VIII. PROCESSING AND TRANSMISSION OF INFORMATION

Prof. E. Arthurs
Prof. P. Elias
Prof. R. M. Fano
Prof. J. Granlund
Prof. D. A. Huffman
Prof. M. L. Minsky
Prof. H. Rogers, Jr.
Prof. C. E. Shannon (absent)
Prof. W. M. Siebert
Prof. J. M. Wozencraft
R. M. Bevensee
R. K. Brayton
D. C. Coll
J. E. Cunningham
M. A. Epstein
E. F. Ferretti
R. G. Gallager
T. L. Grettenberg
P. R. Hall
F. C. Hennie III
E. M. Hofstetter
T. Kailath
D. C. Luckham
R. S. Marcus
J. Max
G. E. Mongold, Jr.
G. E. W. Sparrendahl (visitor)
W. A. Youngblood
H. L. Yudkin

A. THE FUNCTION DISPLAY PROGRAM

1. Introduction

The Function Display Program (1) is designed to facilitate the investigation of computer-human synergism – in particular, develop a four-dimensional intuition by such close cooperation. This program provides a means for displaying plots of cross sections of a function of three variables, $f(x_1, x_2, x_3)$, on the cathode-ray tube of the IBM 704 computer. The operator can vary the range and magnification of these plots by appropriate use of the sense switches.

Typical objectives for the operator might be to locate relative maxima or minima or to investigate points that might be singular points. It should be pointed out that the distinctive feature of this program is the rapid communication link from the human being to the computer and from the computer back to the human being. It is desirable to study the efficiency and possible advantages of such computer-human systems. We hope, for example, that the operator of the program will develop an intuition or "feeling" for the four-dimensional surface that a function represents. Then, too, visual inspection and human guidance might speed up the search for relative maxima or minima, especially in badly conditioned cases. There are other problems, especially those of design, in which the displaying of intermediate results, for the purpose of human intervention, could lead to speedier solutions. If this program is a success, a person will be said to have "learned" about a function with the aid of a computer. How much the learning process can be speeded up, in general, by the use of a computer is an interesting question (2).

2. General Description

Any function of three (or fewer) variables can be displayed, as long as a subroutine exists for calculating the value of the function (in floating point) for any values of the independent variables. The function may be thought of as a surface in four-dimensional space – three dimensions for the three independent variables and a fourth for the value of the function. Initially, we consider the region of the four-dimensional space for which $a_1 \leq x_1 < b_1$, $a_2 \leq x_2 < b_2$, $a_3 \leq x_3 < b_3$, $L \leq y = f(x_1, x_2, x_3) \leq M$; that is, for a four-dimensional rectangle. The initial values of the parameters $a_1, b_1, a_2, \ldots, M$ are read in with the program. The operator, by properly setting the sense switches, chooses any two of the independent variables to be held constant temporarily. One of the variables, x_i, that is so chosen is assigned the value $x_i = (a_i + b_i)/2$. The function is then calculated and plotted on the cathode-ray tube as a function of the other independent variable over its range. If for any point the value of the function is greater than M or less than L, the plot is made as if $y = M$ or $y + 1$, respectively.

The operator, on viewing the plot, may decide to move, to expand, or to contract the plot. He can move the plot to the left or right, up or down, or expand or contract it in the x or y directions, and each of these operations can be fast or slow. For each kind of action there is an appropriate setting of the sense switches. Each action changes either a and b (for the current variable) or M and L; and new calculations of the function may be needed. Expansion or contraction is accomplished by holding the center of the x or y range fixed and varying the magnitude of the range. Each action is performed repeatedly until the appropriate "action" switch is pushed down. The two action switches are called "move" and "size" (expand or contract). The other four sense switches (called the "description" switches) are used to describe the manner in which the action is to be performed (as well as to show which is the current variable). Thus the observer sees the plot continually moving or changing size when an action switch is pushed up. The current variable can be changed by properly setting the description switches, and then the action switches. Thus a complete set of operations for viewing any region of the function is provided.

3. Comments on Operation of and Improvements For the Function Display Program

The program has been coded and tested, and various difficulties have been ironed out. It works properly and will be submitted as a SHARE distributed subroutine.

In carrying out the program we found it easier to make a detailed investigation of a region that was entirely within the specified x and y ranges rather than of a region that was not so restricted. For instance, it was easier to locate relative maxima or minima than to investigate poles.

To improve the practicality of this program, it is advisable to include routines for recording certain information on film or typed print-outs. In future routines it may be well to use the fact that people already have good three-dimensional intuition by drawing three-dimensional displays with the use of perspective techniques.

Experience in the Functional Display Program indicates that the 704 computer lacks some facilities for direct communication with the operator. Some suggested improvements are: more sense lights and sense switches, including switches of the momentarily depress type; a direct input typewriter, a "light pen" to sense whether or not a particular spot on the cathode-ray tube is illuminated; and inputs through controls set in analog fashion should be provided. With such devices ranges could be set, and interesting portions of the display (including cross sections in which more than one variable is varying) could be picked out much more rapidly and conveniently.

4. Conclusions

We have not been able to satisfactorily evaluate the efficiency of the computer-human synergism or the possibility of developing four-dimensional intuition from this program because of the small amount of computer time that has thus far been used for running the program. A reasonable evaluation can only be made after several hours of operation by different operators under controlled conditions. Indeed, an evaluation of such an ill-defined concept as "intuition formation" would be exceedingly difficult in any case.

Nevertheless, some statements can be made. The program does allow the operator to view any permissible cross section of a function in a fairly rapid manner, if we consider the limitations of the equipment which have been mentioned. In this sense, the synergism part of the program has been demonstrated: There is a rapid, convenient (for both human being and computer) communication link that carries information in both directions between man and the machine. Neither half of the system has to wait too long for the other to do its job.

It is harder to say anything about the usefulness of the program. It should be pointed out that this is only a first attempt, and many improvements are possible, as we have indicated. Still, it seems clear that this program, or a refined version of it, represents a powerful and untried tool for the researcher which might enhance his intuition and extend his learning capability.

5. Acknowledgment

The idea for a function display program was suggested by Stanislaw Ulam, of the Los Alamos Laboratories. The work was supervised by Professor Dean N. Arden and Professor Edward Arthurs and supported by the Department of Electrical Engineering and the Research Laboratory of Electronics, M. I. T. The facilities of the Computation Center, M. I. T., were used in the work (their program number for this work is M235).

R. S. Marcus

References

1. For a detailed account of this program, see R. S. Marcus, The function display program, Computation Center Report, M. I. T., May 1958.
2. R. S. Marcus, Computers, humans and intelligence, Computation Center, M. I. T., June 1958 (unpublished).

B. SAMPLING THEOREMS FOR LINEAR TIME-VARIANT SYSTEMS

Under certain conditions, a sampling theorem and a corresponding delay-line model can be found for linear time-variant filters. These filters may prove useful in studying communication through multipath media.

First, we introduce some definitions and assumptions, after which the sampling theorem is derived for a particular case; from the sampling theorem a delay-line model is deduced. Finally, the sampling theorem and delay-line model for a second case are briefly discussed.

1. Definitions

One method of characterizing a linear time-variant system is by its impulse response. This is usually given as $h_1(t, \tau)$ – the response at time t to an impulse input at time τ. An equivalent description, which, as will be explained later, is more useful for our purpose, is $h_2(z, \tau)$ – the response after z seconds to an impulse input at time τ. The relation of these forms is given by

$$h_1(t, \tau) = h_2(t - \tau, \tau)$$

$$h_2(z, \tau) = h_1(z + \tau, \tau)$$

For $h_2(z, \tau)$ we have the Fourier transforms

$$H_2(z, j\mu) = \int_{-\infty}^{\infty} h_2(z, \tau) \exp(-j2\pi\mu\tau)\, d\tau$$

$$H_2(j\omega, \tau) = \int_{-\infty}^{\infty} h_2(z, \tau) \exp(-j2\pi\omega z)\, dz$$

$$H_2(j\omega, j\mu) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(z, \tau) \exp(-j2\pi\omega z) \exp(-j2\pi\mu\tau)\, dz\, d\tau$$

where μ and ω have the dimensions of cycles per second. The variable μ corresponds to the system variation; to this variation we assign a lowpass bandwidth W_o. W_o is a measure of the rate at which the system is varying: if $h_2(z, \tau)$ changes rapidly with τ, W_o will be high; if $h_2(z, \tau)$ does not change with τ (i.e., if the system is time-invariant), W_o is zero. On the other hand, ω is the usual frequency variable corresponding to the elapsed time and is the only one that arises in a time-invariant system. The advantage of specifying the system by $h_2(z, \tau)$ is that this direct physical interpretation of the Fourier transform variables is not possible with $h_1(t, \tau)$. With $h_1(t, \tau)$ we have the complication that $t > \tau$, and it is hard to visualize the physical meaning of its Fourier transform with respect to t, with τ fixed. The introduction of $h_2(z, \tau)$ gives the mathematically more tractable constraint of $z > 0$, and now the Fourier transform with respect to z, with τ fixed, has the significance of the usual frequency-domain variable.

Given $h_2(z, \tau)$ and $H_2(j\omega, j\mu)$, we can make different assumptions about z and τ, ω and μ – for example, z time-limited, μ band-limited – and derive sampling theorems. Two practically useful cases, which are the only ones considered here, are

a. ω band-limited $(-W, W)$ and μ band-limited $(-W_o, W_o)$, which is the lowpass case.

b. ω band-limited $\left(f_c - \frac{W}{2}, f_c + \frac{W}{2}\right)$ and μ band-limited $(-W_o, W_o)$, which is the highpass case.

The assumption of band-limited ω can be interpreted as meaning that we are interested in the impulse response of the system over only a finite specified bandwidth, and the delay-line model and sampling theorem that we shall derive will be valid representations of the system over this bandwidth only. This assumption of finite bandwidth may sometimes be necessary because of the limitations of our measuring equipment.

We shall need a pair of definitions for defining Fourier-transform pairs for periodic functions (1):

$$\text{rep}_T\, h(t) = \sum_{n=-\infty}^{\infty} h(t - nT)$$

$$\text{comb}_{1/T} H(f) = \sum_{n=-\infty}^{\infty} H\left(\frac{n}{T}\right) \delta\left(f - \frac{n}{T}\right)$$

By using a Fourier-series expansion of $\text{rep}_T\, h(t)$, we can derive the relation

$$F\{\text{rep}_T\, h(t)\} = \left|\frac{1}{T}\right| \text{comb}_{1/T}\, H(f)$$

That is to say, if a nonperiodic function $h(t)$, which has a transform $H(f)$, is shifted in time by all integral multiples of T and the results are added together, the spectrum of the resulting periodic function will be obtained by picking out the values of $H(f)$ at intervals $1/T$. And conversely,

$$F\{\text{comb}_T\, h(t)\} = \left|\frac{1}{T}\right| \text{rep}_{1/T}\, H(f)$$

Another useful pair of Fourier transforms consists of the rectangular function and its spectrum. Woodward uses the notation

$$\text{rect}\; t = \begin{cases} 1, & |t| < 1/2 \\ 0, & |t| > 1/2 \end{cases}$$

for the pulse, and

sinc f = sin $\pi f/f$ for its spectrum.

2. Sampling Theorem

The method of deriving sampling theorems differs according to whether the region of interest is a lowpass or a highpass frequency range. In both cases, however, it is convenient to use Woodward's compact notation and method of deriving sampling representations (1). This may be regarded as a translation into compact analytical form of the point of view that regards sampling as impulse modulation (2). The great advantage of this view is that it enables the use of conventional Fourier analysis. Although Woodward's method may appear, at first sight, somewhat artificial, a physical justification of the steps can be understood by bearing in mind that sampling is analogous to impulse modulation.

3. Lowpass Case

We assume that the frequency ω is restricted to a band (W, W) around the origin (3). Since μ is restricted to the range $(-W_o, W_o)$, then, clearly, we can write

$$H_2(z, j\mu) = \text{rep}_{2W_o}\, H_2(z, j\mu)\, \text{rect}\, \frac{\mu}{2W_o}$$

Transforming both sides gives

$$h_2(z, \tau) = \text{comb}_{1/2W_o}\, h_2(z, \tau) * \text{sinc}\, 2W_o\tau$$

in which the asterisk denotes convolution. Therefore

$$h_2(z, \tau) = \int_{-\infty}^{\infty} \sum_m h_2(z, s)\, \delta\left(s - \frac{m}{2W_o}\right) \text{sinc}\, 2W_o(\tau - s)\, ds$$

$$= \sum_m h_2\left(z, \frac{m}{2W_o}\right) \text{sinc}\, 2W_o\left(\tau - \frac{m}{2W_o}\right) \tag{1}$$

Next, for the variable ω, which is restricted to $(-W, W)$, we can write

$$H_2(j\omega, \tau) = \text{rep}_{2W}\, H_2(j\omega, \tau)\, \text{rect}\, \frac{\omega}{2W}$$

If we transform both sides, we obtain

$$h_2(z, \tau) = \sum_n h_2\left(\frac{n}{2W}, \tau\right) \text{sinc}\, 2W\left(z - \frac{n}{2W}\right) \tag{2}$$

Substituting for $h_2\left(\frac{n}{2W}, \tau\right)$ from Eq. 1, we then obtain the desired sampling representation:

$$h_2(z,\tau) = \sum_n \sum_m h_2\left(\frac{n}{2W}, \frac{m}{2W_o}\right) \text{sinc } 2W_o\left(\tau - \frac{m}{2W_o}\right) \text{sinc } 2W\left(z - \frac{n}{2W}\right)$$

Using the relations between $h_2(z,\tau)$ and $h_1(t,\tau)$, we can also write

$$h_1(t,\tau) = \sum_n \sum_m h_1\left(\frac{n}{2W} + \frac{m}{2W_o}, \frac{m}{2W_o}\right) \text{sinc } 2W_o\left(\tau - \frac{m}{2W_o}\right) \text{sinc } 2W\left(t - \tau - \frac{n}{2W}\right)$$

4. Delay-Line Equivalent

The usefulness of these relations lies in the fact that they enable us to obtain a delay-line representation of the time-variant system, as shown in Fig. 1, in which the $f_n(t)$ represent time-variant gain controls.

The ideal filter is assumed to have zero phase shift; assumption of linear phase shift will only introduce a constant delay.

The equivalence of the delay line and the original time-variant filter is proved by calculating the impulse response of this model. If an impulse is put in at $t = \tau$, the output will be

$$\sum_n f_n\left(\tau + \frac{n}{2W}\right) 2W \text{ sinc } 2W\left(t - \tau - \frac{n}{2W}\right)$$

$$= \sum_n \sum_m h_2\left(\frac{n}{2W}, \frac{m}{2W_o}\right) \text{sinc } 2W_o\left(\tau - \frac{m}{2W_o}\right) \text{sinc } 2W\left(t - \tau - \frac{n}{2W}\right)$$

$$= \sum_n \sum_m h_2\left(\frac{n}{2W}, \frac{m}{2W_o}\right) \text{sinc } 2W_o\left(\tau - \frac{m}{2W_o}\right) \text{sinc } 2W\left(z - \frac{n}{2W}\right)$$

$$= h_2(z,\tau)$$

when $z = t - \tau$. Since two linear systems that have the same impulse response have the same output for any excitation, it follows that they are equivalent as far as terminal characteristics are concerned. A significant characteristic of a time-variant system is the frequency expansion that it produces, and at this point it is reasonable to inquire how this enters our analysis.

If an input of bandwidth $2W'$ is put into the system, the output bandwidth will be $2W' + 2W_o$ (or less). If the largest allowed input bandwidth is $2W_{in}$, the maximum possible output bandwidth of the system will be $2W_{in} + 2W_o$. Our $2W$ is equal to $2W_{in} + 2W_o$, which is now justified. We know that (time-domain) sampling at intervals closer than the reciprocal of the bandwidth still gives the original function when we reconstruct from these samples by using sinc functions that correspond to the sampling interval that is

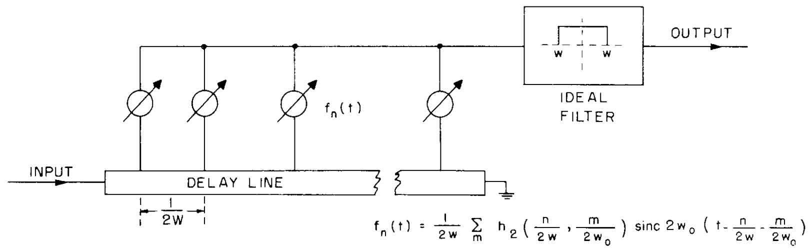


Fig. VIII-1. Delay-line model for the lowpass case.

used (2). Too frequent sampling gives no trouble – it is only when samples are farther apart than $1/2W$ that we are unable to reconstruct the original function from its samples. Thus closer spacing in the delay line (when it does occur) is not a source of error.

Since the impulse response of the delay line is frequency-limited, the line must theoretically be of infinite extent. It can be shown, however, that a finite length can be used at the cost of an error that can be made arbitrarily small by prolonging the line sufficiently. We remark again that this model is not a total equivalent of the time-variant filter; it merely simulates it over a specified frequency range.

5. Highpass Case

In this case the frequency range of interest is not centered at the origin but is a band of high frequencies, say $\left(\omega_c - \frac{W}{2}, \omega_c + \frac{W}{2}\right)$, in which ω_c is the center frequency of the band, and W is its width.

Similarly, by using Woodward's elegant method of deriving bandpass sampling theorems (3), we can obtain

$$h_1(t,\tau) = \sum_n \sum_m h_1\left(\frac{n}{W} + \frac{m}{2W_o}, \frac{m}{2W_o}\right) \text{sinc } 2W_o\left(\tau - \frac{m}{2W_o}\right)$$

$$\text{sinc } W\left(t - \tau - \frac{n}{W}\right) \cos \omega_c\left(t - \tau - \frac{n}{W}\right)$$

$$- \sum_n \sum_m \hat{h}_1\left(\frac{n}{W} + \frac{m}{2W_o}, \frac{m}{2W_o}\right) \text{sinc } 2W_o\left(\tau - \frac{m}{2W_o}\right)$$

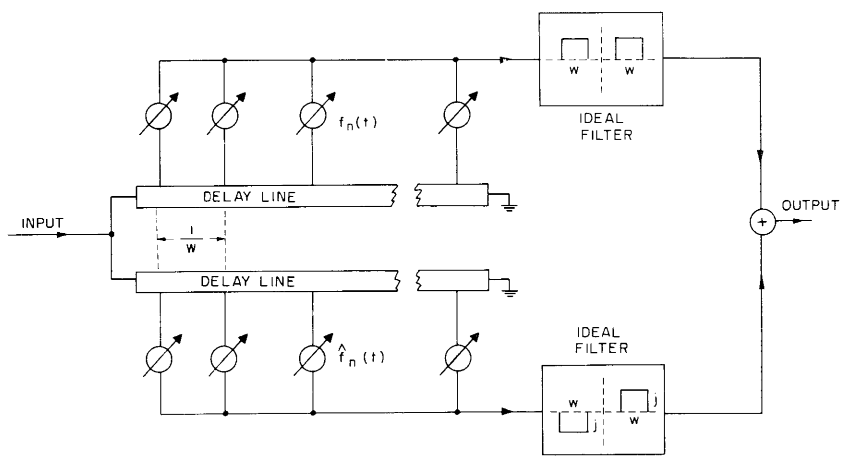
$$\text{sinc } W\left(t - \tau - \frac{n}{W}\right) \text{sinc } \omega_c\left(t - \tau - \frac{n}{W}\right)$$

as a sampling representation for the system impulse response that is valid over the frequency range $\left(\omega_c - \frac{W}{2}, \omega_c + \frac{W}{2}\right)$.

Here $\hat{h}_1(t,\tau)$ is the Hilbert transform of $h_1(t,\tau)$ with respect to t.

$$\hat{h}_1(t, \tau) = \int_{-\infty}^{\infty} \frac{h_1(s + \tau, \tau)ds}{t - \tau - s}$$

Several different, but equivalent, forms can be written in this case, and correspondingly different delay-line models can be obtained. One such model is shown



$$f_n(t) = \frac{1}{2W} \sum_m h_1\left(\frac{n}{W} + \frac{m}{2W_0}, \frac{m}{2W_0}\right) \text{sinc } 2W_0 \left(t - \frac{n}{W} - \frac{m}{2W_0}\right)$$

$$\hat{f}_n(t) = \frac{1}{2W} \sum_m \hat{h}_1\left(\frac{n}{W} + \frac{m}{2W_0}, \frac{m}{2W_0}\right) \text{sinc } 2W_0 \left(t - \frac{n}{W} - \frac{m}{2W_0}\right)$$

Fig. VIII-2. Delay-line model for the highpass case.

in Fig. VIII-2. The filters are assumed to have zero phase shift; as before, the assumption of a linear phase shift serves only to introduce a time delay. Once again, the length of the delay lines is (theoretically) infinite, and the tap spacing is the reciprocal of the "max-max" bandwidth.

T. Kailath

References

1. P. M. Woodward, Probability and Information Theory with Applications to Radar (McGraw-Hill Publishing Company, New York, 1953).

2. W. K. Linvill, Sampled-data control systems studied through comparison of sampling with amplitude modulation, Trans. AIEE, Vol. 70, Part II, p. 1779 (1951).

3. P. M. Woodward, op. cit., p. 31.

4. Ibid., p. 34.

C. PICTURE PROCESSING

The digital system and test pictures that were presented in an earlier report (1) have been used in several exploratory studies of coding pictures for transmission. These studies are summarized in the following sections.

1. Two-Dimensional Second-Difference Pictures

At a sample point in a picture, the two-dimensional second-difference is defined as the sum of the intensities at the preceding, succeeding, and two laterally adjacent sample points, minus four times the intensity at the sample point in question. Transmission of the two-dimensional second-difference for each sample point is sufficient to permit the receiver to reconstruct the picture if the transmitter and receiver have agreed upon a common set of boundary conditions at the start.

If there are long sequences of zero values for the two-dimensional second-difference, the signal can be coded for transmission in terms of run-lengths of zeros and the values of the first nonzero second-difference at the end of each run. To extend the lengths of runs, all two-dimensional second-differences that are smaller in magnitude than a chosen value can be changed to zero. Unfortunately, because of the limited dynamic range of most reproducing systems, the errors introduced into the reconstructed signal at the receiver by the modification of the difference signal are not acceptable.

This difficulty can be avoided by resynthesizing the intensity values at the transmitter from the modified two-dimensional second-difference signal in the same way that the receiver would resynthesize them. Then the two-dimensional second-difference to be transmitted can be obtained from two lines of resynthesized data plus a new line of actual data from the picture. This difference signal is modified by setting all values not greater than the chosen reference criterion to zero. In this method the errors inserted by approximation at any point are not propagated to the remainder of the picture, since they are accounted for in the difference computation along the next line before the next approximation is applied.

For this coding scheme, the source rate can be estimated by using the formula

$$\text{Source-rate estimate} = \frac{H_r + N + 3}{\bar{r}}$$

where H_r is the entropy of the run-length probability distribution, N is the number of binary digits used to specify a sample value, and $\bar{r}$ is the average run length. The 3 appears because the two-dimensional second-difference can be four times as much as the largest intensity value in magnitude and can be either positive or negative. Thus three more digits are required to specify the two-dimensional second-difference than are needed for the intensity itself.

The results obtained by this procedure are given in Table VIII-1; the pictures

(a)

(b)

(c)

(d)

Fig. VIII-3. Pictures reproduced from modified two-dimensional second-difference signals: (a) and (c), criterion 3; (b) and (d), criterion 6.

obtained are shown in Fig. VIII-3. These results are not particularly promising. Although the assumption of a uniform probability distribution for the values of the second-difference is probably far from correct, it does not appear that an exceptionally large saving in channel capacity requirements could be achieved even by accounting for this distribution. However, there is an interesting effect that is especially noticeable in Fig. VIII-3b. Although a grainy effect appears because of the method of approximating

Table VIII-1. Source-Rate Estimates for Two-Dimensional Second-Difference Pictures.

Fig. VIII-3 (N=6)	Criterion (levels)	H_r (bits)	$\bar{r}$ (samples)	Source Rate Estimate (bits/sample)
Picture a	3	1.74	2.16	5.0
Picture b	6	1.96	2.43	4.5
Picture c	3	1.11	1.42	7.3
Picture d	6	1.45	1.65	6.3

the second-differences, it is not particularly objectionable. In some of the pictures reproduced by means of the piecewise-linear approximation of reference 1, the actual intensity errors were much smaller, but undesirable distortions were more noticeable. We may conjecture that it will be necessary to process pictures in at least as many dimensions as we view them, not only to benefit from the statistical relations available for coding, but also to avoid introducing undesirable types of distortion.

W. A. Youngblood

References

1. W. A. Youngblood, Quarterly Progress Report, Jan. 15, 1958, Research Laboratory of Electronics, M.I.T., pp. 95-100.

2. Recoding Pictures by Generation of Lowpass and Correction Signals

The Whirlwind computer was programmed to average the intensity values over blocks of 25 samples (5 samples by 5 samples) to produce what is called the lowpass signal from a picture. This lowpass signal requires specification only in the centers of the blocks if interpolation is used to obtain intervening intensities. The interpolation was carried out by the computer, and the resulting value at each point was subtracted from the actual sample value. If the magnitude of the difference exceeded 5 levels, a correction signal was added, either +10 or −10 levels, as determined by the sign of the error.

An upper bound to the source information rate is numerically equal to the entropy associated with the lowpass signal plus the entropy associated with the correction signal.

Table VIII-2.

Fig. VIII-4	$\Sigma p_i \log p_i$ (bits)	Source Rate Estimate (bits/sample)
Picture a	0.60	0.84
Picture b	1.47	1.71

(a) (b)

Fig. VIII-4. Pictures resynthesized by addition of lowpass signal and recoded highpass signal.

The entropy for the lowpass signal is constant, 6/25 = 0.24, for a 64-level picture. The entropy for the correction signal is $\Sigma p_i \log p_i$, where the p_i are the probabilities of 0, +10, -10, the three values of the correction signal. The results for the pictures shown in Fig. VIII-4 are given in Table VIII-2.

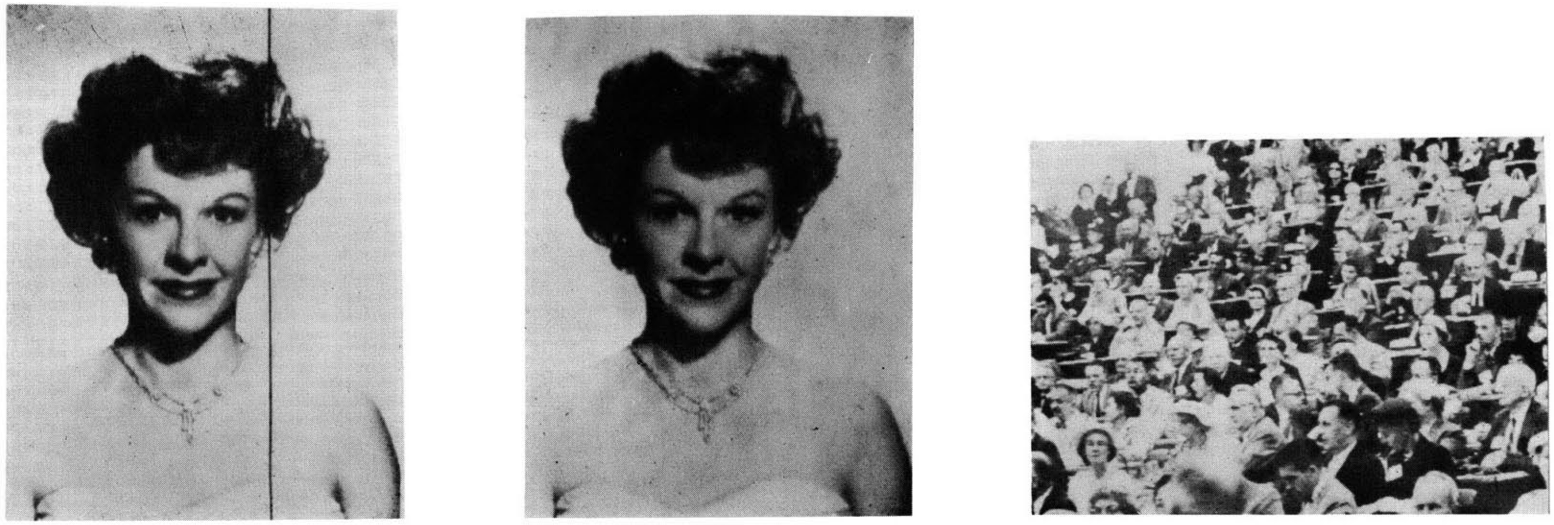
The average number of digits necessary for picture specification depends strongly on the complexity of the picture. More work will be done in this area in an attempt to find optimum criteria for allowed error and for the correction signal magnitudes.

J. E. Cunningham

3. Picture Coding by Linear Interpolation

A simple way to reduce the source rate of a digitalized picture is to reduce the number of quantization levels. However, this process quickly leads to a "staircase" effect in regions where the light intensity is gradually changing. (See upper left-hand corner of picture a, Fig. VIII-5.) One way to avoid this difficulty is to interpolate between levels when the change is gradual. A method for achieving this interpolation has been programmed and tested.

One line of the picture is processed at a time. The line is broken up into runs, each of which includes all of the successive samples in the same coarse quantization level. As shown in Fig. VIII-6, there are sixteen of these. Each run is divided in half. In the first half, the number of samples in the lower half of the coarse level is counted and



(a) (b) (c)

Fig. VIII-5. (a) Four-bit picture. (b) and (c) Pictures processed by linear interpolation.

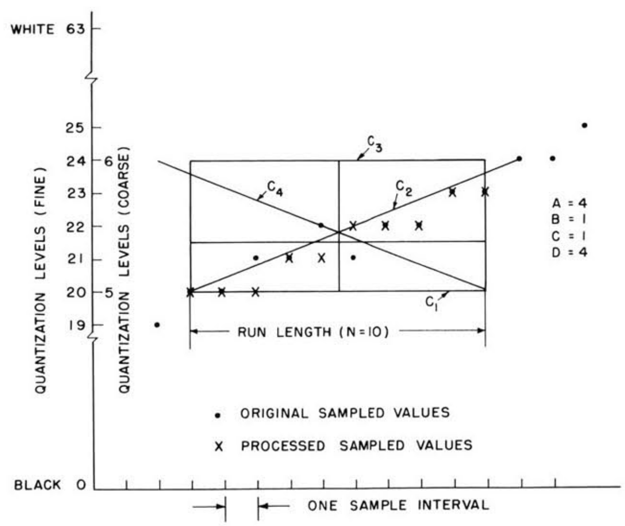


Fig. VIII-6. Processing by linear interpolation.

Table VIII-3. Approximation Criteria.

	$A \geq B$	$A < B$
$C \geq D$	C_1	C_4
$C < D$	C_2	C_3

Table VIII-4. Source-Rate Estimates.

Fig. VIII-5	r	H_r	p	H
(b)	4.21	2.80	.072	1.65
(c)	1.89	1.65	.015	3.00

designated A. (In the example of Fig. VIII-6 these are samples in level 20 or 21.) Likewise, the number in the upper half (here, level 22 or 23) is designated B. Similarly, C and D are the numbers of samples in the lower and upper halves of the second half of the run. As determined by the relationships of A to B and C to D, one of four approximating straight lines, curves C_1, C_2, C_3, or C_4, is chosen (see Table VIII-3) to give the processed samples. (Of course, C_2 and C_4 can only be approximated, because of round-off problems.) If the run-length N, the number of samples in a run, is less than a certain minimum number, N_o, curve C_1 is automatically chosen.

Thus we can specify the processed version of the picture by specifying successively the runs by their coarse level, run-length, and approximating curve. Therefore, a source-rate estimate is

$$H = \frac{H_r + \log Q + p \log C}{\bar{r}}$$

where H_r is the entropy of the run length distribution in bits, Q is the number of coarse levels, C is the number of approximating curves, $\bar{r}$ is the average run length in sample intervals, and p is the fraction of runs for which $N \geq N_o$. For the present experiment, $N_o = 10$; and we have

$$H = \frac{H_r + 4 + 2p}{\bar{r}}$$

Two pictures have been processed (pictures b and c, Fig. VIII-5). The numerical results are given in Table VIII-4. Picture c shows little, if any, distortion when compared with the original picture. In picture b it is seen that some "streaking" along the direction of scan has occurred, but the "staircase" effect seems to have been eliminated. We note that the results are similar to those for the linear approximation method that was used by Youngblood. Further investigation should probably be carried out by two-dimensional methods.

R. S. Marcus

References

1. W. A. Youngblood, Quarterly Progress Report, Jan. 15, 1958, Research Laboratory of Electronics, M.I.T., pp. 95-100.
2. W. A. Youngblood, Estimation of the channel capacity required for picture transmission, Sc.D. Thesis, Department of Electrical Engineering, M.I.T., May 1958.

D. ESTIMATION OF THE CHANNEL CAPACITY REQUIRED FOR PICTURE TRANSMISSION

A thesis, part of which is discussed in Section VIII-C1, was completed and submitted to the Department of Electrical Engineering, M.I.T., in partial fulfillment of the requirements for the degree of Doctor of Science (May 1958).

W. A. Youngblood

E. CODING AND DECODING OF BINARY GROUP CODES

This study was carried out by F. Jelinek. He submitted the results to the Department of Electrical Engineering, M.I.T., as a thesis in partial fulfillment of the requirements for the degree of Master of Science.

R. M. Fano