TELEVISION BANDWIDTH COMPRESSION

By

CHARLES GENE HILBORN, JR.

Bachelor of Science

Louisiana Polytechnic Institute

Ruston, Louisiana

1963

Submitted to the Faculty of the Graduate School of
the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
MASTER OF SCIENCE
May, 1965

TELEVISION BANDWIDTH COMPRESSION

Thesis Approved:

Wm L. Hughes
Thesis Adviser

Arthur M. Breipohl

J H Boggs
Dean of the Graduate School

PREFACE

Even in the early days of television, before mathematical information theory, it was recognized that television signals are inefficient information carriers. The search for methods of more efficiently transmitting the visual information by what is generally known as "television bandwidth compression" has been and will continue for some time in the future to be an interesting and challanging problem.

In Chapters II through V, I have tried to give a systematic treatment to the statistical basis for compression by both theoretical and practical discussions. Chapter VI explores, on the other hand, the visual bases of compression. Although statistical compression on an inter-frame scale is possible, it has not been achieved because of the enormous associated storage difficulties. In Chapters VII and VIII I have suggested an alternative visually coded inter-frame system which simplifies to a large extend the storage problem.

I wish to express my appreciation to Dr. William L. Hughes, my thesis advisor, for his help and guidance, and for our interesting discussions on this subject. I would also like to express my sincere thanks to Dr. Arthur Breipohl for his careful reading and helpful criticisms of the thesis manuscript.

TABLE OF CONTENTS

Chapter		Page
I.	INTRODUCTION	1
II.	CHANNEL CAPACITY	4
	Information Capacity	4
	Capacity of the Television Channel	5
III.	INTRA-FRAME REDUNDANCY	8
	Intutitive Discussion	8
	Picture Statistics	9
	Brightness Distribution	10
	Adjacent Element Correlation	10
IV.	INTRA-FRAME CODING	16
	Velocity Modulation	16
	An "Open-Loop" System: Element Coding . . .	19
	Coding for Vertical Redundancy	24
V.	INTER-FRAME CODING	27
	Intuitive Discussion	27
	Inter-Frame Statistics	28
	Inter-Frame Coding	29
VI.	CODING TO MATCH VISUAL PERCEPTION.	31
	Visual Channel Capacity	31
	Exchange of Contrast Resulution for Spatial Resolution	33
	Continuity of Vision	38
	Exchange of Spatial Resolution for Time-Motion Resolution	40
VII.	A SYSTEM OF INTER-FRAME CODING	42
	Frame Run Coding	42
	Uniform Frame Sampling	46
	Picture Reconstruction	47
	Transmitted Error	48
	Noise Effects	50
	Linear Interpolation	53
	Noise Effects	55

TABLE OF CONTENTS

Chapter		Page
VIII.	INSTRUMENTATION OF AN INTER-FRAME SYSTEM	59
	Simulation	64
IX.	SUMMARY AND CONCLUSIONS	67
	Summary	67
	Suggestions for Future Study	68
	Concluding Remarks	68
BIBLIOGRAPHY		70
APPENDIX		74

LIST OF FIGURES

Figure		Page
1.	Scanning Raster	15
2.	(a) Slope-feedback Encoder, (b) Horizontal Velocity Versus Instantaneous "Detail"	17
3.	Two-channel Variable-velocity System	18
4.	Element Run Coding	21
5.	Variable-velocity element run decoder	22
6.	Two-dimensional Coding, (a) Non-interlaced System (b) Modified Vertical Decision for an Interlaced System	25
7.	Visual Perception Model	32
8.	Quantized Highs Simulation	35
9.	Comparison of Quantized Highs and Synthetic Highs Reproductions	36
10.	Zero-order, First-order, and Contour Interpolation	40
11.	A Frame Run Decision Device	43
12.	Event Redistribution	44
13.	Uniform Frame Sampling Systems	47
14.	Visual Interpolation	53
15.	Linear Interpolation	55
16.	Noise Power Law for Linear Interpolation	57
17.	Transmitted Error/Multiplex System	61
18.	Transmitted Error/Time Scale Change System	61
19.	Element Interpolation Receiver	64
20.	Transmitted Error Simulation	65
21.	Linear Brightness Interpolation Simulation	66

CHAPTER I

INTRODUCTION

The fact that the standard system of transmitting television picture signals requires enormous bandwidth for the actual information subjectively received is almost a cliche. One picture requires a bandwidth of about 4 megacycles - roughly four times the bandwidth of the entire AM broadcast band. It would unquestionably be very useful to find a means of compressing the bandwidth of the picture signal without noticeable degradation.

One approach to the bandwidth problem is as follows: given a certain narrow bandwidth, what is the optimum subjective picture which can be obtained by juggling the scanning parameters and phosphor decay time. This problem has been studied extensively elsewhere [11], and will not be covered here. Furthermore, for the standard 4 Mc. bandwidth, the standard scanning parameters were, to a large extent, optimally chosen. [12]

In this study, we shall be more interested in accepting as "given" the existing standards, and examining how the signal can be further processed (coded) to compress bandwidth. "Bandwidth compression" will be used here in the broad sense of true bandwidth savings, time savings for equal bandwidth, and in general, reducing required "channel capacity" (to be precisely defined).

The Shannon-Hartly law [44] states in the most general sense the terms-of-trade between signal-to-noise ratio and bandwidth for a fixed information channel capacity:

$$C = W \log (P/N + 1)$$

where P is the message power, N is the noise power, and W is the channel bandwidth. To see the futility of compressing bandwidth by increasing signal power, suppose that initially $P/N = 1000$ (30 db). To halve the bandwidth, while keeping C constant, requires increasing P/N by a factor of 1000 - illustrating that the logarithmic relation makes coding to reduce bandwidth by increasing message power an extremely unprofitable exchange. It is well known that exchange in the opposite direction is effectively used in FM and PCM [31].

It is necessary then, to actually reduce the information capacity of the picture channel in order to save bandwidth. This objective can theoretically be accomplished - without noticeable picture deterioration - by discarding data in the signal which, ultimately, do not represent true information. The criteria for selecting what should be discarded fall into two categories. First, it is possible to draw on information theory to determine mathematically the actual information to the allowed channel capacity. The required capacity can be reduced if some form of elastic coding is used which, over some period of time, equalizes the rate of information so that instead of allowing for rare peaks, the channel capacity is brought down toward the average rate.

Second, while much of the data sent over a picture channel is necessary from a mathematical information viewpoint, it often is vastly overspecified or even unusable from a subjective viewpoint - and can therefore be discarded to save channel capacity.

In this thesis, both aspects of channel capacity reduction for bandwidth compression, as well as systems based on them, will be discussed. It is often difficult to decide into which category (if not both) to place a given system. A compression system to be proposed here falls mainly into the "subjective coding" class, although information theory does provide a framework for discussion.

CHAPTER II

CHANNEL CAPACITY

Information capacity

Shannon [44] has defined a quantitative measure of information. If a message is generated by selecting symbols from a set of N symbols, each having a probability $P(x_i)$, the average information or entropy per symbol is given by

$$H_x = -\sum_{i}^{N} P(x_i) \log P(x_i).$$

And for the continuous case, the summation is replaced by integration:

$$H_x = -\int_{-\infty}^{\infty} p(x) \log p(x) \, dx$$

where $p(x)$ is the probability density of x.

The sampling Theorem [44] shows that if a function of time band-limited to frequencies from 0 to W cycles per second, it is completely determined by specifying its value at a set of points $1/2W$ seconds apart.

As a result of this theorem, the rate of information per second of a band-limited signal is given by

$$I_x = 2W H_x$$

where H_x is the entropy per independent sample or degree of freedom.

If a message x is perturbed by independent additive noise n, the rate of information passing through the channel can be shown [44] to be

$$I_c = I_y - I_n,$$

where I_y is the entropy rate of the recieved signal y, and I_n is the entropy of the noise. The channel capacity C is then found by maximizing I with respect to the possible distributions for x:

$$C = \underset{p(x)}{\text{Max}}\ I_y - I_n .$$

Shannon has shown that C defines the highest rate of information through a channel with as small a frequency of errors as desired. If the signal-to-noise ratio is sufficiently large, it suffices to maximize H_x.

Capacity of the Television Channel

With these difinitions we can determine a theoretical upper bound of channel capacity inherent in the standard method of transmitting a TV picture signal.

Let x represent the picture brightness (or analog voltage) at a point. x can vary between fixed limits, x_1 and x_2. With this constraint, H_x is maximized when [44]

$$p(x) = \frac{1}{x_2 - x_1}, \quad x_1 \leq x \leq x_2$$
$$= 0 \quad , \quad \text{elsewhere}$$

which gives the sample point an entropy of

$$\text{Max}\, H_x = \int_{x_1}^{x_2} \frac{\log (x_2 - x_1)}{x_2 - x_1}\, dx = \log (x_2 - x_1) .$$

If the signal is band-limited to W, and the samples are statistically independent, then

$$\text{Max}\ I_x = W \log (x_2 - x_1)^2 .$$

It is also evident that for signals of this type having equal average power, the entropy is maximized when $x_2 = x_1$. That is, a DC component should not be transmitted since it requires power but conveys no information. With this assumption, let

$$x^2 = S = \text{peak power}.$$

Then,

$$\text{Max}\ I_x = W \log 4 S .$$

If, in the channel, white gaussian noise adds to the signal, and is also limited to band W, then the noise entropy is

$$I_n = -2\, W \int_{-\infty}^{\infty} \frac{\exp [-n^2/2N]}{\sqrt{2 \Pi N}} \log [\frac{\exp(-n^2/2N)}{\sqrt{2 \Pi N}}]\, dn$$

$$= W \log_2 2 \Pi e N$$

where N is the variance or average power of n.

From a theorem of Shannon [44], a lower bound and an asymtotic upper bound for I_y are

$$W \log \frac{4S}{e^2} \leq I_y \leq W \log (4S + 2\Pi e N)(I + \varepsilon),$$

where ε tends toward zero as the signal-to-noise ratio becomes large.

This gives a set of bounds for channel capacity:

$$W \log \frac{2S}{\Pi e^2 N} \leq C' \leq W \log \left(\frac{4S + 2\Pi eN}{2\Pi eN}\right)(1 + \varepsilon).$$

If we take the nominally acceptable SNR of 30 db peak signal to RMS noise, $W = 4$ Mc., neglect ε, and use the base two, then

$$25\,(10^6) \leq C' \leq 31\,(10^6) \text{ bits/sec.}$$

To allow for the blanking and synchronizing time in actual TV standards (about 20%) C' should be multiplied by a factor of 0.8. This gives a truer estimate of actual picture channel capacity under existing standards. Hence,

$$21(10)^6 \leq C \leq 25\,(10^6) \text{ bits/sec.}$$

CHAPTER III

INTRA-FRAME REDUNDANCY

Intuitive Discussion

The channel capacity estimate of 25 megabits per second corresponds statistically to an ensemble of pictures having uniform intensity distributions for all elements, and statistically independent element distributions. "Elements" are defined (by the Sampling Theorem) as picture points one Nyquist interval apart.

The subjective impression of a picture with such statistics would be no different from a picture of random noise. Since real pictures never look like this, i.e., there is always pattern (correlation), the theoretically available capacity is never utilized. It follows that if the signal could be coded to remove statistical redundancy the channel capacity could be decreased to approach the average information rate.

If one examines the picture signal produced by the ordinary uniform scanning process, it is evident that statistical redundancies in a single frame can be categorized as follows: deviation of the brightness distribution from uniform, and correlation between adjacent picture elements in both the horizontal and vertical directions.

Picture Statistics

In order to discuss picture statistics it is useful to discuss the source capacity - information in the picture without noise. It is mathematically convenient to quantize brightness into N levels.

If the i-th level occurs with probability P(i), then the information of a single element is

$$H_x = - \sum_{i}^{N} P(i) \log P(i) .$$

The maximum value of this entropy is found when the distribution of P(i) is uniform; i.e., when the probability of all levels 1/N. This maximum value will be denoted by S. Thus,

$$S_x = \log N.$$

It has been observed (7, 14, 37) that about 64 levels are needed.[1] In this case, the element source capacity is 6-bits.

If all elements in a frame are statistically independent, the information in the frame is KS_x, where K is the number of elements per frame (about 200,000). With a frame rate of 30 per second, the picture source capacity is

$$C_s = (6)(30)\,3\,(10^5)$$
$$= 36\,(10^6) \text{ bits/sec.}$$

---

[1] The number of levels needed for a subjectively pleasing picture actually depends upon several factors and will be discussed in Chapter 6.

Brightness Distribution

The redundancy or loss of entropy due to non-uniformity can be defined as

$$R_b = S_x - H_x$$

$$= \log N + \sum_{i}^{N} P(x_i) \log P(x_i) \ .$$

The experimental work of Kretzmer [23] and Schreiber [38] has shown R_b is about one bit when the brightness distribution is taken over one frame, and essentially independent of quantization. The channel capacity savings possible by exploiting this redundancy would not seem to be profitable, since a practical system could not achieve even the modest capacity reduction implied by this figure.

Adjacent Element Correlation

Along a given scan line, the picture signal varies in accordance with element brightness. Since nearly all scenes contain some areas of relatively gradual shading, a prediction that the present element value x is within the same quantum level as the previous element y would often be correct.

The information of two picture elements can be expressed as

$$H_{xy} = - \sum_{i}^{N} \sum_{j}^{N} P(i, j) \log P(i, j)$$

where y is the previous element and x is the present element, and i and j index the values of x and y, respectively. Rewriting

P(i, j),

$$P(i, j) = P(j)\, P(i/j),$$

and defining the conditional entropy,

$$H_{x/y} = - \sum_{i}^{N} \sum_{j}^{N} P(i, j) \log P(i/j)$$

we have

$$H_{xy} = H_x + H_{x/y}\,.$$

If x and y are statistically independent, then

$$P(i/j) = P(i),$$

and

$$H_{x/y} = - \sum_{i}^{N} \sum_{j}^{N} P(i, j) \log P(i)$$

$$= - \sum_{i}^{N} P(i) \log P(i)$$

$$= H_x.$$

Thus, in this case

$$H_{xy} = 2\, H_x.$$

Similarly the entropy of three elements is

$$H_{xyz} = - \sum_{i}^{N} \sum_{j}^{N} \sum_{k}^{N} P(i, j, k) \log P(i, j, k)$$

Rewriting P(i, j, k),

$$P(i, j, k) = P(j, k)\, P(i/j, k)$$

If there exist no correlation between the present element x and the second previous element z, but a correlation does exist between the present and immediately preceding element y (Markoff process), then

$$P(i/j, k) = P(i/j).$$

And[2]

$$H_{xyz} = H_{xy} + H_{x/y}$$

$$H_{xyz} = H_x + 2\, H_{x/y}.$$

The prediction of previous element values has been experimentally studied by Harrison [19] in connection with a system of sending only "mistakes." Experimental analyses of element redundancy have been carried out by Kretzmer [23] and Schreiber [38]. Powers and Staras [33] have analyzed theoretical aspects of element redundancy.

Schreiber's investigations have shown that the third-order correlation between elements are of minor importance - compared with second-order (previous-value) correlation. Consequently, if third- and greater-order horizontal correlations are neglected, the entropy of K-elements is approximated as

$$H_{x_1, \ldots, x_k} = H_x + (K-1)\, H_{x_1/x_2},$$

---

[2] Note that

$$H_y = H_x = H_z, \quad H_{x/y} = H_{y/z}, \quad \text{etc.}$$

and the average entropy per element is found by dividing by K. Taking the limit for many elements, the average entropy per element is

$$\lim_{K \to \infty} \frac{H_x + (K-1)\ H_{x_1/x_2}}{K} = H_{x_1/x_2}$$

or

$$= H_{y/x}\ .$$

The difference between $H_{y/x}$ and the maximum possible value of average entropy per element gives the total horizontal redundancy.

$$R_h = S_x - H_{y/x}\ .$$

It is useful to separate this total into two components:

$$R_h = [S_x - H_x] + [H_x - H_{y/x}]\ .$$

Since the first term is the brightness distribution redundancy R_b, the second term can be attributed to adjacent element correlation. Let it be symbolized by R_{he}. Thus,

$$R_h = R_b + R_{he}.$$

And

$$R_{he} = H_x - H_{y/x}\ .$$

It was seen previously that R_b does not yield the likelihood of significant savings in channel capacity. However the measurements of Schreiber [38] show that for pictures with 64-levels (S_x = 6 bits), "subject A" had a total second-order R_h of 2.64 bits of which R_b was .3 bit, and R_{he} was 2.34 bits. For subject

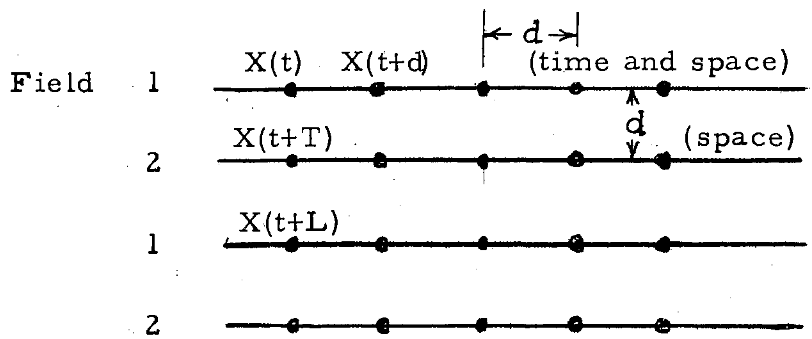
B, the total previous-element redundancy given was 4.15 bits, of which 1.70 bits are brightness distribution redundancy, and 2.45 bits are due to adjacent element correlation. Schreiber's third-order probability distribution measurements yield a slightly higher R_{he} of the 2.90 bits for subject B. Teer [46] has obtained a second-order R_{he} of at least 2.50 bits from Schreiber's figures when a 32-level (5 bits) quantization is applied.

These figures suggest that it is theoretically possible to reduce channel capacity (or bandwidth - other things being equal) by about 50% by coding to exploit only adjacent element previous-value correlation.

Kretzmer [23] measured autocorrelation in various directions on photographic slides; he also obtained brightness probability distributions and the distribution-of-error of previous-value prediction. He gives a total previous-value redundancy of 3.4 bits for a 6-bit signal - in close correspondence with the results of Schreiber. One important finding of Kretzmer is that there seems to be no preferred direction of correlation in most pictures. It follows that the statistical redundancy in the vertical direction - between scan-lines - is about the same as that in the horizontal direction. This gives the possibility of a further - independent - 2:1 compression vertically; or a total of 4:1 compression due only to second-order two-dimensional correlation.

The entropy, redundancy and resulting compression possibilities in this chapter have been based only on brightness information, with the tacit assumption that position information is known. As will be seen in the next chapter, under conditions of compression, this assumption must be modified.

Although vertically adjacent elements may be adjacent in space, the scanning scheme separates them in time by one line-period in a non-interlaced system and one field-period in a 2:1 interlaced system - which is standard in the United States. The situation is depicted diagramatically in Figure 1.



d - Nyquist interval or element distance

L - Line period

T - Field period

Figure 1: Scanning raster

The requirements for long delay or large storage capacity suggests the great difficulties in the practical realization of systems which exploit vertical element correlation.

We have examined some of the measurements and theoretical aspects of single picture statistics. It is of interest, on the other hand, to examine some measurements which are closely matched to practical coding systems, and the success of these systems.

CHAPTER IV

INTRA-FRAME CODING

Velocity Modulation

One of the first television bandwidth compression systems was described (in theory) by Cherry and Gouriet [9] in 1953. This system used variable-velocity scanning in the camera and receiver; the spot velocity being made to vary inversely with the instantanious "picture detail." Picture detail was defined as the modulus of the time derivative of the video waveform.

Measurements [15] of average picture detail for actual television signals gave relative indications of from 3% to 5% for most pictures, where 100% detail corresponded to a full amplitude (black-to-white) sine wave at the highest passband frequency.

The basic variable-velocity system is shown in Figure 2-a. This closed-loop system is known as "slope-feedback coding." The system would be most effective if the instantaneous horizontal spot velocity could be varied from a minimum in areas of maximum detail to infinity in areas of zero detail (flat intensity areas). Because of practical circuit considerations, the velocity is limited to a maximum as shown in Figure 2-b. Since the horizontal sweep velocity in the camera is derived from the transmitted signal, the receiver similarly can derive an equal sweep velocity. Although the system as presented

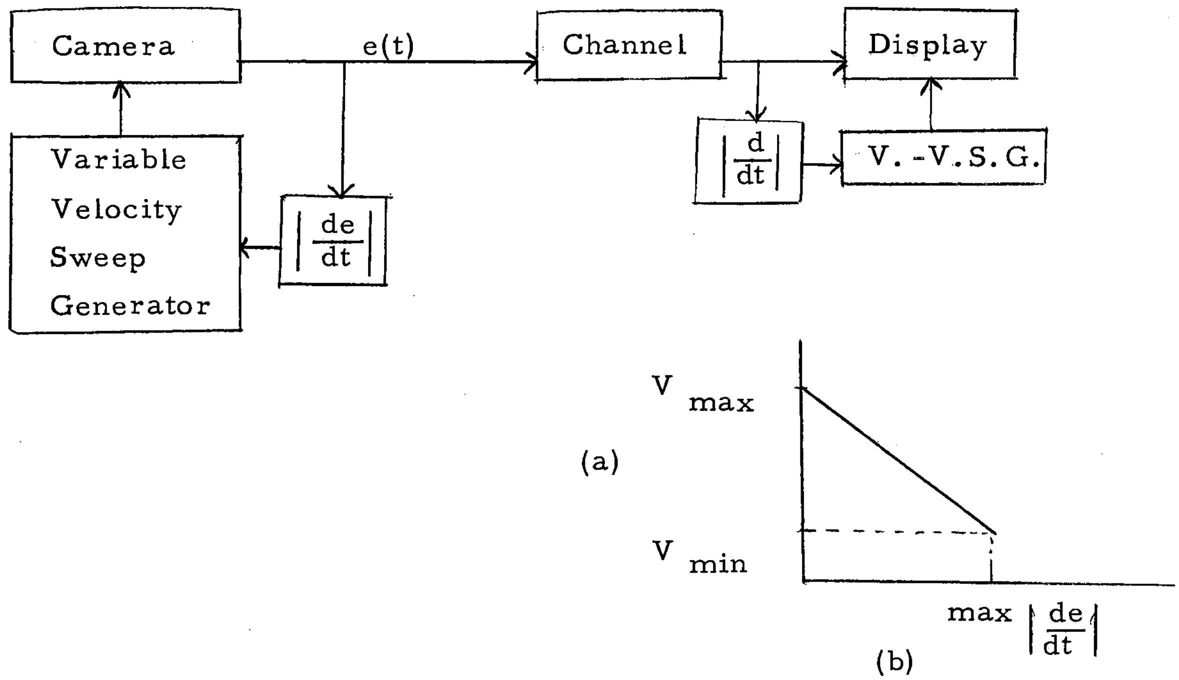


Figure 2. (a) Slope-feedback encoder, (b) Horizontal velocity *versus* instantaneous "detail."

requires a special camera and receiver and is not proposed as a system for processing a standard signal for bandwidth compression, it is conceivable that it could perform such a re-coding operation by replacing the camera and the receiver display by scan-converter tube or other standards conversion equipment.

Experimental results [35] with this system were rather disappointing. Even modest compression (1.7:1) was accompanied by serious geometric distortion and by loss of low-contrast fine detail.

The loss of detail could be expected from the definition or measure of "detail" built into the system. If the minimum velocity (Figure 2-b) is chosen to produce a highest passband signal when a full-contrast maximum resolution transition is occurring in the scanned picture, then a lower contrast maximum resolution transition will set up frequencies outside the allowed passband - which will be lost.

In order to avoid sweeping over fast transition "edges" in the picture before the camera beam sweep velocity can be slowed, the loop bandwidth at the transmitter must be very wide. The receiver loop bandwidth is effectively limited, however, by the limitation on the channel passband, thereby introducing spacial distortion into the received picture.

A later modification [1] reduced the geometric distortion by using a separate position channel. The modified system is shown in Figure 3. Since the position channel carried an actual scanning signal generated by the wideband loop at the transmitter, some reduction of geometric distortion was obtained [1], even with equal or less bandwidth allowed for the position signal than for the brightness signal. An additional bonus is the simplification of the receiver.

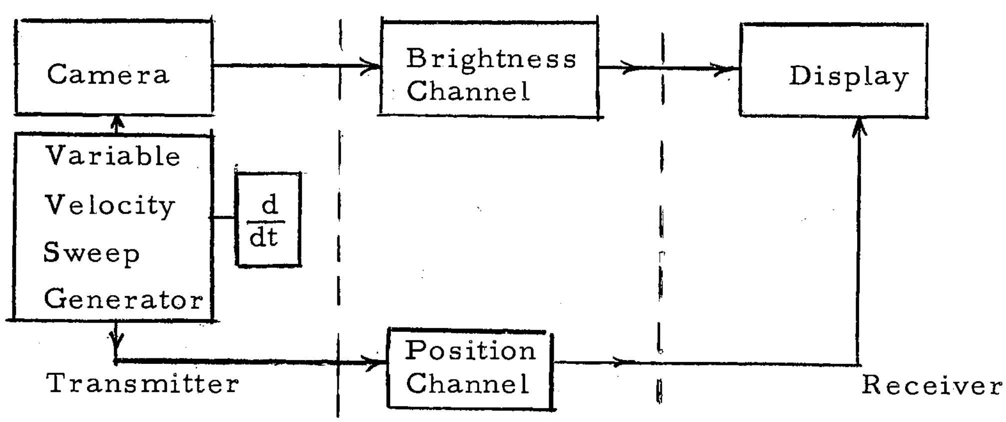


Figure 3. Two-channel variable-velocity system

These improvements are of course purchased at the expense of the bandwidth - or more precisely - the capacity required for the added position signal channel. It is important to say "capacity" because a signal-to-noise ratio of about 45 db. is (subjectively) needed for the position signal, where about 30 db. is acceptable for a signal conveying only brightness [1].

In studying this system, an important point of information theory, which is applicable to all statistical coding requiring epoch re-distribution, becomes clear. With uniform scanning, no position information is needed by the reciever, since the agreed time-space code (the uniform scanning raster) is known and position uncertainties do not exist. However, when the scanning spot velocity departs from an agreed patern, position information associated with that departure must also be sent; i.e., the changes in the "code book" must also be sent.

It is also evident from the study of this system that the detail measure $|de/dt|$ is inadequate as an information measure. Not only is differentiation a hopeless process in the presence of a realistic noise level, but a 100% $|de/dt|$ signal[1] conveys negligible (entropic) information since it requires negligible bandwidth.

An "Open-Loop" System: Element Run Coding

From the early ideas on coding the picture signal by variable-velocity scanning the concept of element run coding has evolved [35, 10]. This "open-loop" system more successfully and more

[1] $A \sin 2\pi f_o t$, where A is maximum amplitude, and f_o is highest frequency.

precisely achieves a uniform information rate - at least on the time interval of one horizontal line.

Element run coding is based upon the idea of transmitting only the picture elements which differ sufficiently from the previously transmitted element.

In the Cherry, at al, system [10], a decision element called a "Detail Detector" is supplied with a conventional picture signal. The Detail Detector produces an output sample element whenever a given element is judged (by a signal-to-noise discriminating inferential process) [25] to differ sufficiently from the previous element or run of "equal" elements. By this process, each horizontal scan line is reduced to "essential" elements. Since correlation reduces the number of essential elements to far less than the total number of samples per line, compression is possible by a uniform re-distribution of these relatively few, but irregularly spaced, sample elements.

The device which accomplishes the uniform re-distribution of essential elements is known as an elastic encoder. The outputs of the eleastic encoder consist of a uniformly spaced set of PCM pulse groups representing the brightness of the sample elements and another set of pulses representing the position information.

In the partial instrumentation of this system [10], the run lengths were restricted to three discrete values - 1, 3, and 9 Nyquist intervals. The elastic encoding becomes a digital shifting process when the signal is processed in digital form. The basic encoding operation is shown in Figure 4.

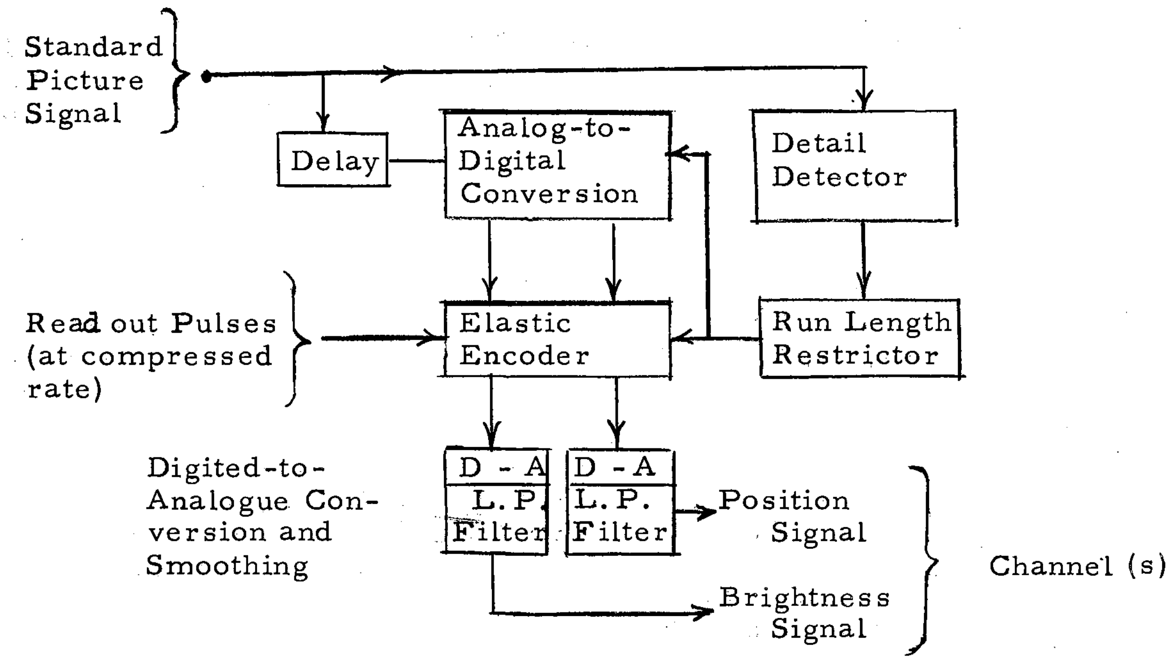


Figure 4. Element Run Coding

A receiver could of course accept the signals which have been encoded by the above process and, by a reverse process, reconstruct a standard picture signal which could then be displayed or re-transmitted in the ordinary manner, or, as the proposers [10] have suggested, a simple receiver could be constructed for directly displaying the picture by using variable-velocity scanning techniques similar to those earlier suggested by Cherry and Gouriet [9]. Geometric distortion should not be a problem because the internal bandwidth of the receiver sweep system could be made as wide as necessary. The basic form of such a receiver is shown in Figure 5.

Actual instrumentation of the system has extended only to a small-scale elastic encoder and to examining the reconstruction of two-level and half-tone pictures from the reduced data at the output of a

Detail Detector.

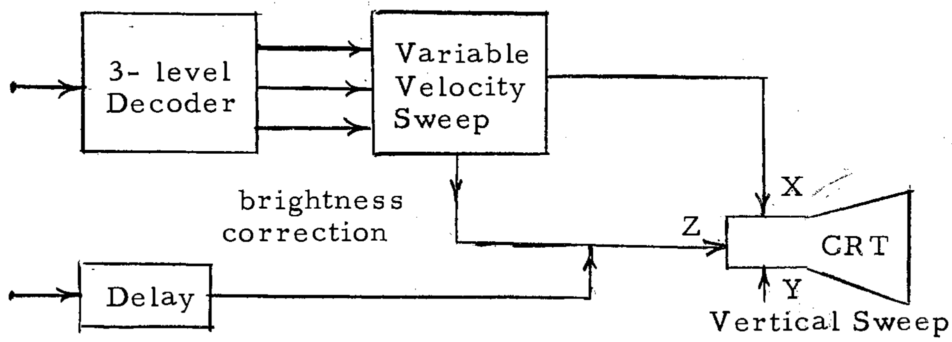


Figure 5. Variable-velocity element run decoder

These studies indicated that satisfactory half-tone pictures are obtainable (by zero-order interpolation) from the reduced data, and that for the pictures used, the data reduction was from 3:1 to 3.5:1. Black and white line drawings could be reconstructed from far less data; the ratios ranging from 6.25:1 to 17.5:1 (by coarser Detail Detector threshold settings).

It appears from these studies that this compression system could be highly successful in line drawing and other two-level picture applications. However, unless the position signal can be more successfully encoded, the system does not offer substantial savings for half-tone pictures. Since the sample-rate for the position signal is the same as for the brightness signal, the two signals require equal bandwidth. The total bandwidth compression possible with the previously mentioned 3.5:1 data reduction is therefore only 1.75:1.

Since the brightness channel can have perhaps 32 to 128 distinguishable levels (depending upon noise), a position signal requiring equal bandwidth but having only three levels - except for its relative noise immunity - uses extravent channel capacity.

Fundamentally, the effectiveness of the system could be improved by allowing more run lengths, i.e., more levels for the position signal. By allowing more run lengths, fewer runs would be required in most lines - thereby further reducing the data rate. Aside from the trivial requirements for additional shift-registers, etc., there remains the fundamental problem that for a given channel noise, each additional level is bought at the price of greater position error rates. There should therefore be a point of diminishing returns where additional compression by finer quantitazation of run lengths is too costly in terms of (subjectively tolerable) position noise.

The best compromise would depend upon the noise and the probability law for run lengths for the particular class of pictures in question. Although not precisely known, run-length distributions are believed [10] to be of the following from:

$$P(n) = (r - 1) \; r^{-n}, \qquad r > 1,$$

where r is a constant parameter and n is the run length in Nyquist samples.

Another approach might be to allow only a few levels to the position signal for the run length encoding process, and the by "brute force" compress the position signal bandwidth by re-coding. For example, if the noise is such that the additional position errors resulting from recoding the three-level signal to nine levels is acceptable, the the bandwidth of the position signal could be reduced by a factor or one-half. With the 3.5:1 data-reduced picture previously mentioned, the total compression possible would be increased from 1.75:1 to 2.33:1.

A more detailed study of these possibilities for improving element run coding is in order, and is perhaps being undertaken by the proposers of the system.

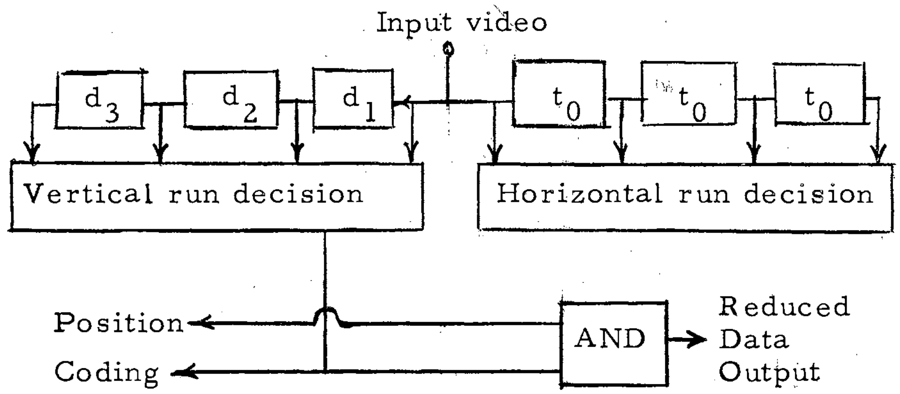
Coding for Vertical Redundancy

As was pointed out previously, correlation in the vertical dimension is about the same as in the hosizontal dimension in most typical pictures. Although no one has proposed or built a practical system for exploiting this vertical redundancy for channel savings, the problem lies more in the areas of technique and economics than in theory.

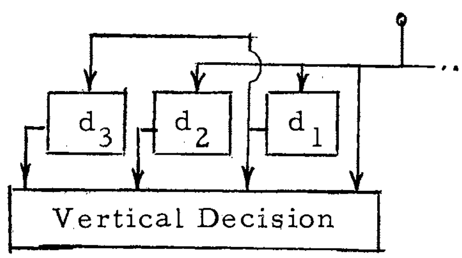
Because of the long delays required for vertical coding of an interlaced picture signal, vertical correlation coding should be more successful when applied to non-interlaced picture signals. For example, a two-dimensional encoder could be built which not only discards previous element value samples by a horizontal "Detail Detector" (element run coding), but also discards those remaining elements which are sufficiently "equal" to previous vertical elements. If the run length decision is determined from the three previous samples, as it was in the inferential Detail Detector of Kubba [25], then three previous sample values are required. Such a system is suggested in Figure 6.

The t_o delays in the horizontal run decision circuit are one Nyquist interval, or .125 μ Sec. for a 4 Mc. picture. If the picture is non-interlaced, the vertical run decision delays d_1, d_2, and d_3 are one line period, or for a <u>field</u> of standard television, 62.5 μ Sec.

However, if we wish to determine runs of actual vertically space-adjacent elements in a standard 2:1 interlaced picture, then delays d_1 and d_3 have to be increased to one field period (16, 650 μ Sec.) and rearranged as shown in Figure 7.



(a) Non-interlaced system



(b) Modified vertical decision for an interlaced system

Figure 6. Two-dimensional coding

An alternate approach for interlaced pictures is to reduce the storage difficulty by using the system of Figure 6 to process only one field at a time and accept the resulting loss of effeciency. The lines of a single standard field are, however, only 28% further apart than horizontal samples in a standard picture [19].

Because of these delay problems, practical experimentation has been very limited. Although many investigators have mentioned the theoretical aspects of vertical or two-dimensional coding processes, e.g. [3, 8, 9, 13, 19, 20, 36, 41, 46, 48, 49], Harrison's linear prediction experiments [19] have been the only attempt at actual realization of removal of vertical redundancy using actual television picture signals.

Perhaps when the storage art has reached a higher state, more interest will be found for two-dimensional picture coding.

CHAPTER V

INTER-FRAME CODING

Intuitive Discussion

The redundancies discussed, to this point, have been associated with regional correlation of intensity in the two dimensions of the picture plane at a fixed time. Television, however, presents a time sequence of pictures, which - due to the nature of visual perception - appears as a continuous scene.

If the individual frames or pictures were formed by a random selection from possible pictures, we would expect to see only a meaningless blurr. The fact that we do, ordinarily, see a coherent scene indicates a considerable correlation in the third-dimension (time) existing between frames.

The implication here is that some form of previous-frame prediction[1] coding might be successful. For example, if a "still" scene is being televised, knowing a particular frame enables perfect prediction of subsequent frames; i.e., subsequent frames convey no actual information. If there is some movement or intensity change in the picture, subsequent frames convey only a relatively small amount of information. If, on the other hand, the "known" frame occurs just

[1] Previous-frame prediction differs from previous-element prediction in that the entire joint state of all elements (a point in picture-variable phase space) is predicted.

prior to a scene change, previous-frame prediction would fail.

The basic idea of coding to exploit inter-frame redundancy is old [22] and has often been mentioned in the literature, e.g. [9, 13, 20, 23, 41, 46, 48] as a possibility. However, no practical system, so far as we know has been built. Perhaps the greatest difficulty has been a need for flexable storage or delay extending over several frames in time.

Inter-Frame Statistics

A second difficulty of inter-frame coding on a statistical basis is the few actual statistics have been measured. Kretzmer [23] measured optically the cross-correlation between adjacent motion picture frames and obtained the figures of 0.80 and 0.86. But what is more important is knowledge of the dynamics or higher order statistics of actual television signals. In order to effectively encode the signal to approach a uniform re-distribution of information [30], an estimate of the necessary queuing storage is needed.

An experimental frame difference signal generator has been constructed by Seyler [42] and Potter [34]. The frame difference signal $g(t)$ is given by

$$g(t) = f(t) - f(t - T),$$

where $f(t)$ is the video signal and T the frame period. In the experimental setup, the difference signal was obtained by using a Vidicon camera tube as a temporary storage device. The Vidicon was subjected to a constant light bias while the input video was applied to modulate the cathode potential; the frame difference signal was obtained on the signal electrode.

The difference signal thus obtained was then sampled at a 10 Mc. rate, and the number of pulses counted whenever

$$\left| g(t) \right| \geq \varepsilon .$$

The pulse-count per frame was recorded.

A record of the difference signal pulse-count was made from ordinary program material, with an ε level of 0.05 of the maximum signal amplitude. Photographs of the count-analogue record clearly show the effects of no motion, little motion, panning, violent motion, and scene changes. Preliminary results (prior to computer processing), indicate that for 1000 frames examined, 50% of the frames had relative counts of less than 15%, and 95% of the frames had relative counts smaller than 50%.

Inter-Frame Coding

A basic approach to inter-frame coding is that after sending an occasional frame, only the difference signal need be sent - within some time interval. A reduction in channel capacity follows if the information rate is re-distributed uniformly over that interval. With the addition of a signal describing the re-distribution, a receiver can then theoretically reconstruct the original picture.

Seyler [41, 42] has deduced from his measurements that in order to benefit from strictly statistical coding based only on the probability of a frame difference signal existing would require averaging or queuing intervals of up to 10 seconds. This interval is based primarily on the frequency of scene changes. Thus, for more manageable or economically justifiable storage times of say 0.1 to 0.2 seconds (three to six frames), it would appear that little savings are possible.

There are, however, some additional considerations which give hope to inter-frame coding. Since nearly all television camera tubes have some "lag" or storage effects, spatial resolution is reduced in areas of the picture involving motion. In addition, there are also the subjective visual effects of exchange between time-motion resolution and spatial detail resolution (to be discussed more fully in Chapters 6 and 7). For these reasons less data are needed to describe a given picture at precisely the times when the frame difference signal arises. However, no law has yet been discovered for this apparently inverse relationship more exact than this intuitive appreciation.

It has also been suggested [41] that instead of transmitting the frame difference signal, it be used by a decision-making element to select for transmission only the frames which differ sufficiently from the previously transmitted one. This concept is known as frame run coding, and is analogous element run coding. It is also implicit that to save channel capacity, the time scale of the selected frames should be uniformly re-distributed over some interval. The previous comments about the necessary queuing interval for frame difference signal coding should apply equally to the interval needed for frame run coding.

CHAPTER VI

CODING TO MATCH VISUAL PERCEPTION

The picture coding problem has been discussed so far only upon a mathematical basis, i.e. the statistical redundancies present in pictures. However, it is improtant to remember that the ultimate "sink" for the image information is the human observer. Data which mathematically are classified as "information" are simply wasted if the observer cannot percieve them. That a system is wasteful of capacity in this sense does not necessarily imply that the quality of the system is too high; it means only that there is an information "mismatch."

Visual Channel Capacity

Recall that the maximum information rate or source capacity for picutes with N brightness levels, K elements, and frame rate R is

$$C_s = (K)(R) \log_2 N \text{ bits/sec.}$$

If we take N to be about 64, then for a standard television picture C_s is about 36 (10^6) bits/sec. It is interesting to compare this figure with estimates for visual channel capacity.

By use of an energy model for the eye, Budridis [6] has estimated a channel capacity of the visual mechanism to be about 2.5 (10^5) bits/second -- two orders of magnitude less than the source entropy for a 64-level picture.

Other investigators [37, 45] have obtained an estimate of from 20 to 50 bits/second, by pattern recognition and reading speed tests. The large difference in estimates can be explained by the different conceptions of the problem which the different investigators used.

In order to discuss these differences, the model of Figure 7 will be postulated. The higher estimate seems to be concerned only with the input-output relations of the first step - and perhaps with a threshold for the second, i.e., with the eye itself and the threshold detection of some aspect of its output.

On the other hand, the psychological investigators, who obtained the lower estimates, must have included the more complex (and slower) operations of step 3, because pattern recognition was involved.

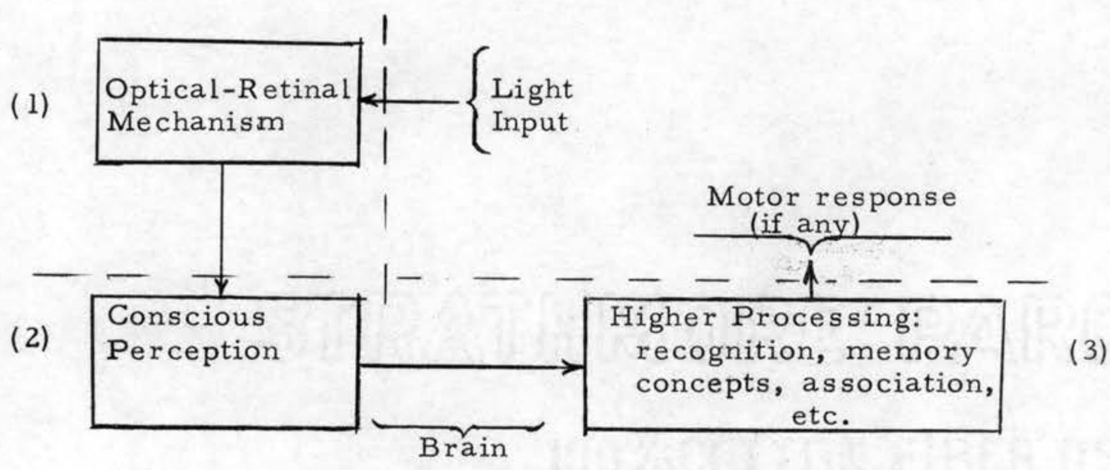


Figure 7. Visual perception model

It may be argued that the lower capacity estimate is better because useful communication necessarily involves cognition. This may be true in a utilitarian sense, but for entertainment purposes, it is desirable to give the illusion of direct visual contact. The presense or absence of the illusion may convey nothing more than just that, but subjectively, it is a very important "bit." Of course imagination permits some departure from the perfect illusion. For example, it is not too difficult to accept black and white pictures, or the limited contrast range and resolution of television.

In accordance with the above discussion, the visual channel capacity for entertainment quality will be taken as nearer the higher estimate of $2.5\ (10^5)$ bits/second. Since the electrical channel capacity is about two orders of magnitude higher than this, yet far from visually perfect, there must be some aspects of the picture which are vastly overspecified. Discovering these over-specifications and exploiting them forms a basis for channel savings.

The general nature of the mismatch between the visual and electrical channels can perhaps be explained in the following manner. The standard television system allows independently for maximum realization of the various image parameters: spatial resolution, brightness resolution, and time resolution. Yet in human vision, the resolutions of these parameters are interrelated. The interrelations are found in several experimental visual laws.

Exchange of Contrast Resolution for Spatial Resolution

Two visual laws relate the threshold visibility of brightness to stimulus area (or visual angle). For small areas less than six minutes of arc on the fovea, an area of angle A is just visible at brightness B when

$$A \cdot B = \text{a constant.}$$

For larger areas, the following law holds:

$$A \cdot B^2 = \text{a constant.}$$

These formulas are known respectively as Ricco's and Piper's Laws [18]. As stated above, they give the absolute thresholds for a dark-adapted eye. They also hold for differential brightness with a given background or adaption level. In this case, B is replaced with the object contrast $|B_s - B_o| / B_o$, where B_s is the object brightness and B_o is the background level.

In terms of picture coding, these laws imply that while large areas require high brightness accuracy or fine quantization, the eye is insensitive to larger brightness errors or coarser quantization in areas of fine detail. Numerous suggestions [17, 21, 24, 37] for compression systems based upon this principle have been made.

Julesz [21] performed computer simulation studies of the bit-rate reductions possible by combining the savings of selectiing samples on a run-length basis with normally fine quantization (element run coding), and then more coarsely re-quantizing the samples in fast-transient regions - where the runs are short.

Kretzmer [24] studied more directly the subjective effects of similar coding in an actually instrumented system. He proposed splitting the spectrum of the picture signal into two or more bands, and quantizing more coarsely the higher bands. Most experiences [7, 14] indicate that if straight quantization is applied, 6- or 7-bit samples (64 or 128 levels) are rquired to portray to the eye gradual shading without visible contouring effects. Thus, the 4-megacycle picture would require (for PCM) a minimum bit-rate of 56 (10^6) bits/second. In his experiment, Kretzmer split the band of the analogue signal into "lows" (below 1/2 Mc.) and "highs" (1/2 - 4 Mc.), and then applied 5-level quantization to the "highs" signal. The experimental setup is shown in Figure 8. No actual encoding was carried out - only the subjective effects of coarse quantization of the highs were observed. The primary subjective objection was that edges were followed by fainter "ghost" images.

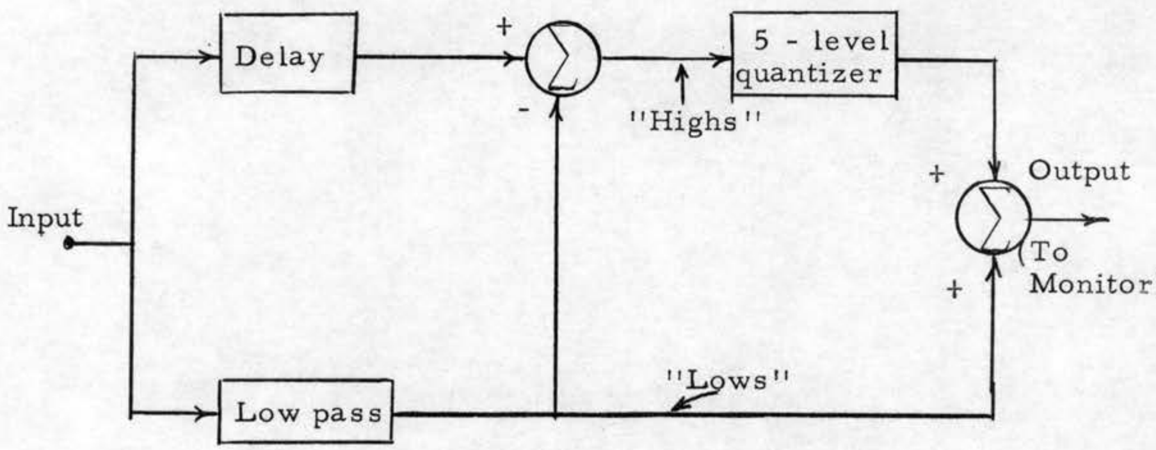


Figure 8. Quantized highs simulation

If the "lows" were quantized to 128 levels and the "highs" to 8 levels, the minimum PCM bit-rate is reduced to 28 (10^6) bits/second. Although without further re-coding no analogue bandwidth is saved, the bit-rate for PCM is reduced 2:1. If noise level permits, the 3-bit highs signal could be re-coded to more levels to reduce bandwidth if transmitted as an analogue signal. Alternately, it was suggested that the "highs" signal be statistically encoded into run-lengths (element run coding).

The system of "Synthetic Highs" of Schreiber and Knapp [39] is somewhat similar to the above proposal except that the quivalent to the quantized "highs" signal is generated by quantizing the "edge" pulses which result from a short-time difference signal. The difference pulses can be uniformly redistributed over some interval for additional (statistical) compression. At the receiver, the edge pulses are decoded and applied to a transversal filter to obtain several standardized height transient signals wich are then added to the "lows."

Since in a picture reproduced by a quantized highs system the high-frequency transients (associated with "edges") are coarsely quantized (stepped), spurious "ghost" outlines trail the true edges. But in the system of Synthetic Highs, the transients can be somewhat in absolute error - yet smooth, and produce no ghost contours. The greater apparent fidelity of Synthetic Highs over Quantized Highs is illustrated by the approximate step-response of each, shown in Figure 9.

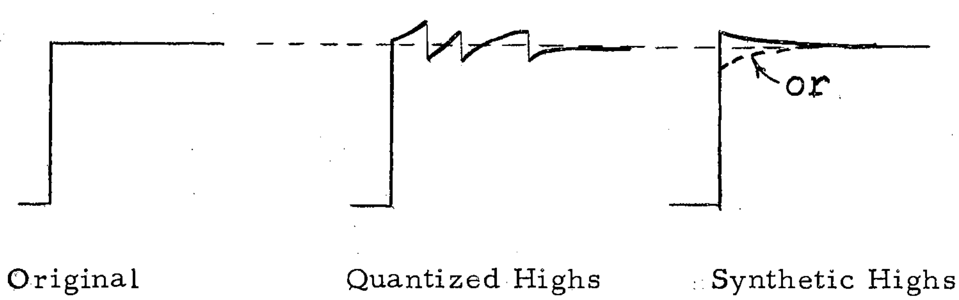


Figure 9. Comparison of quantized highs and synthetic highs reproductions

A third basic variation on the idea of coding for the contrast-detail resolution exchange has been suggested by Graham [17]. The system, known as "differential quantizing," operates on a delta modulation principle where, instead of using a 2-level difference, an 8-level difference is sent. The 8-level quantizer uses a tapered staircase such that areas of gradual change are automatically reproduced with fine quantization, and areas of rapid change (edges) are automatically resolved with coarse quantization. The 30 bit/sample picture coded by this method is comparable to an ordinary 6- or 7-bit picture.

A similar alphabet reduction for picture transmission has been achieved by Roberts [37] by a remarkably different process. Before quantization, pseudo-random noise of one quantum level is added to the analogue picture signal, and at the receiver, identical noise is generated and subtracted from the signal. The visual result is that the error introduced by quantization is transformed into additive noise. By this simple technique, the data rate can be reduced from 6 to 7 bits/sample to 2 or 3 bits/sample (4 to 8 levels).

That this system works, actually depends upon both picture correlation and subjective effects. Because of the large amount of correlation in pictures, the eye can easily distinguish between a certain amount of random noise and the picture. If the pseudo-random noise generators are synchronized to re-start every frame the noise structure would be visibly stationary, and only two-dimentional correlation in the picture is utilized. If the pseudo-random noise cycle runs for several frames before re-starting, then inter-frame (time) picture correlation further aids the eye in signal/noise discrimination. In the Roberts experiment, only single pictures were processed by a computer.

It is also of interest to examine why this system works in explicit terms of the visual law for brightness/detail resolution exchange. Since the visual defects of coarse quantization are caused by the eye's fine contrast resolution in larger areas, the addition of random noise breaks up the large areas so that they exist only as averages, but leaving the eye only very small dots with which to discriminate contrast. Large contrast steps are therefore allowable.

Continuity of Vision

Two aspects of the persistance of vision can be distinguished: first, the threshold ability of the eye to distinguish between separate flashes of light and a continuous light; second, the ability of the eye to distinguish between a rapid sequence of stationary pictures, representing samples of motion, and true continuous motion.

The repetition frequency f_r where flicker is just perceptable is not a fixed number but is proportional to the logarithm of brightness B. This relation is known as the Ferry-Porter law [12]:

$$f_r = 12.57 \log_{10} B + C \text{ cycles/sec.}$$

The constant C depends upon a great many variables of viewing conditions. At any rate, many subjective tests [12] have indicated that under ordinary television viewing conditions a field rate of 50 to 60 per second for maximum brightnesses of about 20 to 100 footlamberts respectively. Thus, the standard field rate of 60 per second (which matches the power frequency) is necessary for flicker free pictures at desirable brightness levels.

The visual requirements on repetition rate for continuity of motion is much easier to satisfy. The threshold of perceptability where separate images have the appearance of continuous motion is primarily a function of image contrast and speed. More pictures per second are required for an increase in either contrast or speed. The rate of 30 frames per second used for television produces satisfactory motion fusion for most scenes, although 24 frames/second is standard for motion pictures. Thus with a field rate of 60 frames/sec. and the frame rate of 30 frames/sec. the standard 2:1 interlace ratio simultaneously satisfies these two continuity requirements.

In actual practice, the theoretical vertical resolution of interlaced pictures is not realized because of finite spot size and imperfect interlace. For this reason, Gabor and Hill [13] have proposed discarding every other field, and interpolating at the receiver. If the previous field is simply repeated or if linear interpolation between fields is employed, the reproduction of non-vertical contours can suffer. To combat this problem Gabor proposed interpolating the missing field by a method known as "contour interpolation." In contour interpolation, not only is the intensity of the missing line linearly interpolated between the previous and next lines, but when "edges" are detected, the horizontal position of the edge in the interpolated line is also interpolated between those of the previous and next lines. The principle is illustrated in comparison to simple repetition and linear interpolation in Figure 10. The shaded functions are intensity versus horizontal position. The center line belongs to the missing field.

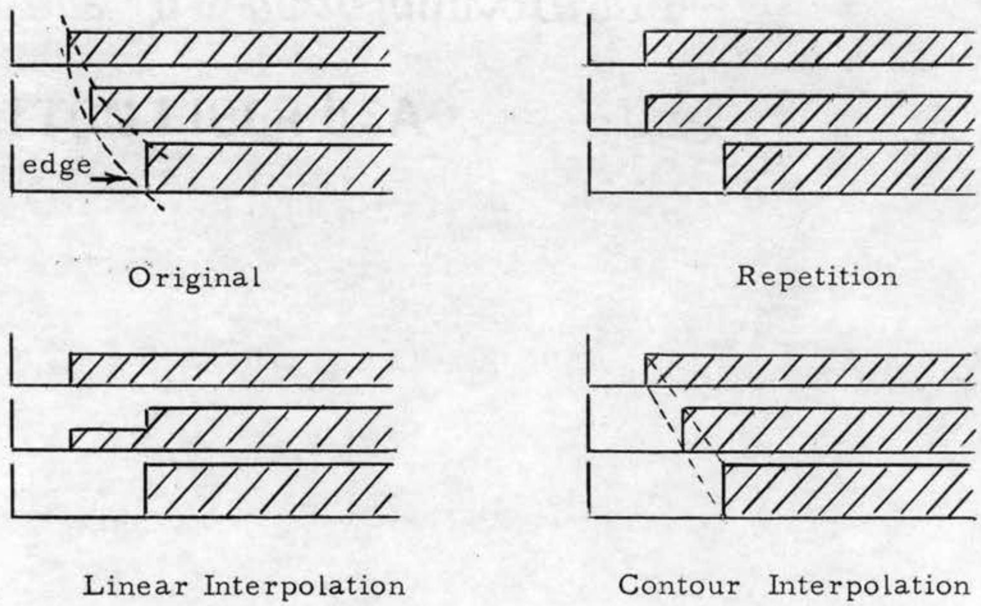


Figure 10. Zero-order, first-order, and linear interpolation

The visual superiority of contour interpolation was demonstrated with a photo-mechanical realization. Electronic realization, while theoretically possible, poses some extremely difficult practical problems.

Exchange of Spatial Resolution for Time-Motion Resolution

While the standard television system is prepared to resolve picture samples of images moving or changing at any rate, or even a different scene each frame, the eye is not. The fact that the eye cannot maintain full acuity when pursuing a moving image is easily demonstrated by a common experience with a motion picture film. Visible blurring of the moving images are seldom noticeable. Yet if a single frame is held stationary and examined, it can be severely blurred because the image was in motion during exposure. The loss of resolution in a film is ordinarily not very noticeable because it closely matches the loss of acuity in the eye.

The loss of aciuty during pursuit of a moving image has been shown to increase as the cube of angular motion [26]. It is thought to result more from imperfect control of the pursuit (hunting) than to simple inability of the eye to move fast enough.

Another visual phenomenon, which is closely related (from an information rate viewpoint) is the temporary loss of acuity following a complete scene change. In studying this effect, Seyler and Budrikis [43] found that if the resolution at the beginning of a new scene in films or television can be reduced by a factor of 20 from the nominal value without being subjectively noticeable if it is gradually restored in about 0.8 second or 24 standard frames.

As was suggested in Chapter V, the possibilities for inter-frame coding may be enormously improved by exploiting the time-motion dependence of visual acuity.

CHAPTER VII

A SYSTEM OF INTER-FRAME CODING

Although some progress toward instrumentation of compression systems exploiting element and/or subjective visual effects has been made, no one can yet begin to build a statistically coded inter-frame system. The system to be suggested here departs somewhat from ideal statistical coding, but does so in such a way as to ease the storage problems and exploit subjective visual effects.

Frame Run Coding

To appreciate the problems of statistical inter-frame coding, consider the basic technique of frame run coding. Suppose there is a decision making device which examines an ordinary picture signal and by some criterion[1] answers the question "Does the present frame (as a whole) differ sufficiently from the previous one?" If the answer is "yes" then the frame in question is sampled (gated). If the answer is "no" it is suppressed. This principle is illustrated in Figure 11.

Presumably for a given decison function and class of pictures there is a probability law for the number of frames that may be selected in an interval. Let P(G) be the probability that a frame

[1] One "reasonable" criterion might be for the mean squared frame difference signal to exceed a threshold.

be sampled. Experimentally this is

$$P(G) = \lim_{N \to \infty} \frac{n}{N},$$

where n is the number of gated frames in an interval of size N.

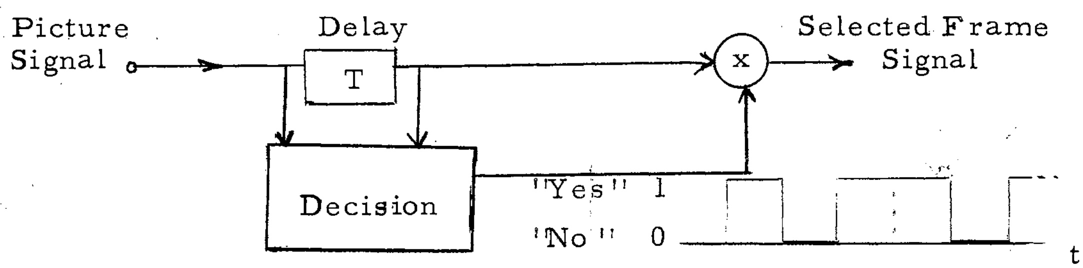


Figure 11. A frame run decision device

P(G) then defines the ideal data reduction by pure frame run coding. That is, for large intervals N, the number n of frames sample converges in probability to NP(G). Without knowing the distribution of n on N, about all that can be said about the required interval length N is to set an upper bound as given by the Chebyshev inequality:

$$P\left(\left| P(G) - \frac{n}{N} \right| \leq \varepsilon \right) \geq 1 - \frac{\text{Var. }(n/N)}{\varepsilon^2} .$$

If it is assumed that n is obtained from independent trials (which it is not), then

$$\text{Var. }(n/N) = \frac{P(G) - P^2(G)}{N},$$

which is maximum for $P(G) = 0.5$. In this case

$$\text{Max. Var}\,(n/N) = \frac{1}{4N} .$$

The assumption of independence makes the distribution of n binomial, but direct use of the binomial distribution cannot be justified. Accordingly, the maximum variance of (1/4N) will be used only in the Chebyshev inequality to obtain a more realistic approximate bound on N. Then

$$P\left(\left|P(G) - \frac{n}{N}\right| \leq \varepsilon\right) \geq \frac{4N\varepsilon^2 - 1}{4N\varepsilon^2} .$$

For example, to make this probability larger than .95, with an ε of 1/10, N must be 500 or more. For the probability of less than 10% error can be as more than .75, then N should be at least 100. These figures imply a rather large queuing interval (in agreement with Seyler's 10 second estimate) on which the uniform time scale re-distribution must be made for actual bandwidth compression. Such a re-distribtuion is illustrated in Figure 12. The "X" blocks represent sampled frames, and T is one frame period.

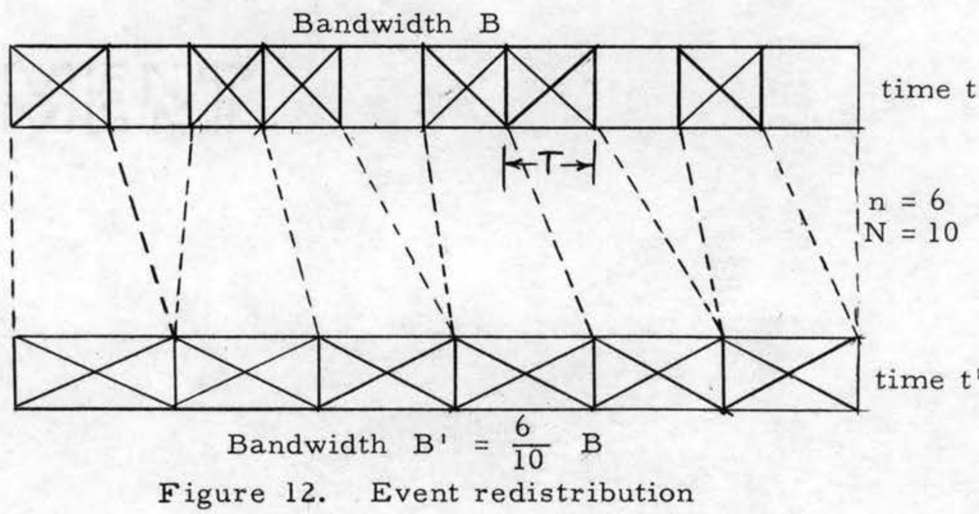


Figure 12. Event redistribution

At the receiver, the original signal could be recovered by restoring the original time scale and ralative position of the gated frames, and simply repeating the previous frame signal during "empty" intervals. If the interval N is chosen to be small (say 5 - 10 frames), the ratio n/N would fluxuate considerably. For a large N (say 300 frames) the n/N ratio should fluxuate only slightly - so that a bandwidth of slightly more than BP(G) should suffice. But the required storage is enormous.

If the interval is restricted to the more "manageable" size of from 5 - 10 frames (say), then it is advantageous to utilize the time-motion spatial resolution exchange effect. Perhaps the simplest procedure here would be to allow some (smaller) bandwidth B' for the re-scaled signal[2], perhaps even less than BP(G). At the times when few frames are gated (little motion etc.), the picture resolution would not suffer. But at other times when n/N exceeds B'/B, the horizontal resolution would be reduced - but not subjectively, since by the visual exchange laws the eye would lose acuity at exactly these times.

While the above procedure would somewhat simplify instrumentation problems over straight statistical frame run coding, these problems are still quite difficult. The storage system at the encoder must have the capacity to "read in" n frames of signal at bandwidth B, and then "read out" at a rate that will "spread" the time scale of n collected frames over the N-frame interval. The decoding system

---

2 The sample rate could be similarly reduced if the transmission is digital .

must have the same capacity, but would operate in reverse, and must have the further ability to repeat the re-scaled signal over the "empty" intervals.

There is little possibility of avoiding the time scale change and still saving effective bandwidth by time multiplexing the n sampled frames with other signals, since n is random and fluxuates between 1 and N. To escape some of these difficulties of non-uniform sampling, the following system of uniform frame sampling is suggested.

Uniform Frame Sampling

Let one frame in every N be gated; the other (N - 1) being suppressed. Assume (for the moment) that a method will be found to recover the lost "information." If the original video bandwidth is B, the gated signal can now be transmitted through a channel with effective bandwidth B/N in one of two ways. First, the bandwidth can actually be compressed by a change of time scale by "spreading" the one sampled frame over the N-frame interval. The problem is considerably simplified over frame run coding because only one frame at a time is processed rather than a random number, and the problem is now reduced to one of standards conversion.

Second, if a channel of bandwidth B is used, N picture signals can be sent through it by time multiplexing, i.e., staggering the sampling times. Of course the frame periods of the N signals must be synchronous.

These two alternative operations are shown in Figure 13. The sequences (s) of sampled frames are recovered at the output of the channel by the inverse operations.

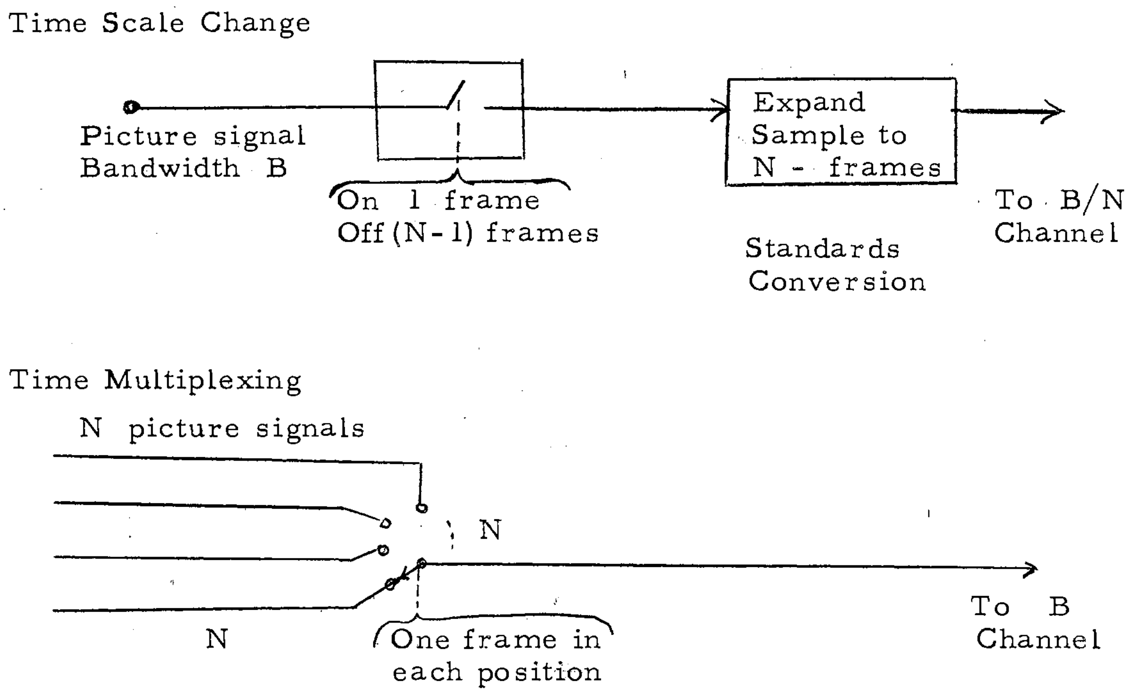


Figure 13. Uniform frame sampling systems

The true problem now arised of how the original (or visually approximate) picture signal is reconstructed from the received sequence of sample frames.

Picture Reconstruction

As a first approximation to the original picture signal, suppose each of the received sample frame signals is simply repeated (N-1) more times to fill in the interval. If f(t) is the original signal, in a particular N-frame interval, let this "zero order" approximation be called $\hat{f}(t)$. Let $f(t)$ be represented as

$$f(t) = \sum_{k=1}^{N} f_k(t) ,$$

where f_k is f in the k-th frame in the interval and zero elsewhere. Then

$$\hat{f}(t) = \sum_{n=1}^{N} f_1(t + T - nT)$$

where T is the frame period.

If $\hat{f}(t)$ were displayed, there would be no visible flicker, since the original field rate is restored; nor would there be any reduction of spatial resolution. With no movement in the pictured scene, the reconstruction $\hat{f}(t)$ would be perfect. However, if the original frame rate were 30 per second, new pictures (sampling any changes) are reconstructed at the rate of only 30/N per second. The subjective impression would be motion discontinutity of "jerkyness." This effect would of course be more pronounced for larger values of N. It is necessary then, to find a method of interpolation which also somehow "fills in" to restore motion continuity. Two approaches to this problem are to be suggested.

Transmitted Error

If, at the encoder, the signal for the gated frame can be somehow stored, it can be compared to each of the subsequent (N-1) suppressed frames in the interval. The difference signal so obtained represents the information lost by not transmitting the (N-1) suppressed frames per interval. In other words, let $\hat{f}(t)$ be generated at the encoder and a difference signal $f_D(t)$ be

$$f_D(t) = f(t) - \hat{f}(t).$$

If f_D could be added to $\hat{f}$ at the receiver, then obviously the original f would be obtained. Of course transmitting f_D as such would defeat any bandwidth savings since it has a spectrum of width B. But what if f_D is passed through a low pass filter with cutoff frequency B' cycles per second ($B' < B$)? Let this filtered difference signal be denoted by $f_{DL}(t)$. Consider the new reconstruction $\hat{F}$, generated at the receiver by adding $\hat{f}$ and f_{DL}:

$$\hat{F}(t) = \hat{f}(t) + f_{DL}(t) .$$

f_{DL} would be transmitted, for a total bandwidth of $(B' + B/N)$. What might the subjective impression of $F(t)$ be?

For the case when $f_D(t) = 0$ (no motion, etc.), then $f(t) = \hat{f}(t)$. So that

$$\hat{F}(t) = f(t),$$

and the reconstruction is of course perfect. When $f_D(t) \neq 0$, then f_{DL} provides a low frequency (therefore low resolution) "fill in" for $\hat{f}$. While this "fill" is of low maximum resolution, it does restore motion continuity. Because of the visual effect of exchange between motion resolution and spatial resolution, the resulting reduction of resolution of moving images should not be subjectively noticeable within some limits on N and B'. Ultimately these limits can be found only by subjective tests.

There are still some difficulties, that can be imagined, which must be dealt with. Suppose a scene change occurs in the middle of an N-frame interval. The low frequency components are fully corrected by f_{DL}, but an error consisting of the high frequency components of $\hat{f}$ persists for the remainder of the interval. This error would cause a brief "ghost" of edges from the old scene. This fault could be corrected by inserting a decision element which would monitor f_{DL} at the receiver, such that if a scene change is detected (by a suitable test on f_{DL}), a filter is applied to $\hat{f}$ to remove the high frequency components and give only "lows." Under these conditions,

$$\hat{F}(t) = \hat{f}_L(t) + f_{DL}(t)$$

$$= f_L(t) .$$

That is, the approximation is reduced to a filtered (low frequency components) version of the true signal f(t). As was discussed in Chapter VI, this procedure is visually permissable as long as full bandwidth is soon restored. Perhaps it would suffice to simply switch in a filter for the remainder of the interval instead of gradually restoring resolution.

Noise Effects

Another difficulty which must be considered is the great enemy of most compression schemes - noise. For the transmitted error system additive noise can be classified as input noise and channel noise. The input noise is that existing with the original picture signal before any processing by the system, and channel noise is that added to each of the two transmitted signals.

First consider the input noise. Let the input f(t) be the sum of signal $s(t)$ and noise $n(t)$:

$$\hat{F}(t) = \hat{f}(t) + f_{DL}(t)$$

$$= \sum_{k=0}^{N-1} [s_1(t-kT) + n_1(t-kT)] + [s(t) + n(t) - \sum_{k=0}^{N-1} s_1(t-kT) + n_1(t-kT) \quad]_L ,$$

where the "L" subscript indicates the bracketed function was filtered to remove spectral components above B'. Since the filter is linear, we have

$$\hat{F}(t) = S_L(t) + \sum_{k=0}^{N-1} [s_1(t-kT) - s_{1L}(t-kT)]$$

$$+ n_L(t) + \sum_{k=0}^{N-1} [n_1(t-kT) - n_{1L}(t-kT)]$$

$$= [s_L(t) + \sum_{k=0}^{N-1} s_{1H}(t-kT)]$$

$$+ [n_L(t) + \sum_{k=0}^{N-1} n_{1H}(t-kT)],$$

where

$$s_1(t-kT) - S_{1L}(t-kT) = S_{1H}(t-kT)$$

and

$$n_1(t-kT) - n_{1L}(t-kT) = n_{1H}(t-kT) .$$

Thus, like the picture, the low frequency components of the input noise are unaffected by the processing of the system; while the high frequency components are those in the sample frame, repeated N times. The high frequency (fine grain) noise may thus become more visible because of its increased time coherence. This may not be serious unless the input noise is large, because the contrast-detail exchange law makes fine grain noise the least visible part [5, 4, 29].

Also, there is always the possibility of reducing the power of input noise by a factor of $1/N$ by obtaining the sample frame signal as an element average over the N frames rather than a single sample. It is a safe assumption that the noise is essentially uncorrelated one frame period away, and that the picture and noise are uncorrelated. As a consequence, the noise power in the sample frame is reduced to $1/N$ of its former value. Thus, the input noise is (theoretically at least) not a problem.

The problem of additive channel noise is more serious. Let the noise added to the difference signal f_{DL} be n_{DL}, and the noise in the sample frame channel be n_S. The total noise n (from the channels) is the sum:

$$n(t) = n_{DL}(t) + n_S(t).$$

n_{DL} has a spectrum extending to B', and has no special characteristics introduced by the system. But since n_S is associated with the sample frame, it occurs in the picture as repeated N-frame blocks. This N-frame repetition will no doubt increase the visibility of n_S. The capacity of the sample frame channel is not therefore truely reduced by $1/N$ from the original picture channel, but something less, since the signal-to-noise ratio must be somewhat increased.

Linear Interpolation

An alternative to transmitting the frame difference error is to actually interpolate at the receiver between sample frames. Consider the difficulties of truely interpolating for motion or other changes in images. In Figure 14 are two sample frames of an image and, in the center a visual linear interpolation.

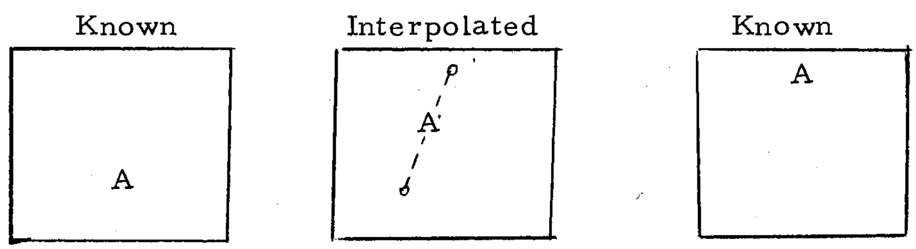


Figure 14. Visual interpolation

This is visually easy with the assumption that the class of pictures is restricted to those where a single figure translates on a white background. The visual interpolation apparently depends upon pattern recognition as well as the mathematical process (of finding the midpoint of a line segment in Figure 14). The problem is of course much more difficult for the more general class of television pictures. One simplified pattern recognition-interpolation process is that of contour interpolation [13], where edge detection is employed. Instead of interpolating between two lines, as was discussed in Chapter VI, the method might be used to interpolate a missing line between two frames. This method would require a rather versatile and extensive storage system in order to accomplish the necessary geometrical transformations, and would still allow only for horizontal motion.

Here an interpolation in only one dimension will be suggested and examined. The approximation $\hat{f}$ to the true signal f is of zero-order for element brightness. As a possibly better approximation consider the next higher order, i.e., a linear "cross fade."

A sample frame is transmitted and N frames later a new sample frame is transmitted. Let an element in the first sample frame have value (brightness) x_1, and let the same element in the new sample have value x_{N+1}. The same element in the k-th frame in the interval has value x_k. This notation will avoid the necessity of explicitly writing the time parameter and keeping track of time shifts. In this notation, then, $\hat{f}$ (t) gives an estimate $\hat{x}_k{}'$ for x_k as follows:

$$\hat{x}_k{}' = \begin{cases} x_1, & k = 1, 2, \ldots N \\ x_{N+1}, & k = 1 + N, 2 + N, \ldots \\ \text{etc.} \end{cases}$$

In this notation, the linear interpolation estimate $\hat{x}_k$ for x_k is

$$\hat{x}_k = \frac{1}{N}(N + 1 - k)\, x_1 + \frac{1}{N}(k - 1)\, x_{N+1}$$
$$k = 1, 2, \ldots N, \; N + 1.$$

Note that $\hat{x}_1 = x_1$, and $\hat{x}_{N+1} = x_{N+1}$. The interpolated values are each a proportional "mixture" of x_1 and x_{N+1}. The process is illustrated with an image in Figure 15. (N = 3).

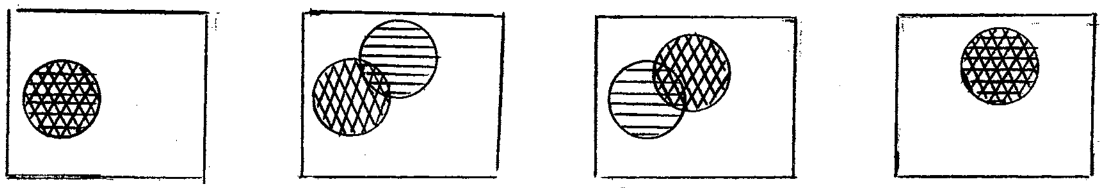


Figure 15. Linear interpolation

In general, higher orders of interpolation of this type would probably give no better results; the elements would still be treated independently. Visually, this interpolation is undoubtedly better than zero-order ($\hat{f}$), since some clues about where the image is "going" are given ahead of time. Unless the movement is very rapid, this cross fade approximates the low resolution "fill" as before; it is very close to the type of blur in the direction of motion that occurs in motion picture films. The number N of frames per sample must as before be found by subjective tests.

Noise Effects

Let x_1 now represent the element noise in the first sample frame, and x_{N+1} the noise of the same element in the next sample. This noise consists of both input noise and channel noise. They will not be distinguished, except to mention that the input noise component can be reduced as before by taking an averaged sample. Let σ^2 be the variance of x_k. It is assumed that

$$E(x_k) = 0, \text{ and } E(x_1 x_{N+1}) = 0.$$

Two measures of the noise in the reconstructed picture will be used: variance, and the (normalized) correlation coefficient between elements in the j-th and k-th frames in the interval. The correlation coefficient is used as a measure of visibility because a repeated noise structure is more visible than an uncorrelated one.

For zero-order interpolation, i.e.

$$\hat{x}_k' = x_1, \qquad K = 1, \ldots N,$$

the variance in the k-th reconstructed frame σ_k^2 is found as

$$\sigma_k^2 = E(x_1^2) = \sigma^2,$$

k in the N frame interval and ρ_{jk} the correlation coefficient (for j, k in the interval) as

$$\rho_{jk} = \frac{E(\hat{x}_j' \hat{x}_k')}{\sigma_j \sigma_k} = \frac{E(x_1^2)}{\sigma^2} = 1.$$

These results were discussed before (qualatatively) as the effects of channel noise in the transmitted error system.

Now, for first-order interpolation, i.e. where

$$\hat{x}_k = \frac{1}{N}(N + 1 - k)\, x_1 + \frac{1}{N}(k - 1)\, x_{N+1},$$

find σ_k^2 and σ_{jk}.

$$\sigma_k^2 = E[(\hat{x}_k - \bar{x}_k)^2] = E(\hat{x}_k^2)$$

$$= E\frac{1}{N}[(N+1-k)^2 x_1^2 + 2(k-1)(N+1-k)x_1 x_{N+1} + (k-1)^2 x_{N+1}^2]$$

$$= \sigma^2 \frac{(N+1-k)^2 + (k-1)^2}{N^2}, \quad k = 1, 2, \ldots N+1.$$

Thus, the linear interpolation process actually reduces the noise power in the reconstructed picture. This law is plotted in Figure 16, for N = 4.

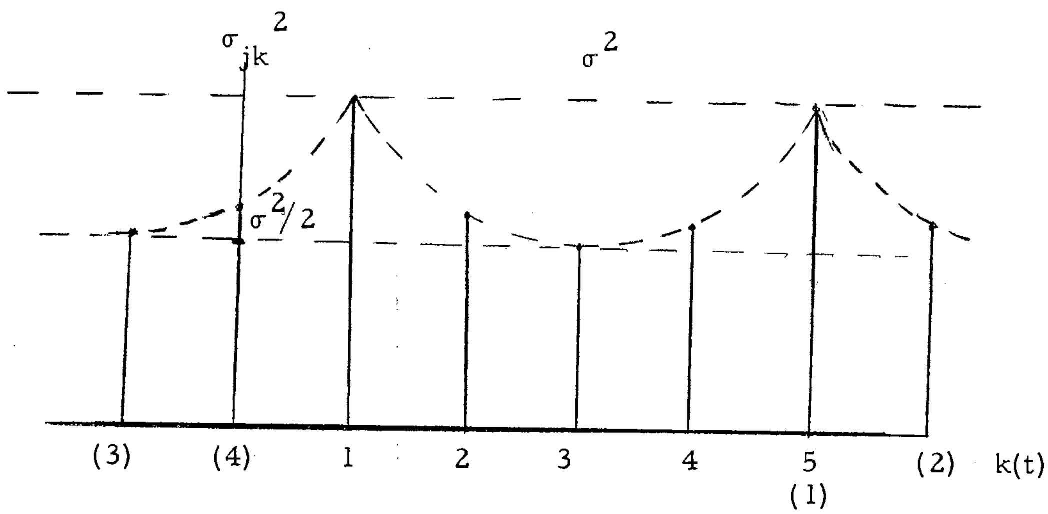


Figure 16. Noise power law for linear interpolation

The correlation coefficient ρ_{jk} is found as follows:

$$\rho_{jk} = \frac{E(\hat{x}_j \hat{x}_k)}{\sigma_j \sigma_k}, \quad \text{since } E(\hat{x}_j) = E(x_j) = 0$$

$$= \frac{(1/N^2)E[(N+1-k)(N+1-j)x_1^2 + (k-1)(j-1)x_{N+1}^2 + C_{jk}]}{(\sigma^2/N^2)[(N+1-k)^2+(k-1)^2]^{1/2}[(N+1-j)^2+(J-1)^2]^{1/2}}$$

where $C_{jk} = (N+1-k)(j-1)x_1 x_{N+1} + (N+1-j)(k-1)x_1 x_{N+1}$, and $E(C_{jk}) = 0$. Thus,

$$\rho_{jk} = \frac{(N+1-k)(N+1-j) + (k-1)(j-1)}{[(N+1-k)^2 + (k-1)^2]^{1/2}[(N+1-j)^2+(j-1)^2]^{1/2}}.$$

To complete the example where N = 4, computing ρ_{jk} gives the following correlation coefficient matrix.

$$(\rho_{jk}) = \begin{pmatrix} 1.000 & 0.948 & 0.707 & 0.316 & 0.000 \\ 0.948 & 1.000 & 0.895 & 0.600 & 0.316 \\ 0.707 & 0.895 & 1.000 & 0.895 & 0.707 \\ 0.316 & 0.600 & 0.895 & 1.000 & 0.948 \\ 0.000 & 0.316 & 0.707 & 0.948 & 1.000 \end{pmatrix}.$$

CHAPTER VIII

INSTRUMENTATION OF AN INTER-FRAME SYSTEM

In the previous chapter, two variations of an inter-frame compression system using uniform frame sampling were proposed. Although little was said about actual hardware, the primary purpose and advantage of uniform frame sampling over statistical frame coding is the simplification of the storage hardware problems.

For transmitted error, two signals must be transmitted: $f_1(t)$ and $f_{DL}(t)$. $f_1(t)$ is simply the sample frame, which may or may not be transmitted with a time scale change. $f_{DL}(t)$ is given by

$$\begin{aligned} f_{DL}(t) &= [f(t) - \hat{f}(t)]_L \\ &= f_L(r) - \hat{f}_L(t), \end{aligned}$$

where

$$\hat{f}(t) = \sum_{n=1}^{N} f_1(t-nT+T).$$

By inspection, $\hat{f}$ is generated by delaying or repeating f_1 N times. Thus, a repeater or recirculating storage of one frame period (1/30) sec.) is required. Since at the transmitter $\hat{f}$ is used only to generate f_{DL}, the storage system need have a bandpass of only B' cycles per second (corresponding to the "L" subscript). This condition will not be true at the receiver where the repeater must operate at full picture bandwidth B.

For such large delay-bandwidth product requirements, electrical delay lines are not practical, and some form of space storage is called for, such as an electrostatic storage tube, an acoustic delay line, or a rotating magnetic storage drum or disc. At present, the storage drum seems best suited. Such units, which store one frame of standard television around the periphery (at 1800 revolutions per minute) and have several tracks, are now available [2, 27, 47].

With the assumption of a magnetic drum storage system, the block diagram of the transmitted error system where no time scale change is used is shown in Figure 17. The (N-1) other sample frame signals are frame period multiplexed with the single sample shown. Details of the sample frame repeater are given in the Appendix. The delay of f(t) before obtaining a difference is necessary to compensate for a slight overall delay in the repeater. This delay requires a second track on the drum and not incidently compensates for distortion in the repeater by equally distorting f(t) . The overall delay of the repeater at the receiver should match that of the repeater at the transmitter. The small delay (electrical network) inserted in the sample frame channel is to compensate for the delay characteristic of the low pass filter.

If an actual time scale change is used, a scan converter is required at the transmitter and the receiver. At the transmitter, the gated sample frame is to be stored in the converter device, and "read out" over the N frame period for a time scale change from t to Nt. The converter at the receiver must reverse the process, i.e., "squeeze" the time scale or rate of events back down to the original. Since the scan converter cannot predict the future, an overall delay of (N-1) T is part of the scan conversion at the receiver. That is, NT is transformed to t + (N-1)T. Consequently it is necessary to also delay $f_{DL}(t)$ for (N-1)T sec., so that the two signals to be added at the receiver are synchronous. The system is diagramed in Figure 18.

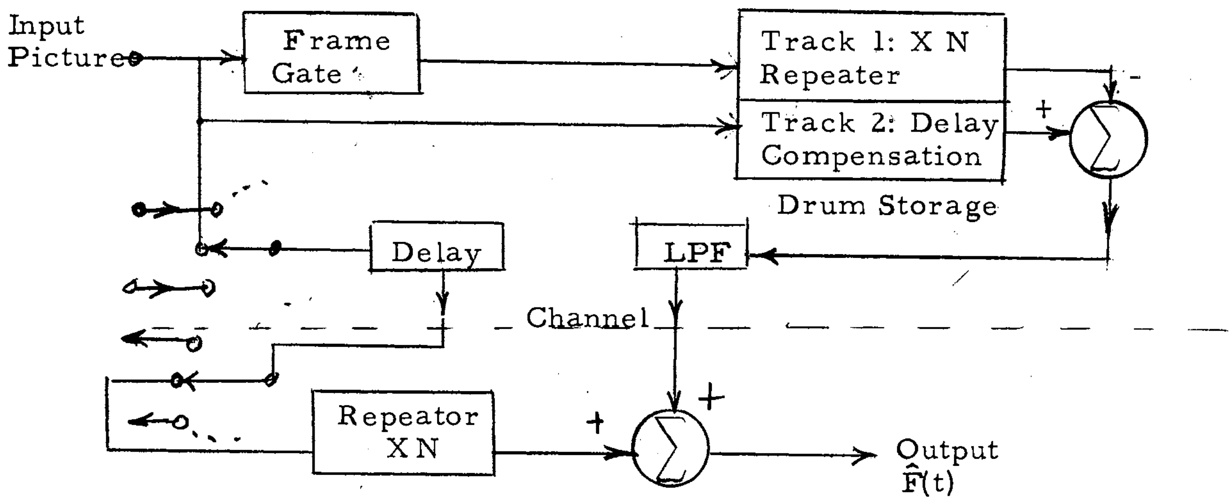


Figure 17. Transmitted error/multiplex system

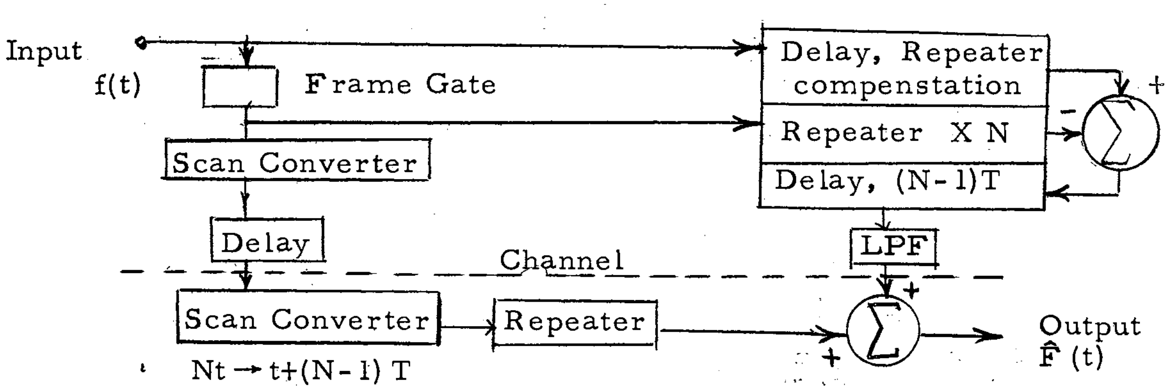


Figure 18. Transmitted error/time scale change system

Again, the details of track connections on an 1800 rpm drum for the delays and repetitions are covered in the Appendix.

One possibly suitable device for the time scale changes is the electrostatic storage tube with simultaneous READ/WRITE capability. For the time scale expansion, the writing and reading beams would each sweep the same raster on the storage surface, but starting simultaneously, the writing sweep cycle is completed in one frame period (and the beam blanked for the remainder of the interval) while the reading beam sweep cycle takes the full N-frame interval. For the time scale compression, the writing beam would be swept at the slow rate while the reading beam is swept at normal (fast) frame rate. To prevent the reading beam from "running out of new picture" the unblanked READ cycle is not started until one T-period before the WRITE cycle is complete. Hence, there is an overall delay of (N-1) T.

A major difficulty to be expected with a storage tube system is erasure time. Single storage tubes operated continuously in the above cycles would either have to have a fully destructive and priming readout, or be capable of a fast ERASE cycle during the vertical blanking period - at present rather difficult reuqirements. Of course two or more storage tubes with erase times too long to satisfy the above could alternate as the active storage device, with the off period used for erasing.

A simple alternative to the use of a storage tube for scan conversion is to face a kinescope with a camera. The camera tube is chosen with sufficient storage [34]; readout is essentially destructive

[34]. For the time scale expansion, the kinescope replaces the operation of the writing beam while the camera tube reads out the image stored on the target with a slow (NT) scan. Similarly, to re-compress the time scale the sweep speeds are reversed and appropriately timed.

There are essentially no additional hardware problems associated with the method of linear interpolation of element brightness. Only the sample frames are transmitted - either by frame multiplexing, or with time scale change. (See Figure 13, page 47.) At the receiver however, the approximation $\hat{x}_k$ is to be generated, where

$$\hat{x}_k = (1/N)[(N + 1 - k)x_1 + (k-1)x_{N+1}].$$

The equation can also be written as

$$\hat{x}_k = x_1 + \frac{k-1}{N}(x_{N+1} - x_1),$$

i.e., the last sample plus a (linearly) weighted difference between the present and last samples.

This formulation suggests the real time receiver operations shown in Figure 19. The repeater makes the samples available for each value of k, and the NT delay present and next sample, or more accurately last and present samples, both available.

The NT second delay requires only two T-period tracks on a drum because of repeated nature of what is being delayed. The repeating operation requires a single track as before. For details see the Appendix.

The preceding comments were intended only to show that it is technically possible to build the specified system(s), and are not of course a detailed design. The question of subjective effects has yet to be answered.

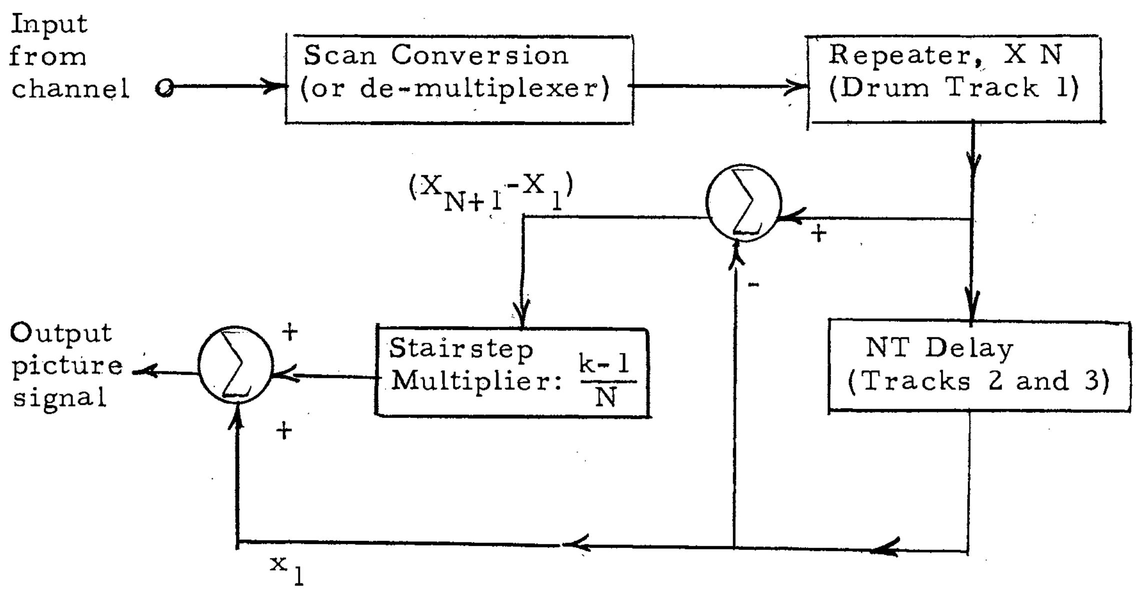


Figure 19. Element interpolation receiver

Simulation

Answering the subjective visual questions ultimately requires experimentation. A good simulation should duplicate exactly the effects of the system. We are somewhat skeptical of anything less than an actual television display of actual program material. A good simulation may therefore involve almost the effort and cost of the actual system. At least, though, neither a time scale change nor a multiplex operation is basic to the two kinds of picture processing to be observed, and can therefore be deleted.

According to these criteria, the system of Figure 20 should provide a good simulation for the effects of the transmitted error system. The primary system parameter N can be varied at will by changing the erase timing on the repeater track of the drum. The bandpass allowed for the error signal is varied by changing the cutoff frequency B' of the low pass filter. Additional input noise could be added at N_1. Sample frame channel noise and difference signal channel noise are accounted for at N_2 and N_3, respectively.

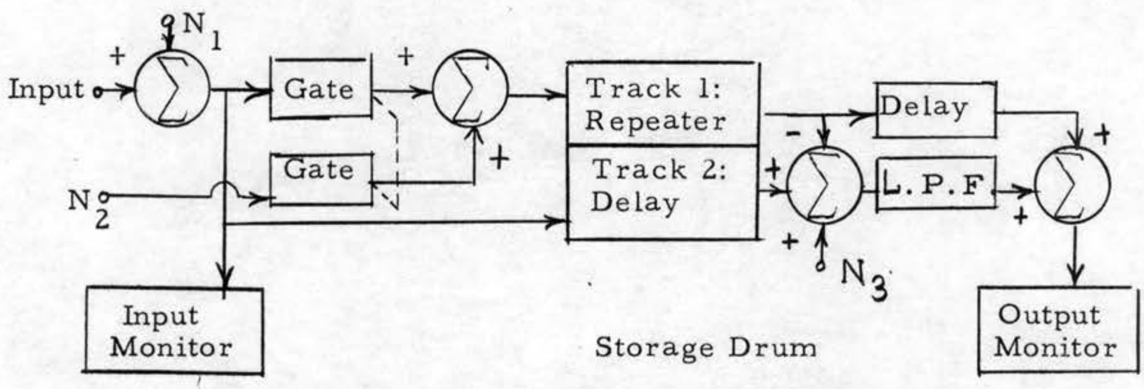


Figure 20. Transmitted error simulation

Similarly, the visual effects of linear element brightness interpolation could be experimentally tested by the system of Figure 21. Following the Frame Gate, the circuit is identical to the actual receiver suggested before. The parameter N can be varied by changing the erase timing cycles on the storage drum tracks and the weights and step number logic in the stairstep multiplier. Input and

and channel noises are accounted for at N_1 and N_2 respectively.

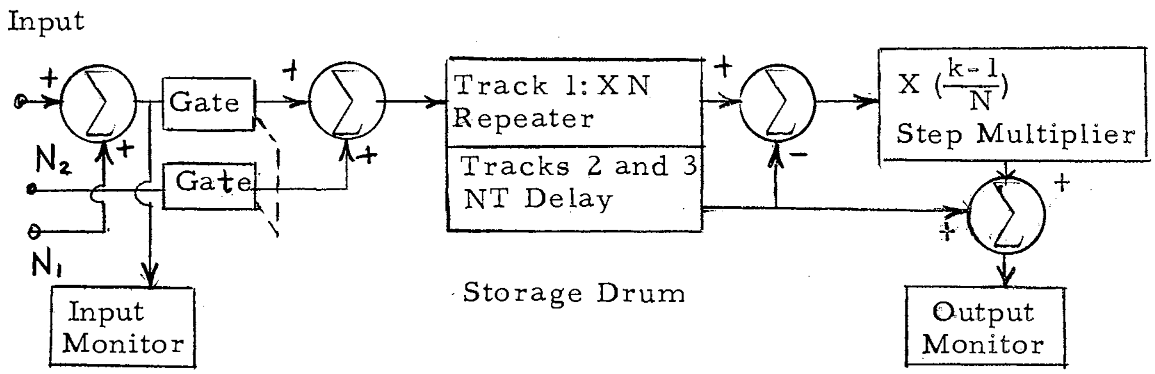


Figure 21. Linear brightness interpolation simulation

CHAPTER IX

SUMMARY AND CONCLUSIONS

The television bandwidth compression system proposed here basically reduces data by transmitting only one sample frame in N. When there is no movement in the pictures, no information is lost, and the original picture can be restored by repeating the sampled frames. With movement in the picture, some error occurs. Two methods of interpolation are suggested for reducing the visibility of this error. Both utilize the visual effect of exhange of spatial detail resolution for time-motion resolution. The two methods are as follows:

1. In addition to the sample frames, a filtered version of of the actual error signal is transmitted and provides a low-resolution fill-in for movement.

or

2. Only the sample frames are transmitted, with the receiver carrying out linear interpolation for element brightness.

Which of these two interpolations is visually superior can at present only be guessed. It was shown that the linerar interpolation method reduces channel noise power, while the transmitted error system does not. Further, the frame-to-frame noise correlation is less for linear interpolation. Thus, on a noise basis, the system of linear interpolation is superior. We guess that it may also be better visually, mostly because the linear interpolation is theoretically "cleaner," i.e., without the "ghost edge" problems of the transmitted error system.

Suggestions for Future Study

It is not easy to predict accurately the response of human vision in a comprex situation. Consequently, further study of the feasibility of the systems suggested here should include actual experimentation. With a magnetic storage drum ststem with at least three tracks, the two simulations decribed in the previous chapter could each be set up and the visual appearance of each type of interpolation for different degrees of compression directly observed.

Concluding Remarks

Statistical compression schemes operating within single frames have in general not been as successful as was once expected. The simple device of a scanning raster provides a priori agreement of the position of each element - leaving only brightness to be specified. On the other hand, as data compression is performed by discarding redundant or correlated brightness data, position information of some sort must also be specified. This important point has not always been given the attention it deserves.

Statistical compression is based upon exploiting the statistical properties of the source. It is at least as profitable to exploit the fidelity criteria of the (human) observer. The maximum resolution of human vision of spatial detail, contrast, and motion are not constant and independent, but vary interdependently. Several successful intra-frame compression ideas are, as we have seen, based directly upon the so-called exchange between detail resolution and contrast resolution. So far, little exploitation of visual effects on an inter-frame basis has been attempted.

To return for a moment to the statistical approach, it seems that statistical compression coding on a frame-to-frame basis holds the promise of more eventual success than does single element coding. While the position information for essential elements in a picture is comparable to the brightness information, an entire frame contains an enormous amount of information compared to the small amount needed to specify the position of the frame on an interval. At the present state of the storage art the storage requirements - while perhaps possible - are rather awkward to satisfy.

The system suggested here could thus be thought of as presently realizable alternative to frame coding on a statistical as well as visual basis.

BIBLIOGRAPHY

1. Beddoes, M. P. "A Two-Channel Method for Compressing the Bandwidth of Television Signals," Proc. I.E.E., Vol. 110, No. 2, February, 1963, pp. 369 - 374.

2. "Behind the Scenes," (Contains a description of a video disc recorder by Siemens and Halske in Germany.) Modern Photography, October, 1963, p. 134.

3. Bell, D. A., "Information Theory and its Engineering Applications," Sir Issac Pitman and Sons Ltd., London, 1953.

4. Below, F., et al, "The Subjective Disturbing Effect of Noise in Television Pictures," E. B. U. Review, Part A, No. 78, April, 1963, pp. 49 - 53.

5. Budrikis, Z. L., "Visual Threshold and the Visibility of Random Noise in T. V." Proc. I. E. E. Australia, December, 1961, pp. 751 - 759.

6. ________. "On the Channel Capacity of the Human Sense of Vision," Proc. I. R.E. E. Australia, April, 1964, pp. 228 - 234.

7. Carbrey, R. L., "Video Transmission Over Telephone Pairs by Pulse Code Modulation" Proc. I.R.E., Vol 48, No. 9, September, 1960, pp. 1546 - 1561.

8. Capon, J., "Bounds to the Entropy of Television Signals," Massachesetts Institute of Technology. Research Laboratory of Electronics, Technical Report No. 296, 1955, pp. 49 - 51.

9. Cherry, E. C. and Gouriet, G. G., "Some Possibilities for the Compression of Television Signals by Recoding," Proc. I. E. E. Paper No. 1401R, January, 1953, (100, Part III, p. 9).

10. Cherry, E. C., et al, "An Experimental Study of the Possible Bandwidth Compression of Visual Image Signals," Proc. I.E.E.E., November, 1963, pp. 1507 - 1517.

11. Deutsch, Sid, "Narrow-Band T. V. Uses Pseudo-Random Scan," Electronics, Vol. 35, No. 17, April 27, 1962, pp. 49 - 51.

12. Fink, D. G., Principles of Television Engineering, Second Edition, (McGraw-Hill, 1948).

13. Gabor, D. and Hill, P. C. J., "Television Band Compression by Contour Interpolation," Proc. I. E. E., May, 1961, pp. 303 - 315.

14. Goodall, W. M., "Television by Pulse Code Modulation," B. S. T. J., Volume 30, pp. 39-49, January, 1951.

15. Gouriet, G. G., "A Method of Measuring Television Picture Detail and its Applications, "Electronic Engineering, 1942, Vol. 24, p. 308.

16. ________. "Bandwidth Compression of a Television Signal," Proc. I.E.E., Vol. 104, pp. 265-272, May, 1957.

17. Graham, R. E., "Subjective Experiments in Visual Communication," I.R.E. National Convention Record, March, 1958, Part A, pp. 100 - 105.

18. Gregory, R. L., "An Experimental Treatment of Vision as an Information Source and Noisy Channel," Third Symposium on Information Theory, London, 1956, pp. 287 - 299.

19. Harrison, C. W., "Experiments With Linear Prediction in Television," B. S. T. J., July, 1952, 765 - 783.

20. Jesty, L. C., "Television as a Communication Problem," Proc. I. E. E., Vol. 99, Part IIIA, 20, 1952, pp. 761 - 770.

21. Julesz, B., "A Method of Coding Based on Edge Detection," B.S.T.J., July, 1959, pp. 1001 - 1020.

22. Kell, R. D., British Patent No. 341811, 1929.

23. Kretzmer, E. R., "Statistics of Television Signals," B.S.T.J., July, 1952, pp. 751 - 763.

24. ________. "Reduced-Alphabet Representations of Television Signals," I.R.E. Convention Record 1956, Part IV, pp. 140-147.

25. Kubba, M. H., "Automatic Picture Detail Detection in the Presence of Random Noise," Proc. I.E.E.E., November, 1963, pp. 1518 - 1523.

26. Ludvigh, E. and Miller, J. W., "Study of Visual Acuity During the Pursuit of Moving Test Objects" Parts I and II, Journal of the Optical Society of America, Vol. 48, No. 11, November, 1958, pp. 799 - 808.

27. "Magnetic Drum System Combines Flexibility, High Storage Capacity," Electronic Design, ("Product Feature" : Description of a Drum Storage System by Transco Products, Venice, California.)

28. Newell, G. F. and Geddes, W. K. E., "Tests of Three Systems of Bandwidth Compression of Television Signals, " Proc. I.E.E., Paper No. 3613E, July, 1961, pp. 311 - 324.

29. ________. "Visibility of Small Luminance Perturbations in Television Displays," Proc. I. E. E., Vol. 110, No. 11, November, 1963.

30. Oliver, B. M., "Efficient Coding." B. S. T. J. Vol 31, July, 1952, pp. 724 - 750.

31. Oliver, B. M., Pierce, J. R., and Shannon, E. E., "Philosophy of PCM," Proc. I.R.E., Vol. 36, pp. 1324 - 1331, November, 1948.

32. Pierce, J. R. and Karlin, J. E., "Reading Rates and the Information Rate of a Human Channel, B. S. T. J., March, 1957.

33. Powers, K. H. and Staras, H., "Some Relations Between Television Picture Redundancy and Bandwidth Requirements," Transactions of the A.I.E.E., September, 1957, pp. 492 - 496.

34. Potter, J. B. "One the Use of the Vidicon Camera Tube as a Video Storage Device. Proc. I.R.E. Australia, December, 1963, pp. 855 - 865.

35. Prasada, B., "Some Possibilities of Picute Signal Bandwidth Compression," I.E.E.E. Transactions on Communication Systems, September, 1963, pp. 315 - 328.

36. Rittermann, M. B., "Application of Autocorrelation Theory to the Video Signal of Television," Sylvania Technologist, 1952, p. 70.

37. Roberts, L. G., "Picture Coding Using Pseudo-Random Noise," I. R. E. Transactions on Information Theory, February, 1962, pp. 145 - 154.

38. Schreiber, W. F., "The Measurement of Third-Order Probability Distributions, "I.R.E. Transactions on Information Theory, Vol. IT. -2, No. 3, September, 1956, p. 94.

39. Schreiber, W. F. and Knapp, C. F. "T. V. Bandwidth Reduction by Digital Coding," I.R.E. Convention Record, 1958, Vol. 6, Part V, pp. 88 - 99.

40. Seyler, A. J., "Visual Communciation and the Psycho-Physics of Vision," Proc. I.R.E. Australia, May, 1962, pp. 291 - 304.

41. ________. "The Coding of Visual Signals to Reduce Channel-Capacity Requirements," I. E. E. Monograph, No. 535 E., July 1962, pp. 676 - 684.

42. ________. "An Experimental Frame Difference Signal Gneerator for the Analysis of Television Signals," Proc. I. R. E. Australia, November, 1963, pp. 797 - 807.

43. Seyler, A. J. and Budrikis, Z. L., "Measurements of Temporal Adaptation to Spatial Detail Vision," Nature, October 17, 1959, pp. 1215 - 1217.

44. Shannon, C. E., "A Mathematical Theory of Communication," B.S.T.J., 1948, Vol. 27, pp. 379 and 623.

45. Syiklai, G. C., "Some Studies in the Speed of Visual Perception," I. R. E. Transactions on Information Theory, September, 1956, pp. 125 - 128.

46. Teer, K., "Investigations into Redundancy and Possible Bandwidth Compression in Television Pictures," Philips Research Reports, 14, pp. 501 - 556, 1959 and 15, pp. 30 - 96, 1960.

47. Wessels, J. H., "A Magnetic Wheel Store for Recording Television Signals," Philips Technical Review, Vol. 22, November, 1960.

48. Wichmann, T., "Optimum Coding of Pictorial Data," I.E.E.E. National Communications Symposium, October 7 - 9, 1963, pp. 18-27.

49. Youngblood, W. A., "Picture Processing," M. I. T. Research Laboratory of Electronics, Quarterly Progress Report, January, 1958, p. 134.

APPENDIX

MAGNETIC STORAGE DRUM CONNECTIONS

Assumptions: It is assumed that tracks are available on the drum or disc as needed. The period of rotation is one frame period T. Angles around the tracks (between heads etc.) are measured as a fraction of 360°.

An FM carrier system is likely to be used, rather than base-band recording, to combat noise, lower head octave ratio, etc. It is not explicitly shown.

Symbols: R = Recording head

P = Playback head

E = Erase head.

XN Sample Frame Repeater

The sample frame repeater shown in Figure 22 involves an over-all delay of A_1 T seconds. The input is a sequence of sample frames. That is, a sample frame is recorded in one revoltuion, and the input is zero until the next sample frame. The erase signal is turned on $(1 - A_2)$ T sec. before the sample frames, and off one period later.

A compensating delay (of A_1T) can be obtained with the same arrangement by simply leaving the erase signal on continuously.

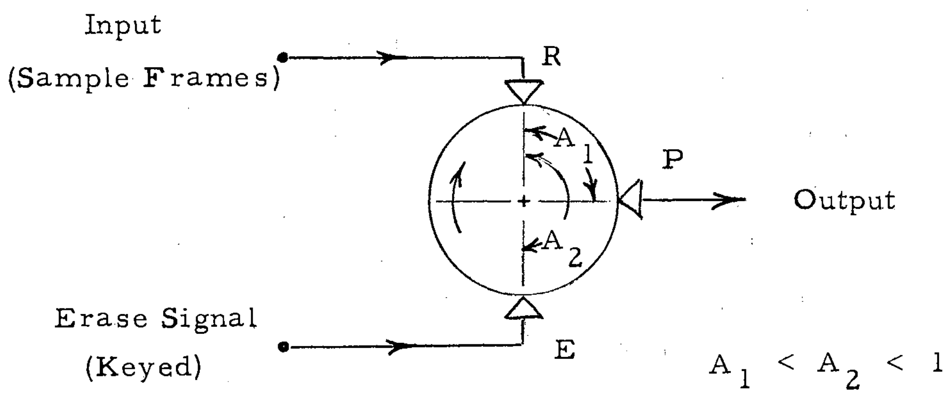


Figure 22. XN sample frame repeater

N-Frame Delay With Redundant Input

The N-frame delay shown in Figure 23 operates as follows. The input is a sequence of repeated sample frames (e.g. the output of the device in Figure 22). The first stage gives a delay of AT (say 3/4 T). The last frame of the output is gated at G and fed to the second track. The second track is a XN repeater (like Figure 22) with an overall delay of BT (say 1/4 T), such that A+B=1. Thus, the output of the second track is the repeated sequence of sample frames - delayed NT seconds.

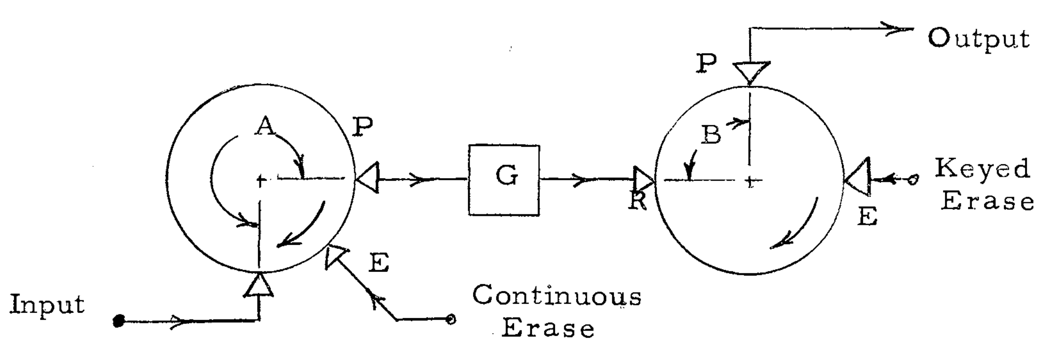


Figure 23. N-frame delay (repeated input)

(N-1)-Frame Delay

The system of Figure 24 is to provide a total delay of (N-1) T seconds to a continuous arbitrary video signal. N tracks are required as shown, subject to the conditions

$$A < \frac{1}{N-1} \quad \text{and} \quad B = (N-1)A,$$

since

$$\text{Delay} = (N-1)(1-A)\,T + BT = (N-1)\,T,$$

and

$$B < 1.$$

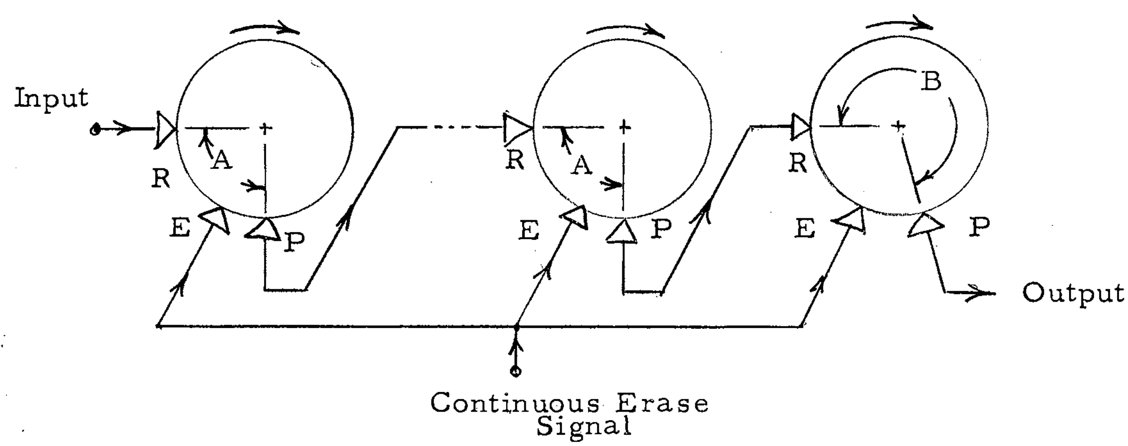


Figure 24. (N-1) delay

VITA

Charles Gene Hilborn, Jr.

Candidate for the Degree of

Master of Science

Thesis: TELEVISION BANDWIDTH COMPRESSION

Major Field: Electrical Engineering

Biographical:

Personal Data: Born in Shreveport, Louisiana, November 29, 1941, the son of Mr. and Mrs. C. G. Hilborn.

Education: Graduated from C. E. Byrd High School, Shreveport, Louisiana in 1959; graduated from Louisiana Polytechnic Institute, Ruston, Louisiana, in August, 1963, with the degree of Bachelor of Science in Electrical Engineering; completed requirements for the degree of Master of Science, with a major in Electrical Engineering, at Oklahoma State University in May, 1965.

Professional Experience: Was employed for the summer of 1962 by the Army Missile Command, at Redstone Arsenal, Alabama, as an Engineering Aid; have been a Graduate Teaching Assistant at Oklahoma State University since September, 1963; student member of the Institute of Electrical and Electronic Engineers; member of Tau Beta Pi and Eta Kappa Nu honorary societies.