DESIGN OF A REAL-TIME FACSIMILE TRANSMISSION SYSTEM

by

KENNETH PAUL WACKS

S.B., Massachusetts Institute of Technology
(1970)

S.M., Massachusetts Institute of Technology
(1970)

E.E., Massachusetts Institute of Technology
(1970)

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August, 1973

Signature of Author _____
Department of Electrical Engineering, August 10, 1973

Certified by _____
Thesis Supervisor

Accepted by _____
Chairman, Departmental Committee on Graduate Students

DESIGN OF A REAL-TIME FACSIMILE TRANSMISSION SYSTEM

by

KENNETH PAUL WACKS

Submitted to the Department of Electrical Engineering on August 10, 1973, in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

ABSTRACT

A feasible and economical system for transmitting a monochrome picture in real-time is developed. The design is constrained to use special purpose hardware and optics in place of a general purpose computer. Picture-dependent processing is eliminated, and the system is adapted to commercial facsimile devices.

The design is evolved from an existing system that requires extensive data processing resources. The transition to the real-time system is examined in great detail. Some of the issues involved include the design of optical anti-aliasing filters, companding, quantizing a non-uniform distribution of data, pseudorandom noise, and the design of one-dimensional noise filters to replace two-dimensional versions.

All theoretical work is supplemented by viewing pictures produced from simulations of the components on a PDP-9 computer. The performance of the system when corrupted by channel noise is investigated. Finally, the transmission at various data compression rates is presented.

THESIS SUPERVISOR: William F. Schreiber
TITLE: Professor of Electrical Engineering

ACKNOWLEDGMENTS

During my graduate studies at M.I.T. I was fortunate to have Professor William F. Schreiber as my research advisor. I wish to thank him for these years of wise counsel and warm friendship.

I appreciate the guidance and suggestions of my thesis readers, Professor Thomas S. Huang and Professor Oleh J. Tretiak.

The assistance provided by numerous students and employees of the Research Laboratory of Electronics at M.I.T. is gratefully acknowledged. This research was supported in part by grants from the Associated Press and the National Institutes of Health.

I would like to extend special thanks to my mother, the proofreader and typist, for infinite patience and expert performance, and to my family for constant support throughout my education.

TABLE OF CONTENTS

LIST OF FIGURES

DESIGN OF A REAL-TIME FACSIMILE TRANSMISSION SYSTEM

Chapter 1

INTRODUCTION

During the past decade proposals for image transmission systems have proliferated. The goal of most of these systems, in the terminology of digital communications, has been to minimize the total number of bits needed for faithful reproduction of a transmitted picture. To this end many highly sophisticated proposals have been developed. The degree of complexity, however, is highly correlated to the cost of simulating such systems and subsequently constructing specific hardware.

Most simulation studies have required a large computer with a fast central processor, abundant main storage, and usually many peripheral storage devices. In addition, those systems that greatly compress the data while maintaining quality have used picture-dependent coding schemes. Certain features of the picture are analyzed, and based on this the particular coding structure is determined. This necessitates storing as much as the entire image while coding, and hinders real-time processing and transmission[39]. This thesis considers an image transmission system with some of these limitations and develops a nearly equivalent system which is practical for real-

time implementation.

In previous research by the author[50] a system of moderate complexity for sending monochrome pictures was developed. It was simulated on an IBM 360/65 and required about 580,000 bytes of main memory, two magnetic tapes, and a disk or drum. Picture dependence, as discussed above, was minor. Processing of a single picture to simulate transmission and reception required a total of 15 minutes on the average. A compression ratio of at least two compared to a slightly modified PCM system was achieved.

This system is described in detail to reveal the processing delays. Each component is analyzed to determine whether real-time implementation will compromise its function. Considerable effort is devoted to eliminating the need for a general purpose computer. Instead, processing can be performed optically or in specially designed hardware employing both analog and digital circuitry. In the latter case our goal is to minimize the amount of computation and memory needed. The initial coding in most image transmission systems consists of line-by-line scanning to produce a stream of data. Direct processing of this data without significant intermediate storage will help meet the stated goals. The source and sink for this stream will be a practical scanner and receiver, such as the instruments used in commercial facsimile systems.

The next chapter presents our design philosophy for picture coding systems. As a new system is evolved, it is important to have

guidelines for evaluating and comparing the performance of various components. The application of these principles to the current design problem is stressed in subsequent chapters.

In Chapter 3 we examine the significant features of various stages in the original system. The transition to the real-time system is then outlined along with methods for simulating the results on a PDP-9 computer. Chapters 4 through 8 explore these issues in detail. The topics in Chapter 4 are the design of two-dimensional anti-aliasing filters with finite duration impulse response, and the effect of logarithm domain processing on such filters. Next, the quantization of a non-uniform data distribution is analyzed statistically to develop a scheme acceptable for picture data. Chapter 6 considers methods for generating pseudorandom noise that is useful for masking undesirable artifacts introduced by quantizing. The visual impact of structured noise as a substitute for white noise is then examined analytically and by simulation. The major barrier to stream processing the data is approached in Chapter 8 where one-dimensional filters are designed to effectively substitute for two-dimensional versions. With the real-time system now specified, the performance in the presence of transmission errors is illustrated in Chapter 9.

Chapter 2

DESIGN PHILOSOPHY

When designing an image transmission system, we are guided primarily by the results from communication theory. Modern communica- theory enables us to divide the problem into tractable subunits, and it provides statistical methods for achieving theoretical limits of performance. The systems presented here are designed for sending pictures to be viewed by humans. Therefore, we seek to maximize subjective quality within the constraints of the transmission environment. As a result, designs may be generated more efficiently by considering the psychophysics of the human visual system along with communication theory. We will now consider some general aspects of communication theory versus psychophysical issues.

In most communications situations there are two limitations on transmission accuracy[52]. One is noise, where this term refers to all disturbances between the transmitter and the receiver. The other is the mismatch between the information source and the communication link. Fortunately the interplay between the limitations of the communication medium and the characteristics of the data to be transmitted can be handled independently with a digital system. Digital processing also affords the practical advantages of greater noise immunity than analog processing, and permits easy manipulation with

computers and digital circuits.

Figure 1 presents an outline of a general digital communication system[10]. The source in the current application is the picture to be sent, while the destination is the reproduced image to be viewed. The channel represents the bottleneck of the system. Here noise makes a match between the receiver input and the transmitter output impossible. This apparent limitation on transmission accuracy can be removed at the expense of a restriction on the data transmission rate. This is the remarkable conclusion of Shannon's theorem[10]. He found that the probability of transmission error, P(e), can be made arbitrarily small provided the rate, R, at which bits enter the channel encoder is less than channel capacity, C:

$$P(e) \propto \exp\left[-TE(R)\right]$$

T is the coding time and E(R) is a decreasing function of R which reaches 0 when R = C.

$$C = \max_{\substack{p(x_i) \\ i=1,\ldots,N}} I(x;y)$$

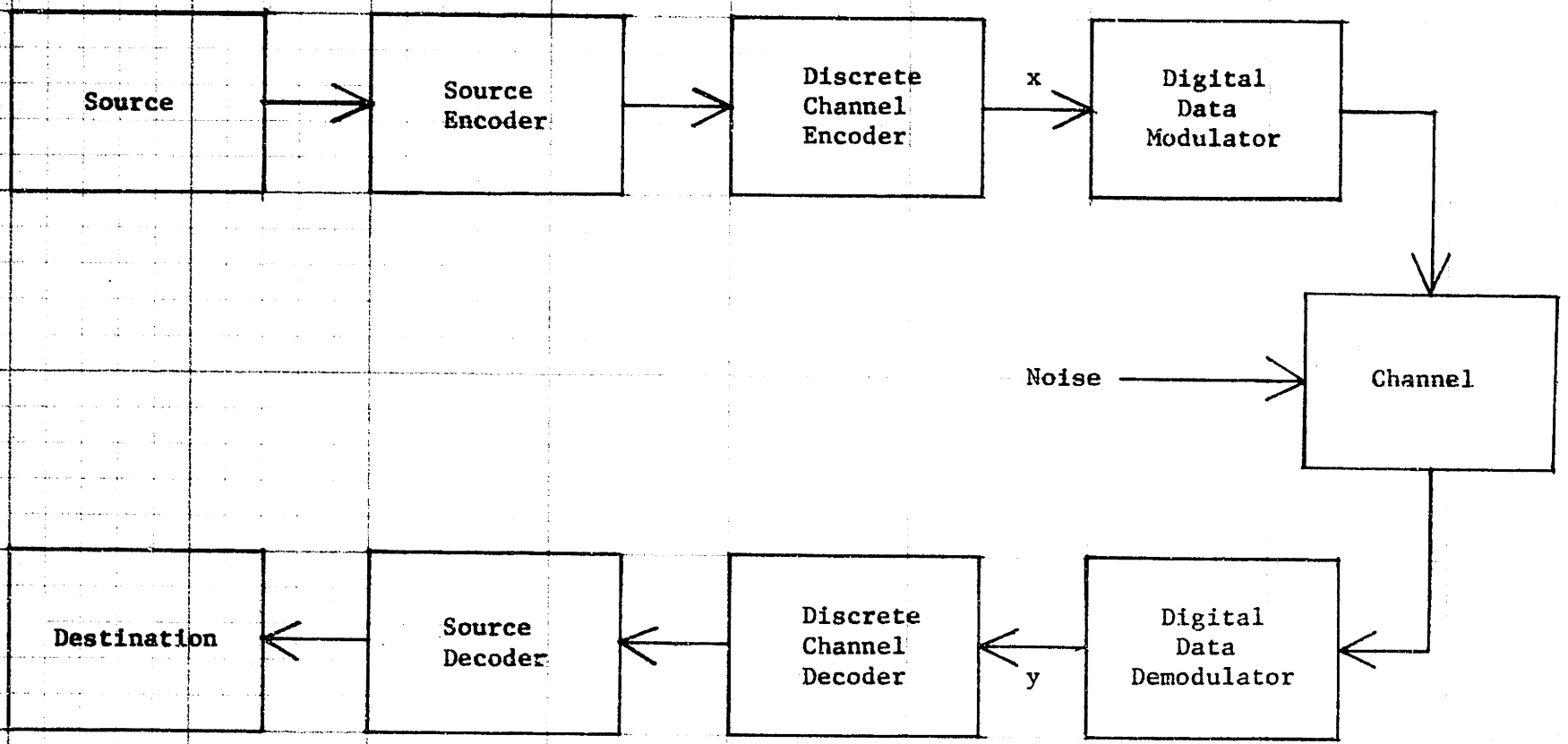
$p(x_i)$ is the probability of the ith channel input symbol

y is the channel output symbol

I is the mutual information

(See Figure 1 for the locations of x and y)

C can be interpreted as a measure of the maximum average amount of uncertainty about the channel input resolved at the receiver by a



The input and output of the channel are continuous time functions.

Digital Communication System

Figure 1

single transmission. Thus accurate transmission of bits entering the channel encoder at a rate less than C is solely a function of the design of the channel encoder and decoder. Defects in this design cause visible flaws in the received picture. This issue of channel noise will be considered in Chapter 9. We now assume a correct design of these units and so must decide how to make best use of at most CT bits to transmit a picture in T seconds.

As is common in most digital image coding schemes, the picture is initially scanned at a high enough rate (say N samples per picture height and width) so no resolution is lost (twice the highest spatial frequency), and each sample is quantized to a sufficient number of bits so an apparently continuous tone scale can be reproduced without coding (usually 6 or 7 bits per sample are needed). This total number of bits must be reduced to CT for reliable transmission. In this way any degradations in the image at the receiver can be controlled by the design of the source coder and source decoder, independent of the channel.

According to information theory the optimum source encoder produces an ensemble of outputs for different inputs such that each of the 2^{CT} possible outputs is equally likely[10]. This maximizes the source entropy:

$$H(x) = \sum_{\substack{\text{all } x \\ \text{possible}}} p(x) \log \frac{1}{p(x)}$$

p(x) is the probability of input x.

This approach requires extensive statistical analysis of all possible input pictures, which is not feasible. Alternatively, we can consider the goal of the communication system in terms of the ultimate destination, the observer. Rather than recreate the image point-by-point as accurately as possible, data are transmitted which will create an image that appears subjectively similar to the original.

By coding to this fidelity criterion, preferential treatment is accorded to certain classes of picture data. The advantages of combining visual analysis with numerical measure will be illustrated as the system is developed.

Chapter 3

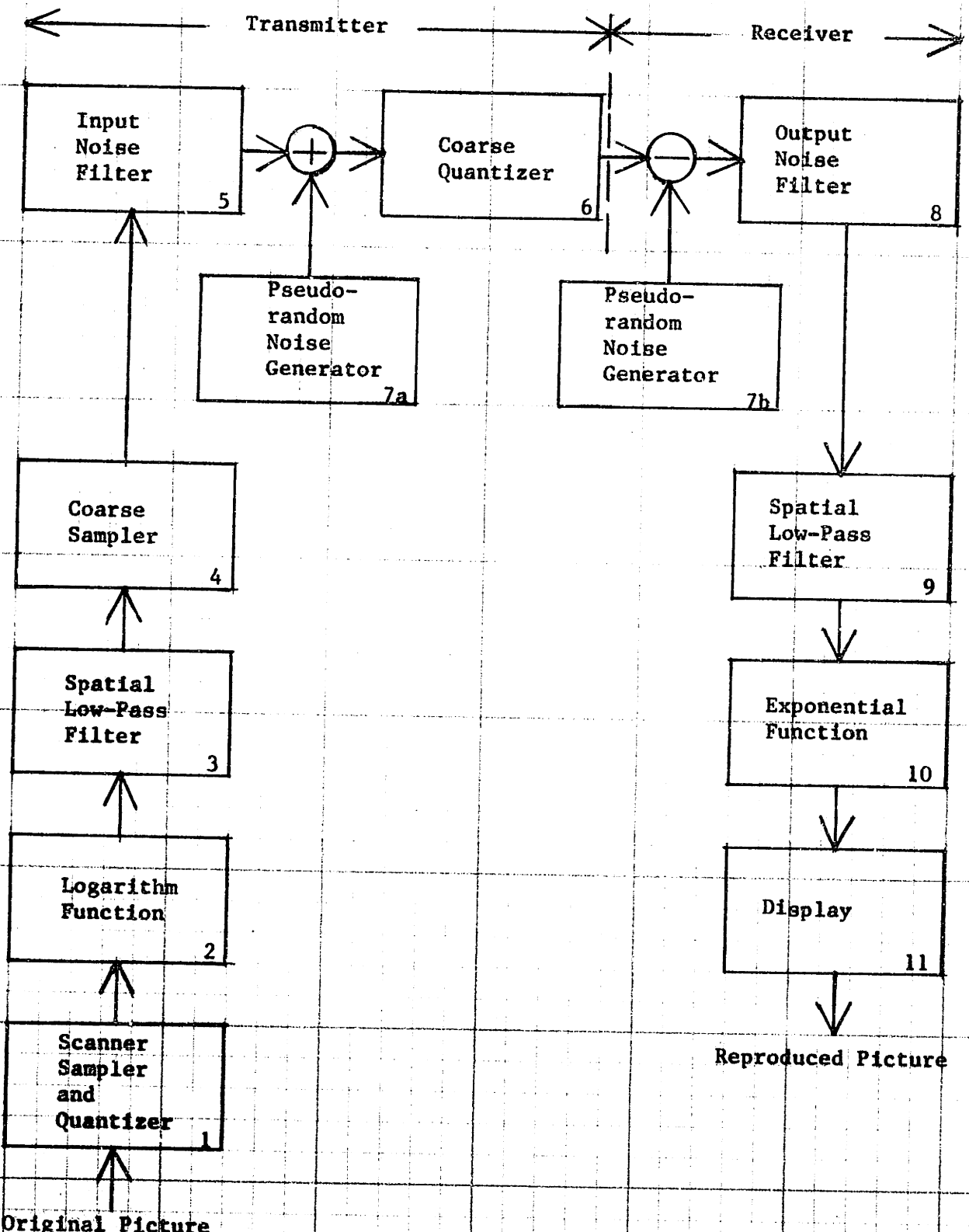
SYSTEM DESCRIPTION

A block diagram of the original system is presented in Figure 2. Blocks 1 through 7a (refer to the figure for the block numbers) constitute the source coder at the transmitter while the remaining blocks represent the source decoder at the receiver. The assumption made in directly connecting the transmitter to the receiver is that a channel of sufficient capacity together with an appropriate channel encoder and decoder are available to convey the coded digital data with no error. The output of blocks 1 and 10 is a square array of N x N samples of the original picture. Each sample is quantized to one of 2^B levels so the entire picture is represented by a total of N x N x B bits. The following stages of the transmitter compress this data to a square array of S x S samples (S<N) with each sample quantized to 2^Q bits (Q<B). Thus CT = S x S x Q bits are transmitted. The choice of S versus Q for a given CT was determined by optimizing the subjective quality of the reproduction.

Significant features of the various blocks are described in the following paragraphs. Further details may be found in the previous research report[50].

Scanner and Display (blocks 1 and 11)--

These functions are embodied in a laboratory device built by



The Original System

Figure 2

the Cognitive Information Processing Group (CIPG) at M.I.T. Data are sampled and quantized as described above by employing a flying-spot scanner controlled by a digital system. The data are written out and read from magnetic tape in a format compatible with the IBM 360 computer. An example of a scanned picture is presented in Figure 5(a) (N = 256, B = 8).

Spatial Low-pass Filters and Coarse Sampler (blocks 3, 4, and 9)--

In order to extract an array of S x S points from the original N x N values, the sampling theorem requires that the data first be band-limited to at most S/2 cycles per picture height and width; otherwise aliasing would occur. The use of appropriate sharp cutoff filters, however, introduces undesirable artifacts, termed "ringing". This is evident in Figure 5(b) where S = 64. This ringing is a visual display of the Gibbs phenomenon. Filters were designed as a compromise between this psychophysical sensitivity and system theory requirements. At the transmitter is a Gaussian low-pass filter:

$$H(u,v) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{u^2 + v^2}{2\sigma^2}\right]$$

where σ was chosen so that 90% of the energy of the impulse response of the filter is at spatial frequencies less than the cutoff frequency of the corresponding sharp cutoff filter. The coarse sampler uses an algorithm where each retained point is an average of all nearest neighbors, each weighted according to its proximity to the retained point. The spatial low-pass filter in block 9 consists of a linear

interpolator. The net result is shown in Figure 5(c).

Sample Quantizer (block 6)--

The most direct method for reducing the number of bits per sample from B to Q is to set the B-Q least significant bits to zero. This method is used for pulse code modulation (PCM) transmission. Figure 6(a) was processed in this way to reduce the number of bits from 8 to 3. The spurious plateaus of brightness commonly known as quantization noise, were eliminated with a scheme developed by Roberts[37]. Pseudorandom noise with amplitude uniformly distributed between ± ½ of a coarse quantization level is added to the data before the coarse quantizer and the same signal is subtracted at the receiver. This effectively converts the quantization noise to less visible "snow" over the entire picture and makes the quantization noise independent of the picture data, as can be seen in Figure 6(b).

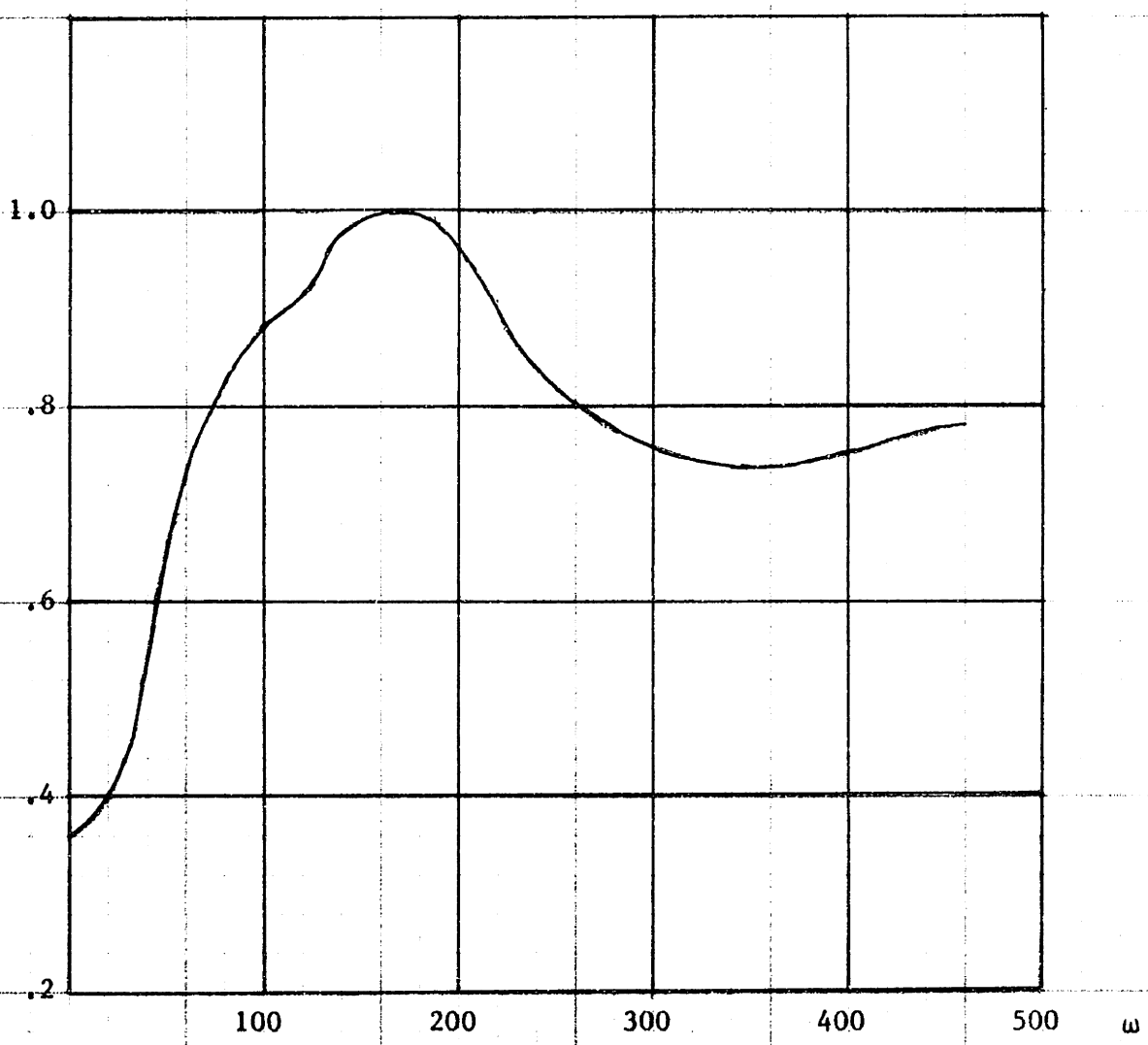
Logarithm and Exponential Functions (blocks 2 and 10)--

Psychophysical experiments have been performed to determine the relationship between the sensitivity of the eye and the incoming intensity. When the eye is adapted to the local luminance level, the perceived intensity is proportional to the logarithm of the actual intensity, as stated in the Weber-Fechner law[6]. This implies that when viewing a photograph under usual conditions, the eye is more sensitive to details in the darker areas than in lighter regions. Preprocessing the image data by the logarithm function (termed companding) thus effectively distributes errors introduced in later

stages (quantization and digital filtering noise, for example) uniformly over the perceived brightness range. The exponential function restores the image to the original tone scale. Figure 6(c) demonstrates this by processing Figure 6(b) in the logarithm domain.

Noise Filters (blocks 5 and 8)--

In order to reduce the visibility of the noise (the snow) the psychophysics of vision was again considered. Figure 3 shows the frequency response of the human visual system, determined[28] by measuring the threshold of visibility of a sine wave grating. Based on the works of Graham[16] and Post[35], filters were constructed to alter the originally flat spectrum of the noise in order to concentrate the noise in those frequency regimes to which the eye is relatively insensitive, namely, low frequencies. As illustrated in Figure 4 the input filter, which is the inverse of the output filter, pre-emphasizes those frequency components in the picture which are reduced by the output filters, so the spectrum of the image data is unchanged. These filter designs agree with those derived by Stockham[46] directly from psychophysical experiments. Figure 6(d) shows the results of combining all the components, except the sample reduction, for transmission at 3 bits per sample. This illustrates that whenever some degradation is unavoidable, the specific type introduced critically affects the perceived quality of the picture. This scheme exchanges artifacts in the form of ringing and contours for the more acceptable ones of blurriness and snow.

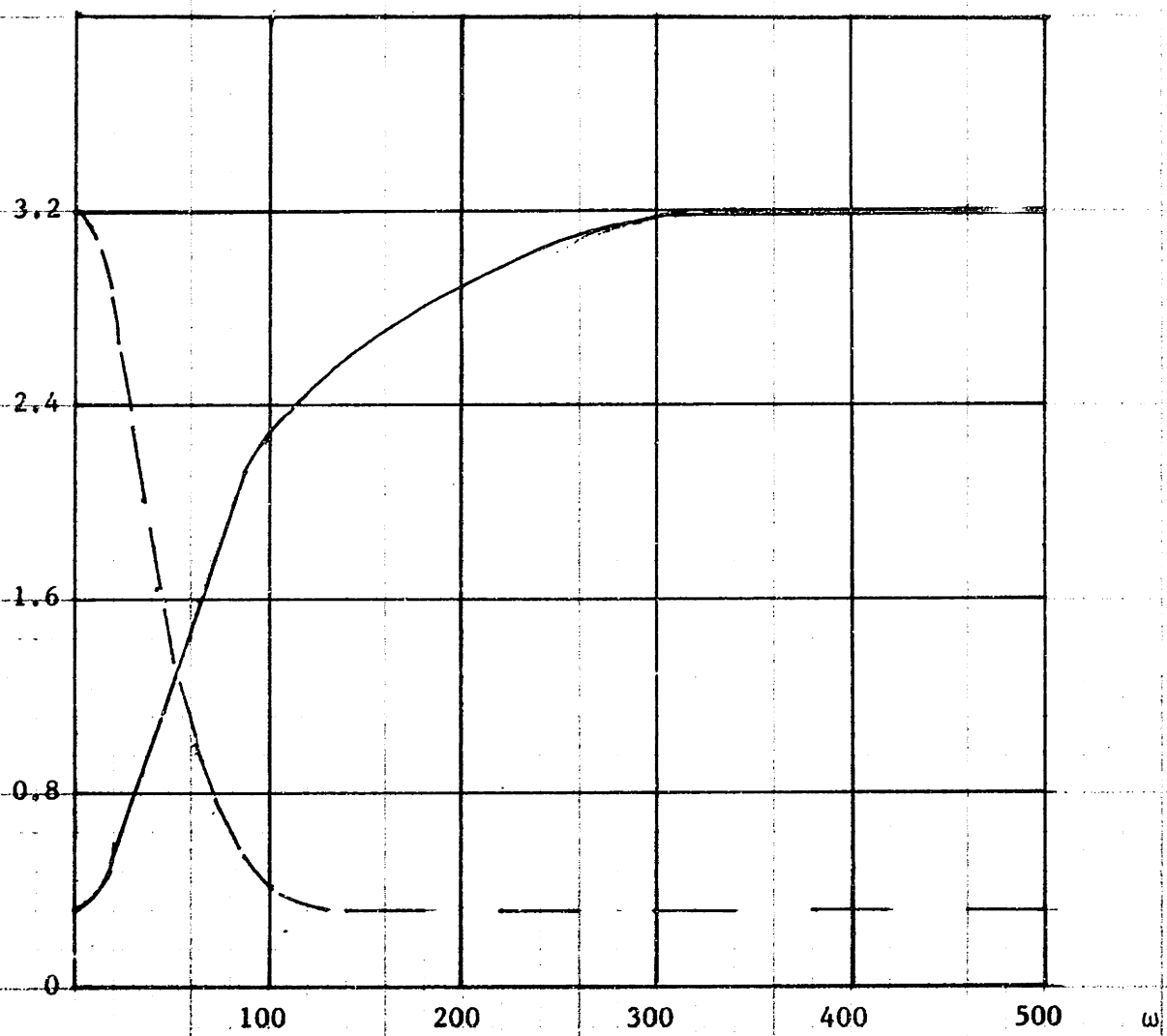


ω: Spatial frequency= $2\pi x$ cycles per picture height

Sine Wave Amplitude Response of the Visual System

at a Distance 6 Times the Picture Height

Figure 3



ω: radians/picture height; $\omega = 2\pi\sqrt{u^2+v^2}$

____ Input noise filter (H_a)

---- Output noise filter (H_b)

$$H_a(u,v) = 3.20 - 0.934 \exp\left(-\pi^2\left(\frac{u^2+v^2}{1445}\right)\right) - 1.96 \exp\left(-\pi^2\left(\frac{u^2+v^2}{3819}\right)\right)$$

$$H_b(u,v) = 0.3125 + 2.77 \exp\left(-\pi^2\left(\frac{u^2+v^2}{445}\right)\right) + 0.164 \exp\left(-\pi^2\left(\frac{u^2+v^2}{5120}\right)\right)$$

Frequency Response of the Input Noise Filter

and the Output Noise Filter

Figure 4

(a)

(b)

(c)

Effects of Coarse Sampling Filters

Figure 5

(a)

(b)

(c)

(d)

Coarse Quantization Techniques

Figure 6

In order to achieve the goal of an economically feasible facsimile transmission system, a revision of the original system, as shown in Figure 7, was designed. The following paragraphs outline the motivations for the various changes and the methods used to simulate the components on a PDP-9 computer. The details of the development and step-by-step comparisons with the original system are presented in later chapters.

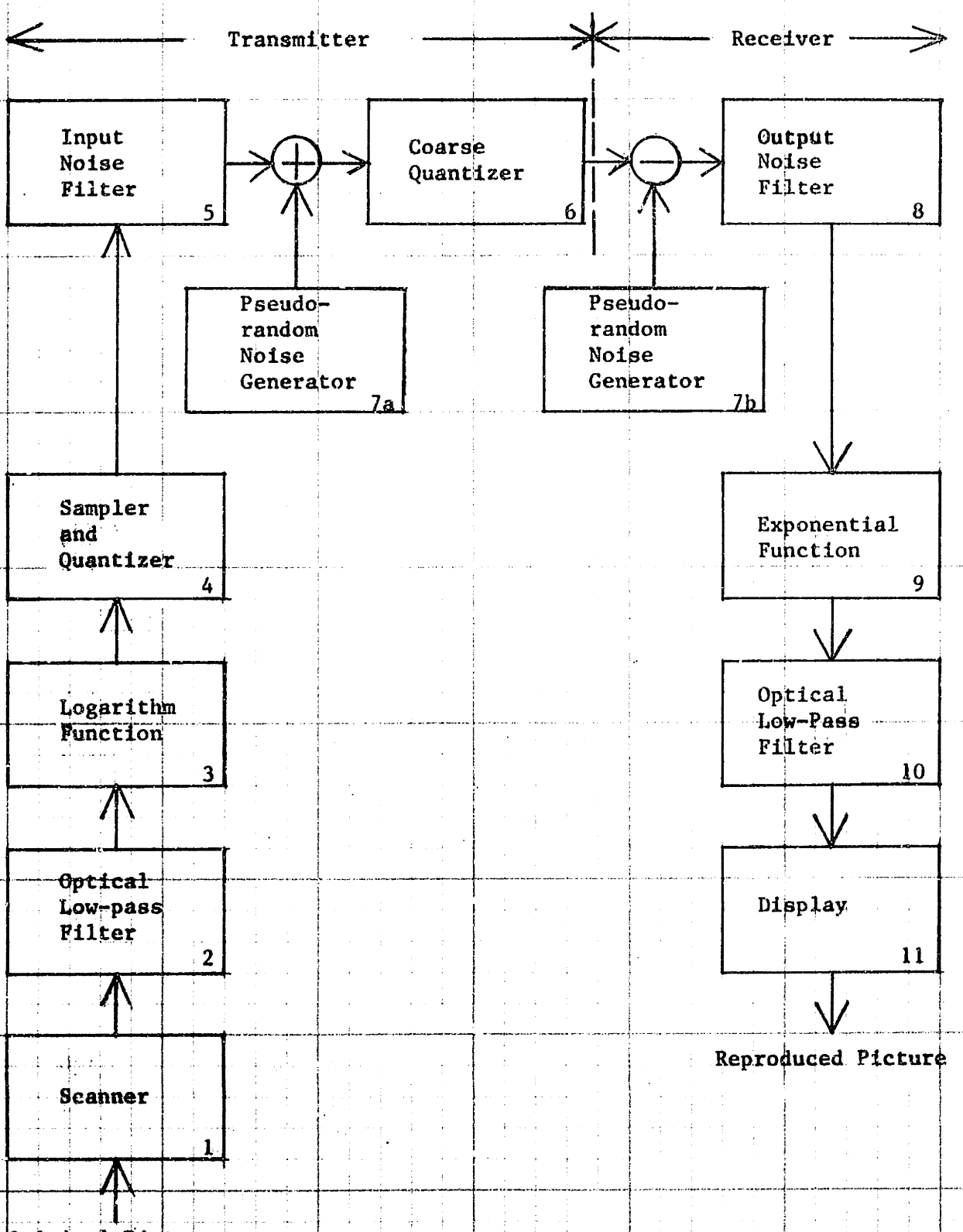
Scanner and Display (blocks 1 and 11)--

The flying-spot scanning system used in the original configuration is replaced by a facsimile scanner and display developed at CIPG. It employs lasers and precise matching of the optics to the characteristics of the display medium, dry silver paper. The format of operation is compatible with commercial facsimile systems. As the new system was evolved, the flying-spot scanner was interfaced to the PDP-9 and still used for convenience.

Sampling Process and Companding (blocks 2, 3, 4, 9 and 10)--

Since facsimile image transducers are being used, it is convenient to implement the spatial low-pass filters optically. This can be accomplished by modifying the scanning beam so it projects the impulse response of the desired filter on the picture. If the photodetector integrates the illumination received over an area A, it will be recording the function g(x,y) which is related to original image, f(x,y) thus:

$$g(x,y) = \int_A \int f(a,b)h(x-a,y-b)dadb$$



The Revised System

Figure 7

where h(x,y) is the impulse response of the filter. The output g(x,y) is identical to the convolution of the input with the filter providing A covers the extent of the impulse response.

In the next chapter the impulse response of this filter will be reexamined to confirm the choice in the original system. The filter at the transmitter has a Gaussian-shaped impulse response with most of its energy concentrated in a small region. The diameter of this area is the maximum distance between samples that can be employed in the sampler without losing image information from the scanner. The sampler and quantizer are hardwired inside the scanner unit and produce output data in the same format as that generated by the coarse sampler in the original system. At the display end the beam produces linear interpolation by being appropriately triangular in shape and modulated by the image data. Since the impulse responses of both of these filters are never negative, each can be realized with a single positive luminance beam. If they had negative lobes, optical filtering would still be possible by scanning simultaneously with two beams of cross-polarized light[43]. One is modulated by the positive portions of the impulse response and the other by the negative lobes. Two cross-polarized photodetectors record the reflected light and feed signals into differential inputs of an amplifier, where the signal resulting from the second beam is subtracted from that corresponding to the beam with the positive impulse information. For monochrome transmission colored beams and color sensitive photodiodes could sub-

stitute for the polarization scheme.

The logarithm and exponentiation functions are also incorporated in the transducer units as analog amplifiers. The use of optical processing has removed the low-pass filters from the log domain. As explained previously, the object of working in the log domain is to psychophysically mask noise. However, the optical filters are assumed relatively noiseless compared to the prime source of noise: the coarse quantization. This hypothesis will be examined more fully in the next chapter.

These revisions place no restriction on the design of the sampling and companding stages. Thus for computer simulation the techniques employed in the original system are applicable. These include convolution via the discrete Fourier transform and two-dimensional processing on the image stored in memory or on disk.

Quantizing Process (block 6)--

Coarse quantization can cause significant distortion in the transmitted picture. The optimum placement of quantizing levels is complicated by the noise filter in block 5. It alters the distribution of the data from the original known range (0 through 255 for the flying-spot scanner). In chapter 5 we investigate methods for quantizing arbitrary signals. Our aim is to find a system that minimizes visible defects and requires as few parameter adjustments for each picture as possible.

Pseudorandom Noise (blocks 7a and 7b)--

Chapter 6 covers pseudorandom noise (PRN) generators developed especially for image processing. A few researchers[26,27] have recently proposed that certain types of structured noise would be less visible than white noise. The psychophysical aspects of this are analyzed in Chapter 7 to determine the relevance to the revised system. Also economical methods for producing structured noise using the results from Chapter 6 are presented.

Noise Filters (blocks 5 and 8)--

The input to the prequantization filter is a stream of numbers. To implement the original noise filters would require storage of an amount of data equal to the extent of the impulse responses of the filters (say r by r) to generate one output point. This effectively means that r rows of data, each consisting of S samples, must be stored. Since r is about 60 this is both uneconomical and introduces a significant processing delay. Therefore, much effort was devoted to restricting the vertical extent of the impulse responses while still retaining the effectiveness of the filters. The analytic and experimental approaches considered are presented in Chapter 8.

Chapter 4

SAMPLING PROCESS

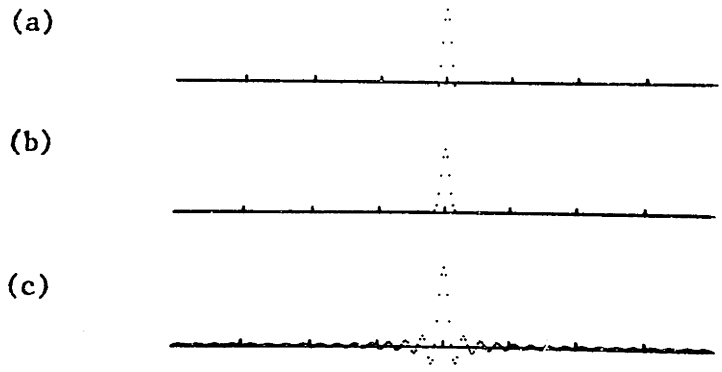
Since the original research was reported in 1970, new techniques for designing digital filters have been published. The application of these results to the low-pass filters was investigated. Before comparisons with the original filters are presented, the fundamental issues in digital filter design are discussed.

The mathematical methods for digital filter design depend significantly on whether the impulse response is of finite duration or infinite duration[13]. This in turn may limit implementation procedures. Recursive filters have feedback and hence can implement any impulse response, while the absence of feedback in nonrecursive filters restricts their use to finite duration impulse responses. The simulation of the low-pass filters with the discrete Fourier transform was nonrecursive. But this is appropriate in the current system since a finite impulse response size is required for optical filtering.

Since the filter is specified in the frequency domain, it is convenient to incorporate the finite duration impulse response constraint there. For a digital filter, the frequency response is periodic because the impulse response is discrete. If the frequency response is also sampled, the impulse response will be periodic, and may be treated as the periodic version of the desired impulse

response. However, the impulse response may now be aliased. This was the case with the ideal low-pass filter and caused the ringing to be more pronounced. Aliasing can be minimized by smoothing the frequency response[45]. The replacement of an ideal low-pass filter with a Gaussian shape accomplished this. Thus, although the primary purpose for the substitution was to eliminate the Gibbs phenomenon, it had the additional advantage, for optical implementation, of limiting the extent of the impulse response. This can be seen by comparing the cross-sectional displays in Figures 8(c) and 8(b). Both functions have the same area and are designed to bandlimit a 256 point signal to 64 cycles per dimension.

Another approach to the requirement for an impulse response with finite duration was developed by Hofstetter[18]. He chose a particular frequency characteristic to optimize via mathematical programming while allowing other characteristics some error tolerance. Fiascinaro[8] has extended this work to two-dimensional filters by designing a low-pass filter with the smallest possible transition bandwidth for fixed passband and stopband ripple of 5%. The impulse response of his filter is shown in Figure 8(a) with the same parameters as the others. Pictures processed with this filter were indistinguishable from those with the Gaussian. Therefore the Gaussian filter was retained. In addition optical implementation of Fiascinaro's filter would be complicated by the negative values at the tails of the impulse response.



(a) Fiascinaro filter

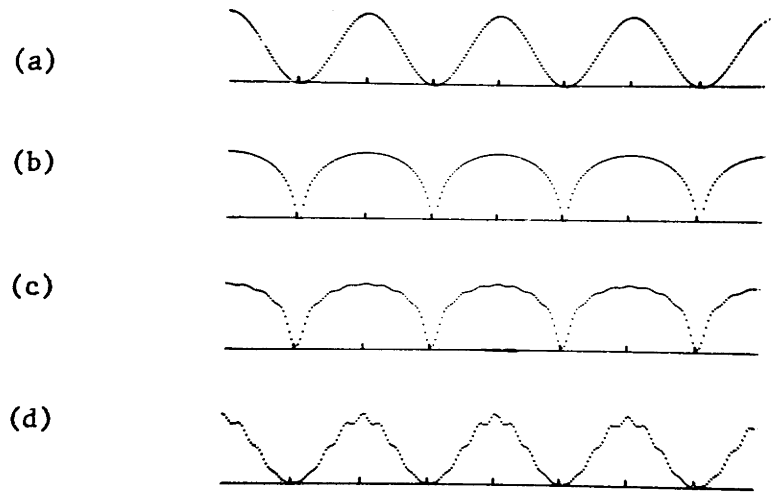
(b) Gaussian low-pass filter

(c) Ideal low-pass filter

Impulse Responses of Anti-Aliasing Filters

Figure 8

In the system description, the placement of the logarithm converter after the low-pass filter was justified by the minimal amount of noise due to optical processing. There are in fact further arguments to support this change. The purpose of the filter is to avoid aliasing when coarse sampling. A picture coarsely sampled to S x S samples can contain information at frequencies as high as S/2 cycles per dimension. Aliasing occurs when information from frequencies above S/2 cycles per dimension are mapped into the retained data. To control the frequency content of the sampled picture, the low-pass filter greatly attenuates these high frequency components. However, the spectrum of the logarithm of data is not simply related to the original spectrum. Information well within the passband of the filter may be distorted if it is first logged, as evident in Figure 9. Trace (a) is the original 256 point sine wave of 4 cycles per dimension, (b) is its logarithm (scaled for display), (c) is composed of the frequency components of (b) up to 32 cycles per dimension, and (d) is the exponentiated result. Thus logging after filtering affords better control over the information extracted from the picture.



(a) Original signal

(b) Logged signal

(c) Logged and filtered signal

(d) Trace (c) exponentiated

Filtering in the Log Domain

Figure 9

Chapter 5

QUANTIZING PROCESS

This chapter concerns designs for the coarse quantizer. The theoretical aspects are presented first. Then the current situation is examined to determine where this theory can be applied to yield a useful design.

Bruce[5] considered the statistical problem of quantizing a random variable into one of 2^Q levels. His criterion was the minimization of the mean of a function which measured the error (E) between the quantized and the unquantized data.

$$E = \int_{-\infty}^{\infty} g\big(x-f(x)\big)\ p(x)dx$$

g is the error function

f is the quantizing function

p(x) is the probability density of input x

Minimizing E involved determining where the quantization levels (termed representational levels by Limb[26]) should be placed within x and which ranges of x should be mapped into a particular level. E can also be expressed as:

$$E = \sum_{i=0}^{2^Q-1} \int_{x_i}^{x_{i+1}} g(x-y_{i+1})\ p(x)dx$$

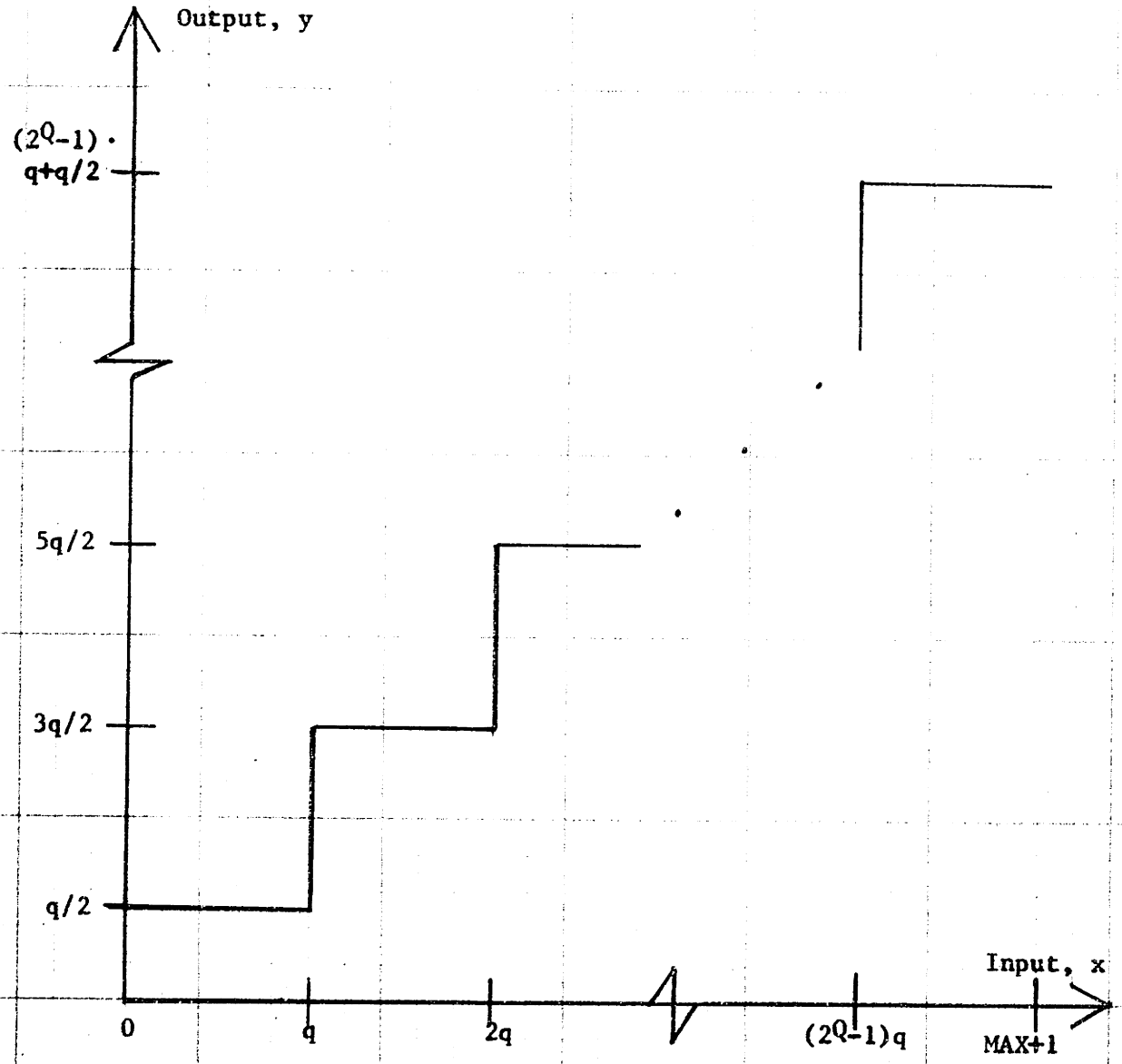
$$x_i \leq y_i+1 \leq x_i+1$$

$$x_0 \leq x_1 \leq x_2 \leq \ldots \leq x_{2Q-1} \leq x_{2Q}$$

(Limb calls the x_i's decision levels.) Bruce used the systematic methods of dynamic programming to minimize E for specific distributions of x and error functions. Max[29] analyzed the situation where $g(z) = z^2$ and found that for any x distribution, the decision levels should be halfway between the representational levels, but the representational levels can be found in general only by numerical methods. This is still true if the representational levels are constrained to be uniformly separated. However, when x has uniform density, the results are straightforward, as shown in Figure 10.

Since the operation of a typical analog-to-digital converter can be modeled by Figure 10, much research has been devoted to the optimum application of a uniform quantizer to a nonuniform density. Max compared the optimum quantizer to the best uniform quantizer for a Gaussian random variable with unity variance. Surprisingly he found that there was less than an 8.5% increase in the mean square error when the quantizer had 8 or fewer levels uniformly spaced rather than optimally spaced. The parameters that were adjusted in the former case were the levels to which the input was clipped (0 and MAX in Figure 10).

Goblick[12] related the number of bits transmitted per sample to the mean square error (ME) and the variance (V) of a Gaussian distribution. For binary coding of 2^Q levels, Q bits are sent, and it is



2^Q = Number of quantization levels

$0 \leq \text{Input} \leq \text{MAX}$

$q = \frac{\text{MAX}+1}{2^Q}$ = quantization level size

Optimum Quantizer for Input with Uniform Density

Figure 10

approximately valid that

$$Q = 0.125 + 0.6 \log(V/ME) \quad (5.1).$$

He also considered statistically recoding these bits so fewer could be sent without loss of information. This is entropy coding, and it yields a reduction in bits whenever each quantizing level is not equally likely. This is the case when uniformly quantizing a non-uniform distribution.

Let us consider the converse issue: how many quantizing levels (N) could be represented for minimum ME if entropy coding to Q bits were used? Goblick showed that with entropy coding

$$Q = 0.25 + 0.5 \log(V/ME) \quad (5.2)$$

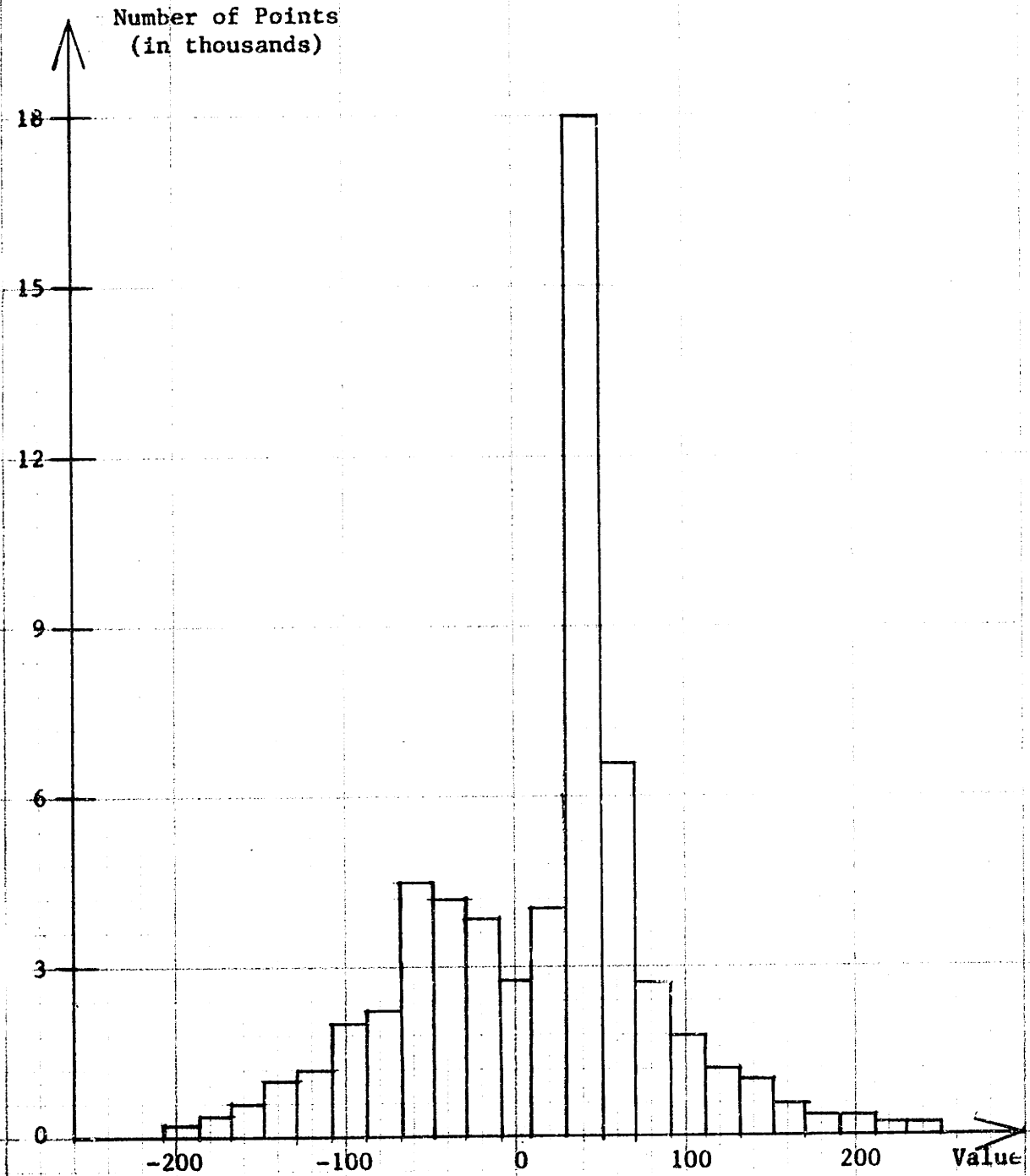
independent of N for N ranging from 2 to 140 if $ME/V < 0.5$. By first replacing Q with log N in (5.1) and combining with (5.2) to eliminate V/ME we find that

$$N = 2^{(1.2Q-0.175)}.$$

This means that for 3 bit entropy transmission, a uniform quantizer could have at most 10 levels, instead of 8 with binary coding. The remarkable efficiency of uniform quantization is particularly evident in (5.2). The second term on the right is Shannon's rate distortion bound, i.e., the minimum entropy for any coder. So an appropriate uniform quantizer requires only another ¼ bit. Gish[11] extended this investigation to a wide class of common distributions and also found that the uniform quantizer is only a fraction of a bit above the rate distortion bound.

Figure 11 shows a typical distribution of data at the output of the input noise filter. By attenuating the D.C. level (u=v=0), this filter allows the image data to range over negative as well as positive values. This fact, coupled with the amplification of high frequency components, implies that the filter outputs of relatively large positive or negative magnitude correspond to brightness transitions in the image, where high frequencies are concentrated. It would be convenient if uniform quantization could be applied to this distribution, which is similar to that analyzed theoretically. The critical test of this proposal is simulation with pictures, especially since the validity of the mean square error criterion has never been established with respect to apparent image quality.

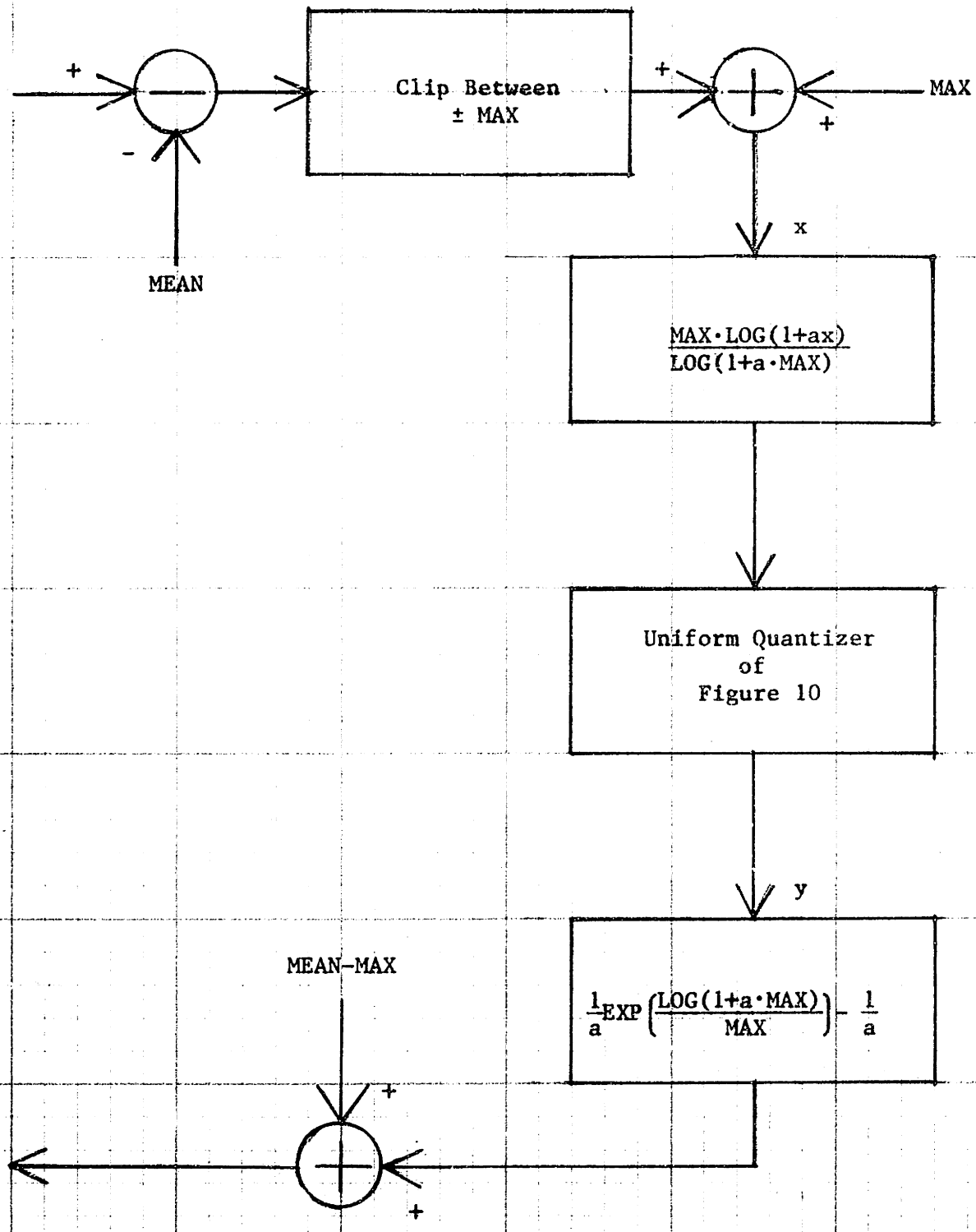
In order to experiment with various quantizers, the system shown in Figure 12 was programmed on the PDP-9 computer. The parameter in the logarithm and exponentiation, a, determined the degree of nonlinearity in the quantizer. For $a \ll 1$, the quantizer is nearly uniform. The data were clipped between MEAN $\pm$ MAX before quantizing. This structure was used as the coarse quantizer in the system. No sample reduction was included so the effect of the quantizer could be isolated. The original two-dimensional noise filters and those to be presented in Chapter 8 were all investigated. After numerous tests it was concluded that the choice of MEAN and MAX, i.e., the clipping levels, had a far greater effect on picture quality than the structure of the quantizer applied within these levels. The uniform quantizer



Histogram of Data from the

Input Noise Filter

Figure 11



Experimental Coarse Quantizer

Figure 12

performed adequately and was retained, especially since it is particularly convenient for a real-time system.

With a uniform quantizer, Roberts[37] demonstrated that the processing done in the steps between the noise filters is equivalent to limiting the data range by clipping to

$$(\text{MEAN} + \text{MAX} - q/2,\ \text{MEAN} - \text{MAX} + q/2)$$

q is the coarse quantization interval

and adding noise distributed uniformly over a range of q. For the variety of pictures in Figure 32, MEAN = 50 was found appropriate. To determine the effect of clipping, Figure 5(a) was clipped to 50 ± MAX for MAX = 32, 64, 128 and 256 using the one-dimensional noise filters (the first version of one-dimensional noise filters to be presented in Chapter 8) to produce Figure 13. Listed in this figure is the standard deviation, D, of each picture from the original, computed as a percentage of the dynamic range of 256. Notice that these values do not convey the threshold effects in the visibility of the distortions for MAX less than 64. This illustrates the problem of mean-type measurements when applied to image perception, where localized disturbances may distract from the entire picture.

When the complete Roberts scheme for Q = 3 is included by adding the appropriate noise, we see the tradeoffs in Figure 14. As the clipping level is increased, the nonlinear distortions decrease, but the overall noise level increases. For these one-dimensional noise filters, Figure 14(b) would be chosen as a compromise where the noise

MAX = 32
D = 9.72%
(a)

MAX = 64
D = 6.42%
(b)

MAX = 128
D = 3.46%
(c)

MAX = 256
D = 3.14%
(d)

Clipping level = 50 ± MAX

Clipping Level Adjustments

Figure 13

MAX = 32
D = 9.54%
(a)

MAX = 64
D = 6.87%
(b)

MAX = 128
D = 6.60%
(c)

MAX = 256
D = 11.90%
(d)

Clipping level = 50 ± MAX
Noise extent = 2 · MAX/8

Clipping Effects Versus Noise

Figure 14

is just sufficient to mask the nonlinearities.

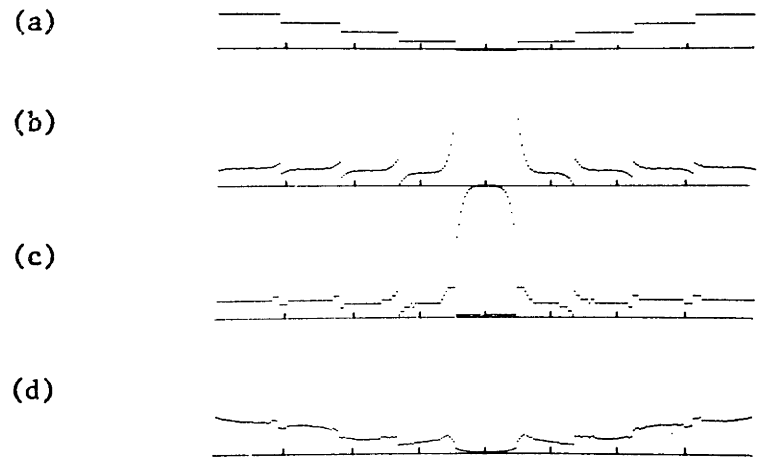
During these experiments the step wedge graphed in Figure 15(a) was examined. It is included here to illustrate the actions of various stages of the system. Figure 15(b), at the output of the transmitter noise filter shows:

a) the attenuation of the mean value of the input due to the low DC response of the filter;

b) the amplification by the filter of the high frequencies at the level transitions; and

c) the preferential treatment of the darker areas in the center of the trace due to processing in the logarithm domain.

To produce Figure 15(c), the data shown above was clipped to the range (0, 120) and uniformly quantized to 8 levels. As a result of (c) there is more detail preserved in the receiver recreation, Figure 15(d), in the darker portions. This step wedge picture or any other test pattern was not used to determine system performance or to adjust parameters. The response of the human visual system to an artificial subject is not indicative of the response to the more typical scenes in Figure 33, which the system is being designed to transmit.



(a) Original step wedge

(b) Output of input noise filter

(c) Output of coarse quantizer

(d) Processed data

Quantizing a Step Wedge

Figure 15

Chapter 6

NOISE GENERATION AND MEASUREMENT

Pseudorandom noise (PRN) is the prime component in the Roberts technique. This noise should have a flat distribution between ±1/2 of a coarse quantization and should be statistically independent on a sample-by-sample basis. The last requirement implies a flat power spectrum. This chapter investigates methods for generating PRN and for measuring its characteristics.

A particularly convenient scheme for generating random numbers in a computer is the power-residue method[25]. Each successive random number is obtained from the low-order b bits of the product of the "seed" and the previous number. The appropriate seed depends on b, which usually corresponds to the computer word size. The initial value is any odd b-bit number, and 2^{b-2} terms are produced before the sequence repeats. This PRN generator was implemented on a PDP-9 computer with b = 18. The highest 6 bits from each sample were retained to produce integers ranging from -32 to +31. The histogram of 1024 values had an rms deviation of 3.5 from the computed (and expected) mean of 16 samples/integer in the range generated.

The power density spectrum of this sequence was calculated for frequencies up to 256 cycles/dimension (the highest frequency that can be represented by a picture composed of 512 x 512 samples) by the

following method:

a) We compute the unbiased autocovariance estimate, a(k), of the noise, x(k),

$$a(k) = \frac{1}{1024-|k|} \sum_{r=1}^{1024-|k|} x(r)\, x(r+k) - \left(\frac{1}{1024} \sum_{r=1}^{1024} x(r) \right)^2 ,$$

where k ranges from -255 to +256.

b) We window the estimate with a Gaussian function having 90% of its area between k = ±64.

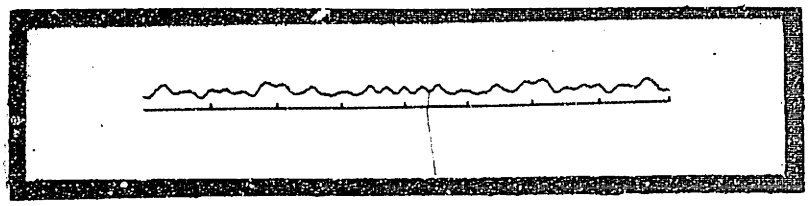
c) The estimate of the power density spectrum is the discrete Fourier transorm of the windowed autocovariance estimate.
It is evident from Figure 16(a), that the spectrum is relatively flat.

The noise was displayed on the flying-spot scanner. In Figure 17(a) each of 512 x 512 samples represents 128 + 4 X (noise value). Since the apparent randomness of the noise is the critical factor in picture coding, the vertical stripes in the figure make this method unacceptable. This problem was eliminated without altering the shape of the noise spectrum by generating 5 extra random numbers at the end of each row. The results of using this scheme are shown in Figure 17(b).

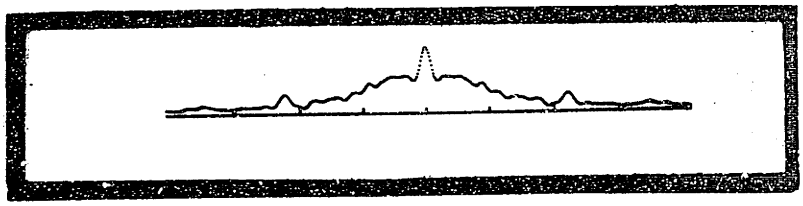
To implement an image transmission system in special-purpose hardware, there is a more economical PRN generator than the power-residue method. This generator utilizes a shift register with feedback of the exclusive OR of selected cells to simulate an irreducible

primitive polynomial[33]. For an 18-bit shift register the simplest arrangement is the exclusive OR of cells 1 and 8 to feed cell 18, which was simulated on the PDP-9 computer. The computed power density spectrum is shown in Figure 16(b). The high-frequency roll-off occurred because the register was shifted once between samples so that the low-order bits remained correlated from point to point. Figure 17(c) was generated by using these values. Comparison with Figure 17(b) demonstrates the correlation of the samples. To decorrelate the bits composing the integers, every other bit was retained after each shift. As shown in Figure 16(c), this did not solve the problem. The relatively flat spectrum of Figure 16(d) was finally produced by shifting the register 6 times between samples composed of the high-order bits. The corresponding noise picture appeared to be similar to Figure 17(b). (Calculations on the histogram yielded an rms deviation of 4.0 from a mean of 16.)

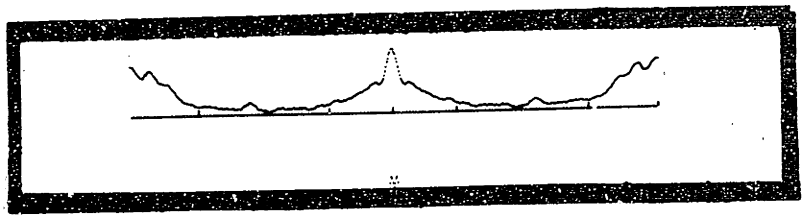
This study illustrates the importance of exploiting the inherent pattern-recognition capabilities of the human visual system when designing image processing components. Not only can design calculations be conveniently corroborated, but unnecessary computation can be avoided if a proposal is unacceptable visually.



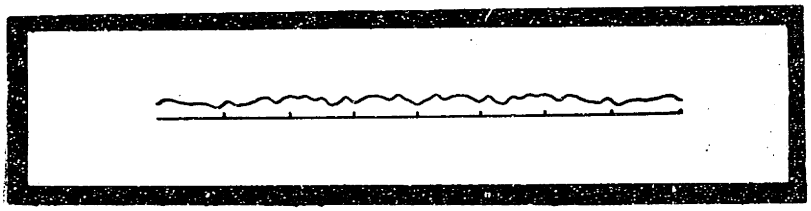
(a)



(b)



(c)

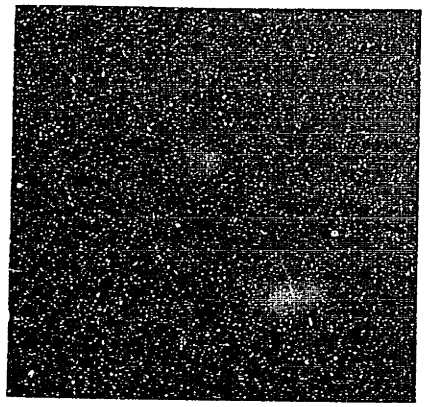


(d)

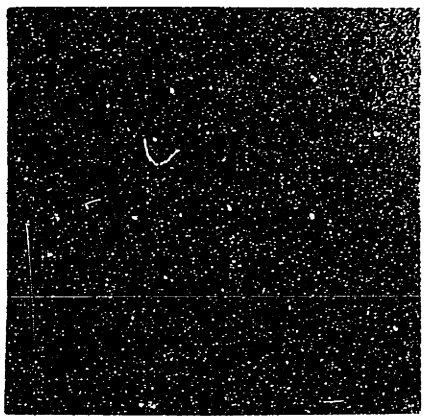
Frequency scale: -512 cycles to +512 cycles, with DC at the center.

Power Density Spectra

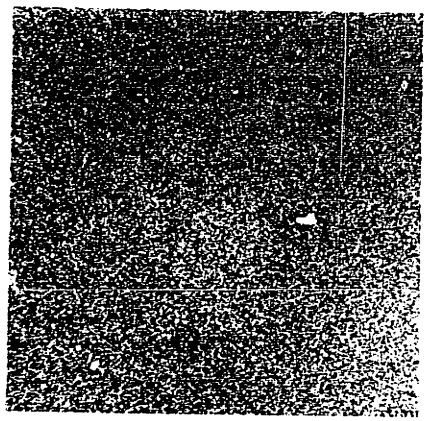
Figure 16



(a)



(b)



(c)

Performance of Pseudorandom Noise Generators

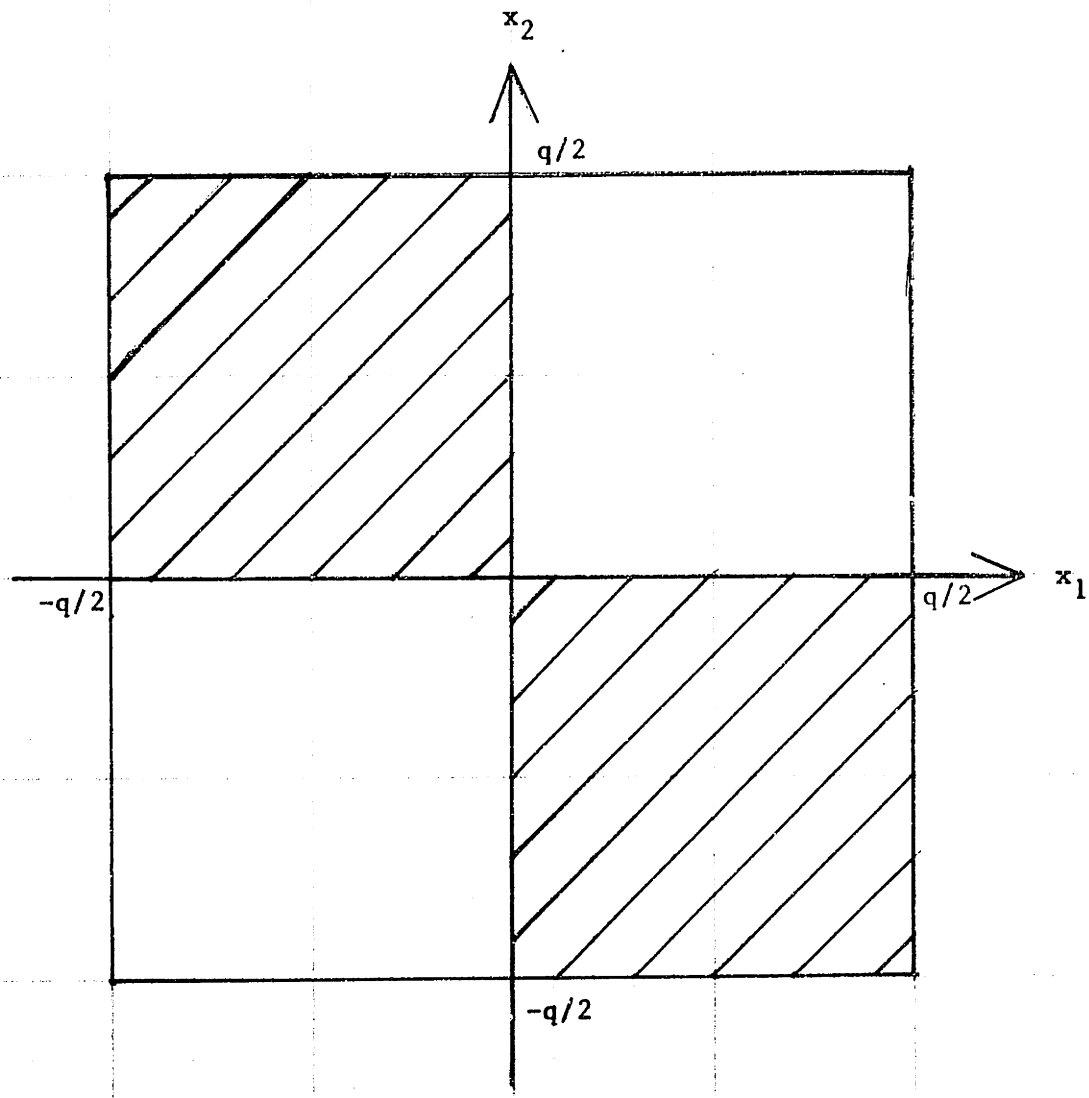
Figure 17

Chapter 7

STRUCTURED NOISE

Research conducted independently by Limb[26] and Lippel[27] has considered the possibility of manipulating the statistics of PRN to improve the performance of the Roberts technique. Their proposals are discussed in this chapter and the results of simulating them are presented.

Limb studied the frequency response of the human visual system (Figure 3) and reasoned that since high frequencies are attenuated, a less visible noise should be concentrated in this frequency region. In the spatial domain this implies that the values of successive noise points should have large differences, rather than differences that are equally likly large or small. As an example, the ideal nonrandom signal would alternate between +q/2 and -q/2 on successive samples (q is the quantization interval) so its power spectrum has nonzero values only at ±S/2 cycles (S x S samples in the picture). A random process which approaches this has a joint probability density for adjacent samples given in Figure 18. Limb described a class of noise signals with similar negative correlation properties. As seen in the examples of Figure 19, the range in which noise may occur changes from sample to sample in a cyclic way; hence, Limb named these "n-step dither patterns", where n is the number of steps per cycle.



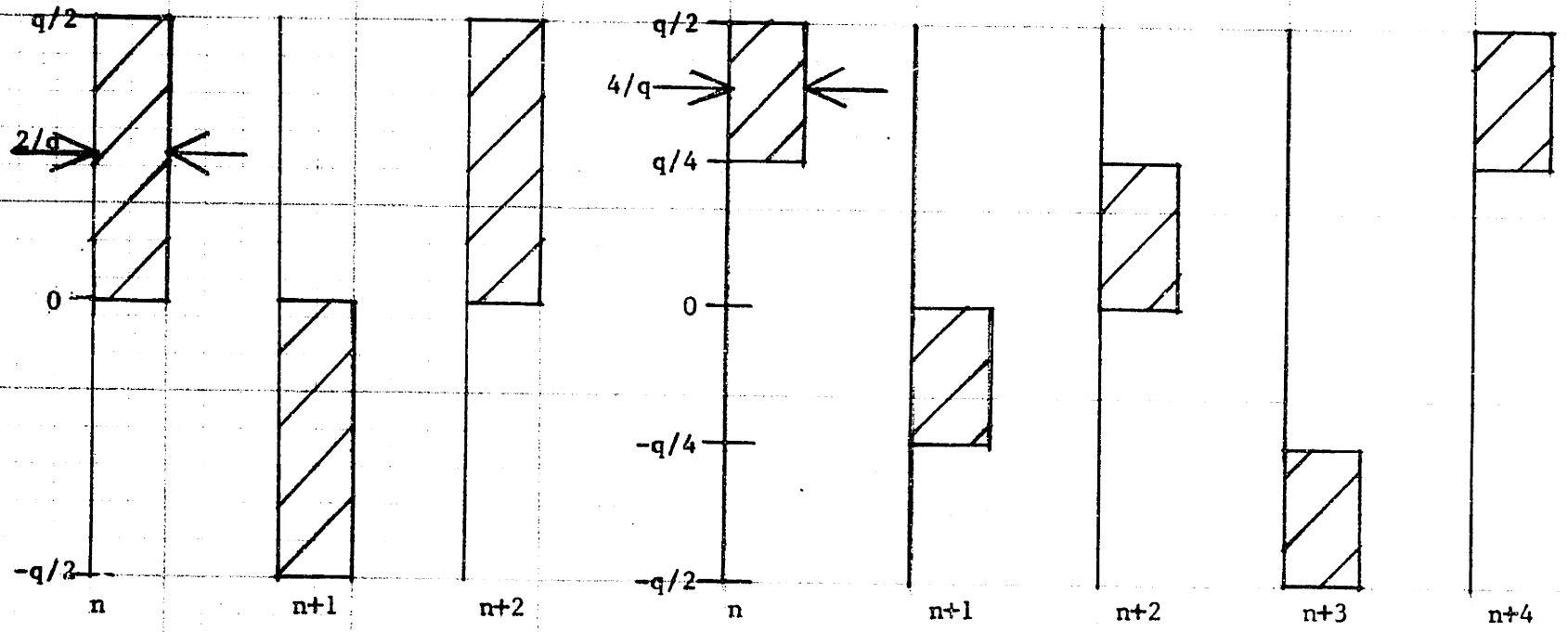
q is the quantization level size

Amplitude = $2/q^2$ in shaded region, and 0 elsewhere

Joint Probability Density for 2 Successive Points

of 2-Step Dither

Figure 18



Shaded area represents probability density

2-Step and 4-Step Dither

Figure 19

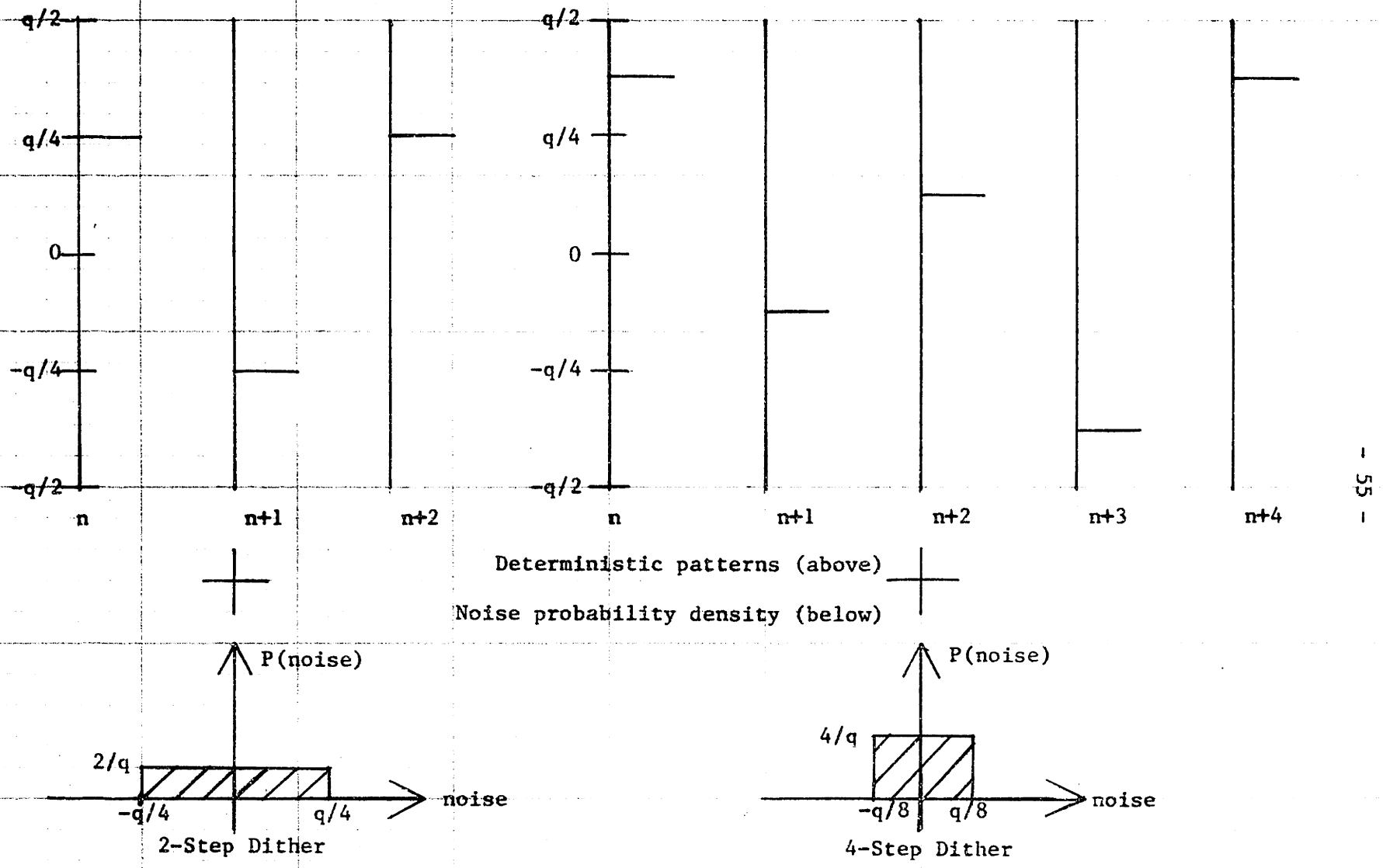
Limb compared various patterns by calculating the rms and mean modulus (absolute value) error after the system output was passed through a filter with the frequency response of the visual system. In Chapter 5 we saw that mean-type measurements do not reveal the nature of the distortion involved, which may significantly affect picture quality. We will therefore consider another approach to analyzing this system. Figure 20 shows alternate representations for 2 and 4-step dither patterns as sums of deterministic and random patterns. Thus a picture processed with an n-step random pattern consists of the original image plus a periodic pattern with n points per period plus random uncorrelated noise uniformly distributed between $\pm q/2n$. This scheme then adds noise with mean power $1/n^2$ of that used by Roberts at the expense of introducing a periodic pattern.

For a 2-step sequence only one periodic pattern exists $(+q/4, -q/4, +q/4, \ldots)$. For higher n, Limb sought sequences which had large changes from point to point. Of the following 4-step patterns, for example, (7.1) is more desirable:

$$-3q/8,\ \ q/8,\ -q/8,\ 3q/8,\ -3q/8, \ldots \qquad (7.1)$$

$$-3q/8,\ -q/8,\ \ q/8,\ 3q/8,\ -3q/8, \ldots \qquad (7.2)$$

Computations (in Appendix A) of the discrete Fourier transform (DFT) of these two patterns show that the one which is better from the spatial domain criterion has more power concentrated at higher frequencies. Each pattern was combined with PRN of amplitude $\pm q/8$ $(q = 32)$ to produce the desired noise. The spectrum of each was calculated by the



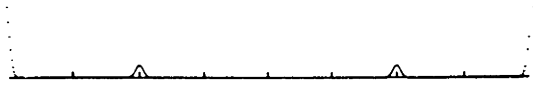
Alternate Representation of Dither Patterns

Figure 20

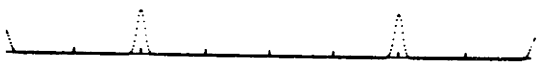
method in Chapter 6 to generate Figure 21, which confirms the computations in Appendix A. In both graphs the noise level is not discernible from the frequency axis because its average amplitude is 1/64 of the peak in trace (a).

Since these sequences are being applied to two-dimensional pictures, the actual frequency goal must be considered in the Fourier transform plane. To handle this problem Limb presented sequences which met his requirements in the horizontal and vertical directions for 2 steps (Figure 22(a)) and 4 steps (Figure 22(b)), and concluded that ideal patterns for larger n do not exist. An 8-step two-dimensional pattern was derived by Lippel. By refering to the ancient notion of magic squares, he found a pattern where every 2 x 2 subsquare and every horizontal, vertical, and diagonal run of four elements has zero sum, as can be seen in Figure 22(c). The noise added to make this an 8-step random sequence has mean power only 1/64th of the original Roberts dither, but the periodic part of the signal has power down at S/8 cycles per picture width. The appropriate PRN was added to each sequence, and the results were applied to 3 bit transmission without sample reduction, logarithm processing, or noise filters (see Figures 23(a) through 23(c)). Essentially these noise patterns are transfering some white noise power at all frequencies to a few spectral lines. To assess the relevance of this, it is important to reconsider the purpose of the noise in quantization.

The significant achievement of the Roberts technique is the



Spectrum of pattern 7.1
(a)



Spectrum of pattern 7.2
(b)

Frequency scale: -512 cycles to +512 cycles, with DC at the center.

Spectra of Structured Noise

Figure 21

-q/2	q/2
q/2	-q/2

2-Step sequence
(a)

-3q/8	q/8
3q/8	-q/8

4-Step sequence (desirable)
(b)

-7q/16	7q/16	-5q/16	5q/16
-3q/16	-3q/16	q/16	-q/16
5q/16	-5q/16	7q/16	-7q/16
-q/16	q/16	-3q/16	3q/16

8-Step sequence
(c)

One period is shown for each sequence

n-Step Sequences for Two Dimensions

Figure 22

2-Step sequence
(a)

4-Step sequence
(b)

8-Step sequence
(c)

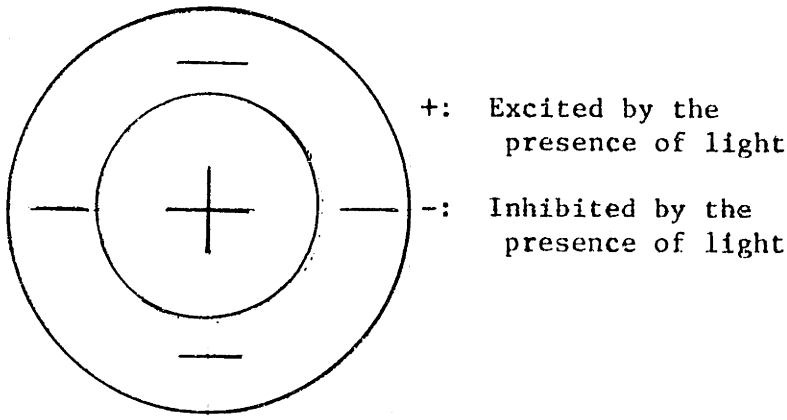
4-Bit quantization
without structured noise
(d)

Quantization Using Structured Noise

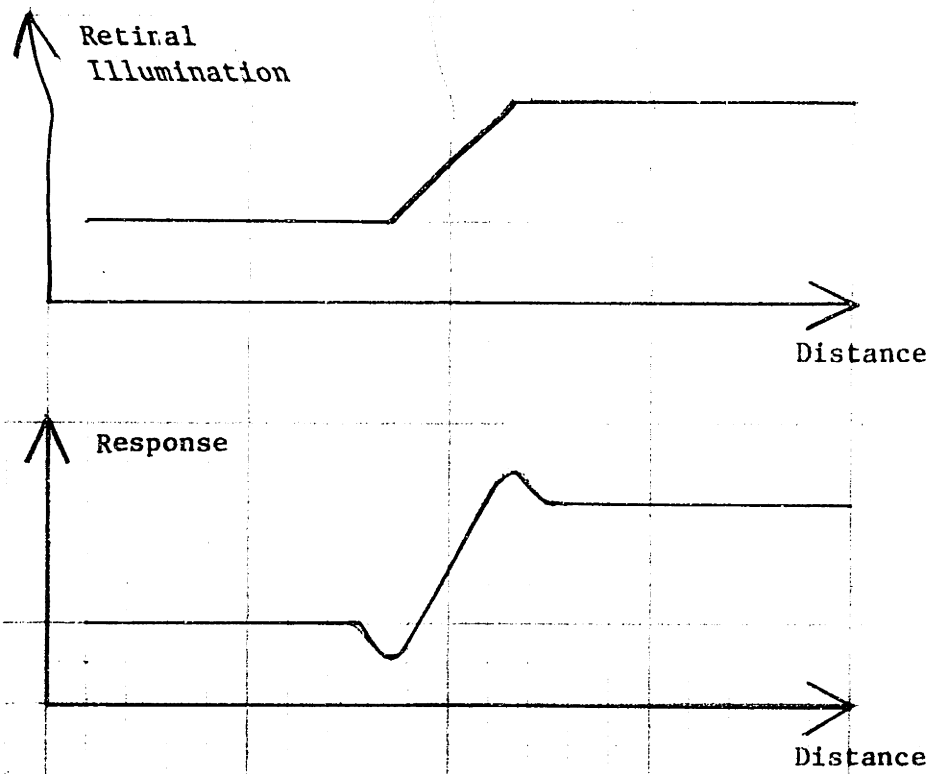
Figure 23

elimination of spurious contours and the substitution of noise with no apparent structure. The marked improvement in picture quality this affords occurs because the coding accomodates a fundamental aspect of the visual system.

Electrophysiological studies of the retinae of animals ranging from invertebrates to mammals have revealed[51] that many individual receptors are interconnected to produce receptive fields (response as a function of position on the retina) as illustrated in Figure 24(a). This means that there is a single ganglion cell which is stimulated by the excitation of receptors located in the "+" region, while light falling on the receptors in the "-" region inhibits the firing of the ganglion cell. The presence of these receptive fields (termed Kuffler units) throughout the visual field causes the eye to suppress information about nonchanging luminances, while transmitting to higher visual processes information about changes. Figure 24(b) shows the effect this has on the perception of a luminance transition. This apparent edge enhancement is called a Mach band[6]. Kuffler type units of various shapes and orientations are being used to describe the abstraction of information that occurs in the higher visual processes. Hubel and Wiesel[24] have found cells in the visual cortex of cats that have receptive fields maximally stimulated by straight lines in specific orientations, and others that serve as corner detectors. This tends to support the view that much preliminary form recognition involves contour perception[53].



Receptive field of a single ganglion cell
(a)



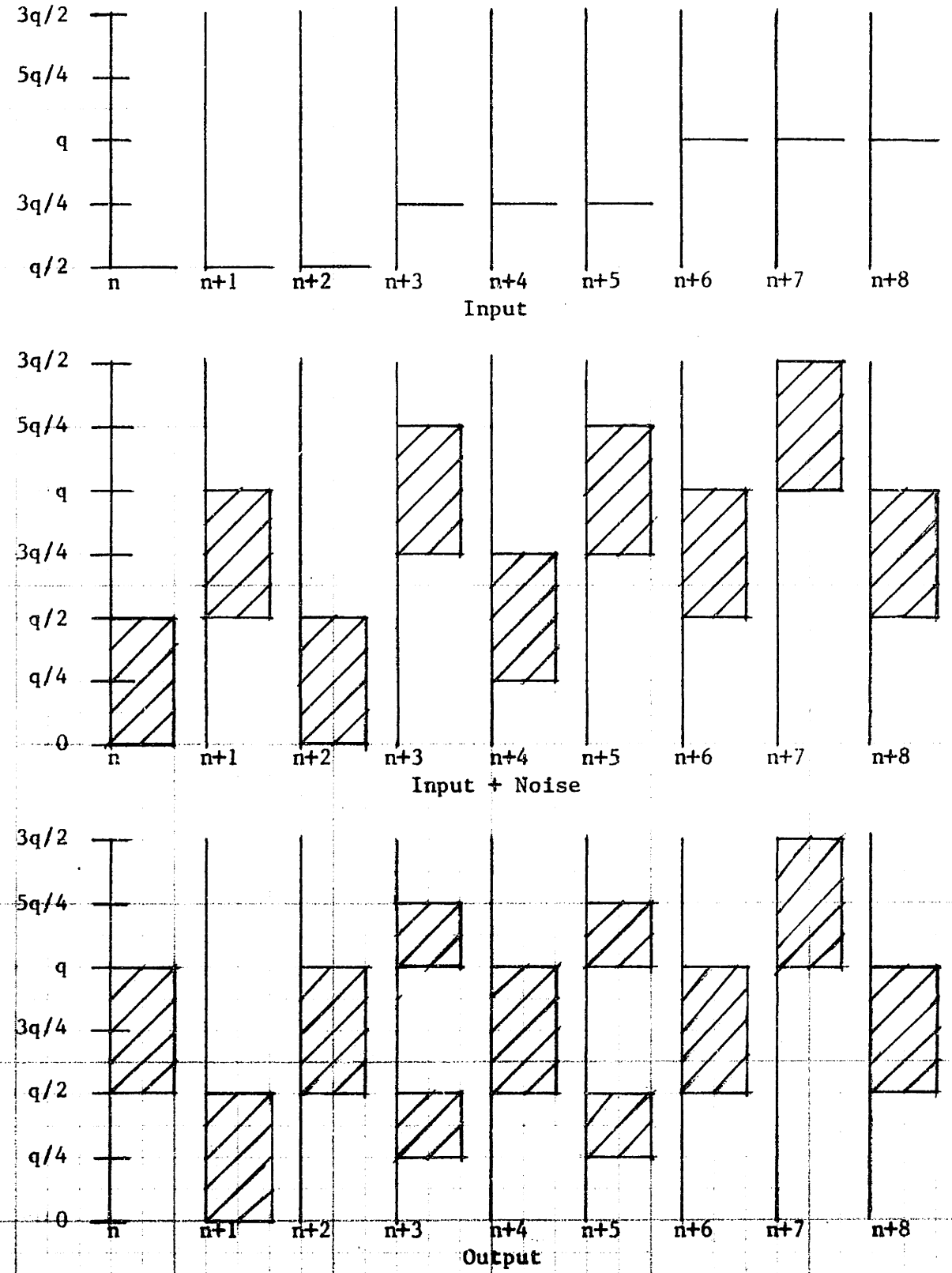
Edge perception of human vision
(b)

Psychophysical Factors in Human Vision

Figure 24

Thus a noise signal with structure is more likely to pass through the information filters of the human visual system and distract from the perception of the picture to which it is added. Mitchell recently[31] tested this hypothesis by having subjects compare the visibility of narrow band noise with wide band noise. He found that as the bandwidth decreases (keeping the total energy in the picture constant), the noise visibility increases by almost a factor of two.

This is why there is not a striking decrease in apparent noise as the patterns get more complex. The greatest change occurs between no pattern (Figure 6(b)) and the 2-step pattern, where the white noise component is ¼ the power of that with no pattern. This noise component is equivalent to the unpatterned noise employed in 4 bit transmission, which was used to produce Figure 23(d). There still is a subtle, but important, difference in the overall performance of a system with patterned versus unpatterned noise. As mentioned in Chapter 5, the noise added to the data in the Roberts quantizer is the same for almost all inputs. However, the n-step noise changes form as the data level changes. This is illustrated in Figure 25 for a 2-step process. Careful examination of Figures 23(a) through 23(d) reveals this data dependence of the noise. Thus structured noise is not employed in the revised transmission system.



Shaded area represents range of noise
Processing of Data with 2-Step Noise

Figure 25

Chapter 8

NOISE FILTERS

The primary hindrances to stream coding of the picture data were the two-dimensional noise filters. As mentioned in Chapter 3, these filters derived a single output point from data in about 60 adjacent rows. This chapter presents the development of filters which require many fewer rows to achieve the effect of the original filters.

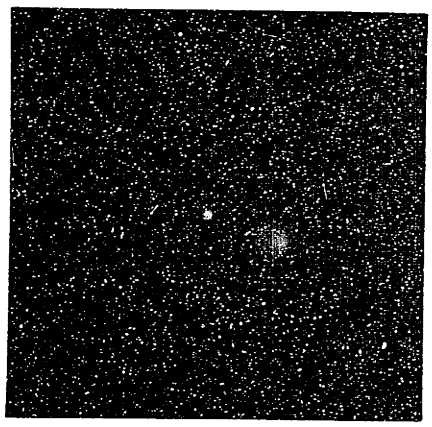
As the first approach to the problem, one-dimensional filters were obtained from the two-dimensional frequency specifications by using the response along the u-axis. The system was simulated without sample reduction, and the quantizing parameters (discussed in Chapter 5) were optimized for 3 bit transmission. As is evident in Figure 28(a), the one-dimensional structure of the noise is particularly visible.

In view of the psychophysical sensitivity to structured noise, the filters were redesigned to include a limited number of rows in the impulse response of the filters. Also the horizontal frequency characteristic was retained by requiring that the u-axis response be unchanged. These goals were formulated in a constrained minimization problem, where we sought to minimize the mean square error between the new and the old frequency responses. Appendix B contains the computations involving Lagrangian methods. The solution states that each

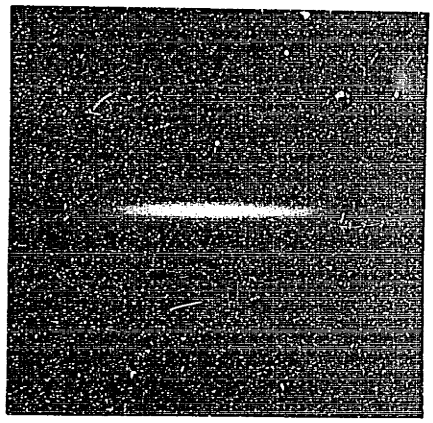
retained row in the impulse response should be a sum of the original row plus a weighted sum of all rows omitted. However, there is a fundamental difficulty in applying this to the noise filters. The resulting filters are not the inverse of each other for the versions considered, including 3, 5, and 7 rows.

The problem was reformulated to solve for both filters simultaneously while including the inverse requirement. As the computation in Appendix C reveals, the mathematics soon became intractable. Instead, the output noise filter designed in Appendix B was used to derive the input filter. The output filter was not directly invertible because its frequency response had negative lobes as a result of the sharp amplitude change where the rows of data in the impulse response were forced to zero. The impulse response was windowed until the frequency response remained non-negative. The coefficients (W_r) selected for the rows are listed in Figure 26 under the corresponding displays of the magnitude of the frequency response for output filters with 3, 5, and 7 rows. Now that the modified input noise filter could be obtained, the system was configured to produce the 3 bit transmissions shown in Figure 28 (b) through Figure 28(d).

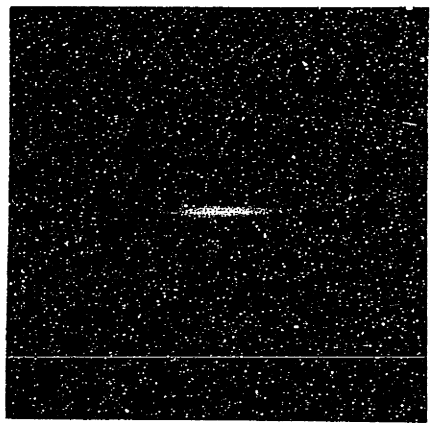
Figure 27(b) presents the noise as shaped by the one-dimensional output filter employed for Figure 28(a). In light of Mitchell's findings (discussed in Chapter 7), the noise is particularly visible because the lack of vertical filtering introduces the striped structure. It was theorized that if the filtering in the



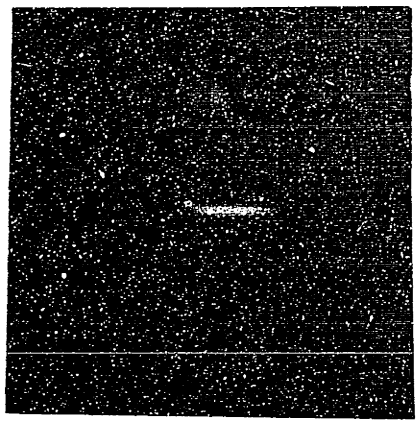
Original filter
(a)



Impulse response limited to
3 rows windowed:
$W_1 = W_{-1} = 0.5$
$W_0 = 1.0$
(b)



Impulse response limited to
5 rows windowed:
$W_2 = W_{-2} = 0.3$
$W_1 = W_{-1} = 0.75$
$W_0 = 1.0$
(c)

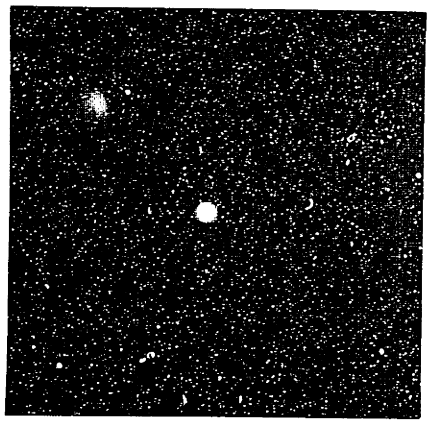


Impulse response limited to
7 rows windowed:
$W_3 = W_{-3} = 0.2$
$W_2 = W_{-2} = 0.4$
$W_1 = W_{-1} = 0.8$
$W_0 = 1.0$
(d)

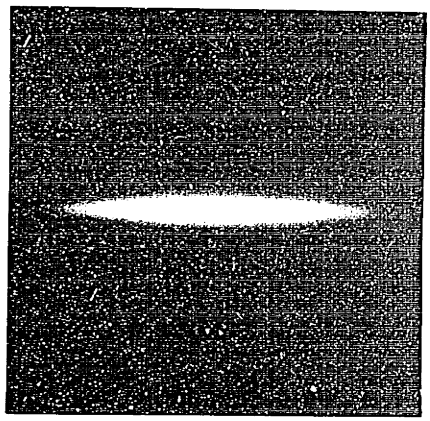
The v axis is horizontal and the u axis is vertical

Magnitude of the Frequency Response of Various Output Noise Filters

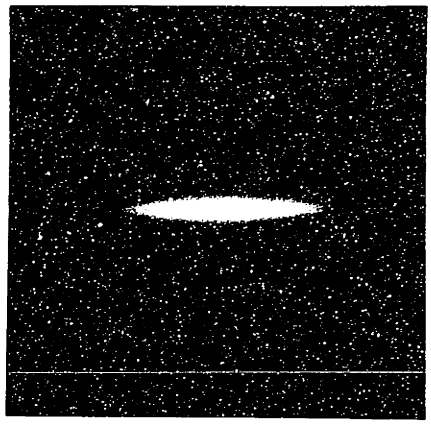
Figure 26



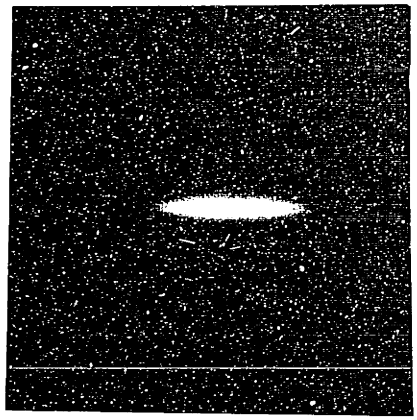
Original filter
(a)



Impulse response limited to
3 rows windowed:
$W_1 = W_{-1} =$
$W_0 =$
(b)



Impulse response limited to
5 rows windowed:
$W_2 = W_{-2} = 0.3$
$W_1 = W_{-1} = 0.75$
$W_0 = 1.0$
(c)



Impulse response limited to
7 rows windowed:
$W_3 = W_{-3} = 0.2$
$W_2 = W_{-2} =$
$W_1 = W_{-1} =$
$W_0 =$
(d)

The v axis is horizontal and the u axis is vertical

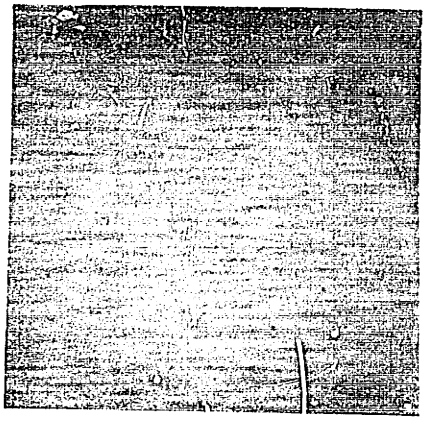
Magnitude of the Frequency Response of Various Output Noise Filters

Figure 26

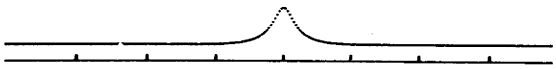
horizontal direction was reduced, perhaps there would be an overall decrease in the noise visibility. After some experimentation, the one-dimensional frequency response was altered by taking the square root of the original values and scaling so the DC response was unchanged, as illustrated in Figure 27(c). As was done in Figure 27(b), noise was passed through this new output filter to yield Figure 27(d). Although the change in noise visibility is rather small, these new filters produced remarkably better pictures, as can be seen in Figure 28(e). This improvement resulted primarily because the clipping level in the quantizer could be reduced from MAX = 128 in Figure 28(a) to MAX = 64 in Figure 28(e) without introducing distortions. Thus the noise amplitude could be out by a factor of 2. Similar results were obtained for all the test subjects in Figure 33. The minimal amount of horizontal structure remaining with these new filters could be virtually removed by expanding the impulse responses to 3 rows.

An additional experiment was performed to compare the performance of these two versions of one-dimensional noise filters with entropy coding. The data produced by the input noise filter was altered until its entropy was less than 3 bits, then the results were entered into the output noise filter. To reduce the entropy the entire range of data was quantized by setting least significant bits to zero without clipping any data. For both versions of the filters an entropy of 2.9 bits was achieved by fixing the values of the 3

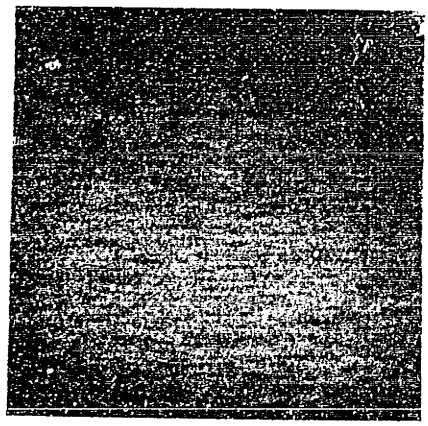
Frequency response of original filter
(DC in center)
(a)



Noise through filter in (a)
(b)



Frequency response of new filter
(DC in center)
(c)



Noise through filter in (c)
(d)

Noise Through One-Dimensional Output Noise Filters

Figure 27

One-dimensional filter
(a)

3-row impulse response
(b)

5-row impulse response
(c)

7-row impulse response
(d)

New one-dimensional filter
(e)

Processing With Various Noise Filters

Figure 28

least significant bits. The results shown in Figure 29 demonstrate the advantages of the modified filters. Furthermore, they indicate that the real-time quantizing proposal yields results that are close in quality to entropy coding.

For real-time implementation these new filters solve the stream processing requirement. For simulation on the PDP-9 computer, the fast Fourier transform (FFT) of each row of the picture was multiplied by the frequency response of the appropriate filter, and the inverse FFT yielded the filtered data. Rabiner[36] discussed the tradeoffs involved in various hardware realizations of such filters and concluded that this FFT method involves fewer multiplications and additions than a recursive or a direct convolution scheme when there are more than 32 samples per row.

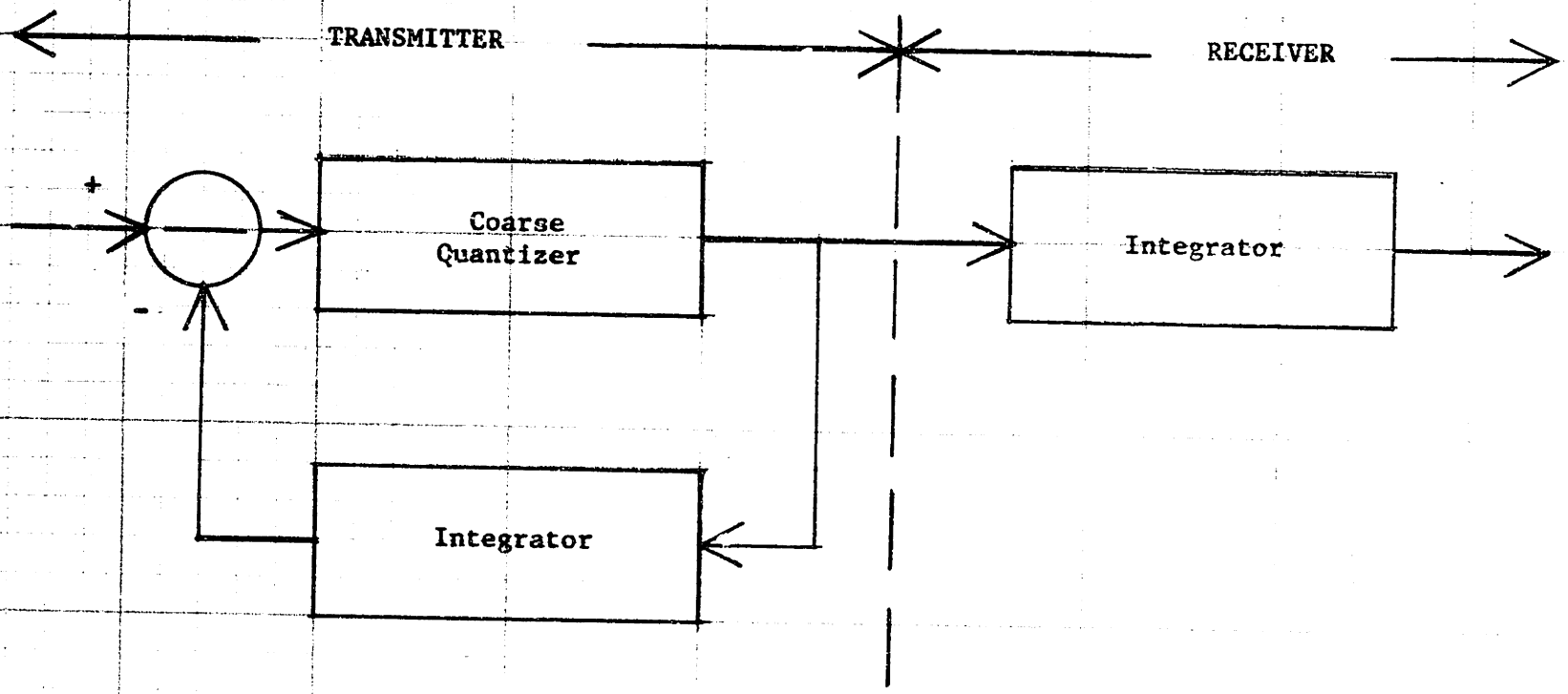
With the one-dimensional noise filters, this facsimile transmission system can be compared to a differential pulse code modulation system (DPCM)[40], as drawn in Figure 30. The feedback network at the transmitter serves two purposes. One is to quantize differences between samples in order to exploit the correlation of adjacent elements. Secondly, the quantization noise associated with each sample depends on previous outputs so it is independent of the brightness level[4]. In the current scheme, this independence is controlled directly by the pseudorandom noise generators. The input and output noise filters, chosen according to psychophysical criteria, are in fact approximately a differentiator and an integrator respectively.

Original one-dimensional noise filters
(a)

Modified one-dimensional noise filters
(b)

Entropy Coding

Figure 29



A Differential Pulse Code Modulation Transmission System

Figure 30

We have here greater flexibility in controlling various characteristics than in DPCM.

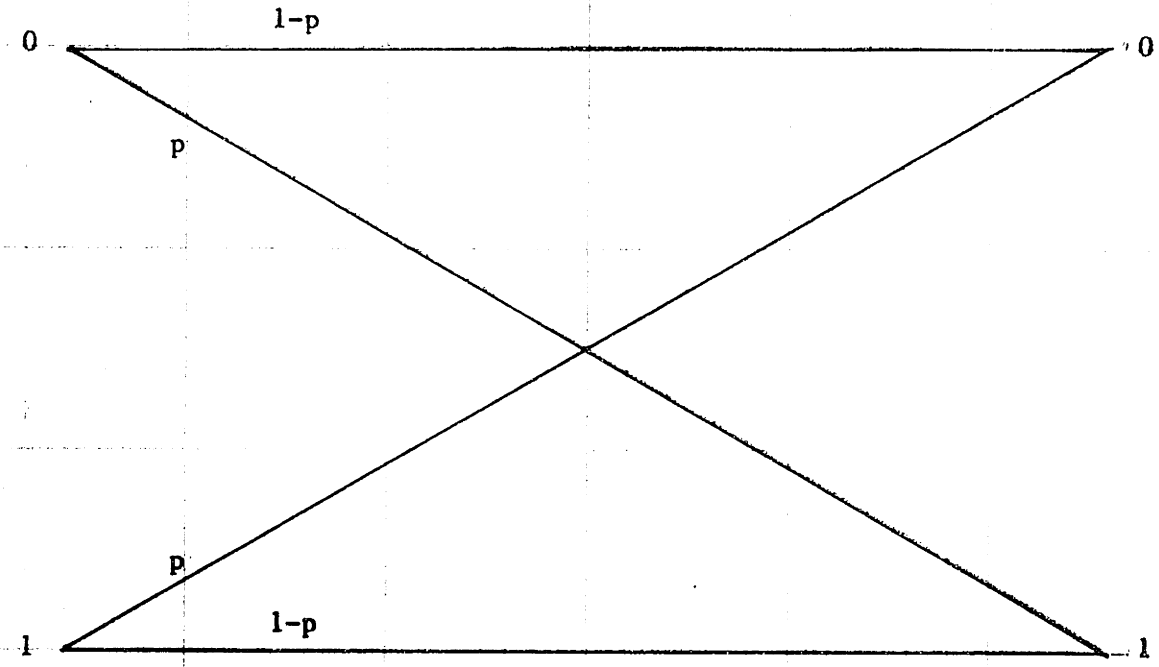
Chapter 9

CHANNEL NOISE

The concept of dividing a communications system into source coder and channel coder provides a powerful design approach. When appropriate channel coding can guarantee perfect transmission of up to C bits per second, the only constraint on the source coder design is the maximum output rate. However, sophisticated channel coding is not usually employed because of the complexity and transmission delay involved. It then becomes very important to analyze how channel errors affect the information content of the source coder data.

To assess the performance of the current system with channel noise, a binary symmetric channel with error rates of 10^{-4} and 10^{-2} was simulated on the PDP-9 computer. A binary symmetric channel, diagrammed in Figure 31, is an appropriate model for a link with additive white noise independent of the data. Channels that involve switching or radio links are plagued by noise that comes in bursts. If the duration of the burst almost equals the length of all the data, retransmission is necessary. However, for relatively short bursts, the data may be scrambled at the transmitter and unscrambled at the receiver so the net effect is that of a binary symmetric channel[10].

Before entering the channel, the data from the quantizer are mapped into a set of binary numbers by tagging the lowest quantizing



Binary Symmetric Channel

Figure 31

level (out of 2^Q levels) as 0 and the highest as 2^{Q-1}. The reverse of this mapping is performed as the data enter the source decoder. Figures 32(a) and 32(b) show 3 bit transmissions at the two error rates considered. Comparing these with Figure 28(e), we see that the effect of a single error is localized as a small dot. This should be contrasted with DPCM, where a single error can damage an entire line of data[3].

Error rate = 10^{-4}
(a)

Error rate = 10^{-2}
(b)

Transmission with Channel Noise

Figure 32

Chapter 10

CONCLUSIONS

In order to develop a real-time facsimile transmission system, a number of constraints and tradeoffs were imposed on the original design. Most significantly, the noise filters were changed from two to one-dimensional while attempting to maintain the same reduction in noise visibility. In addition a uniform quantizer plus non-linear clipping, both with fixed parameters, were introduced in the quantizing process. In the sampling process the filtering was removed from the logarithm domain. The impact on performance can be fully assessed only by processing actual pictures. Those in Figure 33 were chosen as test subjects because they represent a wide range of typical television or newspaper scenes.

Both the original system and the real-time system were simulated on the PDP-9 computer for a selected set of total bit transmission capacities (channel capacity multiplied by transmission time, CT). For each capacity value and subject, the choice of S, the number of samples per dimension, versus Q, the number of bits per sample, was adjusted to maximize picture quality ($CT = S \times S \times Q$). The results for the new system, as recorded by the flying-spot display, are presented in Figure 34. Work is in progress to interface the system to a laser facsimile scanner and receiver.

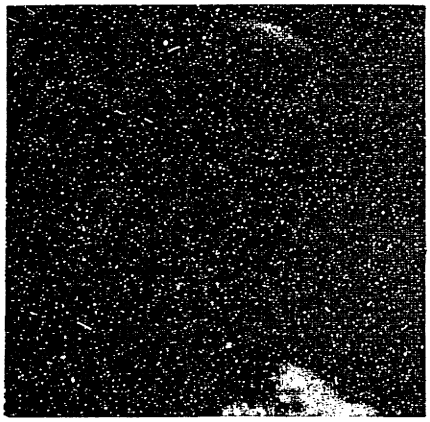
(a)

(b)

(c)

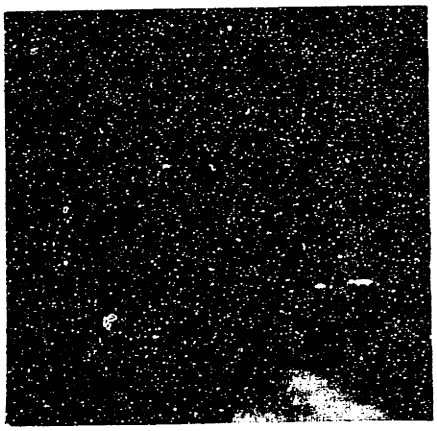
Test Subjects

Figure 33



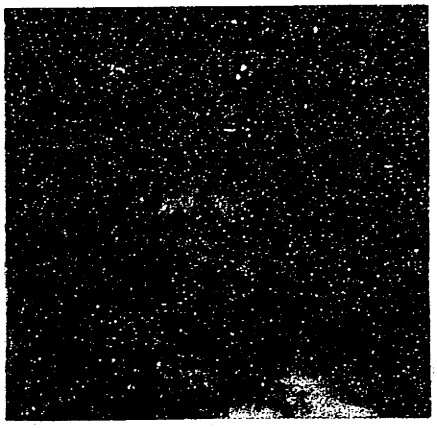
CT = 2^{18}, S = 256, Q = 4
(a)

CT = 2^{18}, S = 256, Q = 4
(d)



CT = 2^{17}, S = 209, Q = 3
(b)

CT = 2^{17}, S = 209, Q = 3
(e)



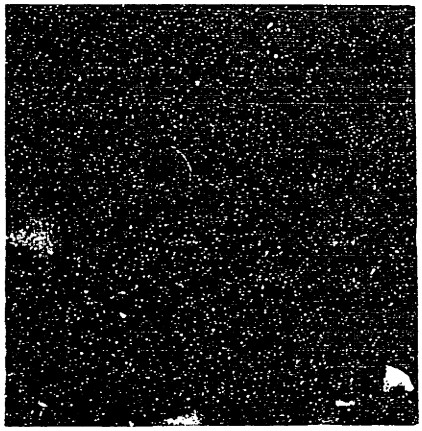
CT = 2^{16}, S = 148, Q = 3
(c)

CT = 2^{16}, S = 181, Q = 2
(f)

Performance of the Real-Time System

Figure 34

CT = 2^{18}, S = 256, Q = 4
(g)



CT = 2^{17}, S = 256, Q = 2
(h)

CT = 2^{16}, S = 181, Q = 2
(i)

Performance of the Real-Time System

Figure 34 (continued)

The performance of the real-time system nearly matches that of the original system. Therefore, the comparison scheme developed by the author in previous research[50] is not applicable here. It is noteworthy to compare these results with the theoretical work of Sakrison[38]. He proved that the rate distortion bound for line-by-line processing is about three times that for two-dimensional methods. However, his distortion measure effectively minimized the mean square difference between the input to the receiver noise filter and the output of the transmitter noise filter. The problems encountered with a mean square criterion in Chapter 5 confirm Sakrison's reservations about the correspondence between his distortion measure and subjective judgments of quality.

Extensions of this research into image enhancement may be fruitful. By including stages in the system that intentionally "distort" the image, it should be possible to compensate for losses in quality due to data compression. A successful technique has been developed by Schreiber[42] to improve Wirephoto pictures by increasing the micro contrast of the image to compensate for the loss of macro contrast in the print medium.

Although actual hardware development of this real-time facsimile transmission system may entail some modifications, the feasibility of such a system has been demonstrated.

BIBLIOGRAPHY

1. Anderson, G. B. and Huang, T. S., "Errors in Frequency-Domain Processing of Images," Spring Joint Computer Conference, 1969.

2. Anderson, G. B., Images and the Fourier Transform, S.M. Thesis, Massachusetts Institute of Technology, May, 1967.

3. Arguello, R. J., Sellner, H. R., and Stuller, J. A., "The Effect of Channel Errors in the Differential Pulse-Code-Modulation Transmission of Sampled Imagery," IEEE Transactions on Communication Technology, COM-19:6, December, 1971, pp. 926-933.

4. Brainard, R. C. and Candy, J. C., "Direct-Feedback Coders: Design and Performance with Television Signals," Proceedings of the IEEE, May, 1969.

5. Bruce, J. D., An Investigation of Optimum Quantizations, Ph.D. Thesis, Massachusetts Institute of Technology, May, 1964.

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

7. Evans, R. M., An Introduction to Color, New York, John Wiley and Sons, Inc., 1948.

8. Fiascinaro, J. G., Two-Dimensional Nonrecursive Digital Filters, Ph.D. Thesis, Massachusetts Institute of Technology, May, 1973.

9. Fink, D. G., ed., Color Television Standards, New York, McGraw-Hill Book Co., Inc., 1955.

10. Gallager, R. G., Information Theory and Reliable Communication, New York, John Wiley and Sons, Inc., 1968.

11. Gish, H. and Pierce, J. N., "Asymptotically Efficient Quantizing," IEEE Transactions on Information Theory, IT-14:5, September, 1968, pp. 676-683.

12. Goblick, T. J., Jr. and Holsinger, "Analog Source Digitization: A Comparison of Theory and Practice," IEEE Transactions on Information Theory, IT-12:2, April, 1967, pp. 323-326.

13. Gold, B. and Rader, C. M., Digital Processing of Signals, New York, McGraw-Hill, Inc., 1969, chapter 6.

14. Graham, D. N., "Image Transmission by Two-Dimensional Contour Coding," Proceedings of the IEEE, 55:3, March,1967, pp. 336-346.

15. Graham, D. N., Image Transmission Systems, 6.39 notes, Massachusetts Institute of Technology, May, 1961.

16. Graham, D. N., Two-Dimensional Filtering to Reduce the Effect of Quantizing Noise in Television, S.M. Thesis, Massachusetts Institute of Technology, May, 1966.

17. Groneman, U. F., Coding Color Pictures, Technical Report No. 422, Research Laboratory of Electronics, Massachusetts Institute of Technology, June, 1964.

18. Hofstetter, E. M., "A New Technique for the Design of Nonrecursive Digital Filters," Technical Note 1970-42 of Lincoln Lab, Massachusetts Institute of Technology, December 15, 1970.

19. Huang, T. S. Tretiak, O. J., Prasada, B., and Yamaguchi, Y., "Design Considerations in PCM Transmission of Low-Resolution Monochrome Still Pictures," Proceedings of the IEEE, 55:3, March 1967, pp. 331-335.

20. Huang, T. S., "Digital Picture Coding," Proceedings of the National Electronics Conference, Vol. XXII, 1966.

21. Huang, T. S., "PCM Picture Transmission," IEEE Spectrum, Vol. 2, December, 1965, pp. 57-63.

22. Huang, T. S. and Tretiak, O. J., "Research in Picture Processing," chapter 3 of Optical and Electro-optical Information Processing, ed. J. T. Tippett, Cambridge, M.I.T. Press, 1965.

23. Huang, T. S., "Restoration of Atmospherically Degraded Images," ed. S. Morgan, Report of N.A.S. Summer Study, Woods Hole, Massachusetts, July, 1966, app. 22, 23.

24. Hubel, D. H. and Wiesel, T. N., "Receptive Fields and Functional Architecture in Two Nonstriate Visual Areas (18 and 19) of the Cat," Journal of Neurophysiology, 28, 1965, pp. 228-289.

25. IBM, Random Number Generation and Testing," GC20-8011-0, December, 1969.

26. Limb, J. I., "Design of Dither Waveforms for Quantized Visual Signals," Bell System Technical Journal, 49:7, September, 1969, pp. 2556-2599.

27. Lippel, B., Kurland, M., and Marsh, A. H., "Ordered Dither Patterns for Coarse Quantization of Pictures," Proceedings of the IEEE, 59:3, March, 1971, pp. 429-431.

28. Lowery, E. M. and DePalma, J. J., "Sine Wave Response of the Visual System, I The Mach Phenomenon," Journal of the Optical Society of America, 51, July, 1961, p. 740.

29. Max, J., "Quantizing for Minimum Distortion," IRE Transactions on Information Theory, IT-6:3, March, 1960, pp. 7-12.

30. Mees, E. E. K., and James, T. H., eds., The Theory of the Photographic Process, New York, The Macmillan Company, 1966.

31. Mitchell, O. R., Jr., The Effect of Spatial Frequency on the Visibility of Unstructured Patterns, Ph.D. Thesis, Massachusetts Institute of Technology, September, 1972.

32. Mueller, C. C., Rudolf, M., and the editors of Life, Light and Vision, New York, Time, Inc., 1966.

33. Petersen, D. E., Error Correcting Codes, Cambridge, M.I.T. Press, 1967.

34. Petersen, D. P. and Middleton, D., "Sampling and Reconstruction of Wave-Number-Limited Functions in N-Dimensional Euclidean Spaces," Information and Control, 5:4, December, 1962, pp. 279-323.

35. Post, E. A., Computer Simulation of a Television Quantizing Noise Filtering System, S.B. Thesis, Massachusetts Institute of Technology, May, 1966.

36. Rabiner, L. R. and Schafer, R. W., "Recursive and Nonrecursive Realizations of Digital Filters Designed by Frequency Sampling Techniques," IEEE Transactions on Audio and Electroacoustics, AU-19:3, September, 1971, pp. 200-207.

37. Roberts, L. G., "Picture Coding Using Pseudo-Random Noise," IRE Transactions on Information Theory, IT-8:2, February, 1962, pp. 145-154.

38. Sakrison, D. J. and Algazi, V. R., "Comparison of Line-by-Line and Two-Dimensional Encoding of Random Images," IEEE Transactions on Information Theory, IT-17:4, July, 1971, pp. 386-398.

39. Schreiber, W. F., Huang, T. S., and Tretiak, O. J., "Contour Coding of Images," 1968 Wescon Technical Papers, Session 8.

40. Schreiber, W. F., "Picture Coding," Proceedings of the IEEE, 55:3, March, 1967, pp. 320-330.

41. Schreiber, W. F. and Knapp, E. F., "T. V. Bandwidth Reduction by Digital Coding," IRE National Convention Record, Part 4, 88, 1958.

42. Schreiber, W. F., "Wirephoto Quality Improvement by Unsharp Masking," Pattern Recognition, 2:2, May 1970, pp. 117-121.

43. Shack, R. V., "Image Processing by an Optical Analog Device," Pattern Recognition, 2:2, May, 1970, pp. 123-126.

44. Siebert, W. M., Signals and Systems, 6.05 notes, Massachusetts Institute of Technology, February, 1966, chapter 4.

45. Siegel, J., Design of Nonrecursive Approximations to Digital Filters with Discontinuous Frequency Responses, Sc.D. Thesis, Massachusetts Institute of Technology, June, 1972.

46. Stockham, T. G., Jr., Intra-Frame Encoding for Monochrome Images by Means of a Psychophysical Model Based on Nonlinear Filtering of Multiplied Signals, presented at the Symposium on Picture Bandwidth Compression at the Massachusetts Institute of Technology, April 4, 1969.

47. Thompson, M. J., "Digital Encoding for Video-telephone Transmission," 1968 Nerem Record.

48. Tretiak, O. J., Image Processing, 6.613 notes, Massachusetts Institute of Technology, February, 1970 (unpublished).

49. Tretiak, O. J., The Picture Sampling Process, Sc.D. Thesis, Massachusetts Institute of Technology, September, 1963.

50. Wacks, K. P., Analyzing the Performance of an Image Processing System, E.E. Thesis, Massachusetts Institute of Technolofy, May, 1970.

51. Wagner, H. G., MacNichol, R. F., and Wolbarsht, M. L., "Functional Basis for 'On' and 'Off' Center Receptive Fields in Retina," Journal of the Optical Society of America, 53, 1963, pp. 66-70.

52. Wyner, A. D., "Another Look at the Coding Theorem of Information Theory - A Tutorial," Proceedings of the IEEE, 58:6, pp. 894-913, June, 1970.

53. Zusne, L., Visual Perception of Form, New York, Academic Press, 1970.

54. Zworkin, V. K. and Morton, G. A., Television, New York, John Wiley and Sons, Inc., 1954.

Appendix A

DFT FOR TWO 4-STEP SEQUENCES

In Chapter 7 we discuss the power spectrum of two 4-step sequences used to produce structured noise. The computations of the spectra via the discrete Fourier transform (DFT)are presented here.

Given a function, f(n), defined for N points, n = 0 to N - 1, the DFT, F(k), is calculated thus:

$$F(k) = \sum_{n=0}^{N-1} f(n) \exp(-j2\pi nk/N), \quad 0 \leq k < N$$

Because f(n) is a 4-step sequence, it is periodic with a period of 4:

$$f(n + 4k) = f(n)$$

for k integer where f(n) is defined. It follows that F(k) = 0 for $k \neq bN/4$ for b integer. Therefore, we need to calculate only

$$F(bN/4) = \sum_{n=0}^{N-1} f(n) \exp(-j\pi nb/2)$$

for b = 0, 1, 2, and 3. The power density spectrum components, S(k), are the square of the magnitude of the corresponding DFT components.

Sequence 7.1 with q = 8 for convenience is:

$$f(n) = -3, 1, -1, 3, -3, 1, \ldots, 3$$

The corresponding DFT and power density spectrum are:

$F(0) = 0$ $\qquad$ $S(0) = 0$

$F(N/4) = (-1+j)N/2$ $\qquad$ $S(N/4) = N^2/2$

$F(N/2) = 2N$ $\qquad$ $S(N/2) = 4N^2$

$F(3N/4) = (-1-j)N/2$ $\qquad$ $S(3N/4) = N^2/2$

Sequence 7.2 with q = 8 as above:

$$f(n) = -3, -1, 1, 3, -3, -1, \ldots, 3$$

The corresponding DFT and power density spectrum are:

$F(0) = 0$ $\qquad$ $S(0) = 0$

$F(N/4) = (-1+j)N$ $\qquad$ $S(N/4) = 2N^2$

$F(N/2) = -N$ $\qquad$ $S(N/2) = N^2$

$F(3N/4) = (-1-j)N$ $\qquad$ $S(3N/4) = 2N^2$

Appendix B

FILTER DESIGN WITH CONSTRAINTS

In Chapter 8 we seek filter designs which limit the vertical extent of the impulse response of a two-dimensional filter while retaining certain frequency characteristics. If F(u,v) is the desired frequency response defined for $0 \leq u, v \leq N - 1$ and G(u,v) is the desired response of the filter with an impulse response, g(n,m), defined for $0 \leq n, m \leq N - 1$, we require:

g(n,m) = 0 for m > r and m < N - r, yielding 2r + 1 non-zero rows

$$G(u,0) = F(u,0)$$

The following Lagrangian formulation minimizes the mean square difference between F and G subject to these constraints:

$$L = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \left[G(u,v) - F(u,v)\right]^2 + \sum_{u=0}^{N-1} \lambda(u) \left[G(u,0) - F(u,0)\right]$$

$$G(u,v) = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} g(n,m) \exp\left[-w(un + vm)\right],$$

where $w = j2\pi/N$.

g(n,m) is determined by the solution to the following set of simultaneous equations:

$$\partial L/\partial g(x,y) = 0 \qquad \text{for } 0 \leq x, y \leq N - 1 \qquad \text{(B.1)}$$

$$\partial L/\partial \lambda(x) = 0 \qquad \text{for } 0 \leq x \leq N - 1 \qquad \text{(B.2)}$$

We consider the first term in L and where appropriate use the fact that

$$f(n,m) = f^*(n,m)$$

$$g(n,m) = g^*(n,m)$$

$$L_1 = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\left(G(u,v) - F(u,v)\right)^2$$

$$= \sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\left\{\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} g(n,m)\ \exp\left[-w(un + vm)\right] - F(u,v)\right\}\cdot$$

$$\left\{\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} g(n,m)\ \exp\left[+w(un + vm)\right] - F^*(u,v)\right\}$$

$$= \sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\left\{\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}\sum_{p=0}^{N-1}\sum_{q=0}^{N-1} g(n,m)\ g(p,q)\ \exp\left[-wu(n - p)\right.\right.$$

$$\left.-wv(m - q)\right] - F^*(u,v)\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} g(n,m)\ \exp\left[-w(un + vm)\right]$$

$$\left.-F(u,v)\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} g(n,m)\ \exp\left[+w(un + vm)\right] + |F(u,v)|^2\right\}$$

$$\partial L_1/\partial g(x,y) =$$

$$\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\partial/\partial g(x,y)\left\{g(x,y)\sum_{\substack{p=0\\p\neq x}}^{N-1}\sum_{\substack{q=0\\q\neq y}}^{N-1} g(p,q)\ \exp\left[-wu(x - p)\right.\right.$$

$$\left.-wv(y-q)\right] + g(x,y)\sum_{\substack{n=0\\n\neq x}}^{N-1}\sum_{\substack{m=0\\m\neq y}}^{N-1} g(n,m)\ \exp\left[-wu(n - x)\right.$$

$$\left.\left.-wv(m - y)\right] + g^2(x,y)\right\} - F^*(u,v)\ \exp\left[-w(ux + vy)\right]$$

$$-F(u,v)\ \exp\left[+w(un + vm)\right]$$

$$= \sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\left\{\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} g(n,m)\ \exp\left(+w(un+vm)\right)\ \exp\left(-w(ux+vy)\right)\right.$$

$$\left.+\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} g(n,m)\ \exp\left(-w(un+vm)\right)\ \exp\left(+w(ux+vy)\right)\right\}$$

$$-N^2 f^*(n,m) - N^2 f(n,m)$$

$$= \sum_{u=0}^{N-1}\sum_{v=0}^{N-1} G^*(u,v)\ \exp\left(-w(ux+vy)\right) + G(u,v)\ \exp\left(+w(ux+vy)\right)$$

$$-2N^2 f(n,m)$$

$$= 2N^2\left(g(x,y) - f(x,y)\right) \qquad (B.3)$$

$$L_2 = \sum_{u=0}^{N-1} \lambda(u)\left(G(u,0) - F(u,0)\right)$$

$$\partial L_2/\partial g(x,y) = \sum_{u=0}^{N-1} \lambda(u)\ \partial/\partial g(x,y)\left[\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} g(n,m)\ \exp(-wun)\right]$$

$$= \sum_{u=0}^{N-1} \lambda(u)\ \exp(-wux) \qquad (B.4)$$

Combining B.3 and B.4 in B.1:

$$g(x,y) = f(x,y) - 1/2N^2\left[\sum_{r=0}^{N-1} (u)\ \exp(-wux)\right] \qquad (B.5)$$

Define $L(x) = 1/2N^2\left[\sum_{r=0}^{N-1} \lambda(u)\ \exp(-wux)\right]$

B.2 implies $G(u,0) - F(u,0) = 0$

$$\sum_{n=0}^{N-1}\left[\sum_{m=0}^{N-1} g(n,m) - f(n,m)\right]\ \exp(-wun) = 0$$

We apply the constraint that

$$g(n,m) = 0 \text{ for } m \geq r,\ m \leq N-r$$

$$\left(\sum_{m=0}^{r} + \sum_{n=N-r}^{N-1} \right) g(n,m) - \sum_{m=0}^{N-1} f(n,m) = 0$$

Substituting B.5

$$\left(\sum_{m=0}^{r} + \sum_{m=N-r}^{N-1} \right) \left[f(n,m) + L(n) \right] - \sum_{m=0}^{N-1} f(n,m) = 0$$

$$(2r + 1)\ L(n) = \sum_{m=0}^{N-1} f(n,m) - \left(\sum_{m=0}^{r} + \sum_{m=N-r}^{N-1} \right) f(n,m)$$

$$L(n) = 1/(2r + 1) \sum_{m=r+1}^{N-r-1} f(n,m)$$

Finally we find that

$$g(n,m) = f(n,m) + 1/(2r + 1) \sum_{y=r+1}^{N-r-1} f(n,y)$$

$$\text{for } m \leq r,\ m \geq N-r$$

Appendix C

FILTER DESIGN WITH CONSTRAINSTS ON THE INVERSE

This filter design problem is similar to that of Appendix B, except that we seek to minimize the difference between

$$G(u,v) \text{ and } F(u,v)$$

$$1/G(u,v) \text{ and } 1/F(u,v)$$

Without considering the requirement that $G(u,0) = F(u,0)$,

$$L = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \left(G(u,v) - F(u,v)\right)^2 + \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \left(1/G(u,v) - 1/F(u,v)\right)^2$$

$$\partial L/\partial g(x,y) = 0$$

From Appendix B, (define L_1 as the first term)

$$\partial L_1/\partial g(x,y) = 2N^2\left(g(x,y) - f(x,y)\right)$$

We consider the second term, L_2

$$\partial L_2/\partial g(x,y) =$$

$$\sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \partial/\partial g(x,y) \left[1/|G(u,v)|^2 - 1/G(u,v)F^*(u,v) - 1/G^*(u,v)F(u,v) + 1/|F(u,v)|^2\right]$$

$$= \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \left\{1/|G(u,v)|^4 \; \partial/\partial g(x,y) \left[|G(u,v)|^2\right] - \partial/\partial g(x,y) \left[G(u,v)\right]/\left[G^2(u,v)F^*(u,v)\right] - \partial/\partial g(x,y) \left[G^*(u,v)\right]/\left[G^{*2}(u,v)F(u,v)\right]\right\}$$

$$= \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \left\{ 1/|G(u,v)|^4 \left[G^*(u,v) \exp\left(-w(un + vm)\right) \right.\right.$$

$$\left.\left. + G(u,v) \exp\left(+w(un + vm)\right) \right] \right.$$

$$- \exp\left(-w(un + vm)\right) / \left(G^2(u,v)F^*(u,v)\right)$$

$$\left. - \exp\left(+w(un + vm)\right) / \left(G^{*2}(u,v)F(u,v)\right) \right\}$$

Thus

$$\partial L/\partial g(x,y) =$$

$$2N^2\left(g(x,y) - f(x,y)\right) + \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} (S + S^*) = 0$$

where $S = \exp\left(-w(un + vm)\right) \left\{ 1/|G(u,v)|^2 - 1/\left(G(u,v)F^*(u,v)\right) \right\} / G(u,v)$

To solve this G and F should be expressed in the space domain. This yields N coupled non-linear equations which must be combined with the constraint on the vertical extent of g(n,m) to find the non-zero values of this function. At this point the problem is quite intractable even for computer solution.

BIOGRAPHICAL NOTE

Kenneth Paul Wacks is a native of Massachusetts. In 1970 he received the degrees of E.E., S.M., and S.B. from the Massachusetts Institute of Technology. He has been a teaching assistant in courses on network theory and electromechanics. During his doctorate program he was supported by a fellowship from the Fannie and John Hertz Foundation.

Mr. Wacks is a member of Tau Beta Pi, Eta Kappa Nu, Society of the Sigma Xi, and the IEEE.

Publications

"Pseudorandom Noise Generation for Use in Image Transmission," Quarterly Progress Report No. 109, Research Laboratory of Electronics, Massachusetts Institute of Technology, April 15, 1973.

"Psychophysical Considerations in the Design of an Image Transmission System," National Telecommunications Conference, December 6, 1972.

"Analyzing the Performance of an Image Processing System," Quarterly Progress Report No. 98, Research Laboratory of Electronics, Massachusetts Institute of Technology, July 15, 1970.

"Photomultiplier Gate for Stimulated-Spontaneous Light Scattering Discrimination," The Review of Scientific Instruments, 38:7, July, 1967 (coauthored with Dr. Francesco DeMartini).