^{N-2} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ v^{N-1} & v^{N-2} & v^{N-3} & \cdots & 1 \end{pmatrix} \tag{5.56}$$

and

$$\underline{C}_h = s_h^2 \begin{pmatrix} 1 & h & h^2 & \cdots & h^{N-1} \\ h & 1 & h & \cdots & h^{N-2} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ h^{N-1} & h^{N-2} & h^{N-3} & \cdots & 1 \end{pmatrix} \tag{5.57}$$

where s_v^2 is the variance of the pixels in each column and s_h^2 is the variance of the pixels in each row. After performing a separable unitary transformation $\underline{H}$ <*> on the image $\underline{F}$, the corresponding row and column correlation matrices in the transform domain can be found by

---

<*> For a separable transformation,

$$\underline{H} = \underline{H}_v \otimes \underline{H}_h$$

so that

$$\underline{\mathcal{F}} = \underline{H}_v \underline{F} \underline{H}_h$$

$$\underline{C}_f = \underline{H}\underline{C}_f \underline{H}^{*T}$$

$$= (\underline{H}_v \otimes \underline{H}_h)(\underline{C}_v \otimes \underline{C}_h)(\underline{H}_v \otimes \underline{H}_h)^{*T}$$

$$= (\underline{H}_v \underline{C}_v \underline{H}_v^{*T}) \otimes (\underline{H}_h \underline{C}_h \underline{H}_h^{*T})$$

$$= \underline{C}_v \otimes \underline{C}_h \qquad (5.58)$$

The matrices $\underline{C}_v$ and $\underline{C}_h$ can now be used to find the transform domain variance matrix, as

$$\underline{V} = \underline{V}_h \underline{V}_v^{*T} \qquad (5.59)$$

where the vectors $\underline{V}_h$ and $\underline{V}$ consist of the diagonal elements of $\underline{C}_h$ and $\underline{C}_v$, respectively.

5.4.2.4 Quantization Levels

Specification of the variance matrix permits the determination of an optimum bit assignment. However, this provides only the number of levels to which each component is quantized and does not provide the locations of these levels. For the analysis presented here, the placement of these levels is arbitrary, but is chosen to be the Max quantization levels to provide the minimum mean-square quantization error.

5.4.2.5 Zonal Coded Images

For the transform image coding simulations, the correlations were chosen to be h=Ø.95 and v=Ø.93. These values most closely modeled the correlations of the images chosen for the simulations, and resulted in the smallest quantization errors and reconstruction errors. Figure 5-7 shows the three "original" images which were used for the simulations: each image is an array of 256x256 pixels, with the intensity at each pixel quantized to one of 256 (8 bits) grey levels. The images were Haar transformed in blocks of 16x16 pixels, so that a 16x16 variance matrix was required to compute the bit assignment. Figure 5-8a displays the "girl" image after being zonal Haar transform coded with an average of 1.Ø bits according to the bit assignment shown in fig. 5-6. (This image actually shows the Max restoration levels after the image has been quantized according to the Max decision levels.) Figure 5-8b shows the reconstructed version of this image after it was restored according to the techniques of Sec. 5.1. The restoration technique was applied recursively to the quantized samples since it is only capable of restoring one sample at a time; the current best estimates of the other samples were used to obtain the estimate of the sample being restored. The initial estimates of the samples were chosen to be the Max restoration levels. The procedure, in

GIRL

COUPLE

MOON

Figure 5-7. Original images used for image coding simulations. Each image consists of 256x256 pixels, with each pixel quantized to 8 bits.

a. Quantized image

b. Restored image

Figure 5-8. Minimum mean-square error restoration of Haar transformed, one bit zonal quantized image.

essence, begins with the Max restored image as an initial guess and attempts to improve it by utilizing the known correlation of the transform domain samples. One iteration of this restoration procedure was utilized, because it was experimentally determined that more iterations were of negligible benefit. The mean-square error was reduced by 10% as a result of the restoration.

The "girl" image was next quantized to an average of 0.5 bits according to the zonal bit assignment shown in fig. 5-9. Figure 5-10a shows the Max restored version of this coded image and fig. 5-10b shows the version reconstructed according to the technique discussed above. In this example, the reconstructed version has 19% less mean-square error. The "couple" and the "moon" pictures were also quantized to 0.5 bits and restored. The results are shown in fig. 5-10. The resultant mean-square errors are summarized in table 5-3. The restoration technique of Sec. 5.1 is seen to provide a significant decrease in mean-square error in all cases.

5.4.2.6 Visual Coded Images

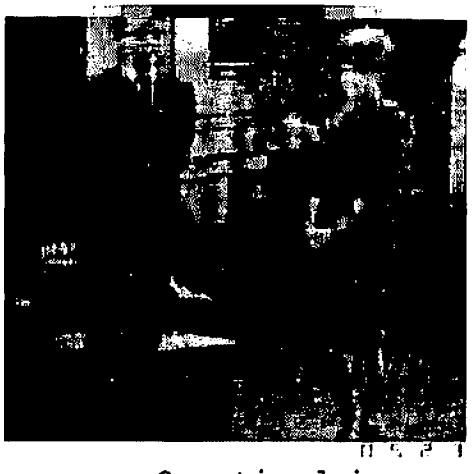
Subjectively, the reconstructed images of figures 5-8b, 5-10b, 5-10d, and 5-10f appear to be much smoother and less noisy than the corresponding Max restored images of figures 5-8a, 5-10a, 5-10c, and 5-10e. However, they

7	5	4	4	3	3	3	3	2	2	2	2	2	2	2	2
4	3	2	2	1	1	1	1	1	1	1	1	1	1	1	1
3	2	1	1	1	1	1	1	0	0	0	0	0	0	0	0
3	2	1	1	1	1	1	1	0	0	0	0	0	0	0	0
3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 5-9. Haar transform domain zonal quantizing bit assignment for correlation factors h=0.95 and v=0.93 and for 0.5 bits per pixel.

a. Quantized image

b. Restored image



c. Quantized image

d. Restored image

e. Quantized image

f. Restored image

Figure 5-10. Minimum mean-square error restoration of Haar transformed, 0.5 bit zonal quantized image.

TABLE 5-3

NORMALIZED MEAN-SQUARE ERROR
FOR ZONAL CODED IMAGES

Picture	Error/Pixel after Quantization	Error/Pixel after Restoration	Improvement
Girl 1.0 bit/pixel	1.024%	0.922%	9.9%
Girl 0.5 bit/pixel	1.866%	1.509%	19.2%
Couple 0.5 bit/pixel	1.793%	1.412%	21.2%
Moon 0.5 bit/pixel	1.523%	1.158%	24.0%

sometimes appear more blurred, as is very evident in comparing figures 5-10e and 5-10f. An improvement in mean-square error apparently does not correspond to an improvement in subjective quality in all cases. Hence an error measure is required in which numerical results match subjective results.

This has been provided by modeling the error measure after the human visual system. Mannos and Sakrison [14] have derived a nonlinear error criterion which achieves this objective. They found that the human visual system is sensitive to approximately the cube root of incident light intensities. It is also most sensitive to middle spatial frequencies, near eight cycles per degree of arc subtended at the eye. Hence, to apply this error measure, an image has been processed according to the block diagram in fig. 5-11. The (i,j)th component of the filter function shown there is chosen to be

$$T_{ij}=(.05+.18525r)\exp\{-(.07125r)^{1.1}\} \qquad (5.60)$$

where

$$r=(i^2+j^2)^{1/2} \qquad (5.61)$$

This filter was applied to the Hadamard domain sequencies of the image, rather than the Fourier domain frequencies,

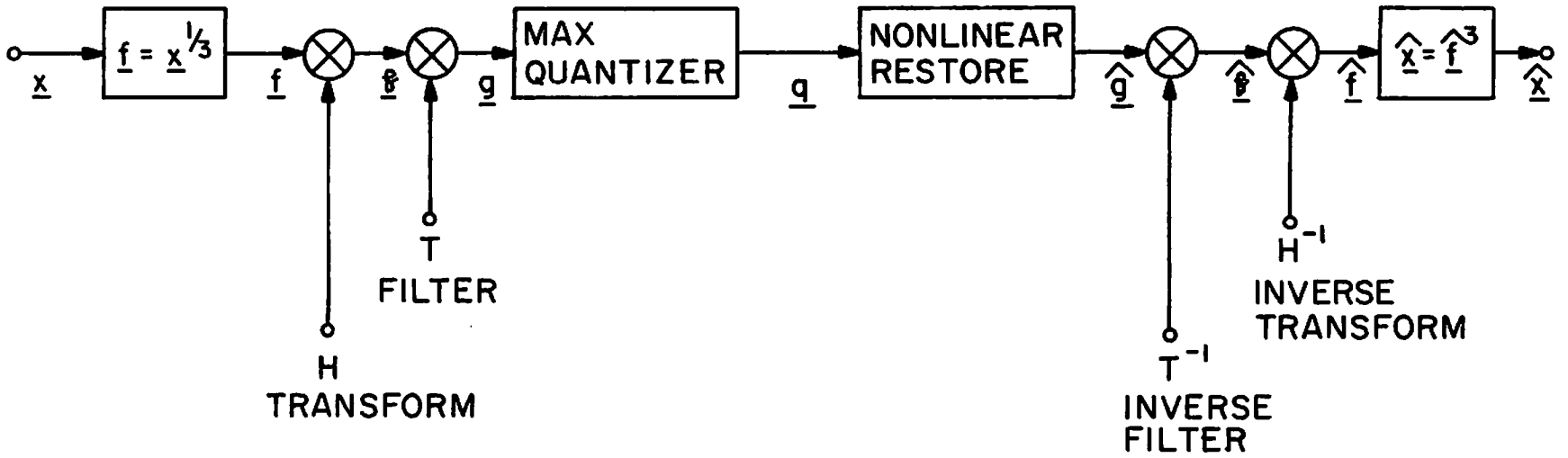


Figure 5-11. Coding and restoration technique for a nonlinear error criterion.

because of the close similarities between sequencies and frequencies. Figures 5-12a and 5-12b show the results of the procedure for a 0.5 bit zonal coded Hadamard transformed "moon" image, with and without the reconstruction operation, respectively. The bit assignment, shown in fig. 5-13, was calculated by choosing h=0.96 and v=0.944 as the spatial, cube-root domain, correlations. The resultant Hadamard domain variance matrix was then scalar filtered by a multiplication by the square of the filter function T. The mean-square error is reduced by 5.5% due to the reconstruction operation. There is also a noticeable subjective improvement after the restoration. Figures 5-12c, 5-12d, 5-12e, and 5-12f display similar results for the "girl" and "couple" pictures. Table 5-4 summarizes the errors for the restorations of these pictures.

5.4.2.7 Zonal Coded Color Images

The restoration technique has also been applied in an experiment in which a color image is encoded. In this experiment, a color image is first transformed to the YIQ color coordinate system and then quantized according to the bit assignment indicated below for a typical block of four pixels.

a. Visual quantized image

b. Visual restored image

c. Visual quantized image

d. Visual restored image

e. Visual quantized image

f. Visual restored image

Figure 5-12. Restoration of 0.5 bit Hadamard transformed zonal quantized images according to a visual error criterion.

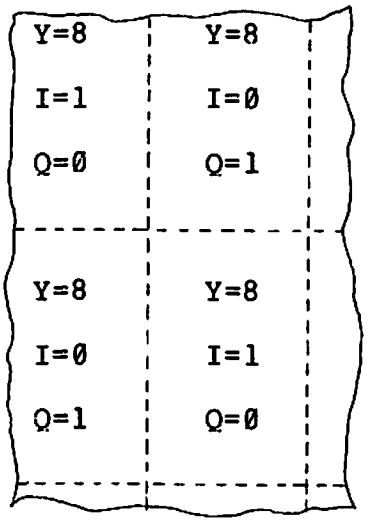
5	4	3	3	3	3	3	3	2	2	2	2	2	2	2	2
4	3	2	2	1	1	2	2	1	1	1	1	1	1	1	1
3	2	1	1	1	1	1	1	0	0	0	0	0	0	0	0
3	2	1	1	1	1	1	1	0	0	0	0	0	0	0	0
3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
3	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0
3	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 5-13. A visual quantizing bit assignment in the Hadamard transform domain for spatial, cube-root domain, correlation factors h=0.96 and v=0.944, and for 0.5 bits per pixel.

TABLE 5-4

NORMALIZED MEAN-SQUARE ERROR
FOR VISUAL CODED IMAGES

Picture	Error/Pixel after Quantization	Error/Pixel after Restoration	Improvement
Girl 0.5 bit/pixel	1.21%	1.11%	8.3%
Couple 0.5 bit/pixel	1.94%	1.79%	7.5%
Moon 0.5 bit/pixel	0.73%	0.69%	5.5%



Each pixel is hence coded with an average of nine bits per pixel, compared to an original coding assignment of 24 bits per pixel. The restoration technique of Sec. 5.1 provides a decrease of 42% in mean-square error in this case and an improvement in subjective quality.

5.5 Summary

Gaussian data which are correlated and which have been coarsely quantized are amenable to being restored by the techniques outlined in this chapter. By choosing a suitable error criterion, zonal transform coded images can be analytically and, in many cases, subjectively improved so that they more faithfully reproduce the details of an original image. The restoration techniques have been found to be most successful when the transform is inefficient, i.e., when some statistical correlation remains between the transform domain samples, and when the quantization is

coarse.

REFERENCES

1. R. E. Curry, Estimation and Control with Quantized Measurements, The M.I.T. Press, Cambridge, Massachusetts, 1970, pp. 21,22,101-103.

2. H. L. Van Trees, Detection, Estimation, and Modulation Theory, Part I, John Wiley and Sons, Inc., New York, 1968, pp. 178-182.

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

4. C. R. Rao, Linear Statistical Inference and Its Applications, John Wiley and Sons, Inc., New York, 1965, p. 29.

5. V. I. Krylov, Approximate Calculation of Integrals, Macmillan, New York/London, 1962, pp. 129-130 and 343-346.

6. J. Max, "Quantizing for Minimum Distortion," IEEE Transactions on Information Theory, vol. IT-6, March 1960, pp. 7-12.

7. W. K. Pratt and H. C. Andrews, "Two-dimensional Transform Coding for Images," presented at the 1969 IEEE International Symposium on Information Theory, January 1969.

8. W. K. Pratt, "Karhunen-Loeve Transform Coding of Images," presented at the 1970 IEEE International Symposium on Information Theory, June 1970.

9. P. A. Wintz, "Transform Picture Coding," Proceedings of the IEEE, vol. 60, July 1972, pp. 809-820.

10. J. J. Y. Huang and P. M. Schultheiss, "Block Quantization of Correlated Gaussian Random Variables," IEEE Transactions on Communications Systems, September 1963, pp. 289-296.

11. L. E. Franks, "A Model for the Random Video Process," Bell System Technical Journal, vol. 45, April 1966.

12. T. S. Huang, "The Subjective Effect of Two-dimensional Pictorial Noise," IEEE Transactions on Information Theory, vol. IT-11, January 1965, pp. 43-53.

13. W. K. Pratt, "Vector Space Formulation of Two-Dimensional Signal Processing Operations," Journal of Computer Graphics and Image Processing, Academic Press, March 1975.

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

15. F. A. Graybill, Introduction to Matrices with Applications in Statistics, Wadsworth Publishing Company, Inc., Belmont, California, 1969, p. 197.

CHAPTER 6

RESTORATION OF QUANTIZED LAPLACIAN SAMPLES

Laplacian density functions arise in the stochastic modelling of certain kinds of communications systems. Specifically, the output of a differential pulse code modulation (DPCM) system can often be modelled as a laplacian random process. When a laplacian random variable is quantized, as in a DPCM system, some of the original information about the variable is lost. If an estimate of the continuous laplacian variable is then made, based only upon the output of the quantizer, the estimate will usually be poor. However, if the characteristics of the random process are known and are utilized also, then the estimate can be improved and a reconstruction of the original signal will be attained. The estimation equation derived in Chapter 3 (eq. 3.6) provides a means for accomplishing this restoration. Use of this equation requires knowledge of both the quantizer structure and the multidimensional probability density function of the input to the quantizer. An approximation to a multidimensional laplacian density function was derived in Chapter 4. This density function can then be utilized in the estimation equation, together with arbitrary quantization parameters, to obtain a restoration of quantized laplacian samples. The next section presents in detail the solution to the estimation

equation for this case. The results of that solution are then applied to the restoration of DPCM coded images.

6.1 Laplacian Quantization Estimator

A multidimensional laplacian probability density function, as derived in Chapter 4, can be approximated by

$$p(\underline{x})=\frac{|\underline{C}|^{-1/2}}{2^{N/2}}\exp\{-\sqrt{2}\sum_{i=1}^{N}\left|\sum_{k=1}^{N}g_{ik}x_k\right|\} \tag{6.1}$$

where $\underline{C}$ is the NxN correlation matrix of $\underline{x}$, g_{ik} represents the (i,k)th element of $\underline{G}$ as given by

$$\underline{G}=\underline{E}\underline{\Lambda}^{-1/2}\underline{E}^{T} \tag{6.2}$$

and $\underline{E}$ and $\underline{\Lambda}$ are the matrices of eigenvectors and eigenvalues of $\underline{C}$, respectively. This density function can then be utilized in eq. 3.6 to obtain a minimum mean-square error estimate of a quantized N-vector of laplacian samples, $\underline{x}$. Now, it is assumed that $\underline{x}$ is quantized to a region in N-space, R. It is also assumed that each component of $\underline{x}$ is quantized individually, so that R is rectangular (R can then be denoted $[\underline{a},\underline{b}]$). Then the estimate, $\underline{y}$, of $\underline{x}$, given that $\underline{x} \in R$, is

$$\underline{y}=\frac{\int_R \underline{x}\ \exp\{-\sqrt{2}\sum_{i=1}^{N}\left|\sum_{k=1}^{N} g_{ik}x_k\right|\}d\underline{x}}{\int_R \exp\{-\sqrt{2}\sum_{i=1}^{N}\left|\sum_{k=1}^{N} g_{ik}x_k\right|\}d\underline{x}} \tag{6.3}$$

A general solution to these integrals exists, but is rather complicated to state. Special cases of practical interest are considered in the following paragraphs.

6.1.1 Scalar Case

When the components of $\underline{x}$ are uncorrelated, or when they are restored individually, then eq. 6.3 can be decomposed into a product of one-dimensional integrals which can be solved separately. Considering one of the components of $\underline{x}$ (for a unit variance),

$$y_i=\frac{\int_{a_i}^{b_i} x_i \exp\{-\sqrt{2}|x_i|\}dx_i}{\int_{a_i}^{b_i} \exp\{-\sqrt{2}|x_i|\}dx_i}$$

or

$$y_i=\frac{(|b_i|+\frac{1}{\sqrt{2}})\exp\{-\sqrt{2}|b_i|\}-(|a_i|+\frac{1}{\sqrt{2}})\exp\{-\sqrt{2}|a_i|\}}{\frac{b_i}{|b_i|}(\exp\{-\sqrt{2}|b_i|\}-1)-\frac{a_i}{|a_i|}(\exp\{-\sqrt{2}|a_i|\}-1)} \tag{6.4}$$

If a_i and b_i are chosen to be Max decision levels, then y_i is the corresponding Max restoration level.

6.1.2 Two-dimensional Case

In two dimensions, eq. 6.3 yields simultaneous estimates of two quantized and correlated laplacian variables. The solution is lengthy, however, so the details are omitted and only the results are shown here. To simplify the notation, let the exponential terms in the two-dimensional version of eq. 6.3 be written as

$$\exp\{-|fx_1 - gx_2| - |gx_1 - fx_2|\}$$

and let the quantization intervals be $a \leq x_1 \leq b$ and $c \leq x_2 \leq d$. The constants f and g are based on the correlation between x_1 and x_2. Also let $s(\cdot)$ denote the sign or signum function. Then the numerator of y_1 is

$$y_1(\text{num}) = \frac{\{b[s(gd-fb)f+s(fd-gb)g]-1\}\exp\{-|gd-fb|-|fd-gb|\}}{[s(gd-fb)f+s(fd-gb)g]^2\,[s(gb-fd)f+s(fb-gd)g]}$$

$$+ \frac{\{b[s(gc-fb)f+s(fc-gb)g]-1\}\exp\{-|gc-fb|-|fc-gb|\}}{[s(gc-fb)f+s(fc-gb)g]^2\,[s(fc-gb)f+s(gc-fb)g]}$$

$$+ \frac{\{a[s(gd-fa)f+s(fd-ga)g]-1\}\exp\{-|gd-fa|-|fd-ga|\}}{[s(gd-fa)f+s(fd-ga)g]^2\,[s(fd-ga)f+s(gd-fa)g]}$$

$$
- \frac{\{a[s(gc-fa)f+s(fc-ga)g]-1\}\exp\{-|gc-fa|-|fc-ga|\}}{[s(gc-fa)f+s(fc-ga)g]^2[s(fc-ga)f+s(gc-fa)g]}
$$

$$
+ \frac{[s(gd-fb)-s(gd-fa)]gf^2[|d(f^2-g^2)/f|+3]\exp\{-|d(f^2-g^2)/f|\}}{(f^2-g^2)^3}
$$

$$
+ \frac{[s(gc-fa)-s(gc-fb)]gf^2[|c(f^2-g^2)/f|+3]\exp\{-|c(f^2-g^2)/f|\}}{(f^2-g^2)^3}
$$

$$
- \frac{[s(fd-gb)-s(fd-ga)]fg^2[|d(f^2-g^2)/g|+3]\exp\{-|d(f^2-g^2)/g|\}}{(f^2-g^2)^3}
$$

$$
- \frac{[s(fc-ga)-s(fc-gb)]fg^2[|c(f^2-g^2)/g|+3]\exp\{-|c(f^2-g^2)/g|\}}{(f^2-g^2)^3}
$$

$$
+ \frac{[s(gd-fb)-s(gc-fb)]g^3[|b(f^2-g^2)/g|+1]\exp\{-|b(f^2-g^2)/g|\}}{(f^2-g^2)^3}
$$

$$
- \frac{[s(fd-gb)-s(fc-gb)]f^3[|b(f^2-g^2)/f|+1]\exp\{-|b(f^2-g^2)/f|\}}{(f^2-g^2)^3}
$$

$$
- \frac{[s(gd-fa)-s(gc-fa)]g^3[|a(f^2-g^2)/g|+1]\exp\{-|a(f^2-g^2)/g|\}}{(f^2-g^2)^3}
$$

$$
+ \frac{[s(fd-ga)-s(fc-ga)]f^3[|a(f^2-g^2)/f|+1]\exp\{-|a(f^2-g^2)/f|\}}{(f^2-g^2)^3}
$$

(6.5)

The denominator is

$$y_1(\text{denom}) = \frac{\exp\{-|fb-gd|-|gb-fd|\}}{[s(gd-fb)f+s(fd-gb)g][s(gb-fd)f+s(fb-gd)g]}$$

$$- \frac{\exp\{-|fb-gc|-|gb-fc|\}}{[s(gc-fb)f+s(fc-gb)g][s(gb-fc)f+s(fb-gc)g]}$$

$$- \frac{\exp\{-|fa-gd|-|ga-fd|\}}{[s(gd-fa)f+s(fd-ga)g][s(ga-fd)f+s(fa-gd)g]}$$

$$+ \frac{\exp\{-|fa-gc|-|ga-fc|\}}{[s(gc-fa)f+s(fc-ga)g][s(ga-fc)f+s(fa-gc)g]}$$

$$- \frac{[s(fd-gb)-s(fc-gb)]f^2\exp\{-|b(f^2-g^2)/f|\}b/f}{(f^2-g^2)|b(f^2-g^2)/f|}$$

$$- \frac{[s(gc-fb)-s(gd-fb)]g^2\exp\{-|b(f^2-g^2)/g|\}b/g}{(f^2-g^2)|b(f^2-g^2)/g|}$$

$$+ \frac{[s(fd-ga)-s(fc-ga)]f^2\exp\{-|a(f^2-g^2)/f|\}a/f}{(f^2-g^2)|a(f^2-g^2)/f|}$$

$$+ \frac{[s(gc-fa)-s(gd-fa)]g^2\exp\{-|a(f^2-g^2)/g|\}a/g}{(f^2-g^2)|a(f^2-g^2)/g|}$$

$$+ \frac{[s(gd-fb)-s(gd-fa)]f^2\exp\{-|d(f^2-g^2)/f|\}d/f}{(f^2-g^2)|d(f^2-g^2)/f|}$$

$$- \frac{[s(gc-fb)-s(gc-fa)]f^2\exp\{-|c(f^2-g^2)/f|\}c/f}{(f^2-g^2)|c(f^2-g^2)/f|}$$

$$- \frac{[s(fd-gb)-s(fd-ga)]g^2 \exp\{-|d(f^2-g^2)/g|\}d/g}{(f^2-g^2)|d(f^2-g^2)/g|}$$

$$+ \frac{[s(fc-gb)-s(fc-ga)]g^2 \exp\{-|c(f^2-g^2)/g|\}c/g}{(f^2-g^2)|c(f^2-g^2)/g|} \qquad (6.6)$$

Hence,

$$y_1 = \frac{y_1(\text{num})}{y_1(\text{denom})} \qquad (6.7)$$

The other component, y_2, can be found in a similar manner. Expressing the constants f and g in terms of the correlation between x_1 and x_2 reveals

$$f = \frac{\sqrt{1+r} + \sqrt{1-r}}{\sqrt{2(1-r^2)}} \qquad (6.8)$$

and

$$g = \frac{\sqrt{1+r} - \sqrt{1-r}}{\sqrt{2(1-r^2)}} \qquad (6.9)$$

where

$$r = E\{x_1 x_2\} \qquad (6.10)$$

It can be seen that as the correlation, r, approaches zero, g also approaches zero and the solution (equations 6.5 to

6.7) becomes the same as that shown in eq. 6.4 for the scalar case. For non-zero r, equations 6.5 to 6.7 provide a minimum mean-square error restoration of two correlated laplacian variables which have been quantized to the arbitrary intervals $a \leq x_1 \leq b$ and $c \leq x_2 \leq d$.

6.2 Covariance of the Laplacian Estimator

The performance of the estimator found in equations 6.5 to 6.7 can be analyzed by computing its conditional covariance matrix. Equation 3.10 contains a general expression for finding this covariance matrix which, in this case, must be solved for a two-dimensional laplacian probability density function. A general solution for this covariance matrix has been obtained, but is too complicated to be shown here. Instead, the covariance matrix to be derived is for the special case of one-bit quantization only.

The two quantization intervals for the one-bit quantizer are chosen to be $[0,\infty)$ and $(-\infty,0)$. This choice is the optimum one for quantizing individual laplacian samples according to a mean-square error criterion. There are now four possible rectangular regions, R_m for $m=1,2,3,4$, into which $\underline{x}=(x_1,x_2)$ can be quantized. Within each region there is a restoration point which can be found by means of equations 6.5 to 6.7. The restoration points

are dependent on the correlation, r, between x_1 and x_2; on whether x_i is quantized to a positive or a negative interval, denoted by $s(x_i)\lessgtr 0$; and on the variance of x_i, denoted as σ_i^2. Then the restoration points are

1. for $rs(x_1)s(x_2)\geq 0$

$$y_1 = \frac{\sigma_1(1+2|r|)s(x_1)}{2(2\sqrt{1+|r|}-\sqrt{1-|r|})} \tag{6.11a}$$

$$y_2 = \frac{\sigma_2(1+2|r|)s(x_2)}{2(2\sqrt{1+|r|}-\sqrt{1-|r|})} \tag{6.11b}$$

2. for $rs(x_1)s(x_2)<0$

$$y_1 = \sigma_1\sqrt{\frac{1-|r|}{2}}s(x_1) \tag{6.12a}$$

$$y_2 = \sigma_2\sqrt{\frac{1-|r|}{2}}s(x_2) \tag{6.12b}$$

These values for $\underline{y}$ can now be substituted into eq. 3.10 and the two-dimensional integrals evaluated to obtain the covariance matrix for each of the two cases (assuming $\sigma_1=\sigma_2=1$):

1. for $rs(x_1)s(x_2)\geq 0$

$$\operatorname{cov}\{\underline{x} \mid \underline{x} \in R\} = \begin{pmatrix} \dfrac{2\sqrt{1-r^2}-(1-|r|)^2}{2\sqrt{1-r^2}-(1-|r|)} & \dfrac{|r|\sqrt{1-r^2}+2(1-|r|)^2}{4[2\sqrt{1-r^2}-(1-|r|)]} \\ \dfrac{|r|\sqrt{1-r^2}+2(1-|r|)^2}{4[2\sqrt{1-r^2}-(1-|r|)]} & \dfrac{2\sqrt{1-r^2}-(1-|r|)^2}{2\sqrt{1-r^2}-(1-|r|)} \end{pmatrix}$$

$$- \begin{bmatrix} \dfrac{(1+2|r|)s(x_1)}{\sqrt{2}(2\sqrt{1+|r|}-\sqrt{1-|r|})} \\ \dfrac{(1+2|r|)s(x_2)}{\sqrt{2}(2\sqrt{1+|r|}-\sqrt{1-|r|})} \end{bmatrix} \begin{bmatrix} \dfrac{(1+2|r|)s(x_1)}{\sqrt{2}(2\sqrt{1+|r|}-\sqrt{1-|r|})} & \dfrac{(1+2|r|)s(x_2)}{\sqrt{2}(2\sqrt{1+|r|}-\sqrt{1-|r|})} \end{bmatrix}$$

$$= \begin{pmatrix} \dfrac{4(|r|-2)\sqrt{1-r^2}-2r^2+9}{2(5+3|r|-4\sqrt{1-r^2})} & \dfrac{(4-5|r|)\sqrt{1-r^2}-8r^2-2|r|-4}{4(5+3|r|-4\sqrt{1-r^2})} \\ \dfrac{(4-5|r|)\sqrt{1-r^2}-8r^2-2|r|-4}{4(5+3|r|-4\sqrt{1-r^2})} & \dfrac{4(|r|-2)\sqrt{1-r^2}-2r^2+9}{2(5+3|r|-4\sqrt{1-r^2})} \end{pmatrix} \quad (6.13)$$

2. for $rs(x_1)s(x_2)<0$

$$\operatorname{cov}\{\underline{x} \mid \underline{x} \in R\} = \begin{pmatrix} 1-r & -\frac{1-r}{2} \\ -\frac{1-r}{2} & 1-r \end{pmatrix} - \begin{pmatrix} \frac{1-r}{2} & -\frac{1-r}{2} \\ -\frac{1-r}{2} & \frac{1-r}{2} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1-r}{2} & 0 \\ 0 & \frac{1-r}{2} \end{pmatrix} \tag{6.14}$$

The probability of occurrence of each quantization region is also dependent on the correlation. These probabilities are found by integrating the two-dimensional probability density function over each region. The results are

1. for $r \geq 0$

$$\Pr\{s(x_1)s(x_2) \geq 0\} = \frac{1}{2} - \frac{1}{4}\sqrt{\frac{1-r}{1+r}} \tag{6.15a}$$

$$\Pr\{s(x_1)s(x_2) < 0\} = \frac{1}{4}\sqrt{\frac{1-r}{1+r}} \tag{6.15b}$$

2. for $r<0$

$$\Pr\{s(x_1)s(x_2)\geq 0\}=\frac{1}{4}\sqrt{\frac{1+r}{1-r}} \tag{6.16a}$$

$$\Pr\{s(x_1)s(x_2)<0\}=\frac{1}{2}-\frac{1}{4}\sqrt{\frac{1+r}{1-r}} \tag{6.16b}$$

These probabilities, together with the covariance matrix, $\underline{C}$, for the original vector $\underline{x}$,

$$\underline{C}=\begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix} \tag{6.17}$$

and the expressions for the restoration points listed in equations 6.11 and 6.12, can be used in eq. 3.13 to obtain an expression for the total mean-square error <*>:

$$\mathcal{E}=\mathrm{Tr}\{\underline{C}-\sum_{m=1}^{4}\underline{y}_m\underline{y}_m^{T}\Pr(\underline{x}\in R_m)\} \tag{3.13}$$

which becomes

<*> The total mean-square error is the same whether r is positive or negative, so only the results for $r\geq 0$ are shown.

$$\hat{\mathcal{E}}=\mathrm{Tr}\left\{\underline{C}-\begin{pmatrix}\frac{3r^2+6r+(2-2r)\sqrt{1-r^2}}{8+8r-4\sqrt{1-r^2}} & \frac{5r^2+2r+2-(2-2r)\sqrt{1-r^2}}{8+8r-4\sqrt{1-r^2}}\\ \frac{5r^2+2r+2-(2-2r)\sqrt{1-r^2}}{8+8r-4\sqrt{1-r^2}} & \frac{3r^2+6r+(2-2r)\sqrt{1-r^2}}{8+8r-4\sqrt{1-r^2}}\end{pmatrix}\right\}$$

$$=\mathrm{Tr}\begin{pmatrix}\frac{8+2r-3r^2+(6-2r)\sqrt{1-r^2}}{8+8r-4\sqrt{1-r^2}} & \frac{3r^2+6r-2+(2-2r)\sqrt{1-r^2}}{8+8r-4\sqrt{1-r^2}}\\ \frac{3r^2+6r-2+(2-2r)\sqrt{1-r^2}}{8+8r-4\sqrt{1-r^2}} & \frac{8+2r-3r^2+(6-2r)\sqrt{1-r^2}}{8+8r-4\sqrt{1-r^2}}\end{pmatrix}$$

$$=\frac{8+2r-3r^2-(6-2r)\sqrt{1-r^2}}{4+4r-2\sqrt{1-r^2}} \qquad (6.18)$$

This error is plotted in fig. 6-1 as a function of the correlation r. Also shown is the mean-square error that results when the correlation between any two quantized laplacian samples is not utilized in the reconstruction of these samples. The use of the correlation information is seen to substantially reduce the resultant mean-square error. This holds true independent of the sign of the correlation, but is only true when there is at least a moderate amount of correlation. The next section shows that deltamodulation image coders satisfy this requirement, so that a reconstruction of the coded image can then be

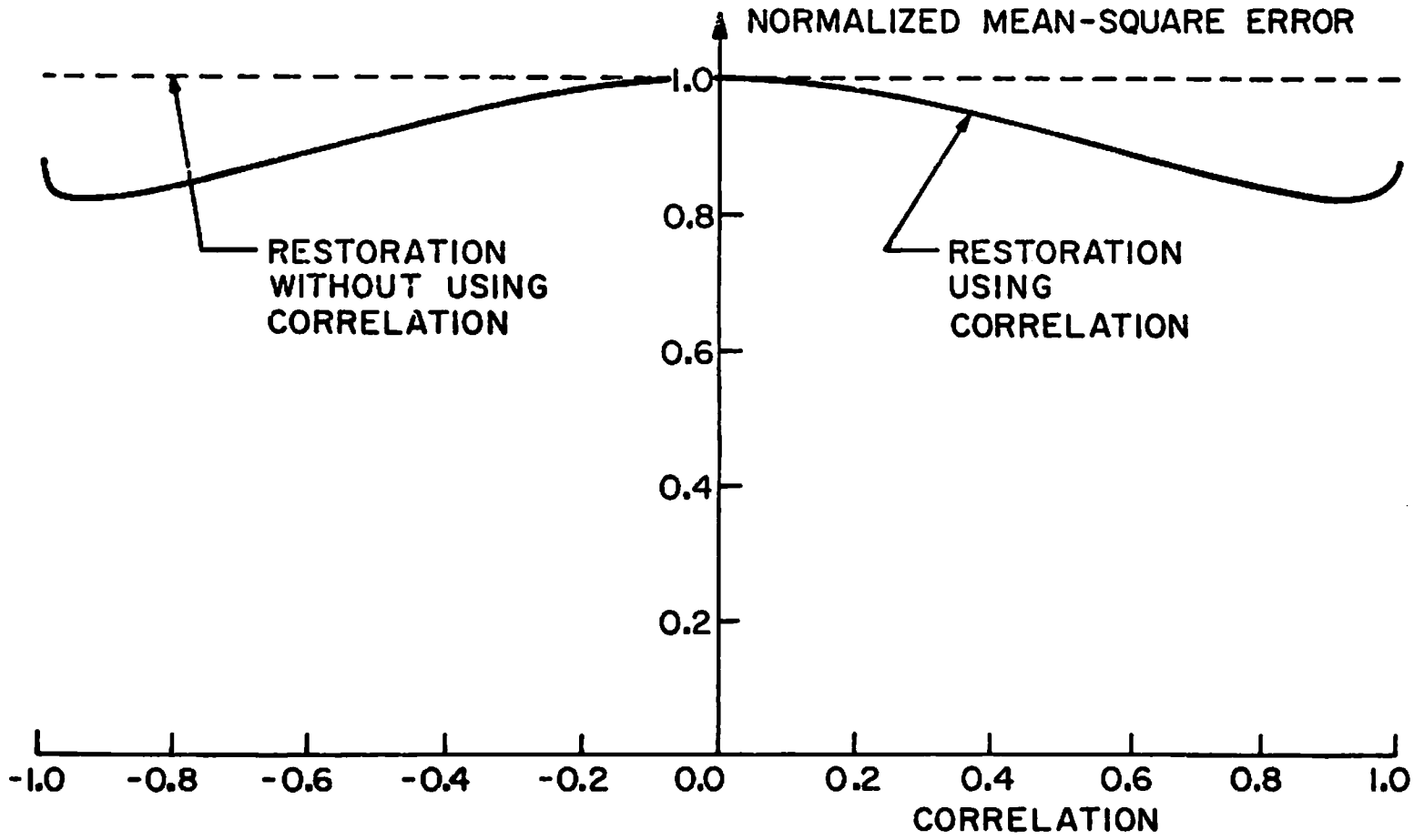


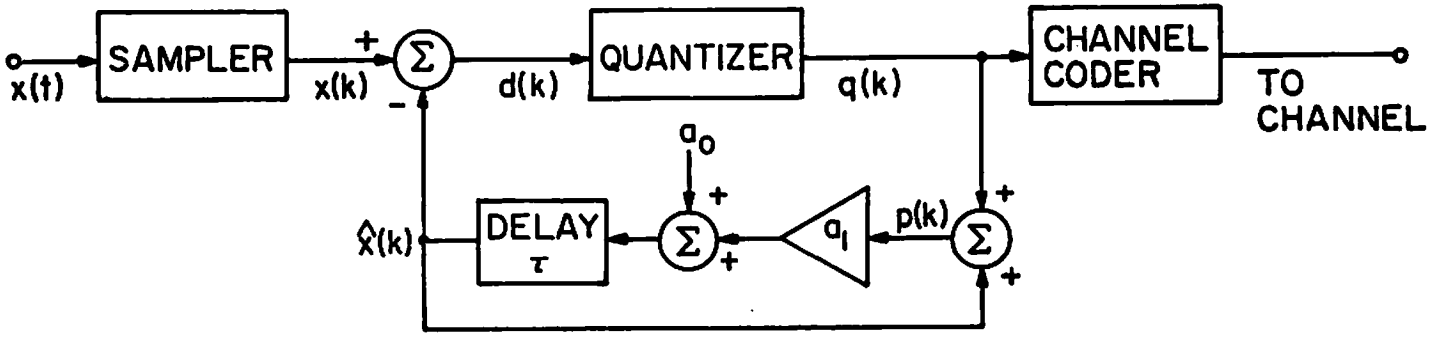
Figure 6-1. Restoration of correlated pairs of laplacian samples that have been quantized to one bit.

obtained.

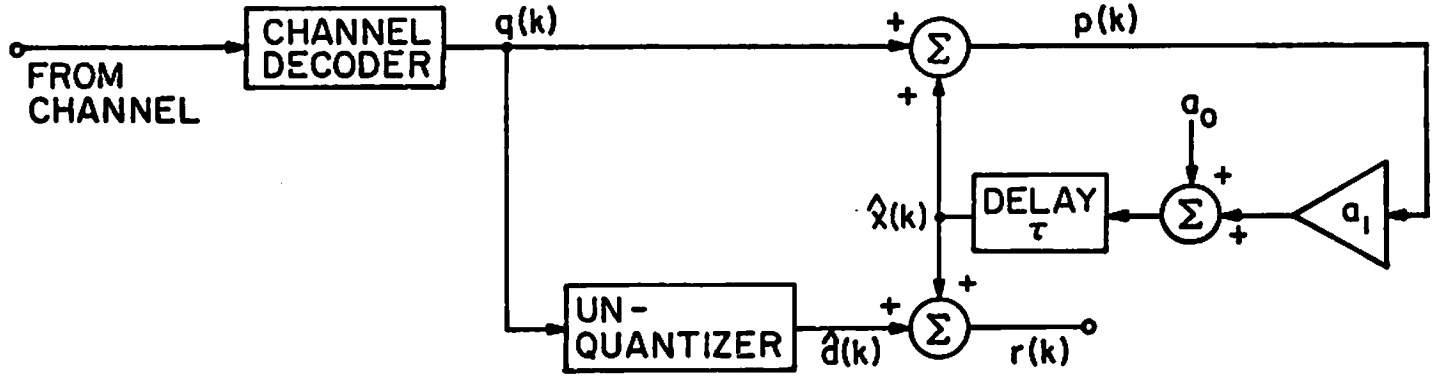
6.3 DPCM and Deltamodulation Image Coding

Differential pulse code modulation (DPCM) is a common technique for achieving bandwidth compression in digital systems. In this method differences between successive signal samples, rather than the signals themselves, are transmitted. Compression occurs because adjacent samples are often very similar, and transmitting only signal differences removes some of this redundancy. A block diagram of a typical DPCM system is shown in fig. 6-2. An essential component of this system is the quantizer, because it permits a bandwidth compression. The coarser the quantization, the greater the compression, but also the greater the degradation of the reconstructed signal. To minimize this degradation, a reconstruction must utilize all of the knowledge that is available about the signal, such as the quantization levels, the signal distribution, and any correlation which remains after the differencing operation. Section 6.1 presented a means for achieving an optimal solution to this restoration problem. This section applies that solution to the reconstruction of DPCM coded images.

In the DPCM system of fig. 6-2, the difference signal $d(k)$ is formed by subtracting a prediction of the current



(a) Transmitter



(b) Receiver

Figure 6-2. Spatial predictive DPCM coding system with quantization restoration.

signal from the current signal itself. The first-order predictor shown in the figure can be characterized by specifying the constants a_0 and a_1. These constants are usually chosen to minimize the mean-square prediction error. The prediction error can be written as

$$\mathcal{E}=E\{[x(k)-(a_1x(k-1)+a_0)]\} \tag{6.19}$$

This expression can be minimized by simultaneously solving the two equations

$$\frac{\partial\mathcal{E}}{\partial a_0}=2a_0-2m+2a_1m=0 \tag{6.20a}$$

and

$$\frac{\partial\mathcal{E}}{\partial a_1}=2a_1s^2-2rs^2+2a_0m=0 \tag{6.20b}$$

where

$$m=E\{x(k)\} \tag{6.21}$$

$$s^2=E\{[x(k)]^2\} \tag{6.22}$$

and

$$r=\frac{E\{x(k)x(k-1)\}}{s^2} \tag{6.23}$$

The results are

$$a_0=\frac{ms^2(1-r)}{s^2-m^2} \tag{6.24}$$

and

$$a_1=\frac{rs^2-m^2}{s^2-m^2} \tag{6.25}$$

The reconstruction unit in the block diagram of the receiver in fig. 6-2 is a device which attempts to reduce the effects of the quantizer. The particular form of this device is based on a priori knowledge of the quantizer and the statistics of the quantizer input. This input is the difference signal, which has been found to have a laplacian distribution [1], described by

$$p(x)=\frac{1}{\sqrt{2}\,\sigma}\exp(-|x|/2\sigma) \tag{6.26}$$

The distribution of the difference signal for the "girl" image (see fig. 6-6) is shown in fig. 6-3. It can be seen that a laplacian could model this distribution quite well. However, it has been found that the difference signals of images are correlated, so that an accurate statistical

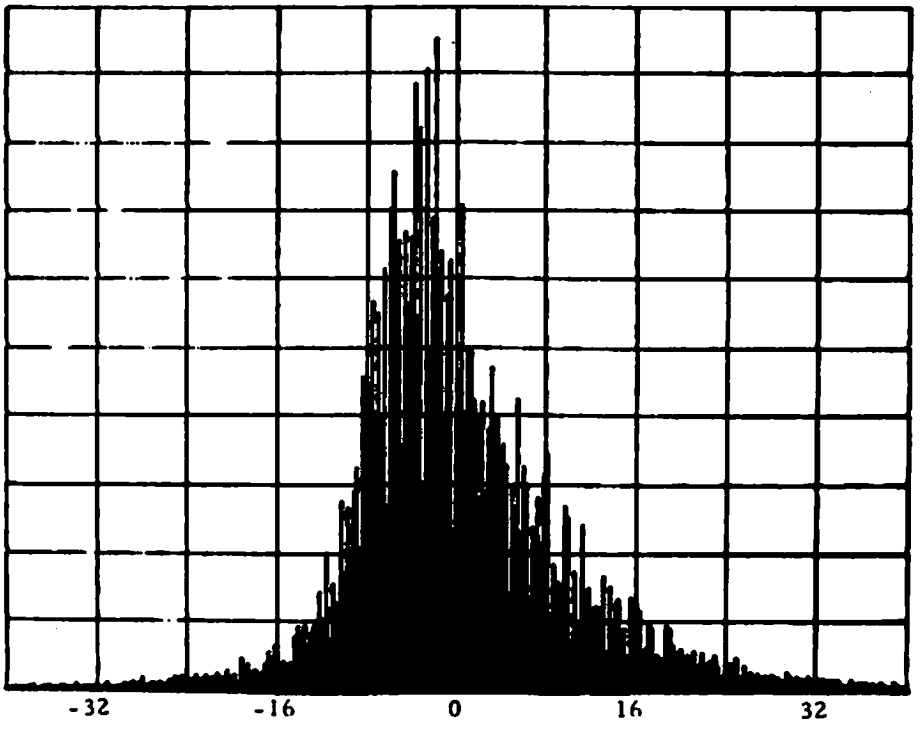


Figure 6-3. Histogram of the DPCM signal for the "girl" picture

representation of them must also account for this fact. Figure 6-4 shows the actual two-dimensional distribution of the DPCM coded "girl". For this image, the average correlation for adjacent difference samples has been measured as 0.4. These samples can then be modeled by a correlated two-dimensional laplacian density, written as

$$p(x,y)=\frac{1}{2\sigma_x\sigma_y\sqrt{1-r^2}}\exp\{-\frac{1}{\sqrt{2(1-r^2)}}(\left|\frac{ax}{\sigma_x}-\frac{by}{\sigma_y}\right|+\left|\frac{ay}{\sigma_y}-\frac{bx}{\sigma_x}\right|)\} \quad (6.27)$$

where

$$\sigma_x^2=E\{x^2\} \quad (6.28)$$

$$\sigma_y^2=E\{y^2\} \quad (6.29)$$

$$r=\frac{E\{xy\}}{\sigma_x\sigma_y} \quad (6.30)$$

$$a=\sqrt{1+r}+\sqrt{1-r} \quad (6.31)$$

$$b=\sqrt{1+r}-\sqrt{1-r} \quad (6.32)$$

Figure 6-5 contains a plot of this density function for $r=0.4$ and $\sigma_x=\sigma_y$.

The two-dimensional laplacian distribution is seen to provide an accurate model for the DPCM samples. These

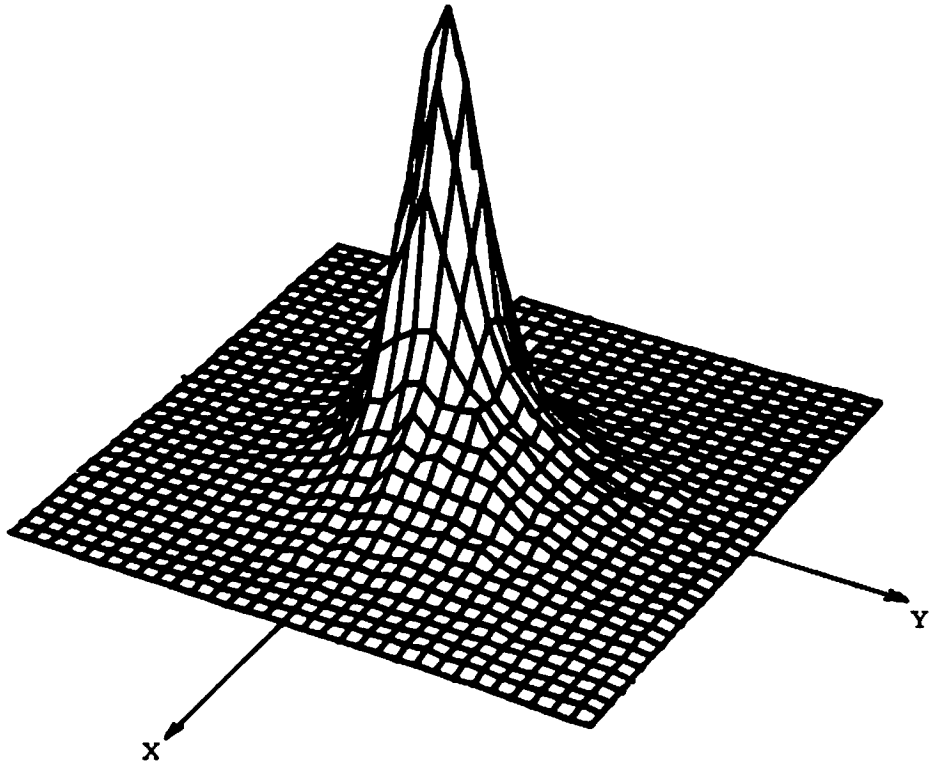


FIGURE 6-4. TWO-DIMENSIONAL HISTOGRAM OF THE DPCM CODED "GIRL" PICTURE

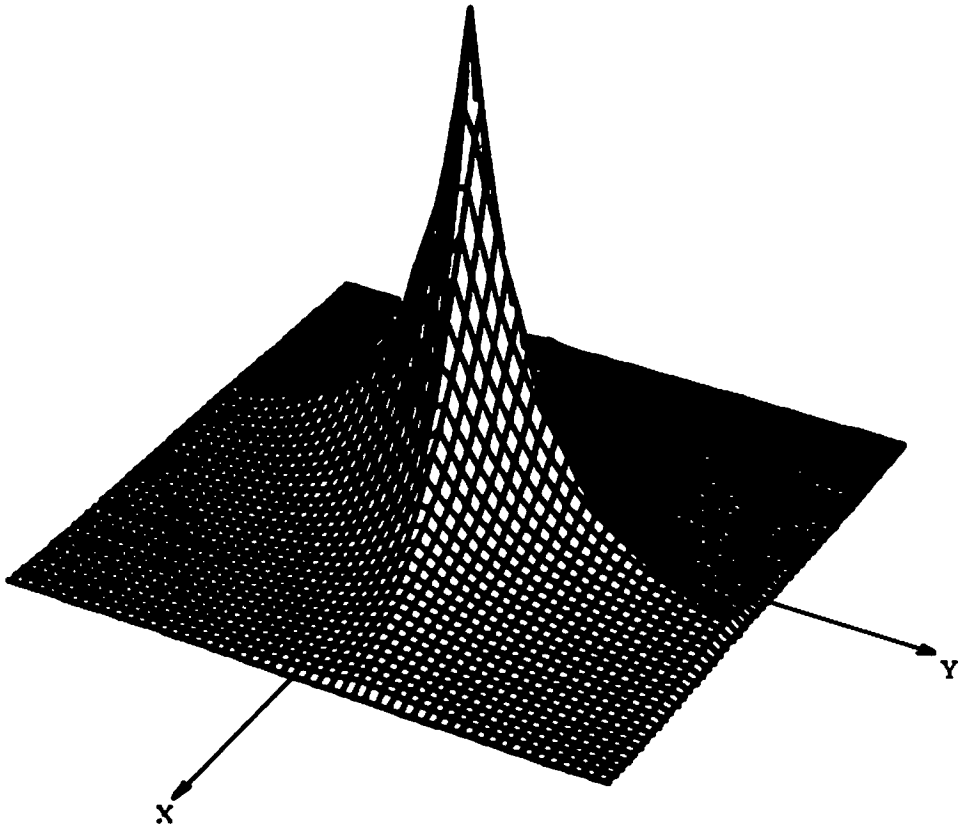


FIGURE 6-5. TWO-DIMENSIONAL LAPLACIAN DENSITY FUNCTION USED TO MODEL THE TWO-DIMENSIONAL DPCM SIGNAL SHOWN IN FIG. 6-4; CORRELATION =0.4

samples are quantized before they are transmitted through the channel. A minimum mean-square error restoration of the quantization can be obtained by utilizing equations 6.5 to 6.7. This solution is complicated, but can be simplified for the case of deltamodulation (one bit quantization). In this case x and y are each quantized to the intervals $[0,\infty)$ or $(-\infty,0)$, or equivalently, as positive or negative. The results are contained in equations 6.11 and 6.12.

Applying this restoration to the quantized image in fig. 6-6a results in a mean-square error reduction of 12%. Subjectively, as the restored image in fig. 6-6b shows, there is less apparent noise and more discernible detail. Equations 6.5 to 6.7 have also been utilized directly to restore samples quantized to general regions. Specifically, the results were applied to the two and three bit coded images shown in figures 6-6c and 6-6e, respectively. The reconstructed images in figures 6-6d and 6-6f exhibit both a reduction in mean-square error and a subjective improvement in quality. The subjective improvement is less apparent in these pictures, however, because the quantization itself is less noticeable. Table 6-1 shows the quantization intervals that were used to code these images [2]. The intervals were chosen to minimize the mean-square quantization error that occurs when an individual laplacian sample is quantized.

a. One bit encoded

b. One bit restored

c. Two bits encoded

d. Two bits restored

e. Three bits encoded

f. Three bits restored

Figure 6-6. Minimum mean-square error restoration of DPCM encoded images using two adjacent pixels.

TABLE 6-1

QUANTIZATION INTERVALS FOR SIGNALS WITH A LAPLACIAN DISTRIBUTION CHOSEN ACCORDING TO A MINIMUM MEAN-SQUARE ERROR CRITERION

Bits	Quantization Interval		Restoration Point
0	$-\infty$	∞	0.0
1	0.0	∞	0.707
2	0.0	1.102	0.395
	1.102	∞	1.810
3	0.0	0.504	0.222
	0.504	1.181	0.785
	1.181	2.285	1.576
	2.285	∞	2.994
4	0.0	0.266	0.126
	0.266	0.566	0.407
	0.566	0.910	0.726
	0.910	1.317	1.095
	1.317	1.821	1.540
	1.821	2.499	2.103
	2.499	3.605	2.895
	3.605	∞	4.316

The image reconstructions shown in fig. 6-6 applied the techniques summarized in equations 6.5 to 6.7 by using the quantized information about a previous sample to restore a current sample. By solving eq. 6.3 in three dimensions rather than in just two, quantized information about two other samples can be used in the reconstruction of a third quantized sample. The DPCM image samples are thus modelled by a correlated three-dimensional laplacian density function. A form of this three-dimensional solution was then employed to obtain the reconstructed image shown in fig. 6-7a. In this image, the sample immediately preceding, and the one immediately following, the sample being restored were used in the reconstruction. The resultant image has 39% lower mean-square error than the one-bit quantized image shown in fig. 6-6a. Figure 6-7b presents a similar result, except the samples immediately above and below the current sample were utilized in the reconstruction. In this case, the mean-square error improvement is 18%. Both image reconstructions also reveal a distinct subjective improvement--there is a reduction in visual noise in constant luminance areas of the image, together with a decrease in slope overload <*> at edges within the image.

Thus the technique described above provides an effective method for restoring DPCM coded images, particularly when the quantization is coarse and the DPCM

a. Restoration using three horizontal pixels

b. Restoration using three vertical pixels

Figure 6-7. Minimum mean -square error restorations of one bit (deltamodulation) encoded image using three adjacent pixels.

samples are correlated.

REFERENCES

1. J. B. O'Neal, Jr., "Predictive Quantizing Systems (Differential Pulse Code Modulation) for the Transmission of Television Signals," Bell System Technical Journal, Vol. 45, May/June 1966, pp. 689-721.

2. M. D. Paez and T. H. Glisson, "Minimum Mean-Squared-Error Quantization in Speech PCM and DPCM Systems," IEEE Transactions on Communications, April 1972, pp. 225-230.

<*> In DPCM image coding, slope overload arises whenever the DPCM system cannot follow video signals with fast rise-times, such as those that occur at the edges within a picture.

CHAPTER 7

QUANTIZATION AND RESTORATION OF RAYLEIGH SAMPLES

The last two chapters have shown that the nonlinear quantization restoration equations of Chapter 3 can be solved for both gaussian and laplacian quantizer inputs. This chapter extends these results to quantizer inputs which can be described by Rayleigh probability density functions. It will be shown that a Rayleigh distribution can accurately model the intensity distribution of an image, so that quantized images can be reconstructed by the techniques developed herein.

7.1 Rayleigh Densities in PCM Image Coding

In the PCM (pulse code modulation) coding of an image, the image is first sampled at an array of points known as pixels. The value assigned to each pixel is a measure of the light intensity at that point. Each light intensity is then quantized and coded, for either digital transmission or storage. Because it is proportional to the square of the magnitude of an electric field, light intensity is a real and non-negative quantity. The distribution of light intensities for an image can be characterized by a histogram ranging from black (zero) to white. For most natural images, there are many more dark pixels than bright

pixels, and the envelope of the histogram tends to fall off exponentially at higher brightness levels [1]. Image histograms measured by Kretzmer [2] and Stockham [3] exhibit these properties. Figure 7-1 shows that the histogram of the "girl" image, shown originally in fig. 5-7, also possesses these characteristics.

In the past, image intensity distributions have been well-modelled by Rayleigh, log-normal, and exponential probability density functions [4]. In this chapter, a variation of a Rayleigh distribution extended to many dimensions is utilized to model image intensities. A need for a multidimensional distribution arises because of the usefulness of including pixel-to-pixel correlations in the model. A correlated multidimensional Rayleigh distribution, derived in Chapter 4 as eq. 4.34, can be rewritten as

$$p(\underline{x}) = |\underline{C}|^{-1/2} \prod_{i=1}^{N} (\underline{h}_i^T \underline{x}) U(\underline{h}_i^T \underline{x}) \exp\{-\tfrac{1}{2} \underline{x}^T \underline{C}^{-1} \underline{x}\} \qquad (7.1)$$

where $\underline{C}$ is the correlation matrix of $\underline{x}$ and the vectors $\underline{h}_i$ are a function of the eigenvalues, λ_i, and eigenvectors, $\underline{e}_i$, of $\underline{C}$. If the eigenvectors are arranged column-wise into a matrix $\underline{E}$ having elements e_{ij}, then the components of the vectors $\underline{h}_i$ can be found from

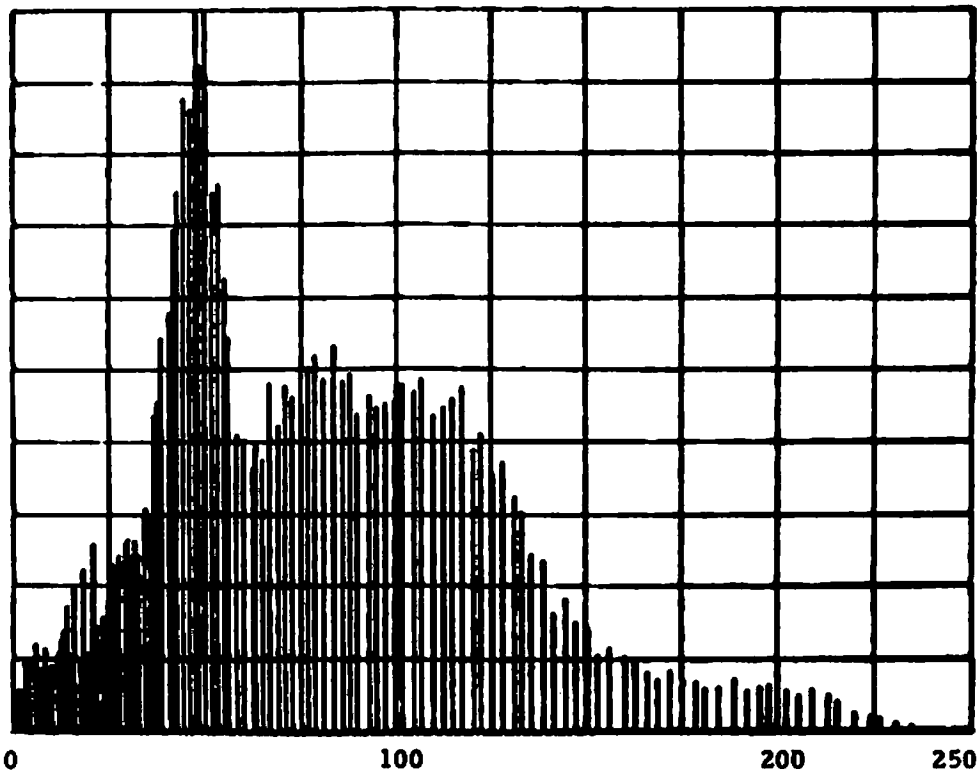


Figure 7-1. Histogram of the light intensities of the "girl" image

$$h_{ik} = \sum_{j=1}^{N} \frac{e_{ij} e_{kj}}{\sqrt{\lambda_j}} \qquad (7.2)$$

The average correlation between adjacent pixels for an image such as the "girl" has been measured to be Ø.95. A one-dimensional marginal probability distribution of eq. 7.1 for N=2 and a correlation of Ø.95 is shown in fig. 7-2 (an expression for this marginal distribution can be found in eq. 4.45). This distribution is seen to provide an accurate model for the intensity distribution of fig. 7-1. A two-dimensional histogram of the "girl" image, obtained by plotting pairs of adjacent pixel intensities, is shown in fig. 7-3. This histogram can be closely modelled by the correlated two-dimensional Rayleigh density function shown in fig. 7-4. (Only the positive quadrants of figures 7-3 and 7-4 are shown since both the histogram and its Rayleigh model are identically zero elsewhere.) This statistical model for image intensities can then be used for the restoration of quantized and coded values of these intensities.

7.2 Estimation of Quantized Rayleigh Samples

In order to obtain a restoration of quantized Rayleigh samples, such as those that result from the PCM coding of an image, eq. 3.6 must be solved using eq. 7.1 as a statistical model of the underlying Rayleigh random

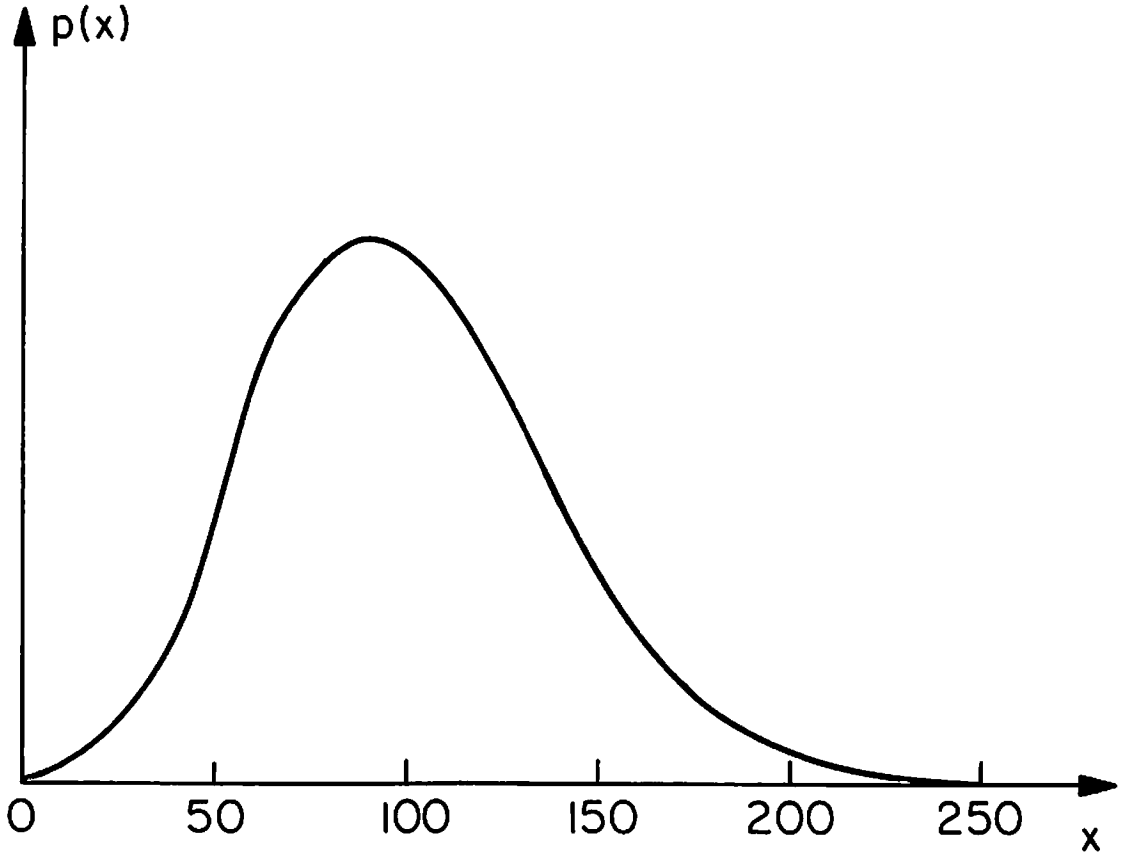


Figure 7-2. Marginal distribution of a two-dimensional Rayleigh probability density function with correlation factor = 0.95.

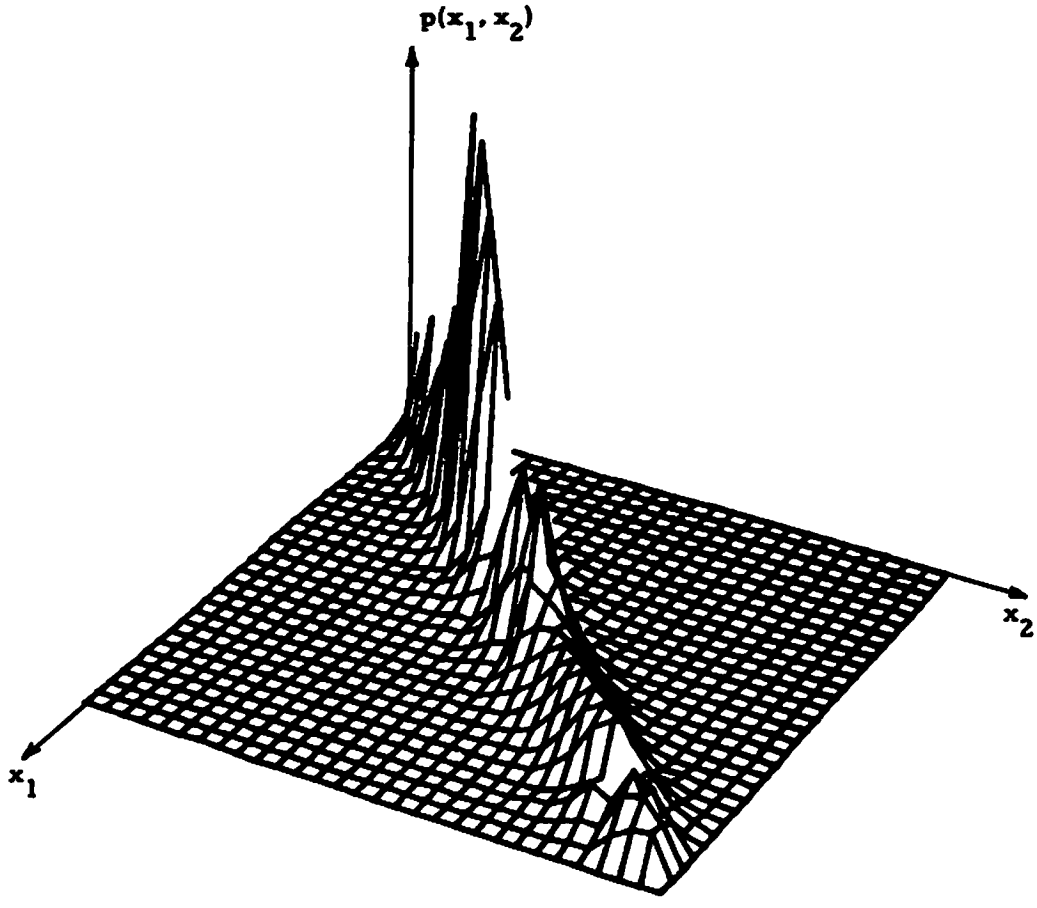


FIGURE 7-3. TWO-DIMENSIONAL HISTOGRAM OF LIGHT INTENSITIES FROM THE "GIRL" IMAGE

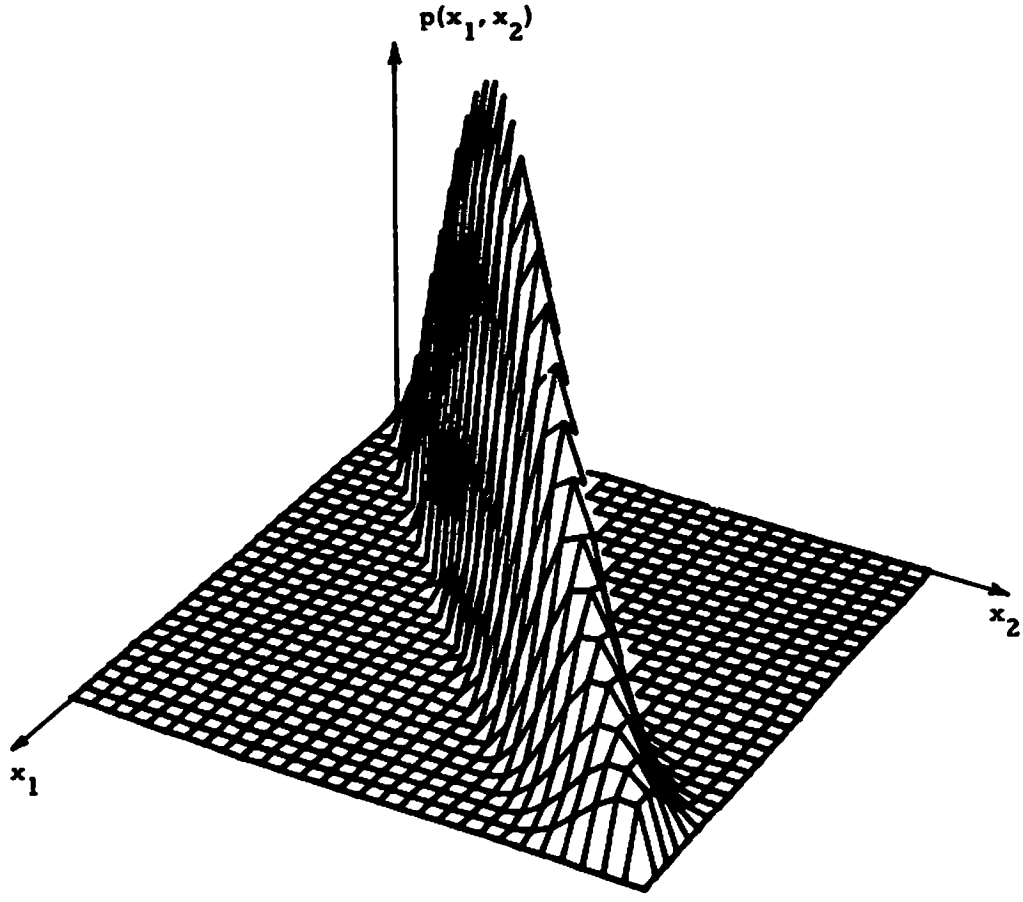


FIGURE 7-4. TWO-DIMENSIONAL CORRELATED RAYLEIGH PROBABILITY DENSITY FUNCTION USED TO MODEL THE INTENSITY DISTRIBUTION OF THE "GIRL" IMAGE; CORRELATION = 0.95

process. Unfortunately, an explicit solution to eq. 3.6 does not exist for this choice of input distribution, $p(\underline{x})$. However, solutions can be found for several special cases of quantization regions and correlation matrices. Also, an approximate solution can be obtained, using a method similar to Curry's [5], which is valid for very fine quantization.

The approximate solution can be found by means of a power-series expansion of $p(\underline{x})$. By retaining terms of the expansion up to fourth-order, the approximation is accurate enough to restore the outputs of a wide variety of quantizers. A sufficient condition for the approximation to hold is that the size of the quantization intervals must be small compared to the variance of a component quantized to one of these intervals. A multidimensional Taylor series is used to expand $p(\underline{x})$ about the midpoint, $\underline{g}$, of a rectangular quantization region, R. Let the region be defined as

$$R=\{\underline{x} \mid \underline{x} \in [\underline{a},\underline{b}]\} \tag{7.3}$$

where $\underline{a}$, $\underline{b}$, and $\underline{x}$ are N-vectors, so that its midpoint is

$$\underline{g}=(\underline{a}+\underline{b})/2 \tag{7.4}$$

The size of the quantization region can be expressed in

terms of a vector of half-widths as

$$\underline{d}=(\underline{b}-\underline{a})/2 \tag{7.5}$$

The Taylor-series expansion of $p(\underline{x})$ about $\underline{g}$ is then

$$p(\underline{x})=p(\underline{g})+\left[\frac{\partial p(\underline{x})}{\partial \underline{x}}\bigg|_{\underline{g}}\right]^T(\underline{x}-\underline{g})+\frac{1}{2}(\underline{x}-\underline{g})^T\frac{\partial^2 p(\underline{x})}{\partial \underline{x}\partial \underline{x}^T}\bigg|_{\underline{g}}(\underline{x}-\underline{g})$$

$$+\frac{1}{6}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}\frac{\partial^3 p(\underline{x})}{\partial x_i \partial x_j \partial x_k}\bigg|_{\underline{g}}(x_i-g_i)(x_j-g_j)(x_k-g_k)+\ \ldots \tag{7.6}$$

where $p(\underline{x})$ is given by eq. 7.1. The terms of this expansion are evaluated individually as follows:

$$\frac{\partial p(\underline{x})}{\partial \underline{x}}=p(\underline{x})\left\{\sum_{i=1}^{N}\underline{h}_i\left[\frac{1}{\underline{h}_i^T\underline{x}}+\frac{\delta(\underline{h}_i^T\underline{x})}{U(\underline{h}_i^T\underline{x})}\right]-\underline{C}^{-1}\underline{x}\right\} \tag{7.7}$$

where $\delta(\cdot)$ represents the Dirac delta function, defined as the derivative of the unit step function $U(\cdot)$, and

$$\underline{h}_i^T=[h_{i1}\quad h_{i2}\ \ldots\ h_{iN}] \tag{7.8}$$

This vector represents the (i)th row of the matrix $\underline{H}$. Next, define the vector $\underline{w}$ such that its components are

$$w_i = \frac{1}{\underline{h}_i^T \underline{g}} + \frac{\delta(\underline{h}_i^T \underline{g})}{U(\underline{h}_i^T \underline{g})} \tag{7.9}$$

so that eq. 7.7 can be rewritten as

$$\left.\frac{\partial p(\underline{x})}{\partial \underline{x}}\right|_{\underline{g}} = p(\underline{g})(\underline{H}^T \underline{w} - \underline{C}^{-1} \underline{g}) \tag{7.10}$$

The second derivative can be computed in a similar manner as

$$\left.\frac{\partial}{\partial \underline{x}}\left[\frac{\partial p(\underline{x})}{\partial \underline{x}}\right]^T\right|_{\underline{g}} = p(\underline{g})\left[(\underline{H}^T \underline{w} - \underline{C}^{-1} \underline{g})(\underline{H}^T \underline{w} - \underline{C}^{-1} \underline{g})^T - \underline{C}^{-1} - \underline{H}^T \underline{W} \underline{H}\right] \tag{7.11}$$

where $\underline{W}$ is a diagonal matrix having components

$$W_{ii} = \frac{1}{(\underline{h}_i^T \underline{g})} - \frac{\delta'(\underline{h}_i^T \underline{g})}{U(\underline{h}_i^T \underline{g})} + \frac{\delta(\underline{h}_i^T \underline{g}) U(\underline{h}_i^T \underline{g})}{[U(\underline{h}_i^T \underline{g})]^2} \tag{7.12}$$

Now let $\underline{z} = \underline{x} - \underline{g}$. Then the denominator of the restoration equation (eq. 3.6), after retaining the first three terms of the expansion, becomes

$$\int_{\underline{a}}^{\underline{b}} p(\underline{x})\,d\underline{x} \approx \int_{-\underline{d}}^{\underline{d}} p(\underline{g})\{1+(\underline{H}^T\underline{w}-\underline{C}^{-1}\underline{g})^T\underline{z}$$

$$+\tfrac{1}{2}\underline{z}^T[(\underline{H}^T\underline{w}-\underline{C}^{-1}\underline{g})(\underline{H}^T\underline{w}-\underline{C}^{-1}\underline{g})^T-\underline{C}^{-1}-\underline{H}^T\underline{W}\underline{H}]\underline{z}\}d\underline{z}$$

$$=p(\underline{g})V(1+\tfrac{1}{2}\sum_{i=1}^{N} B_{ii}\frac{d_i^2}{3}) \tag{7.13}$$

where

$$\underline{B}=[(\underline{H}^T\underline{w}-\underline{C}^{-1}\underline{g})(\underline{H}^T\underline{w}-\underline{C}^{-1}\underline{g})^T-\underline{C}^{-1}-\underline{H}^T\underline{W}\underline{H}] \tag{7.14}$$

and

$$V=\prod_{i=1}^{N}(2d_i) \tag{7.15}$$

The quantity V is the volume of the rectangular quantization region R, and $\underline{B}$ is a matrix based on parameters of this quantization region.

Next, the numerator of eq. 3.6 can be approximated by

$$\int_{\underline{a}}^{\underline{b}} \underline{x} p(\underline{x}) d\underline{x} = \int_{-\underline{d}}^{\underline{d}} (\underline{g}+\underline{z}) p(\underline{g}+\underline{z}) d\underline{z}$$

$$\approx \underline{g} \int_{\underline{a}}^{\underline{b}} p(\underline{x}) d\underline{x} + \int_{-\underline{d}}^{\underline{d}} \underline{z} p(\underline{g}) \{1 + (\underline{H}^T \underline{w} - \underline{C}^{-1} \underline{g})^T \underline{z} + \tfrac{1}{2} \underline{z}^T \underline{B} \underline{z}\} d\underline{z}$$

$$\approx p(\underline{g}) V [1 + \tfrac{1}{2} \sum_{i=1}^{N} B_{ii} \frac{d_i^2}{3} + \frac{1}{3} (\underline{H}^T \underline{w} - \underline{C}^{-1} \underline{g}) \underline{D}] \qquad (7.16)$$

where the diagonal matrix $\underline{D}$ has components

$$D_{ii} = d_i^2 \qquad (7.17)$$

Thus, the conditional mean estimate of a vector $\underline{x}$, quantized to the region R, is

$$\underline{y} = \underline{g} + \frac{(\underline{H}^T \underline{w} - \underline{C}^{-1} \underline{g}) \underline{D}}{3 + \frac{1}{2} \mathrm{Tr}\{\underline{B}\underline{D}\}} \qquad (7.18)$$

This estimate is seen to consist of a correction term added to the midpoint of the quantization region. Since the estimate is a function of a particular quantization region, the estimation process is a nonlinear operation. The estimation results are valid for

$$2d_i < s_i \qquad (7.19)$$

for i=1,2,...,N, where s_i^2 is the variance of the (i)th component of $\underline{x}$. It has been experimentally determined that an image must be coded with more than four bits per pixel for the solution given by eq. 7.18 to be useful.

For very coarse quantization, the relation in eq. 7.19 is not satisfied, so another solution to eq. 3.6 must be obtained which is valid for this situation. A solution can be found by utilizing a method developed in Chapter 5. The essence of this method is that the components of a quantized vector are restored individually, based on--

1. the quantization interval of the component being restored;
2. estimates of the remaining quantized components (not the quantization intervals of the remaining components).

The initial estimates can be chosen to be the midpoints of the quantization intervals. Alternately, they can be chosen as the solution to the one-dimensional version of eq. 3.6, which can be solved exactly. This latter choice is the more accurate one, and is the choice that will be made here. As new estimates of the components are found using the method of Chapter 5, they are used to replace the initial estimates. The technique is thus iterative.

Keeping these concepts in mind, consider a typical component of eq. 3.6. This component can be solved

approximately by

$$y_i = \frac{\int_{a_i}^{b_i} x_i\, p(\underline{x})\, d\underline{x}}{\int_{a_i}^{b_i} p(\underline{x})\, d\underline{x}}$$

$$\approx \frac{\int_{a_i}^{b_i} x_i\, p(t_1, t_2, \ldots, t_{i-1}, x_i, t_{i+1}, \ldots, t_N)\, dx_i}{\int_{a_i}^{b_i} p(t_1, t_2, \ldots, t_{i-1}, x_i, t_{i+1}, \ldots, t_N)\, dx_i} \qquad (7.20)$$

where

$$t_k = \frac{\int_{a_k}^{b_k} x_k\, p(x_k)\, dx_k}{\int_{a_k}^{b_k} p(x_k)\, dx_k} \qquad (7.21)$$

for $k \neq i$. A straightforward integration of this last equation yields

$$t_k = \frac{\sqrt{\frac{\pi}{2}}\, s_k \left(\operatorname{erf} \frac{b_k}{\sqrt{2} s_k} - \operatorname{erf} \frac{a_k}{\sqrt{2} s_k} \right) - \left(b \exp\left\{ -\frac{b_k^2}{2 s_k^2} \right\} - a \exp\left\{ -\frac{a_k^2}{2 s_k^2} \right\} \right)}{\exp\left\{ -\frac{a_k^2}{2 s_k^2} \right\} - \exp\left\{ -\frac{b_k^2}{2 s_k^2} \right\}} \qquad (7.22)$$

where s_k^2 is the variance of x_k. Unfortunately, eq. 7.20 cannot be integrated so easily. When the expression for $p(\underline{x})$, as defined by eq. 7.1, is substituted into this equation, the estimate becomes

$$y_i = \frac{\int_{a_i}^{b_i} x_i \prod_{k=1}^{N} \underline{h}_i^T \underline{t} U(\underline{h}_i^T \underline{t}) \exp\{-\frac{1}{2}\underline{t}^T \underline{C}^{-1} \underline{t}\} dx_i}{\int_{a_i}^{b_i} \prod_{k=1}^{N} \underline{h}_i^T \underline{t} U(\underline{h}_i^T \underline{t}) \exp\{-\frac{1}{2}\underline{t}^T \underline{C}^{-1} \underline{t}\} dx_i} \qquad (7.23)$$

where

$$\underline{t}^T = [t_1 \quad t_2 \quad \cdots \quad t_{i-1} \quad x_i \quad t_{i+1} \quad \cdots \quad t_N] \qquad (7.24)$$

Proceeding as in equations 5.16 to 5.19, common factors can be cancelled from both the numerator and denominator, leaving

$$y_i = \frac{\int_{a_i}^{b_i} x_i \prod_{k=1}^{N} (h_{ki}x_i + H_k) U(h_{ki}x_i + H_k) \exp\{\frac{-1}{2r_{ii}}(r_{ii}x_i + R_i)^2\} dx_i}{\int_{a_i}^{b_i} \prod_{k=1}^{N} (h_{ki}x_i + H_k) U(h_{ki}x_i + H_k) \exp\{\frac{-1}{2r_{ii}}(r_{ii}x_i + R_i)^2\} dx_i} \qquad (7.25)$$

where r_{ij} is the (i,j)th element of $\underline{C}^{-1}$,

$$H_k = \sum_{\substack{j=1 \\ j\neq i}}^{N} h_{kj}\, t_j \tag{7.26}$$

and

$$R_i = \sum_{\substack{j=1 \\ j\neq i}}^{N} r_{ij}\, t_j \tag{7.27}$$

Making the substitution

$$z = \frac{1}{\sqrt{2r_{ii}}}\left(r_{ii}\, x_i + R_i\right) \tag{7.28}$$

yields

$$y_i = \frac{\displaystyle\int_u^v \left(\sqrt{2r_{ii}}\, z - R_i\right) \prod_{k=1}^{N} \left[\frac{h_{ki}}{r_{ii}}\left(\sqrt{2r_{ii}}\, z - R_i\right) + H_k\right] \exp\{-z^2\}\, dz}{r_{ii} \displaystyle\int_u^v \prod_{k=1}^{N} \left[\frac{h_{ki}}{r_{ii}}\left(\sqrt{2r_{ii}}\, z - R_i\right) + H_k\right] \exp\{-z^2\}\, dz} \tag{7.29}$$

where

$$u = \max_{\substack{k \\ h_{ki}>0}} \left\{ \frac{1}{\sqrt{2r_{ii}}}\left(r_{ii}\, a_i + R_i\right),\ \frac{R_i}{\sqrt{2r_{ii}}} - \frac{H_k}{h_{ki}}\sqrt{\frac{r_{ii}}{2}} \right\} \tag{7.30a}$$

$$v = \min_{\substack{k \\ h_{ki}<0}} \left\{ \frac{1}{\sqrt{2r_{ii}}}\left(r_{ii}\, b_i + R_i\right),\ \frac{R_i}{\sqrt{2r_{ii}}} - \frac{H_k}{h_{ki}}\sqrt{\frac{r_{ii}}{2}} \right\} \tag{7.30b}$$

The expressions which multiply the exponential terms in eq. 7.29 are seen to be Nth-order polynomials. Writing

them as such, reveals that

$$y_i = \frac{\sqrt{2}\int_u^v z\sum_{k=0}^{N} g_k z^k \exp\{-z^2\}dz}{\sqrt{r_{ii}}\int_u^v \sum_{k=0}^{N} g_k z^k \exp\{-z^2\}dz} - \frac{R_i}{r_{ii}} \qquad (7.31)$$

The integrations can finally be performed to yield <*,**>

$$\text{num.} = \sqrt{2}\left[\sqrt{\pi}\,\text{erf}(z)\sum_{k=1}^{\lfloor\frac{N+1}{2}\rfloor}\frac{g_{2k-1}(2k-1)!!}{2^k} - \exp\{-z^2\}\sum_{k=0}^{N} g_k k!!\sum_{j=0}^{\lfloor\frac{k}{2}\rfloor}\frac{z^{k-2j}}{2^j(k-2j)!!}\right]_u^v \qquad (7.32a)$$

$$\text{denom.} = \sqrt{r_{ii}}\left[\sqrt{\pi}\,\text{erf}(z)\sum_{k=0}^{\lfloor\frac{N}{2}\rfloor}\frac{g_{2k}(2k-1)!!}{2^k} - \exp\{-z^2\}\sum_{k=1}^{N} g_k (k-1)!!\sum_{j=0}^{\lfloor\frac{k-1}{2}\rfloor}\frac{z^{k-2j-1}}{2^j(k-2j-1)!!}\right]_u^v \qquad (7.32b)$$

so that

<*> The notation $\lfloor x \rfloor$ denotes the smallest integer $\leq x$.

<**> The notation k!! denotes the product

$$k!! = \begin{cases} k\cdot(k-2)\cdot\cdot\cdot 3\cdot 1 & \text{for k odd} \\ k\cdot(k-2)\cdot\cdot\cdot 4\cdot 2 & \text{for k even} \end{cases}$$

$$y_i = \frac{\text{num.}}{\text{denom.}} - \frac{R_i}{r_{ii}} \qquad (7.33)$$

When $\underline{C}=\underline{I}$ or, equivalently, when x_i is being restored independently of the other components of $\underline{x}$, the estimate becomes $y_i=t_i$, as given by eq. 7.22. When $\underline{C}=\underline{I}$, the estimate in eq. 7.33 is used in the recursive manner described previously.

7.3 Error Covariance of Rayleigh Estimator

The estimate given by eq. 7.33 provides a minimum mean-square error reconstruction of a quantized Rayleigh variable. The performance of this estimate can be determined by a calculation of the estimation covariance. A general expression for this covariance was derived as eq. 3.10. This expression can be solved, for a Rayleigh probability density function, by the same technique that was used in equations 5.36 to 5.37 for a gaussian distribution. Denoting the (i,i)th component of the estimation covariance (the only non-zero component because of the recursive nature of the estimation procedure described in Sec. 7.2) as e_{ii}, then

$$e_{ii}=\frac{\int_{a_i}^{b_i} x_i \; p(\underline{t})\,dx_i}{\int_{a_i}^{b_i} p(\underline{t})\,dx_i} - y_i^{\,2} \tag{7.34}$$

or

$$e_{ii}=\frac{e_{num}}{e_{denom}} - \left(y_i+\frac{R_i}{r_{ii}}\right)^2 \tag{7.35}$$

where

$$e_{num}= 2\left[\sqrt{\pi}\,\mathrm{erf}(z)\sum_{k=1}^{\left\lfloor\frac{N+2}{2}\right\rfloor}\frac{g_{2k-2}(2k-1)!!}{2^k}\right.$$

$$\left.-\exp\{-z^2\}\sum_{k=\emptyset}^{N} g_k(k+1)!!\sum_{j=\emptyset}^{\left\lfloor\frac{k+1}{2}\right\rfloor}\frac{z^{k-2j+1}}{2^j(k-2j+1)!!}\right]_u^v \tag{7.36}$$

$$e_{denom}= r_{ii}\left[\sqrt{\pi}\,\mathrm{erf}(z)\sum_{k=\emptyset}^{\left\lfloor\frac{N}{2}\right\rfloor}\frac{g_{2k}(2k-1)!!}{2^k}\right.$$

$$\left.-\exp\{-z^2\}\sum_{k=1}^{N} g_k(k-1)!!\sum_{j=\emptyset}^{\left\lfloor\frac{k-1}{2}\right\rfloor}\frac{z^{k-2j-1}}{2^j(k-2j-1)!!}\right]_u^v \tag{7.37}$$

This error covariance for the (i)th component is related to the estimates of the other components, t_k for $k\neq i$, by means of both R_i, as given by eq. 7.27, and H_k for $k=1,2,\ldots,N$,

as given by eq. 7.26. The expression in eq. 7.35 hence specifies the conditional covariance of y_i, conditioned on the other estimates, t_k.

Because the components, t_k, are random variables, the conditioning can be removed by integrating eq. 7.35 with respect to the distribution of the t_k. Equivalently, the dependence on the t_k can be removed by integrating with respect to R_i and H_k, which are functions of the t_k. The first and second moments of R_i and H_k can be calculated as

$$E\{R_i\} = \sum_{\substack{j=1 \\ j\neq i}}^{N} r_{ij} E\{t_j\} = \sqrt{\frac{\pi}{2}} \sum_{\substack{j=1 \\ j\neq i}}^{N} r_{ij} \tag{7.38}$$

$$E\{R_i^2\} = -r_{ii} \sum_{\substack{j=1 \\ j\neq i}}^{N} r_{ij} C_{ji} \tag{7.39}$$

$$E\{H_k\} = \sqrt{\frac{\pi}{2}} \sum_{\substack{j=1 \\ j\neq i}}^{N} h_{kj} \qquad k=1,2,\ldots,N \tag{7.40}$$

$$E\{H_k^2\} = \sum_{\substack{n=1 \\ n\neq i}}^{N} h_{kn} \sum_{\substack{j=1 \\ j\neq i}}^{N} C_{nj} h_{kj} \qquad k=1,2,\ldots,N \tag{7.41}$$

It can be shown, by means of lengthy derivations, that

$$E\{R_i\} \propto E\{H_k\} \tag{7.42}$$

and

$$E\{R_i^2\} \propto E\{H_k^2\} \tag{7.43}$$

for k=1,2,...,N. Since both R_i and H_k consist of a sum of many random variables, it can be argued that they have a gaussian distribution. Then, under this assumption, the two moments found in equations 7.38 and 7.39 are sufficient to characterize this distribution, now denoted as $p(R_i)$. Hence, the unconditioned error covariance can be found from

$$e = \int e_{ii} p(R_i) dR_i \tag{7.44}$$

However, due to the complicated form of the constituents of this expression, the error covariance cannot be evaluated for general cases. Rather, specific cases of the estimation covariance must be considered on an individual basis. An approximation can then be obtained numerically.

7.4 Simulation and Restoration of PCM Coded Images

To determine the utility and effectiveness of the restoration technique developed in Sec. 7.2, the technique was computer-simulated and then applied to images which had been quantized according to various bit assignments. The images were quantized in intensity using Max decision and restoration levels. These levels are shown in table 7-1. They were calculated by means of eq. 2.3, using a

TABLE 7-1

QUANTIZATION INTERVALS FOR SIGNALS WITH A RAYLEIGH DISTRIBUTION CHOSEN ACCORDING TO A MINIMUM MEAN-SQUARE ERROR CRITERION

Bits	Quantization Interval		Restoration Point
0	0.0	∞	1.253
1	0.0	1.375	0.829
	1.375	∞	1.920
2	0.0	0.822	0.529
	0.822	1.420	1.114
	1.420	2.127	1.725
	2.127	∞	2.529
3	0.0	0.499	0.329
	0.499	0.825	0.670
	0.825	1.135	0.980
	1.135	1.453	1.290
	1.453	1.800	1.617
	1.800	2.208	1.984
	2.208	2.760	2.433
	2.760	∞	3.086

one-dimensional Rayleigh probability density function. Each pixel was quantized individually, i.e., independently of its neighboring pixels. The results of this quantizing scheme, when applied to the "girl" image, are shown in figures 7-5a and 7-5c for one bit per pixel and two bits per pixel quantizers, respectively.

Figures 7-5a and 7-5c were a posteriori restored by utilizing eq. 7.33 with N=5. Using this value for N means that the restoration of one pixel is achieved by utilizing the information that is available about four other pixels. A five-dimensional correlated Rayleigh probability density function is thus used to model the image for this restoration procedure. The five pixels were chosen and ordered as shown in fig. 7-6. These particular pixels allow maximum use to be made of the correlation between adjacent pixels. The 5x5 correlation matrix of these pixels is also shown in fig. 7-6. This correlation matrix is a necessary input to the restoration equation.

The restoration provides an improvement because it permits many possible output intensity levels to occur, even though, as in the one bit case, there are only two input levels <*>. The restored images for one bit and for two bits are shown in figures 7-5b and 7-5d, respectively. The restoration of the one bit quantized image results in a decrease of 20.8% in mean-square error. The restored two

a. One bit/pixel PCM

b. Restored one bit/pixel

c. Two bits/pixel PCM

d. Restored two bits/pixel

Figure 7-5. Minimum mean-square error reconstructions of PCM coded images.

a. Pixels chosen for the restoration of pixel no. 1

1.00	0.97	0.96	0.97	0.96
0.97	1.00	0.94	0.93	0.95
0.96	0.94	1.00	0.95	0.90
0.97	0.93	0.95	1.00	0.94
0.96	0.95	0.90	0.94	1.00

b. Correlation matrix of the ordered pixels

Figure 7-6. Typical ordering of five PCM coded pixels for restoration, and their corresponding correlation matrix.

bit image has 28.6% lower mean-square error than its quantized version. Each restored image exhibits a subtle, but noticeable, visual improvement over its corresponding quantized image. The restored versions appear subjectively to be more "real" due to the extra intensity levels that result from the restoration. It should be emphasized that these improvements were obtained a posteriori, utilizing only the correlation matrix of a typical sampled image, together with the quantized image to be restored.

<*> For this one bit case, since each of five quantized pixels can be in one of two possible states, there are 32 possible output values for the restored pixel.

REFERENCES

1. W. K. Pratt, Digital Image Processing, chap. 7, to be published.

2. E. R. Kretzmer, "Statistics of Television Signals," Bell System Technical Journal, vol. 31, July 1952, pp. 751-763.

3. T. G. Stockham, Jr., "Image Processing in the Context of a Visual Model," Proceedings of the IEEE, vol. 60, July 1972, pp. 828-842.

4. Pratt, chap. 1.

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

CHAPTER 8

RESTORATION OF BINARY SYMMETRIC CHANNEL ERRORS

The previous chapters have presented and analyzed techniques for restoring the output of a quantizer so that the result more accurately matches the quantizer's input. The restorations are based essentially upon exact knowledge of the quantizer output. A similar, but more difficult problem results when the quantizer output is not known exactly. This could occur, for example, when the quantizer output is transmitted over a noisy channel. The first section in this chapter explores the effect of channel errors on the restorations derived previously. The next section examines a technique that statistically compensates for the effect of channel errors.

8.1 Effects of Channel Errors on Quantized Signals

In this analysis, channel errors are assumed to arise in the context of a binary symmetric channel (BSC) [1]. The characteristics of this type of channel are shown in fig. 8-1. The channel is discrete and memoryless and can be specified by a transition probability assignment $P(j|k)$, for $j,k=0,1$, as

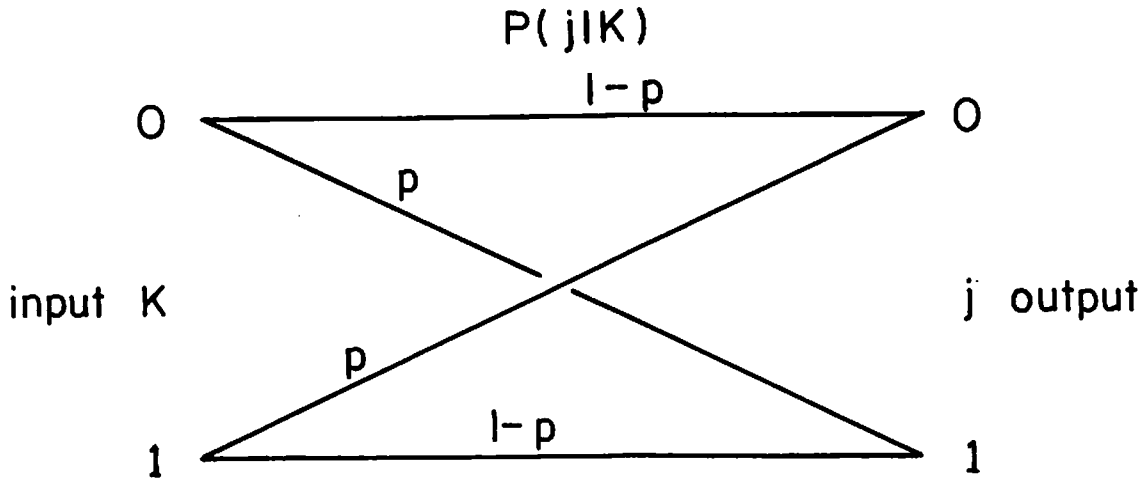


Figure 8-1. Transition probabilities for a binary symmetric channel.

$$\underline{P}=\{P(j|k)\}=\begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix} \tag{8.1}$$

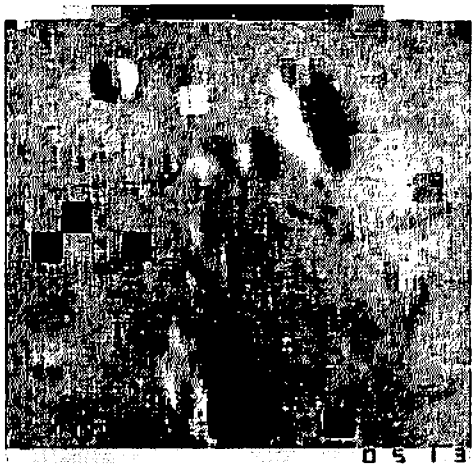
Since the channel is memoryless, the probability of an output sequence $\underline{z}=(z_1,z_2,\ldots,z_N)$, given an input sequence $\underline{x}=(x_1,x_2,\ldots,x_N)$, is given by

$$P(\underline{z}|\underline{x})=\prod_{i=1}^{N} P(z_i|x_i) \tag{8.2}$$

Based on this definition, a BSC was simulated by means of a computer, with the channel error probability, p, chosen to be 0.01. The simulated channel was then applied to transform coded images. The three "original" images shown in fig. 5-7 were zonal transform coded in 16x16 blocks, as described in Sec. 5.4. The quantized transform domain components were encoded by assigning each a binary code word. The resulting sequence of binary digits was operated on by the simulated channel. The error-corrupted bit stream was then either decoded directly, as shown in figures 8-2a, 8-2c, and 8-2e, or restored by the techniques of Chapter 5 to reduce the effects of the quantization process (see fig. 8-3 for a schematic of this procedure). The decoded images with the quantization effects reduced are shown in figures 8-2b, 8-2d, and 8-2f.

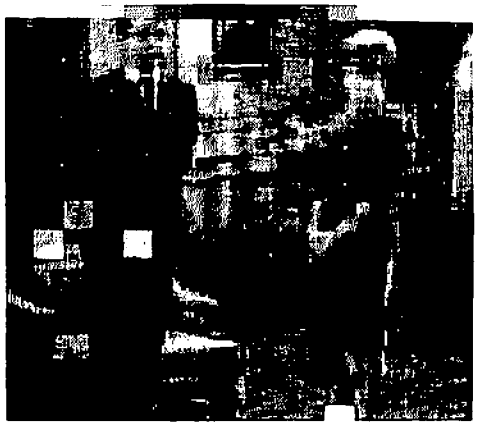
a. Max quantized

b. Restored



c. Max quantized

d. Restored



e. Max quantized

f. Restored

Figure 8-2. Minimum mean-square error restoration of Haar transformed 0.5 bit zonal quantized images transmitted through a BSC with an error probability = 0.01.

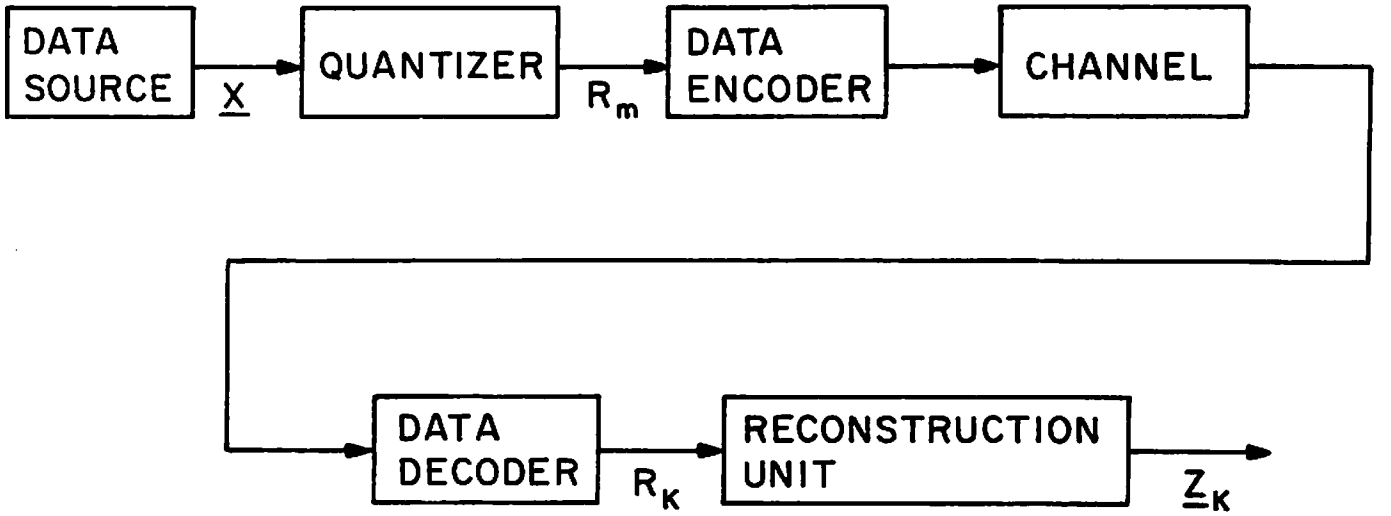


Figure 8-3. Data system used to model the effects of channel errors on the quantization restoration process.

Bit errors in transform coding that arise due to a binary symmetric channel are seen to result in an emphasis of the block structure and a subjective error that extends over the entire block. This latter effect occurs because inverse transforming a block containing an error distributes this error over all the resultant image domain components. However, what is important to note from fig. 8-2 is that channel errors affect quantized and quantization-restored images identically. The reconstruction techniques derived in Chapters 5, 6, and 7 are thus insensitive to channel errors. Since they provide visual and mean-square error improvements in noise-free cases, they can be utilized equally well in noisy environments.

8.2 Reconstruction of Quantized and Transmitted Signals

The previous section demonstrated that channel errors do not adversely affect the performance of the restoration techniques derived earlier. However, these techniques do nothing to ameliorate the effects of the channel errors. This is because the fundamental restoration equation presented in eq. 3.6 was derived without any consideration of channel structure. By including the channel structure in the derivation, the resultant restoration technique can simultaneously reduce the effects of the quantization process and mitigate the effects of channel errors.

The output of a data source (this output could consist of DPCM samples, PCM samples, or transform domain samples) is denoted by $\underline{x}=(x_1,x_2,\ldots,x_N)$ and described by a probability density function $p(\underline{x})$. The reconstruction of $\underline{x}$, after $\underline{x}$ has been quantized to one of M regions and channel-error corrupted, is denoted by $\underline{z}_k=(z_1,z_2,\ldots,z_N)_k$ for k=1,2,...,M (refer to fig. 8-3). The mean-square error that results from this process is

$$\mathcal{E}=\sum_{k=1}^{M}\sum_{m=1}^{M}P(m|k)\int_{R_m}(\underline{x}-\underline{z}_k)(\underline{x}-\underline{z}_k)^T p(\underline{x})\,d\underline{x} \tag{8.3}$$

This error can be minimized by proper choice of the restoration points, $\underline{z}_k$. Setting the partial derivatives of this error with respect to $\underline{z}_k$ equal to zero yields

$$\underline{z}_k=\frac{\sum_{m=1}^{M}P(m|k)\int_{R_m}\underline{x}p(\underline{x})\,d\underline{x}}{\sum_{m=1}^{M}P(m|k)\int_{R_m}p(\underline{x})\,d\underline{x}} \tag{8.4}$$

for k=1,2,...,M. This expression is the noisy channel version of eq. 3.6 and provides a minimum mean-square error estimate of the input to a quantizer based on the output of a noisy channel, the characteristics of the quantizer, and the a priori statistics of the input. This equation is

also a multidimensional version of a result first derived in [2]. For a noiseless channel, the channel matrix $\underline{P}$ becomes the identity matrix and eq. 8.4 reduces to eq. 3.6. When the probability volume integrals in the denominator of eq. 8.4 are all equal, which is approximately true for Max quantization, the restoration equation simplifies to

$$\underline{z}_k = \sum_{m=1}^{M} P(m|k) \frac{\int_{R_m} \underline{x} p(\underline{x}) d\underline{x}}{\int_{R_m} p(\underline{x}) d\underline{x}} \tag{8.5}$$

or

$$\underline{z}_k = \sum_{m=1}^{M} P(m|k) \underline{y}_m \tag{8.6}$$

where $\underline{y}_m$ is given by eq. 3.6. This result holds for maximum output entropy quantizers and two-level symmetrical quantizers, and is approximately correct for many other types.

A signal that has been quantized and then transmitted over a noisy channel can thus be optimally restored by utilizing eq. 8.4. The restoration solutions found earlier for gaussian, laplacian, and Rayleigh probability density functions (see equations 5.23, 6.7, and 7.33, respectively) can be substituted directly into eq. 8.4 once the transition matrix for the channel has been determined. The

resultant estimator can then be used to restore the outputs of transform, DPCM, and PCM coders that have been degraded by channel errors.

REFERENCES

1. R. G. Gallager, Information Theory and Reliable Communications, John Wiley and Sons, Inc., New York, 1968, p. 73.

2. A. J. Kurtenbach and P. A. Wintz, "Quantizing for Noisy Channels," IEEE Transactions on Communication Technology, vol. COM-17, April 1969, pp. 291-302.

CHAPTER 9

CONCLUSIONS, AND TOPICS FOR FUTURE RESEARCH

This dissertation has described a means for reconstructing quantized signals according to a minimum mean-square error criterion. The method, which is entirely a posteriori, is based on a priori statistical information about the original, unquantized signals. In a broad sense, the reconstruction technique described is applicable to any coding system for which the necessary a priori statistical information is available, since any coding system can be considered as a special case of a quantizer. In other words, assigning one of a finite number of code words to a random variable is equivalent to quantizing that variable to one of a finite number of intervals. Further, block encoding a string of random variables is the same as quantizing a random vector to a generalized region in space.

No attempt has been made in this work to optimize the location or choice of these generalized quantization regions, but only to utilize arbitrary, given regions in an optimal manner to obtain a signal reconstruction. Finding an optimum set of regions to vector-quantize a string of random variables remains an unexplored area, but one in which fruitful research can be conducted.

The techniques described herein, in all cases, achieved a reduction in mean-square quantization error when compared to a Max restoration scheme. The reason for this improvement over a method which purportedly provides the "minimum mean-square error," is that Max's scheme is memoryless while the restoration described herein requires memory. The memory requirement arises in the form of knowledge of the quantized samples that surround a sample being restored. This added information, which must be stored in memory, permits the improvement in mean-square error. The restoration techniques derived in this dissertation, therefore, represent a generalization of Max's restoration results to many variables.

From an information theory standpoint, this use of memory allows a coding system to operate closer to the rate distortion theory bound. Also utilizing memory to perform source-encoding--equivalent to vector quantization, as discussed previously--can provide an even closer approach to this bound. However, in this work a memoryless encoding scheme has been employed with its implicit assumption of rectangular quantization regions.

The reconstruction technique developed herein has specifically been applied to three common types of image coding systems--transform coders, DPCM coders, and PCM coders. These coding systems have been statistically

characterized by gaussian, laplacian, and Rayleigh probability density functions, respectively. However, to completely characterize the coders requires correlated multidimensional versions of these density functions. Heretofore, multidimensional versions have not existed for either laplacian or Rayleigh distributions. This deficiency motivated the general technique derived in Chapter 4 for generating correlated multidimensional density functions from desired marginal distributions and correlation functions. In essence, this allows the correlation which exists between the variables in many coding systems to be utilized in a statistical restoration of their coded outputs.

A determination of the minimum mean-square error restoration point for a quantized vector has often proven difficult, because a multidimensional integration of a complicated probability function is required. However, this difficulty has been surmounted by a novel, recursive approach. This approach permits a reduction to only a single integration which can then be evaluated analytically. The solutions which resulted from this approach have been applied to quantized images, and only one recursion of the technique has been found to be necessary for an image restoration. The restoration procedure can thus be readily implemented.

A decrease in mean-square error was obtained in all simulations of this restoration procedure. However, a corresponding subjective improvement was not always observed. This ancillary result substantiates an observation made by many others: lower mean-square error does not always correspond to a visual improvement. Repeating the image restorations in the context of a weighted error criterion, with the weighting chosen according to characteristics of the human visual system, resulted in both an analytical and a subjective improvement. It is suggested that further image coding reconstructions be performed with respect to a weighted error criterion.

A fundamental limitation imposed on the images restored in this dissertation was an assumption of statistical stationarity. This restriction is very basic, because the reconstruction techniques herein are completely dependent on the choice of a statistical model for the underlying random process. Since images have been found to be inherently nonstationary, an assumption of stationarity limits the reconstruction performance. An adaptive technique could remove the limitation and improve the performance. This has not yet been investigated, but could prove to be a productive area of research.

Another problem that was discussed only briefly in

this work concerns the reconstruction of quantized signals in the presence of channel noise. This is a much more difficult problem than noise-free quantization restoration and warrants further study.