THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 5, May-June, 1971
Printed in U.S.A.

Some Computer Experiments in Picture Processing for Bandwidth Reduction

By H. J. LANDAU and D. SLEPIAN

(Manuscript received January 4, 1971)

Some computer experiments in processing still pictures for bandwidth reduction are described. In the scheme studied, a picture is partitioned into subpictures each of which is encoded separately. A subpicture is expressed as a linear combination of a finite set of specially chosen basis subpictures. Quantized versions of the coefficients of this expansion are transmitted as binary digits. Using this procedure, we were able to obtain pictures of good quality using approximately 2 bits per picture-element; we were unable to do so at lower bit-rates.

Some general comments on the encoding of pictures are included.

I. INTRODUCTION

This paper describes some computer experiments in picture processing carried out by us during the winter of 1969 and the spring of 1970. Our goal was to explore a particular method for the efficient encoding of typical *Picturephone*® scenes into binary digits. The experiments involved still pictures only. We first describe the experiments and their results, then follow with some general comments on the encoding of pictures. These comments are intended to explain our motivation for the particular investigation undertaken.

II. THE EXPERIMENTS

By means of the TAPEX unit at Murray Hill,[1,2,3] a photograph can be represented in digital form suitable for handling by the GE-635 computer. Specifically, the picture is scanned from top to bottom along n_1 horizontal lines, the light intensity being sampled n_2 times along each line; every sample is then quantized to the nearest one of 2^k equally spaced amplitude levels, and recorded on a digital tape as a single k-bit integer. Conversely, given the digital tape, each k-bit integer is replaced by its corresponding amplitude value, which is then regarded as a sample, taken at the Nyquist rate, of a bandlimited waveform. The reconstructed waveform controls the beam intensity of successive lines traced by a scanning cathode-ray oscilloscope. The oscilloscope face is photographed.

In all our experiments, the values used for the above quantities were $n_1 = n_2 = 256$, and $k = 10$. Figure 1a is an original photograph; Fig. 1b is the result of converting the picture into binary digits on tape and reconstructing via TAPEX. Comparison shows that, with the parameters as chosen, the digital representation is of customary television quality. It has, of course, the inevitable raster lines.*

In processing a picture, we first converted each k-tuple of binary digits into the corresponding integer, and subtracted 2^{k-1}. The resulting integers, lying in the range $(-2^{k-1}, 2^{k-1} - 1)$, are called *picture elements*, and we denote by X_{ij} the picture element obtained from the jth sample of the ith line of the picture. For computer processing, we regard a picture as an $n_1 \times n_2$ matrix of picture-elements. In our experiments, we further partitioned the $n_1 \times n_2$-element picture by a square grid (as in Fig. 2) into $n_1 \times n_2/m^2$ square subpictures, each having m picture-elements on a side. These subpictures were encoded independently, one at a time, by the scheme described below.

We view the $M = m^2$ picture-elements of a subpicture, when read out row by row from left to right, as the components of an M-dimensional vector $\mathbf{Y}$, which represents the subpicture. For example,

$$\mathbf{Y} = \begin{bmatrix} X_{11} \\ X_{12} \\ \vdots \\ X_{m1} \\ \vdots \\ X_{mm} \end{bmatrix}$$

is the vector representing the top-left subpicture of Fig. 2. To describe such vectors, it is natural to introduce a basis. We therefore choose M orthonormal M-dimensional basis vectors $\mathbf{b}_1, \cdots, \mathbf{b}_M$. These remain fixed, and determine the particular encoding scheme under discussion.

* (*Added in proof.*) The half-tone method of picture reproduction used in printing this article of necessity obscures some of the detail visible in the original TAPEX photographs. Copies of these photographs will be sent on request.

Fig. 1—(a) Original photograph; (b) Reconstruction of (a) via TAPEX, using 10 bits per picture-element.

We may now expand **Y** in terms of the basis vectors, to obtain

$$\mathbf{Y} = \sum_{1}^{M} c_j \mathbf{b}_j \tag{1}$$

where, by orthonormality of the $\mathbf{b}_j$,

$$c_j = \mathbf{b}_j \cdot \mathbf{Y}, \qquad j = 1, \cdots, M. \tag{2}$$

We then quantize c_j into one of r_j different values, denoting by $\hat{c}_j$ the quantized version of c_j. To transmit these quantized coefficients of a subpicture in the simplest possible way, i.e., by encoding them independently without exploiting the statistical distribution of their values, requires $r = \sum_1^M [\log_2 r_j]$ binary digits. We take the number

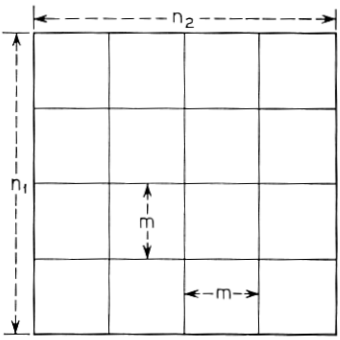

Fig. 2—Partition of pictures into subpictures.

$R = r/m^2$, which is the number of bits used per picture-element, as a measure of the efficiency (or bandwidth) of the encoding scheme.

To reconstruct a picture, we suppose the quantized coefficients are known, and obtain a reconstructed subpicture code vector $\hat{\mathbf{Y}}$ from the recipe

$$\hat{\mathbf{Y}} = \sum \hat{c}_i \mathbf{b}_i \ .$$

The components of $\hat{\mathbf{Y}}$ are quantized to the nearest integer value in the range $[-2^{k-1}, 2^{k-1} - 1]$. These are the picture-elements $\hat{X}_{ij}$ of the reconstituted picture. A photograph is obtained from these values using the TAPEX unit in the manner already described.

Fig. 3—(a) 10 bits per picture-element; (b) Reconstruction of (a) by means of the Hadamard basis, using 2 bits per picture-element; (c) Reconstruction of (a) by means of differential PCM, using 3 bits per picture-element.

Our experiment consisted of choosing various basis vectors and various quantization rules for the expansion coefficients c_j. We also experimented with making nonlinear transformations on the picture-elements before and after bandwidth compression processing. None of the various types of companding we tried yielded better results than were obtained without companding. In most of our work, subpictures of $M = 16$ picture-elements ($m = 4$) were used; a few experiments were run with $m = 8$.

We were able to obtain pictures of good quality with a rate $R = 2$ bits per picture-element, but were unable to do so at lower rates. Figure 3a repeats the 10 bit per picture-element photograph of Fig. 1b. Figure 3b shows a reconstructed picture with $R = 2$ bits per picture-element obtained with a scheme using $m = 4$ and the Hadamard basis described below. We also simulated on the computer the differential PCM scheme employed in *Picturephone* coding which uses 3 bits per picture-element.[4] Figure 3c shows the result of this simulation; it compares favorably with Fig. 3b. Two different subjects are treated analogously in Figs. 4 and 5.

Although the subpicture encoding achieves a one-third decrease in rate, the differential PCM scheme is far easier to instrument. From our experience, it seems unlikely that good pictures can be obtained with the subpicture scheme at rates much less than 2 bits per picture-element.

III. COMMENTS ON COMPRESSION

To avoid needless complications, in all that follows we shall think of a picture in discrete terms, i.e., as a finite collection of picture-elements, each of which can assume finitely many different values.

How many bits must one use to transmit the picture of Fig. 1b? The answer is, of course, zero. It is a single picture. The question is not an interesting one. More pertinently, we can ask how many bits per picture are required on the average to transmit long strings of pictures drawn from a given ensemble of pictures. Since a picture source can be regarded as producing sequences of picture elements, each of which can assume one of K different values, evidently we can transmit all possible pictures perfectly by using $[\log K]$ bits per picture-element. For any reduction of the bit-rate below this value, we must capitalize on one or both of these facts:

(*i*) not all pictures are produced with equal probability by the source, nor are they produced independently;

(*ii*) the observer does not require all pictures to be reproduced exactly.

Fig. 4—(a) 10 bits per picture-element; (b) Reconstruction of (a) by means of the Hadamard basis, using 2 bits per picture-element; (c) Reconstruction of (a) by means of differential PCM, using 3 bits per picture-element.

The question of how to take advantage of such considerations has been much studied in information theory, and methods are known in principle for computing the answer. A calculation of the *entropy* of the picture ensemble describes how far it is possible to reduce the bit-rate, and still maintain perfect reconstruction, by exploiting source redundancies; this minimum rate is determined solely by the statistics of the ensemble, and has nothing to do with the nature of pictures, vision, or the observer.

Determination of the entropy of a picture source does not solve the problem of real interest to workers in picture transmission; for pictures, as they are usually presented by a source, have more detail and resolution than the observer can utilize. Thus pictures that are counted as different in the source ensemble may be indistinguishable to the viewer who wishes to reconstruct them only to some set limit of accuracy, i.e., to achieve some "level of fidelity". The minimum bit-rate, given as a function of both the particular fidelity criterion adopted and the source statistics, may also be computed, and is called the *rate distortion function.* As with entropy, this minimum rate is achievable only in the limit of more and more complicated encoding processes.

Conceptually, rate distortion theory formulates carefully and answers completely the foremost question in the TV coder's mind: "How many bits do I need?" In actuality, it doesn't do very much for him

Fig. 5—(a) 10 bits per picture-element; (b) Reconstruction of (a) by means of the Hadamard basis, using 2 bits per picture-element; (c) Reconstruction of (a) by means of differential PCM, using 3 bits per picture-element.

at all. To see why this is so, we must look just a bit closer at the formalism of rate distortion theory.

The general theory presupposes a source that produces infinite strings of symbols, each symbol being drawn from a K-letter alphabet. The "values" of the letters play no role in the theory, so for convenience we suppose that they are the integers $1, 2, \cdots, K$. Denote a typical string produced by the source by $\cdots X_{-1}, X_0, X_1, X_2 \cdots$, where each X is one of the integers from 1 to K. A measure is placed upon the set of such infinite strings in such a way that we can regard the Xs as random variables and meaningfully ask and answer such questions as "What is the probability that $X_0 = 3$, $X_2 = 1$, and $X_3 = 5$?" There are many technicalities involved in specifying this measure, but they need not concern us here. We are also given a numerical-valued distortion function $\delta(j, k) \geqq 0$ which gives the distortion when a transmitted letter j is reconstructed as letter k.

Let us now consider transmitting the strings produced by the source by encoding them in the following manner. We break the source strings up into blocks of n successive symbols. Since each block is composed of n source symbols and each symbol can be one of K different integers, there are $B = K^n$ different blocks possible. We suppose that a dictionary is provided, which lists for each one of these B blocks a special block called its representative block. As successive strings of n source letters are produced by the source, each is looked up in the dictionary and encoded into its representative block. If the letters of a block are $x_1, x_2, \cdots, x_n$ and the letters of the corresponding representative block taken from the dictionary are $y_1, y_2, \cdots, y_n$, we take the quantity $D = \sum_1^n \delta(x_i, y_i)$ as the distortion per block. We take d = average D/n as the level of distortion achieved with the given code book, where the average is over all source strings.

Let us now fix the number of representative blocks in the dictionary at 2^L. Some code books translating the B blocks into 2^L representative blocks will yield smaller values for the distortion d than will others. We denote by $d(L, n)$ the smallest distortion obtainable by any such code book. Note now that since there are only 2^L representative blocks in these code books, we could use L binary digits to transmit each representative block name to a destination. We would achieve distortion $d(L, n)$ and be transmitting at a rate

$$R = (L/n) \text{ bits/(source symbol)}.$$

Now fix R and write $\bar{d}(R) = \lim_{n \to \infty} d(nR, n)$. This function gives the smallest distortion obtainable for a fixed binary rate R that can be had in the limit of arbitrarily large code books. The inverse function $R(\bar{d})$ which gives the smallest binary rate per source letter that will yield a given distortion $\bar{d}$ is called the *rate distortion function*. Information theory shows how $R(\bar{d})$ can be calculated in principle from the symbol distortion function $\delta(j, k)$ and the measure assigned to the source. We do not display these complex formulas here.

How can we apply this to picture transmission? There are two obvious different methods of identifying the source symbol X with a quantity of interest in picture coding. The method most satisfying conceptually is to identify the random variable X with an entire picture. This is possible since there are only finitely many different pictures, due to our assumption that a picture is composed of $n_1 \times n_2$ picture elements, each taking one of 2^k values. We number the possible pictures and take the picture numbers as the values of X. The distortion function $\delta(j, k)$, which we must now describe to apply the theory, measures our dissatisfaction at having picture j reproduced as picture k. Conceptually such a measure exists, but we know little about it. In our experiments, we would have to prescribe it for $(2^{10\times256\times256})^2$ pairs of values of j and k.

To compute a value for the rate distortion function, we require in addition a measure on the source symbols: at a minimum, this involves assigning a probability distribution, bearing some relation to what will be observed in practical transmission, to the $2^{10\times256\times256} = 10^{197,283}$ different possible pictures. This task seems quite beyond us now. For to obtain a histogram empirically is out of the question: at 30 frames per second, one sees only 10^9 frames per year, and if the different possible pictures were run off in sequence at this rate, it would take $10^{197,274}$ years to view them all. On the other hand, to specify the distribution theoretically requires more understanding of the situation than we now have.

Indeed, being able to describe a reasonable distribution for the possible pictures goes a long way towards solving the problem of efficient coding. We suspect that a reasonably good description would assign probability $1/N$ to each of N of the pictures, and zero to the rest, with N small indeed compared to $10^{197,283}$. If we could describe this set well, we could encode using log N bits/picture. But which are the "likely" pictures? For *Picturephone* service or commercial broadcast television, intuition suggests that chaotic pictures, in which adjacent picture-elements jump about between extreme values, would be classified as unlikely. Likely pictures are, roughly speaking, made up of regions of nearly constant brightness. The brightness might change considerably from one region to the next, but there cannot be too many small regions, or we are back to unlikely chaos, nor can the boundaries of the region be too wild or fast-turning. However, the enormous number of possibilities involved prevents an accurate description.

A more tractable application of the general theory comes from associating the source symbol X with a picture-element. Now, however, the distortion function $\delta(j, k)$ measures our displeasure at having the jth level of brightness for a picture-element reproduced as the kth level of brightness. This is an excessively local measure of picture fidelity and is probably quite remote from the criteria used by human observers.

In summary, rate distortion theory tells us that to encode efficiently we must pay attention to the more likely pictures (or sequences of pictures), and that we must replace these in groups by representative values which yield an acceptable distortion. The theory tells us how to calculate the minimum bit-rate needed to achieve a given level of fidelity, but to carry out the calculation we need to know the distortion function $\delta(j, k)$ and the measure that gives a statistical description of the source. In picture coding we have at present very meager knowledge concerning these quantities. Any new understanding of either will undoubtedly lead to improved practical coding schemes. It will take a great deal of understanding, however, to know these quantities well enough to allow a calculation of a rate distortion function in which one can have much confidence.

IV. MOTIVATION FOR THE EXPERIMENTS

As we have argued, we lack the information required to bring the full force of rate distortion theory to bear on picture coding. Nevertheless, it was the general approach of rate distortion theory that led to the encoding scheme of our experiments. We wanted an encoding procedure that also would derive from considerations of likelihood and fidelity, but that would be manageable in practice. Accordingly, we began by focusing on subsections of the picture.

Although we cannot characterize adequately those entire pictures that are likely, perhaps we can do so for subpictures. How large must a section of a TV picture be before we can describe it as a likely subpicture or an unlikely subpicture? If we look at a single picture-element, every value is "likely." If we look at two adjacent picture-elements, again we must say that any pair of values is "likely." If we consider square subpictures of $m \times m$ picture-elements, most observers feel that for $m = 4$ they can already classify some subpictures as likely parts of the *Picturephone* or TV ensemble and others as less likely. The 4×4 checkerboard pattern, where adjacent picture elements oscillate between extreme values, seems unlikely: the uniform 4×4 subpicture seems highly likely.

We decided then to break a picture into $m \times m$ subpictures and to encode each subpicture independently. If m is large enough, not much compression potential will be lost by neglecting the correlation between subpictures, for the chaos of structure that we intuitively feel to be unlikely in the TV ensemble is of a within-subpicture scale. The factor driving us to choose a small value of m is the need for a reasonable number of pictures to distinguish among on a probabilistic basis.

Even with $m = 4$ and $k = 10$, there are $2^{10 \times 16} = 10^{48}$ different subpictures, so that it is out of the question to use a Huffman-Fano code, or other dictionary-like code, to take advantage of the unequal probabilities of the various subpictures. We seek some other scheme. A natural idea here is to represent the subpicture in terms of some coordinates that can be treated independently and that are related to the probability measure on the subpictures.

We begin by interpreting the subpicture as a vector, in the manner described in Section II. Suppose that an M-dimensional subpicture vector $\mathbf{Y}$ is to be expanded on M linearly independent basis vectors $\mathbf{b}_i$ as in equation (1), but that only $J < M$ of the cs will be used (exactly) to reconstruct an approximation $\hat{\mathbf{Y}}$ to $\mathbf{Y}$. Thus

$$\mathbf{Y} = \sum_1^M c_i \mathbf{b}_i \,, \qquad \hat{\mathbf{Y}} = \sum_1^J c_{\alpha_i} \mathbf{b}_{\alpha_i} \,.$$

where $\alpha_1, \alpha_2, \ldots, \alpha_J$ are distinct integers from the list $1, 2, \ldots, M$. What basis vectors should be used, and which J coefficients retained, in order to minimize the mean squared error between $\mathbf{Y}$ and $\hat{\mathbf{Y}}$? The answer to this problem is well known. Let $\mathbf{Y} = (y_1, y_2, \cdots, y_M)$ and denote the covariances of the components of $\mathbf{Y}$ by $\rho_{ij} = Ey_iy_j$. Let ρ be the $M \times M$ matrix with elements ρ_{ij}. The basis vectors that solve the above problem are the eigenvectors of ρ having the J largest eigenvalues. Basis vectors chosen in this way are known as a Karhunen-Loève basis. The fidelity criterion implicit here is that of mean-square error—one not very adequate when applied to pictures.

We began our experiments with finding the Karhunen-Loève basis for the picture of Fig. 1, by determining empirically the 16×16 covariance matrix for the 4×4 subpictures. We discovered, as expected, that the ρ_{ij} were all extremely close to 1, expressing the high likelihood of uniform brightness over so small a square. When each $\rho_{ij} = 1$, the eigenvalue problem is degenerate: the first eigenvector has all components equal and an eigenvalue of 1, while the remaining eigenvectors are indeterminate and correspond to an eigenvalue of 0. Thus we felt no great confidence in our determination of the Karhunen-Loève vectors, believing that it was probably unstable, and turned instead to the following intuitive justification for introducing another basis, which we call the Hadamard basis.

Intuitively, a subpicture with constant brightness for all picture-elements is a very likely subpicture. Part (1) of Figure 6 depicts a

4×4 array of picture elements all having value $+ 1/4$. Another likely subpicture has a vertical edge running down its middle. Part (2) of Figure 6 depicts such a case, where the picture elements on the left have value $+ 1/4$ and those on the right have value $- 1/4$. If now we form a basis vector, $\mathbf{b}_1$, from Fig. 6(1) and $\mathbf{b}_2$ from Fig. 6(2), we see that linear combinations $\mathbf{Y} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2$ give all possible subpictures having a center vertical transition between two regions of uniform brightness.

Continuing this train of thought, we are led to seek 16 linearly independent subpictures of decreasing likelihood that will serve as a basis on which to expand an arbitrary subpicture. Such a basis, chosen so that the vectors are orthonormal, is shown in Fig. 6. If each subpicture $\mathbf{Y}$ of a large picture is expanded on this basis, so that

$$\mathbf{Y} = \sum_1^{16} c_i \mathbf{b}_i ,$$

we would expect frequently to find the higher coefficients, say c_{10}, c_{11}, etc., to have values near zero. The coefficient c_1, which gives the average brightness of the pictures, would be expected to have a large variance—higher coefficients, a much smaller variance. Table I lists the ratios $\xi_i = \sigma_i^2/\sigma_1^2$, where σ_i^2 is the variance of c_i, as determined empirically from all the subpictures of Fig. 1b. The results agree remarkably well with intuition. That c_1, which has more than ten times the variance of any other coefficient, does indeed contain a great deal of the essence of the original picture is seen from Fig. 7b, which shows the picture resulting from the reconstruction $\hat{\mathbf{Y}} = c_1\mathbf{b}_1$, next to the original photo-

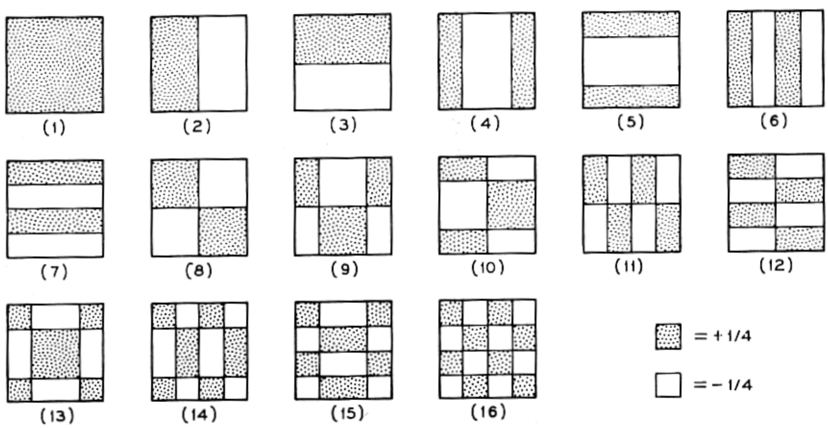

Fig. 6—The Hadamard basis.

Table I—Coefficient Variances

i	ξ_i	i	ξ_i
1	1.00	9	0.024
2	0.098	10	0.024
3	0.087	11	0.020
4	0.035	12	0.022
5	0.038	13	0.019
6	0.051	14	0.015
7	0.048	15	0.016
8	0.034	16	0.014

graph 7a. Figure 7b allows one to see clearly the size of the subpictures used in the experiments.

Our choice of the Hadamard basis is thus dictated by plausible guesses about the probabilities of subpictures. Furthermore, it is not inconsistent with the Kahunen-Loève procedure, since the correlation matrix is so nearly singular. Finally, it has an important practical advantage: since its components each have value $\pm$ 1/4, the computation of the coefficients can be carried out by simple switching. In our experiments, we found the results of processing with the Karhunen-Loève basis to be no better than those obtained with the Hadamard basis, and so, for reasons of simplicity, judged the latter to be superior.

Since our object is to reduce the bit-rate, we must adopt some scheme of quantization for the coefficients. This will lead to an approximate reconstruction of the subpicture, and considerations of fidelity must guide us in our choice of rules. One such quantization scheme—keeping some of the coefficients exactly and dropping the remainder altogether—gives rise to the Karhunen-Loève problem. We adopted a different procedure, based on quantizing successive coefficients more

Fig. 7—(a) Original photograph; (b) Reconstruction of (a) by means of c_1 only.

Table II—Quantization of Coefficients

i	r_i	i	r_i
1	64	9	4
2	16	10	4
3	16	11	0
4	8	12	0
5	8	13	0
6	8	14	0
7	8	15	0
8	4	16	0

and more coarsely. Two arguments led us to this. Firstly, since the lower coefficients have more variability, reproducing these more accurately helps reduce the mean-square error for the more probable pictures. Secondly, the higher coefficients tend to be large mainly when the subpicture has a very "busy" or chaotic nature; we guessed the detail of that chaos to be less important to the viewer than the existence of chaos. Thus the fidelity criterion behind our encoding contains an element of the characteristics of observers, in addition to considerations of mean-square error.

Much experimenting bore out the general truth of these suppositions. Table II gives the number of quanta, r_i, used for c_i in Figs. 3b, 4b, and 5b. The quantization of a given c was carried out by dividing its range into disjoint intervals whose endpoints are called *cut points* and by associating with each interval a *representative value*. The quantized value of c is the representative value associated with the interval in which c lies. Table III lists the cut points and representative values used to obtain Figs. 3b, 4b, and 5b.

We carried out over 100 experiments in which the r_i, the cut points, and representative values were varied over considerable ranges; details are available on request. Although we ultimately settled on the configuration described in Tables II and III, the number of possibilities to be explored is so large that we have no great confidence that we have found the best values for the parameters. On the other hand, based on our experience we would judge it unlikely that significant improvements can be made with this scheme by further changes of parameter values.

V. PICTURE REPRODUCTION AND QUALITY JUDGMENT*

The development of ordinary photographic film, as well as the characteristics of analog devices such as scanners and picture tubes,

* (*Added in proof.*) The comments of this section refer to the original TAPEX photographs, rather than to their reproductions in this article. See footnote on p. 1526.

Table III—Cut Points and Representative Values used for Quantizing the Coefficients $c_1, \cdots, c_{16}$.

Possible values for the c_k lie in the range ± 2048.0, and are integer multiples of 0.25.

c_1: The range between −1940 and +1900 was divided into 64 intervals of length 60, the midpoint of each serving as representative point for the interval. The first and last representative points also represented any values of c_1 outside this range.

The cut points and representative values for $c_2 - c_{10}$ were generated from curves of the form $y = k\sqrt{x}$, as follows:

c_2, c_3:	Cut points	± 0		8.6		34.4		77.4		137.5
	Repr. values	±	2.1		19.3		53.7		105	
	Cut points	±	215		309		421			
	Repr. values	± 174		260		363		483.		
c_4, c_5, c_6, c_7:	Cut points	± 0		15.6		62.4		140.4		
	Repr. values	±	3.9		35.2		97.6		191.	
c_8, c_9, c_{10}:	Cut points	± 0		62.5						
	Repr. values	±	15.6		141.					

$c_{11}, \cdots, c_{16}$: dropped

are sensitive to many parameters and can vary noticeably over time. The result is that in processing and reproducing photographs it is extremely difficult to maintain rigid control of contrast and average gray level. Yet these two quantities strongly influence the viewer's judgment of the quality of a picture.

In our subpicture encoding scheme, information about overall contrast and gray level is contained almost entirely in the values of c_1. We are convinced by experiments performed that the quantization of this coefficient, as described by Tables II and III, is sufficiently fine to render these characteristics faithfully. We therefore believe that what variation in contrast exists on Figs. 1, 3, 4, 5, and 7 is attributable to vagaries of the reproduction process and not to failures of the encoding, which are evidenced by inaccuracies in edges and texture. Accordingly, the reader should attempt to subtract out the differences in contrast among the photographs and should judge the quality of our scheme by examination of detail in Figs. 3b, 4b, and 5b.

VI. RELATED WORK

At about the same time that the present experiments were carried out, rather similar investigations were independently conducted elsewhere by other workers.[5,6,7] While related to our work, these studies differ from it somewhat in detail of execution and very much in theoretical approach.

VII. ACKNOWLEDGMENT

We are greatly indebted to C. A. Sjursen for his continual help in scanning and reproducing pictures on the TAPEX unit.

REFERENCES

1. Anderson, W. A., Swartzwelder, J. D., Weller, D. R., "System and Programming Aspects of TAPEX," unpublished work, March 1, 1966.
2. Camlet, J. V., "Remote High-Speed Data Station," unpublished work, January 16, 1967.
3. Giordano, P. P., "Variable Speed Scanner," unpublished work, March 7, 1962.
4. Limb, J. O., and Mounts, F. W., "Digital Differential Quantizer for Television," B.S.T.J., *48,* No. 7, part 3 (September 1969), pp. 2583–2599.
5. Tasto, M., and Wintz, P. A., "Picture Bandwidth Compression by Adaptive Block Quantization," Report TR-EE 70-14, School of Electrical Engineering, Purdue University, Lafayette, Indiana, July 1970.
6. Huang, T. S., and Woods, J. W., "Picture Bandwidth Compression by Block Quantization," presented at 1969 IEEE International Symposium on Information Theory, Ellenville, N. Y.
7. Pratt, W. W., and Andrews, H. C., "Transform Processing and Coding of Images," USCEE Report 341, University of Southern California, Los Angeles, Calif., March 1969.