

IMAGE CLASSIFICATION AND CODING FOR TRANSMISSION OVER COMMUNICATION CHANNELS

ABSTRACT

The varieties of documents which might be encountered in an image transmission system are first classified by human inspection, and then by information content in terms of ϵ -entropy. The term " ϵ -entropy" is used here to specify the bits of information per page required for a particular resolution " ϵ ". Then by statistical properties such as variance, root-mean-square, auto-correlation function, spectral power density and run length distribution curves are used to obtain the same classification. The application of these latter properties to simple machine implementation is reviewed for possible use in adaptive scanning systems for image transmission.

The different potential methods of coding facsimile data for data transmission are tabulated and reviewed. The systems are classified by the major type of document for which they might be useful. The choice of modulation system is also discussed as it interacts with the coding system.

Then a set of limiting conditions for an elementary image transmission are postulated. The list of possible systems is then screened to fit these limitations and the potentialities of the remaining scanning systems are reviewed. Then a feasibility model being built in San Jose ASDD Laboratory is described.

F. B. Wood, Editor

FOREWORD

This report consists of a compilation of notes edited by F. B. Wood, which were written or suggested by the following members of the department: R. M. Bennett, P. R. Daher, P. D. Dodd, G. F. Grometer, O. F. Meyer, I. D. Provazek, C. E. Schlaepfer and F. B. Wood. Some analyses made by Professor N. M. Abramson are also included.

IMAGE CLASSIFICATION
AND CODING FOR TRANSMISSION OVER COMMUNICATION
CHANNELS

ABSTRACT

FOREWORD

TABLE OF CONTENTS
AND
LIST OF TABLES AND FIGURES

I. INTRODUCTION. - Objective, reference to earlier time-bandwidth compression work for facsimile and television, conditions established for this study, and objectives of the feasibility model. References.

II. IMAGE CLASSIFICATION

A. BASIC CLASSES OF DOCUMENTS. - Continuous tone, type, and line drawings.

Table II-A ϵ -Entropy of Number of Bits of Information per page for Different Classes of Document.

Table II-B Notes on the Calculation of ϵ -Entropy and Redundancy Range for Examples of Table II-A.

B. BASIC CONCEPTS INFORMATION THEORY APPLIED TO CODING.

1. THEORY

Fig. 2.1 Sampling of Waveform

Fig. 2.2 Waveform Spectrum

2. OBJECT

3. METHOD.

Fig. 2.3 Trend of Development of Coding Scheme
(Redundancy)

Fig. 2.4 Trend of Development of Coding Scheme
(Equipment Cost)

4. SUMMARY OF PROBLEMS IN CODING

Fig. 2.5A Signal Conversion Model

5. EXAMPLE OF MODIFIED HUFFMAN CODING OF RUN LENGTHS

Table II-C Modified Huffman Coding of Run-Lengths.

Fig. 2.5B Comparison of Several Coding Schemes on
Average Code Word Length.

C. AN ILLUSTRATIVE EXAMPLE OF REDUCING REDUNDANCY.

Fig. 2.6 Process of Recoding Facsimile Data for
Compression.

D. PRECAUTIONS IN REFERRING THE COMPRESSION RATIO TO A PROPER BASE.

Fig. 2.7 Graphical Representation of Example and
Limits on Redundancy Removal for Typed
Documents.

III. THEORETICAL APPROACHES TO DEFINING INFORMATION CONTENT OF DIFFERENT CLASSES OF DOCUMENTS WITH EXAMPLES.

A. GRAPHICAL COMPARISON AND DESCRIPTION CAPACITY.

Fig. 3A.1 Detailed Classification of Documents

Fig. 3A.2 List of Speed Ranges for Ideal Adoptive Image
Transmission System.

B. SCANNING RESOLUTION

C. DEFINITION OF ϵ -Entropy

D. SAMPLE CALCULATION OF UPPER AND LOWER BOUNDS ON ϵ -Entropy

Table III Calculation of Sample Points on the ϵ -Entropy Graph

E. PARAMETERS POTENTIALLY SUITABLE FOR AUTOMATED DOCUMENT CLASSIFICATION (ONE DIMENSIONAL SPACE)

1. Logical Check for Presence or Absence of Signal
2. Fourier Spectrum of Signal
 - Fig. 3E.1 Gray Scale Calibration of Image Dissection
3. Elementary Statistical Properties
4. Higher Order Statistical Properties
5. Alternative Discrete Statistical Properties
 - Fig. 3E.2 Sample Waveform with Parameter Identification
 - Fig. 3E.3 Sample Output of Image Quality Measurement
6. Range of Change of Derivatives
7. Preliminary Discussion of Decision Properties
 - Fig. 3E.4 Signal Properties for Different Classes of Document
 - Fig. 3E.5 Sample Type
 - Fig. 3E.6 Sample Scan Line
 - Fig. 3E.7 Black and White Facsimile Signal
 - Fig. 3E.8 Random Telegraph Signal
 - Fig. 3E.9 Auto Correlation Function of Random Telegraph Signal ($a = 0.425$)
 - Fig. 3E.10 Spectral Power Density
 - Fig. 3E.11 Spectral Power Density Curves
 - Fig. 3E.12 Auto Correlation Function of Type Scan

F. GENERAL FORM OF ADAPTIVE SCANNING SYSTEM

Fig. 3F. 1 Possible Adaptive System of Different Codes Stepwise Approaching the Lower Boundary

Fig. 3F. 2 Adaptive Scanning System

Fig. 3F. 3 Adaptive Scanning System Using Variable Speed

G. PARAMETERS AND TECHNIQUES FOR TWO-DIMENSIONAL IMAGE PROCESSING

1. Introduction

2. General Model

Fig. 3G. 1 Model of Image Space and Object Space

Fig. 3G. 2 Coherent Source

Fig. 3G. 3 Incoherent Source

3. Fourier Transform Analysis

4. Impulse Response Analysis:
Coherent Source Systems

Fig. 3G. 4 Geometry of Diffraction Pattern

Fig. 3G. 5 Geometry of Coherent Source

5. Impulse Response:
Incoherent Source Systems

6. Resolution

Fig. 3G. 6 Resolution of Cube

Fig. 3G. 7 Resolution of Cylinder

7. Scanning

Fig. 3G. 8 Scanning

Fig. 3G. 9 Small Segment of Scanning Space

Fig. 3G. 10 Stimuli Function

Fig. 3G. 11 Pseudo-Periodic Signal

8. Applications
9. Bibliography
10. Symbols and Notation

IV. CATALOG OF DIFFERENT APPROACHES TO IMAGE TRANSMISSION

Table IV. Image Scanning and Coding Methods

A. GENERAL CASE, CONTINUOUS TONE

1. Basic Facsimile
2. Variable Velocity
 - Fig. 4.2A Variable Velocity TV System
 - Fig. 4.2B Spatial Frequency Curves
3. Partial Transmission
 - Fig. 4.3A Interpolation Method
 - Fig. 4.3B Edge Detection
 - Fig. 4.3C Error Tunneling Effect
4. Predictive Coding
5. Band Splitting
 - Fig. 4.5 Band Splitter
6. Analog Compression Coding
 - Fig. 4.6-1 Analog Compression
 - through-
 - Fig. 4.6-10
7. Deviation and Rate Coding
 - Fig. 4.7-1 Approximation of Waveform with a_o and a_h .
 - Fig. 4.7-2 System Diagram for Coding Rate and Deviation

Fig. 4.7-3 Frequency Detector

8. Adaptive Scanning with Bandwidth Switching

Fig. 4.8-1 Two Data Bandwidth

Fig. 4.8-2 Error Detection and Control Bandwidth

Fig. 4.8-3 Bandwidth Use vs Time

9. Optical Processing: (a) Electro-Optical Crystal, (b) Masks, (c) O-Modulation

Fig. 4.9-1 Schematic View of Photographic Auto Correlation

Fig. 4.9-2 Cross-Correlation Matrix of Alphabet Letter

Fig. 4.9-3 Grating at Angle $\phi = 90^\circ - \theta$

Fig. 4.9-4 Image of θ -Modulation Object in Fraunhofer Diffraction Plane.

B. HANDWRITTEN, TYPED, AND PRINTED TEXT DOCUMENTS

10. Run-Length Coding

Fig. 4.10-1 Counter and Buffer for Run-Length Codes

Fig. 4.10-2 Run Length Coding Storage

Fig. 4.10-3 Run Length Distribution- Typing

Fig. 4.10-3 Run Length Distribution - Typing and Graphic Image

Fig. 4.10-5 Sample Scan

Fig. 4.10-6 Logic

Fig. 4.10-7 Buffer Function

11. Adaptive Run Length Coding

Fig. 4.11-1 Scan and Pre-Scan

12. Buffer-Controlled Scanning for Run Length Coding

Fig. 4.12-1 Low Inertia Document Scanner

13. FM/PM Adaptive System

Fig. 4.13-1 Fiber-Optics Connection for Adaptive Scan

Fig. 4.13-2 Frequency (for FM) and Scanning Velocity vs Bits/Run.

14. Pattern Coding (Segments)

Fig. 4.14-1 Segments

15. Pattern Coding (Dots)

Fig. 4.15-1 Pattern Coding Matrix of Fiber-Optics

16. Full Line Scan

Fig. 4.16-1 Full Line Scan

17. Coordinate Coding

Fig. 4.17-1 Sector Layout for Coordinate Coding

18. Two-Dimensional Line Scan

Fig. 4.18-1 Two Dimensional Line Scan:
Reading and Coding

Fig. 4.18-2 Transmission Encoder and Receiving Decoder

19. Matrix Processing

Fig. 4.19-1 Matrix Processor

20. Optimum Block Length

(Details in Section V)

C. Line Drawings

21. Coordinate Engineering Drawing Compression

22. Cyclic Code Compression

23. Analog from Biological Systems

V. AN EXAMPLE OF IMAGE TRANSMISSION USING SPATIAL REDUNDANCY
AND OPTIMUM FIXED BLOCK LENGTH CODING (FEASIBILITY MODEL)

A. THEORETICAL BACKGROUND

Fig. 5.1 Optimum Block Length

Fig. 5.2 Distortion of Run-Length Distribution Due to Block Coding

B. DESCRIPTION OF MODEL

Fig. 5.3 Image Coder and Data Smoother

Fig. 5.4 Alternative Buffer Scheme

C. EXPERIMENTAL RESULTS

Fig. 5.5 Block-Length Coding Compression Results

D. EXPONENTIAL APPROXIMATION TO RUN LENGTH PROBABILITY

Fig. 5.6 Run Length

Fig. 5.7 Exponential Run Length Distribution & Block Coding

VI. CONCLUSIONS AND STATUS OF IMAGE TRANSMISSION

APPENDIX A: DESCRIPTION OF A SCANNER AS SIMULATION BY A
FORTRAN PROGRAM

1. Introduction

2. The Program

Fig. A. 1 The Scanner

Table A-1 Coding of Horizontal Speed Relationship

Table A-2 Coding of Vertical Skip Relationship

Fig. A. 2 The Image

3. Supplement - A General Program

Fig. A. 3 An Adaptive Scanning System

APPENDIX B. MODULATION SYSTEM

APPENDIX C: SPECTRA OF SOME PULSE SIGNALS

Fig. C. 1 Example of $y(t)$

Fig. C. 2 Rectangular Pulses

Fig. C. 3 Example of Power Spectral Density

Fig. C. 4 Random Telegraph Signal

Fig. C. 5 Generalized Random Telegraph Signal

Fig. C. 6 Power Spectral Density of Generalized Random Telegraph Signal for Two Values of P .

APPENDIX D: POWER SPECTRUM OF A RANDOM PROCESS GENERATED BY SCANNING EQUI-SPACED TYPE

Fig. D. 1 Type Specimens Under Analysis

Fig. D. 2 Scan of Line of Type

Fig. D. 3 Map of Allowable Type Positions

Fig. D. 4 Waveforms

Fig. D. 5 Autocorrelation of S_1

Fig. D. 6 Autocorrelation Function

Fig. D. 7 Spectral Density of First Term

Fig. D. 8 Power Spectrum of Square-Wave Modulated Random Process

IMAGE CLASSIFICATION AND CODING FOR TRANSMISSION OVER COMMUNICATION CHANNELS

I. INTRODUCTION

The function of this report is to summarize our studies of image classification, the possibility of automated classification for adaptive scanning systems, and to examine the major alternative coding systems. The objective is to develop a simple adaptive document scanning system which will adjust to various types of black and white documents restricted to two levels for efficient compression of the image data compared to straight facsimile systems. However to insure the most generality of results, the study first considers the most general cases of documents before making the more narrow study of the specific class.

Image transmission by linear scanning as in ordinary facsimile can be very slow or expensive, depending upon how trade-off of time and bandwidth is handled. Since most images have a large redundancy it is profitable to have a code compression system to minimize the time-bandwidth product. Although many images have high redundancy permitting savings in the time-bandwidth product of the order of magnitude of ten to one hundred, practical coding systems usually result in a saving factor of three to five, because of the great variation in signal distribution properties from image to image. Even with this saving most systems require large buffers

An improvement in time-bandwidth compression would be possible with an adaptive system. If a prescan device could determine what kind of a

document was being processed, the logic could switch to the appropriate compression coding system. To be able to consider such an adaptive system, first a knowledge of the types of documents and second, means of automatically identifying the class of documents are needed. These two stages: classification, and identification of the class of a document are examined in Sections II and III of this report. This is a problem which was suggested and defined by R. M. Bennett.

The principal references to the earlier work on Facsimile and Television Compression done in London, Polytechnic Institute of Brooklyn, Bell Telephone Laboratories, M.I.T., and Ford Instrument Company are listed in the bibliography, 1.1 - 1.8. A catalog of image coding systems considered on this project is included as Section IV in this report. They are grouped under sub-headings corresponding to their principal class of document application.

The procedure used was first to consider all potential coding systems and then narrow them down to those meeting certain specific conditions. The general method of attack was proposed by R. M. Bennett. Specific proposals were made by R. M. Bennett, P. R. Esher, P. D. Dodd, G. F. Grometer, O. M. Meyer, L. P. Provasek, C. E. Schlaepfer, and F. B. Wood. All the systems considered are tabulated in Table ^{IV} and are discussed in detail in Section IV.

The following set of conditions were used to screen the different proposals in order to plan the feasibility model described in Section V:

1. The problem is
 - a) to obtain intelligence from an image (line drawing, typed, printed

and handwritten information).

b) to convert this information into suitable form for transmission (say, over a 3000 baud line at a rate of n documents per unit of time).

c) to receive and reconvert the information into the form in which all transitions between black and white are coded with a tolerance of $\pm 1/2 \epsilon$ of the original spatial width, ϵ is the resolution of the scanning system.

2. The solution(s) to the problem will contain answers to:

a) what minimum amount of information can be set by with?

b) how can we strip the maximum amount of redundancy from the image?

c) what are the buffering requirements?

d) what is economic trade-off between complexity of terminal equipment and transmission costs.

3. The ultimate implementation will minimize transmission time (or bandwidth), complexity, cost, etc., and will hopefully present a novel approach.

4. The following restrictions apply to the reading and writing techniques:

a) reading and writing must be done incrementally provide a serial or continuous flow of information

b) a prescan at a variable distance ahead of the encoder scanning spot can be used.

c) multiple scans consisting of a row or scanning spot in a column or diagonal can be used in reading, both rows and writing

5. The first stage of the project was the preparation of a standard facsimile test pattern. This was done by taking a series of a standard. A two stage method of identification could be developed or modified to test the quality of the transmission.

The objective of the current project was to develop a single coding system which will be adaptable to various changes in grouping of lines vertically, and also to skew and to cover the range of handwriting. Different parts of the image could be described to have a feature by the use of a vertical line. A coding system then the detail of the image could be described. The first stage of the project was developed. The two stage method of identification could be developed. The two stage method of identification could be developed. The two stage method of identification could be developed.

asked the system to be flexible and to be able to handle a flexible laboratory type to experimentally with new ideas and to be able to handle data.

~~the~~ the system implementation may be some of the things that are done. The system implementation may be some of the things that are done. The system implementation may be some of the things that are done.

15

REFERENCES

1. A. E. Laemmel, "Coding Processes for Bandwidth Reduction in Picture Transmission", Report R 246-51, Microwave Research Institute, Polytechnic Institute of Brooklyn, Aug. 30, 1951.
2. B. M. Oliver, "Efficient Coding", Bell System Technical Journal, Vol 31, pp. 724 - 750, July 1952.
3. E. C. Cherry, G. G. Gouliant, "Some Possibilities for the Compression of Television Signals by Recoding", pp. 328 - 353, Communication Theory, London 1953
4. A. E. Laemmel, "Techniques for Increasing Facsimile Transmission Speed", 9/1/54
5. S. Deutsch, "A Note on Some Statistics Concerning Type-written or Printed Material", IRE Transactions on Information Theory, Vol. 17-3, pp. 147 - 148, June 1957.
6. W. S. Michael, W. O. Fleckenstein, and E. R. Kretzmer, "A Coded Facsimile System," 1957 IRE WESCON Convention Record, Part 2, p. 84.
7. H. L. Hoynes, D.T.Hage, "Stop-Go Scanning Saves Spectrum Space" Electronics 31:39, pp. 848, September 26, 1958
8. H. Wyle, T. Erb, and R. Banow, "Reduced Time Facsimile

These are listed in the bibliography of the report of the author at the end of the volume.

Transmission by Digital Coding", IRE Trans. on Comm. Systems

Vol. CS-9, No. 3, Sept. 1961, pp. 215-222.

9.

L. G. Roberts, "Picture Coding Using Pseudo-Random Noise,"

IRE Trans. on Inf. Theory, Feb., 1962, IT-8, No. 2, p. 145

II. IMAGE CLASSIFICATION AND REDUNDANCY

A. BASIC CLASSES OF DOCUMENTS

Documents can be easily put into three elementary categories of:

- (1) Continuous tone (including photographs)
- (2) Handwritten, typed, and printed text, or
- (3) Line Drawings

by human inspection. The numbers of bits per 8-1/2" by 11" page to describe these types of documents for 125 lines per inch is described in Table II-A. The choice of 125 lines per inch was based upon the use of a resolution of 0.02 centimeters or 50 lines per cm which gives 1.5×10^6 binary digits per page. For 100 lines per inch the 8-1/2" x 11" page would have 10^6 bits. The number of bits per page is defined as the entropy of the page. Since the entropy is a function of the resolution (ϵ), we specify " ϵ -entropy" to make certain that the resolution is defined. In mathematics, ϵ -entropy has a more general definition required by the extension to coordinate systems other than rectangular coordinates. For example, if an optical image is described in elliptic coordinates referred to a diffracting aperture, the physical size of a unit element of area or of volume changes with position in the field. In this report we use only the special case of rectangular coordinates in reference to " ϵ -entropy."

The column marked "redundancy range" in Table 1 has three ratios for each entry. The first is the binary digits per page (from basic resolution) divided by the upper bound. For typed pages this is $1.5 \times 10^6 / 3.5 \times 10^5 = 4.3$. This is the order of magnitude of time band-width compression that can be obtained by simple straightforward coding techniques using a single scan of the input data, when a large buffer

The second figure in the "redundancy range" column is a typical value of code compression that can be obtained from particular sets of statistics. If a code is designed to realize this greater compression on the basis of the statistics of documents in a given office or business with a given type style, the attempted use of the compression code with another set of documents may result in less efficient operation. To realize this greater efficiency, adaptive coding systems must be considered. The judicious choice of direction in designing adaptive systems requires the determination of the condition for minimum buffer requirements and the maximum overlapping of hardware elements for the different modes of operation in the adaptive system. This model is based upon at least two scans of the input data.

The third figure in the "redundancy range" column is based upon the lower bound of bits per page. To realize this upper bound on code compression (i.e. lower bound on bits per page to be sent over data channel) the logic must be increased to permit a sequence of many scans. In fact as we push beyond the second value in the redundancy column our problem is being transformed into the character recognition problem. This can be done by a human key-punch operator or by the 1418 optical reader.

For line drawings the "redundancy range" column has a similar meaning, however there are some additional restrictions. The second value, namely 115, is for a particular line drawing which may be simpler than the typical range. The third value, 710, is an approximate value for conversion in BLODI computer computer language, which is the character recognition equivalent for the line drawings encountered is Bell Telephone Labs circuit diagrams. To represent this requires a human operator to examine the drawing and transform it into the format that can be represented in BLODI.

P. D. Dodd has suggested an intermediate level adaptive system for image transmission where the human preparation of the original document includes typing the code designation of which code-compression system is applicable to the document.

TABLE II-A

ε-Entropy of Number of Bits of Information
Per Page for Different Classes of Documents.

	Bits/Page (Upper Bound)	Bits/Page (Lower Bound)	Redundancy Range ** (relative to 1.5×10^6 bits)
Photographs (a)† (10 levels of grey)	1.5×10^6 decimal digits or Hartleys 5×10^6 binary digits or Binitis (d)	1.0×10^6 Binitis* (e)	(m)
Typed or Printed (b) (2 levels)	3.5×10^5 (f)	3.2×10^4 (h)	(i) (g) (k) 4.3/12.8/47.0
Line Drawing (c) (2 levels)	6×10^4 (i)	2.1×10^3	(l) (j) (n) 25/115/710
			one/two/many
			Number of Scans

*estimated value

** $\epsilon = 0.02$ cm

+ References (a)-(m) listed in Table II-B.

TABLE II-B

Notes on the Calculation of E-Entropy and Redundancy Range for Examples of Table II-A.

<u>Case</u>	<u>Notes</u>
(a)	Grey scale resolution of ten levels with scanning resolution of 120 lines per inch.
(b)	Two levels: black and white, dot-point and pic (1/2 dot) used with blank space allowance corresponding to 100% black.
(c)	Two levels: black and white. Simple arithmetic logic electric circuit block diagram used for these estimates.
(d)	I (binits) = I (Hartleys) / $\log 2$ = 10.3
(e)	Estimated from the lower value of I obtained at 100 lines/inch, needed in picture coding mentioned in L. J. Roberts, "Picture Coding Using Pseudo-Random Noise," <u>IRE Trans. on Info. Theory</u> Feb. 1962, IT-8, No. 2, p. 107.
(f)	Calculated by description method as technique for typing with margin eliminated (40 lines/inch). Also comes close to F Instrument Co. analysis, <u>IRE Trans. on Comm. Systems</u> Vol. CS-9, No. 3, Sept. 1961, p. 215.
(g)	Ratio of black to white on typed page 7.4% gives 13.5 reduction. Ratio of 12.8 comes from BTL study, W. S. Michael, 1957 IRE WESCON CONV. RECORD, Part 3, p. 94, using simplified variable length code approximation Shannon-Fano code.
(h)	4.75 bit per character alphabetic code requiring character recognition logic techniques.
(i)	Simplified variable length code designed for typing statistics, used on BTL line drawing (art) 1957 Wescon.
(j)	Huffman code designed for line drawing in 13-0-12, 13-1-12.
(k)	Translation of BTL Circuit Block Diagram to Compiler Language (BJTJ, 1961, p. 76)
(m)	No comparable data available. However, compression ratio of 2.85:1 for TV pictures has been reported by J. A. Cherry (<u>Electronic News</u> , 9/11/62)

B. SOME BASIC CONCEPTS OF INFORMATION THEORY APPLIED TO IMAGE COMPRESSION

1. THEORY

Consider a scanned image: Let $s(t)$ be the ensemble of scanner output waveforms, one member of which is shown in Figure 2.1 and let $S(f)$ be the spectrum of $s(t)$ as shown in Figure 2.2 (note that if the image data is black white only, the signal $s(t)$ will be some form of random square wave). If $S(f)$ is band limited as shown, $s(t)$ is completely described as a vector in $2FT$ dimensional signal space.

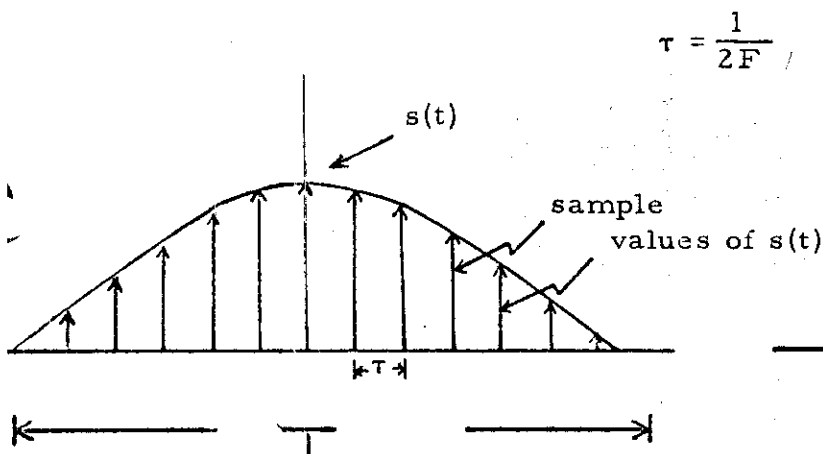


Figure 2.1: Scanner Output

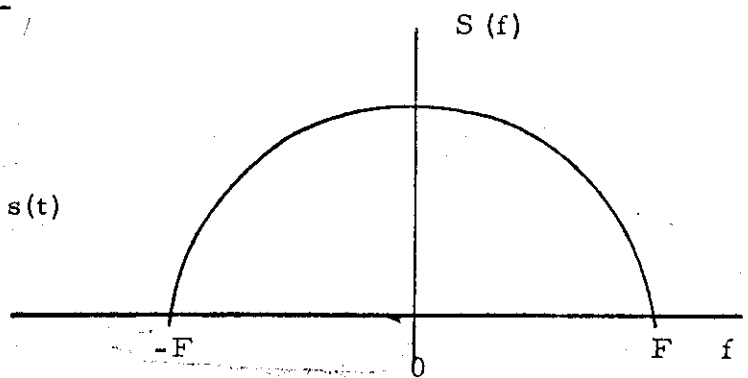


Figure 2.2: Spectrum of Scanner Output

Now quantize signal samples into n discrete levels so that the n -dimensional sample space contains sample points s . Information theory gives us a measure of the average amount of information (entropy) conveyed by each signal sample. Thus, if no knowledge of the past or future is available, the average amount of information required to specify a particular present sample $s(o)$ is

$$H(\sigma) = - \sum_i P(s_i) \log P(s_i) \text{ bits}$$

where $P(s_i)$ is the probability of the present sample value s_i
 σ is the space of sample values.

An upper bound on the entropy is

$$H(\sigma) \leq \log n$$

with equality if and only if all sample values are equally likely. If, however we can utilize a knowledge of the past and future samples, it is possible to reduce the amount of information required to specify the present sample. This reduced amount of information is

$$H(\sigma/\rho) = - \sum_{ij} P(s_i, r_j) \log P(s_i/r_j) \text{ bits}$$

where σ is the space of the present sample values, s_i
 ρ is the space of all past and future sample values r_j

(Note that ρ is the set $(\rho - \infty, \dots, \rho - 2, \rho - 1, \rho_1, \rho_2, \dots, \rho_\infty)$

(It can be shown that $H(\sigma/\rho) < H(\sigma)$ when the samples of s are correlated - see Fano, "Transmission of Information", J. Wiley, 1961, p. 45).

We thus define the redundancy as the average mutual information

$$\text{Red}(\sigma) = I(\sigma;\rho) = H(\sigma) - H(\sigma/\rho).$$

Thus, if we can take a source which is transmitting information at an information rate

$$R_1 = \frac{H(\sigma/\rho)}{\tau} \text{ bits per sec.}$$

and remove the redundant information from each sample so that successive samples are uncorrelated, we could transmit the source samples at the higher information rate

$$R_2 = \frac{H(\sigma)}{\tau} \text{ bits per sec.}$$

and thus transmit more information at the same sample rate or bandwidth.

Further, if we make all sample values equally likely through some transformation, we can make R_c a maximum - i. e., we can make optimum use of the channel bandwidth.

OBJECT

The object is then to convert the primary source distribution from arbitrary statistics to the statistics of an original, uncorrelated samples in order to relieve a secondary source with maximum entropy - i. e., a maximum amount of information conveyed per sample. Compression coding is the technique whereby this transformation is achieved.

Method:

- A) ~~Shannon-Fano~~ Code: converts peaked first order distribution of a primary source with $\phi(\tau) = \delta(\tau)$ to a uniform distribution in the secondary source - thus achieving maximum source entropy. $\phi(\tau)$ is the autocorrelation function.
- B) De-correlation coding: Converts arbitrary second, third, . . . nth order statistics (correlation) into a source in which the individual samples are independent, and the first order statistics are peaked in order that Shannon-Fano coding may be utilized.

Thus, several conversion stages are involved, first to de-correlate the source, then to encode it such that each sample (set of bits), or symbol transmits a maximum amount of information.

~~Unresolved: This source distribution does not give maximum C, however. Does this mean a third conversion from uniform statistics to Gaussian statistics for transmission over continuous channel?~~

Thus, the image conversion problem falls into four categories of coding to conserve power and time or bandwidth (ξ thus $\$$)

(1) De-correlation - compression coding: - (hereinafter called source compression coding) here no fidelity is lost - the noiseless channel can transmit just enough information to completely specify the source.

(2) Correlation coding: here, only part of the source is transmitted (e.g. alternate samples) - the known (pre-programmed) statistics of the source are used at the receiver to "fill in the blanks". There is some loss of fidelity here since the interpolation or extrapolation process will always involve some error.

(3) Bandwidth Compression Coding: here, the statistics of the source is not considered. Novel modulation schemes are used to compress bandwidth (at perhaps an expense of signal-to-noise ratio).

(4) Analog Coding: ^{or Signal Design} this is the analog equivalent of (1) above. Although this area is relatively unexplored, it should be expected to lend itself more readily to analog sources and would possibly provide greater B. W. savings. Properties and examples of analog codes should be investigated. Signal characteristics such as area, slope, moments, discontinuities,

frequency, etc. are used to convey information. Present continuous modulation systems AM, FM, SSB, A, etc. are examples of analog coding. However it is expected that there is a far greater class of analog transformations which will yield good B. W. and power savings and be more compatible with the source.

Various coding schemes might be compared by several graphical means such as redundancy removed vs fidelity ratio in Fig 2.3

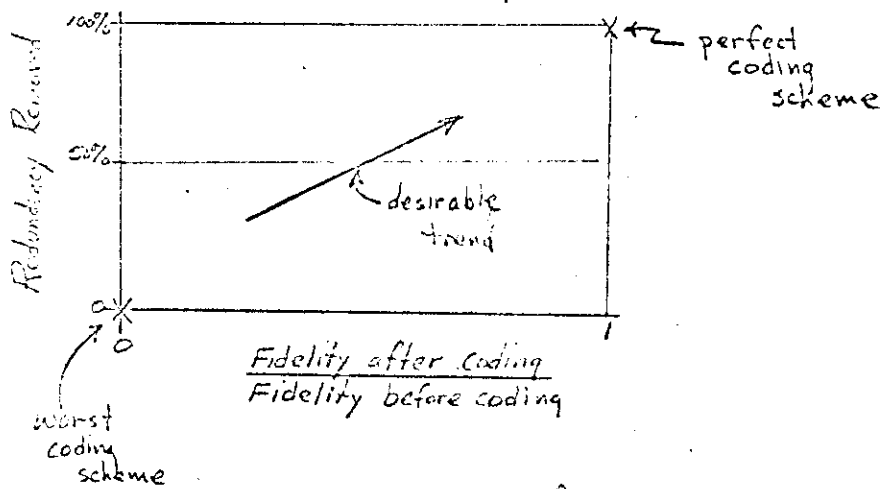


Fig 2.3 Trend of Development of Coding Scheme (Redundancy)

and equipment cost ratio vs redundancy removed in Fig 2.4

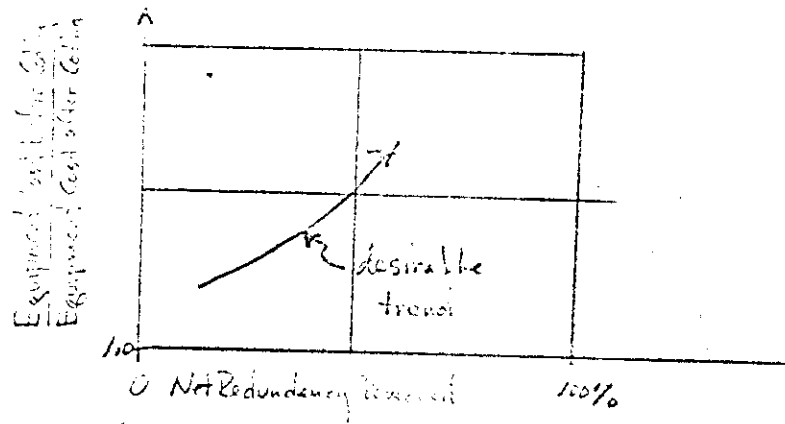


Fig 2.4 Trend of Development of Coding Scheme (Equipment Cost)

4. Summary of Problems in Coding:

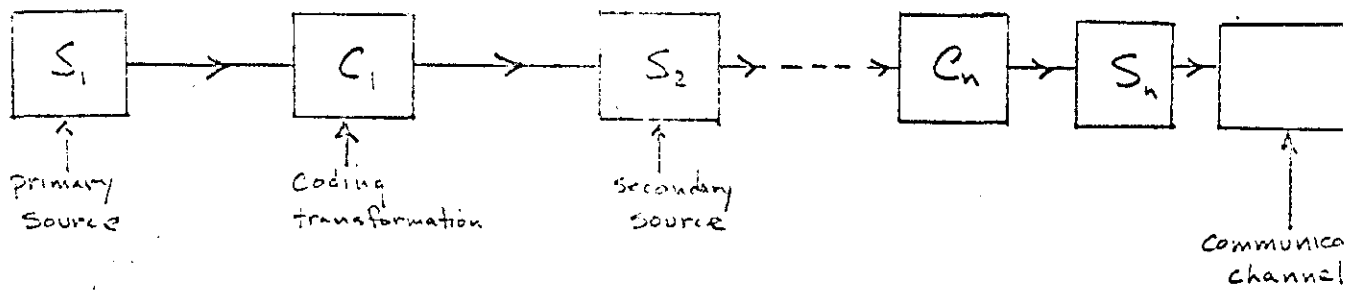
- (1) Is there an optimum transformation for "decorrelation" - - one that is easily implemented by, say, FSRs? *wavy scribble*
- (2) What other transformation besides "prediction" ^{and "differencing"} (ref. Oliver, Harrison, BSTJ) are suitable for de-correlating? How are they implemented?
- (3) "Analog" equivalent to digital coding? *scribble*
- (4) Hybrid schemes (ref. Schreiber - Technicolor) i.e., partly digital, partly analog.
- (5) Comparison of above - choose "best" method and implement.

NOTE: Two fundamental hypotheses should be observed:

- (1) The most efficient coding system is one which completely removes all redundancy from the source leaving only sufficient information to completely reconstruct the source. Then, error correction and detection coding is inserted such that the code structure matches the noise characteristics of the transmission channel. This "structure matching" characteristic of error correction/detection coding is a better method of providing redundancy for error-free reception than only removing part of the original source redundancy.
- (2) If a transformation (T) is required to remove all signal redundancy at the transmitter, then $(T)^{-1}$ must be realizable

and provided at the receiver for proper reconstruction of the source (decoding).

The general signal conversion model is shown below in Fig 2.5



$[T]$ = transformation between S_1 & S_n

$[T]$ = function of C_1, C_2, \dots, C_n



Fig 2.5A Signal Conversion Model

~~P. W. Dodd~~

~~2/27/62~~

5. EXAMPLE OF MODIFIED HUFFMAN CODING OF RUN LENGTHS

Consider the following illustrative example of variable-length compression coding. We utilize the run-length statistics in Michael, et al, "A Coded Facsimile System", IRE Wescon Convention Record, Part 2 1957. In order to avoid complex coding of an extremely large alphabet, we compression code only the ten most probable run-lengths plus some special symbols. All others are coded with a special prefix plus a ten-digit binary number to designate the count. Table I (II-B) shows the results of Huffman coding this source.

Table II-C.

Example of a Modified Huffman Coding
of Run Lengths

<u>Symbol</u>	<u>prob. / occurrence</u>	<u>Modified Huffman Code</u>
"Black"	.200	10
2	.180	000
3	.180	001
4	.100	110
5	.070	0100
"All others" *	.066	0101
6	.050	0111
"Margin"	.050	1110
7	.031	01100
1	.025	11110
8	.022	11111
9	.015	011010
10	.010	011011

* Note "All others" - all other run lengths greater than
have 0101 + 10 digits run length.

The amount of comparison that can be achieved with a coding of this type can be computed from the average run length, T

$$T = \sum_i L_i P_i$$

where

L_i = length of code word corresponding to i th symbol

P_i = probability of occurrence of i th symbol.

Figure ^{2.53} shows a graph of the average length of this code, compared with the average length of the code described in Michael, et al. For comparison, the estimated average length of straight facsimile, a Shannon-Fano coding of the source (non-optimum) and representation of run length by a straight 10 binit binary number are also given.

COMPARISON OF SEVERAL CODING SCHEMES

ON AVERAGE CODE WORD LENGTH

For frequency: $\bar{L} = \sum_{i=1}^{850} i P(i)$

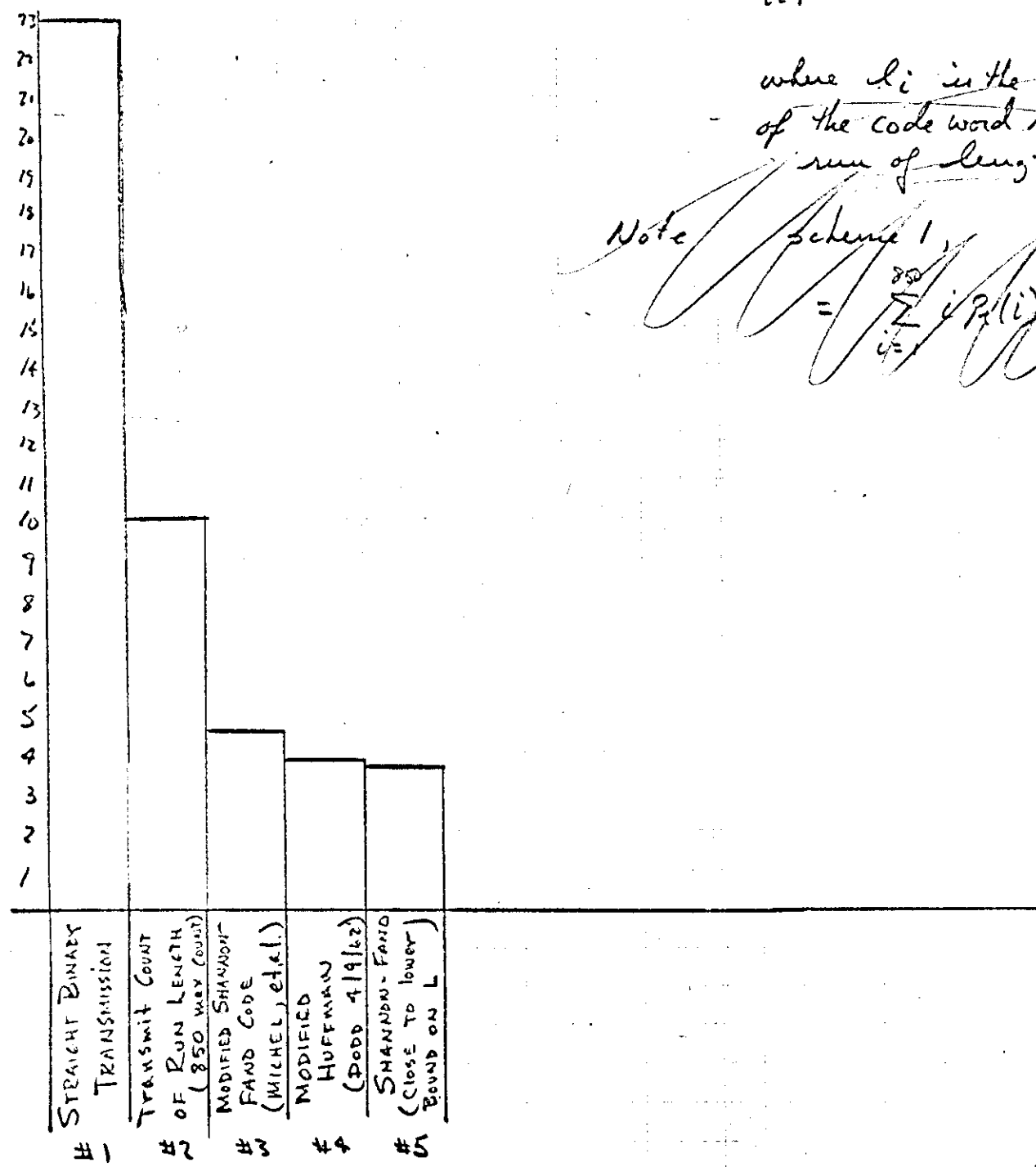
$$\bar{L} = \frac{\sum_{i=1}^{850} i P(i)}{\sum_{i=1}^{850} P(i)}$$

where i is the length of the code word represents run of length i

Note Scheme 1,

$$= \sum_{i=1}^{850} i P(i)$$

AVERAGE CODE WORD LENGTH



CODING SCHEME

Ty 2.5B

P. D. Dodd
4/13/62

c. An Illustrative Example of Reducing Redundancy

An $8\frac{1}{2}'' \times 11''$ page of Teletype is considered, first by ordinary facsimile scanning, and then through successive stages of code compression through character recognition. Elite type is assumed except where noted.

A straight facsimile scan is represented in Fig. 2.6A for $\epsilon = 0.02$ or 125 lines/inch giving $1060 \times 1375 = 1.455 \times 10^6$ bits/page. The average length of a facsimile scan is $e = \sum l_p(\epsilon) = 1060$ bits.

The results of counting run lengths of solid black or white are shown in Fig. 2.6B. The average run length is k bits and there are n runs per line in the average, making $n \times h$ runs per page.

A recoding of each string of zeros or ones into a ten-bit binary count gives the compression indicated in Fig. 2.6C. Examination of the probability distribution of the number of digits required to represent the run length counts indicates a recoding could improve the efficiency. Fig. 2.6D illustrates how a Shannon-Fano, Huffman, or simplified variable length code could reduce the average number of bits per run (k). Part of the saving would be lost due to addition of bits for synchronization (SYNC)

If s lines are grouped together by using s fiber optic scanning heads the vertical line scan (h) are divided by s , and the average run length count (k) is increased but making a net increase in compression, if a suitable code is employed as shown in Fig. 2.6E. This compression is typical of the lower bound on information per page for optimum recoding of typed documents.

* Elite type: 6 characters per vertical inch, 12 characters per horizontal inch.
Pica type: 6 characters vertical; 10 char/in horizontal

To obtain an ultimate limit, let us examine the case where we have the logic required for character recognition. This is shown in Fig. 2.6F both for elite and pica type. The ultimate limit, but not easily realizable is shown in Fig. 2.6G for character recognition based upon the average entropy per character of the English alphabet of 4.75 bits/letter.

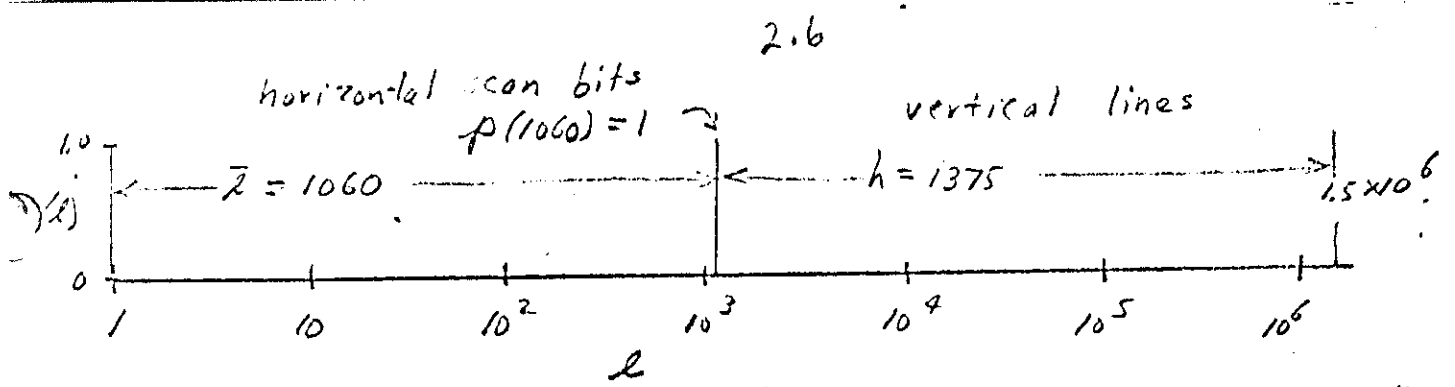


Fig. 2.6A Straight Facsimile Scan. ($\epsilon = 0.02$ cm, 125 lines/inch)

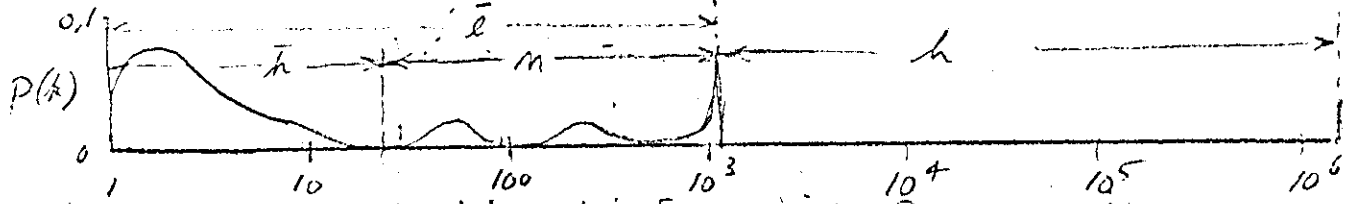


Fig. 2.6B Revision of Line Scans into Runs of All Zeros and All Ones.

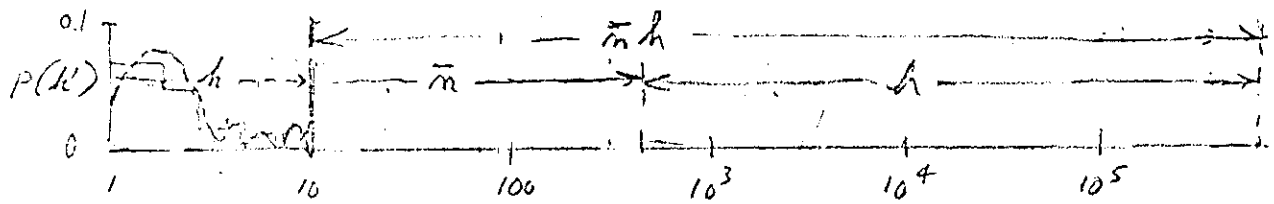


Fig. 2.6C - Recoding of Run Lengths into Binary Counts (10 bits)

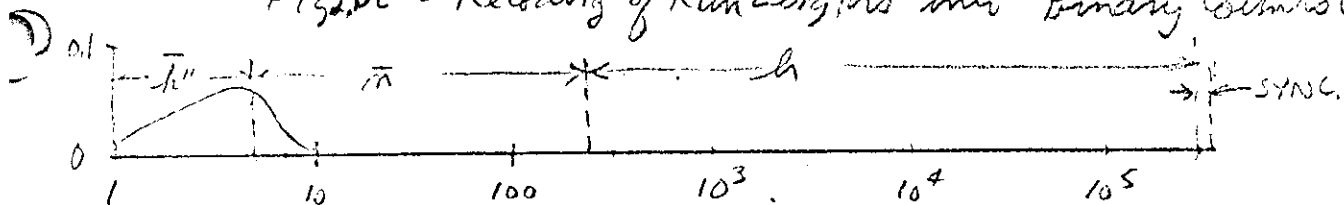


Fig. 2.6D - Optimum Recoding of Run Lengths.

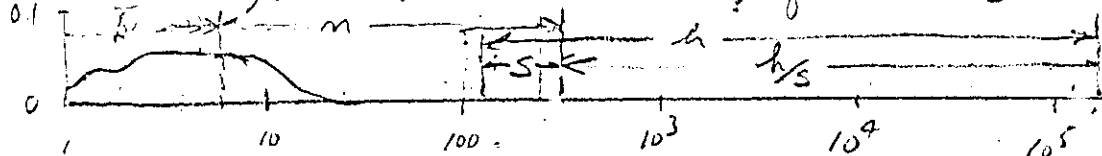


Fig. 2.6E - Multiple Line Scan to Use Vertical (Spatial) Redundancy.

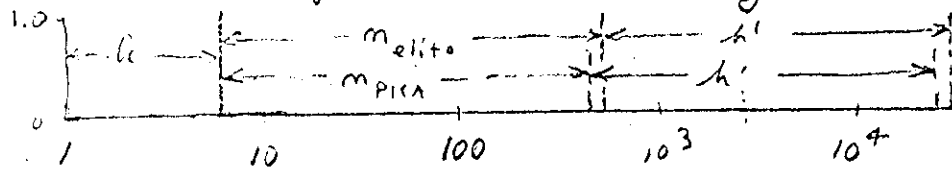


Fig. 2.6F - Character Recognition. (6 bits)

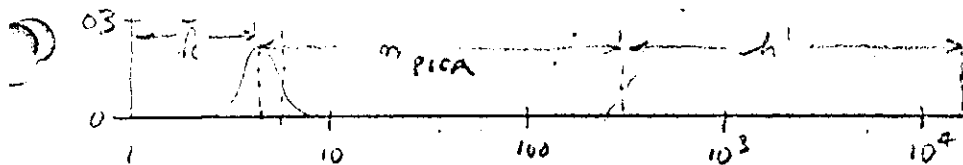


Fig. 2.6G - Character Recognition ($\bar{l} = 4.75$ bits)

C. Precautions in Referring the Compression Ratio to a Proper Base

Although the redundancy ranges in Table I are marked "relative to 1.5×10^6 bits", there is the possibility that redundancy or compression ratios computed in reference to other standards may appear inconsistent. In the next section a more precise analysis will be made in terms of " ξ -entropy" where " ξ " is the scanning resolution.

Before proceeding to the detailed study, a graphical plot in Fig. 2.7 will be used to show the difference in compression ratios computed in regard to different values of " ξ ". This example will deal with typed copy only.

Fig. 2.7 parts (A) - (C) show the bits per page of $\xi = 0.031$, 0.020, 0.0125, i. e., for 100, 125, and 200 lines/inch respectively. Fig. 2.7(d) shows an approximate method of finding order of magnitude compression possible. The $\xi = 0.031$ reference is used as a base, then the entropy is reduced by the fraction of space on the page devoted to margins, then further reduced by the space occupied by blank lines. Then a further rough approximation is made by using the ratio of white to black space in the typed lines. This gives a compression ratio of 14 which is slightly higher than the actual cases of 10, 11 and 128 in Figs. 2.7I, 2.7J, and 2.7H.

Note that if one started with $\xi = 0.0125$, a compression ratio of 37.8 would be obtained for PICA type in Fig. 2.7I. To avoid this

2.3

ambiguity, we must determine the largest ϵ for which the characters can be duplicated. If exact duplication is not necessary, but the requirement is that the images must be clearly and unambiguously recognizable, then the solid character of Fig. 2N^{2.7N} could be allowed to deteriorate to the dotted character of Fig. 2P^{2.7P}. By this criterion the limiting ϵ is 0.031, setting the compression ratio at 10 in Fig. 2I^{2.7I}.

The lines in Fig. 2E-G^{2.7} represent a case that is half typing, a quarter line drawings and a quarter blank. Fig. 2G^{2.7} is the combined results of E, F, and G. For $\epsilon = 0.020$ the compression ratio is 25 and for $\epsilon = 0.031$ it is 17. Further detail is needed on the required resolution to determine which compression ratio is a true limit.

Fig. 2J^{2.7} illustrates an important point ~~namely~~ namely the compression ratio of 11 for elite type is not increased when pica type is used, because ϵ changes as indicated.

Figs. 2(K, L, M)^{2.7} shows the limits attainable for character recognition when one can afford the computer logic required for character recognition in addition to accepting the restrictions of a fixed alphabet instead of being able to process any image of a given resolution.

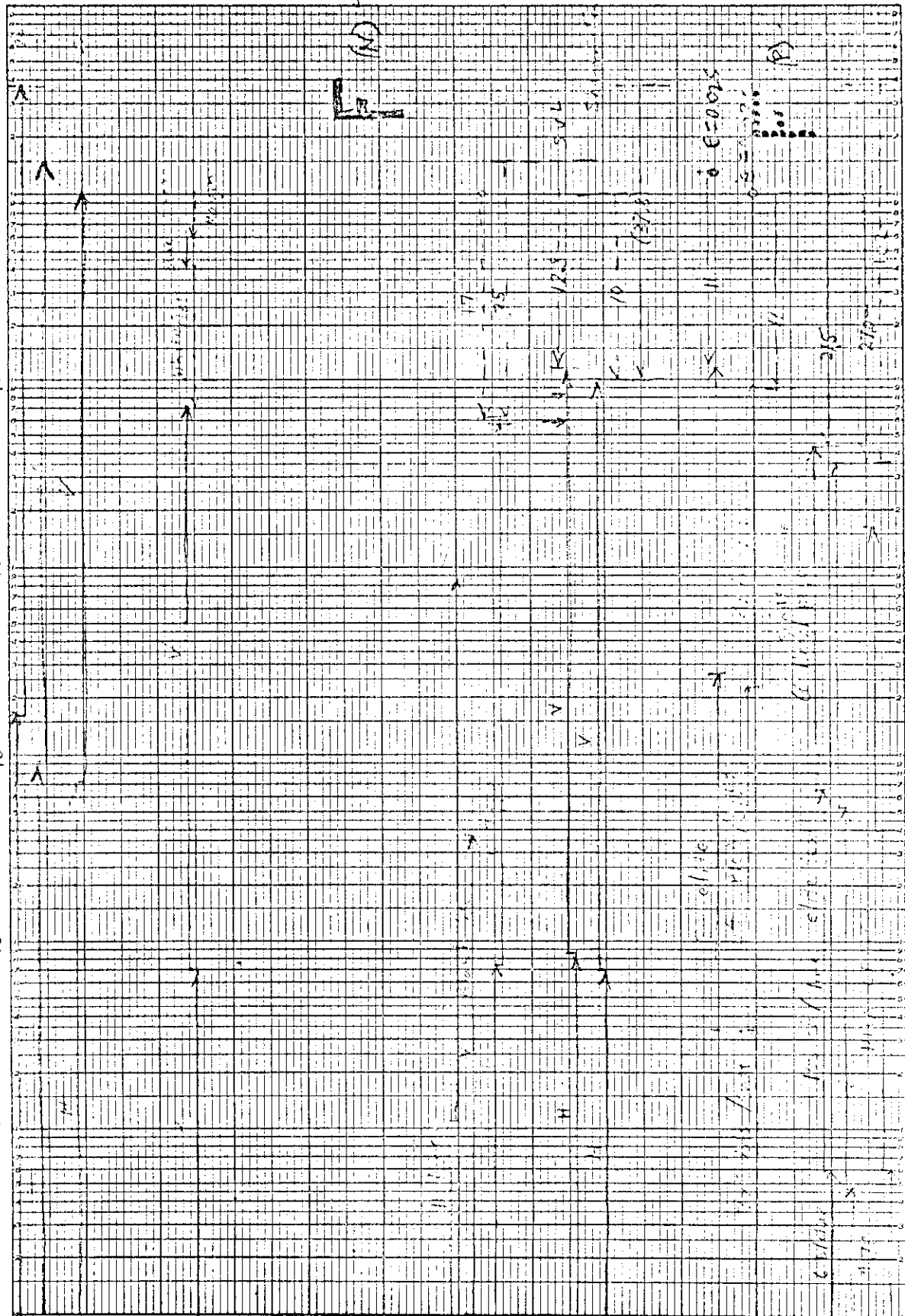
7-3-62 FBSW

F100 8/10
F110 12/10



10¹ 10² 10³ 10⁴ 10⁵ 10⁶ 10⁷

(A) (B) (C) (D) (E) (F) (G) (H) (I) (J) (K) (L) (M)



$\epsilon = 0.0125$
 $\epsilon = 0.02$
 $\epsilon = 0.031$

III. THEORETICAL APPROACHES TO DEFINING INFORMATION CONTENT OF DIFFERENT CLASSES OF DOCUMENTS

There are several approaches to calculating the information content. If we have a way of estimating the number of different images (w) that can be represented by a page with a specified resolution we can call this number the description capacity (w). If we take each of these possible images as equally likely, i. e., $p = 1/w$, then, the entropy is

$$I = \sum_{i=1}^w p_i \log p_i = -w (1/w \log w) = -\log w.$$

Since we must specify a particular resolution " ϵ " it is more consistent to designate the information as " ϵ -entropy" for proper identification. If we take the logarithm to the base two the units of " ϵ -entropy" will be binary digits or bits.

Since we do not have detailed knowledge of the conditional probabilities of black and white spots in documents, we must approach the problem by finding coding systems which when used with experimental statistics give upper bounds on ϵ -entropy. For lower bounds we look for character recognition examples or simple documents for which we can construct a message from the document can be redrawn at the receiver.

There are other properties such as the mean value, r. m. s. value, variance, and the auto correlation function which have empirical

relationships to the information content and hence the class of document. Therefore the automatic computation of certain statistical properties of an image could lead to the classification of the document in an adaptive system. We shall consider the following aspects.

- (a) Description capacity and Graphical comparison
- (b) Scanning Resolution
- (c) Definition of ϵ -entropy
- (d) Sample calculation of upper and lower bounds on ϵ -entropy
- (e) Parameter potentially suitable for automated document classification (one dimensional)
- (f) Parameters and techniques for two-dimensional image processing.

GRAPHICAL COMPARISON OF

A. Description Capacity

These are some preliminary comments on the relevancy of McLachlan's description mechanics

to the classification of images. Description mechanics applies to physical systems, including images but lacks generality beyond providing upper bounds.

The paper by Dan McLachlan of ~~SR~~ ^{Stanford Research Institute} on "Description Mechanics," provides a useful method of comparison of the information per page under different conditions ^{3.2} ~~(2)~~. The "descriptive capacity" is a static concept describing the number of different patterns or messages that can be represented by a given area with a specified resolution and set of restraints. The information content is defined as:

$$I = \log_2 W \quad (1)$$

where W is the descriptive capacity.

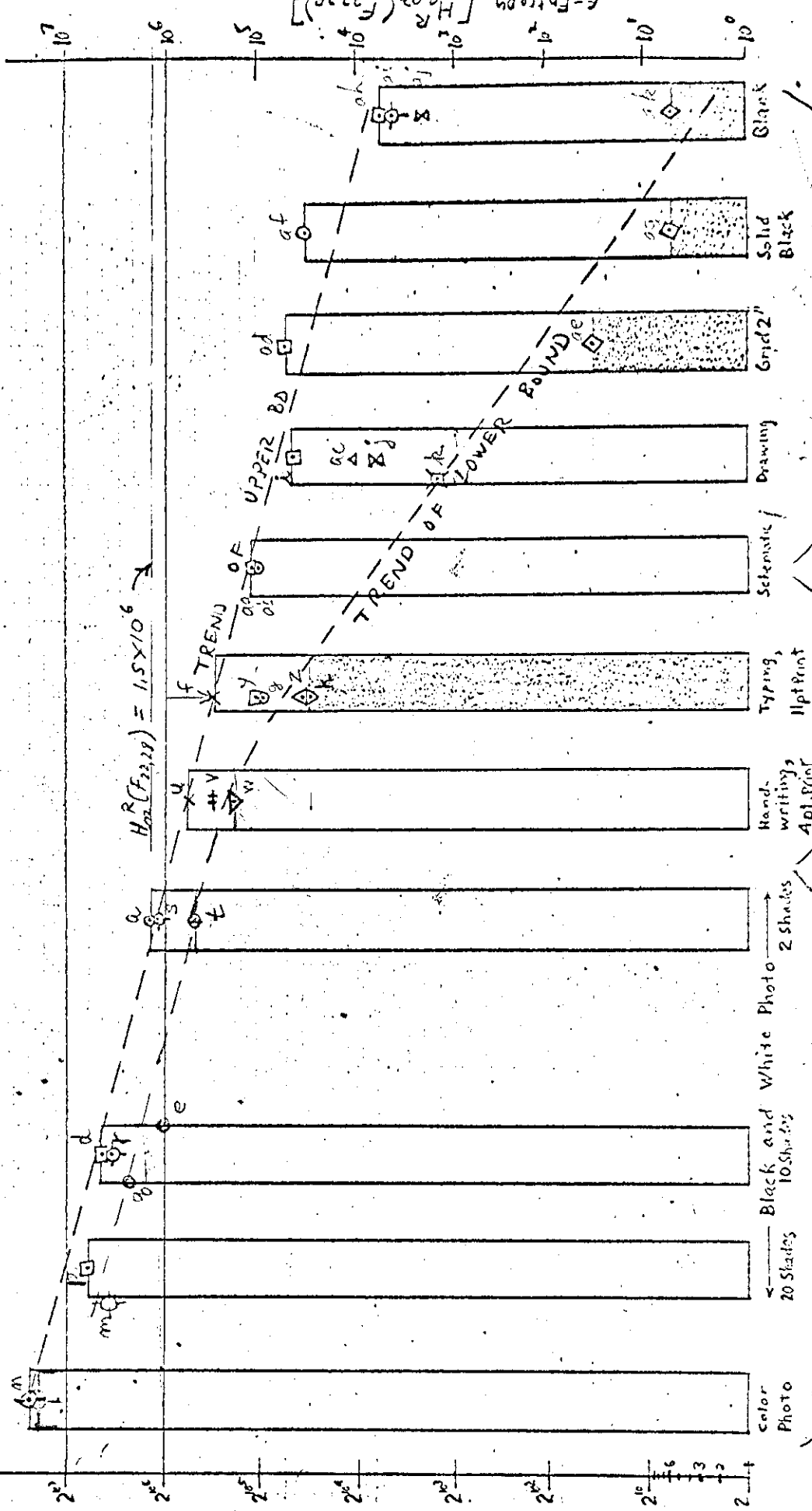
Figure 3 has the descriptive capacity, W , plotted on a logarithmic scale for the range $W = 2^1$ to $W = 2^{10^7}$ or for the information range $I = 1$ bit to $I = 10^7$ bits.

B C. Scanning Resolution

In these sample plots, a resolution of 0.02 cm or 50 lines/cm (125 lines/inch) is used. For an 8-1/2" x 11" page, $N = 1.506 \times 10^6$ cells or bits per page or (1080 x 1400). For an 8" by 10" page, $N = 1.292 \times 10^6$.

When reduced on a microfilm, half frame of 17 x 23 mm, it corresponds to:

$$\frac{1080}{17} = 64 \text{ lines/mm or } 640 \text{ lines/cm}$$



Drawings (c)

Typical or Printed Characters (h)

Photographs

Fig 3A1-1. f-Entropy (Entropy) vs. Description Capacity

(Note higher resolution such as 150 lines/mm and 200 lines/mm need also to be considered for microfilm.

In this example the resolution of 0.02 centimeters determines the ϵ in Vituskis analysis. The page size determines the subsection of the set F.

The potential range of limits to scan a page with ideal adaptive transmission system are tabulated in Fig 3A.2

8 1/2" x 11" Page: ϵ -entropy $H_{0.02}^R(F_{27,78}) = 1.5 \times 10$

Equivalent (1) Scan Speed v in/sec	Bit Rate (3) bits/sec	Ideal Bandwidth W (Hz)	Transmission Time per page					
			Facsimile 2-tone	Typing	Schematic	Grid	B	
8	1000	1	0.5	{ 25 min } { 9 min }	{ 5.6 min } { 37 sec }	{ 2.5 min } { 8.3 sec }	{ 37 sec } { 0.15 sec }	
24	3000	3		{ 8.3 min } { 3 min }	{ 1.9 min } { 12 sec }	{ 50 sec } { 2.8 sec }	{ 12 sec } { 0.050 sec }	
120	15,000	15		{ 1.66 min } { 36 sec }	{ 0.38 min } { 2.5 sec }	{ 10 sec } { 0.56 sec }	{ 2.4 sec } { 0.010 sec }	
382	48,000	48		{ 31 sec } { 11 sec }	{ 7 sec } { 0.775 sec }	{ 3.1 sec } { 0.173 sec }		
764	96,000	96		{ 15.5 sec } { 5.5 sec }	{ 3.5 sec } { 0.388 sec }			
8,000	10^6	1000	500	{ 1.5 sec } { 0.540 sec }				
32,000	4×10^6	4000	2000	{ 0.374 sec } { 0.135 sec }				

Co
pl
[2.
x
[6.
x

(1) $\epsilon = 0.02 \rightarrow 125$ bits/in
 $(8 \text{ in/sec})(125 \text{ bits/in}) = 1000$ bits/sec

For one sample per bit time the base frequency, $f_0 = B/2$.
 i.e. for $B = 1000$ bits/sec, $f_0 = 500$ cycles/sec.

- (2) Lower bound ^{easy} not realizable except possibly with computers at terminal.
- (3) Taken as half the noiseless binary channel capacity, i.e. w instead of $2w$, to allow for practical modulation system.

Fig. 3A.2 List of Speed Ranges for Ideal Adaptive Image Transmission System

C.D. Definition of ϵ -Entropy for Image Classification.

At present ^{we} ~~A~~ shall use Vituskin's analysis as simplified cases, and use McLachlan's description mechanics to obtain some upper bounds on the description capacity of certain types of images. For other images particular codes will be used to get upper and lower bounds in Fig 3A1

The relative ϵ -entropy concept applies to classes of mathematical spaces which give indications of being useful in image classification, but require further development to reach practical cases.

A Feinstein in a memo of Sept. 21, 1961, pointed out a special case of a theorem from A. G. Vituskin⁽¹⁰⁾. The English translation of this book is now available. Examination indicates that Vituskin's work may have more relevance to the image classification problem than previously thought.

Vituskin's concept of relative ϵ -entropy corresponds to the logarithm of McLachlan's descriptive capacity of a page⁽¹¹⁾. McLachlan carries his analysis over a range of physical systems, while Vituskin carried his analysis over a range of mathematical spaces. The

convergence of various series representations of functions permits the introduction of the concept of absolute ϵ -entropy of metrical spaces and the ϵ -capacity of metrical spaces.

Let F_{abc} be a physical space such as a plane defined by x-y coordinates, and let F_{ab} a sub-section defined by a specific dimension such as a 8-1/2" x 11" page, i. e. $a = 8-1/2$ " and $b = 11$ ". Let \emptyset be the third dimension defining 2^\emptyset levels, i. e. $\emptyset = 1$ for black and white;

$\emptyset = 5$ for 32 levels in a color system. Let $S_\epsilon^{\emptyset} (F_{ab})$ be the coordinate net ^{covering} ~~wherever~~ the space F_{ab} with resolution ϵ . Then let W_ϵ^{\emptyset} be the descriptive capacity ^{of the} ~~of the~~ elements of the minimum number of points in F_{ab} of the ϵ -net $S_\epsilon^{\emptyset} (F_{ab})$. The description capacity is the total number of different images that can be described by $S_\epsilon^{\emptyset} (F_{ab})$.

$$\text{Let } I_\epsilon^{\emptyset} (F_{ab}) = \log_2 W_\epsilon^{\emptyset} (F_{ab}). \quad (2)$$

$I_\epsilon^{\emptyset} (F_{ab})$ is the ϵ -entropy of the subset of set= F_{ab} with resolution ϵ . The smallest resolvable spot is a square of side ϵ . This is analogous to Vituskin:

$$H_c^{\emptyset} (F) = \log_2 N_\epsilon^{\emptyset} (F). \quad (3)$$

and $W_\epsilon^{\emptyset} (F_{ab})$ corresponds to MacLachlan's descriptive capacity, W .

^D Upper and Lower Bounds on E-Entropy Determined by Sample Cases.

~~2. A. Several Methods of Calculation Using Descriptive Capacity.~~

From McLachlan we have the concept of "descriptive capacity,"
W as the number of different images that can be described by a
given space with a given resolution. For example, a plane surface
divided up into N cells with M possible occupants could take on

$$W_G = NM \tag{4}$$

different patterns or descriptions. This can be derived by noting
the first occupant M_1 can be in any one of the N cells, likewise the
second, etc., through the M_m -th occupant, such that the possible
number of description is: N, for M_1 ; $N \cdot N$ for M_1 & M_2 ; $N \cdot N \cdot N$ for
 M_1, M_2 & M_3 , and N^m for M_1 ----- M_m .

If the N cells are divided into q groups, each of which has a
description capacity W_q , then the total description capacity is:

$$W_G = \prod_q W_q \tag{5}$$

For N cells with either binary "0" or "1" allowed,

$q = N, W_q = 2$, so

$$W_G = \prod_N 2 = 2^N \tag{6}$$

Thus for the binary case in rectangular coordinates the descriptive
capacity is

$$W = 2^N \tag{7}$$

and the E-Entropy is

$$I_{R, 0.02}^{(F22, 28)} = \log_2 W_{0.02}^{R(F22, 28)} = \log_2 2^N = N,$$

$$N = 1.5 \times 10^6 \text{ bits.}$$

This result could have been obtained directly without bothering with
"descriptive capacity" or "E-Entropy." The reason for introducing
these concepts is to lay the base for possible generalization and
specialization.

If we were operating in elliptic coordinates instead of rectangular coordinates, the physical area of a unit "square" in elliptic coordinates would change with position in the field. In such a case Vituskins's concept of E-entropy extends the to give the entropy for coordinate systems with such varying scales.

When we have restraints such as dependent probabilities or extensions to multilevels such as 10 or 20 levels of gray scale, or extension to color photographs, a good number of cases can be easily defined by use of McLachlan's "descriptive capacity."

For black and white photographs with 10 levels of sensitivity, Mr. Lachlan finds

$$W_G = \frac{N!}{\prod_{k=1}^S (M_k)!} \tag{8}$$

by using the simplifying assumption* that

$$M_k = M_j = \dots = N/S$$

This reduces the formula to:

$$W_G = \frac{N!}{(N/S)!^S} \tag{9}$$

This reduces by Stirling formula to:

$$W_G \approx \frac{S(N + S/2)}{(2\pi N)^{S/2}} \tag{10}$$

For N = 1.5 X 10⁶, S = 10

$$W_G = \frac{10(1.5 \times 10^6 + 5)}{(3\pi \times 10^6)^5} = \frac{10^{1.5 \times 10^6}}{1034.87} = 10^{(1.5 \times 10^6 + 5 - 3)}$$

10^{1.5 X 10⁶} = 24.

* This assumption doesn't make much difference in the first few significant figures in this case. We could have set W_q = 10, then by

$$W = \frac{N}{\prod W_q} = 10^{1.5 \times 10^6}$$

9

For color photographs an approximate value can be estimated using $S = 40$ in the above formulas, which give

$$W = 106 \times 10^6 = 22 \times 10^7$$

~~B. Graphical Representation of Bounds on Information Required to Represent Different Classes of Documents.~~

Now it is desirable to expand the document classification system of Table I in a way that will indicate the potential compression on a finer scale. To do this a bar graph has been constructed in Fig 3 in which a vertical bar represents the logarithm of the E-entropy or information required to describe each class of documents. The bars are placed in approximate order of decreasing E-entropy across the page. The class of photographs is divided into color; black & white, 20 shades, 10 shades, and 2 shades. The typed or printed character class is divided into small print and hard writing; typing and large print; and overlaps with the drawings class in regard to schematic drawings.

The drawing class overlaps with the typing class in regard to schematics and then covers line drawings; graph paper guides; solid block; and blank paper.

The ordinate scale of Fig. 1 is given two ways: on the left the units are description capacity on a log-log scale; on the right the ϵ -entropy or bits of information on a log scale. The upper bound points are calculated from the use of a coding system using only one scan across the information. Where the code ^(2, 3, 4) ~~of use~~ for a broader class of code, than the particular sub-class being examined. The lower bound points are calculated for the most restrictive conditions such as the use of a code designed for the particular sub-class and where a series of iterative steps are allowed in the encoding. This means that the lower bound for the typed characters and drawings correspond to the character recognition problem. Trial calculations have shown that adaptive systems using a prescan and a readout scan generally fall halfway between the upper and lower bounds on a log scale.

In this report, the general definition of ϵ -entropy is defined in more restrictive sense as:

$$H_{\epsilon}(F_{a,b}) = \log_2 W_{\epsilon}(F_{a,b})$$

analogous to the ϵ -entropy defined by Vituskin⁽¹⁾². W_{ϵ} is the description capacity or number of possible messages in the list when described by coordinate net \mathcal{O} with resolution ϵ over the space F with boundary a, b .

3A1

In the examples of Fig. A the right hand scale gives the ϵ -entropy $H_{0.02}^R(F_{22,28})$ which means a resolution of 0.02 centimeters or 50 lines per centimeter or 125 lines per inch. The superscript R means a rectangular coordinate system is used. The bound (22, 28) one 22 by 28 centimeter which is approximately 8-1/2 x 11". With no restraints on the image, the basic value of the ϵ -entropy for an 8-1/2 x 11" page is

$$H_{0.02}^R(F_{22,28}) = 1.5 \times 10^6 \text{ bits.}$$

This is drawn in on Fig. 1 for reference.

- 3.1 10. A. G. Vituskin, "Theory of the Transmission and Processing of Information", London: Pergamon Press (1961), Trans from Russian of "Estimates of the Complexity of Tabulation Functions".
- 3.2 11. Dan McLachlan, Jr., "Description Mechanics" Information and Control, 1, 240 - 266 (1958).
- 3.3 12. Prudhom and Faber,
IBM Report 17-037, April 1961

P.F. PARAMETERS POTENTIALLY SUITABLE FOR AUTOMATED
DOCUMENT CLASSIFICATION (0 to 100%)

To determine what operations ^{the prescanning logic} ~~the analyser (An) in Fig. 3~~ must

do, we list the properties of the signals that might be considered, before testing a specific system.

1. Logical Check for Presence or Absence of Signal

The analyser must be able to prescan for blank rows and or columns. This could be overlapped with other logic in the system.

2. Fourier Spectrum of Signal

An approximation to the Fourier spectrum of a signal could be calculated by a Fourier series approximation implemented with a set of delay lines and adders.

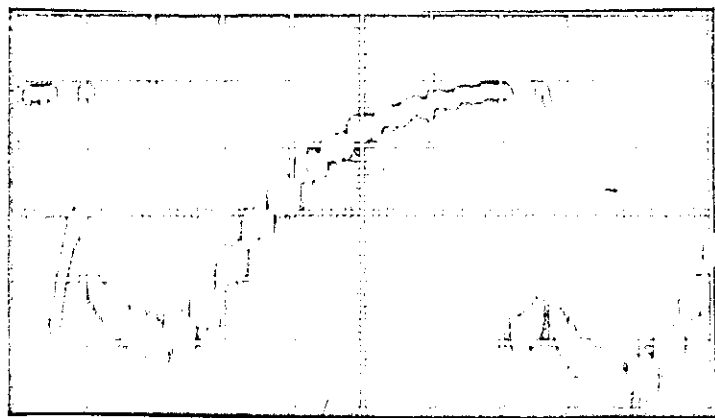
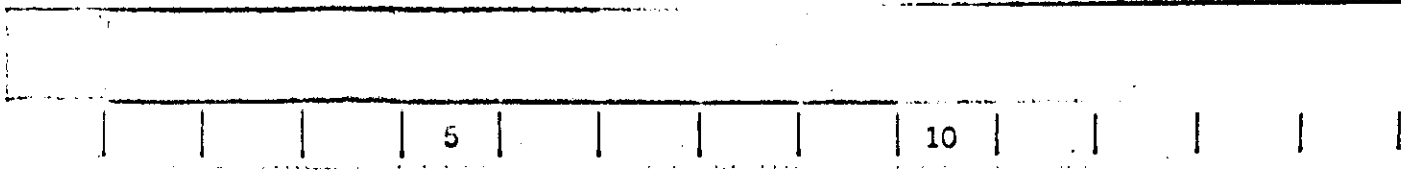
The direct computation of the Fourier spectrum has some limitations here. First the electrical signal from a typical phototube is proportional to the intensity or the square of the basic function of black and white. For example, in Fig. 3 a scan of the range of gray scale steps is shown to be square law. Thus the Fourier Series or Fourier spectrum analysis would have to be preceded by a square-root operation.

$$x(t) = \sqrt{1/2} (t).$$

If $y(t)$ is limited to values of 1 and 0 as in 2 level (black and white) then since:

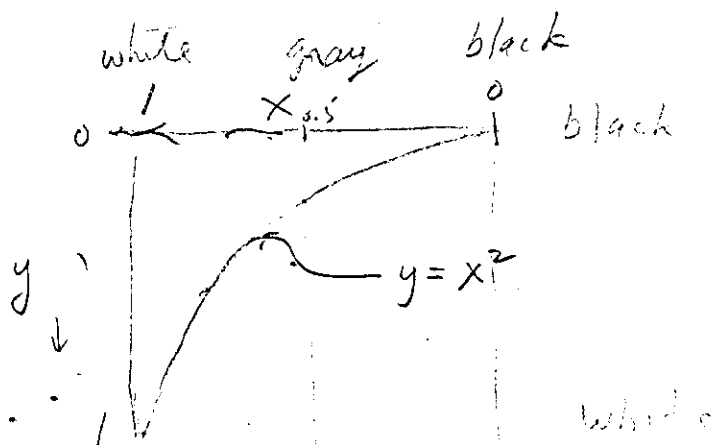
WHITE

00



- Calibration of
output of
analog device
using IRE
Std gray

00



x = level of gray, linear, 0 = black, 1 = white
 y = current from photo tube, proportional to intensity of illumination (x^2).

Fig 4 - Gray scale calibration of analog device
3E.11

00

$$0^{1/2} = 0, \quad \text{and } 1^{1/2} = 1,$$

$$x(f) = y(f)$$

and the Fourier analysis can be used without using the square-root operation.

3. Elementary Statistical Properties

Elementary statistical properties of the signal such as the mean (\bar{x}), the mean square ($\overline{x^2}$) and the variance (σ^2) could differentiate between some classes of documents.

$$\bar{x} = \sum_k p_k x_k = \int_{-\infty}^{+\infty} x p(x) dx$$

$$\overline{x^2} = \sum_k p_k x_k^2 = \int_{-\infty}^{+\infty} x^2 p(x) dx$$

$$\sigma^2 = \sum_k p_k (x_k - \bar{x})^2 = \int_{-\infty}^{+\infty} (x - \bar{x})^2 p(x) dx$$

When these parameters are all dropped to zero an absence of signal is indicated, eliminating the need for a separate logical check of paragraph a, unless other factors require separate logic.

4. Higher Order Statistical Properties

Differences between random telegraph signals and the signals from scanning typing can be more easily differentiated by the auto correlation function and related statistical parameters.

Auto correlation function:

$$R_x(\tau) = E[x(t)x(t+\tau)] = \int x(t)x(t+\tau) dt$$

co variance:

$$\text{Cov}[x_i, x_j] = \sum_i [p(x_i) x_i] x_j, \text{ where } j = i+b$$

correlation coefficient:

$$\rho(x_i, x_j) = \frac{\text{Cov}[x_i, x_j] - \bar{x}^2}{\sigma^2}$$

power spectral density:

$$S_x(f) = \int_{-\infty}^{+\infty} R_x(t) e^{-j\omega t} dt$$

This could also be derived from the Fourier transform $G(f)$. A possible direction to explore is the use of an approximation to the auto-correlation function to determine type of document, and then take the transform to get an approximation to the power spectral density for indicating the best scanning speed.

5. Alternative Discrete Statistical Properties

A set of logical circuits which compile an approximate distribution of white and black run lengths could compare the pattern with the pattern upon which a shannon-fano or huffman code was based, could determine which code of a set came closest to the observed statistics. There is also the possibility of accumulating statistics with memistor-like elements which can generate new codes.

A system for measuring the quality of an image has been developed by Roy A. Jensen*. This system averages a series of signals like that shown in Fig. 4A from scanning a line of a document. The average value of seven parameters shown in Fig. 4B can be scaled off of the graphical plot of the output.

The correspondence between these analog properties and the properties discussed so far are as follows: The correspondence between resolutions of scanning ϵ and image quality parameters depends upon whether one wishes to recognize the characters or accurately reproduce the images. If the object is to accurately reproduce the image,

* Roy A. Jensen, "An Analog System for the Measurement of Image Quality."

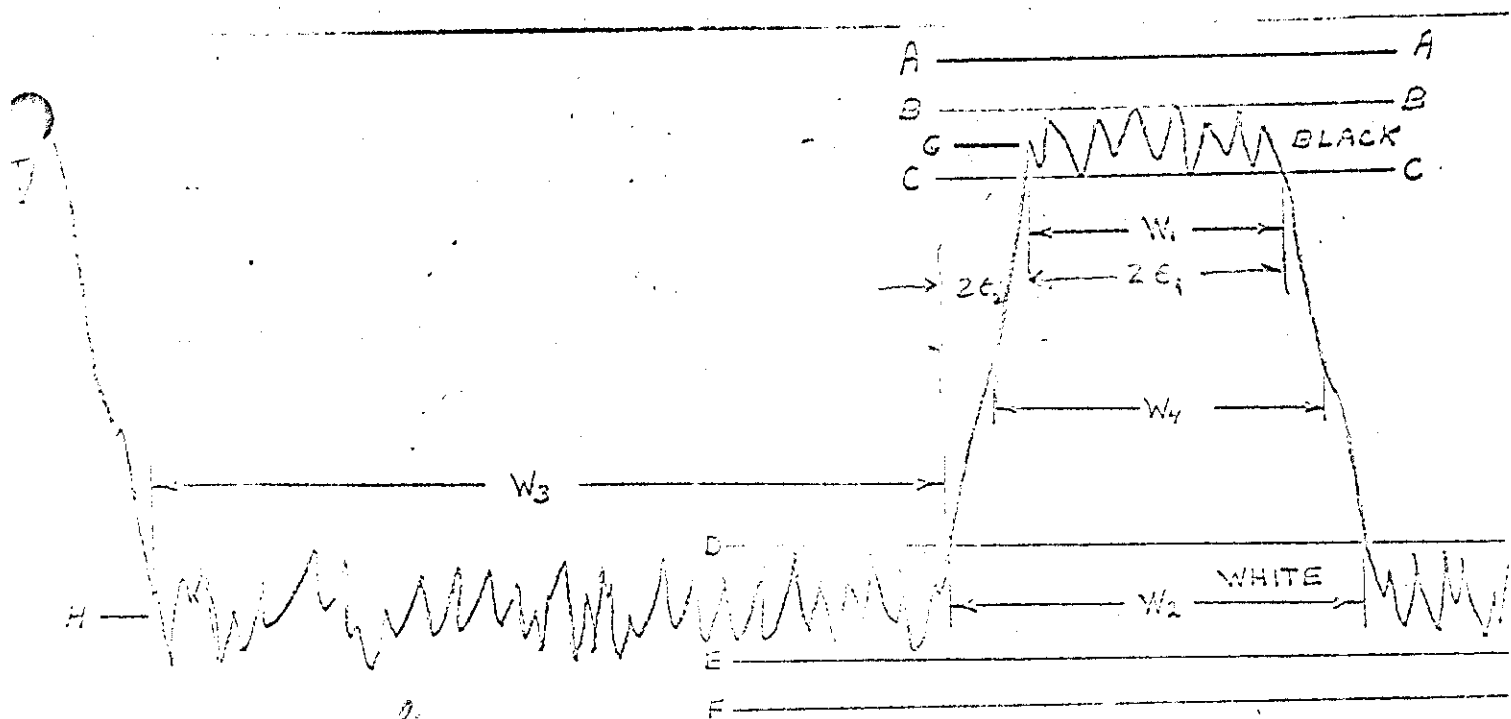
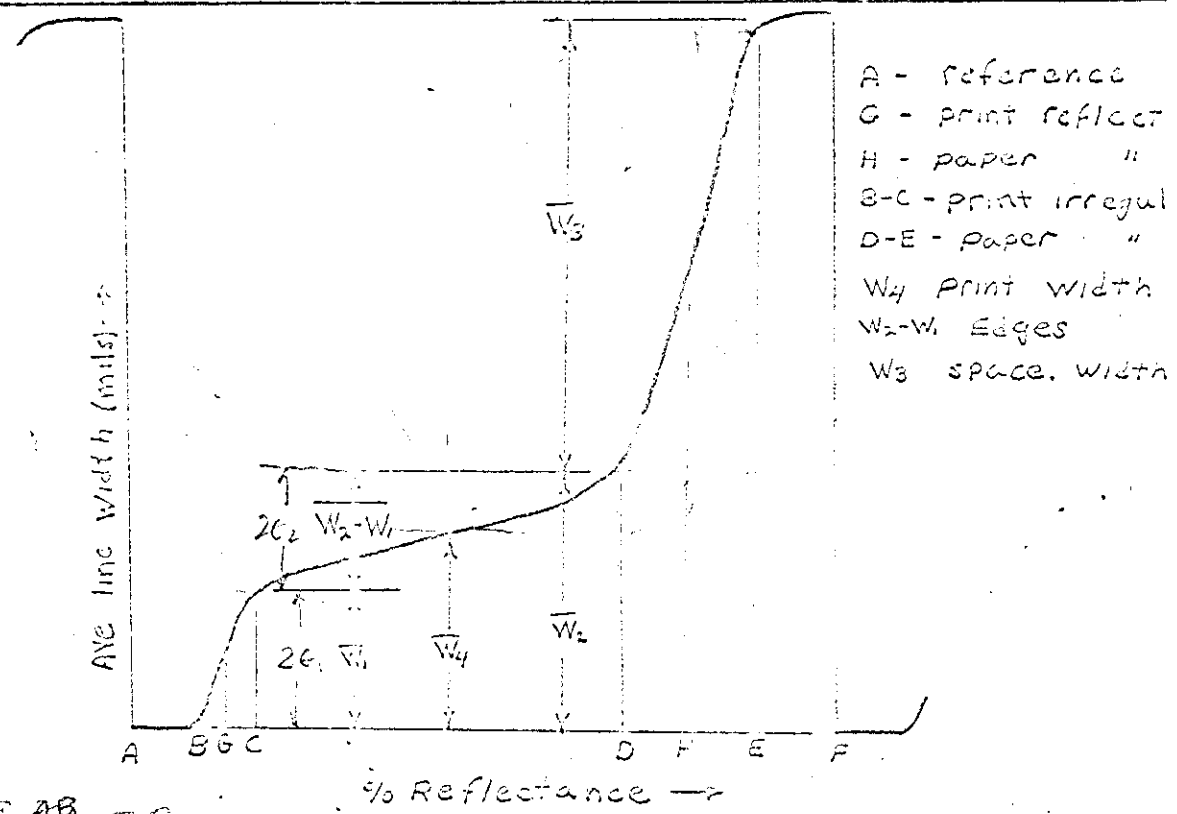


FIGURE 4A 3E.2



- A - reference
- G - print reflect
- H - paper "
- B-C - print irregul
- D-E - paper "
- W4 Print width
- W2-W1 Edges
- W3 space. width

FIGURE 4B 3E.3

$$\epsilon = \epsilon_2 = (w_2 - w_1)/2$$

and if it is to reproduce recognizable characters with some blurring tolerated the requirements can be relaxed so

$$\epsilon = \epsilon_1 = w_1/2 .$$

The signal to noise ratio is approximately $S/N = H-G/D-E$ for the paper irregularity, or $S/N = H-G/B-C$ for the print irregularity.

6. Range of Change of Derivatives

When the image pattern has high redundancy such that the image signal can be approximated by segments of sine waves for 2-level black and white (or deviation and rate of change for multi-level), the a-c component of the signal approaches:

$$x(t) = a_k \cos \left\{ \frac{2\pi k}{t_0} t + \phi_k \right\}$$

A parameter θ can be defined

$$\begin{aligned} \theta(x(t)) &= x^2(t) + \left(\frac{t_0}{2\pi k} \right)^2 \left(\frac{dx(t)}{dt} \right)^2 + \sum (t) = \\ &= a_k^2 + \sum (t) \end{aligned}$$

For $f_0 = k/t_0$, $\sum (t) = 0$, so $\theta(x(t)) = a_k^2$,

then $d\theta/dt = 0$ for $f = f_0$.

otherwise $\frac{d\theta}{dt} = \frac{d\sum}{dt}$.

An analysis circuit which determines $\theta(x(t))$ is a constant would define the frequency of the sine wave. An additional unit would

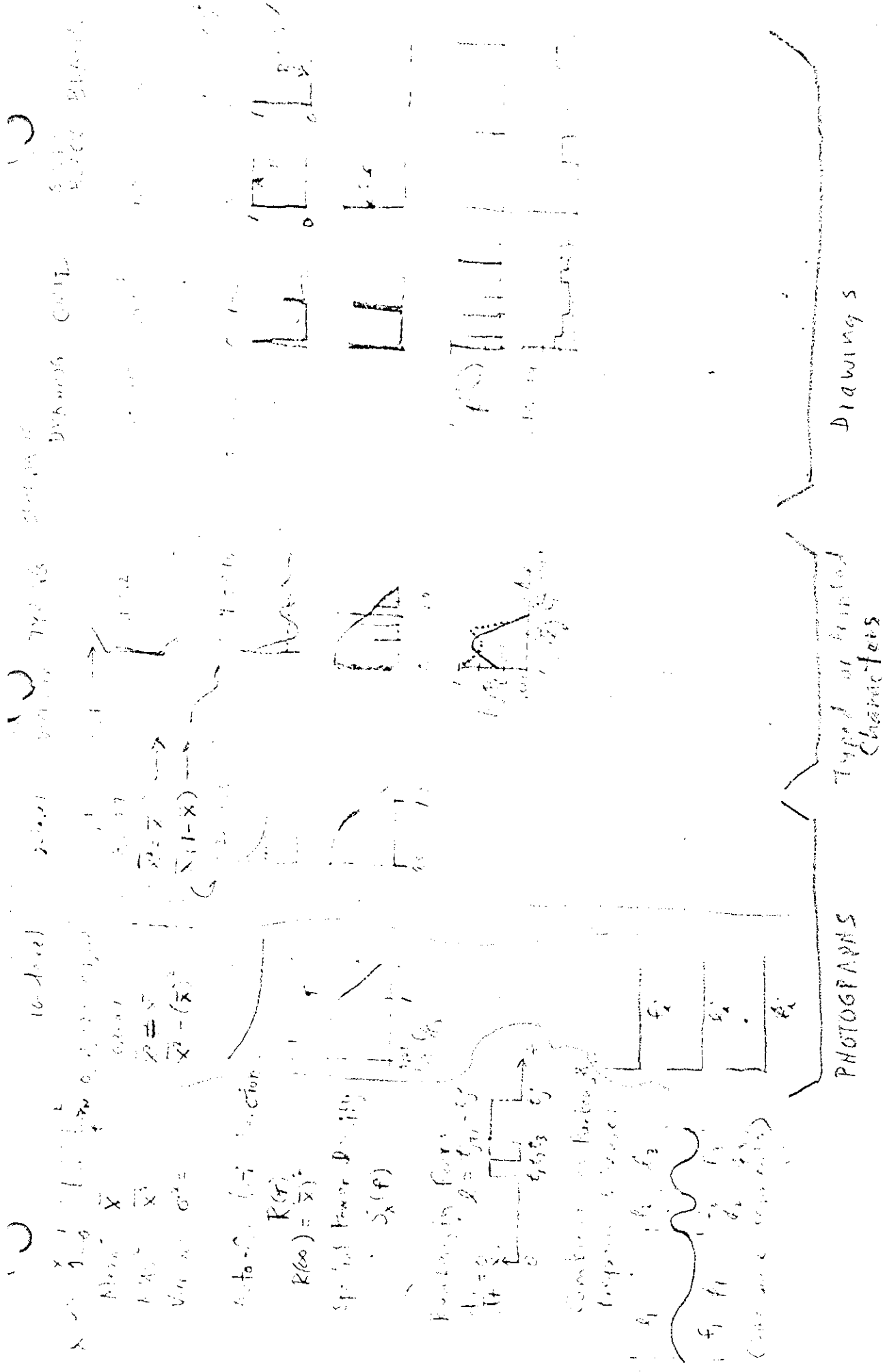


Fig. 3.18

have to be added to determine the phase. This outlook may lead to a special case of a spectrum analysis where the gain of an amplifier following a differentiator is swept instead of the frequency, so that the output of the differentiator-amplifier and the original signal after each are squared separately always equals the constant a_{\downarrow} after addition.

7. Preliminary Discussion of Decision Properties

The information in bits per page for different types of documents is plotted in Fig. ^{3B} 4. A resolution of $\epsilon = 0.02$ centimeters is assumed. For black and white, ~~two~~ i. e. two levels four different codes are proposed. Code A for Blank, Solid Black and Graph Paper Guides. Code B for Grids Drawings, and Circuit Schematics, Code C for typing and Printing, Code D for Handwriting and Special Images with Fine Detail.

The slope of the lines in Fig. 4 indicates that within a given code variation in print size could be accommodated by changing to ~~the~~ scanning speed.

An additional code E is shown to indicate future extension to include 10 shades of gray to include photographs.

The present objectives define more specifically how an adaptive system could distinguish between the different types of documents. Now we can examine these different properties to see how significant they are.

The mean value (\bar{x}) in Fig. ~~4~~ ^{3B} is excellent for distinguishing between blank and solid black. However the range of the mean value is not distinct enough at the transitions between the different groups of types. For two levels (black and white) the R. M. S. value and the variance, being equal to or simply related to the mean, do not provide any better criterion.

The auto-correlation function does provide a clear cut distinction, but is more complicated to implement. For the column in which we have sample values of the auto-correlation function there is a clear distinction for each class of documents.

Sample calculation of some of the functions in Fig. 3E4 are illustrated in Figs. -3E5 through 3E12.

The outline of a magnified image of the sample type of Fig. 3E5a is shown at 11 times full size in Fig. 3E5b. A sample line scan AA' is marked across the characters in Fig. 3E5b. The digital representation of the scan line is marked below the character.

A sample waveform for a section of the scan is shown in Fig. 3E6. The values of the mean (\bar{x}), mean square ($\overline{x^2}$), square of the mean (\bar{x}^2), and the variance (σ_x^2) are all marked on the waveform. These values are representative values from a small sample only.

For comparison the waveform and corresponding statistical properties of a random telegraph signal are shown in Fig. 3E7. While the mean value (\bar{x}) for the random telegraph signal is 0.5, it runs as low as 0.16 for typed text. ✓

A sample autocorrelation function of the random telegraph signal of Fig. 3E7 is plotted in Fig. 3E9 for $a = 0.425$. The spectral power density for the same is plotted in Fig. 3E10. A family spectral power densities as functions of a and p are plotted in Fig. 3E11 for the generalized random telegraph signal. The formulas for these parameters are derived in Appendix C: Spectra of Some Pulse Signals.

A sample autocorrelation function for the type scan of Fig. 3E5bd is shown in Fig. 3E12. Of the function were computed over a larger sample of the irregularities would smooth out except for inherent steps due to the average line width, average character width, and average word length.

3.2

the radii

3/4" scan

Fig. 7 - Sample Type
3ESA

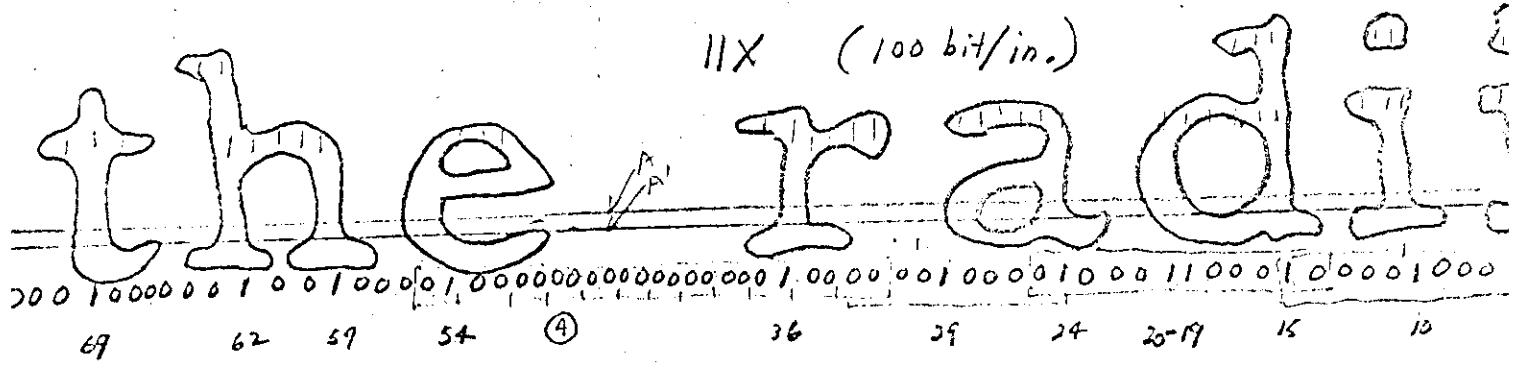


Fig 8 - Sample Scan Line
3FB



Fig 9. Black & White Facsimile Signal
3ET

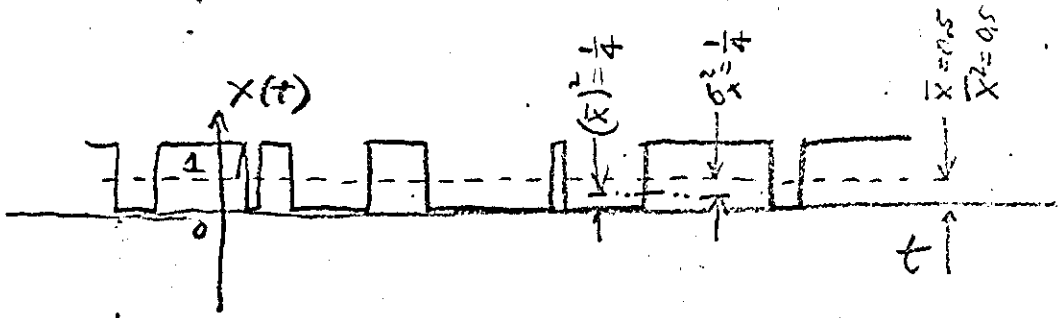
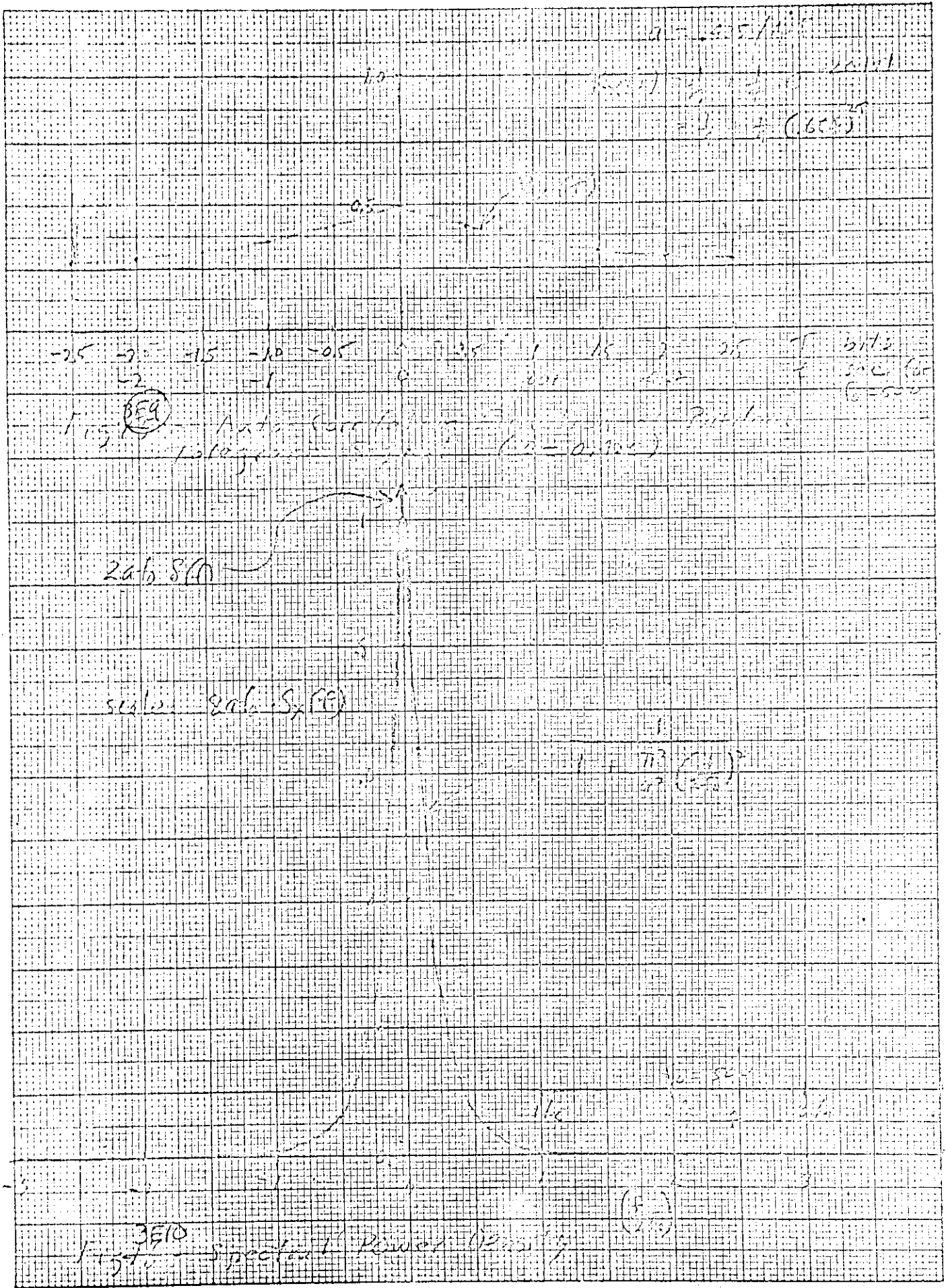


Fig 10. Random Telegraph Signal
3EF

3.23 / 3.24

CHARLES BRUNING COMPANY, INC.
MADE IN U.S.A.

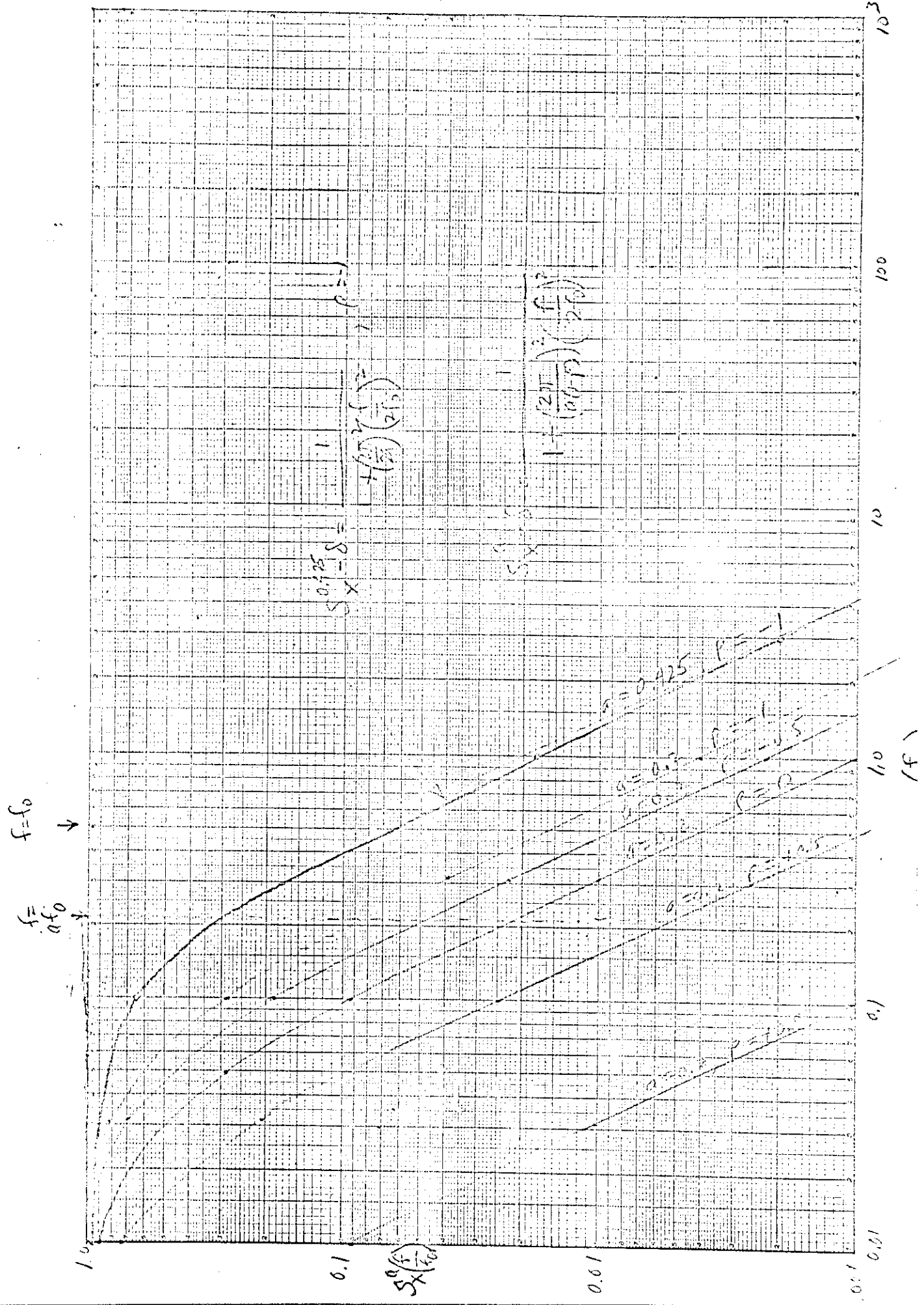
BRUNING 700-26
10 X 10 TO 1/2 INCH



3F11

3.25

12-



$f = f_0$
 $f = a f_0$

$a = 0.125$
 $a = 0.25$
 $a = 0.5$
 $a = 1$
 $a = 2$

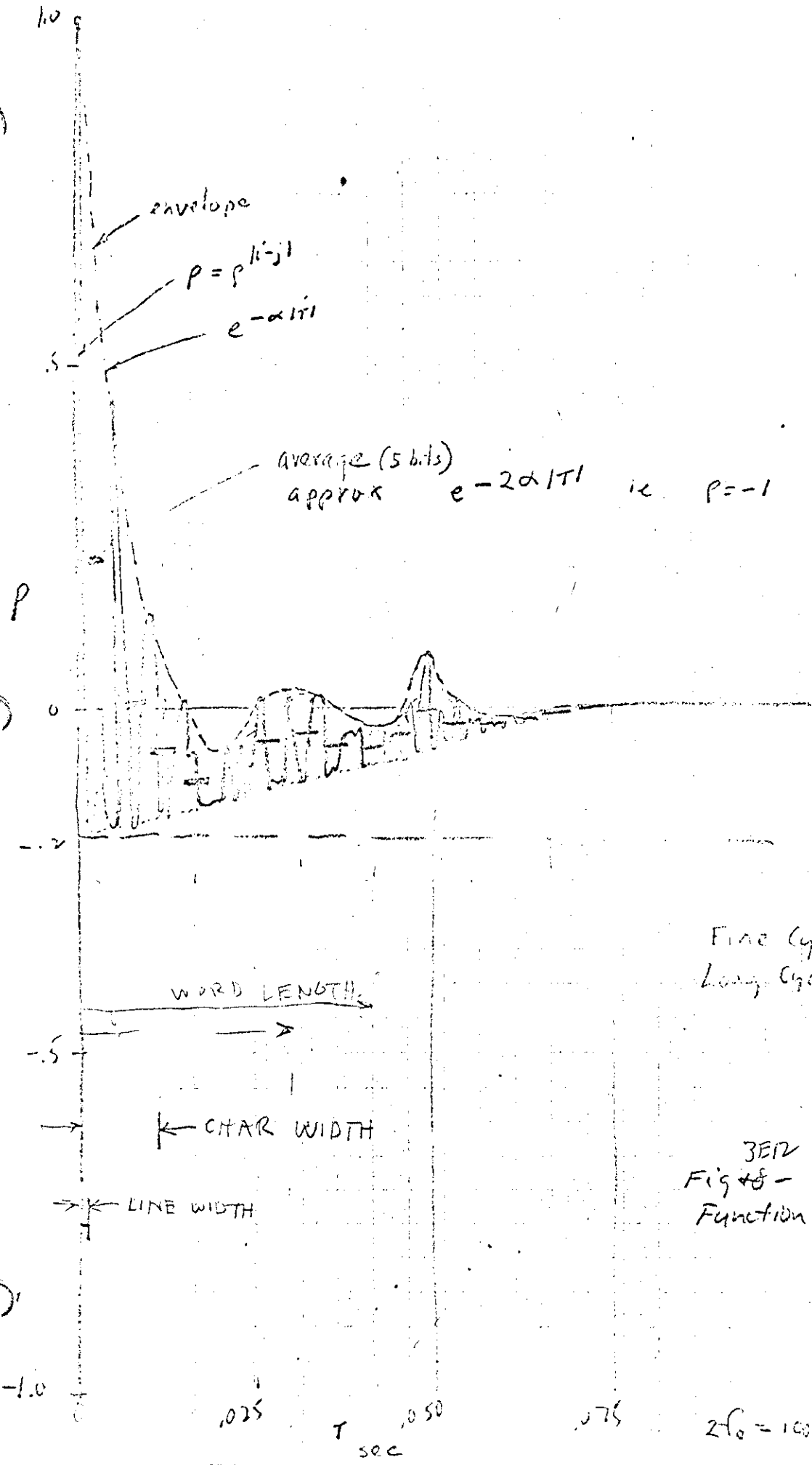
$$S_x = 0.125$$

$$f = \left(\frac{f}{f_0}\right)^2 \left(\frac{f_0}{a f_0}\right)$$

$$S_x = 0.25$$

$$f = \left(\frac{f}{f_0}\right)^2 \left(\frac{f_0}{a f_0}\right)$$

$$S_x = 0.5$$



Fine Cycle 5 bits
 Long Cycle 25 bits

3E12
 Fig 48 - Autocorrelation
 Function of Type Scan

$2f_0 = 1000 \text{ bits/sec.}$

~~V. POTENTIAL APPLICATION OF IMAGE CLASSIFICATION
TO ADAPTIVE SCANNING OF IMAGES~~

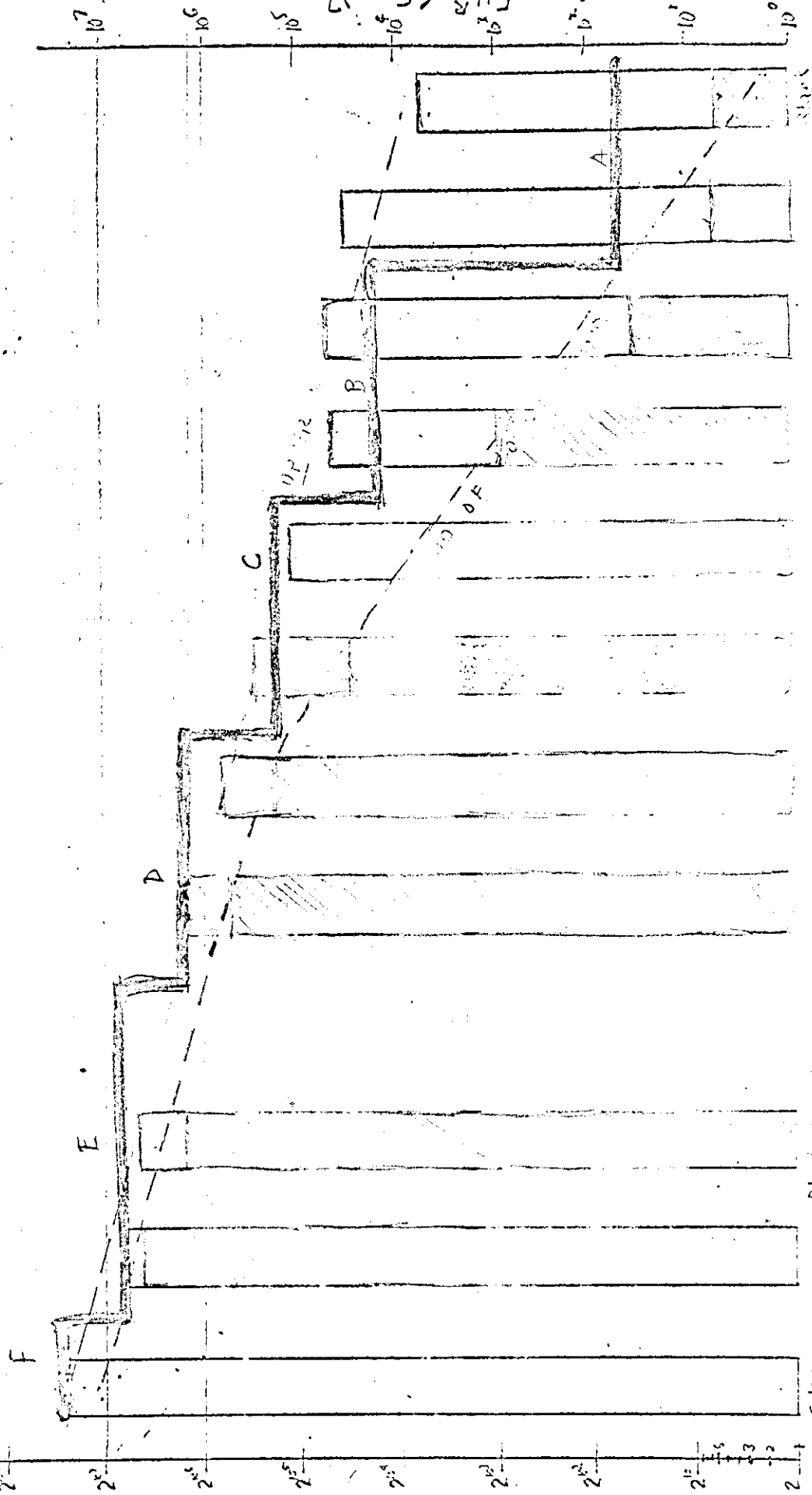
X. F. General Form of an Adaptive Scanning System

A possible adaptive system using different codes is illustrated in Figs. ^{3F1} 4 and ^{3F2} 3, Fig. ^{3F1} 2 is a reproduction of the general limits of Fig. ^{3A1} 4 with a stepwise curve ABCDEF added to show how the use of different codes could approach the lower bound of ϵ -entropy. Fig. ^{3F2} 2 is a block diagram of such an adaptive image scanning system illustrating the principal features required in such a system.

The step function curve with flat sections A, B, C, D, E, and F in Fig. ^{3F1} 2 illustrated how an adaptive system might handle the range of documents with different coding systems. An adaptive system has the advantage of stepwise coming closer to the lower bound of the entropy of the document. A fixed coding system designed for the statistics of a certain resolution or ϵ becomes inefficient when batches of documents departing from the average statistics enter the system.

The adaptive might only use a few of the different codes if some of the types of documents occurred only rarely. For example, if code D is required 10% of the time, code C 60% and Code B 29%, and Code A 1% the ultimate compression ratio would be

$$R = \frac{1}{0.10 \times 1 + 0.60 \times 0.1 + 0.29 \times .01 + .01 \times .0003}$$



Description (Capacity [W])

Color Photo
20 Slides 105.4.14

151 is right before system of all in 6th S. 10000
to lower (B) 10000

$$= \frac{1}{0.1 + .06 + .0029 + .000003} = \frac{1}{.162903} = 6.14$$

If only D and C were used:

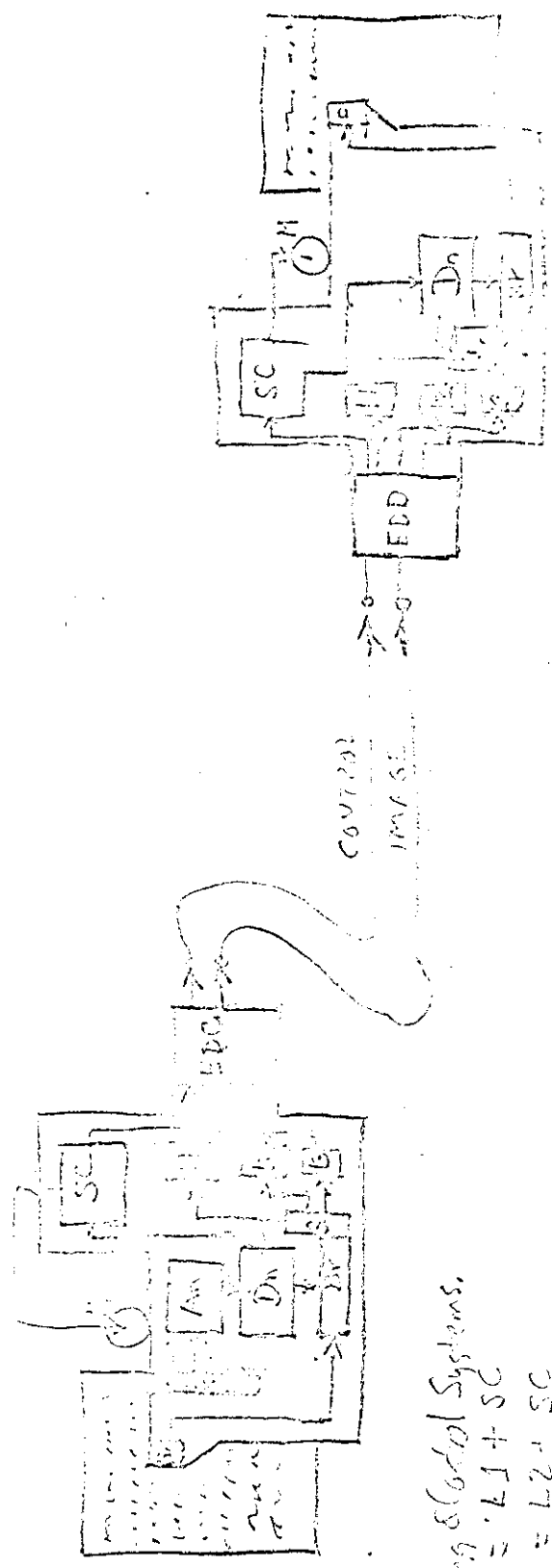
$$R = \frac{1}{0.10 \times 1 + .90 \times 0.1} = \frac{1}{0.19} = 5.27$$

In this case it might be more economical to settle for logic set C and D with compresses 5.3 instead of adding logic sets B and A to increase the compressor to 6.1.

The general form of an adaptive image transmission system is shown in Fig. 3. The various elements such as the scanner, pre-scanner analyser, decision logic, buffer, and speed and control logic are identified in Fig. 3.

A variation of the classification system of Fig. 2 is shown in Fig. 4 where a variable speed scan is proposed for bringing the adaptive scanning ϵ -entropy curve closer to the lower bound.

The variable speed scan can reduce further the information per page that needs to be transmitted, however it must add some speed information, so that an analysis must be made of the document statistics to determine if the additional saving is worthwhile. Additional feature of the adaptive speed system is that it may reduce the amount of logic by permitting Code C to be used in part of the B and D region through speed scaling. In this case the adaptive speed feature would not make added savings in time-bandwidth, but would permit less logic and buffer storage.



Coding & Control Systems.

- D = L1 + SC
- C = L2 + SC
- B = L3 + SC
- A = SC

- SC = Scanner
- PS = Pre-Scanner
- SM = Scanning Motor
- ED = Error Detector
- P = Printer
- AN = Analysor
- DK = Decision Logic
- BU = Buffer
- SC = Switch
- LI, L2, L3 = Logical Devices for Coding & Decoding of Image Information.
- EDC = Error Correction Control
- EDD = Error Detection Detector

Fig 3F2. Adaptive Control System

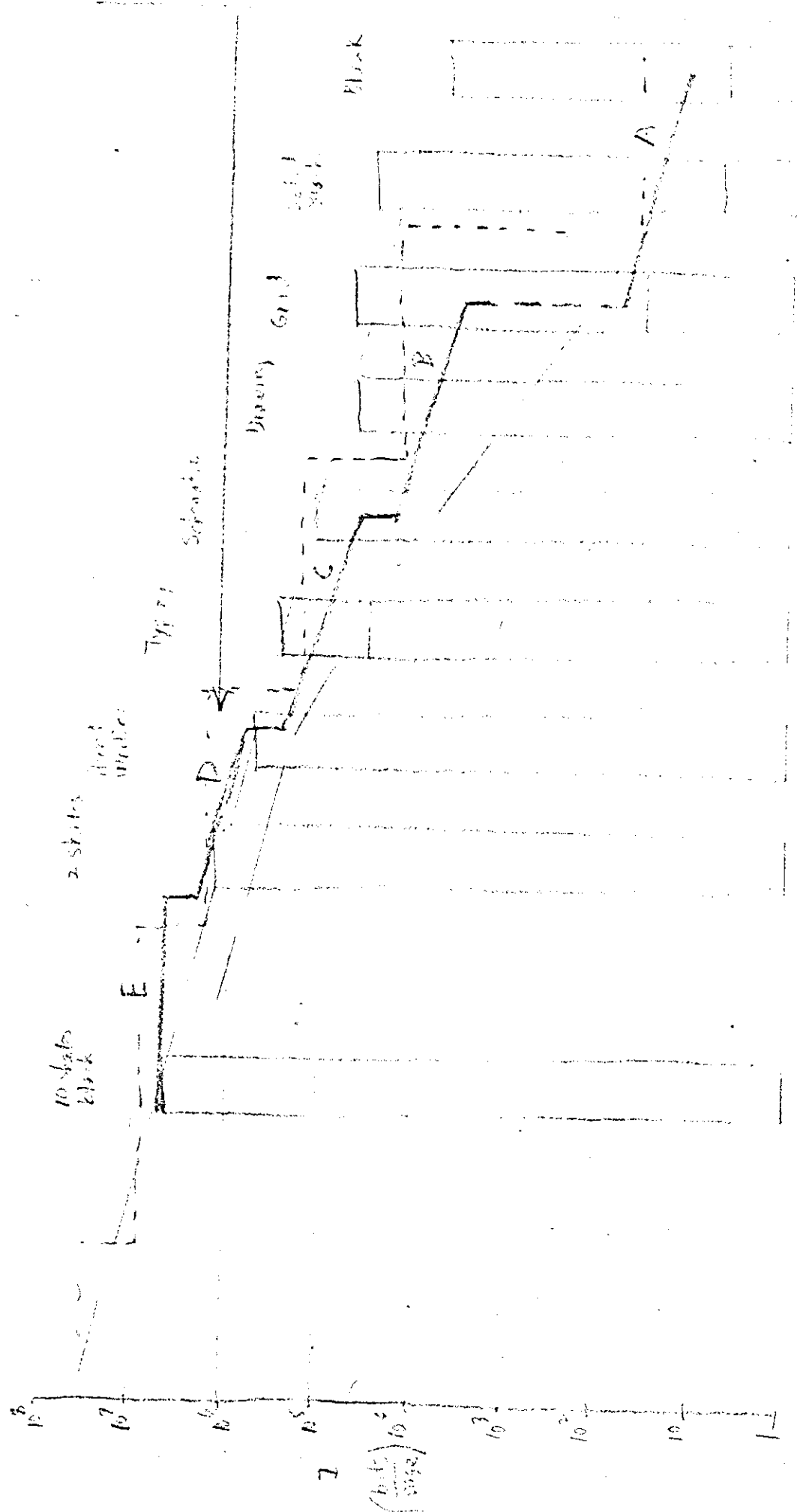


Fig 3F3 A Laplacian Scanning System Using Variable Speed

G. PARAMETERS AND TECHNIQUES FOR TWO-DIMENSIONAL IMAGE PROCESSING

1. INTRODUCTION

One possible method of implementing an image information compression system is through optical signal processing (e.g. optical prediction - subtraction schemes to be discussed later). It seems reasonable to process the signal in its original form, i.e., a two dimensional image - before converting to an electrical medium. As we shall see, many of the difficulties encountered in electrical signal processing may be avoided in the spatial domain.

This section is intended as a brief review of the application of linear system theory to optical or spatial systems analysis and synthesis. Although no specific references are made in the text, much of the material presented herein can be found in original form in the papers listed in the bibliography.

2. GENERAL MODEL

A general model for discussion is shown in Figure 3G1. We can define a point field function $s(x, y)$ on the object space, This is the

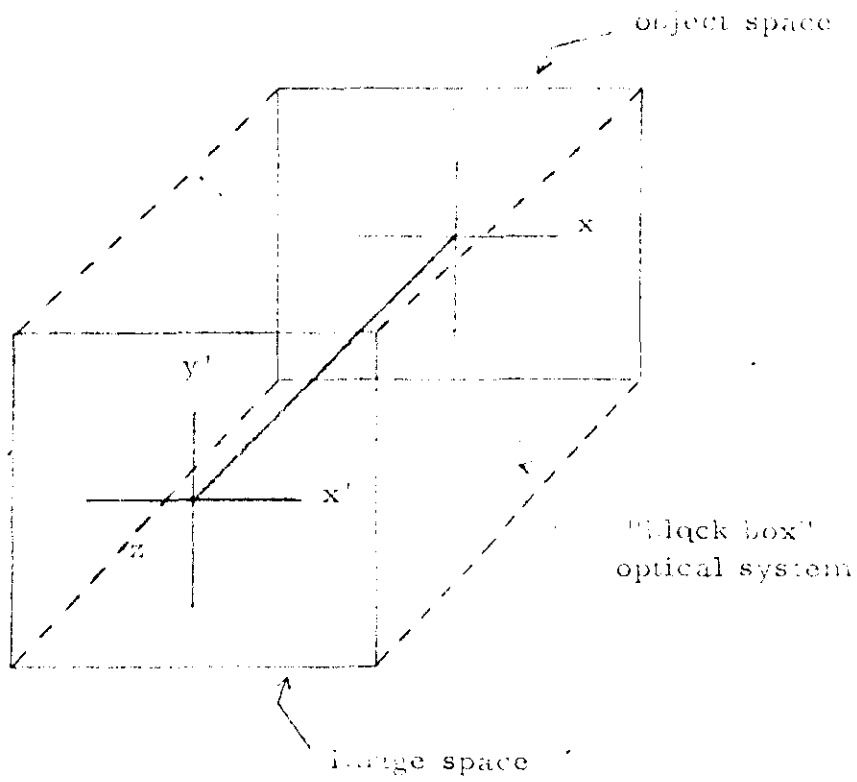


Fig. 34.1 Model of Image Space and Object Space.

stimulus to the "black box" system. The response to this stimulus is given by $r(x', y')$ on the image space. This is also a two-dimensional point function; i. e., it is a surface describing the image which is a result of processing the object by a "black box".

We now assume that the optical system is linear and exhibits spatial invariance, i. e., assume that

- a) if $s_1(x, y)$ produces $r_1(x', y')$ and $s_2(x, y)$ produces $r_2(x', y')$, then $s_1 + s_2$ produces $r_1 + r_2$, and
- b) if $s(x, y)$ produces $r(x', y')$, then $s(x - \xi, y - \eta)$ produces $r(x' - \xi, y' - \eta)$.

This assumption, if violated over the large space, may usually be salvaged by dividing the space into smaller subspaces where the assumption does hold. With the assumption of linearity and spatial invariance we can always define a 2 dimensional weighting function $w(x, y)$ that relates the stimulus and response by a cross-correlation operation

$$r(x', y') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(x - x', y - y') s(x, y) dx dy \quad (1)$$

Now, if the stimulus is impulsive, i. e. if

$$s(x, y) = w \delta(x, y) = \delta(x) \delta(y) \quad (2)$$

(where the generalized-function concept of the impulse is extended to two dimensions), then, due to the sifting property of the impulse, the response is

$$r(x', y') = w(-x', -y') \quad (3)$$

This impulse or "point" response is given a special designation, h :

$$w(-x', -y') = h(x', y') \quad (4)$$

and we see, analogous to temporal filter systems, the impulse response is the spatial inverse of the weighting function. The impulse response and arbitrary stimulus are thus related to the arbitrary response by convolution,

$$\begin{aligned}
r'(x', y') &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x' - x, y' - y) s(x, y) dx dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) s(x' - x, y' - y) dx dy. \quad (5)
\end{aligned}$$

Before going further, a word should be injected regarding the physical meaning of the previously defined functions. The meaning depends on whether we are considering the object as a coherent or incoherent light source.

(1) Coherent source - e.g., transmitted illumination of the object as in Figure ^{3G2} 2. Here, the stimulus function $s(x, y)$ is an amplitude function $s_a(x, y)$ because amplitudes of coherent light beams add.

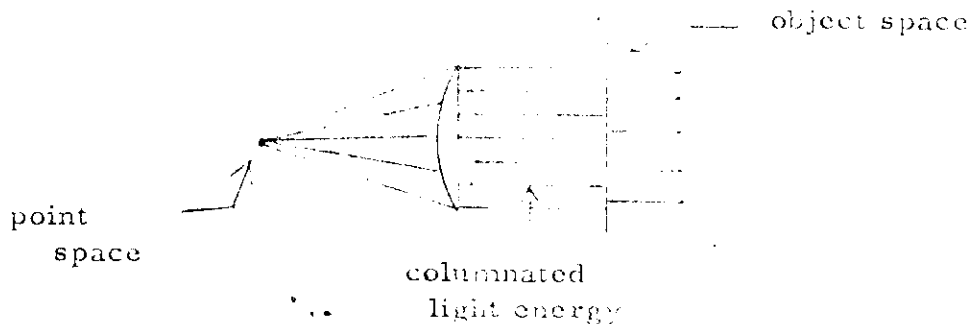


FIGURE ~~2-5H~~ (3G2) Coherent Source

Then the response function $r(x, y)$ must also be an amplitude function $r_a(x, y)$ related to the stimulus by correlation with the amplitude weighting function. The amplitude weighting function $w_a(x, y)$ could be a diffraction

pattern of the lens system as will be discussed in a later section of this paper. The functions s_a and r_a will in general be complex, since both amplitude and phase are important in the determination of the resultant amplitude (e. g., the functions r and s may describe the electric field component of a plane-polarized light wave, i. e.

$$r_a(x, y, t) = \vec{E}(x, y, t) = i \vec{E}_{ox}(x, y) e^{j(\omega t - \alpha(x, y))} + j \vec{E}_{oy}(x, y) e^{j(\omega t - \beta(x, y))}.$$

Note that since the linear relationship in coherent source systems exists for amplitudes, the intensity functions $|s_a|^2$, $|w_a|^2$, $|r_a|^2$ are in general related in a non-linear fashion.

(2) Incoherent Sources - e. g., reflected illumination of the object as in Figure 3. In incoherent source systems, the intensities of light beams

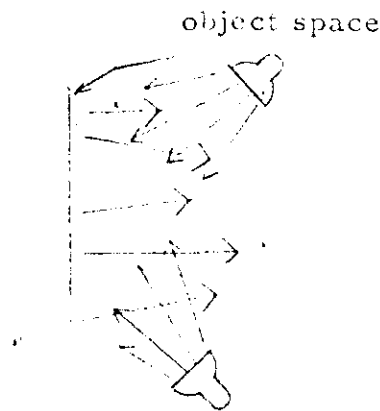


FIGURE 3 36.3

add, so that the response, stimulus, and weighting functions are intensity distributions r_i , s_i , w_i related by the correlation,

$$r_i(x', y') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w_i(x - x', y - y') s(x, y) dx dy \quad (6)$$

Having made the above distinctions concerning the meaning of r , w , and s we will normally drop the subscripts a and i since the meaning will follow directly from a statement of which type of source is being used.

3. Fourier Transform Analysis

Since stimulus and response are related by the convolution operator

$$h(x, y) \otimes s(x, y) = r(x, y) \quad (7)$$

and they are functions of space variables, it appears that considerable physical insight and mathematical simplification would result from an application of Fourier analysis. Thus, we define,

$$Q(v_x, v_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} q(x, y) e^{-j2\pi(v_x x + v_y y)} dx dy \quad (8)$$

and

$$q(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Q(v_x, v_y) e^{j2\pi(v_x x + v_y y)} dv_x dv_y \quad (9)$$

as the two-dimensional Fourier transform pair.

We may then compute the transform equivalent of (1) and (5) as,

$$R(v_x, v_y) = W^*(v_x, v_y) S(v_x, v_y) \quad (10)$$

and

$$R(v_x, v_y) = H(v_x, v_y) S(v_x, v_y) \quad (11)$$

the capital letters denote Fourier transforms of the lower case letters. Equations (11) or (12) provide a basis for synthesis of a desired response from a given stimulus - i. e., concepts such as spatial filtering, equalization, detection, etc. can be utilized. $H(v_x, v_y)$ is termed the transfer function of the optical system - it relates stimulus and response by an algebraic equation.

4 Impulse Response Analysis: Coherent Source Systems

Obviously, the determination of the impulse response function is of prime import. This can usually be done by either of two methods:

- (1) by analysis
- (2) by experiment - measurement of the (approximate) impulse response.

The first method is useful only in the case of simple optical systems. The second is very useful in more complicated systems where analysis becomes cumbersome, if not impossible. In this section we will consider the first method only.

As an example, consider the simple one-dimensional optical system shown in Figure ³⁶⁴ 4 (the distribution is of infinite extent in the y-direction). Consider the distribution $g_a(x)$ to be resulting from a coherent point source followed by an aperture mask of varying transmittance. Thus $g_a(x)$ represents a complex wavefront distribution;

$$g_a(x) = a(x) e^{j(\omega t - \phi(x))} \quad \omega = \frac{2\pi c}{\lambda} \quad (12)$$

where $a(x)$ is the magnitude and $\phi(x)$ is the phase of the light beams. To compute the amplitude distribution at P, we consider the effect of differential elements dx at x . The wave in the direction r from dx will be

$$g_a(x) e^{-j 2\pi r/\lambda} dx. \quad (13)$$

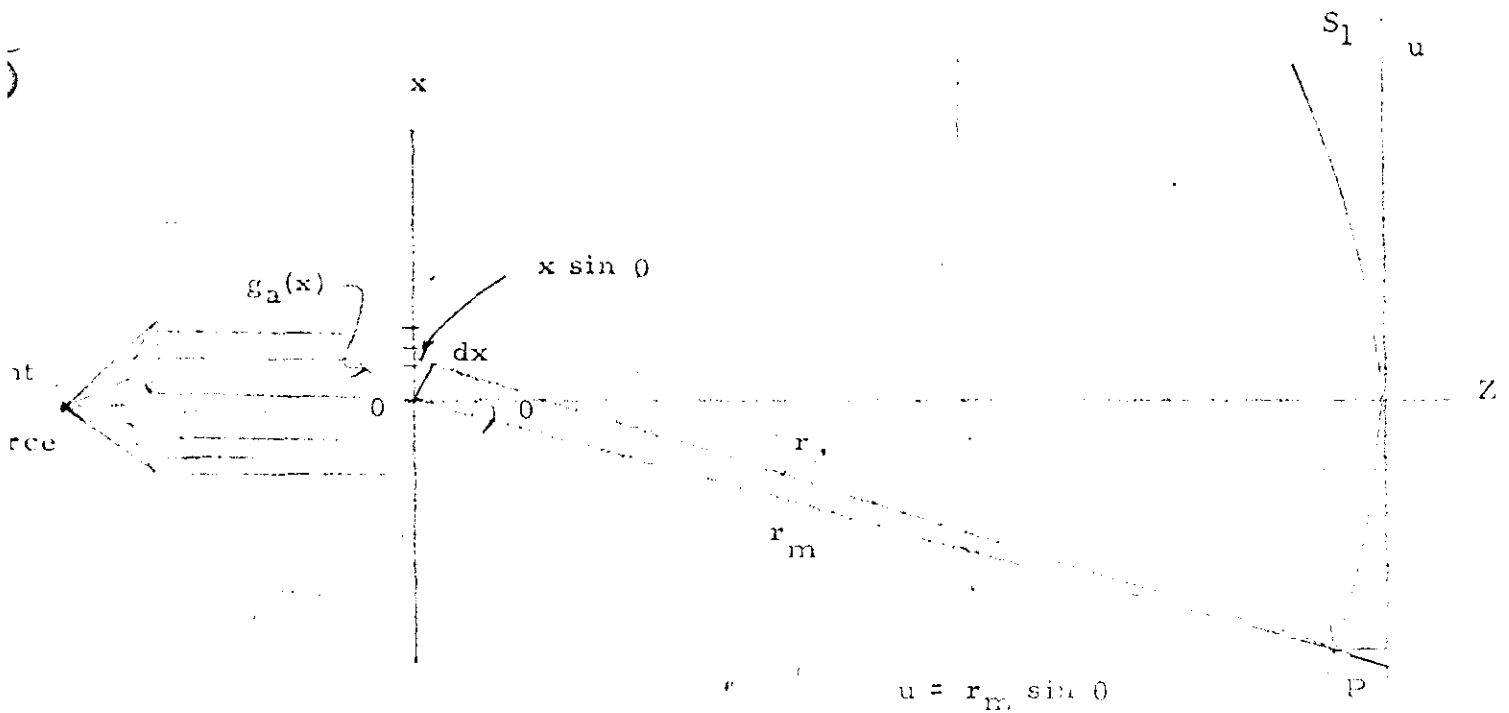


FIGURE 8-5 3G4 Geometry of Fraunhofer Diffraction

If $g_a(x)$ has small enough variance and we consider

$$r_m \gg \text{var} [g_a(x)] \quad (14)$$

(Fraunhofer diffraction), then to a good approximation, (considering the appropriate columnating lens to be along S_1)

$$r = r_m + x \sin \theta = r_m + x u / r_m \quad (15)$$

where $u = r_m \sin \theta$. Then the contribution of amplitude from dx at P (or u)

$$e^{-j2\pi r_m / \lambda} g_a(x) e^{-j2\pi x u / r_m \lambda} dx \quad (16)$$

and the total contribution at P may be obtained by summing the infinitesimal contributions at all x

$$e^{-j2\pi \frac{r_m}{\lambda}} \int_{-\infty}^{\infty} g_a(x) e^{-j2\pi x \frac{u}{r_m \lambda}} dx = f_a\left(\frac{u}{r_m \lambda}\right) = f_a(\xi) \quad (17)$$

We thus see that the amplitude distribution (henceforth called the diffraction pattern) of radiation U at a distance r_m from the origin of x is given by the Fourier transform of the aperture distribution. The function $g_a(x)$ may be recovered from f_a by inverse Fourier transformation (provided the spread of f_a is small)

$$g_a(x) = K \int_{-\infty}^{\infty} f_a(\xi) e^{j2\pi x \xi} d\xi \quad (18)$$

Obviously, the diffraction pattern represents the impulse response of the system, i. e.

$$f_a(\xi) = h_a(u) = w(-u) \quad (19)$$

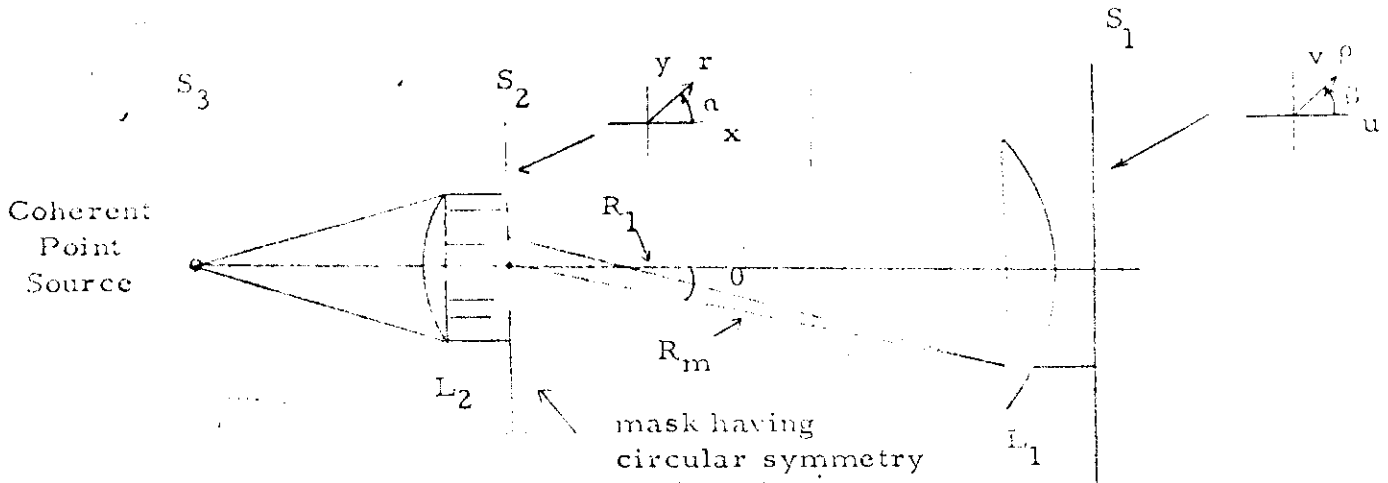
on the line u .

The derivation presented above may easily be extended to two dimensions: the point response or diffraction pattern is the two dimensional Fourier transform of the aperture distribution (neglecting aberrations and other lens distortions), i. e.

$$h_a(\zeta, \eta) = K \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_a(\gamma, \rho) e^{-j2\pi(\gamma\zeta + \rho\eta)} d\gamma d\rho \quad (20)$$

where the independent variables are appropriately chosen and scaled as either spatial or angular variables depending on how and where lenses are used.

In cases where circular symmetry exists, i. e., in many lens systems, the 2-dimensional Fourier transform becomes the Hankel transform. For example, consider the geometry depicted in Fig. ³⁶⁵5.



$$R_m \sin \theta = \rho$$

FIGURE ~~5-5~~

365 Symmetry of Coherent Source

We can make the following substitutions:

$$\hat{g}_a(x, y) = g_a(r)$$

$$x + jy = r e^{j\alpha}$$

$$r^2 = x^2 + y^2$$

$$\hat{f}_a(u, v) = f_a(\rho)$$

$$u + jv = \rho e^{j\beta}$$

$$\rho^2 = u^2 + v^2$$

We now change the transform formula to polar coordinates and integrate over the angular variable.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{g}(x, y) e^{-j2\pi(ux+yv)} dx dy = \int_0^{\infty} r dr \int_0^{2\pi} g(r) e^{-j2\pi(\rho r \cos\alpha \cos\beta + \rho r \sin\alpha \sin\beta)} d\alpha \quad (22)$$

$$= \int_0^{\infty} r dr \int_0^{2\pi} g(r) e^{-j2\pi\rho r \cos(\alpha - \beta)} d\alpha \quad (23)$$

$$\text{But } J_0(z) = \frac{1}{2\pi} \int_0^{2\pi} e^{-jz \cos \phi} d\phi \quad (24)$$

so that the Fourier transform pair for circular symmetry may be written as the zero order Hankel transform pair (Fourier transform with a zero order Bessel function kernel.)

$$\begin{aligned} f(\rho) &= 2\pi \int_0^{\infty} g(r) J_0(2\pi\rho r) r dr \\ g(r) &= 2\pi \int_0^{\infty} f(\rho) J_0(2\pi\rho r) \rho d\rho \end{aligned} \quad (25)$$

Thus, in the case of circular symmetry, the aperture distribution and diffraction pattern are a Hankel transform pair: in other words, the impulse response is the Hankel transform of the aperture distribution.

It should be noted that for many simple coherent source systems, the transfer function $H(v_x, v_y)$ can be obtained by applying a Fourier transformation to the diffraction pattern. The resulting distribution is (for even symmetric distribution) precisely that of the aperture distribution with a scale factor change.

§. Impulse Response: Incoherent Source Systems

In incoherent source systems, where radiation intensities add, we must compute the intensity point response. In general, the impulse response depends on geometry and lens quality of the system. Thus, we may obtain the intensity impulse response from the amplitude impulse response by the relation

$$h_i(x, y) = |h_a(x, y)|^2 = h_a(x, y) h_a^*(x, y) \quad (26)$$

In situations where $h_a(x, y)$ is the diffraction pattern, we can easily compute the point response as the squared magnitude of the coherent source amplitude diffraction pattern. The transfer function may then be found by Fourier transforming this squared modulus function, or from the autocorrelation theorem:

$$|F(\eta)|^2 = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f^*(z) f(z - \gamma) dz \right] e^{-j2\pi\gamma\eta} d\gamma \quad (27)$$

we can show that

$$H_i(\nu_x, \nu_y) = K_1 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_a^*(\alpha, \beta) g_a(\alpha - \nu_x, \beta - \nu_y) d\alpha d\beta \quad (28)$$

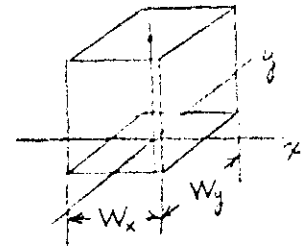
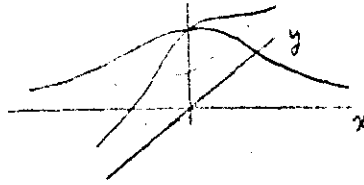
where α, β are suitably scaled variables. Thus, the transfer function for incoherent source systems can be obtained by autocorrelating the aperture distribution.

It should be noted here that, since the wavelength of the light source enters into the calculation, we must contemplate its meaning for non-monochromatic sources. A good practice would probably be to use the mean or "carrier" wavelength of the (usually narrowband) source in the computations.

6. Resolution

The resolution area or measure of the point response spread of the optical system can be defined in several ways - all extensions of the one-dimensional case. Four examples are depicted in Figures 6 and 7. ^{366 367} These definitions may sometimes be more simply expressed in the transform domain, as shown in the Figures.

Equivalent Width:

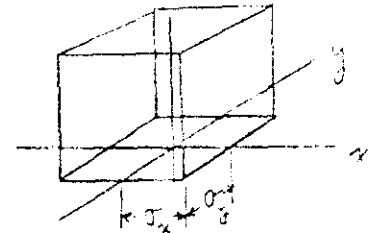
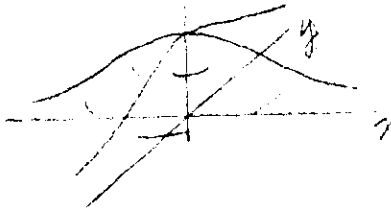


$$\left\{ \begin{array}{l} W_x = \frac{\int_{-\infty}^{\infty} h(x, 0) dx}{h(0, 0)} \\ W_y = \frac{\int_{-\infty}^{\infty} h(0, y) dy}{h(0, 0)} \end{array} \right\} \quad (29)$$

or

$$\left\{ W_{xy} = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) dx dy}{h(0, 0)} \right\} \quad \left[\begin{array}{l} \text{Square Case} \\ \frac{1}{2} \end{array} \right] \quad (30)$$

Deviation:



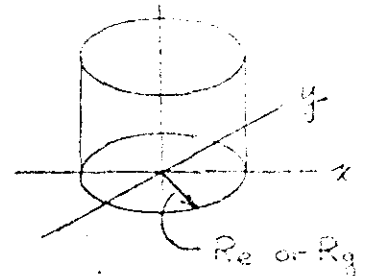
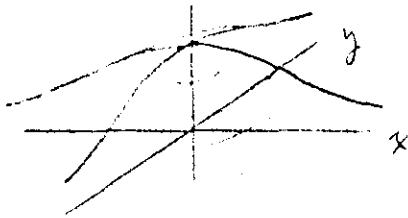
$$\sigma_x^2 = \frac{\int_{-\infty}^{\infty} (x-\bar{x})^2 h(x, y) dx dy}{\int_{-\infty}^{\infty} h(x, y) dx dy} ; \quad \sigma_y^2 = \frac{\int_{-\infty}^{\infty} (y-\bar{y})^2 h(x, y) dy dx}{\int_{-\infty}^{\infty} h(x, y) dy dx}$$

(31)

FIGURE 6 - Resolution 361

Equivalent radius:

$$R_e = \left[\frac{1}{4\pi} \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) dx dy}{h(0, 0)} \right]^{-1/2} = \left[\frac{1}{4\pi} \frac{H(0, 0)}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(v_x, v_y) dv_x dv_y} \right]^{-1/2} \quad (32)$$



Radius of Gyration:

$$R_g = \frac{1}{4\pi} \sigma_{xy} = \frac{1}{4\pi} \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [(x-\bar{x})^2 + (y-\bar{y})^2] h(x, y) dx dy}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) dx dy}$$

$$= - \frac{1}{16\pi^3 H(0, 0)} \left[H''_{v_x v_x}(0, 0) + H''_{v_y v_y}(0, 0) + \frac{(H'_{v_x}(0, 0))^2}{H(0, 0)} + \frac{(H'_{v_y}(0, 0))^2}{H(0, 0)} \right]$$

where

$$H'_{v_x}(v_x, v_y) = \frac{\partial H}{\partial v_x} \quad (33)$$

FIGURE 7 - Resolution

367

7
4. Scanning

Consider again the geometry depicted in Figure 1. The relation between object and image spaces is given by the convolution integral,

$$r(x', y') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x'-x, y'-y) s(x, y) dx dy \quad (34)$$

↑ (Impulse response)

or the correlation integral

$$r(x', y') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(x-x', y-y') s(x, y) ds dy \quad (35)$$

↑ (Weighting function)

If the process is merely the formation of an image on the space defined by (x', y') , the integral can be viewed as the superposition of a large number of point responses - - the object is resolved into many points or dots and convolution with the point response gives the image signal. However, in the discussion of scanning, it will be more convenient to use the weighting function interpretation. With the object considered stationary, the action of the scanning process is to move the weighting function (w_s) across the object and integrate to get the image as a function of the two dimensional shift, (x', y') as shown one-dimensionally in Figure 8.

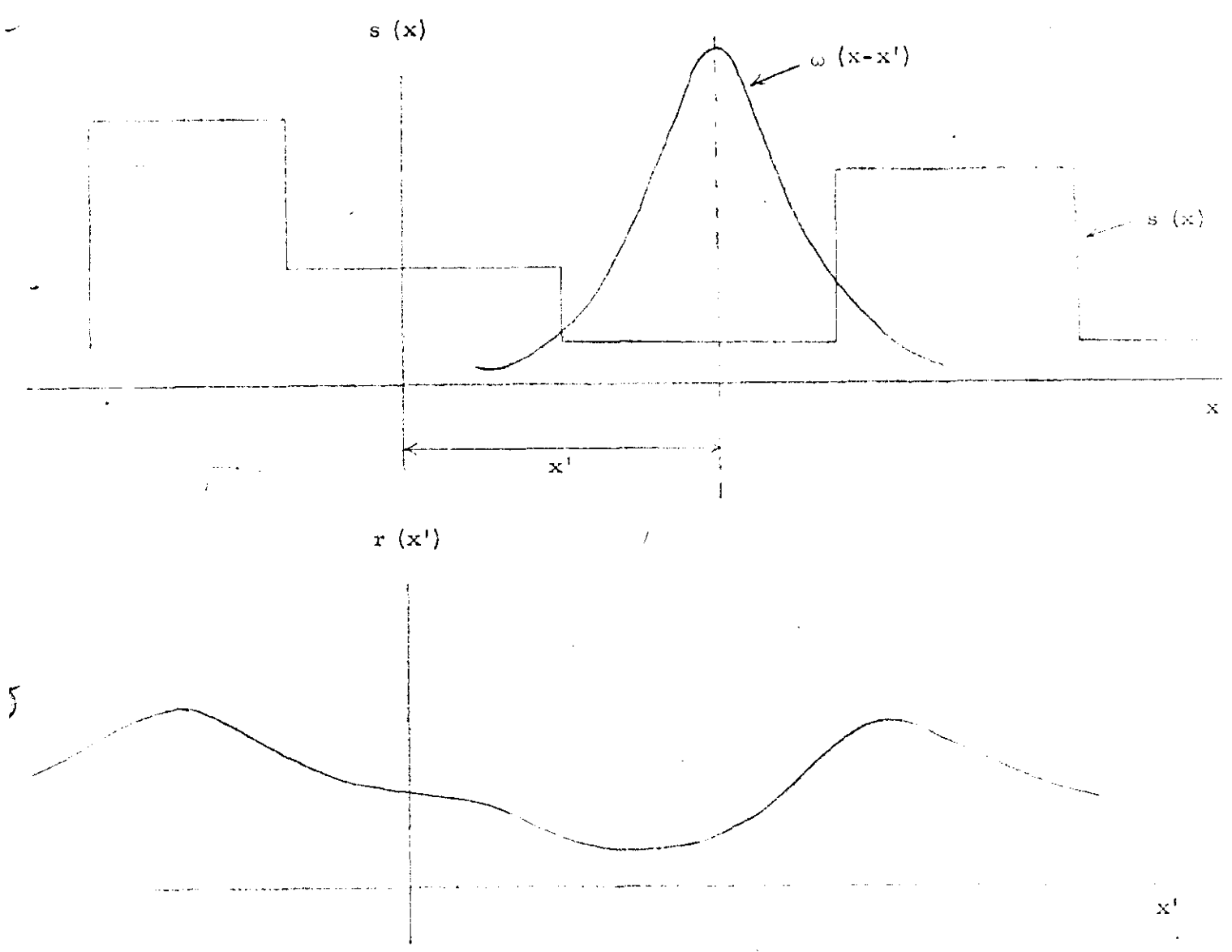


FIGURE 8.5.8 *35.8* Scanning

A simplification in the integral can be obtained (34) by making use of one of the definitions of resolution given in the preceding section. We assume that $h(x, y)$ is even - symmetric in both the x and y directions, otherwise the resolution definitions should be modified to take the lack of

symmetry into account. Using, for example, the definition of equation (31), we have

$$w(x, y) = h(0, 0) \Pi\left(\frac{x}{2\sigma_x}, \frac{y}{2\sigma_y}\right) = h(0, 0) \Pi\left(\frac{x}{2\sigma_x}\right) \Pi\left(\frac{y}{2\sigma_y}\right) \quad (36)$$

where $\Pi(z)$ is the rectangle function defined by

$$\Pi(z) = \begin{cases} 1 & \text{for } |z| < 1/2 \\ 1/2 & \text{for } |z| = 1/2 \\ 0 & \text{for } |z| > 1/2 \end{cases} \quad (37)$$

Then, the scanning integral (35) becomes more simply

$$r(x', y') = \int_{x'-\sigma_x}^{x'+\sigma_x} \int_{y'-\sigma_y}^{y'+\sigma_y} s(x, y) dy dx \quad (38)$$

which is a two-dimensional running mean (without interval normalization).

Since (38) can be expressed as a convolution, this is equivalent in the transform domain to multiplication by the two dimensional sinc function, i. e.,

$$R(v_x, v_y) = 4\sigma_x\sigma_y \text{sinc}(2\sigma_x v_x) \text{sinc}(\sigma_y v_y) S(v_x, v_y). \quad (39)$$

Now we are faced squarely with the problem of connecting the two-dimensional image signal space into a one dimensional waveform. This waveshape will depend strongly on the method of scanning used (i. e., TV roster-type interlaced, y -direction roster, circular, cycloidal, etc.). We will not try to explore the various methods here, but will construct an example to show one approach for determining this waveform.

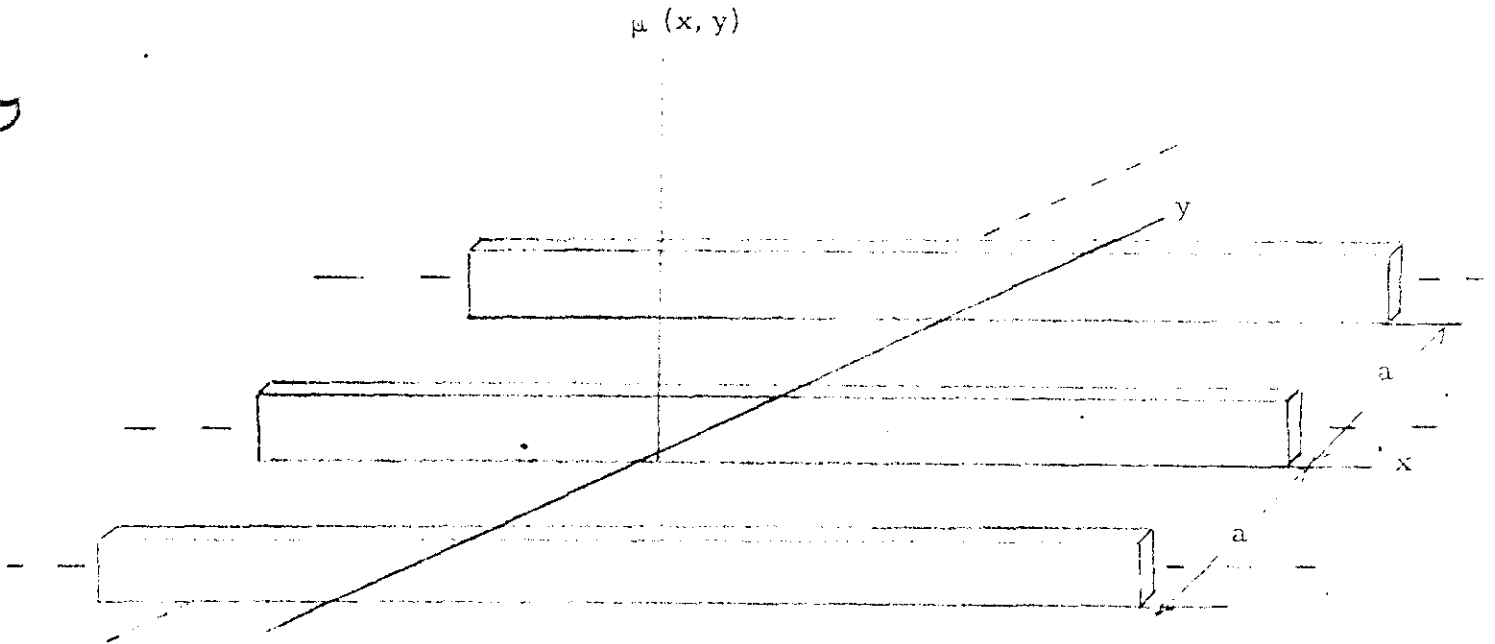
Example: Let the object be scanned by a system having the weighting function

$$w(x, y) = \text{rect}\left(\frac{x}{2\sigma_x}\right) \text{rect}\left(\frac{y}{2\sigma_y}\right) \quad (40)$$

We further define a scanning or image sampling function

$$\mu(x, y) = \frac{1}{a} \text{III}\left(\frac{y}{a}\right) \text{rect}\left(\frac{x}{b}\right) \text{rect}\left(\frac{y}{c}\right) \quad (41)$$

(a is the scanning interval; b, c are the dimensions of the document printing) which describes the scanning method - - for this example; the scanning is mathematically the "grate" function of Figure 9. Thus, if we scan with uniform velocity ($\dot{x}' = \dot{x}$) in the x-direction from $-b/2$ to $b/2$ and in



369
FIGURE 9: Small Segment of Scanning Space

discrete steps from $c/2$ to $-c/2$ in the y direction we can describe the resulting image as

$$r_s(x', y') = \left[\text{rect} \left(\frac{x}{2\sigma_x} \right) \text{rect} \left(\frac{y}{2\sigma_y} \right) \otimes s(x, y) \right] \mu(x', y') \quad (42)$$

Since the width of $r(x, y)$ can be no greater in the x -direction than $b + 2\sigma_x$, we can hypothetically place the grates end for end and let the time axis be divided into sections of duration

$$T = \frac{b + 2\sigma_x}{\dot{x}} \quad (43)$$

such that

$$t = \frac{x + \frac{b}{2}}{\dot{x}} + (n-1)T \quad (44)$$

$$n = \left\lfloor \frac{t}{T} \right\rfloor + 1 \quad (45)$$

where $\lfloor Z \rfloor$ = greatest integer less than Z .

Thus, the formal mathematical description of the time response is

$$v_r(t) = \int_{x' - \sigma_x}^{x' + \sigma_x} \int_{y' - \sigma_y}^{y' + \sigma_y} s(x, y) dx dy \quad \left. \begin{array}{l} \text{where} \\ x' = \dot{x} [t - (n-1)T] - \frac{b}{2} \\ y' = \frac{nc}{2} - a(n-1) \end{array} \right\} \quad (46)$$

We see that $v_r(t)$ is pseudo-periodic; similar values of x correspond to values of t at periodic intervals nT . If, for example, the document is a printed page, the waveform will exhibit pseudo-periodicity as a result of the spaces between lines and letters. Let us assume that the stimulus

is of the form

$$s(x, y) = \text{rect}\left(\frac{x}{a}, \frac{y}{\beta}\right)$$

as depicted in Figure 10. Then the resulting waveform will be that of

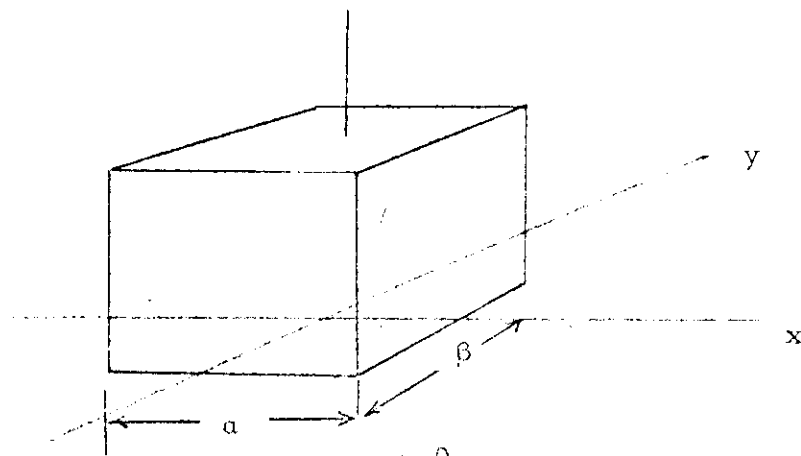


FIGURE 10 *3510 Stimulus Function*

Figure 11. Thus, we see in the simple case illustrated that a non-periodic

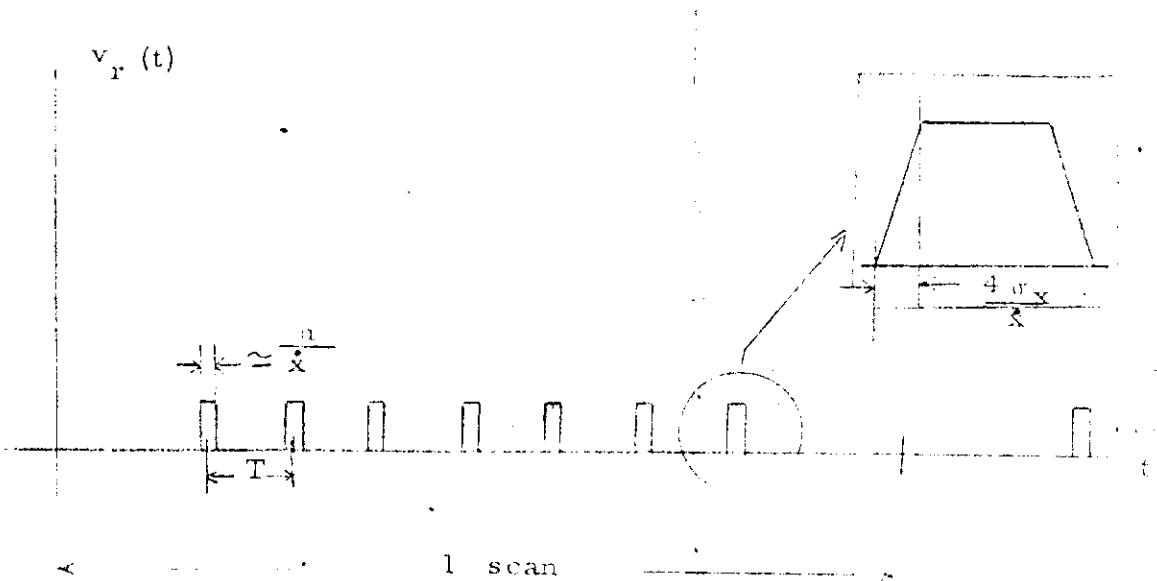


FIGURE 11 *3511 Pseudo-Periodic Signal*

signal in the spatial domain gives rise to a pseudo periodic signal in the time domain.

In the usual case of random spatial signals, we might do well to consider the spatially sampled waveform $r_s(x', y')$ as yielding an ensemble of time waveforms

$$v_r(t_n, n)$$

where v_r is the random variable and n is the raster number or sample number (actually, v_r is only an approximate of the ensemble since it is not of infinite extent. Then we compute the auto-correlation of the process described by v_r and deal in the frequency domain with the power density spectrum. We might expect sharp peaks in the spectrum due to periodicity.

The relation between spatial frequency and temporal frequency is rather abstruse. Evidently we can perform a Fourier transformation on the right-hand side of (46) but justification (considering the work involved) is questionable. Evidently $f \neq v_x \dot{x}$ so that no simple relation exists between the spatial frequency component, and the temporal frequency domain of the waveform. The non-linear operational characteristic of sampling certainly explains this fact, but may also provide a clue to some intermediate transformation which would facilitate a simpler transition between the domains of spatial and temporal frequency. Perhaps other transformations can be found which take advantage of the pseudo-periodic nature of the scan-waveform to effect bandwidth and time compression.

In summary, several problems for further study are revealed by the foregoing analysis:

- (1) To find a more definite relationship between spatial and temporal frequency, enabling prediction of the temporal wave-

form spectrum from a knowledge of the scanning function and object spatial signal spectrum.

(2) Synthesis of a suitable scanning transformation which effects desirable time and bandwidth compression. Thus, perhaps $\mu(x', y')$ is a function of the signal $s(x, y)$ - i. e., it is a nonlinear adaptive transformation. (variable velocity scan).

8. Applications

The preceding theory indicates that many of the concepts of signal processing in electronic systems theory can be carried over to the two dimensional signals found in optical systems. Indeed, some effort has been directed toward development of spatial filtering techniques as described in some of the references in the bibliography. In many cases the filter synthesis is more readily achieved in the optical case than in the electrical analog. For example, electrical filters are synthesized using lumped constant parameter elements; this restricts the form of the transfer function to the quotient of two finite polynomials. However in optical systems, any transfer function can be realized as the autocorrelation of some aperture distribution (equation (2b)) which in turn may be realized by appropriately shading a transparent film.

Further, the usual realizability restriction on impulse response in temporal systems, i. e.,

$$h(t) = 0 \text{ for } t < 0$$

does not find counterpart in spatial systems since an impulse response can exist to any side of the stimulus.

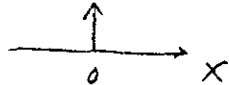
In summary, the techniques of equalization, mean square filtering, and matched filtering have been used effectively in optical systems to improve resolution and contrast, and to discriminate against background noise such as graininess. Spatial sampling techniques can be easily extended from their temporal counterparts to further effect signal processing - e. g., any raster scanning system is a spatial sampled data system, and may be further sampled to yield a series of dots. The required raster and dot intervals are determined by the spatial frequency bandwidth of the two dimensional signal. Further, many optical transfer functions may be realized where similar electrical realizations are impossible. It may even be possible to realize low pass type complex valued impulse responses in optical systems (e. g. using polarized light sources) - a clear impossibility when dealing with electrical waveforms which are always real. Thus, most of linear system theory used extensively by the electrical engineer becomes a pervasive tool which gives new dimension to the analysis and synthesis of optical systems.

9. BIBLIOGRAPHY

This is not intended as complete, but provides a filtered sample of the pertinent work done in this field. Other references can be found in these papers.

1. O'Neill, "Spatial Filtering in Optics", IRE PGIT Transactions, June 1956.
2. Cheatham, T. P. and Kohlenberg, A., "Optical Filters - Their Equivalence to and Difference from Electrical Networks," IRE Convention Record, 1954, part 4.
3. Elias, P., Grey, D.S., and Robinson, D. Z., "Fourier Treatment of Optical Processes," Journal Opt. Soc. America, February 1952
4. Kovasynay, L.S.G., and Joseph, H.M., "Image Processing," Proc. IRE, May 1955.
5. O'Neill, E.L., "Transfer Function for an Annular Aperture," Journal Opt. Soc. America, April 1956.
6. Elias, P., "Optics and Communication Theory," Journal Opt. Soc. America, April 1953.

10. SYMBOLS AND NOTATIONS

1. $\delta(x)$ is the unit impulse symbol \rightarrow 

2. \otimes is notation for convolution operation between two functions,

i. e.

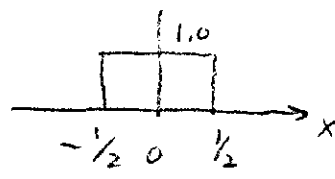
$$f(x) \otimes g(x) = \int_{-\infty}^{+\infty} f(x'-x) g(x) dx = \int_{-\infty}^{+\infty} f(x) g(x'-x) dx$$

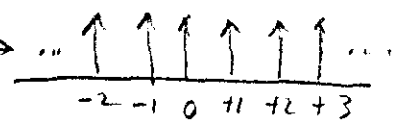
3. Fourier spectra (forward transforms) are always represented in caps while the functions (inverse transforms) are represented in lower case, e. g.

$$g(x) = \int_{-\infty}^{+\infty} G(f) e^{j2\pi fx} df$$

$$g(f) = \int_{-\infty}^{+\infty} g(x) e^{-j2\pi fx} dx$$

4. Sinc $x = \frac{\sin \pi x}{\pi x}$

5. $\Pi(x) = \begin{cases} 1 & \text{for } x < 1/2 \\ 1/2 & \text{for } x = 1/2 \\ 0 & \text{for } x > 1/2 \end{cases} \rightarrow$ 

6. $\text{III}(x) = \sum_{n=-\infty}^{+\infty} \delta(x-n), \quad n = 0, 1, 2, \dots \rightarrow$ 

IV. GENERAL LIST OF IMAGE CODING SYSTEM

The various potential image coding systems are listed in Table IV. Then each system is described on a separate page with supplementary notes where available. The systems are grouped in accordance with the principal use such as: continuous tone, text, and drawings. Some systems are more universal and are listed under the general case preceding the above mentioned sections. The reference numbers in Table IV are not the same file numbers used during the research and development work, but are new numbers for this report.

Table IV. IMAGE SCANNING AND CODING
METHODS

Name	Notes	Ref. No.
Section A: CONTINUOUS TONE DOCUMENTS		
Facsimile	A basic scan of 50 lines per centimeter or 1.5×10^6 binary digits per 8-1/2 x 11 page is used for reference.	1
Variable Velocity	Maximum scanning velocity is reduced in proportion to the amount of picture detail.	2
Partial Transmission	A fraction of the signal samples are transmitted, the gaps are filled by interpolation.	3
Predictive Coding	Difference signal between predicted and observed signal is transmitted.	4
Band Splitting	Two signals are used, one for the low frequency, and one for the high frequency or edge information.	5
Analog Compression Coding	Sampling, quantization, and interlocking amplitude and position pulse trains.	6
Deviation and Rate Coding	Signal is approximated by first two terms of Fourier series: a_0 (deviation) and a_K (rate), where f_K is adaptive.	7
Bandwidth Switching With Adaptive Speed Scan	Error Detection redundancy and control information would be stored and then added during low bandwidth periods of line scan.	8

Name	Notes	Ref. No.
------	-------	----------

Optical Processing	Two alternatives: Photoelectric crystal transformations, and optical matrix operations. For image plane processing (e- modulation).	9
--------------------	---	---

Section B: HANDWRITTEN, TYPED OR PRINTED TEXT.

Run-length Coding	Counting and recoding of runs of black and white.	10
-------------------	---	----

Adaptive Run-length	Switches between different run-length codes for different documents.	11
---------------------	--	----

Buffer-Controlling Scanning to RLC	A run-length decoder scanner is required to control the buffer.	12
------------------------------------	---	----

FM/PM Adaptive System	Frequency is proportional to distance between transitions. Transitions between black and white coded by phase.	13
-----------------------	--	----

Pattern Coding (Letter Segment Coding)	Recognizes vertical horizontal lines various curved segments.	14
--	---	----

Pattern Coding (Letter Sample Spot Coding)	Samples a subset of matrix covering character.	15
--	--	----

Full Line Scan	Signal proportional to total fraction of space that is black and is then recoded.	16
----------------	---	----

Coordinate Coding	Areas where black occurs are coded by section identification with run-length coding inside section.	
-------------------	---	--

Name	Notes	Ref. No.
Two Dimensional Line Scan	Prescan measures vertical and horizontal redundancy and groups reading fibers for optimum type.	18
Matrix Processing	Optical matrix of fibers reads data which is reduced by digital filter in network.	19
Optimum Fixed Block Length Coding	Fixed block length run-length coding where length is optimum	20 (20 with 10 No. 2)
Section C: LINE DRAWING		
Coordinate Engineering Drawing Compression	Type of line and end points coded	21
Cyclic Code Compression	Adaptation of feedback shift registers to compression coding.	22
	(Appendix on FER, compression of coloration, with recurrent code)	
Analogy from Biological Systems	Consideration of error code compression and multichannel in neural system of rabbit as potential analogy.	23

SYSTEM #1: BASIC FACSIMILE

Description: Limited to binary representation of each square of size $\epsilon \times \epsilon$. For $\epsilon = 0.02$ centimeters, ϵ -entropy is 1.5×10^6 bits per $8\text{-}1/2 \times 11$ " page, $\epsilon = 0.02$ corresponds to 125 lines/inch.

Typical Usage: UPI, Times Facsimile, Alden and Murison use straight facsimile with analog signal for photographs. Western Union uses telefax system.

References:

Charles R. Jones, Facsimile, N.Y. Murray Hill Books (1949),
Part One: What Facsimile Is; part two: How Facsimile Works,
Part three: Present-Day Facsimile Systems, and Part Four:
Servicing.

Radio Corp. of America, Radio Facsimile, An assemblage of
papers from engineers of RCA laboratories, N.Y. RCA

Institutes Technical Press Vol. I, October 1938.

*Note: For conversion of analog signal to 20-level digital signal, the ϵ -entropy is 5×10^6 binary digits.

Edward Parker, "Bibliography of Literature References on
FACSIMILE TRANSMITTERS AND RECEIVERS, IBM Kingston
N.Y., Report TP 60-4507, N. W. 1960, (IBM - San Jose File
BIB-648)

Dept. 673, "Bibliography of Literature References on High
Speed Electronic Facsimile (Telephotography) Code 05.11,
IBM, Endicott, N. Y., Feb. 15, 1954, Revised May 16, 1958
Revised Aug. 28, 1957, (IBM San Jose File No. BIB-777)

SYSTEM # 2: VARIABLE VELOCITY

Description: Variable velocity scanning has been proposed by various people for compressing the bandwidth required for facsimile and television signals. The basic technique is to reduce the scanning speed in proportion to the increase in picture details. The 1953 proposal of E. C. Cherry and G. G. Gouriet contain a good analysis of variable velocity scanning. ⁽¹³²⁾

References:

E. C. Cherry and G. G. Gouriet, "Some Possibilities for the Compression of Television Signals by Recording" Second London Symposium; Butterworths Scientific Publication (1953); in Communication Theory edited by Willis Jackson, pp. 328 - 353.

E. C. Cherry and G. G. Gouriet, paper of same title as above in Proc. IEE, January 1953.

Analysis of Variable Velocity Scanning

The following analysis is abstracted from the Cherry & Gouriet paper, ref. 2

(Re: Cherry and Gouriet, "Some possibilities for the Compression of TV Signals by Recoding", Proc. IEE, January 1953).

This scheme is essentially a velocity-modulation scheme where a specific maximum scanning velocity is reduced in proportion to the amount of picture detail. The basic criterion is the slope modulus,

$$\left| \frac{dv}{dx} \right|$$

where v is the waveform representing picture intensity along a line.

A parameter which measures the average amount of detail in a line or field, L , is the detail factor, D_f , where

$$D_f = \frac{1}{L} \int_L \left| \frac{dv}{dx} \right| dx.$$

(Note: the authors show that this factor is related to source conditional entropy - i. e., the entropy of a first order Markov source - provided the conditional probabilities are exponentially distributed).

The basic scheme is shown in ^{Fig 4.2A} ~~the diagram~~ below from Cherry and Gouriet:

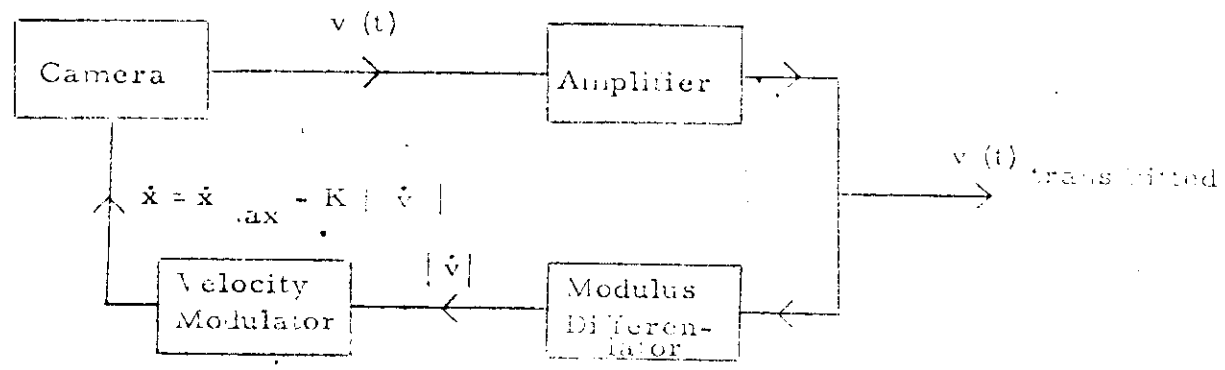


Fig 4.2A Variable Velocity TV System

The instantaneous velocity can be expressed in either of two ways - temporally or spatially

$$\dot{x} = \dot{x}_{max} \cdot \dots \Rightarrow \text{(temporal)}$$

$$\dot{x} = \dot{x}_{max} \cdot \frac{1}{1 + \dots \frac{d}{dx}} \Rightarrow \text{(spatial)}$$

It is shown that for two level intensity pictures, the bandwidth com-

pression factor, F_c ,

$$F_c = 1 - \frac{W}{W_0} = 1 - \frac{1 + (n-1)D_f}{n} = \frac{(n-1)}{n} (1 - D_f)$$

where W_0 is the constant velocity bandwidth,

W is the variable velocity bandwidth, and n is the ratio of maximum to minimum velocity (note that the authors use the factor W/W_0 , but it is more reasonable to use one minus this quantity).

When the scheme is extended to grey-scale pictures, the compression is reduced; thus if λ is the number of quantum levels in the

picture, the bandwidth compression factor becomes

$$F_L = \frac{(n-1)}{n} [1 - (l-1) = D_F]$$

Note that this is the bandwidth compression assuming that all transitions are of amplitude $v_{max} / (l-1)$ larger transitions will have greater compression factors - so that this scheme achieves bandwidth compression at the expense of dynamic range. The reason for this is shown in the ~~figure below~~ *Fig 4.2B:*

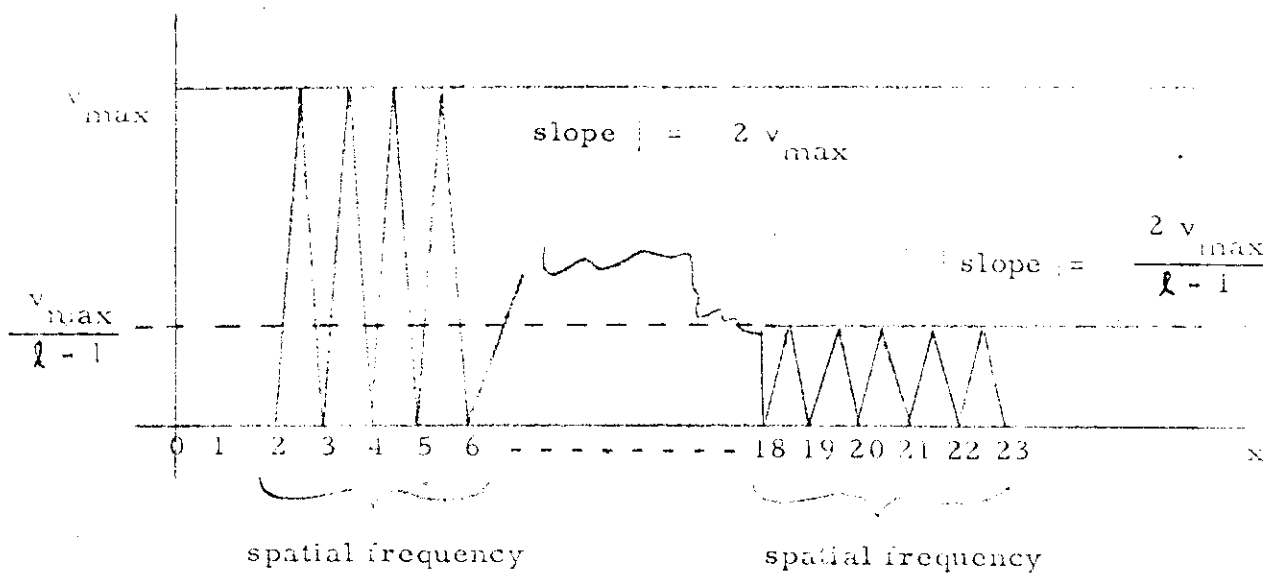


Fig 4.2B Spatial Frequency Curves $= v_1$ $= v_1$

As the summer proceeds along x , it encounters the first series of transitions from $x = 0$ to $x = 6$. The velocity x_1 is adjusted so that the temporal frequency is

$$f_1 = x_1 v_1$$

where $\dot{x}_1 = x_{\max} - K \frac{dv}{dt} = \dot{x}_{\max} - 2K v_{\max}$

However when the scanner reaches the detail in the interval

$$18 < x < 23$$

the velocity \dot{x}_2 is greater than \dot{x}_1 since

$$\dot{x}_2 = \dot{x}_{\max} - \frac{2 v_{\max}}{1 - 1}$$

But the spatial frequency is the same, so that the temporal frequency $f_2 > f_1$ and less bandwidth compression is possible for the same spatial frequency detail.

Comment: One obvious solution is to make the scanning velocity constant for all sections having the same spatial frequency or transition duration. Thus, the instantaneous velocity should be modified by the amplitude of the transition involved in such a manner that it is reduced only for increased spatial frequency. Thus the detail factor that Cherry and Gouriet have chosen becomes the fly in the ointment for continuous contrast pictures. A better choice might be

$$\text{Detail figure } F_D = \frac{N_T}{L}$$

where N_T is the number of transitions (quantum jumps greater than or equal to 1) in the interval L . Then we might change the instantaneous

velocity according to

$$\dot{x} = \dot{x}_{\max} - K \frac{N_T}{L}$$

Another possibility is to convert the multilevel picture to a two-level, variable spatial frequency picture which can be variable-velocity scanned (or variable time-base scanned). This could be done instantaneously by using, for example, pulse width modulation where the width of the pulse corresponds to the sample height. Then we must scan rapidly over sections of the time scale having constant width pulses and reduce speed over sections where the width is changing. I have no definite hardware system in mind yet, but the principle would be analogous to variable velocity scanning of two-level signals. However here we are applying a variable scale factor to the time base of a two-level waveform, i.e, we perform a time base transformation, (obviously a reasonable signal transformation of the PWM signal must be found to realize this).

Note: The authors discuss a modification of the variable-velocity scan scheme which would be more applicable to grey-scale pictures. In this scheme signal temporal frequency is fed back instead of slope to control the speed of scanning a previously quantized picture. No details regarding potential compression are given.

In January 1962 Newell and Geddes concluded at an IEE meeting that bandwidth compression for TV was not very practical (G. F. Newell and W. K. E. Geddes, "Tests of Three Systems of Bandwidth Compression of Television Signals," Proc. IEE, Part B, Vol. 109, July 1962, pp. 311-324.

Dr. E. C. cherry questioned their conclusions in the discussion.

In June 1962, Dr. Colin Cherry announced that successful compression of a TV picture (half-tone) of a girl had been demonstrated with a compression ratio of 2.85 to 1.

(Electronic News, Monday, June 11, 1962, page 23, "UK Team Reported Nearing Practical Visual Telephone.") He did not disclose any technical details, so we do not know if his present feasibility models use the variable velocity techniques. Dr. Cherry further disclosed that for simple block and white drawings are typescript a compression ratio of 20 to 1 has been demonstrated. His press release stated that information compression depends on the establishment of "edges" of different picture density, which results in pulses which are unequally spaced. Then the pulses are more equally spaced by a coder.

SYSTEM #3:

PARTIAL TRANSMISSION

~~3. Partial Transmission~~

In this type of scheme only a fraction of the signal samples are transmitted - linear interpolation is used to reconstruct the signal (with some loss in fidelity) at the receiver.

- (i) One reference to a scheme of this type is Huang, T.S., "A Method of Picture Coding", Quarterly Prog. Rep., MIT Res. Lab. of Elect., April 15, 1960, p. 109. Results are included in the reference. When every other sample was deleted (sampling at half the Nyquist rate) and averaging (interpolating) was used to "fill in the blanks", a staircase effect was observed (horizontal or vertical lines). The scheme was modified such that extra points near sudden contour changes (as determined by a change in intensity over a threshold) were

~~4A.11~~ 4-3.2
~~-7.2~~

sent along with the half-frequency samples. This increases the source rate such that

$$R(x) > 3 \text{ bits per sample}$$

(64 level or 6 bit samples to start with), but the staircase effect is eliminated. A further extension of the scheme was incorporated where every fourth sample point of every fourth line was sent (spatial sampling frequency = 1/4 Nyquist rate) and in between points were sent according to four threshold criteria (i. e. extra points sent only if contour change was abrupt enough). In this manner, the source rate was reduced from 6 bits/sample to 1.55 bits per sample with only a small loss in sharpness of the picture! Thus the fidelity loss which we expect using this method manifests itself in picture fuzziness - loss of fine detail. Comment: if we could use a scheme of this type together with some statistical methods or band-splitting to transmit the fine detail, we might realize a tremendous saving in bandwidth over ordinary PCM. Note also that a disadvantage of the scheme described by Huang is the high buffer capacity required to store and compare four lines of information. Perhaps some sort of queuing system could be devised to reduce costs.

(ii) (Ref: Youngblood, "Picture Processing", MIT Electronics Lab Quarterly Progress Report, January 15, 1958). In this scheme, picture samples are eliminated by an error criterion established with reference to the diagram ^{Fig 4.3A} ~~below~~ (from Youngblood).

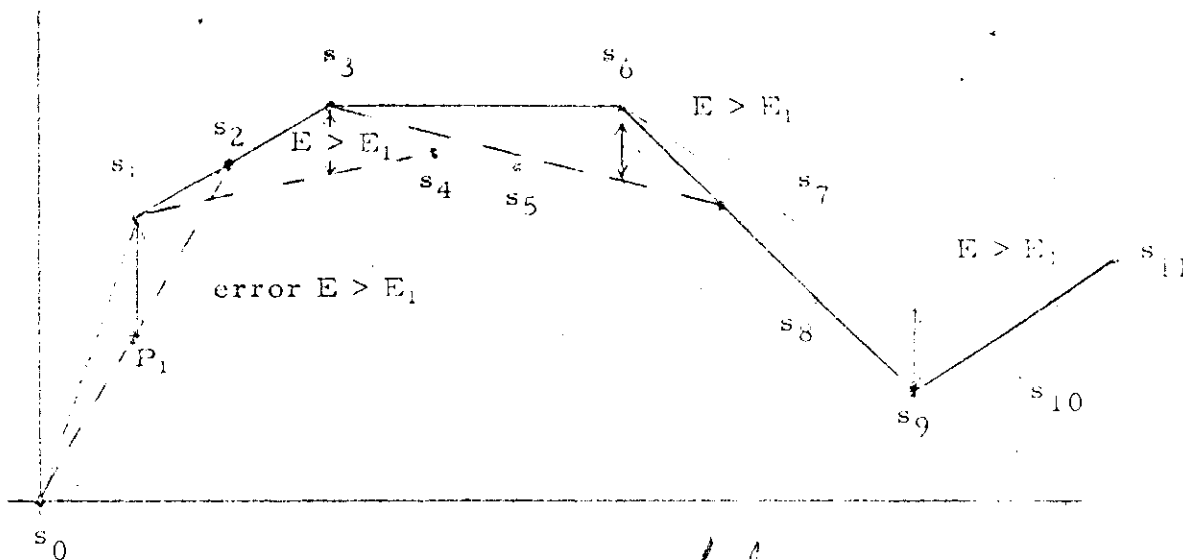


Fig. 4.3A Interpolation method

A trial straight line (dotted) is drawn from the origin (s_0) to the sample s_2 . If the projected value of s_1 , i.e., P_1 is such that the error;

$$E = s_1 - P_1$$

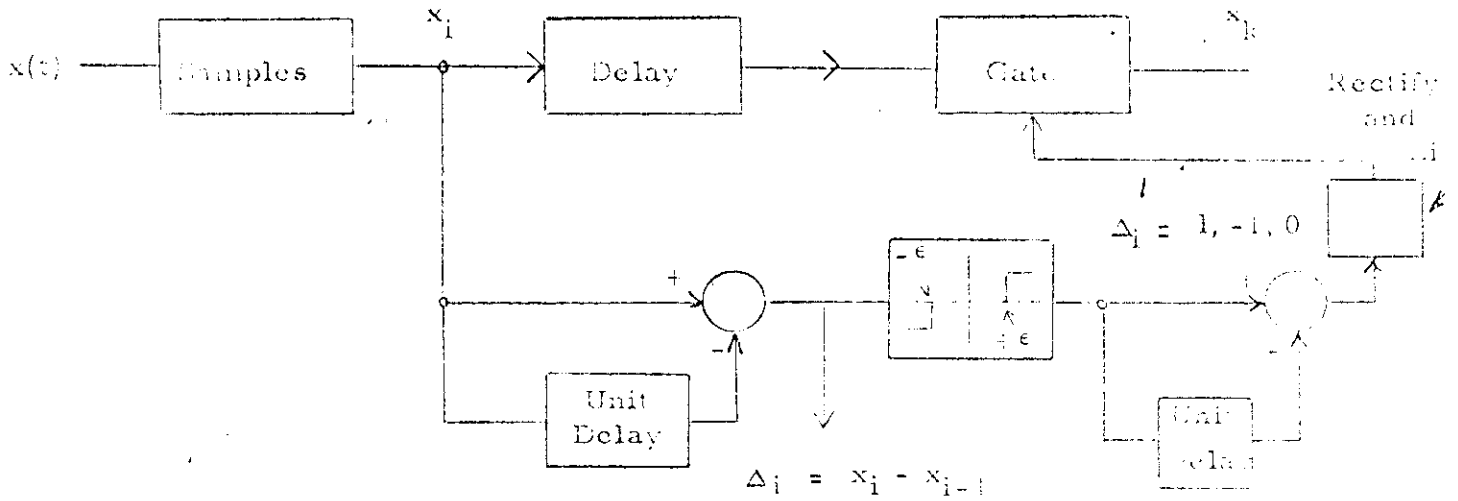
is greater than a certain threshold, E_1 , then the sample point s_1 is transmitted. If

$$E < E_1$$

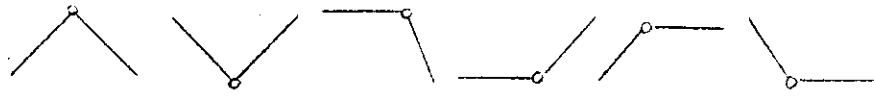
then the point is not transmitted. The process is continued (using the last transmitted sample, in this case s_0 , as the origin of the trial straight line) until $E > E_1$ for the line terminating on s_i . Then s_{i-1} is transmitted and the new origin of the line is s_{i-1} . Thus the straight line approximation to the actual curve is obtained - only the corner

points are transmitted and the receiver linearly interpolates between them. Since the information is sent at a varying rate (rapidly changing intensity levels yield many samples per second, slowly changing levels yield few samples per second) information must be sent with the sample value which will enable the receiver to determine its position. The number of samples between corner points (sample run length) is the information which is sent with the sample amplitude which enables its correct positioning. Run length (sample interval) probabilities are computed and the lengths are coded to minimize the average code word length. With six-level quantization, source rates could be reduced from 6 bits/sample to 3.34 bits/sample for the worst case (crowd picture) with little decrease in fidelity. The results were obtained by computer simulation.

(iii) A closely related scheme is given by Julesz (Ref: Julesz, "A Method of Coding TV Signals Based On Edge Detection", BSTJ, July, 1959). Here, the edge or corner points are detected by first quantizing the first difference of the signal (differenced over the sampling interval) then differencing again. The output of the second differencing operation will be 1 if a corner point is present and zero otherwise. Note that the "error threshold" of Youngblood's scheme is incorporated in the quantizing levels set for the first difference as shown in the diagram below, Fig 4.3 (from Julesz).



Corner Points:
(transmitted)



$\Delta_{i-1} \neq \Delta_i$

Non-transmitted Points:

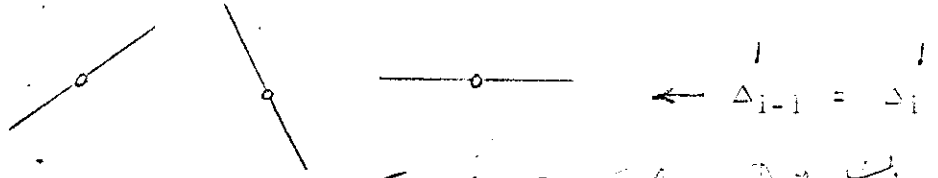


Fig 4.3B: Edge Detection

The larger the quantizing levels, the fewer the samples sent out, of course, the fidelity is greater. Extra points were sent in the vicinity of sharp intensity transitions to assure correct positioning of those transitions in the reproduced picture, and also extra points were sent to eliminate the tunneling effect shown below.

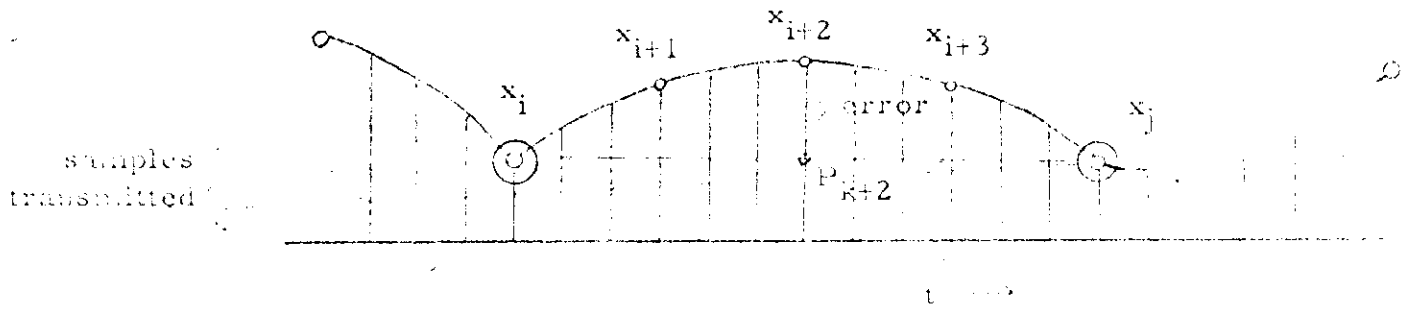


Fig 4.3C Error Tunneling Effect,

This effect occurs because the difference is taken over only one sampling interval. Thus, in areas of slowly changing intensity, the difference between successive samples may not be great enough to exceed the quantizing level and a cumulative error can build between transmitted corner points (Note that this problem does not afflict Youngblood's scheme since the same origin for the test line is kept until the error (cumulative or otherwise) is exceeded).

The extra points to be sent are determined by setting a threshold on the error between the interpolated point (c. g. P_{k+2} ^{in Fig 4.3C}) and the actual data. The number of extra points which have to be sent is extremely small. The results obtained again by computer simulation are a reduction from 3.15 bits per sample to 2.99 bits per sample for the worst case investigated (i. e. wearing load sportshirt!) using Huffman coding for the distance between samples. Storage requirements are considerable (about 500 samples for a fairly smooth scene), indicating that the average rate (i. e., the rate at which we want to transmit the samples) is much lower than the maximum rate.

Comments: Again, queuing techniques might be used to advantage in meeting buffer storage requirements. The Youngblood scheme seems to have an advantage since it eliminates tunneling effect as well as transition position ambiguity (usually) although it is relatively difficult to instrument. An extension of the above scheme would be to set thresholds on the second and higher order back differences in an attempt to avoid tunneling and transition position ambiguity. This might be simpler to instrument than Julesz's scheme.

SYSTEM 710 - PREDICTIVE CODING

~~Correlation Techniques - Prediction~~

(Refs: Oliver, "Efficient Coding," BSTJ, July 1952; Harrison, "Experiments with Linear Prediction in TV", BSTJ, July 1952; Graham, "Communication Theory Applied to Television Coding", Asta-Electronics, v. 2, 1957 - 58, p. 333 - 343; Graham, "Predictive Quantizing of TV Signals," 1958 Wescon Conv. Record, pt. 4, 147 - 152)

Since information theory (and common sense) tell us that, in order to use a communication channel most efficiently, we should send message samples which are uncorrelated, we should look toward techniques which reduce this correlation. Run-length coding ^{black and white runs} does not solve the problem of correlation between runs - the chief problem in continuous

7

7

gray level images (where the runs are all short and nearly equally likely). Prediction - subtraction is one technique which has been proposed to "decorrelate" the source. A diagram of a basic scheme is shown ~~below~~ in Fig 4.4A:

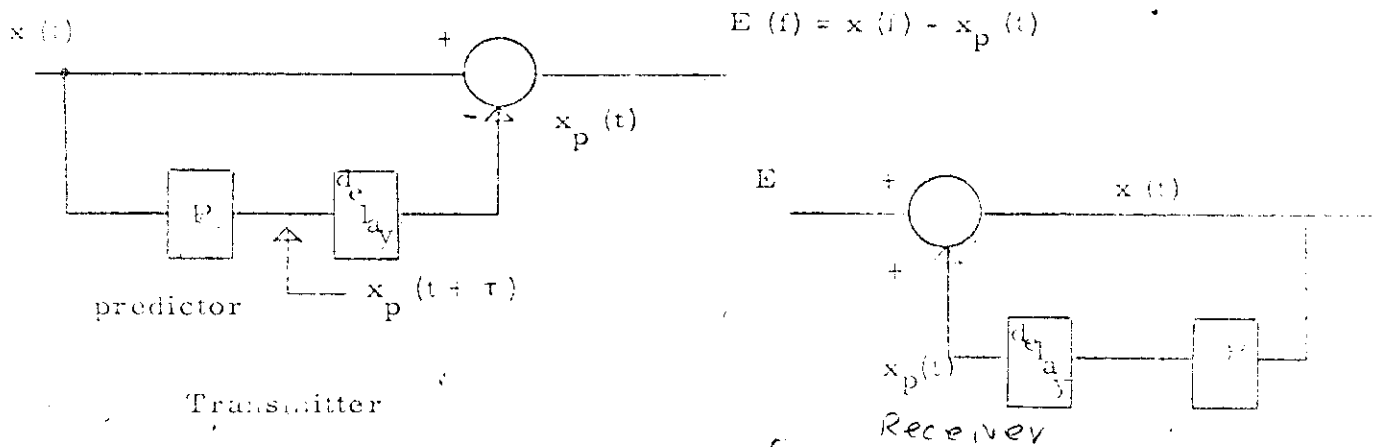


Fig 4.4A Predictive Coding

The predictor is presumed to have memory if desired - it computes the next value of $x(t)$ based on appropriately weighted combinations of the previous samples. If the next value $x_p(t + \tau)$ is computed as the weighted sum of previous values, the term "linear-prediction" is used. The samples of the "error" signal $E(t)$ are less correlated than the samples of $x(t)$, i.e. there is less redundancy, due to intersample correlation. Further the power required to transmit the signal is reduced since

$$\overline{E^2} < \overline{x^2}$$

Thus a bandwidth saving for the same signal-to-noise ratio can be achieved.

~~Stat 4-9, 3~~
~~-1.0~~

Another way of looking at the action of predictive-subtractive decorrelation is to consider the conditional distributions

$$P(j / i)$$

where j is the present sample and i is the previous sample (or previous samples). The fact that the conditional distribution is peaked means that considerable correlation is present (Recall that

$$I(J; I) = \sum_i \sum_j P(i, j) \log \frac{P(i, j)}{P(j)}$$

is a measure of the redundancy. If $I(J; I)$ is large as for peaked conditional distributions, there is a considerable amount of information about j provided by i . The decorrelation procedure makes the conditional distribution of the output $P(k/l)$ less dependent on l (less peaked with l) and yields a peaked simple distribution $P(k)$ for the amplitude of the output where the mean and mode occur near zero. Thus, the signal power is reduced virtually to that of the deviation component of the error signal. The peaked error amplitude distribution can then be coded by a Huffman code for optimum transmission on the Gaussian channel.

One basic problem with this type of coding is the consequence of errors (e. g. impulse) on the ability to reconstruct the original waveform. For example, if the error signal were reduced to zero, the reproduced picture line would appear split and dc level shifted everywhere below the

~~5-2~~ 4-9.4
5-2

error. This could have serious consequences on the reproduced picture quality unless the peak error signal amplitude was an extremely small fraction of the original signal amplitude (in general, the maximum peak to peak deviation of the error signal can be twice the original signal since the error can be either positive or negative from zero - only the mean square value is reduced). A possible solution might be to inject error correction in the form of dc level set information (this could be an extremely low-frequency signal requiring very narrow bandwidth - e.g. A. M. sine wave or shallow FM sine wave).

Note that the peak power required by the transmitter is at least the same as that required to transmit the original signal since the error amplitude can be as much as twice the signal amplitude. The effect of the Huffman coding is to alleviate this condition by recoding the quantized error amplitude such that the (hopefully) infrequent large peak amplitudes are transmitted by a larger number of bits than the smaller more frequent amplitude. We might also consider analog coding of the error signal - dc level set signal at this point.

A third problem results from the fact that in all systems considered in the literature, emphasis was placed on power reduction.

i.e., if we let

$$\rho = \frac{E^2}{x^2}$$

~~9-4.5~~ 9-4.5
6-2

the problem is to minimize ρ (or E^2). However the minimum ρ does not necessarily correspond to minimizing I (J/I) or maximizing the redundancy removal.

In fact, one problem might be stated as follows: to choose some transformation [T] which minimizes the width of the autocorrelation function.

Graham feels that the quantizing scheme is as significant in bandwidth reduction as the coding transformations. Thus, he combines a scheme for quantizing the information in a suitable manner (to be discussed later) with a modification of predictive-subtractive coding.

4-5.1

SYSTEM #5 : BAND SPLITTING

4. Band Splitting

In this general classification, schemes are considered where the video signal is split into two waveforms - one representing the low frequency information and the other representing the high frequency or "edge" information. With the signal split, each portion can be processed most efficiently by taking account of its own peculiarities. There are two processing means which have been extensively investigated:

(a) Edge Detection

(b) Special Quantizing Techniques

(a) Edge Detection

Ref: Schreiber and Knapp, "TV Bandwidth Reduction by Digital Coding", IRE Convention Record (Information Theory), 1958; Schreiber, Knapp and Kay, "Synthetic Highs - An Experimental TV Bandwidth

Reduction System;" Presented at the 84th Convention of the SMPTE, Detroit, Michigan, October 23, 1958). In this scheme, the low frequency analog information is PCM^d in the ordinary fashion. The high frequency information represents sharp contrast changes or edge information. The position and amplitude of these edges must be transmitted with extreme precision, otherwise serious degradation of the picture will result. The high frequency information is differentiated (or differenced over one sample interval) creating a series of spikes in time representing the amplitude and position of the edges. The spike amplitude is quantized into 7 levels - 3 positive, 3 negative and one zero. The seven levels are represented by a four digit number. The samples between edges represent a "run" of almost constant intensity level. This run length is coded by a 5 digit number (provision for runs longer than 32 is made by adding extra bits to the nominally 5-digit number as required - thus this scheme is a crude form of modified Huffman coding). Each edge is thus represented by a 9-digit number. Since the edge numbers are generated at a varying rate, they are buffered and transmitted at a rate near the average number of edges per second (about $.9 \times 10^6$ per second). The pcm^d low frequency information is transmitted (128 level quantization) at 4×10^6 bauds and the "highs" at 8.1×10^6 bauds for a total of 12.1×10^6 bauds compared with 56×10^6 bauds required for ordinary (7-bit) PCM.

The receiver decodes the edge position by counting and converts

~~4A, 24~~ 4-5.3
- - -

the 4 digit edge amplitude number to a synthetic analog "highs" signal which, when added to the analog converted "lows" signal forms the original video signal.

Discussion: The scheme is predominantly digital and does not lend itself to analog comparison (the authors make a rough attempt to compare the scheme on an analog basis using a signal which requires an 8-level threshold detector - lotsa luck Charlie!) However the idea of band-splitting and separate processing for each band is a sound one, and perhaps we should look at certain pulse signal classes which might convey the "highs" information satisfactorily in an analog fashion - e. g., by slopes, derivatives at certain points, moments, pulse widths, etc. It may be asking a great deal of a pulse to convey, by its shape, 8 or more bits of information (i. e. edge height and position) which may be readily and cheaply detected, but the rewards would be significant.

Another interesting point is that of run-length-coding of continuous intensity picture information. As the authors imply, we may consider the successive signal samples as a sequence of 7-digit numbers, for example. We now count, not the run length of identical numbers, but the run length of various order digits in the numbers, i. e., the run of the 7 digit, etc. It should be noted that this concept could be extended to runs of patterns of digits, e. g. the 2^7 , 2^6 , 2^5 digits taken together. We might be able to achieve tremendous source entropy reductions in this

nanner, using a modified Huffman code on the pattern runs (provided enough cheap storage were available at both the sending and receiving ends, we might be able to extend the concept to two dimensions and code pattern area sizes).

One disadvantage of the scheme is, of course, the storage requirements. Here, as with previous schemes, queuing techniques may prove useful (the authors used electrostatic storage techniques).

Another disadvantage of their scheme was the use of the same fixed-length (5 digit) code to specify edge position. It would seem, in view of Kretzmer and Laemmel's work on run-length coding that more optimum codes requiring only slightly more hardware could have been used.

(b) Special Quantizing Techniques

(Ref: Kretzmer, "Reduced-Alphabet Representation of TV Signals", IRE Conv. Record, 1956, Pt. 4). We now consider a system which includes the information sink (human) in the system design. Picture fidelity is really a subjective quantity and two characteristics of the human sink have been exploited:

- (1) The eye is very sensitive to small errors resulting from amplitude quantization of almost constant intensity areas, (e.g. in a nearly constant intensity area, the eye is very sensitive to coarse quantization of the intensity).

(2) The eye tolerates large errors in intensity when large (sudden) intensity changes occur - this is equivalent to turning up (or down) the "contrast" switch on your TV set - the eye is not too bothered by the magnitude of the setting - within limits). Thus Kretzmer divides the video signal into several bands - quantizing the low frequency bands more finely (corresponding to constant areas of illumination and the high frequency bands more coarsely. The table below (from Kretzmer) summarizes the scheme (Fig 4.5).

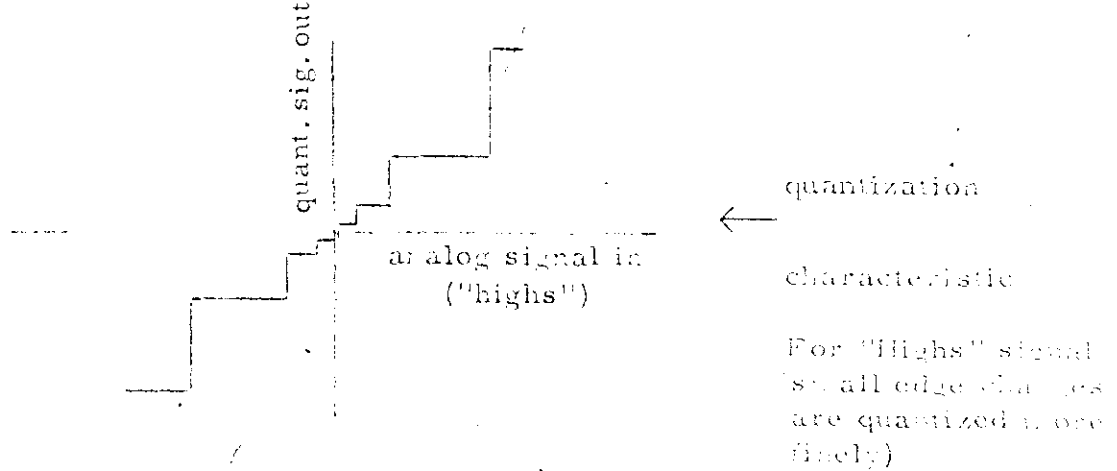
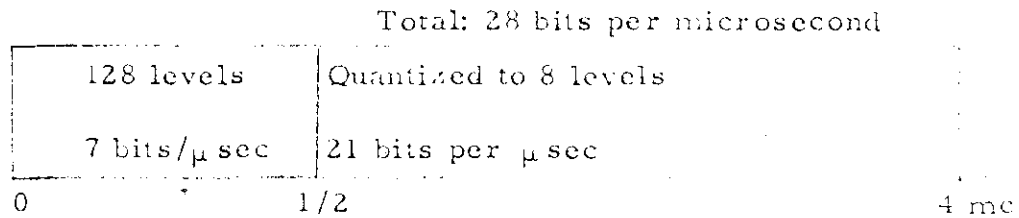
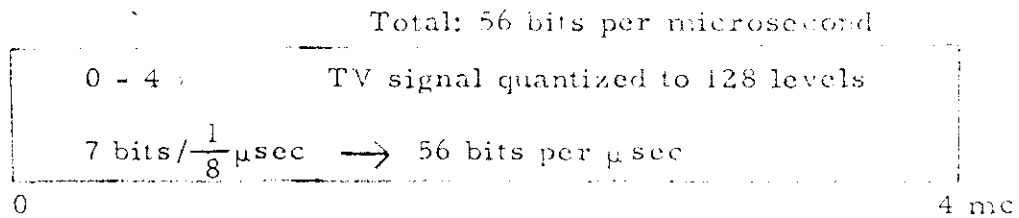


Fig 4.5 Band Splitting

Thus, by quantizing each portion of the signal according to its own need, the source information rate is reduced by a factor of two. Note also that the tapered quantizing characteristic gives more weight or importance to small intensity level changes and less to large ones. Thus, the psychophysical characteristics of vision enter into the system design in two ways: band-splitting and tapered quantization for the high frequency information.

Discussion: Although the results of using this technique do display some shortcomings (spurious contours or ghost images appear as a result of quantizing some of the high-frequency transients) the technique yields a 2:1 bandwidth reduction over ordinary PCM using remarkably little hardware (filter, delay line, subtractor - no memory). Thus, any scheme which we incorporate should certainly include the human observer as a sink if remote picture printing is the ultimate objective - certainly the remote printer itself may be the ultimate limiting factor as an information sink and we should design our system so that it transmits the information with no greater fidelity than the printer can reproduce. If however, the information sink is a computer, these special quantization techniques may not be applicable, since large cumulative errors in computer processing of pictures should be avoided.

Analog Compression Coding

(Ref: Gouriet, "Bandwidth Compression of a TV Signal"; Proc. IRE, P & B, 1957) This scheme is really the analog equivalent of some of the edge amplitude-position coding schemes discussed earlier. Briefly, the analog intensity information is first sampled, then quantized (although the author doesn't mention it, the quantization could be in fewer unequal steps to allow for the psycho-physical characteristics of the eye). The quantized sample values are held over to the next sample forming a step waveform (see figures 1, 2 (from Gouriet) ^{labeled 4.6-1 and 4.6-2 in this sec} ~~at end of this section~~). The step waveform is next differentiated, rectified and clipped forming the sequence of equal-height pulses in figure 4 ^(4.6-4). These in turn gate a saw-tooth generator giving the output waveform of figure 5 ^(4.6-5). The peak amplitude of this waveform at the sample points of figure 4 is proportional to the distance between samples. Thus, sampling of this waveform at this time yield a string of pulses (figure 6 ^{or 4.6-6}) whose amplitude is proportional to the distance between samples. The two pulse trains giving sample amplitude and position (figures 6 and 7 ^{or figs 4.6-6 and 4.6-7}) are now time compressed by buffering to yield a string of pulses occurring at equal-spaced intervals - the interval being the average rate of pulses per second for the field (figures 8 and 9 ^{or figs 4.6-8 and 4.6-9}). These two pulse trains may then be low-pass filtered and sent over separate channels, or interleaved, low-pass filtered and transmitted over twice the bandwidth of the two channel sample. The decoder then product-samples the received waveform(s) and derives

4.6-9 4-5.2
 (5)

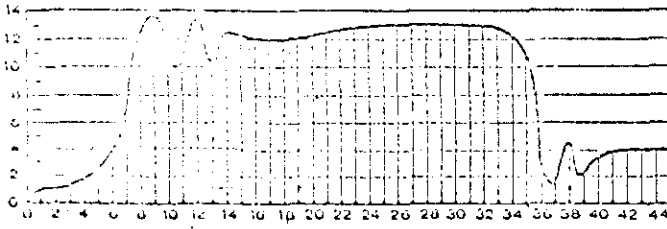


FIGURE 4.6-1 4.6-1

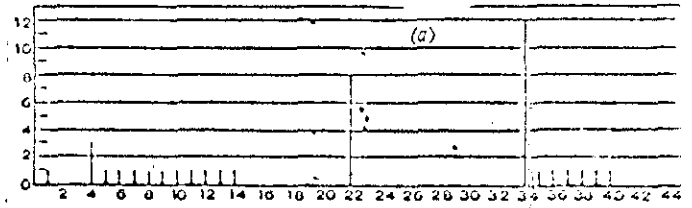


FIGURE 4.6-6

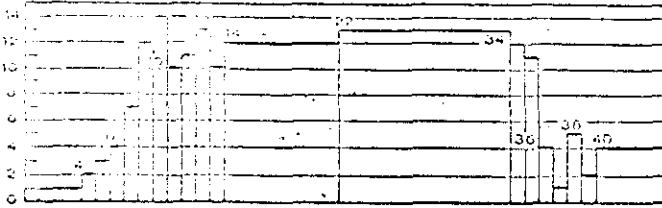


FIGURE 4.6-2

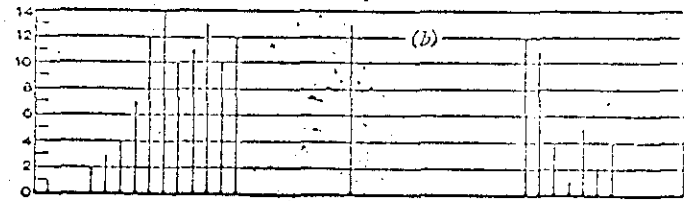


FIGURE 4.6-7

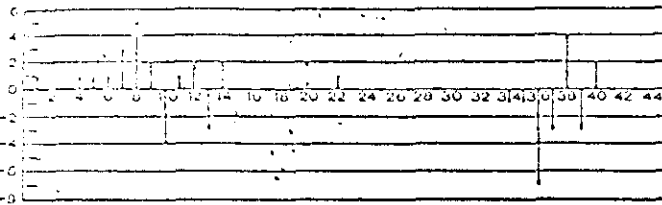


FIGURE 4.6-3

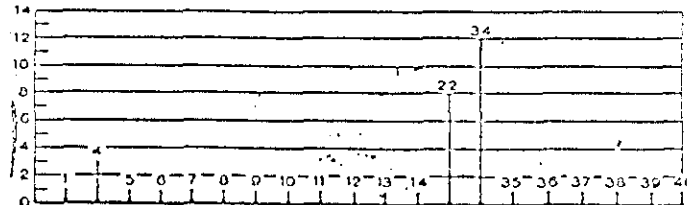


FIGURE 4.6-8

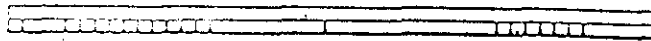


FIGURE 4.6-4

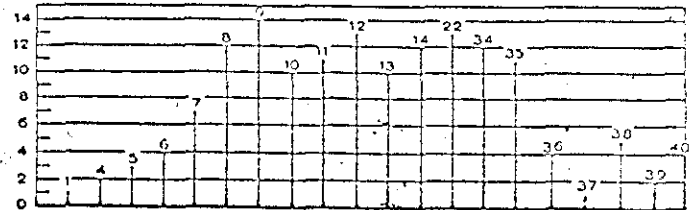


FIGURE 4.6-9

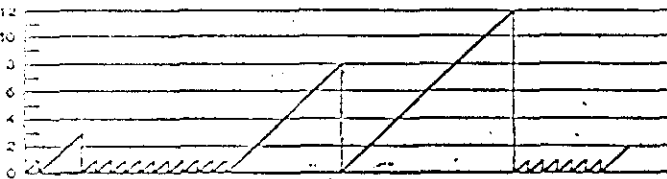


FIGURE 4.6-5

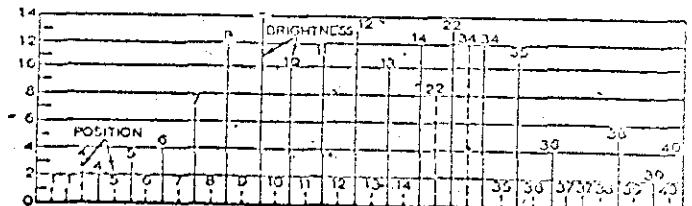


FIGURE 4.6-10

the amplitude-position information. If the velocity of the printing mechanism at the receiver is proportional to the height of the position pulse, the dots will occur in the correct relative position again. The printer would maintain the same level of intensity in the region between samples (with some memory, we could interpolate).

Unfortunately, the bandwidth reduction is achieved at the expense of a higher required signal-to-noise ratio, although the proportion is not one to one so that some signal compression is achieved (note that ideally, we would like a bandwidth compression with no increase in SNR so that bandwidth compression corresponds to a smaller required channel capacity in direct proportion:

$$C \approx W (SNR)^{DB/2}$$

The reason for the increase in required SNR stems from the stringent requirements on the accuracy of the position information. If we require an rms error of one sample duration with negligible cumulative error, we would need a 72 db SNR for 6:1 bandwidth compression (based on 100 samples per line, 100 transmitted samples per line, i.e. 2 samples per transition). This stringent requirement on SNR can be reduced at the expense of bandwidth or fidelity, but is probably the limiting factor in the scheme. CONCLUSION: Another reason for the high SNR requirement is that the positional information has not been optimally coded. If we could code the more probable inter-sample durations with a smaller voltage

values, a lower average-power "position" signal should result. This would then reduce the SNR requirements for the same B.W. compression. Note that the scheme provides a good example of "analog coding". We see that it may well be possible to code information with analog waveforms in a statistical manner similar to the discrete case. Question: Does analog coding really amount to r -ary (discrete) coding where one r -ary digit is used per sample and $r \rightarrow \infty$ (i.e. we are able to code an infinite number of quantum levels as in the continuous case). This viewpoint poses interesting possibilities: could we not derive a new analog waveform which is a coded version of the original by considering two or more r -ary digits per sample? This would reduce the required number of quantum levels per sample since we double-up or triple-up on our coding, i.e., use more than 1 r -ary digit per sample. The analog waveform derived from the double-sampled function (e.g. by low-pass filtering) would then require twice the bandwidth of the original (Nyquist rate sampled) function, but a good deal less than the PCM^d waveform. This might provide a compromise between requiring a large bandwidth but high-fidelity PCM transmission and smaller bandwidth, lower fidelity analog transmission. The number of quantum levels demanded of the detector is reduced from the single r -ary digit per sample case.

Another improvement on the Gouriet scheme would be to send not square-wave data, but ramp-type data (note that ramps are a more

efficient approximation to any waveform than square waves. This would cut down the number of samples required to specify a transition from two to one. The ramp could be specified by two numbers: its angle and length (or the horizontal and vertical components). Note that this scheme is merely the extension of previously discussed schemes to the analog case, and the disadvantages associated with those digital schemes would also apply to this scheme.

SYSTEM #7: DEVIATION AND RATE CODING

Description: The waveform of photographic images is approximated by two parameters, the deviation and the rate. The deviation and rate are the d. c. term and the most significant term of a Fourier series expansion of the signal over a short period.

Source: This system was proposed by R. M. Bennett after a study of some sample photographic images indicated. There are large sections of photographs which can be approximated this way.

Proposed Procedure: An integrating circuit would repeatedly integrate and read out the values of the deviation or mean value of the signal. The signal would also go to a set of filters in parallel. The filter with the largest output would determine the value of k to use, In Fig. 4-7.1 a sample scan line is shown with a linear wave above (dotted line). Although this example is taken to have a linear relationship between waveform and degree of gray, a practical device would probably be based on a square-law response to match the example in Fig. 3E1. Above Fig. 4.71, the value of (d, k) which would be transmitted as shown. In addition some timing or phase information must be transmitted. Either the time interval must be specified with a convention as to the starting phase at transitions or the starting phase

4-7.2

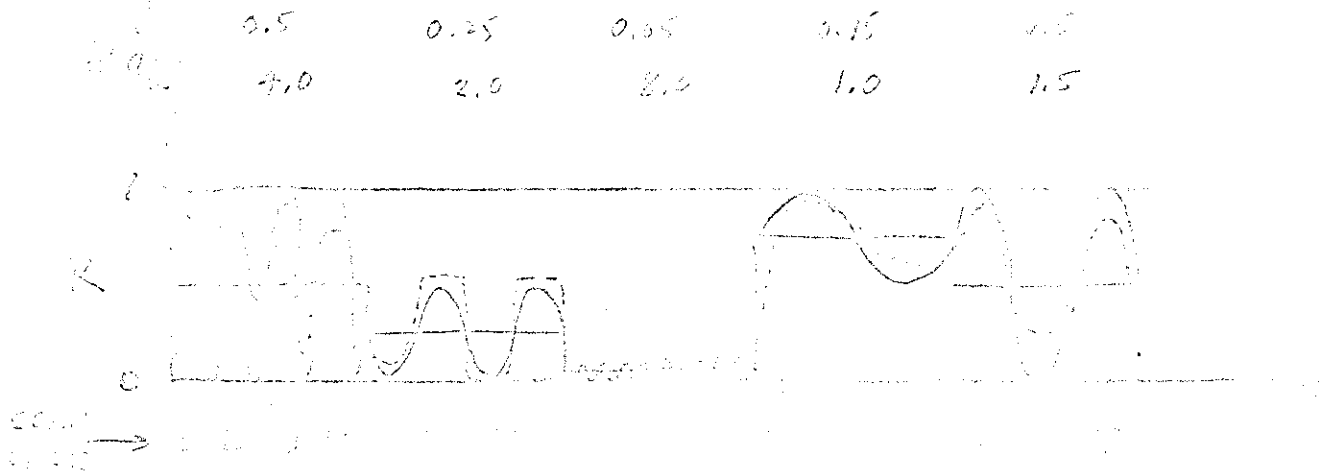


Fig 4-7-1 Approximation of waveform with $a_0 = 0.5$

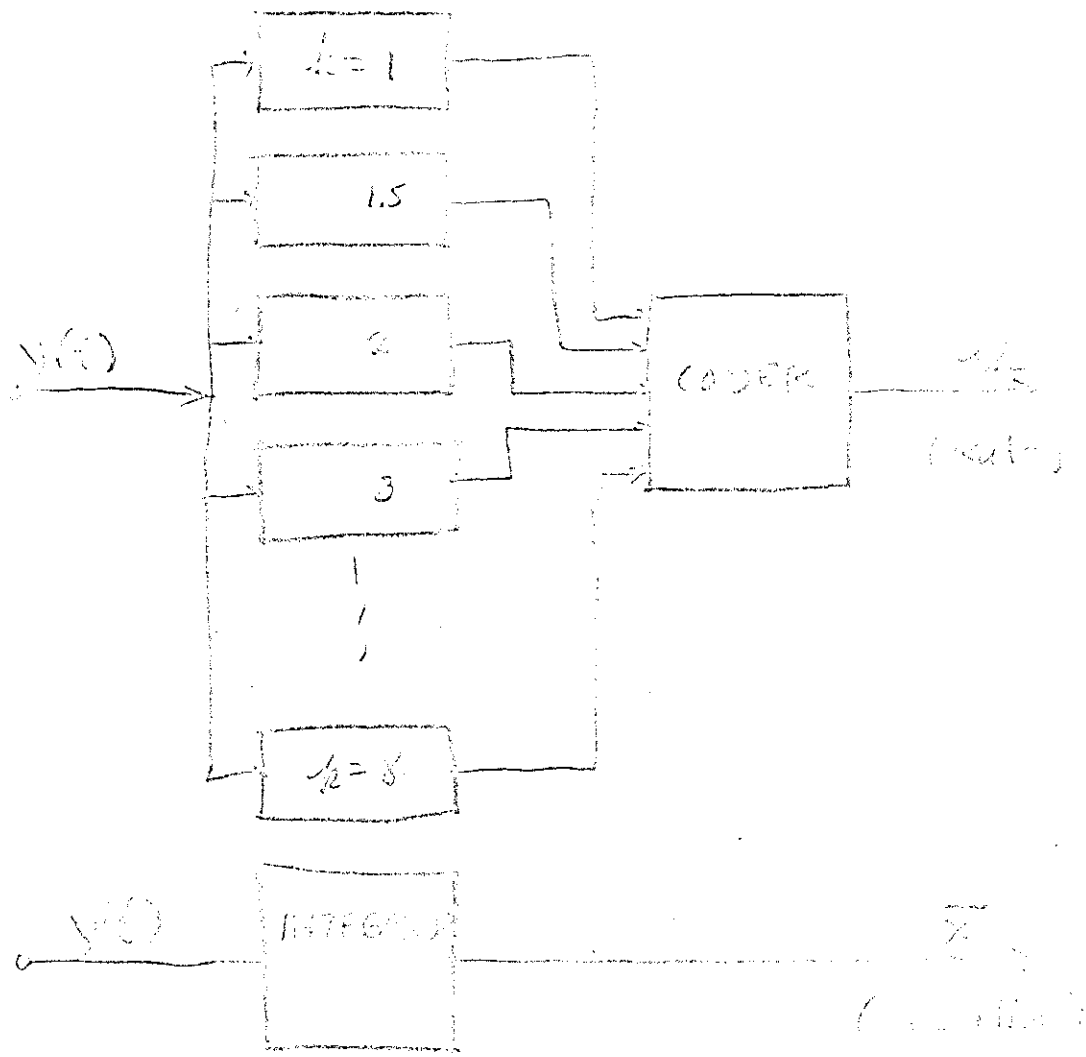


Fig 4-7-2 system diagram for Coding Rate and Deviation

must be transmitted with a uniform time interval agreed upon.

A straight-forward way to implement this system would be to set up a bank of filters as is shown in Fig. 4.7-2. In this block diagram the coder selects the filter having the largest output and then sends out the code for that frequency.

To develop a simpler implementation consider (1) what other circuits could be used in place of filters, and (2) whether the filters or other circuits can be replaced by a single swept-frequency unit like a spectrum analyser.

To get at the first question, we examine the Fourier series representation of the waveform:

$$x(t) = \sum_{j-} a_j \cos \left(\frac{2\pi j}{T_0} t + \phi_j \right)$$

which in this class of photographs can be approximated by

$$x(t) = a_0 + a_k \cos \left(\frac{2\pi k}{T_0} t + \phi_k \right)$$

Let $Y(t) = x(t) - a_0 = a_k \cos \left(\frac{2\pi k}{T_0} t + \phi_k \right)$ and define

$$\Theta(y(t)) = y^2(t) + \left(\frac{T_0}{2\pi k} \right)^2 \left(\frac{dy(t)}{dt} \right)^2 .$$

Then

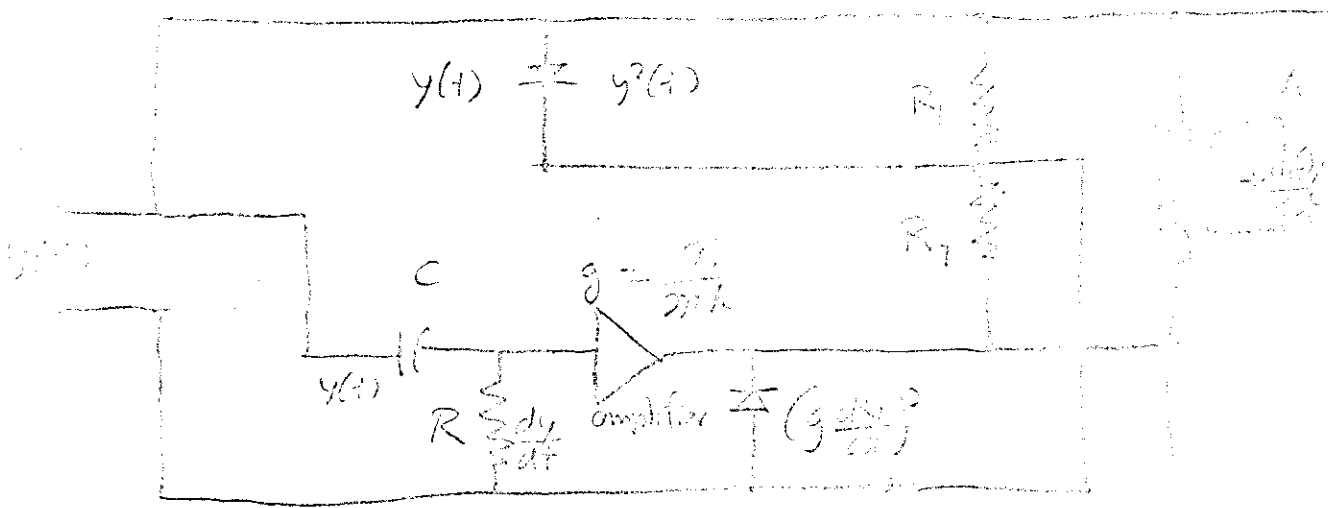
$$\Theta(y(t)) = \sum_{k=1}^{\infty} a_k^2 ,$$

since $\Theta_k = a_k^2 \left(\cos^2 \left(\frac{2\pi k}{T_0} t + \phi_k \right) + \sin^2 \left(\frac{2\pi k}{T_0} t + \phi_k \right) \right) = a_k^2 .$

The next step is to find a physical circuit producing the parameter $\hat{\theta}_k(t)$. Such a circuit is shown in Fig. 4.7-3. When these proposed frequency detector circuits are used in parallel the one with the highest output at B gives the $\hat{\theta}_k$ closest to the true waveform. The one with the minimum output (non-fluctuation) at output A would also identify the proper rate.

The next step in reduction of components would be to use one frequency-detector with the gain (g) controlled by a sawtooth wave as in a spectrum analyser.

This system had not been constructed at the time of closing of the image transmission project.



$\theta(t)$

Fig 4.7-3 Frequency Detector.

SYSTEM #8: ADAPTIVE SCANNING WITH BANDWIDTH SWITCHING

Description: *

Prescan adjusts scan speed and selects which of two frequencies f_1 and f_2 are to be used where $f_1 < f_2$. Redundancy for error detection is transmitted at f_2 when f_1 had data, and vice versa.

The prescan system could be a set of fibre optic paths connected by logic to select the scanning speed and hence the frequency.

Illustration

A possible application might be an electronic scanning system such as a flying spot scanner used on large type (18 point) printed page with illustrations occurring often, having detail equivalent to 4-point type. Suppose the large type required a bandwidth of 20 Kc/s centered at 12 kc/s and the fine detail a bandwidth of 48 kc/s centered at 32 kc/s.

For intermediate resolution the prescan could control the scanning speed to center the signal about f_1 or f_2 , putting the control signals and error detection redundancy into $f_b = f_1$ into a delay line circulator to be heterodyned with local oscillator f_p during the next f_1 period as is shown in Figs. 4.8-1 through 4.8-3. If the whole section of a page was in fine detail, the scanner would have to periodically stop to insert the control data.

*This is a system proposed by R. M. Bennett

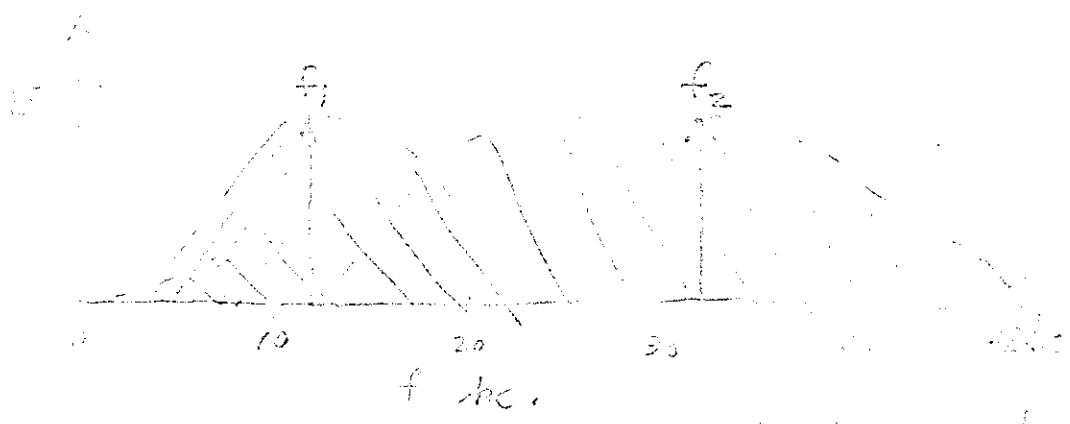


Fig 4.8-1 Two Data Distributions.

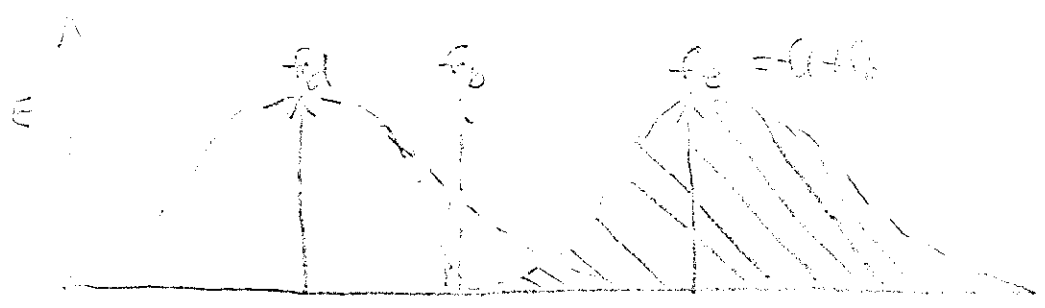


Fig 4.8-2 Frequency Detection & Center Bandwidth.

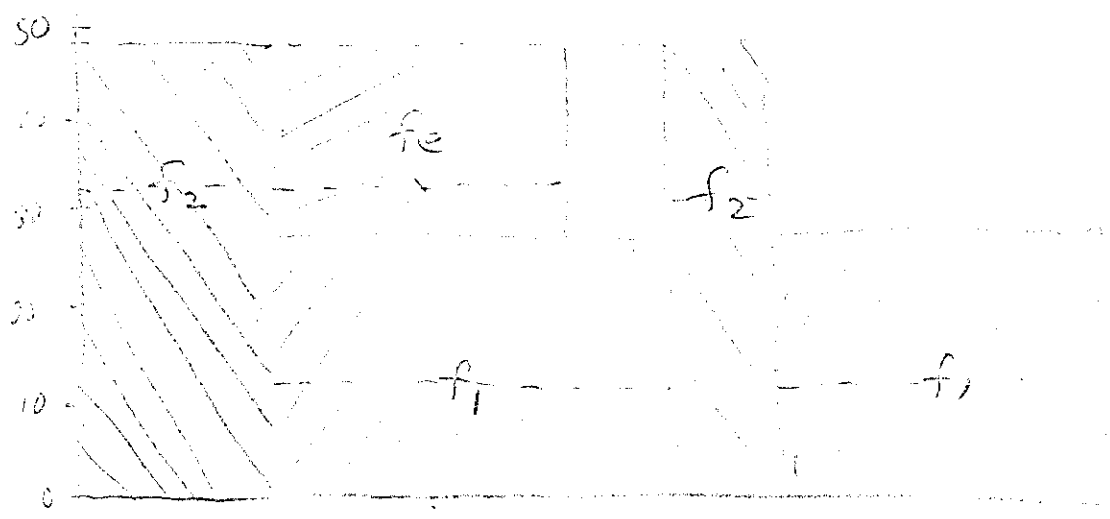


Fig 4.8-3 Bandwidth Use vs. Time

SYSTEM #9: OPTICAL PROCESSING

Description:

Two potential systems of optical processing were considered:

(a) One spatial dimension auto-correlation functions and Fourier transforms obtained by sending a light beam through an optically active crystal modulated by the function $h(t-\tau)$ or $\cos \omega t$ in the form of a voltage across the crystal which achieve multiplication by rotation of the plane of polarization.

(b) Two spatial dimension auto-correlation functions through passing of two-dimensional light beams through masks.

(c) Θ -modulation: Modulation of the image by grating sections at angle Θ which puts images at angle Θ in the Fraumtrope diffraction plane.

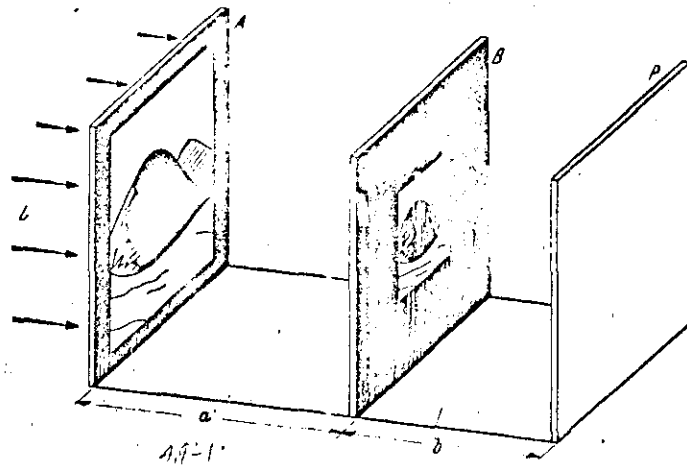
Status of Optical Processing Systems:

(a) The one spatial dimension system are believed to be covered by a series of patent applications by Mr. Wilmotte. These are essentially a substitution of optical circuits in place of analog electric networks or digital computing circuits. For example, see Kerr Electro-Optic effect.

(b) For compression coding the two spatial dimension techniques look more interesting. These are described by a paper of W. Meyer-Eppler (with G. Darius in Third London Symposium, Information Theory, London Butterworth's Sai.Pub. (1956), Two-Dimensional Photographic Auto-

Correlation of picture and alphabet letters, pp. 34 - 36).

The basic method of obtaining photographic auto-correlation diagrams is shown in Fig. 4.9-1 and a sample set of auto-correlations and cross-correlations functions from Meyer-Eppler and shown in Fig. 4.9-2.



4.9-1
Figure 1. Schematic view of photographic auto-correlator.
L, light source; A, B, transparencies (slides); P, photographic plate.

	A	B	C	D	E
A					
B					
C					
D					
E					

4.9-2
Figure 2. Cross-correlation matrix of alphabet letters.

(c) θ -Modulation

Dr. A. Lohmann* has described a method of modulating an optical image through superimposing a diffraction grating at angle θ . The spectra of different parts of the image constructed of diffraction gratings at different angles θ give Fourier transforms in the Frankfer diffraction plane at the corresponding angles θ , where $\theta = 90^\circ - \theta$. The θ -modulation technique offers potential applications to image transmission which have not been fully explored yet.

*IBM - GPD Report in preparation.

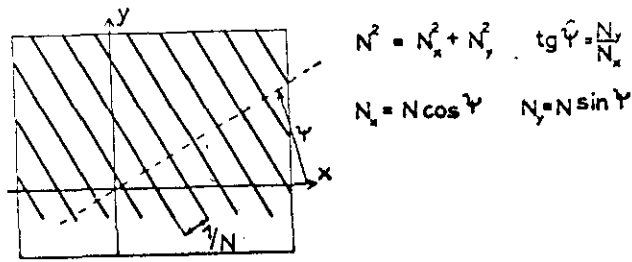


Fig. 4.9-3 Grating at Angle $\phi = 90^\circ - \Theta$

(from A. Lohmann, "Einführung in die Ubertregungstheorie der Optischen Affeldung," pp. 5 - 10, in Photographische Korrespondenz & Sonderheft (Kohn 1958) Darmstadt 1959, Verlag Dr. Othmar Helwith).

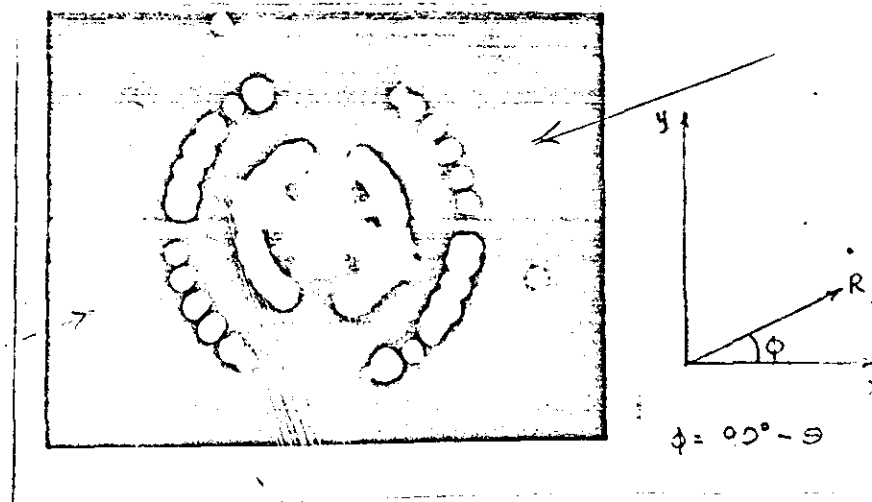


Fig. 4.9-4 Image of Θ -Modulation Object in the Fraunhofer Diffraction plane showing 1st, 2nd and 3rd orders. (photo by B. Margenstern, T Hochschunch, Braunschweig, April 1962)

Although neither of these systems were tried experimentally, they warrant further investigation. Meyer-Eppler discusses the symmetry properties of characters and their correlation function. These symmetry properties may be significant in designing special cases of image coding and decoding.

SYSTEM #10 : RUN LENGTH CODING

1. Run Length Coding

(Refs: Michel, Fleckenstein, Kretzmer, "A Coded Facsimile System", 1957 Wescon Conv. Record, p. 2; Laemmle and Brogan, "Coded Transmission of Information," ASTIA 36785; Prudhom and Faber, "Bandwidth Compression for the Transmission of Wide-Band Facsimile Signals", IBM - ASD Technical Report #17-037).

Used primarily for black and white signals (e.g. printed page and line drawings) the scheme is based on the fact that the probability distribution for a run of black dots in a scanning line is peaked. Thus, a code such as a Huffman code can be used to reduce the average number of bits required to specify a run. Thus, the average number of bits required to specify a page is reduced. This results in either a bandwidth or transmission time saving. In practice a modified Huffman code is used to compress the source in order to minimize equipment requirements.

An extension of run length coding is suggested by Oliver ("Efficient Coding", BSTJ, July 1952) where any commonly occurring sequence of symbols may be considered a "run" and treated by the methods outlined above.

Run length coding, although quite efficient for printed page sources, is a poor method for picture processing because it does not reduce inter-sample correlation (however see section 4(a) for modification).

Obviously, then, other methods of source encoding which reduce inter-sample redundancy must be investigated. However run length coding may be expected to yield good results where large solid-color areas or patterns appear quite frequently, e.g., line drawings. Note that run-length coding is a first-order approach to optimum coding for a Markov source since we are coding certain sequences of source symbols according to their probability of occurrence. However true Markov source coding would involve coding sequences of runs as well as the runs themselves. Wyle, et al, (Wyle, Erb, and Barrow, Reduced-Time Facsimile Transmission by Digital Coding, Transactions P. G. C. S., September 1961) have found that source compression factors of 3 - 4 are feasible using run-length coding. However the buffering requirements are quite high since the information is to be transmitted at the average rates.

Typical Compression Ratios:	Typing	4.3	(1)
	Line Drawings (Art)	25	(2)
	(Simple Lines)	115	(3)

- 1. End Instrument Co. IRE Trans. CS-9, No. 3, Sept 61 p. 215
- 2. BTL WESCON
- 3. IBM ~~77-037~~ Report 17-037

The above compression ratios are based on large buffers (up to 1600 bit storage). An improvement to bring the practical run down for run length coding would be a system requiring less buffer storage.

Fig 4.10-1 indicates the general nature of the counter and storage required.

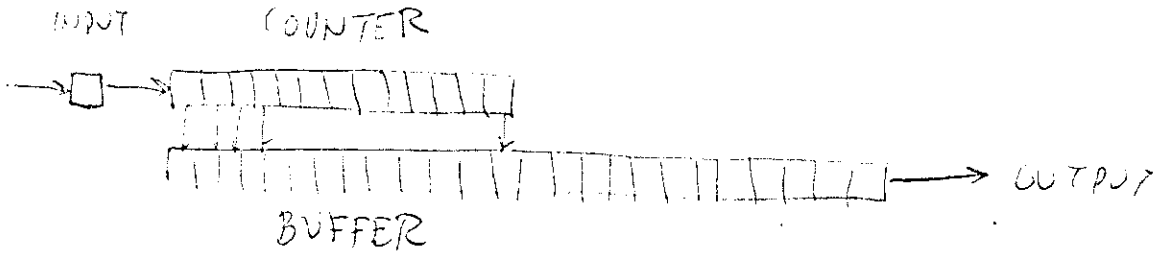
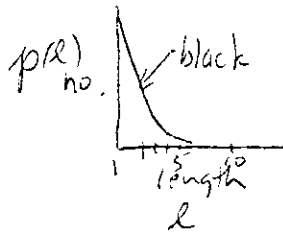
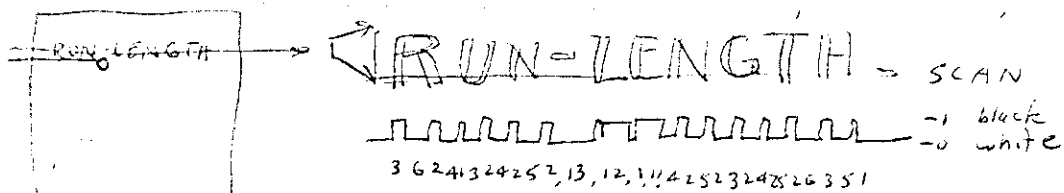
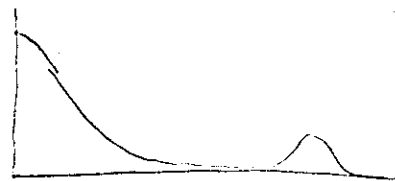


Fig 4.10-1 Counter & Buffer for Run-length Coding



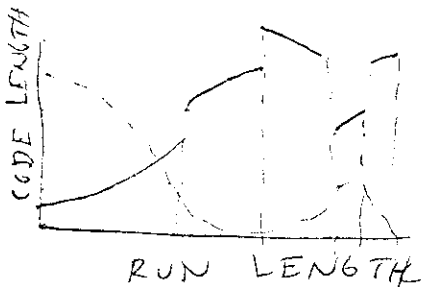
Run-length distribution separated for Black & white



Combined, Black & white

RECODE
SO THAT

$r(l)$



$\sum r(l) \cdot p(l)$ is minimum.

Fig 4.10-2 Run length coding stages

The group in the country and reading of
numbers being an illustration in Fig 4.10-2.
The statistics for several sources data on numbers
are plotted in Figs 4.10-3, and 4.10-4.

Possible implementation of run length
by delay-line or by cone are discussed in the
notes accompanying Figures 4.10-5 through 4.10-7.

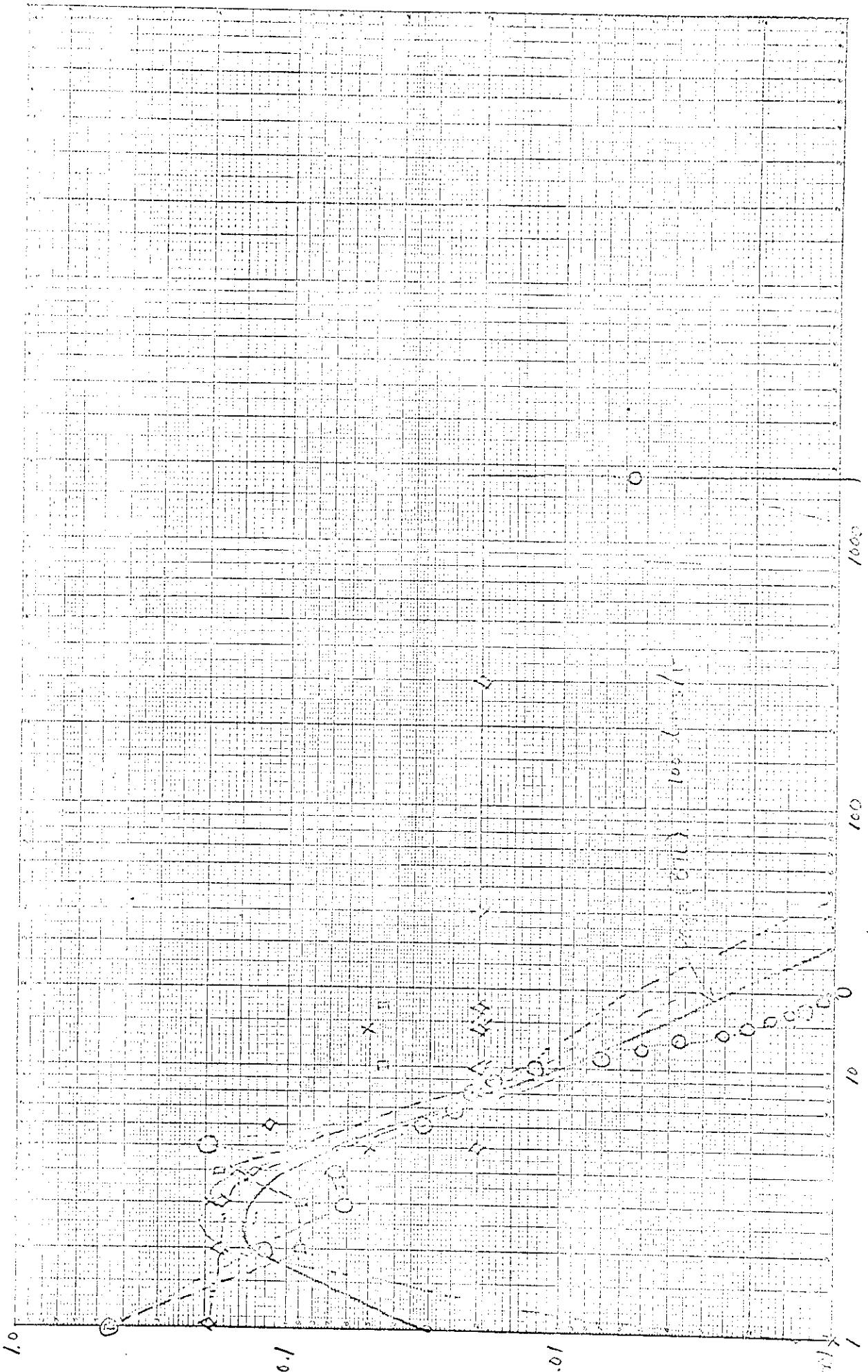
7

7

7

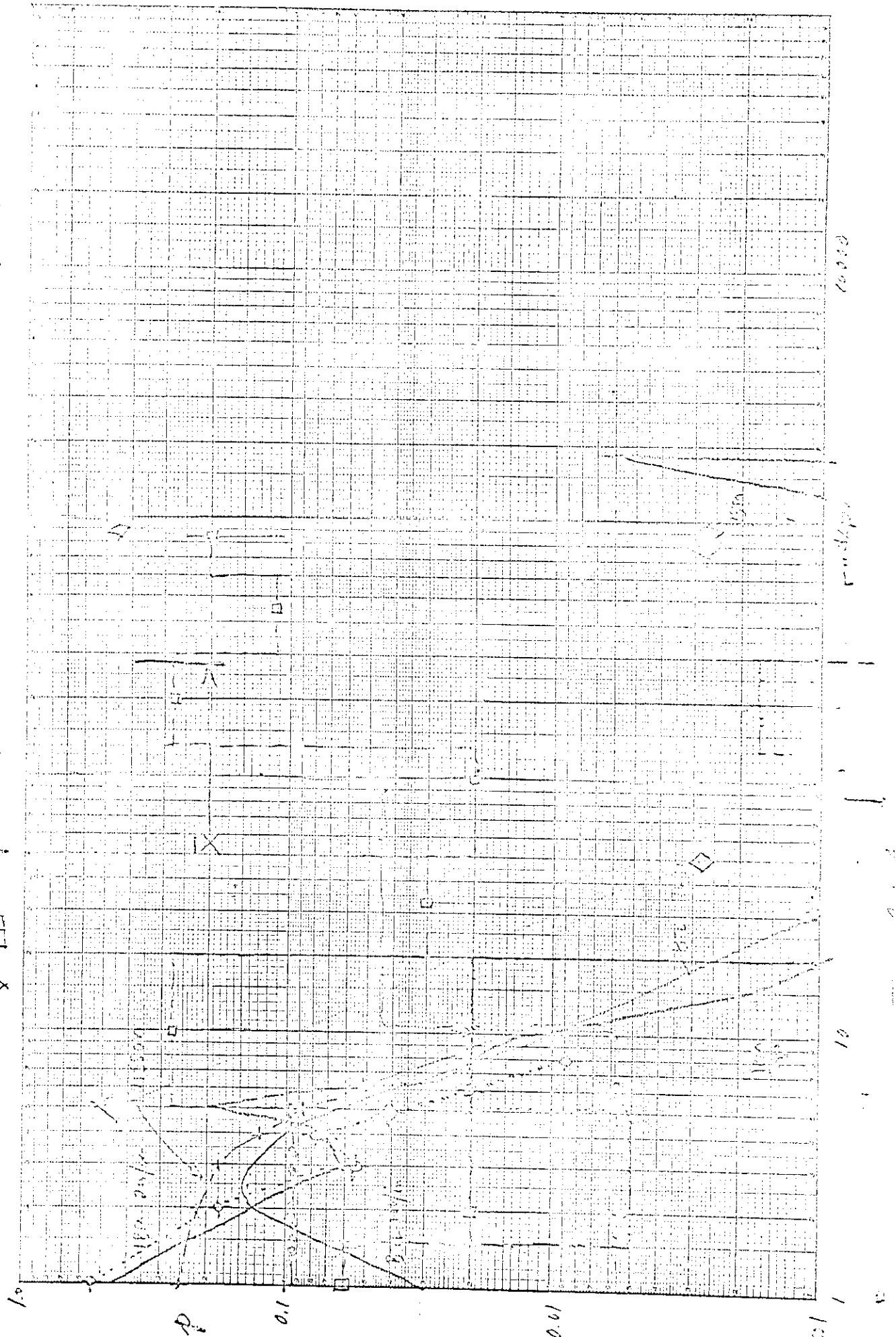
7

O 13M 17-027 200 km/yr
 — 500 100 km/yr
 D Line A 75 bits
 X Line B 75 bits
 ◊ 3 Lines A,B,C 225 bits



Run length
4.10.3

17-087 p.17 100/100 + Poisson Approx. (Fig.13)
 p. 4-9 1/2 Tolgryuk (Fig.13)



5-A

Description: RUN-LENGTH BUFFER & CODER

① RUN-LENGTHS ARE CONVERTED TO STORE COUNTS, WITH THE COUNT OCCURRING AT THE END OF THE RUN-LENGTH. THEREFORE THE TIME POSITION OF THE COUNT IS KNOWN WHEN THE VALUE OF THE COUNT IS KNOWN, E.G. THE WHITE COUNT 12 OCCURS 7-12 STORES TIMES LATER THAN PREVIOUS BLACK COUNT 2 IN THE EXAMPLE BELOW.

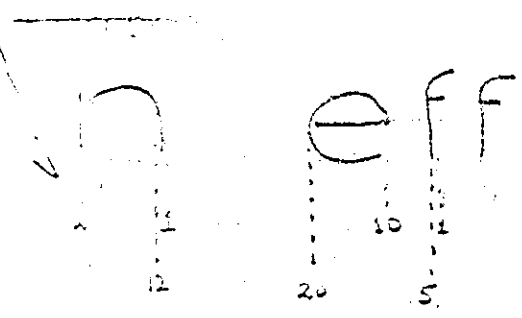


Fig 4.10-5 Sample Scan

② A MULTIPLE INPUT DELAY (LINE) BUFFER CAN "EQUALISE" THE COUNT POSITIONS BY ALLOWING THE LOW COUNTS TO UNDERGO THE LARGEST DELAY, THE HIGH COUNTS THE LEAST DELAY. THE OUTPUT OF THE BUFFER CONSISTS OF COUNTS WHICH ARE REDISTRIBUTED INTO AN EVEN, PERIODIC FORM.

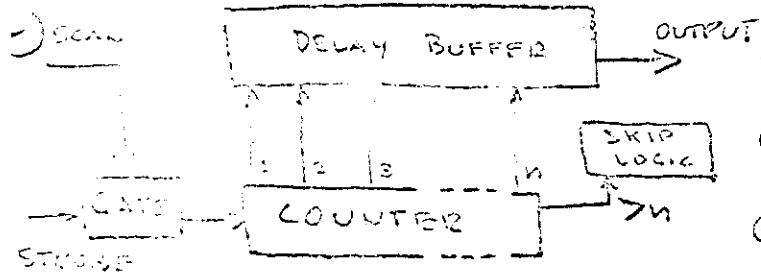


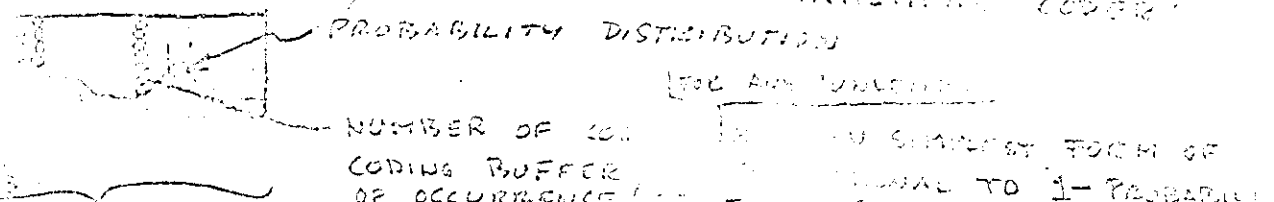
Fig 4.10-6 Logic

③ IMPLEMENTATION: CORES
④ THE DECODING (DE-BUFFERING) AT THE RECEIVING END WOULD BE THE INVERSE.

COMMENTS:

IN ORDER TO LIMIT THE NUMBER OF INPUTS (M) TO, E.G. 32, THE DOCUMENT IS SECTIONED INTO 32-STORE-WIDTH COLUMNS. THIS ALTERS THE RUN-LENGTH PROBABILITY DISTRIBUTION (PROB = 0 FOR M > 32) SINCE NO VALID RUN-LENGTH EXISTS FOR OVER 32 STORES LENGTHS. A >32 INDICATION MAY BE USED AS A FLAG.

THE BUFFER MAY DOUBLE AS STATISTICAL CODER



SYSTEM #11 Adaptive Run-Length Coding

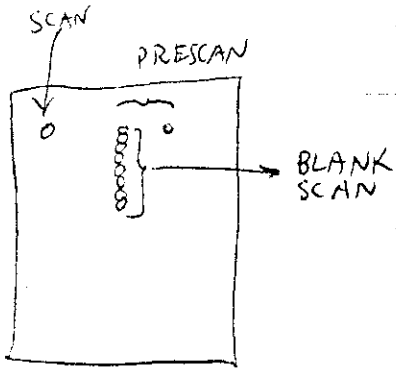


Fig 4.11-1 Scan & Pre Scan

Multiple Scanning Head: vertical group sensors in the black scan allows for the skipping of groups of blank lines.

The two horizontal pre-scan sensors measure the signals for some point of the autocorrelation generation for

controlling the speed of scan to fit the rate to optimize the rate of scan to match the coding system.

A sample graphical illustration of the compression theoretically available from this system is shown by item F - I in Fig. 2.7.

4-12.1

SYSTEM #12: BUFFER-CONTROLLED SCANNING FOR RUN LENGTH CODING

In order to utilize more fully the capability of a particular channel for image transmission, it has been proposed that redundancy be removed from facsimile scanning by encoding run lengths of black and white (RLC). By the very nature of this type of compression, there is considerable mismatch between the fixed, channel information rate and the variable, encoder-output rate. Although this problem can be minimized by designing a special purpose machine to process only a particular kind of document (e.g. typed letters or engineering drawings or weather maps), a general purpose machine to process all kinds of documents would be desirable. Several approaches to this problem have been proposed; the one to be described here is a compromise for optimum cost and performance.

First however, the other methods, with their disadvantages, should be discussed. Of course, the most obvious solution is to provide a large, expensive buffer to allow for the greatest possible discrepancies in information rates. The opposite extreme buffers only one run length and by repeated, high speed scanning, advances to the following run length only after the buffer has been unloaded by the channel. This scheme requires very complex, as well as high speed, logic. The difference in rates could be provided for by starting and stopping the scanning; however, the inertia of even a flying spot is too great for high speed scanning. Merely stopping the feed (vertical scan) for a full buffer and allowing the channel to "run dry" for an empty buffer is both wasteful and complex to resynchronize. A system which velocity-serve scans on instantaneous picture-detail is quite complex from an overall scanner-to-display viewpoint.

An optimum system, then, is one which will maintain a constant channel rate and which will scan any type of document. The system proposed here utilizes a moderately sized buffer, the instantaneous loading of which can be monitored for controlling the rate of scanning. By rigidly coupling the feed to the scan, the system does not require resynchronization. Clocking for the encoding process can be derived from the variable speed scanner. Of course, a similar system can be used at the decoding-display terminal. Because the buffer can complement servo response requirement, no extreme burden is placed in either area; hence, each should be well within its state of the art.

From notes of S. F. Gornelle (8/3/62)

By the proper choice of scanning mechanism and multiplexing, the concept of buffer-controlled scanning can be applied to a wide range of channel frequencies. For high frequency channels, the more or less conventional flying spot scanner can be employed. However, because the low speed, conventional facsimile scanner is designed for synchronous operation, it is usually of high mass and, hence, of low frequency response.

Therefore, a novel type of scanner has been conceived whereby, a document is wrapped around a stationary, transparent tube and scanned by two canted mirrors which rotate on-axis. The optical signal is reflected to a photoelectric transducer which is also on-axis but not rotating. A light source is focussed on the document in a similar manner. The entire assembly inside the tube is geared to the scan driver to feed along the document (See ^{T.S. 7.13-1} attached figure). A feature of this design is that the scanned document, usually of irregular size and dogeared, need not be moved. All the elements of fast response, variable speed scanning are present in this concept. The mechanical speeds required are compatible with potential RLC compression factors and telephone-line bit rates.

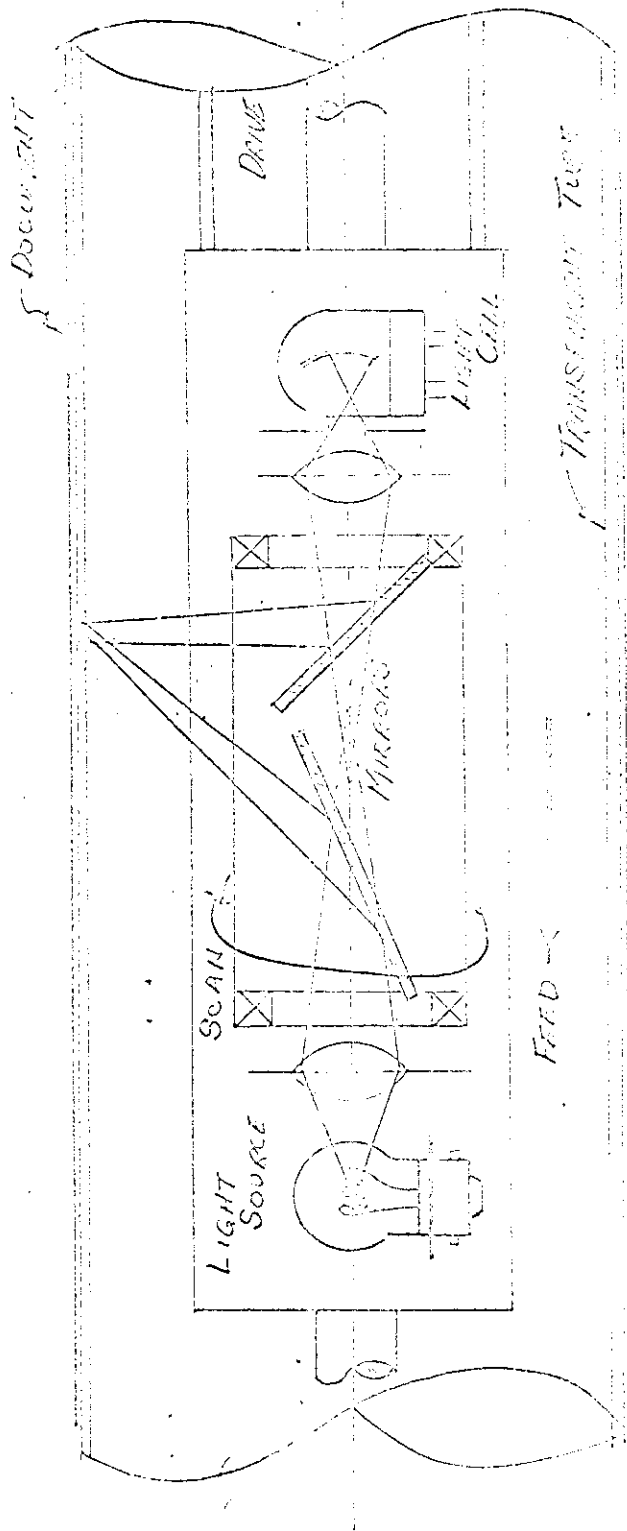


FIG 4-12-1 LOW-INERTIA DOCUMENT SCANNER

See 4-12-5 for size

SECRET

SYSTEM #13: FREQUENCY MODULATION -
PHASE MODULATION ADAPTIVE SYSTEM

Description: This is a proposed system in which a set of fiber optic rods connected to phototransistors and logic determine the approximate distance to the next transition between white and black.

The speed of the scanner is transmitted by the frequency as an FM signal. The time of transition between black and white or white and black is indicated by 180° phase shifts (PM).

A possible fiber optic grouping and logic is shown in Fig. 4.13-1. The correspondence between Frequency, Scanning Velocity, and Run Length is illustrated in Fig. 4.13-2.

4-13, 2

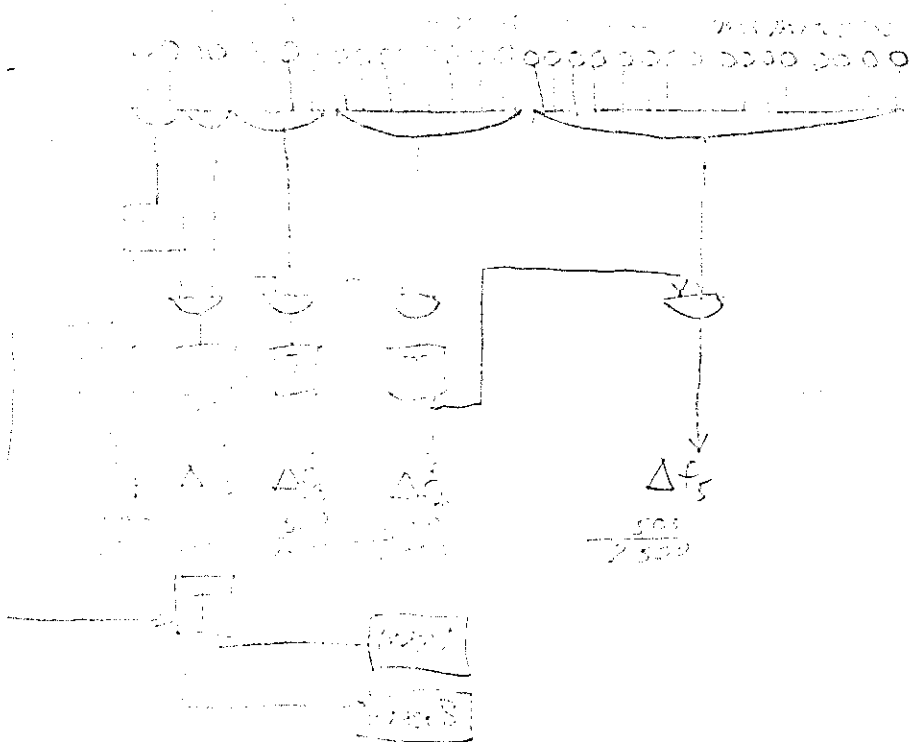


Fig 4.13-1 Fiber-Optic Communication
Amplification

1000 1000 1000 1000

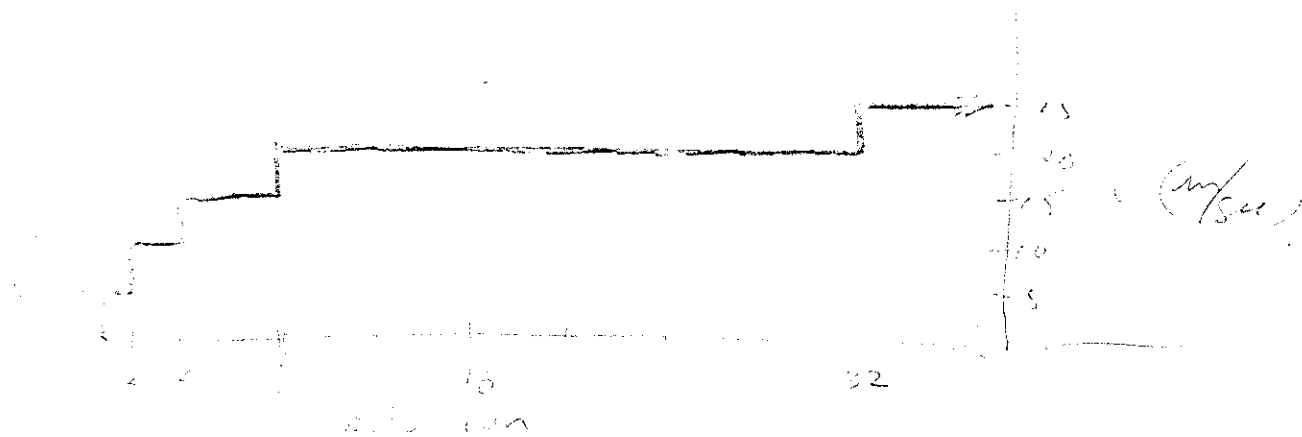


Fig 4.13-2 Frequency and distance
versus distance

SYSTEM #14: PATTERN CODING (SEGMENTS)

Description: Break all letters, graphics, etc., into basic pattern elements as is shown in Fig. 4.14-1. Code element and position (or angle) into binary code. This is similar to the handwriting and pattern recognition system of M. Eden. (IRE Trans. IT, Feb. 1962)

This scheme, although seemingly "way out" might be easily instrumented with some clever choice of the basic patterns to be recognized and techniques for recognition. Common factors to all should be utilized and registration should be immaterial.

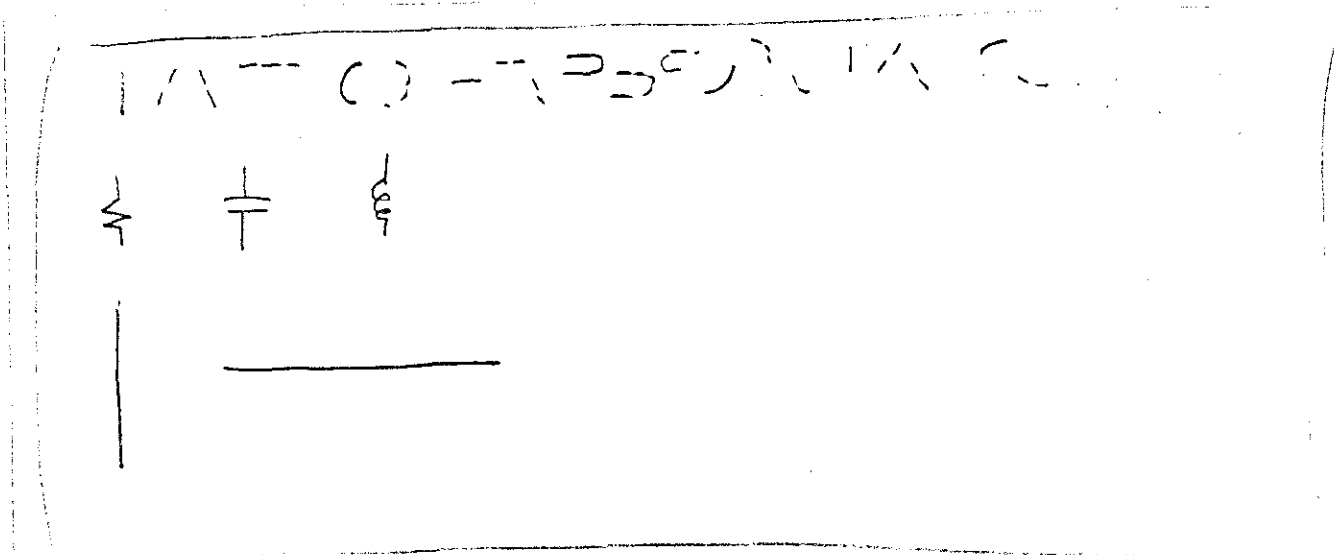


Fig 4.14-1. Segments

A related system was considered, but not examined in detail - -
TOPOLOGICAL TRANSFORMATION. This applies to system where
the format of the characters may be distorted but still recognizable.

SYSTEM #15 - PATTERN CODING (DOTS)

Description: This is closer to a pattern recognition system than a true compression coding system. A prescan would line up an 8 x 8 fiber optic matrix over a character as in Fig. 4.15-1(a). The character would energize 12 of the 64 fiber optics, of which 5 would belong to the subset of 10 used in Fig. 4.15-1C. For limited type fonts this system could effect a compression. It would require a read only memory at the receiver to fill in the lines through the dots.

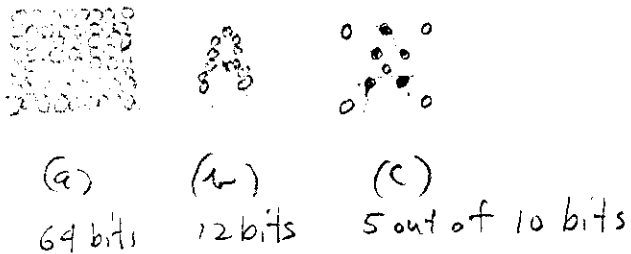


Fig 4.15-1 Pattern Coding Matrix of Fiber Optics.

4-16.1

SYSTEM #16: FULL LINE SCAN

- DOCUMENT ONLY (NOT APPL. TO)

(1) DOCUMENT SCANNED WITH VERTICAL SLOT OF 0.25 INCHES HEIGHT,
OPTIMAL SHAPE:

(2) THE PHOTOCELL OUTPUT CAN
BE CLASSIFIED AS A SEQUENCE
OF SLOPE-AMPLITUDE PARAMETERS.

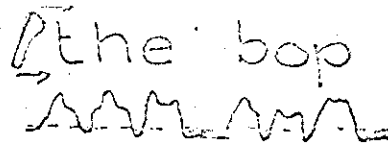


Fig 4.16-1 Full Line Scan

(3a) SLOPE-AMPLITUDE PARAMETERS ARE THEN CODED AND
MITTED.

(3b) POSSIBLY ANALOG INFO CAN BE MODULATED DIRECTLY.
IN CASE 3a, 3b, ^{WHOLE} ONE LINE OF PRINT IS TRANSMITTED
ONE LINE AT A TIME.

(4) THE SLOT WOULD HAVE TO DIFFERENTIATE BETWEEN
b/p AND q/c/d

(5) THE ANALOGY TO VOICE RECOGNITION SCHEMES
(SHOE BOX) IS APPARENT, BUT THIS FULL
LINE SCAN HAS SEVERAL SIMPLIFYING FACTORS
OVER THE SHOE BOX SYSTEM:

4-16.2

a) CHARACTER INTERSPACE EASILY
DETECTABLE (DARK PORTIONS IN DRAWINGS
ABOVE)

b) INDEPENDENCE FROM AMPLITUDE
(TYPE FONT, IN THIS CASE), ACHIEVABLE
BY DESCRIPTORS:

S⁺ STEEP POS. SLOPE

S⁻ " " NEG. SLOPE

P FLAT PEDESTAL AT NON-PEAK AMPLITUDE

SO THAT CHARACTERS DEFINED
BY SERIES OF DESCRIPTORS:

(S⁺, S⁻, P, S⁻) (S⁺ P S⁺ S⁻ P S⁻) etc.

6. TOTAL OF 64, or 50, COMBINATIONS REQUIRED
(> 7 bits per character)

SYSTEM #17 COORDINATE CODING

Description: Coordinates of Block Areas are Transmitted. Horizontal portion of document divided into 30 sections - 30 "bits" long. Each section described by a 5-bit code word. Vertical position in one scan - width increments - at end of each line scan a sub-head to "one inch one scan" generated.

As long as no block detected by scanner, only "Advance" code transmitted. When black detected, sector code and data string transmitted - if succeeding sector also contains black its data string transmitted immediately without section box.

Scanning rate could be about 100 times transmitting rate but scanner would be controlled by buffer, i. e., as buffer becomes filled scanner is slowed or stopped.

(see Fig. 4.17-1)

Comments: Total blank document could be described to transmitted in about 2 seconds. A closely printed document could probably then or more.

Since means of distinguishing between codes loads and data must be devised.

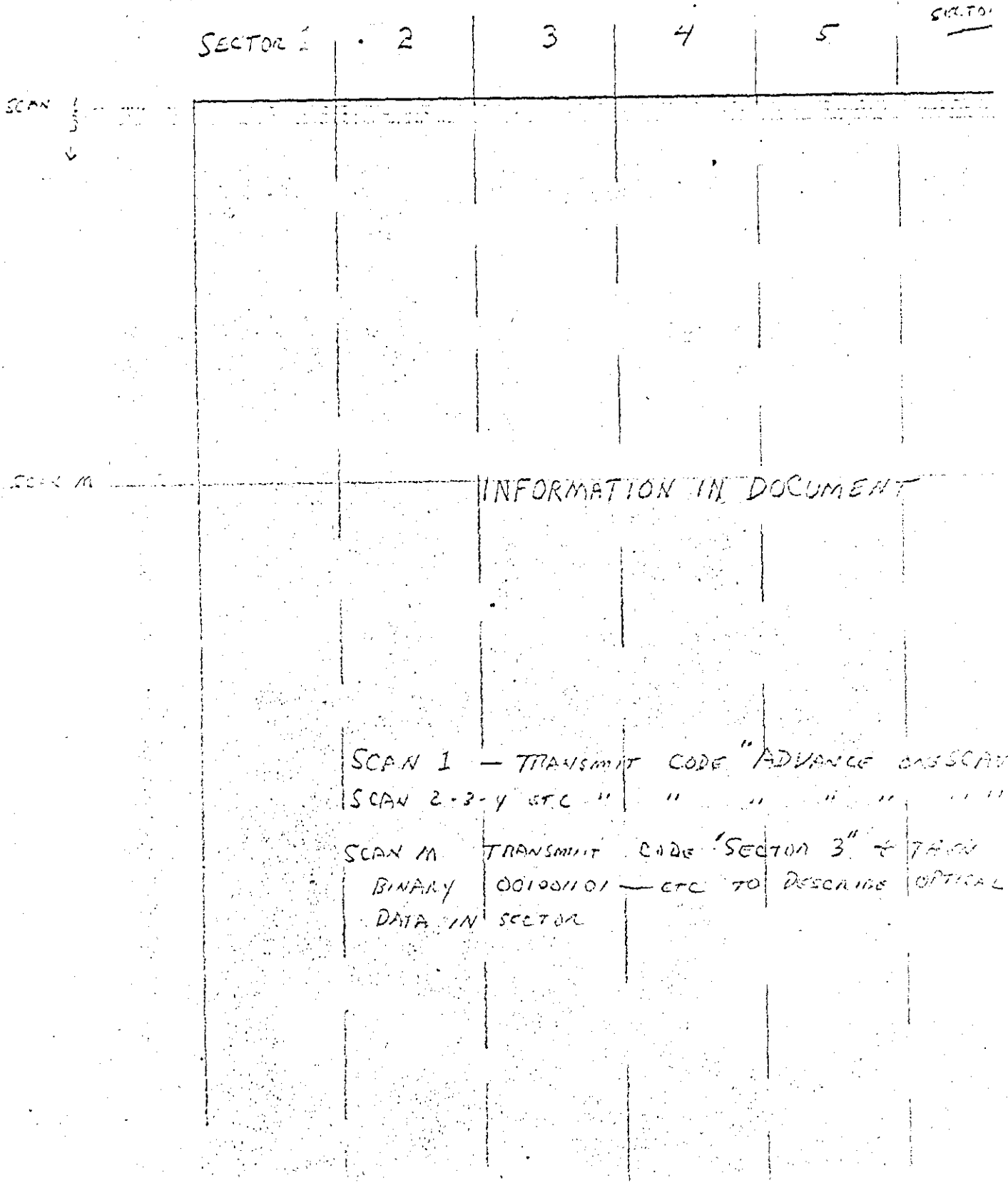


Fig 4.17-1 Sector layout for local units

SYSTEM #18 - TWO DIMENSIONAL LINE SCAN

Description: Prescan measures vertical and horizontal redundancy and connects reading fibers for optimum type (i. e. horizontal scan or grouped into 5 vertical dots together).

This system would scan horizontal lines and prescan 5 lines ahead for state count and partial scan for blanks next 6 lines. The logic would determine whether the next line would be processed vertical or horizontal.

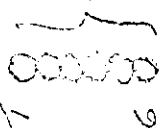
This system may involve features of: Facsimile AM-SSB plus Run Length Coding, Adaptive, plus Coordinates of Sections.

The basic feature of this proposed system is that the lower case body of typed lines is coded vertically, while the hats, dots, and tails are coded horizontally.

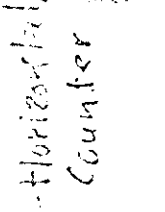
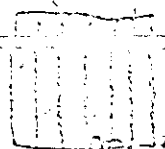
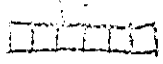
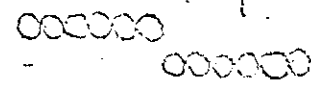
TWO-DIMENSIONAL LINE SCAN

the

READ

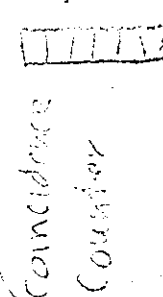


WILL



Hbr. Signal
Indicators

LOGIC
Vert.
Hbr.
SKIP



Coincidence
Counter

CLASSIFY

CLASSIFY

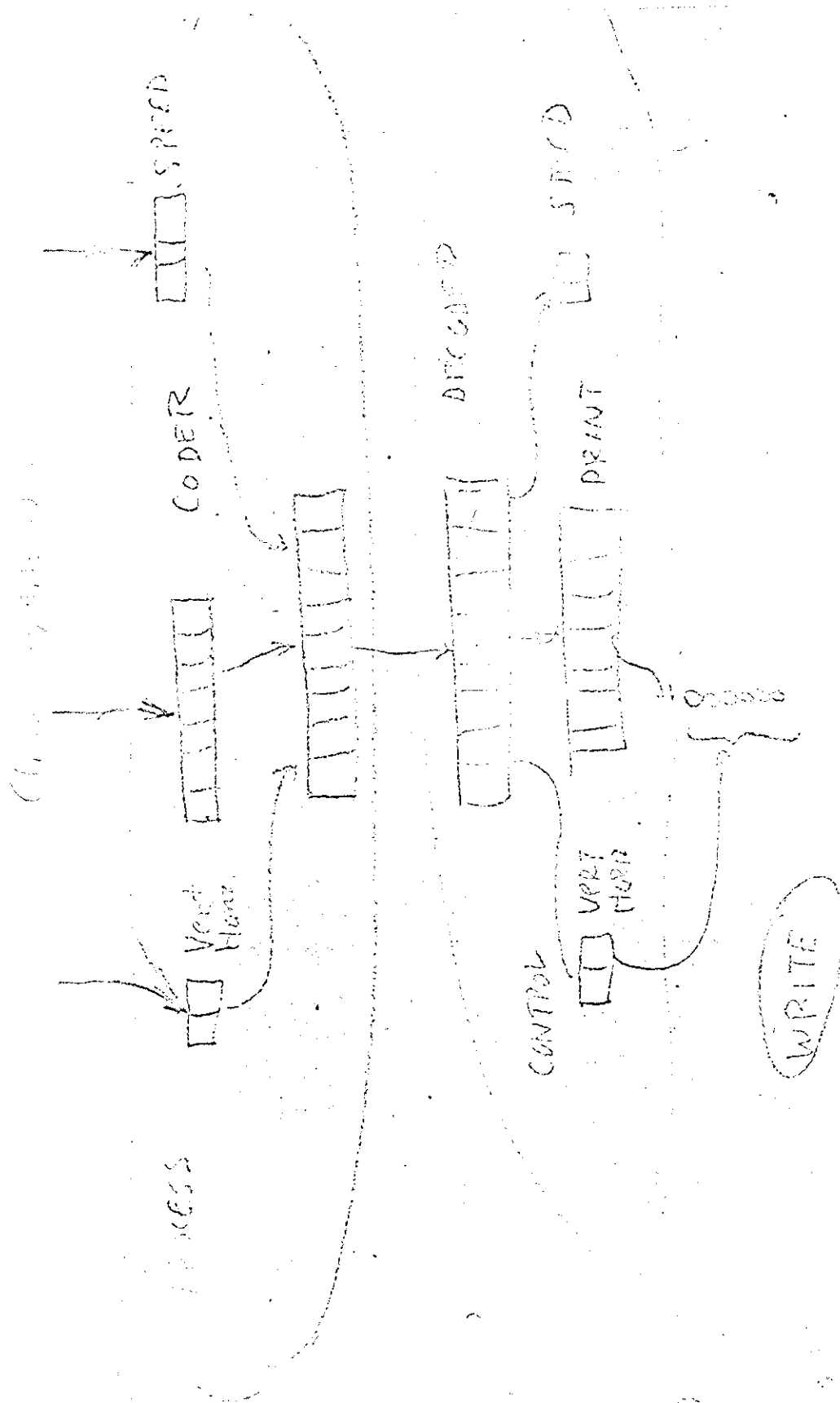


Figure 1.1 Computer System

SYSTEM #19: MATRIX PROCESSING

In this scheme, we provide an optical scanning "head" which consists of an array of optical fibers as shown in Fig. 4.19-1. We

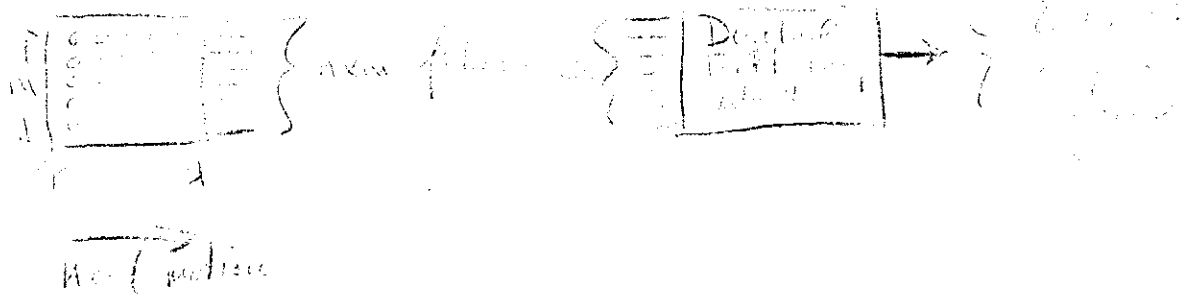


Fig 4.19-1 Matrix Processing

combine the fiber outputs in a digital filtering network (sequential combinational logic) such that the spatial redundancy in the scanned image is removed. One such scheme consists of two adjacent fibers whose outputs are summed modulo 2 such that "0" is produced if their outputs are the same and "1" is produced if their outputs are different. This is a form of cell-to-cell prediction to reduce redundancy. If it is desired to use only a portion of the $m \times n$ array (e.g. the diagonal)

Then the other fiber outputs might be dropped. The object is to convert undesirable run-length statistics (e. g. monotone decreasing) to more desirable statistics (e. g. highlight at longer runs) so that Huffman coding will result in greater compression factors. Some of the fiber outputs could be parallel outputs to be converted to serial at a later point. Thus, due to the head motion, we might be scanning little vertical strips along a print line (this is one method of changing the run length best solution).

The receiver might have a similar head with both read and write elements. The read elements would view the information already written out and, together with the redundant-free information received, compute what information should be written at any particular instance. There, an "increase" type digital filter would be provided at the receiver which would put the redundancy back in the transmitted information.

SYSTEM #20: OPTIMUM BLOCK LENGTH CODING

Since this system was incorporated in the feasibility model,
the description will be found in Section V.

NOTE

Certain cyclic code and recurrent code techniques were considered in this section, but were moved on to the line drawing section, because they were not found to be capable of reasonable use for typed or handwritten documents.

SYSTEM #21: COORDINATE ENGINEERING DRAWING COMPRESSION

Goal:

I. DRAWINGS:

- A. OLD FASHIONED - ARTY, FINE DETAIL, DIFFICULT TO COMP.
- B. "SIMPLIFIED DRAFTING PRACTICE" ea., (BC). POSSIBLE
- C. FUTURISTIC - DESIGN DWG FOR IMAGE TRANSMISS.

II PICKUP:

- A. FOLLOW LINES WITH SERVO CONTROL - LOOK AT STRAIGHT
- B. FUTURISTIC - INFO STORED ON TAPE ~ POSSIBLE
- C. SCAN WIDE TO FIND LINES - SPIRAL AT END POINTS

III BUFFERING:

- A. DATA - GATHER SLOWLY - BUFFERS BATCH DATA - ONE
- B. INSTRUCTION - STORED LINE ENDS & FEW INTERMEDIATE. PROHIBIT SCANNING & LINE FOLLOWING LOOPS
- C. FOLLOW RULES OF SCANNING BY REF. TO BUFFER

III ENCODING ALTERNATIVES:

- A. XY ANALOG SIGNALS ~ (TELEAUTOGRAPH) - POOR
- B. XY ANALOG - QUANTIZE - DIGITAL
- C. END POINTS + LINE CURVATURE (S OR C) + TYPE + ID EACH LINE
- D. EP + LENGTH + SLOPE + CURVATURE + TYPE + ID
- E. COMPRESS EP BY REF. TO COMMON EP ONLY & LIST ALL STRAIGHTS & CIRCLES SEPARATELY COMPLEX TO DIM'N.

V CHANNEL

- A. SEND BATCHES

VI RULES ~ COULD BE QUITE SOPHISTICATED

- A. LIST LINES BY TYPE
- B. ALL ARCS CCW
- C. ETC

VII OUTPUT

- A. X-Y RECORDER - MECH
- B. CRT

Description:

Image transmission of engineering drawings could be accomplished by using an encoding scheme quite different from that used in transmitting printed material due to the nature of these drawings. Generally speaking, engineering drawings consist of coordinate points joined by relatively few kinds of lines; e. g., straight lines, arcs, electronic components, dimensions. The coordinate encoding method, which can be quite efficient compared to run-length coding, recognizes and utilizes this fact.

From its beginning until recently, the art of drafting has been primarily that - - an art. Great emphasis was placed on drawing fine details and perfect letters. However, since the end of World War II, designers have come to realize drafting is not a fine art - - that one picture is not necessarily worth a thousand words. Rather, many times a brief, quickly typed (and transmitted, incidentally) note supplants a complex drawing. A modern systematic approach to drafting has been adopted by the General Electric Company ⁽¹⁾. It is to this latter type of drafting that image transmission would be most readily applied. In fact, because the lead time of this transmission system approximates the life of a typical engineering drawing (three to six years), a new method of drafting, which is also compatible with mechanical design automation, can be created for optimum encoding and transmission.

Facsimile image transmission by standard techniques is accomplished by dividing an area into a large number of arbitrarily small units, defining each unit by position, and describing it by color. Rather than transmitting each dot or specifying the number of like dots in a line, this system would transmit only the locations of line end points (nodes) and information defining the material between these nodes. In order to reduce the total number of kinds of lines transmitted, code lengths could be reduced by using control panels or stored program for each major kind of drawing. Typical kinds of drawings are electrical schematic, block diagram, plumbing schematic or layout, patent, mechanical, welding, architectural, and structural.

Two drawings - - a still life line drawing of a fruit bowl, and an electrical schematic were studied to determine what compression factors are possible using this method. (See-attached-sheets) These

particular drawings were chosen because they had previously been analyzed by Bell Telephone Laboratories. (2) Two fairly general codes were used with the still life drawing: single line encoding and multiple line encoding. By the first method the alphabet consisted of four geometric shapes: straight lines, curved lines approximated by arcs of radius less than 10.2 inches, sectioning bounded by polygons, and sectioning bounded by arcs and lines. To be consistent with the previous analysis, the 8-1/2" x 11" page was divided into 10-mil dots. Therefore 21 binits were required to encode each coordinate. Arcs can be defined by any one of several methods, the optimum to be determined by the implementation scheme. Single line encoding required the transmission of 4,023 binits representing compression factor of 12.5 times from Shannon-Fano run-length coding or 230 times from straight facsimile. When nodes were shared by two, three, or four lines, the compression factor increased to 15 and 280, respectively.

A special code was devised for the electrical schematic. Only those components which were present were allowed in the scheme; therefore, the resulting compression factors represent an optimum rather than a generalized electrical schematic condition. The alphabet consisted of three separate types of words: two-node connectors, straight-line end points, and fixed length notes. Each word was made up of: data bits which describe the coordinates and other unique information, and a uniquely decodable prefix which specified the word and its length. Two-node connectors consisted of lines, chokes, resistors, etc. that can be described by type and two node positions. Straight-line end points are lines that emanate from a dot and terminate in a special symbol such as an emitter, a relay contact, or an arrowhead. A separate prefix was used for those notes of a unique length, and the location of each note was specified by one coordinate.

Two different methods of prefix coding were tried: variable length Shannon-Fano, and fixed length five-bit. Although the former resulted in fewer total binits transmitted (19,270) it would be much easier to implement the slightly longer (19,615 binits) five-bit prefix. This difference is small because a major portion of each word consists of data bits (42 to 141 binits) per word. Coordinate compression of this electrical schematic resulted in compression factor from Shannon-Fano run length and straight facsimile of 6.25 and 47, respectively.

These preliminary analyses were performed to approximate compression factors with little regard for implementation. Unless a storage medium, such as paper tape, is prepared during the drafting procedure, scanning might prove to be very difficult.

Possibly a servoactuated line follower with associated rules and logic could be developed. Output documents could be prepared with a modified X-Y recorder controlled by an analog or digital computer, available commercially. (3)

The compression factors obtained indicate that further work should be done to exploit this method of image transmission.

REFERENCES

1. Healy, William L. and Rau, Arthur H., Simplified Drafting Practice, John Wiley & Sons, Inc., New York, 1953
Copyright by General Electric Company.
2. Michel, W. S., Fleckenstein, W. O., and Kretzmer, E. R., "A Coded Facsimile System" 1957 IRE Wescon Convention Record, Part II. Bell Telephone Laboratories.
3. Kliever, W. H., and Little, H., "Numerical Control Produces Drawings," Control Engineering, May, 1962, Universal Drafting Machine Corp.

SYSTEM # >22: CYCLIC CODE COMPRESSION

Description: If we consider a blank page as the correct message and the dots of block as errors and run the digital representation of a page through a Bose-Chandhum,-Hockingem error-correcting decoder to obtain the "corrector" digits. Then we transmit the corrector digits which from a message having a number of digits equal to the parity bits only. At the receiving end the running of the "corrector" parity bits through the decoder-corrector circuits would generate correction bits corresponding to the block dots. The limitations on number of errors that can be corrected by a reasonable amount of redundancy in a code limit this case to line drawings, with a limited amount of text for labels.

Reference and Example:

Edwin Weiss has developed a formulation of the relationship between compression and coding*. Weiss states that if each block of data contains less than e ones and there exists an e -error correcting (n, k) code A , then a compression of $(n) \rightarrow (n-k)$ is possible.

For 64 pica characters per line with 4 transitions per character' or 256 bits per line of 850 bits or 1700 bits

Divide line of 850 bits into sectors of 127 bits

i. e. $n = 127, e = 256/7 = 37$

for $t = 31, n = 127, k = 8.$

$$R = 127-8/127 = 1/1.07$$

For 14 sections of 127 bits @ lines per inch or 1700 bits/line

$e = 256/14 = 18+, t = 15$

$(n, k) = (127, 36) \quad R = 127/127-36 = 127/91 = 1.4,$

*Edwin Weiss, "Compressional Coding" Trans IRE, Vol. 17-18, No. 3
April 1962, pp. 256-7

Appendix on FSR Compression of Characters:

Cyclic code compression works for drawings, but not for typing. Methods were explored of sending the parity bits plus a transposition code for transposing the columns of the FSR at the receiver, plus a starting point code to indicate where in the cycle of state of the FSR, the initial point should be. Examples of this type either filed in the holes in character like A and B or lost the crop bars in characters like A. For further possibilities see I. S. Reed and R. M. Stewarts "Note on the Existence of Perfect Maps" Trans. IRE, Vol. II, January 1962, pp. 10-12, Appendix on Recurrent Codes.

Recurrent codes such as Hagelbarger codes were considered for compression of line drawings. It was found that the restriction to "bursts" correction put on too many restraints on the system to be workable.

SYSTEM # 23: ANALOGY FROM BIOLOGICAL SYSTEMS

There are biological systems exhibiting a kind of compression coding in the multiplexing of the axons of the sensory system.

Clyde M. Williams ("Representation of Locality in a Biological Information System," Bull. Math Biophysics - 20, 1958, pp. 217 - 230) discusses the evidence for the existence of complex code compression (or multiplexing systems) in nerve system of the rabbit. For example, 5000 dorsal root axons innervate 100,000 hairs on the rabbit ear, and each hair is innervated by 4 to 6 axons. Under minimal stimuli the exact locality is transmitted to the brain, but under maximal stimulation the details regarding locality are lost.

No direct benefit to engineering studies has come from such biological examples, but as we learn more about biological systems, we may find useful analogies.

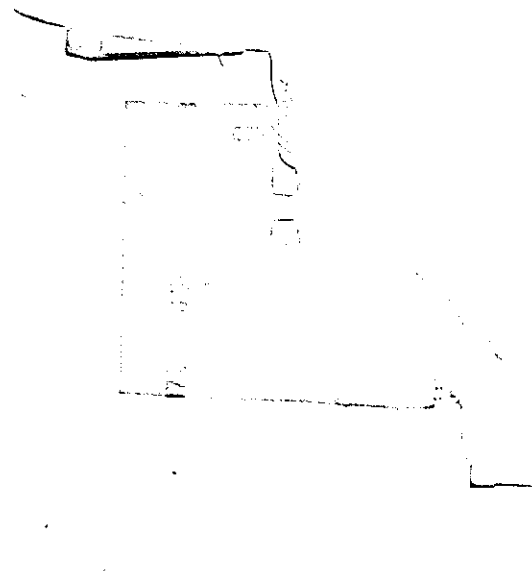
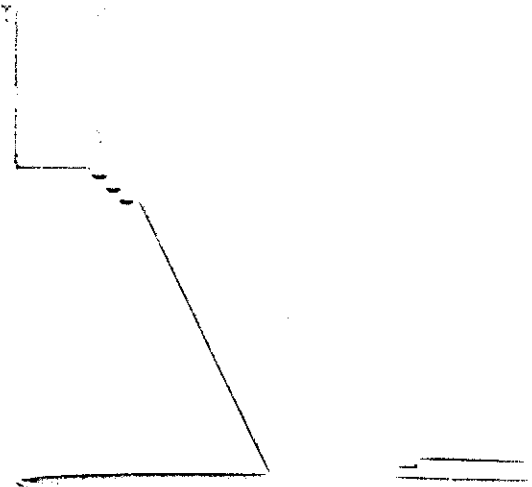
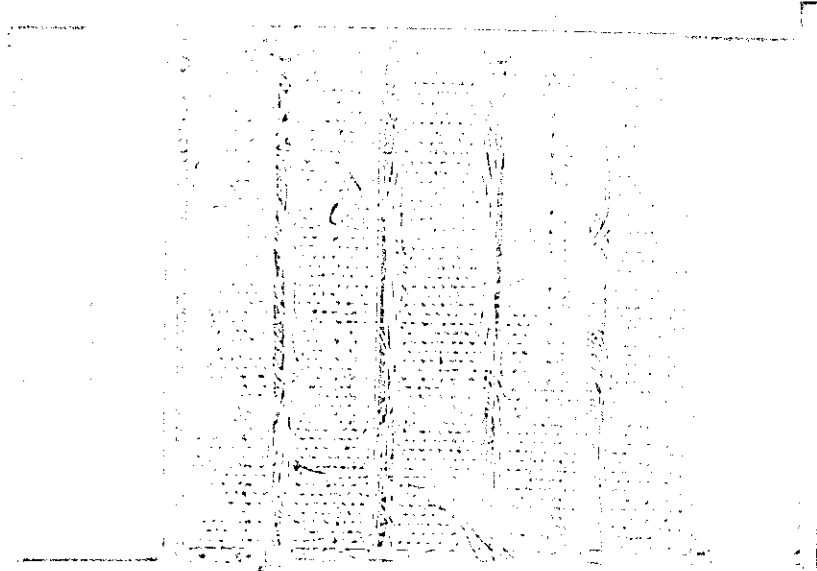
√, FEASIBILITY MODEL - RUN LENGTH CODER AND
SMOOTHING BUFFER

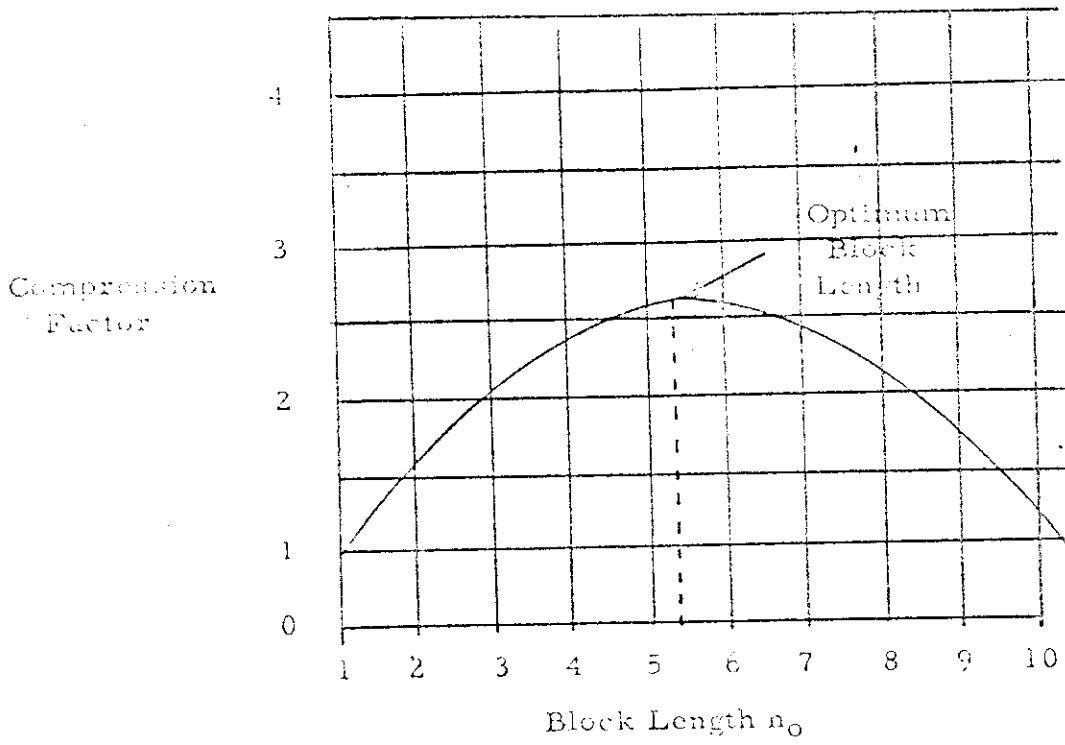
A. THEORETICAL BACKGROUND

Digital block coding of run lengths is a coding scheme where all runs up to a maximum of 2^{n_0-1} are counted by a block of n_0-1 bits and one bit is used to designate the color of the run. This scheme can result in compression only if the length of the run is greater than the length of the block code representing that run length. If the average run length is longer than the block length of the code, then compression will occur on the average. The amount of compression that can be achieved is strongly dependent on the block length of the code. For a particular run length probability distribution, there is an optimum block length as shown in Figure 1. Straight facsimile occurs when the block length $n_0 = 1$ and no compression is achieved. Intermediate block lengths result in moderate compression, but long block lengths result in compression factors less than unity.

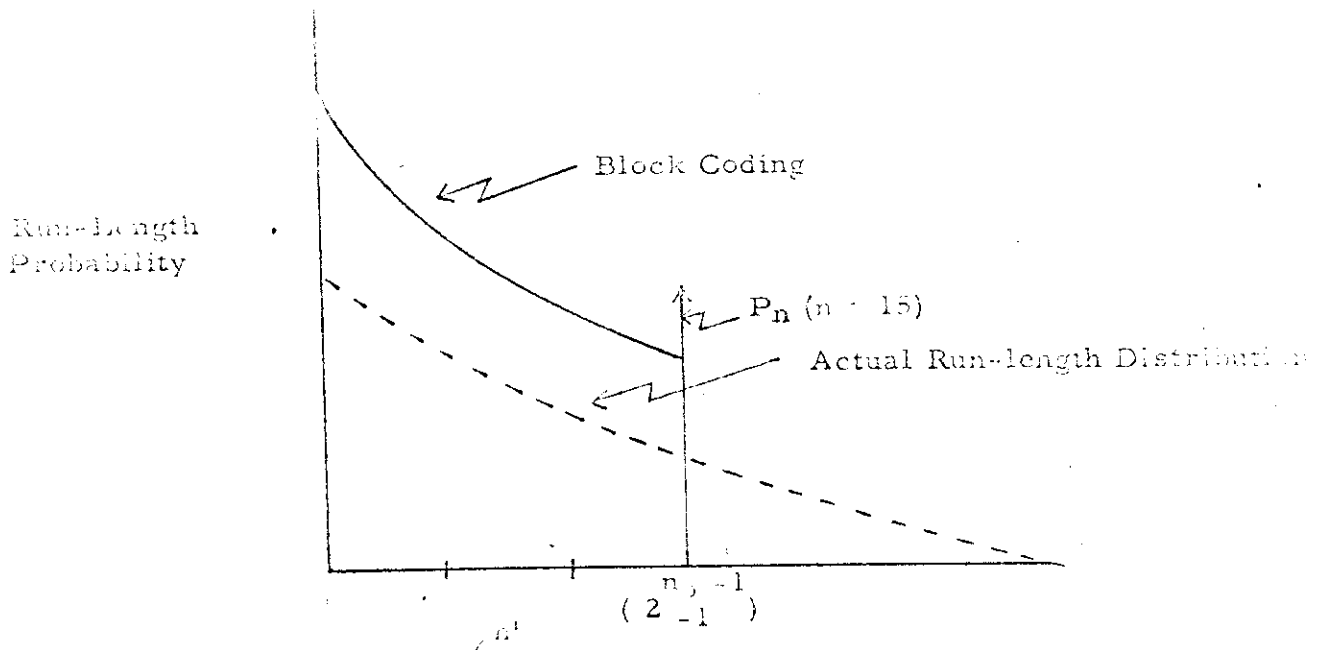
It should be noted that statistical variable-length coding of the run lengths can result in even greater compression factors than block coding. However, variable length coding is difficult to implement and thus is not used in the initial feasibility model. One way of viewing block coding is suggested in Figure 2. Block coding distorts the actual run-length probability distribution so that run-length counts greater than 2^{n_0-1} are impossible. This distorted distribution is the one which must be used for determining the optimum variable-length code (i. e., the Huffman coding of the new run length source, n^1).

Coding of the run length data introduces time gaps in flow of information out of the coder even though the information is being fed





5.1
 FIGURE 4 Optimum Block Length



5.2
 FIGURE 2: Distortion of Run-Length Distribution Due to Block Coding

smoothly into the coder. Compression can occur in real time only if this coded information is fed smoothly to the communication channel. A smoothing buffer performs the duty of eliminating the gaps in the coder output. The operation of this buffer will be described in the hardware section.

In order to experimentally determine the feasibility of an image coding and data smoothing scheme, the system depicted in Figure 5.3 was constructed. Black-white image data is scanned by seven optical fibers in parallel and sampled at a fixed rate. Parallel to serial conversion of the sampled data is performed by a seven stage shift register. This serialized information is then fed into a 4 stage binary counter which counts run lengths of black or white elements up to a maximum run of 15. Since two or more consecutive runs of the same color can occur, a fifth binit is suffixed to determine whether the run was black or white.

The next operation that must be performed by the system is a smoothing of the information flow. The coder emits information in the form of 5 binit words quite sparsely in time, since a long run code word is emitted a long time after the preceding word and a short run code is emitted a short time after the preceding word. In order to use the communication channel efficiently, this spurious emission must be smoothed such that the words are presented to the channel at equi-spaced time intervals. A special smoothing buffer effects a solution to the problem.

The buffer is always filled relative to the right end (output end). Data is entered into one of the input lines by means of a gating network controlled by circuits that keep track of how far to the left the buffer is filled. The data in the buffer is simply shifted right to the transmission line at equal intervals in time. The coded image information can thus be delivered smoothly to the communication channel.

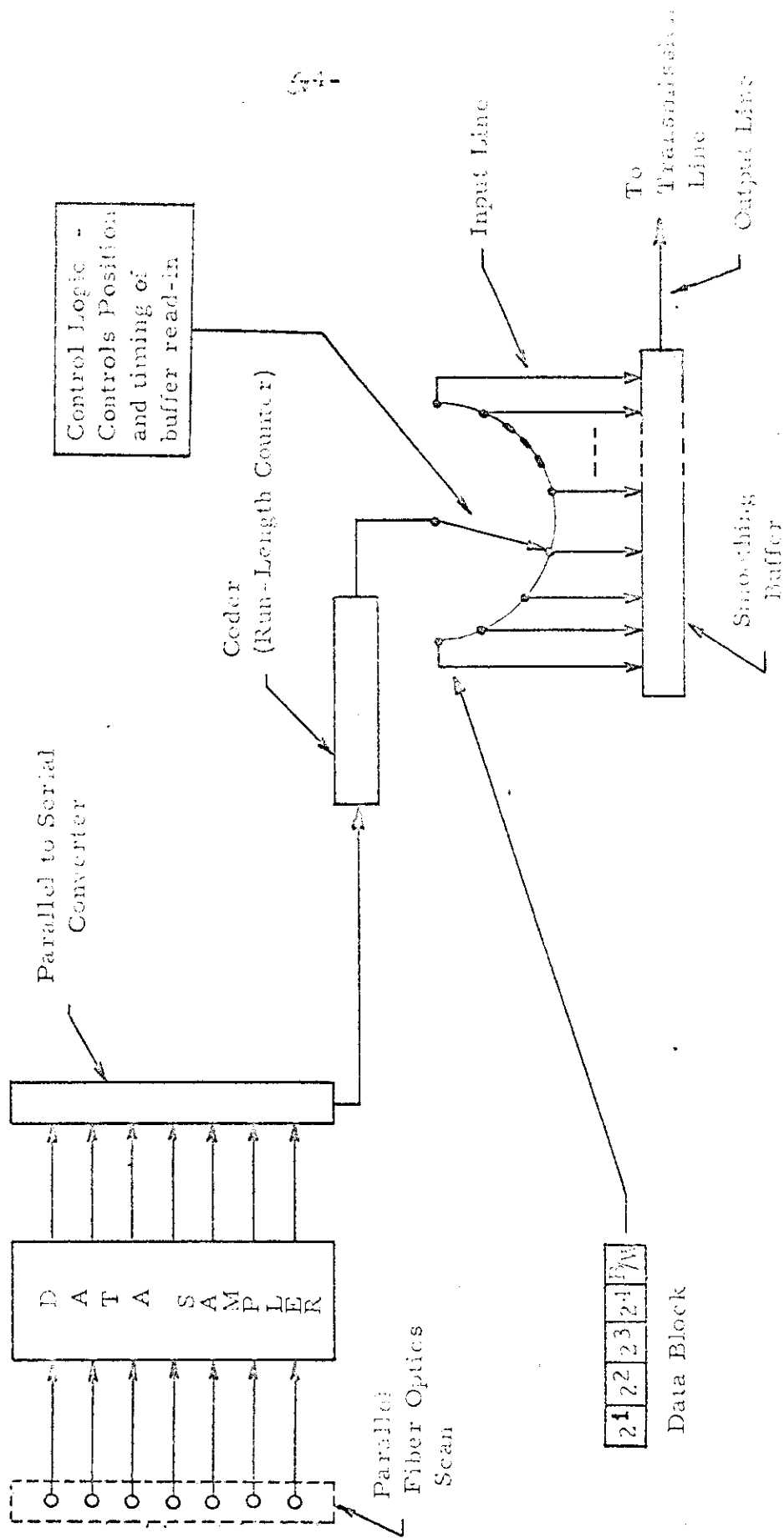
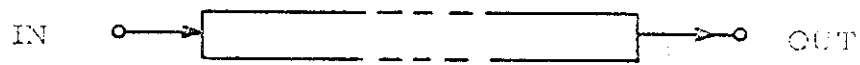


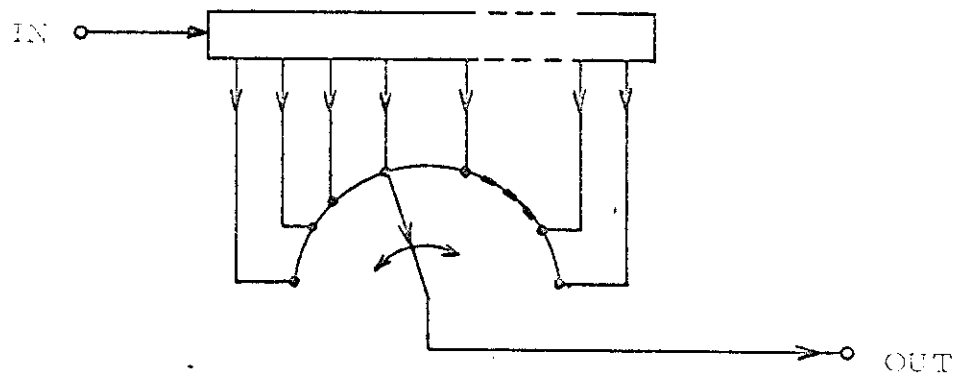
FIG. 10. Inverse Coding and Decoding

It should be noted that there are three other variations of this buffer structure, shown in Figure 4. The fixed position input, fixed position output scheme depicted in Figure 4 (a) requires the detection of "0" bits to determine the position of information being shifted toward the right in the portion of the buffer not already filled. When the "propagating" information reaches the last word in the queue, propagation is ceased so that no information is lost. Shifting of the queue is always inhibited unless an output request occurs. The scheme in Figure 4 (b) is similar to the buffer used in the feasibility model, except that the input position is fixed and the output position is variable. In Figure 4 (c) we eliminate the shifting operation by using variable position control on both input and output, utilizing a static storage medium.

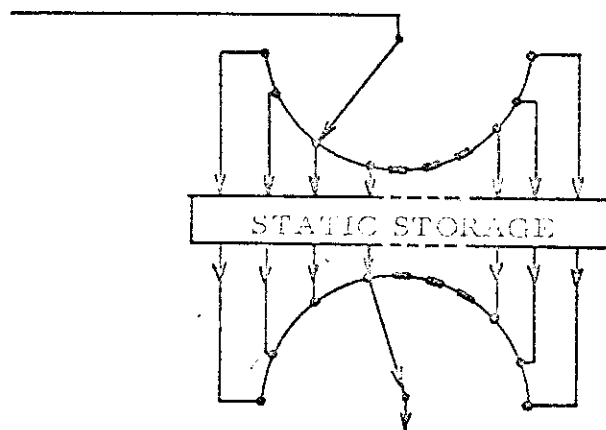
FIGURE 4: Alternative Buffer Schemes



a) Fixed Position Input, Fixed Position Output



b) Fixed Position Input, Variable Position Output



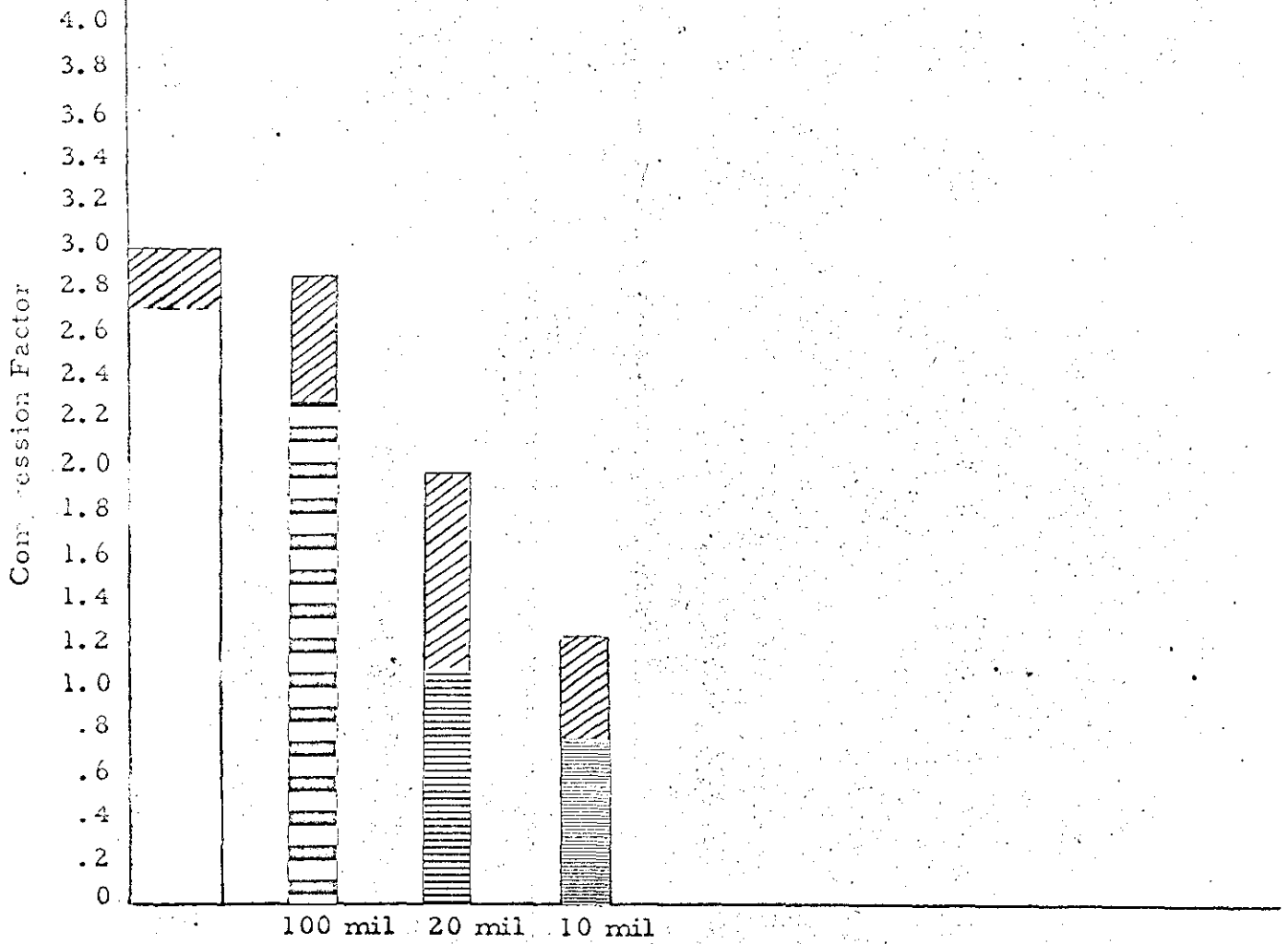
c) Variable Position Input, Variable Position Output

C. EXPERIMENTAL RESULTS

Time did not permit a thorough investigation of the compression capabilities of the system. However some preliminary data was taken to compare the predicted results. A scanner resolution of 6 mils was used to obtain the data in Figure 5. The compression factor is obtained by forming the quotient of bits into the coder and bits out of the coder for one scan line. For blank areas, a compression factor of 3 is expected since we code a run of 15 binitis with 5 binitis. Compression is reduced for shorter run lengths and becomes unity for run lengths of 5. This can be seen in the figure: as the system scans the finer grids, compression is reduced.

Several factors ^{may} explain the discrepancy between the predicted and experimental results. If the grid bars are slightly skewed with respect to the vertical line of 7 fibers, run lengths shorter than 5 will be frequently encountered when scanning the finer grids, thus reducing compression. Further, if some of the fibers are out of line, short runs will be encountered when the sampling time occurs near a black-white transition. Another factor is that one of the scanner output voltages could not be maintained at the desired level for switching succeeding circuits during the test. This would introduce errors in the data and produce shorter runs when scanning the "blank" pattern, as well as the other periodic patterns:

Despite the discrepancy between some of the experimental and theoretical results, the model does prove feasibility of the coder-buffer system in achieving moderate compression of black-white image data. It is believed that block codes and a short smoothing buffer of the type described in this section (although perhaps implemented with low cost core logic) used in conjunction with a variable speed controlled scanner of the type discussed elsewhere in the report would provide a moderate compression terminal at a reasonable cost.



3

FIGURE 5 Block-Length Coding Compression Results

D. EXPONENTIAL APPROXIMATION TO RUN LENGTH PROBABILITY

The run length code used in the feasibility model described in the preceding sections was derived by fitting a Poisson probability distribution to the Mohansic run-length data for a sample set of typed pages. Print-outs of the statistics used in IBM Report 17-037 (Arthur L. Prudhom and Ulbe Faber, "Band-Width Compression for the Transmission of Wide-Band Facsimile Signals, " 4/21/61) were generously made available to us by A. L. Prudhom.

If we assume a random signal with the transitions between black and white randomly distributed in accordance with the Poisson distribution

$$P(n, T) = (a T)^n e^{-a T} / n! ,$$

where n is the number of transitions per interval of time T (seconds), and a is the average number of transitions per second.

The probability of a run length is derived as follows:

$$\text{Let } P_1 = P_r \text{ (no transition in the time } T) = e^{-a T}$$

$$\text{and } p_2 = P_r \text{ (no transitions in the time } \Delta T) = e^{-a \Delta T}$$

$$\begin{aligned} \text{then } P_r \text{ (run length equal to } T) &= \lim_{\Delta T \rightarrow 0} P_1 (1 - P_2) \\ &= e^{-a T} (1 - e^{-a \Delta T}) \approx e^{-a T} (1 - 1 + a \Delta T - \dots) \\ &= a e^{-a T} \Delta T \end{aligned}$$

$$\text{Then } P_{re} = P_r / \Delta T = a e^{-a T} .$$

Let $T = \ell \Delta T$, then the mean time:

$$T = \int_0^{\infty} T a e^{-a T} dT = 1/a = \bar{\ell} \Delta T; \quad a = 1/\bar{\ell} \Delta T$$

$$P(\bar{\ell}) = P_{r\ell} \Delta T = a e^{-a T} \Delta T = a e^{-T/\bar{\ell} \Delta T} = 1/\bar{\ell} e^{-\ell/\bar{\ell}},$$

where ℓ is the number of bit times in the run.

The Mohansic statistics are plotted in Fig. 5.6. Points have been added to indicate the effect of limiting the blocks to 15, 31, 63, etc. bits. Two Poisson derived curves have been added, one for $\bar{\ell} = 3$ and the other for $\bar{\ell} = 4$.

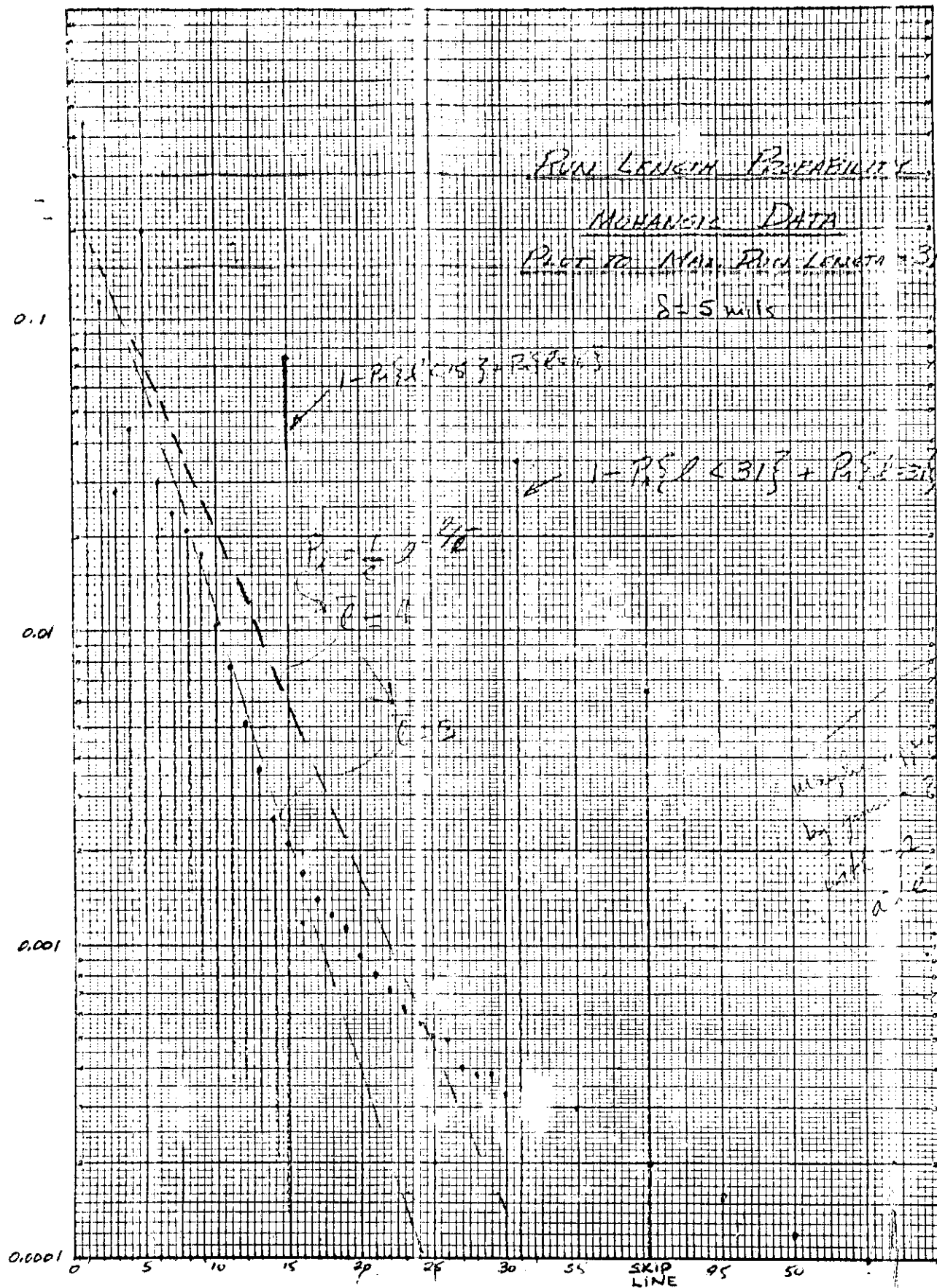
Below run lengths of 15, the exponential curve for $\bar{\ell} = 3$ fits better and above 15 the curve for $\bar{\ell} = 4$ approaches the distribution better.

The block coding method may best be described with reference to Figure 5.7. Optimum (Huffman) compression coding of the run length source yields code word lengths which are a close approximate to

$$-\log_2 P(\ell)$$

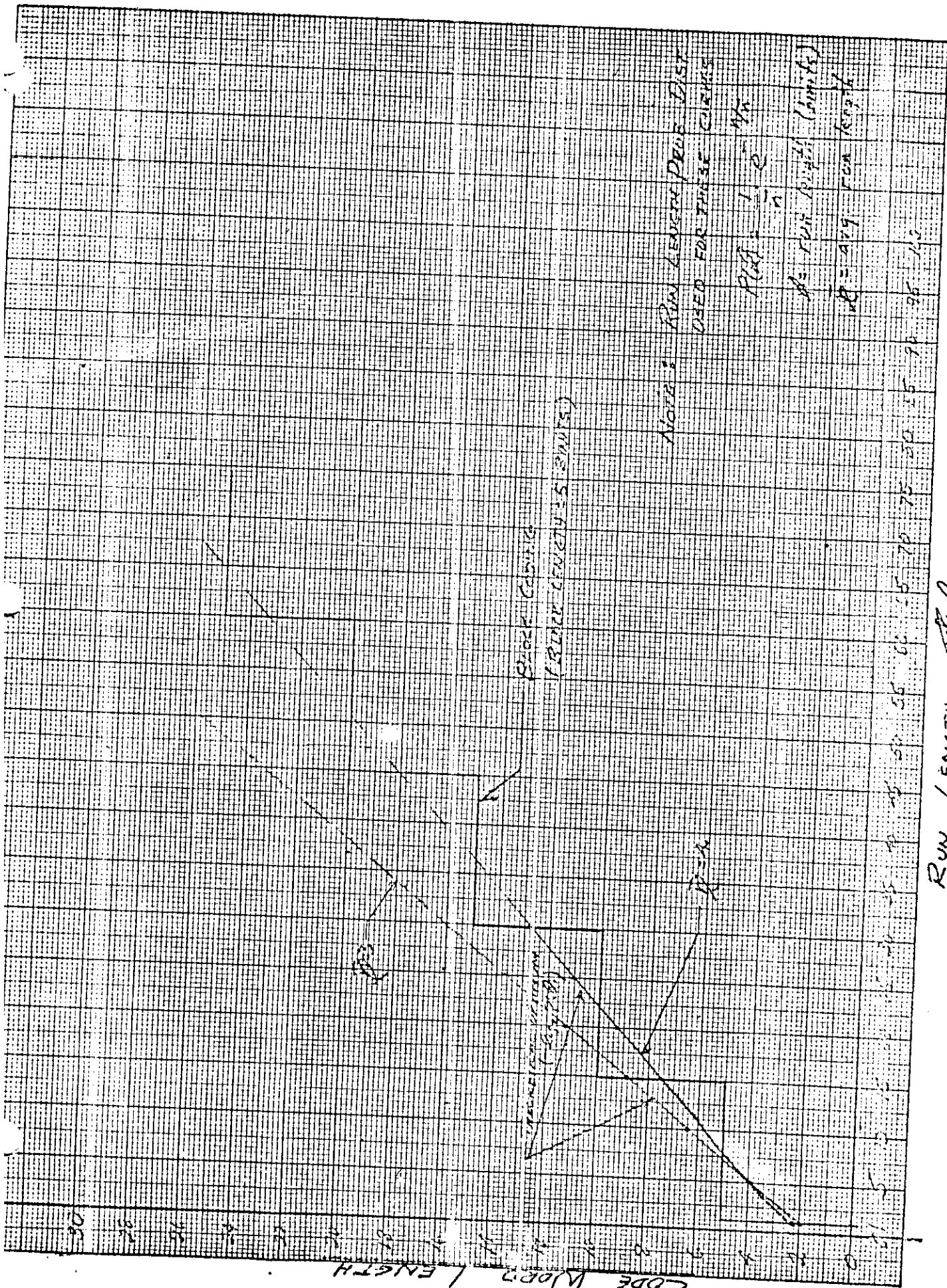
plotted for $\bar{\ell} = 3, 4$. The number of code word bits required to code a given run length for the block coding method is given by the staircase function in Figure 5.7. Note that this forms a first-order approximation to optimum coding for the $\bar{\ell} = 4$ exponential distribution. This method, although less economical in its use of bits to code the source, offers the advantage of ease of implementation over the Huffman code structure.

RUN LENGTH PROBABILITY



1.95.6 RUN LENGTH

Using diagrams such as this as a guide, special variations of block coding may be adapted to achieve second order approximations to the ideal code structure with arbitrary source statistics.



P. D. DODD
AUG 161 1923

RUN LENGTH, FEET

FIG 5.7

VI. CONCLUSION

The basic classes of documents, namely photograph, typing and drawings can be further subdivided into subclasses by the difference in the amount of redundancy which reduces the number of bits of information required per page to represent a document. The information required to represent a page is defined as the ϵ -entropy in order to define the information per page for a specific resolution " ϵ ".

A set of upper and lower bounds on ϵ -entropy are developed for different types of documents. The upper bounds represent the amount of information per page needed for simple coding systems which code on one scan of a line. The lower bounds represent the information per page needed if many scans of the page are made similar to the stages of logic used in character recognition systems:

From these upper and lower limits it is possible to estimate an intermediate value for system using two scans - a prescan and a reading scan.

This classification system leads to the possibility of an adaptive scanning system in which a prescan measures the statistical properties of the page from which the logic selects the coding system to match the document information statistics.

The choice of a method for image transmission depends upon the scanning technology, compression coding technique, and the modulation systems available. Basic facsimile (#1) is recommended only where fine detail is required without too much concern for cost. Even then a more accurate reproduction could be obtained by using some compression and then adding back the redundancy for error detection and correction.

Variable velocity (#2 and others) scanning extremely desirable to adjust the data rate to fit the available channel capacity. Partial transmission (#3) is potentially useful, but requires some logic at the receiver to fill in gaps between transition points. Predictive coding (#4) offers potential savings, if combined a particular modulation system such as data modulation. For photographs the band splitting techniques (#5) offer some advantages. Analog compression coding techniques (#6) have been fairly well developed for TV picture compression. Deviation and rate coding (#7) is definitely an approximation system, but might be useful for certain classes of pictures where small shifts of detail are not serious. The adaptive scanning with bandwidth switching (#8) might be valuable for predominately mixed text where standard type is interrupted by illustrations and tables.

The three types of optical processing (#9) have considerable potential for replacing some electric circuits with optical-mechanical and electro-optic devices. The state of the art for such devices is not

quite at the stage for economic utilization.

Run length coding (#10) is the old standby of practically all feasibility systems for compression coding. If new techniques of using run length coding can be found which do not require large buffers an advance would be made. Adaptive run-length coding (#11) offers a more efficient use of run-length coding by skipping blanks and adjusting to different run-length distribution. The low inertia type scanner (#12) offers the possibility of a practical buffer-controlled scanner for run-length coding.

Another way of overlapping an adaptive scanning system with a special modulation system is to use fm to indicate the scanning speed and am to indicate phase changes at transition between black and white (#13).

Pattern coding (#14) using segments of characters is in reality a kind of character recognition system and is limited by the recognition logic required. If a certain amount of degradation of character shape can be tolerated, it may be possible to take short cuts which result in some distortion of characters in the found a topological transformation.

Pattern coding (#15) using dots can have some simple cases, but in general it is close to the character recognition problem. Where there are large white spaces, coordinate coding (#17) could be very useful. The proposed two-dimensional line scan (#18) offers a way to realize some benefits from the two-dimensional nature of the code compression problem.

To evaluate the potential of matrix processing (#19) we need a more general theory of "digital filters." Some approaches to this problem can be gained from A. Gill's paper on probability transformers in the Journ. Focus of the Franklin Institute and Reed and Stewart's paper "Note on the Existence of Perfect Maps," IRE Trans. on Information Theory, January 1962.

The optimum block length coding (#20) has been shown feasible through the experimental feasibility model, but insufficient time was devoted to finding the best conditions.

The coordinate engineering drawing compression (#21) is becoming more practical as large companies orient their drawing standards for automated procedures. By establishing some simple restraints on line drawings including a limit of the number of labels allowed, it is possible to use optic code compression (#22) following the theorem developed by Weiss. In considering biological systems (#23) it is instructive to note the compression obtainable in nature under conditions transmitting the outline of maps, but not retaining internal fine structure.

When funds are again available for further studies of image transmission, it would be useful to use the feasibility model to collect new statistics and to test alternative compression schemes.

Description of a Scanner as Simulated by a Fortran Program

P. R. Daher - 9 Apr. 62

1 - Introduction

The variable-speed document scanner has the capability of skipping lines containing no data. Edge effects and fly-back time are neglected. Fig. 1 shows a diagram of the scanner. A is the main scanner, while B is the horizontal pre-scanner (assumed to be the same size as A).

The scanner progresses at high speed until a transition occurs (in this case, the transition will not involve shades of gray); at this time it proceeds at slow speed until both A and B receive the same type of information - fast speed is then resumed. In this case, then, the slow speed will be in effect for one bit (resolution) space. A truth-table summary of speed change conditions is shown in Table 1.

The vertical scanners, labeled 2, 4, 8, and 16 (Fig. 1) encompass the indicated number of spaces as measured in "A" units. As the scanner progresses horizontally, the vertical scanner "looks" for data. The number of lines skipped on fly-back depends on which unit detects the data. If "2" detects data, no lines will be skipped; if "4" does, 2 will be skipped; if "8" does, 6 will be skipped; if "16" does, 14 will be skipped; if no data is detected, then 30 lines will be skipped. See Table 2.

Some means, such as a mechanical stop (or the detection of the scanning spot at a given position), is necessary for detecting the end of a row. A similar requirement exists for the end of a page. In the program, the "physical distance" travelled by the scanner was limited in either direction.

The output information was coded in the following manner. At the end of a given row, sets of coordinates were generated summarizing the following:

- a - the vertical skip distance, and
- b - for each speed change, a summary of the speed value, data value, and distance travelled since the previous change.

The program was sufficiently complicated so that the second part was not attempted in the time allotted. This would have been to simulate a receiver which accepts the data produced above and converts it into the original image.

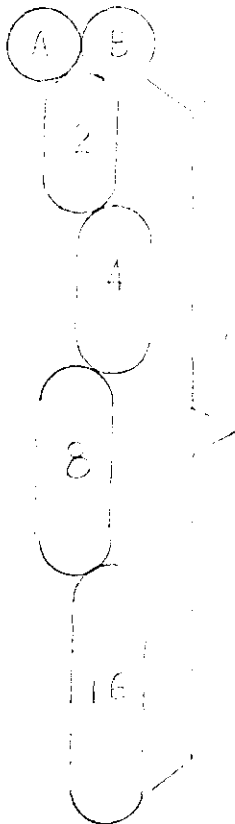
2, II - The Program

Fig. 2 shows the image set up in an array, A_{jk} . Each square represents one unit of scanner resolution. The program may be found in Section I of the enclosure. Comments are written in pencil on the program to help in its interpretation. Section II shows the output. The columns are labelled and comments inserted as required.

Reference: Electronics, V. 31 #31, Sept. 26, 1958, pp. 84 - 88,
- - - a system developed by RCA - - stops momentarily at a transition.

Figure A.1

The Scanner



Vertical Skip
Prescanners

Table A-1

<u>Horizontal speed relationships:</u>	<u>Data Value</u>		<u>Previous Speed</u>	<u>Next Speed</u>
	<u>A</u>	<u>B</u>		
	0	0	Slow	Fast
	0	0	Fast	Fast
	0	1	Slow	Slow
	0	1	Fast	Slow
	1	0	Slow	Slow
	1	0	Fast	Slow
	1	1	Slow	Fast
	1	1	Fast	Fast

0 = blank

1 = black, i.e., when the minimum detectable black area occurs.

Table 2

<u>Vertical skip relationships:</u>	<u>Smallest area picking up data</u>					<u>Skip Distance</u> <u>(No. spaces)</u>
	<u>2</u>	<u>4</u>	<u>8</u>	<u>16</u>	<u>None</u>	
	X					0
		X				2
			X			6
				X		14
					X	30

(1) (2) (3) (4) (5) (6) (7) (8) (9) (10)

(11) (12) (13) (14) (15) (16) (17) (18) (19) (20)

(21) (22) (23) (24) (25) (26) (27) (28) (29) (30)

(31) (32) (33) (34) (35) (36) (37) (38) (39) (40)

(41) (42) (43) (44) (45) (46) (47) (48) (49) (50)

3. Appendix - A General System

Fig. 2 shows a general system (shades of gray included). The scanner progresses at high speed until a transition occurs. The speed will then be reduced in accordance with some relationship to the change in data value.

For example, let $S_h = f(I)$, say $S_h = S \max - K \frac{dI}{dt}$

S_h = Scanning speed

$S \max$ = Maximum speed to be used for blank or solid areas

where

K = Some constant to regulate the sensitivity of S_h to $\frac{dI}{dt}$

I = Intensity of light from the scanned image

Some questions present themselves:

1. How can the slope of the transmitted data cause the receiver scanner to track at the proper speed?
2. Can the slope also be used to vary the intensity of the reproduced image?
3. Should $K = f(S \max)$?
4. How can coding be used to effectively compress messages?
5. Should the scanner output be analog or digital?

Some suggestions regarding the vertical scanner:

1. Use bundles of fibers; logic circuits can provide the decisions regarding skip distances.
2. It may be possible to use larger fibers instead of bundles of smaller ones.

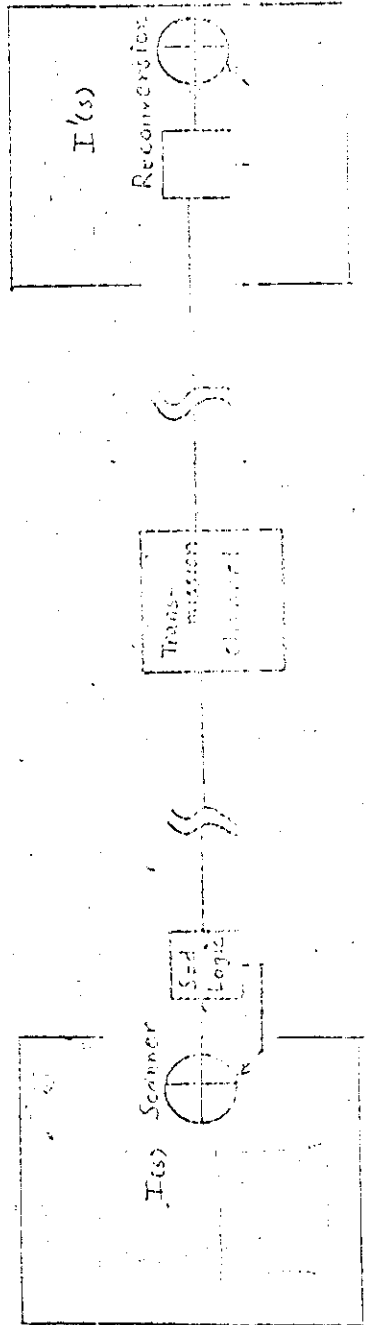


Figure A-3
An Adaptive Scanning System

Appendix B - Modulation Systems

The project did not reach the stage of a systematic study of modulation systems. However the following modulation systems were briefly reviewed:

- (1) Amplitude Modulation - Double Side Band (AM - DSB)
- (2) Amplitude Modulation - Single Side Band (AM-SSB)
- (3) Frequency Modulation (FM and FSK)
- (4) Phase Modulation (PM)
- (5) Phase Modulation - Vestigial Sideband (PM - VSB)
- (6) Delta Modulation (AM)
- (7) Amplitude and Angle Modulation - Single Sideband (AAM - SSB)

For an earlier comparison of different modulation, see the paper by Jelonek at the Second London Symposium in 1952 ("A Comparison of Transmission Systems", in Willis Jackson, Communication Theory, London: Butterworths (1953), pp 44 - 81).

There is a 1958 paper by George J. Kelley which has limited application to this problem but is a good example of the comparison of different modulation systems ("Choosing the Optimum Type of Modulation - A Comparison of

Several Communication Systems", IRE Trans. Comm. Systems, Vol. CS-6, No. 1, June 1958, pp 14 - 21) and includes a useful bibliography.

The work on phase modulation systems in the San Jose IBM-ASDD Laboratory is summarized in reports by E. Hopner ("An Experimental Modulation - Demodulation Scheme for High Speed Data Transmission", IBM Journal, pp 74 - 84, Jan 1959; "Phase Reversal Data Transmission System for Switched and Private Telephone Line Applications", pp 93 - 105, Vol S, No. 2, April 1961), C. M. Melas ("A Transistorized Data Set for Operation over Telephone Lines", IBM Report RJ-159, May 20, 1959).

The development of a vestigial sideband phase modulation system with injected carrier has been reported by G. A. McAuliffe ("Vestigial Sideband Modern", MARP TASK 0903, 3-22-62).

In Deltamodulation the information signal is converted into a sequence of binary pulses in such a manner that a reconstruction of it is obtained by applying the series of binary pulses to a linear network. Essentially the pulses are a measure of the change in input signal and offer the prospect of

integrating the scanning system and the modulation system for a predictive system which transmits the changes in image signal or the difference between the actual signal and a predicted signal. For a detailed description of "delta-modulation" refer to the paper by F. de Jager ("Deltamodulation.- A Method of PCM Transmission Using the 1-Unit Code", in Communication Theory, edited by Willis Jackson, London: Butterworths (1953), pp 119 - 137) and the paper by Lars-Henning Zetterberg ("A Comparative Study of Delta and Pulse Code Modulation", Information Theory (Third London Symposium, 1955) edited by Colin Cherry, London: Butterworths (1956) pp 103 - 110).

P. D. Dodd has proposed a technique using Amplitude and Angle Modulation on single sideband where the signal and its harmonic conjugate are Hilbert transfer of each other by a carrier of a system proposed by K. H. Powers ("The Compatibility Problem in Single-Sideband Transmission", Proc. I.R.E., Vol 48, pp 1431 - 1436, Aug. 1960; SJ-IBM-Lib No. 42983) from the analog case to the digital case.

Powers demonstrated that conveying the analog message in the square of the envelope permits a phase function to be found for which the hybrid wave

occupies a width equal to a conventional single sideband. Dodd has pointed out that simple implementation is required for the corresponding digital case and further that the phase information could be used as a redundancy check.

SPECTRA OF SOME PULSE SIGNALS

Norman Abramson

There are two primary theoretical questions which must be attacked in order to analyze and synthesize systems capable of efficient image transmission and storage. These questions are:

- (a) "What are the possible methods of encoding redundant two-dimensional images into one-dimensional waveforms?" and
 (b) "What are the bandwidth requirements arising out of a solution to question (a)?"

In this note we shall attack the second of these questions.

We shall investigate the bandwidth required by a large variety of pulse signals which might arise from a solution of the coding problems. Clearly the solutions of the two problems we have outlined are not independent. The coding method we shall want to use will depend upon the bandwidth requirements of the output of the encoder. The bandwidth in turn will depend directly upon the coding method selected. An analysis of the bandwidth required by some possible encoder outputs can be expected to suggest efficient coding methods.

First we establish some notation and elementary results which we shall use in the rest of this note. The Fourier transform of a function $g(t)$ will be denoted by $G(f)$

$$G(f) = \int_{-\infty}^{\infty} g(t) e^{-i\omega t} dt \quad (1)$$

$\omega = 2\pi f$

$$g(t) = \int_{-\infty}^{\infty} G(f) e^{i\omega t} df \quad (2)$$

A sample function of a stationary random process, $x(t)$, does not have a Fourier transform in the usual sense. We define the autocorrelation function of a stationary random process as

$$R_x(\tau) = E[x(t+\tau) x(t)] \quad (3)$$

where $E[\cdot]$ indicates an expected or average value. The transform of $R_x(\tau)$ is $S_x(f)$, the power spectral density of $x(t)$. $S_x(f)$ provides an indication of the bandwidth necessary to transmit the random process $x(t)$. If $x(t)$ and $y(t)$ are sample functions of two independent stationary random processes and

$$z(t) = x(t) y(t) \quad (4)$$

$$\text{then } R_z(\tau) = R_x(\tau) R_y(\tau) \quad (5)$$

$$\text{and } S_z(f) = S_x(f) * S_y(f) \quad (6)$$

where $*$ denotes convolution.

Sampling Theorem for Known Functions: Let $g(t)$ have a Fourier Transform $G(f) = 0$ for $|f| > W$.

Then

$$g(t) = \sum_n g_n \text{ sinc } 2W \left(t - \frac{n}{2W} \right) \quad (7)$$

where

$$g_n = g \left(\frac{n}{2W} \right) \quad (8)$$

and

$$\text{sinc } x = \frac{\sin \pi x}{\pi x} \quad (9)$$

Sampling Theorem for Random Functions: Let $x(t)$ be a sample function of a random process with $S_x(f) = 0$ for $|f| > W$. Then

$$x(t) = \sum_n x_n \text{ sinc } 2W \left(t - \frac{n}{2W} \right) \quad (10)$$

where

$$x_n = x \left(\frac{n}{2W} \right) \quad (11)$$

Now assume we have reduced our two-dimensional image to a simple voltage as a function of time by scanning. We can determine the pulse rate necessary to represent this voltage by a direct application of the sampling theorem. If there are no frequencies greater than W cps in the voltage then the voltage can be represented by a sequence of values taken at the Nyquist rate of $2W$ per second. We are interested therefore in determining the power spectral density of random processes of the form

$$y(t) = \sum_n x_n p(t - \frac{n}{2W} + \Theta) \tag{12}$$

where the x_n are samples of the random process $x(t)$ representing our image

$$x_n = x(\frac{n}{2W} - \Theta) \tag{13}$$

where $p(t)$ is some arbitrary pulse shape and where Θ is a random variable introduced only to provide a random phase. The probability density of Θ is assumed uniform from 0 to $\frac{1}{2W}$. For example, if the pulse $p(t)$ is narrow and rectangular $y(t)$ might look like

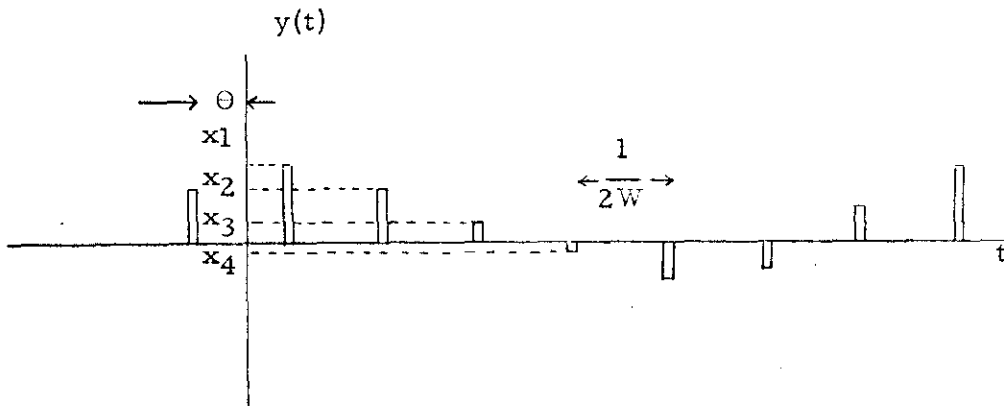


Figure 1 EXAMPLE OF $y(t)$

It is of course also possible for $p(t)$ to be a wide pulse (compared to the pulse separation $\frac{1}{2W}$) so that pulse overlap occurs. We are also interested in the spectral density of $y(t)$ for various statistics of the x_n and for various degrees of correlation between the successive x_n . In order to determine the spectral density $S_y(f)$ we use the trick of introducing the new random process $\Delta(t)$ composed of a periodic train of impulse functions $\delta(t)$ of random phase

$$\Delta(t) = \sum_n \delta(t - \frac{n}{2W} + \Theta) \quad (14)$$

The autocorrelation function of this random process is

$$R_{\Delta}(\tau) = 2W \sum_n \delta(\tau - \frac{n}{2W}) \quad (15)$$

and the power spectral density

$$S_{\Delta}(f) = (2W)^2 \sum_n \delta(f - 2nW) \quad (16)$$

The reason for introducing $\Delta(t)$ is that now we may write $y(t)$ as

$$\begin{aligned} y(t) &= \sum_n x_n p(t - nT + \Theta) \\ &= p(t) * \sum_n x_n \delta(t - nT + \Theta) \\ &= p(t) * [x(t) \sum_n \delta(t - nT + \Theta)] \\ &= p(t) * [x(t) \Delta(t)] \end{aligned} \quad (17)$$

Now using (6) we can immediately obtain $S_y(f)$. Let $P(f)$ be the Fourier transform of $p(t)$.

$$S_y(f) = |P(f)|^2 [S_x(f) * S_{\Delta}(f)] \tag{18}$$

The convolution (18) is easily performed since $S_{\Delta}(f)$ consists of a sequence of δ functions.

$$S_y(f) = |2W P(f)|^2 \sum_n S_x(f - 2nW) \tag{19}$$

Example: To illustrate the application of (19) let the x_n be samples of a white-bandlimited random process with

$$S_x(f) = \begin{cases} \frac{1}{2W} & |f| \leq W \\ 0 & \text{elsewhere} \end{cases} \tag{20}$$

and let $p(t)$ be a rectangular pulse of width T

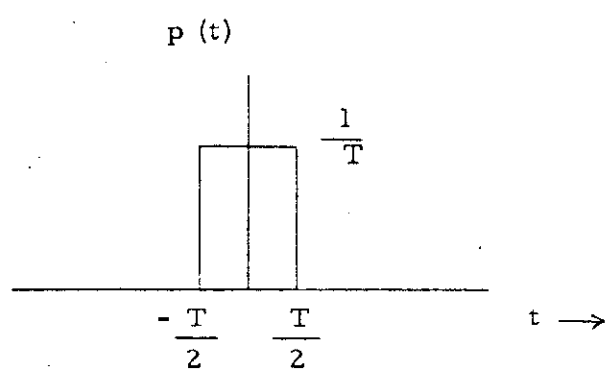


Figure 2 RECTANGULAR PULSE

Then

$$P(f) = \text{sinc } fT \quad (21)$$

and (19) becomes

$$\begin{aligned} S_y(f) &= (2W \text{sinc } fT)^2 \sum_n S_x(f - 2nW) \\ &= 2W \text{sinc}^2 fT \end{aligned} \quad (22)$$

Equation (22) is sketched in figure 3

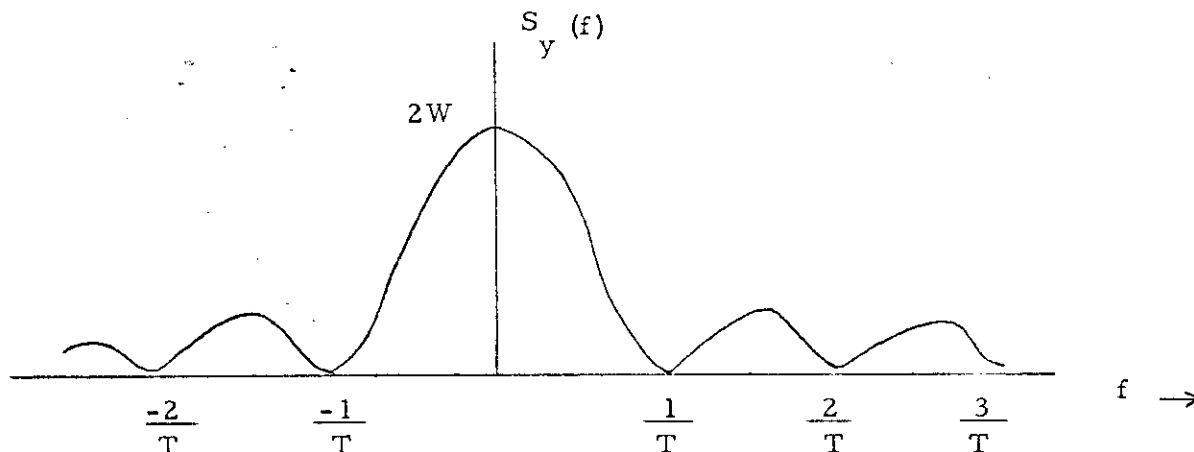


Figure 3 - EXAMPLE OF: POWER SPECTRAL DENSITY

For some applications a more useful form of the power spectral density is obtained by first calculating $R_y(\tau)$. The Fourier transform of $|P(f)|^2$ is $\int_{-\infty}^{\infty} p(t) p(t+\tau) dt$. We define

$$\int_{-\infty}^{\infty} p(t) p(t+\tau) dt \triangleq \rho(\tau) \quad (23)$$

Then we transform (18) to get

$$\begin{aligned}
 R_y(\tau) &= \rho(\tau) * [R_x(\tau) R_{\Delta}(\tau)] \\
 &= \rho(\tau) * [R_x(\tau) \cdot 2W \sum_n \delta(\tau - \frac{n}{2W})] \\
 &= 2W \rho(\tau) * \sum_n R_x(\frac{n}{2W}) \delta(\tau - \frac{n}{2W}) \\
 &= 2W \sum_n R_x(\frac{n}{2W}) \rho(\tau - \frac{n}{2W}) \quad (24)
 \end{aligned}$$

Examples: Several examples are possible to demonstrate the utility of (24). First consider the case of independent pulses with the x_n chosen according to some density $p(x_n)$. Then

$$\begin{aligned}
 R_x(0) &= E[x_n^2] \\
 &= \sigma_x^2 + (\bar{x})^2 \quad (25)
 \end{aligned}$$

where σ_x^2 is the variance of the x_n and \bar{x} is the mean of x_n . We also have

$$\begin{aligned}
 R_x(\frac{n}{2W}) &= E[x(t + \frac{n}{2W}) x(t)] \\
 &= E[x(t + \frac{n}{2W})] E[x(t)] \\
 &= (\bar{x})^2 \quad (26)
 \end{aligned}$$

Now, using (25) and (26) in (24) we get (for independent pulses)

$$R_y(\tau) = 2W \sigma_x^2 \rho(\tau) + 2W (\bar{x})^2 \sum_n \rho(\tau - \frac{n}{2W}) \quad (27)$$

Note that the second term on the right of (27) is periodic. Therefore this term will lead to discrete spectral lines at frequency W and harmonics of W in the power spectral density. Note further that in the common case where $\bar{x} = 0$ this periodic part vanishes, $\sigma_x^2 = \bar{x}^2$ and we are left with

$$\begin{aligned} R_y(\tau) &= 2W \frac{\sigma_x^2}{\bar{x}^2} \rho(\tau) \\ &= 2W \frac{\sigma_x^2}{\bar{x}^2} \int_{-\infty}^{\infty} p(t) p(t + \tau) dt \end{aligned} \quad (28)$$

The power spectral density then takes the simple form

$$S_y(f) = 2W \frac{\sigma_x^2}{\bar{x}^2} |P(f)|^2 \quad (29)$$

As a further example of (24) we take the case where some correlation exists between the amplitudes of adjacent pulses but the amplitudes of non-adjacent pulses are independent. For simplicity we shall take $\bar{x} = 0$. Assume the correlation is given by

$$\begin{aligned} E[x_m^2] &= \frac{\sigma_x^2}{\bar{x}^2} \\ E[x_n x_{n+1}] &= a \frac{\sigma_x^2}{\bar{x}^2} \\ E[x_n x_{n+j}] &= 0 \quad j \geq 2 \end{aligned} \quad (30)$$

Then (24) becomes

$$R_y(\tau) = 2W \frac{\sigma_x^2}{\bar{x}^2} \rho(\tau) + 2W a \frac{\sigma_x^2}{\bar{x}^2} \left[\rho(\tau - \frac{1}{2W}) + \rho(\tau + \frac{1}{2W}) \right] \quad (31)$$

Transforming (31) yields

$$\begin{aligned}
S_y(f) &= 2W \bar{x}^{-2} |P(f)|^2 + W a \bar{x}^{-2} |P(f)|^2 \cos 2\pi \frac{f}{2W} \\
&= 2W \bar{x}^{-2} |P(f)|^2 \left[1 + \frac{a}{2} \cos \frac{\pi f}{W} \right] \quad (32)
\end{aligned}$$

Comparison of (32) with (29) shows the effect on the power spectral density of introducing correlation between adjacent pulses.



Another type of signal which may be of some use in transmitting and storing image information is the random telegraph signal. This signal has been analyzed in several books (e.g., "Random Signals and Noise", by Davenport and Root). We shall define the random telegraph signal and then develop a method of analyzing a generalization of this signal which appears to have some uses in image transmission

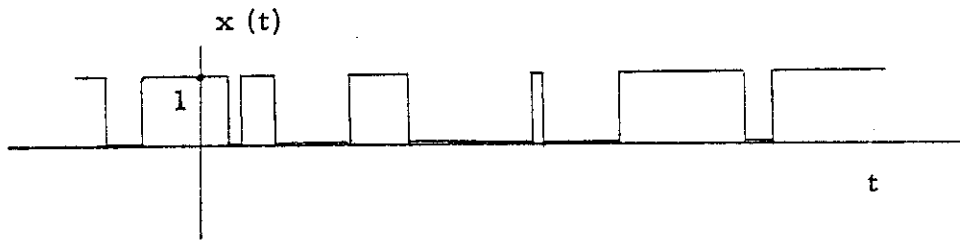


Figure 4 - RANDOM TELEGRAPH SIGNAL

The random telegraph signal is assumed to oscillate between the two states "zero" and "one" at an average rate of a transition per second. The transitions are assumed to occur completely at random; more

precisely the occurrence of transition times obeys the Poisson probability law. If $P(k; \tau)$ is the probability that we will have k transitions (zero to one or one to zero) in a time interval τ , then

$$P(k; \tau) = \frac{(a\tau)^k}{k!} e^{-a\tau} \tag{33}$$

Davenport and Root (page 62) show that

$$R_x(\tau) = \frac{1}{4} + \frac{1}{4} e^{-2a|\tau|} \tag{34}$$

and also that (page 104)

$$S_x(f) = \frac{1}{4} \delta(f) + \frac{a}{4a^2 + \omega^2} \tag{35}$$

We define a generalized random telegraph signal by allowing $x(t)$ to assume one of a continuum of possible values rather than merely zero and one.

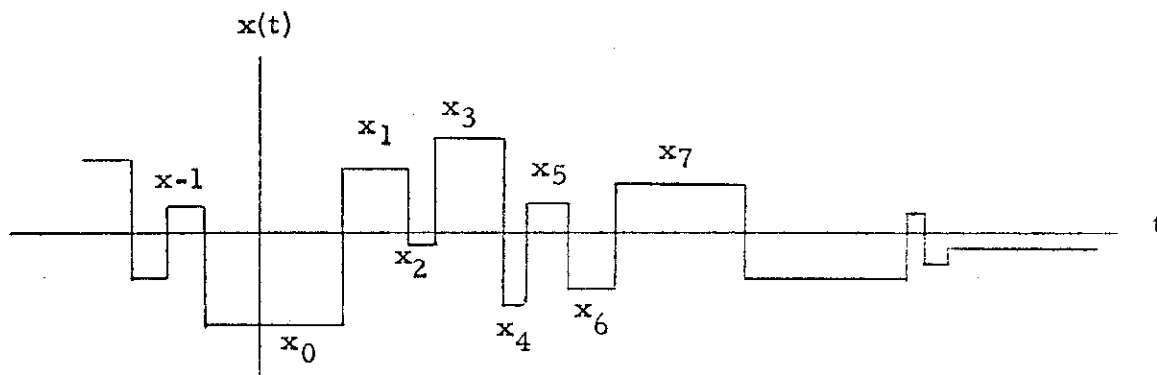


Figure 5 - GENERALIZED RANDOM TELEGRAPH SIGNAL

The values of $x(t)$ after successive transitions are labelled $x_{-1}, x_0, x_1, x_2, \dots$. The x_n are assumed to be random variables

with some density $p(x)$ giving a mean \bar{x} and variance σ_x^2 . It is not assumed that the x_n are necessarily independent. We assume that the correlation between x_i and x_j decreases exponentially with the difference $|i - j|$. The correlation between x_i and x_j is

$$\begin{aligned} \rho_{ij} &= E[(x_i - \bar{x}_i)(x_j - \bar{x}_j)] / \sigma_x^2 \\ &= \rho^{|i-j|} \end{aligned} \tag{36}$$

where ρ is some constant satisfying $|\rho| \leq 1$.

Note that (since $\bar{x}_i = \bar{x}_j = \bar{x}$)

$$\sigma_x^2 \rho_{ij} = E[x_i x_j] - (\bar{x})^2 \tag{37}$$

In order to find $R_x(\tau)$ we focus our attention on the number of transitions of $x(t)$ in an interval τ . For $\tau \geq 0$

$$\begin{aligned} R_x(\tau) &= E[x(t+\tau)x(t)] \\ &= P(0;\tau) E[x_i^2] + P(1;\tau) E[x_i x_{i+1}] \\ &\quad + P(2;\tau) E[x_i x_{i+2}] + \dots \\ &= e^{-a\tau} [\sigma_x^2 + (\bar{x})^2] + \frac{a\tau}{1!} e^{-a\tau} [\sigma_x^2 \rho + (\bar{x})^2] \\ &\quad + \frac{(a\tau)^2}{2!} e^{-a\tau} [\sigma_x^2 \rho^2 + (\bar{x})^2] \\ &\quad + \frac{(a\tau)^3}{3!} e^{-a\tau} [\sigma_x^2 \rho^3 + (\bar{x})^2] + \dots \\ &= e^{-a\tau} \left[\sigma_x^2 + \frac{\rho a\tau}{1!} \sigma_x^2 + \frac{(\rho a\tau)^2}{2!} \sigma_x^2 + \dots \right] \end{aligned}$$

$$\begin{aligned}
& + e^{-a\tau} \left[(\bar{x})^2 + \frac{a\tau}{1!} (\bar{x})^2 + \frac{(a\tau)^2}{2!} (\bar{x})^2 + \dots \right] \\
= & \sigma_x^2 e^{-a\tau} e^{\rho a\tau} + (\bar{x})^2 e^{-a\tau} e^{a\tau} \\
= & \sigma_x^2 e^{-a\tau (1 - \rho)} + (\bar{x})^2
\end{aligned} \tag{38}$$

The previous calculation was valid for $\tau \geq 0$. Since $R_x(\tau)$ is an even function of τ we have

$$R_x(\tau) = \sigma_x^2 e^{-a|\tau|(1-\rho)} + (\bar{x})^2 \tag{39}$$

The classical random telegraph is of course a special case of the generalized random telegraph signal with $\bar{x} = \frac{1}{2}$, $\sigma_x^2 = \frac{1}{4}$ and $\rho = -1$. Note that with these values of \bar{x} , σ_x^2 and ρ (39) reduces to (34).

To find the power spectral density of $x(t)$ we transform (39)

$$S_x(f) = (\bar{x})^2 \delta(f) + \frac{2(1-\rho)a\sigma_x^2}{\omega^2 + (1-\rho)^2 a^2} \tag{40}$$

We may obtain several useful facts from (40). First the mean value of x will affect only the first term -- $(\bar{x})^2 \delta(f)$ -- corresponding to the DC power of $x(t)$. Second, even the variance σ_x^2 has no significant effect on the bandwidth of $x(t)$. The variance appears only as a scale factor on the second part of $S_x(f)$. Finally note that the correlation between the x_n has a most direct effect on the bandwidth. We can obtain quantitative measures of this effect from (40). For a qualitative measure of the effect of ρ we sketch $S_x(f)$ below in figure 6.

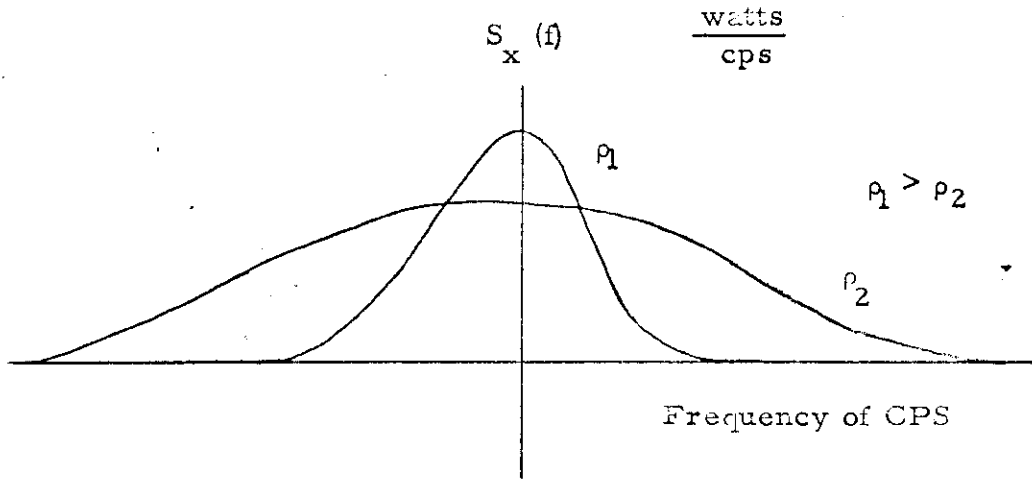


Figure 6 - POWER SPECTRAL DENSITY OF GENERALIZED
RANDOM TELEGRAPH SIGNALS FOR TWO VALUES
OF ρ

$\Phi_{ws}(f) \approx$ watts per cycle of bandwidth

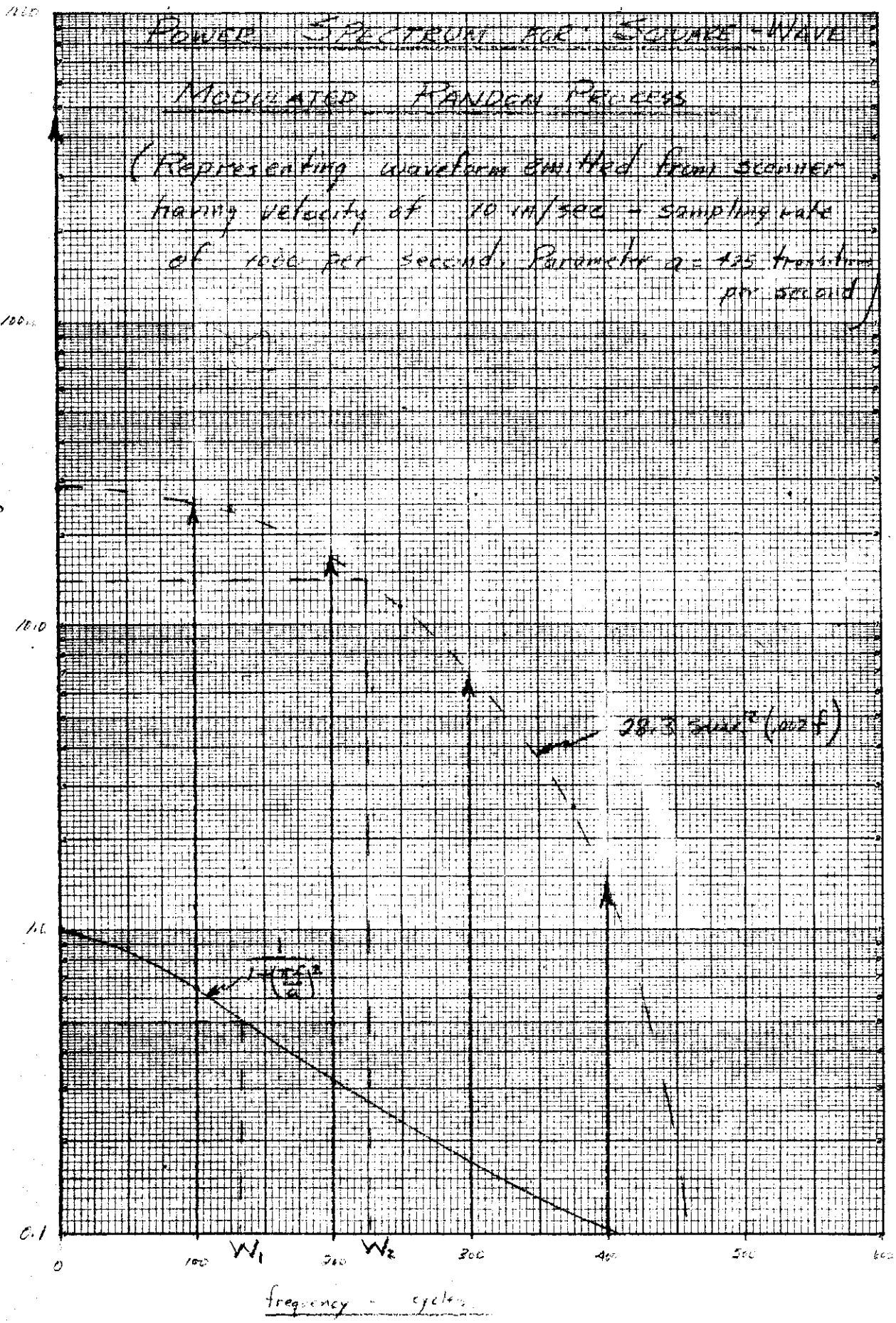


FIGURE D.8