PREDICTIVE ENCODING OF DIGITIZED TV PICTURES

Thesis by

D. K. Sharma

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1975

(Submitted May 19, 1975)

ACKNOWLEDGMENTS

It gives me great pleasure to express my deep gratitude to my advisor, Dr. John R. Pierce, for his guidance, encouragement and inspiration throughout this work. His invaluable suggestions and comments helped initiate and develop the work presented here.

My sincere thanks are due to Mr. C. C. Cutler for providing me with the unique opportunity of working in the extremely friendly and open environment of the picture coding group at Bell Telephone Laboratories, New Jersey. Without using the equipment there, it would not have been possible to see this work through completion. I express my very special thanks to Dr. J. O. Limb for his guidance and friendship during my stay at Bell Telephone Labs. There I had many fruitful discussions with J. Candy, D. J. Connor, P. L. Gordon, B. G. Haskell, A. N. Netravali, B. Prasada and C. B. Rubinstein. My sincere thanks to them all. Gentle help of J. D. Bayer and R. C. Brainard in the computer work is gratefully acknowledged. Thanks are due to E. G. Bowen, M. E. Lukacs and K. A. Walsh for their help in getting the pictures made.

I thank Ruth Stratton for her wonderful work at the typewriter.

Many special thanks are due to Xerox Corporation and Bell Telephone Laboratories, Inc. for their generous financial support during the course of this work.

ABSTRACT

In this report, we consider the problem of efficient intra-frame encoding of digitized TV pictures using Differential Pulse Code Modulation (DPCM) type encoders. Our goal is to reduce the average number of bits used to encode a pixel while subjective quality of the picture still remains acceptable.

To this end, we propose some new predictors and compare them with the existing ones using the error criteria "essential maximum" and "variance" of the prediction errors. The two criteria have also been compared with respect to the subjective quality of the final encoded pictures. We propose to design the quantizers using "mean n-th power" distortion criterion. Effects of "n" on bit rate and subjective quality of the encoded picture have been studied as it varies from 1 to 6. In this way, we achieve reductions up to 6 to 1 in the bit rate with little degradation in the picture quality. We also consider the use of second and third order entropies of the encoded pictures to reduce the bit rate.

Next, we adapt the encoding procedure to local structural variations in the picture by first segmenting it in three areas of varying detail and then using the concepts developed for non-adaptive encoding to encode the three areas differently. The resulting encoder yields better picture quality than before and can be used to encode a broad class of pictures.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 General Background*

In the following we shall focus our attention on a particular kind of communication system for transmitting TV pictures over a digital (i.e., binary) channel. More specifically, we shall consider the source encoding aspect of this transmission system. Since the repertoire of our digital channel is limited to 1's and 0's, we can state our source encoding problem as having to convert the given picture into a stream of 1's and 0's, so that

(a) the entropy of this stream is minimized, and

(b) the distortion present in the picture reconstructed from it is also as small as possible.

To this end, let us assume the brightness value at point (ξ_1,ξ_2) in the picture is represented by $x(\xi_1,\xi_2)$, a smooth function of two spatial coordinates ξ_1 and ξ_2, bandlimited to Ω_H and Ω_V, the maximum horizontal and vertical spatial frequencies present in the picture. This two-dimensional function is converted into a one-dimensional time signal $x(t)$, bandlimited to Ω by the process of scanning. No vertical information is lost if

$$\text{line spacing} \leq \frac{\pi}{\Omega_V} \quad .$$

---

*This section has been included for the convenience of readers not familiar with the field of picture coding.

Next, we sample $x(t)$ at the Nyquist rate so that the resulting samples

$$x_n = x(n\Delta t) \quad , \quad n \geq 0 \quad ,$$

where

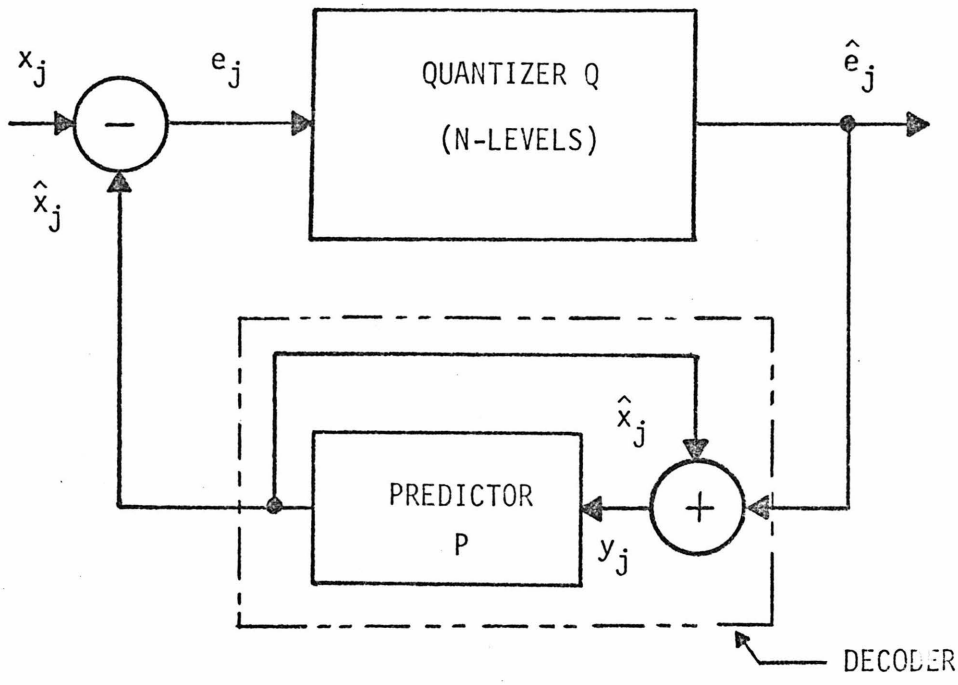
$$\Delta t \leq \frac{\pi}{\Omega} \quad ,$$

represent the picture completely.

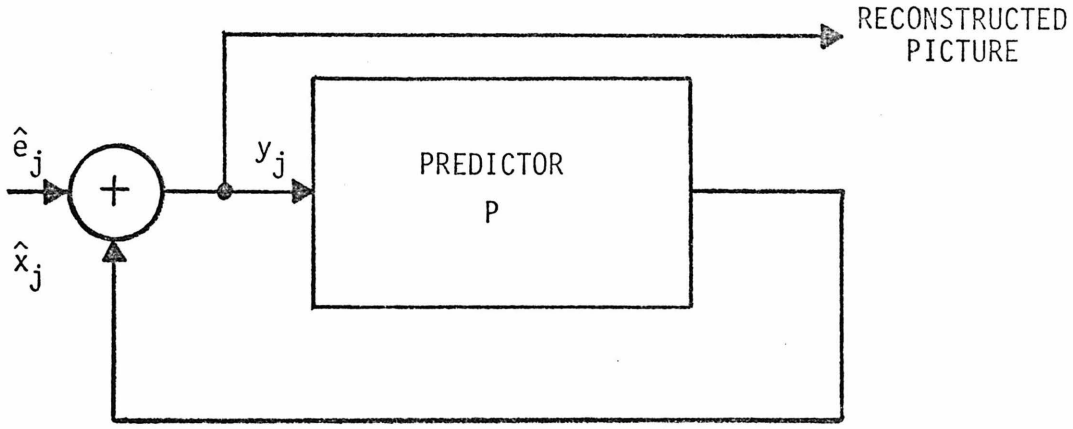
Exact representation (transmission or storage) of these x_n's would require "infinite precision machine", which is not desirable. Hence, we make our first compromise in deciding to throw away some information by dividing the range of values of x_n's in $256(=2^8)$ equal sub-ranges and by representing all the x_n's lying in the m^{th} such sub-range by the number m itself. As a result of this 256-level uniform quantization we obtain 8-bit pulse code modulated (PCM) samples called pixels. We shall refer to them as x_n's from now on. It is well known that distortion caused by such a quantization process is not visible to the human eye, therefore we shall take the ORIGINAL to be a rectangular matrix (256 x 256 in this case) of these 8-bit PCM samples.

1.2 Problem Definition and Solution Outline

In our communication system we shall address ourselves to the differential pulse code modulator (or DPCM encoder) shown in Fig. 1-1. In this predictive kind of encoder, a prediction of the current pixel value is made based upon the encoded values of previous pixels. The difference between the actual and the predicted value of the current pixel, called the prediction error, is then quantized using a quantizer



(A) DPCM ENCODER



(B) DECODER

FIG. 1-1

Q and transmitted. The decoder at the receiver end also makes a similar prediction (which would be the same as the encoder's prediction in the absence of channel errors) and adds to it the received value of the quantized prediction error to obtain the transmitted value of the current pixel. Similarly, the encoder also adds predicted pixel value and the corresponding quantized prediction error to obtain the encoded value and uses it to encode the next pixel. Clearly, to bootstrap the process one needs to transmit the values of the first few pixels. We shall not be concerned about these pixels, as they constitute only a negligible fraction of the total information transmitted. We also assume a noise-free channel; that is, the picture reconstructed at the receiver end would be the same as the encoded picture at the transmitter end.

Let $\hat{x}_j$ denote the prediction of x_j, the current pixel value. Also, let $\hat{e}_j$ be the quantized value of the prediction error e_j. Then

$$e_j = x_j - \hat{x}_j \quad ,$$

$$\hat{e}_j = Q(e_j) \quad ,$$

and

$$\eta_j = Q(e_j) - e_j$$

$$= \hat{e}_j - e_j \quad .$$

where η_j is the quantization error and Q is the quantizing function. Then y_j, the encoded value of x_j, is given by

$$y_j = \hat{x}_j + \hat{e}_j$$

$$= x_j + \eta_j$$

and d_j, the distortion in the j^{th} pixel value, is

$$d_j = y_j - x_j = \eta_j \quad .$$

The quantity that should be minimized is the entropy of $\hat{e}_j$'s, defined by

$$H = -\sum_{i=1}^{N} p_i \log^* p_i$$

where N is the number of levels in the quantizer Q, and p_i is the probability of usage of the i^{th} level of the quantizer.

The second quantity we want to minimize is the subjective distortion measure D_s of the reconstructed picture. The numerical value of D_s is inversely proportional to the viewer's subjective liking of the reconstructed picture.

Thus, ideally speaking, we should solve the following set P of related problems:

(P) Given a value of the subjective distortion measure D_s, find the predictor P and the quantizer Q such that the entropy H of the quantized prediction errors e_j is minimized; or, given a value of H, find P and Q such that D_s is minimized.

---

*Logarithms appearing in definitions of various entropies are all assumed to be to the base 2.

The solution to such a problem requires quantitative knowledge of the distortion measure D_s, which would be based on quantitative description of the visibility of quantization noise n_j as a function of its complex local surroundings. Unfortunately, such a description is lacking due to the lack of a good model for properties of the human visual system including its neurophysiology and psychophysics. In such a case, an approach to solving the problem set P must remain heuristic in nature.

In the first step of the heuristic approach our main emphasis is to reduce the magnitudes of the prediction errors e_j which leads to improved picture quality. We propose a new criterion, namely $MAX(= \max_j\{e_j\})$, which with the $VAR(= \text{variance}_j\{e_j\})$ should be used to compare different predictors. We show that in some cases (discussed in the quantization section) use of the MAX criterion only suffices, and the VAR criterion could be ignored. We also use the difference between the second order and the first order entropies of the e_j's to judge the effectiveness of predictors in decorrelating the picture. Some new predictors are proposed which are compared with the existing ones by computer simulation, using four originals. Some of the new predictors are shown to be more effective than the existing ones. As a matter of side interest, we also establish conditions on the size of the picture and the number of values its pixels can assume, under which third-order entropy of the picture is a meaningful criterion of its redundancy.

In the second step we turn to designing the quantizer Q . We wish to design it in a way that gives us some control over subjective quality and the output entropy. Designing the quantizer to minimize the "mean n^{th} power" distortion enables us to do that. Based on integer arithmetic, an efficient algorithm, having modest storage requirements, is suggested to carry out the design for any n , the power in the distortion criterion. Effects of n on the picture quality and the entropy of quantized prediction error are investigated as it varies from 1 through 6. It is shown as n increases that substantial reduction in the entropy takes place, accompanied by only small distortion appearing in smooth areas of the picture. Quantizers are compared for different originals keeping the same n , and for a fixed original with n varying from 1 to 6.

We show that further reduction in entropy, about 10-25%, is possible by using second-order entropy with respect to the vertical neighbor. Third-order entropies have also been considered.

In the third step we concentrate on adapting the encoding strategy to the local properties of the picture. First of all, an algorithm is presented to segment the picture into three areas of different details. In all the three areas, previously discussed predictors are compared with respect to the MAX and the VAR criteria as in step 1. Finally, quantizers are designed for the three areas using methods of step 2. This demonstrates the gains possible by encoding the three areas differently. Comparison of the quantizers is made for different orignals in each area, as well as for the three areas of one

original. Finally, an ensemble quantizer is designed for each area based on the prediction error statistics of all the originals in that area.

1.3 Brief Literature Survey

Contributions in the areas of standard pulse code modulation, predictive quantizing, quantizer design, and adaptive source encoding are related to our work.

The paper by Oliver, Pierce and Shannon [1] describes standard pulse code modulation. Information theoretic basis for this has been dealt with by Shannon [2]. Oliver [3] and Graham [4] put these concepts in the context of efficient encoding of pictures. Schreiber [5] measured third-order probability distributions of TV pictures.

DPCM systems are primarily based on Cutler's work [6]. Later efforts in predictive quantization are due to Oliver [3], Harrison [7], Elias [8], Graham [9], O'Neal [10], Connor, Pease and Scholes [11], Kummerow [12], and Thoma [13]. They proposed and investigated a number of one-dimensional and two-dimensional predictors about which we shall say more later on. Connor [14] also talked about reducing visibility of transmission errors. Kretzmer [15] determined some statistics of typical TV pictures.

In the area of quantizer design Panter and Dite [16] discuss non-uniform quantization, Max [17] discusses theoretical basis for the optimal quantizer design, and Bruce [18] investigates the globally optimal quantizer. Quantization of TV signals has been dealt with by Bernard [19], and Limb and Mounts [20]. Incorporation of some viewer

criteria in picture coding schemes has been carried out by Graham [9], Limb [21], Candy and Bosworth [22], and Limb [23]. Entropy of quantized signals was measured by Limb [24].

Adapting the coding scheme to local variations in the picture was suggested by Graham [9]. Other studies on adaptive quantization were done by Golding and Schultheiss [25], Wintz and Tasto [26], Brown and Kaminsky [27], Limb [28], Cohen [29], Kummerow [30], and Thomas [13].

CHAPTER 2

PREDICTION

In this chapter we define the prediction problem, following which we discuss a number of predictors and compare them using the MAX and the VAR of prediction errors. In the last section we consider the first and higher order entropies of the original and the "prediction error pictures".

2.1 The Prediction Problem

We define the predicted value $\hat{X}_j$ of a pixel X_j as a linear combination of the brightness values of some neighboring pixels occurring prior to X_j in the same line or in the preceding lines, see Fig. 2-1. We shall refer to this prediction as based on the original or the reconstructed picture, depending on from which picture the neighbors of X_j have been chosen.

Thus to devise a predictor scheme we have to (a) choose some neighbors of X_j whose values will be used in the linear combination and (b) find their respective coefficients so that the difference $e_j = X_j - \hat{X}_j$ is as small as possible. Let us assume we know which neighboring pixels to use to predict the value of X_j. Then to choose their coefficients optimally we need a criterion which could be used to measure the aggregate magnitude of e_j's. One such criterion is VAR = variance of e_j's. This leads to minimum least square predictor and also has the intuitive appeal of minimizing power in the error signal (see [10] and [31]). However, in such an approach we are not able

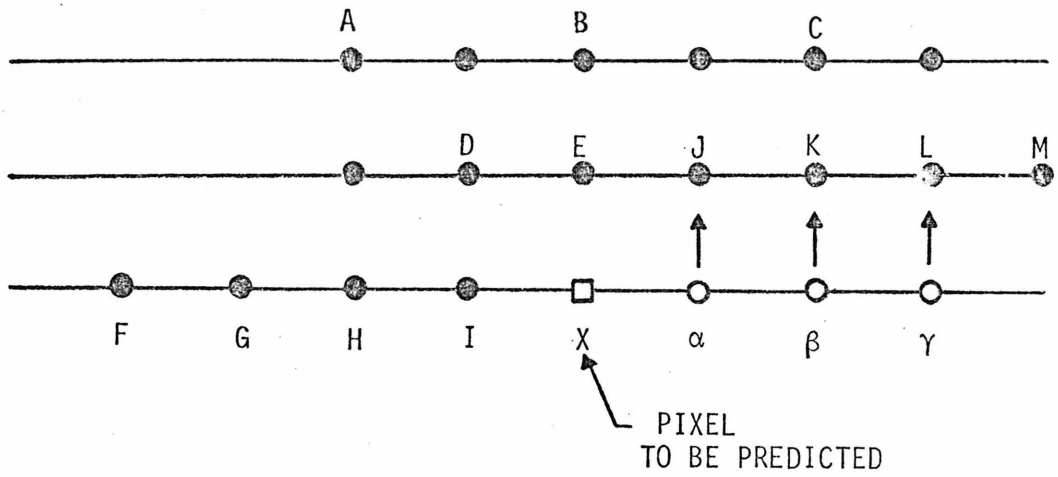


FIG. 2-1

to assess the magnitudes of relatively large errors that occur near sharp edges and the boundaries of various objects in the picture. Such errors, although their frequency of occurrence is small, affect the picture quality greatly. Besides this, the coefficients computed using this approach would be picture dependent, as they depend on the cross-correlations of X_j with its neighbors. With these considerations in mind, we introduce the MAX criterion defined as

$$\text{MAX} = \text{essential maximum of } e_j\text{'s} \ .$$

Here, the essential maximum has been introduced to eliminate the effects of large noise spikes, etc. in the picture, which are much less frequent than the large errors that we are after. To compute the essential maximum of a given set of e_j's we make a plot of their frequency of occurrence versus their magnitudes. Then MAX is the maximum spread of this plot in either direction from 0, after neglecting the isolated clusters of points at both the far ends that are expected to be due to artifacts like noise. We shall show in Chapter 3 that MAX is very useful for designing quantizers that give subjectively better pictures.

Coming back to the problem of finding the coefficients in the linear combination, due to the use of the above two criteria, mathematical tractability of the solution has been lost. Therefore, we resort to intuitive guidelines, and some concepts of extrapolation and interpolation from numerical analysis to find not only the coefficients but also which neighboring pixels to use.

In the following section we present various predictors that we shall compare using the MAX and the VAR criteria in Section 2.3. As will be noted later on also, some of these predictors have already appeared in the literature (see Refs. [7] through [13]).

2.2 The Predictors

Depending upon the underlying principles behind the prediction algorithms, we have divided them into two major classes. As shown in Table 2-1, class 1 contains the predictors based on the principle of extrapolation, while class 2 consists of what we call interpolative type predictors.

Class 1:

This class of predictors contains extrapolative type of predictors and has been divided into three subclasses. Predictors based on one-dimensional extrapolation form classes 1-A and 1-B, while those using two-dimensional extrapolation come under class 1-C.

Under class 1-B, the simplest predictor, i.e., Pred #10, is obtained by using 0^{th} order extrapolation yielding I as the prediction of X. Similarly, linear extrapolation gives us $I + (I - H)$ as the prediction for X in Pred #11. Use of this predictor causes large errors due to overshoot near sharp edges [7]. But instead, if we were to use $I + (I - H)/2$ as the prediction, we would reduce the shape of the extrapolating line to half of its previous value which would reduce the overshoot effect. In essence, we have damped the predictor giving rise to Pred #9, a 1/2-order predictor in the x direction. Note that if we substitute the Taylor expansion about the point X of the

TABLE 2-1

	PRED #	PREDICTED VALUE ($\hat{X}$)	DESCRIPTION
CLASS 1-A	5	(3I+3E-H-B)/4	1/2 order in x and y directions
	6	(3J-C)/2	1/2 order in $+45^{o}$ direction
	7	(3D-A)/2	1/2 order in -45^{o} direction
	8	(3E-B)/2	1/2 order in y direction
	9	(3I-H)/2	1/2 order in x direction
CLASS 1-B	10	I	0^{th} order in x direction (last element prediction)
	11	2I-H	1^{st} order in x direction
	12	3I-3H+G	2^{nd} order in x direction
	13	4I-6H+4G-F	3^{rd} order in x direction
	14	(5I-4H+G)/2	2^{nd} order in x direction (using divided differences)
	15	(16I-15H+6G-F)/6	3^{rd} order in x direction (using divided differences)

TABLE 2-1 (continued)

	PRED #	PREDICTED VALUE ($\hat{X}$)	DESCRIPTION
CLASS 1-C	19	I + (E-D)/2	1/2 order in x direction. Slope being approximated from the previous line
	20	(I+E)/2	0th order in x and y directions
	22	I - D + E	Planar prediction
CLASS 2-A	21	(I+J)/2	1st order interpolative
	23	(-H+4I+4J-K)/6	3rd order interpolative
	24	$(-\frac{45}{17}H + 10I + 8J - \frac{40}{17}K)/13$	3rd order interpolative (kinked along HIXJK)
	25	(G-6H+15I+15J-6K+L)/20	5th order interpolative
CLASS 2-B	16	(-4H+16I+7K-2L)/17	3rd order interpolative (variation 1 - use #21 for α and β)
	17	(-H+4I+4J-3E-K)/3	3rd order interpolative (variation 2 - use #22 for α and β)
	18	(-18H+62I+40K-22L+3M)/65	3rd order interpolative (variation 3 - use #23 for α and β)

brightness function $X(\xi_1,\xi_2)$, not even the first order terms cancel out in the expression $I + (H - I)/2$. Despite this fact, this predictor performs better than #11 as we shall see later. This way we obtain class 1-A containing various 1/2-order predictors in various directions. Pred #5 has been obtained by taking the average of #8 and #9, which extrapolate in the y and x directions, respectively.

Coming back to the other members of class 1-B, Pred #'s 12 and 13 have been obtained by passing second and third order polynomials through three and four previous neighbors of X , respectively, and then using these polynomials to compute the value at X . To obtain Pred #14 we write Taylor expansion of the brightness function about the point I and keep terms only up to the second order. Then we replace the first and second order derivatives in the expansion by the first and second divided differences of the function at the point I, and find the value at point X . Similarly, Pred #15 is obtained by retaining up to third order terms and using up to third divided differences of the function; refer to Appendix B for details.

Class 1-C contains Pred #'s 19, 20 and 22. In #20, average of I and E, the pixel values of the immediate horizontal and vertical neighbors of X are formed, while in #22 a plane is passed through I, E, and D to compute the ordinate $(I - D + E)$ at X as the predicted value. It can also be thought of as similar to #11, except the slope of the extrapolatihg line is approximated from the previous line to obtain $I + (E - D)$, instead of $I + (I - H)$ as the prediction of X . Then Pred #19 is formed, just like #9, by reducing this approximated

slope by half to obtain $I + (E - D)/2$ as the prediction of X .

Class 2:

Class 2 consists of what we call interpolative type predictors. To illustrate the basic idea here we take the example of Pred #23. To predict X let us pretend that α and β are available. Then we pass a third degree polynomial through H, I, α and β whose ordinate at X is

$$X = (-H + 4I + 4\alpha - \beta)/6 \qquad (I)$$

Next we approximate α and β by J and K respectively to obtain

$$\hat{X} = (-H + 4I + 4J - K)/6$$

as the prediction of X . On the other hand, we can approximate α and β by a variety of other methods. For instance using Pred #22 formula we can write

$$\alpha \cong X + J - E \qquad (II)$$

$$\beta \cong \alpha + K - J \qquad (III)$$

Thus we can solve Eqs. I, II and III for $\hat{X}$ as

$$\hat{X} = (-H + 4I + 4J - 3E - K)/3$$

which gives us Pred #17. Pred #'s 16 and 18 are obtained similarly. For details please refer to Appendix B. These three predictors are included under class 2-B. To obtain Pred #24 we pretend that the path HIXJK is a straight line and then pass a cubic through the four known points. It is called kinked third order interpolative predictor.

Pred #25 is similar to #23 except we use fifth degree polynomial for interpolation. Pred #'s 21, 23, 24 and 25 form class 2-A. Among the predictors we have described, Pred #'s 10 through 13, and 19 through 21 appear in Refs [7] to [13] inclusive.

2.3 Comparing Predictors

As discussed in Sec. 2.1, we shall use the MAX and VAR of the prediction errors to compare the various predictors of Table 2-1. In actual encoding, prediction of pixels is to be based on the reconstructed picture for which we need a quantizer also; see Fig. 1-1. In the absence of a quantizer at this stage, we shall look at the MAX and the VAR of the e_j's when prediction is based on the original picture.

We have chosen four originals for our work. They are MOON, CHECKERED GIRL (or just CHECK), KAREN and AERIAL. They appear as PHOTO #'s 1, 2, 3, and 4 in Appendix A. Choice of these pictures is governed by the range of contrast and the amounts of large and small details they contain.

We have simulated all our predictors on the computer and obtained the MAX and the VAR of the prediction errors for each original as shown in Figs. 2-2 and 2-3, respectively. Points in these graphs should be read on the scale which is on the same side of the broken middle line as the points themselves.

In Fig. 2-2, note the low values of MAX for Pred #'s 23, 24, and 25 in KAREN, CHECK and AERIAL. In MOON, the value of MAX is much higher than in other originals except for Pred #'s 8, 20 and 5.

In Fig. 2-3, AERIAL shows relatively large value of

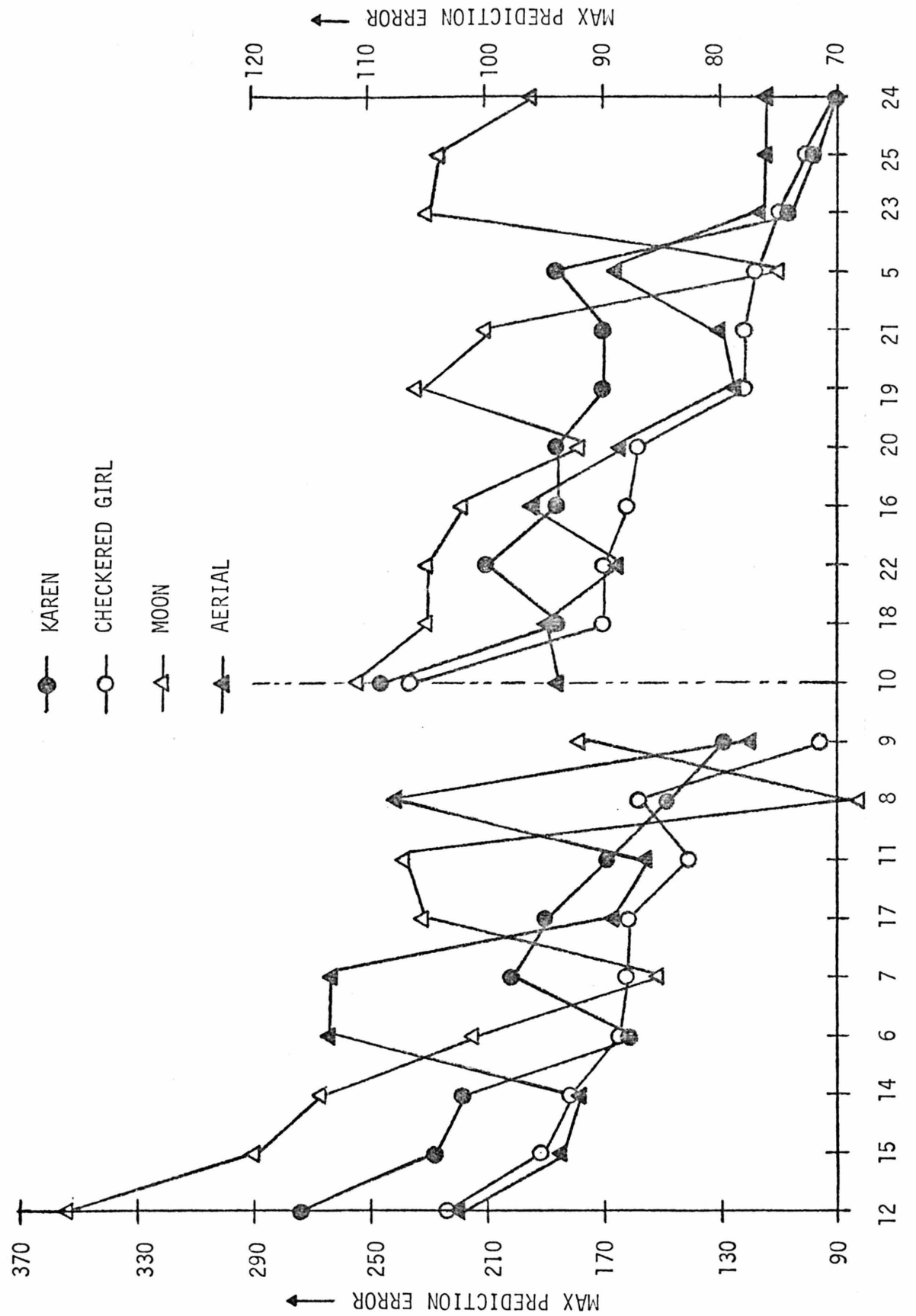


FIG. 2-2 PLOT OF MAX PREDICTION ERROR versus PRED #

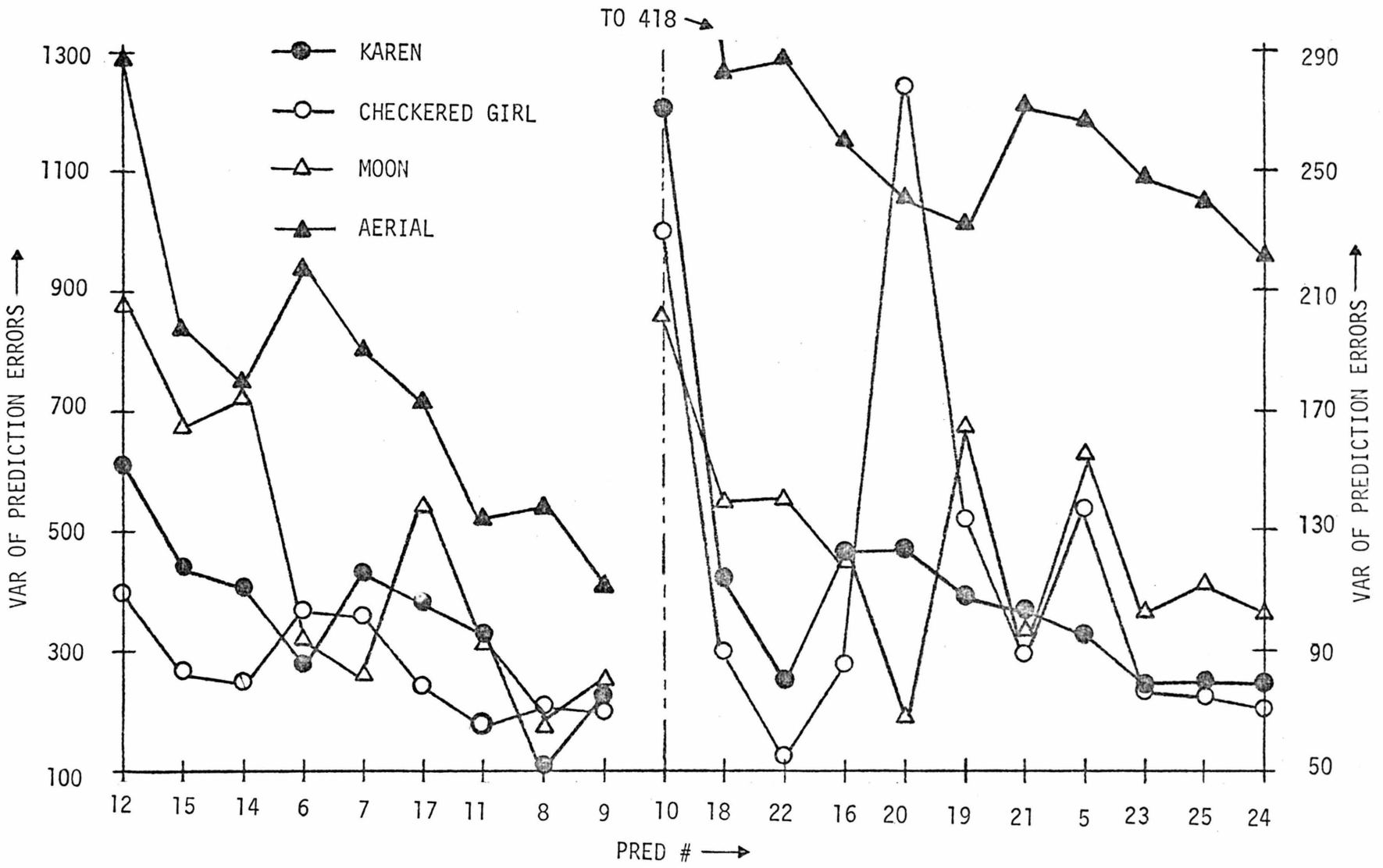


FIG. 2-3 VAR OF PREDICTION ERRORS versus PRED #

VAR. Notice the dip and the peak at Pred #22 and #20 respectively for CHECK, and the dip at #20 for MOON. To estimate the general trends we took the average values of MAX and VAR over all the originals and plotted them against Pred #'s in Fig. 2-4. Note that the predictors have been ordered according to diminishing values of MAX and that we have focussed our attention on Pred #'s 5, 10, and 18-25 only, as the others have larger values of MAX as well as VAR. Using this figure we classify these predictors in four categories in order of decreasing effectiveness with respect to MAX AND/OR VAR. Under each category we present the Pred #'s in order of increasing preference in Table 2-2. Using the third column of this table, we obtain the following list of predictors in order of increasing preference--Pred #'s 10 | 18,16,22 | 20,19,5,21 | 23,25,24 where vertical slashes indicate category boundaries. Clearly, the interpolative predictors of class 2-A (Table 2-1) are among the better ones in the whole lot. At this point we would like to draw the reader's attention to PHOTO #'s 5, 6, 7, and 8 which are prediction error pictures and the histograms corresponding to Pred #'s 10 and 24 for KAREN. Description of these is included in the PHOTO-KEY preceding the photographs. We would also like to point out that the MAX for Pred #24 is determined from PHOTO #8 by neglecting the isolated little clusters at the two far ends. Such clusters are absent in PHOTO #7.

The predictors found to be relatively better in the above discussion are desirable from one more point of view, hitherto undiscussed; that is, propagation of transmission errors. Suppose in the case of

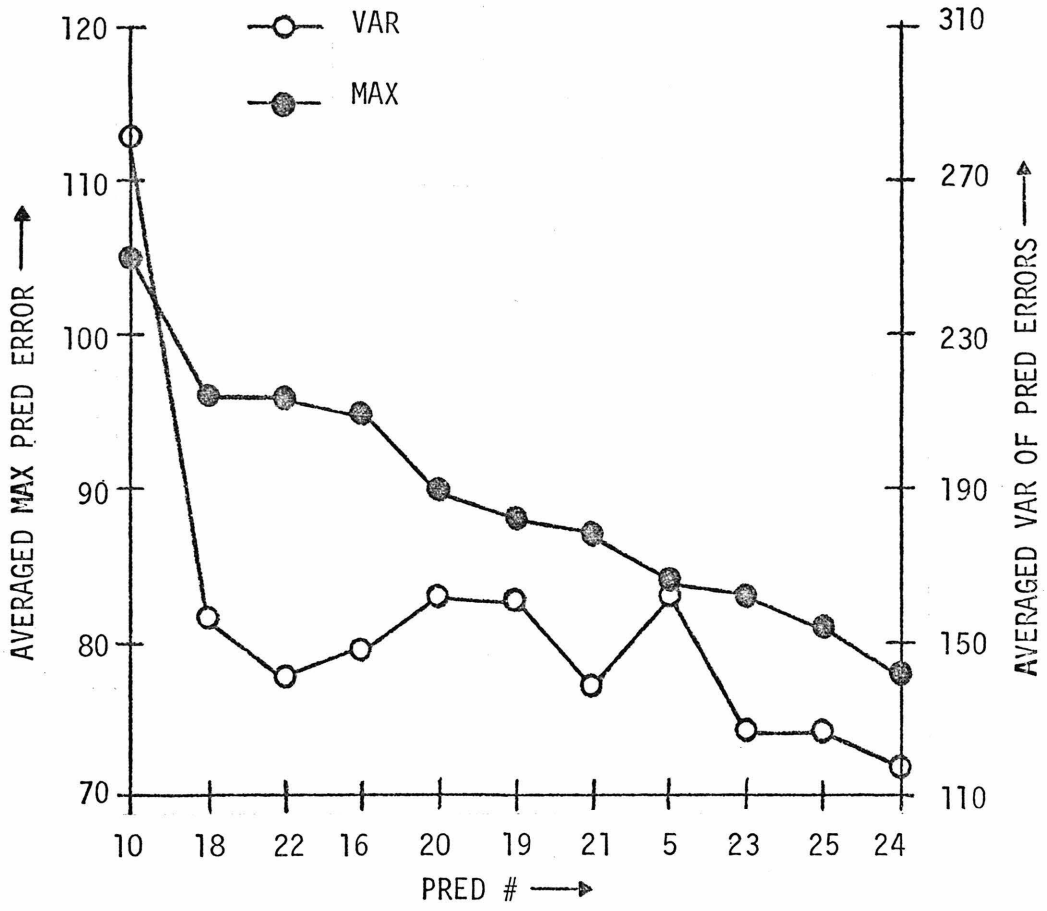


FIG. 2-4 AVERAGED MAX AND VAR OF PREDICTION ERRORS
(AVERAGE TAKEN OVER ALL THE FOUR ORIGINALS)

TABLE 2-2

CLASSIFICATION OF PREDICTORS ACCORDING TO MAX AND/OR VAR IN ORDER OF DECREASING EFFECTIVENESS

Category #	Pred # (w.r.t. MAX)	Pred # (w.r.t. VAR)	Pred # (w.r.t. MAX and VAR)
1	5,23,25,24*	23,25,24	23,25,24
2	20,19,21	16,22,21	20,19,5,21
3	18,22,16	5,20,19,18	18,16,22
4	10	10	10

*Pred #'s are in the order of increasing preference.

Pred #10 ($\hat{X} = I$) element X is received as $X + \Delta X$. Then the prediction error (and hence the reconstruction error) in all the future elements of the same line will also be ΔX. But, if we were to use Pred #21 ($\hat{X} = (I+J)/2$), the errors in the following elements would be $1/2$, $1/2^2$, $1/2^3$, etc. times ΔX; in other words, the overall effect of transmission error would be much less. This property is enjoyed by the predictors in which coefficients of the neighboring elements are less than 1. In particular, Pred #'s 5 and 20, which are averaging extrapolative type and all the predictors of the interpolating type (Class 2 in Table 2-1) except #17, are of this nature.

2.4 Entropy Considerations

For the original picture let us denote the average entropy per pixel by H_o and the average entropy per pixel given its previous neighbor by S_o (loosely referred to as second order entropy also). Then the minimum rate at which we can transmit this picture using one pixel at a time is H_o [2]. But, if we were also to consider each pixel's past "contained" in its previous neighbor, we would reduce the minimum required rate to S_o - a fractional reduction of

$$R_o = \frac{H_o - S_o}{H_o}$$

We can name this quantity as second order redundancy (or just redundancy) in the picture. Obviously this would be higher for a picture in which there is higher element-to-element correlation. Now, in doing prediction on the picture we are, from information theoretic point of

view, making use of this correlation to generate a stream of decorrelated prediction errors. Thus, if H_e and S_e denote respectively the first and the second entropies of the error picture, then the quantity

$$R_r = \frac{H_o - H_e}{H_o}$$

represents the fractional redundancy "removed" by the predictor or the fractional reduction in the minimum channel capacity required to transmit the picture in its modified form. Also,

$$R_e = \frac{H_e - S_e}{H_e}$$

the redundancy still remaining in the error picture, indicates the maximum gain achievable by encoding pixels using their past. Thus the overall obtainable reduction is

$$R_R = \frac{H_o - S_e}{H_o}$$

We also call it the net "removable" redundancy for the given predictor and it is related to R_r and R_e by

$$R_R = R_r + R_e(1 - R_r)$$

We carried out the measurement of all the four R's for the originals and the predictors we have. The first and second order entropies and the redundancies for the originals are tabulated in Table 2-3. Notice the low value of R_o for MOON and AERIAL followed by KAREN and CHECK. R_r's for various predictors are shown in Fig. 2-5. If Pred #10 for KAREN were making optimal use of the previous elements, its R_r

TABLE 2-3

SECOND ORDER ENTROPIES AND REDUNDANCIES

IN THE ORIGINALS

Picture	H_o	S_o	$R_o(=\frac{H_o - S_o}{H_o})$
			(Percent)
MOON	6.707	5.226	22.1
CHECK	7.405	4.426	40.2
KAREN	7.520	4.870	35.2
AERIAL	7.312	5.696	22.1
ENSEMBLE	7.236	5.055	30.1

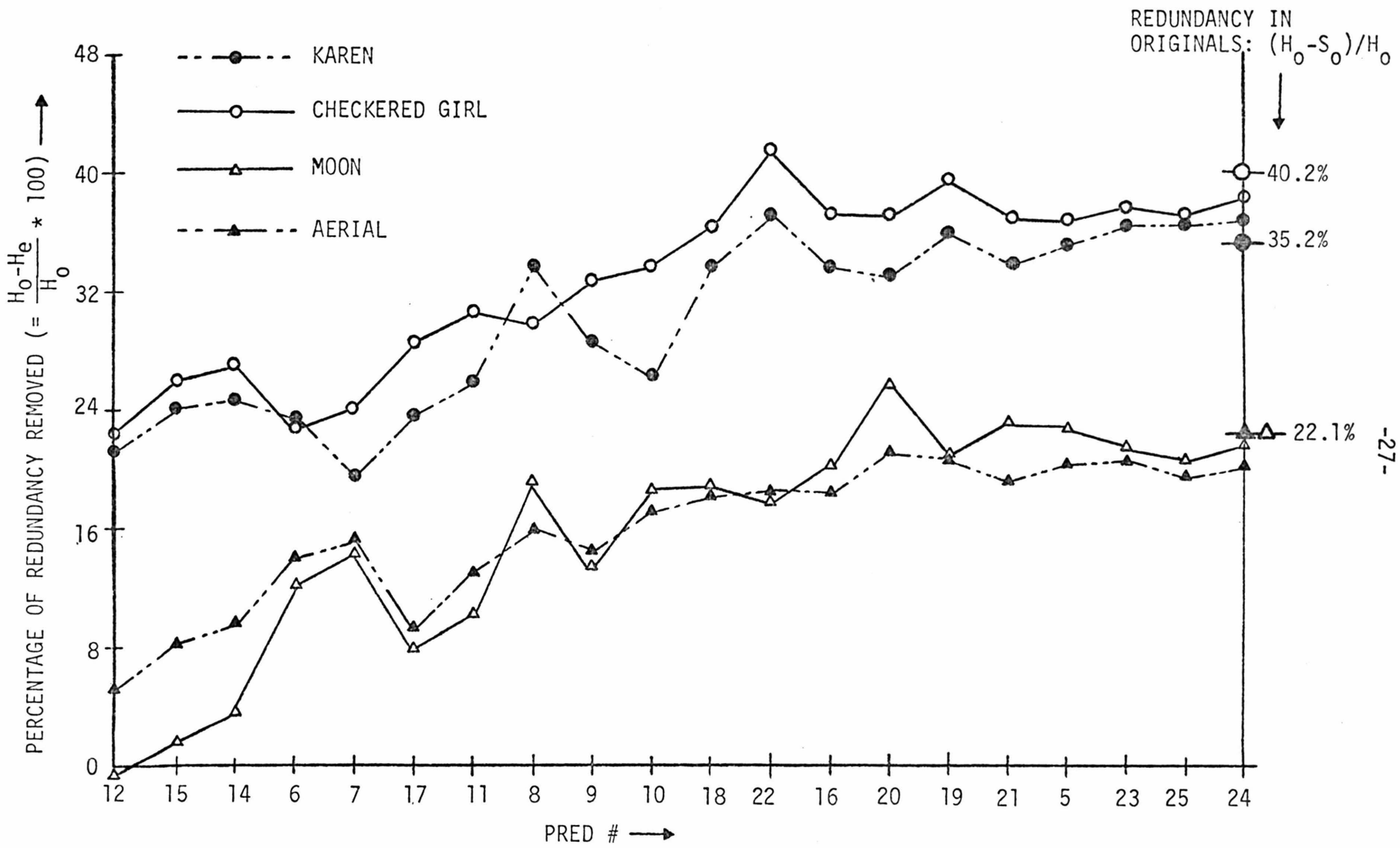


FIG. 2-5. PERCENTAGE OF REDUNDANCY REMOVED AFTER PREDICTION versus PRED #

should be equal to R_o , but on the contrary, it is much less than R_o . R_r for Pred #24 is greater than R_o . This is explicable on the grounds that this predictor uses a number of other neighboring elements while R_o is computed only on the basis of the previous neighbor. On the whole, MOON and AERIAL have similar behavior and such is the case of KAREN and CHECK also. This trend continues for R_e and R_R in Figs. 2-6 and 2-7 respectively. In Fig. 2-8 we show $\hat{R}_r$, $\hat{R}_e$, and $\hat{R}_R$ the ENSEMBLE values defined using $\hat{H}_o$, $\hat{H}_e$, $\hat{S}_o$, and $\hat{S}_e$, the averages over all the four originals of the first and the second order entropies of the original and the error pictures.

The trends in these figures are self-explanatory. In Table 2-4 we have compiled the minimum values of H_e and S_e for all the originals along with the predictors that attain this value. These are the predictors one would use in a completely information preserving scheme to encode the pictures.

We shall finish this chapter with a brief note about higher order entropies. We could try to extend the arguments given above by using third order entropies, in which case we would be counting the number of occurrences of the triplets of intensities corresponding to three neighboring pixels in the picture. For the 8 bit PCM pixels there are 2^{24} such triplets possible while the number of triplets we can observe in our 256x256-pixel originals is about 2^{16} which is much less than their total possible number. That is, the implicit assumption that the probabilities of occurrences of triplets are approximated by their relative frequencies would necessarily be at fault. This allows us to construct a picture (see PHOTO #39. See also Appendix C for more details

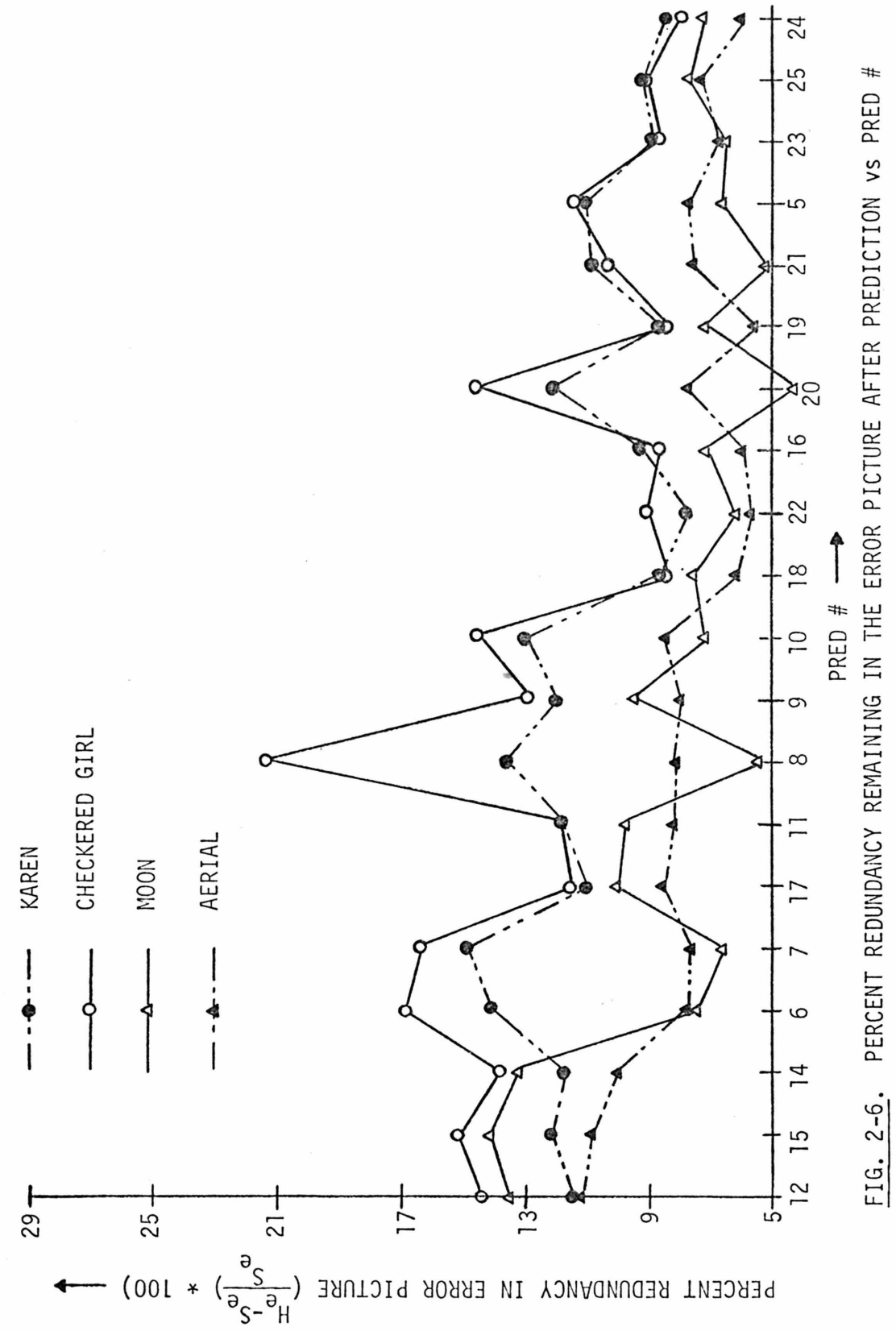


FIG. 2-6. PERCENT REDUNDANCY REMAINING IN THE ERROR PICTURE AFTER PREDICTION vs PRED #

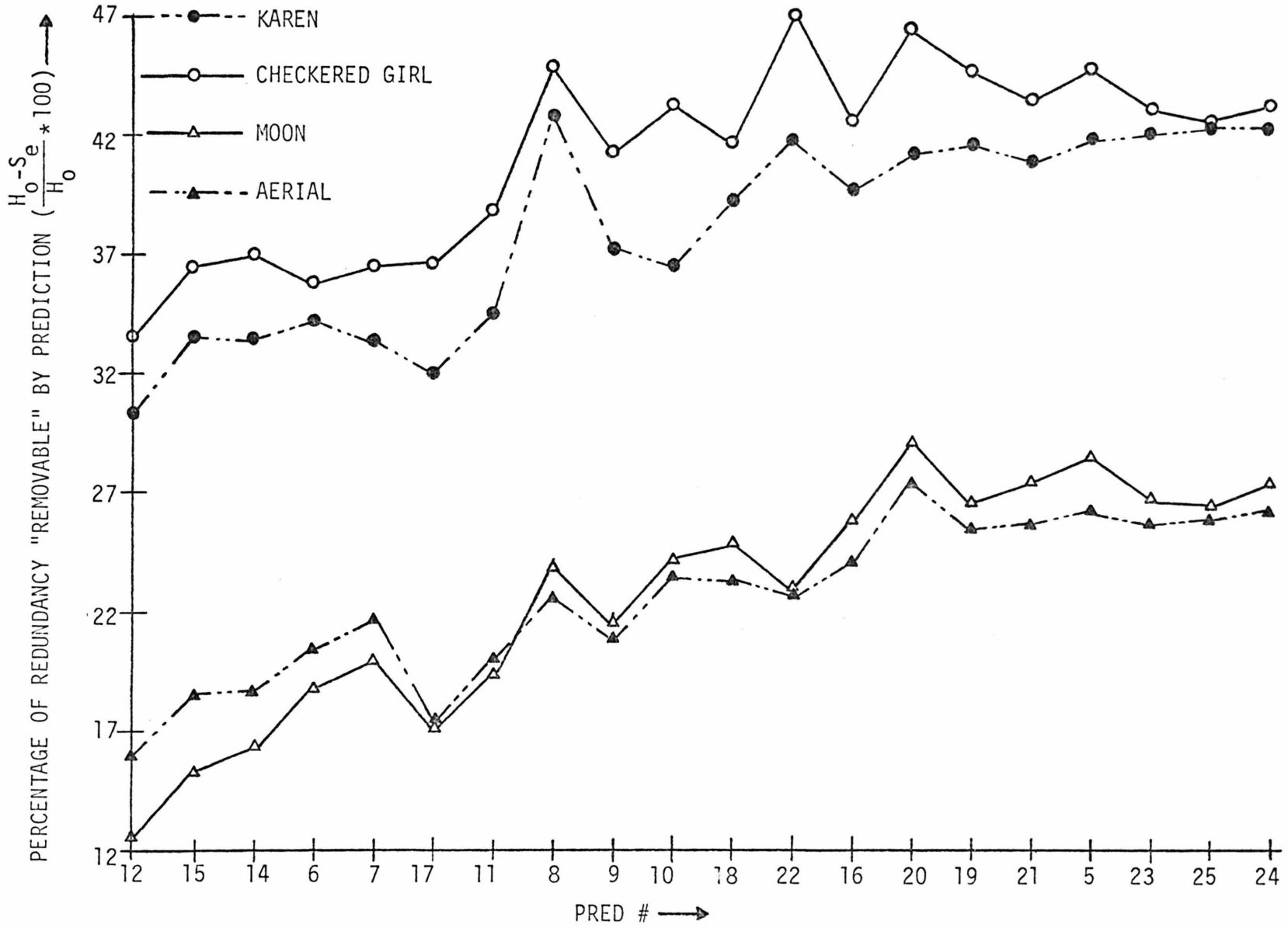


FIG. 2-7. PERCENTAGE OF REDUNDANCY "REMOVABLE" BY PREDICTION FOR A GIVEN PREDICTOR vs PRED #

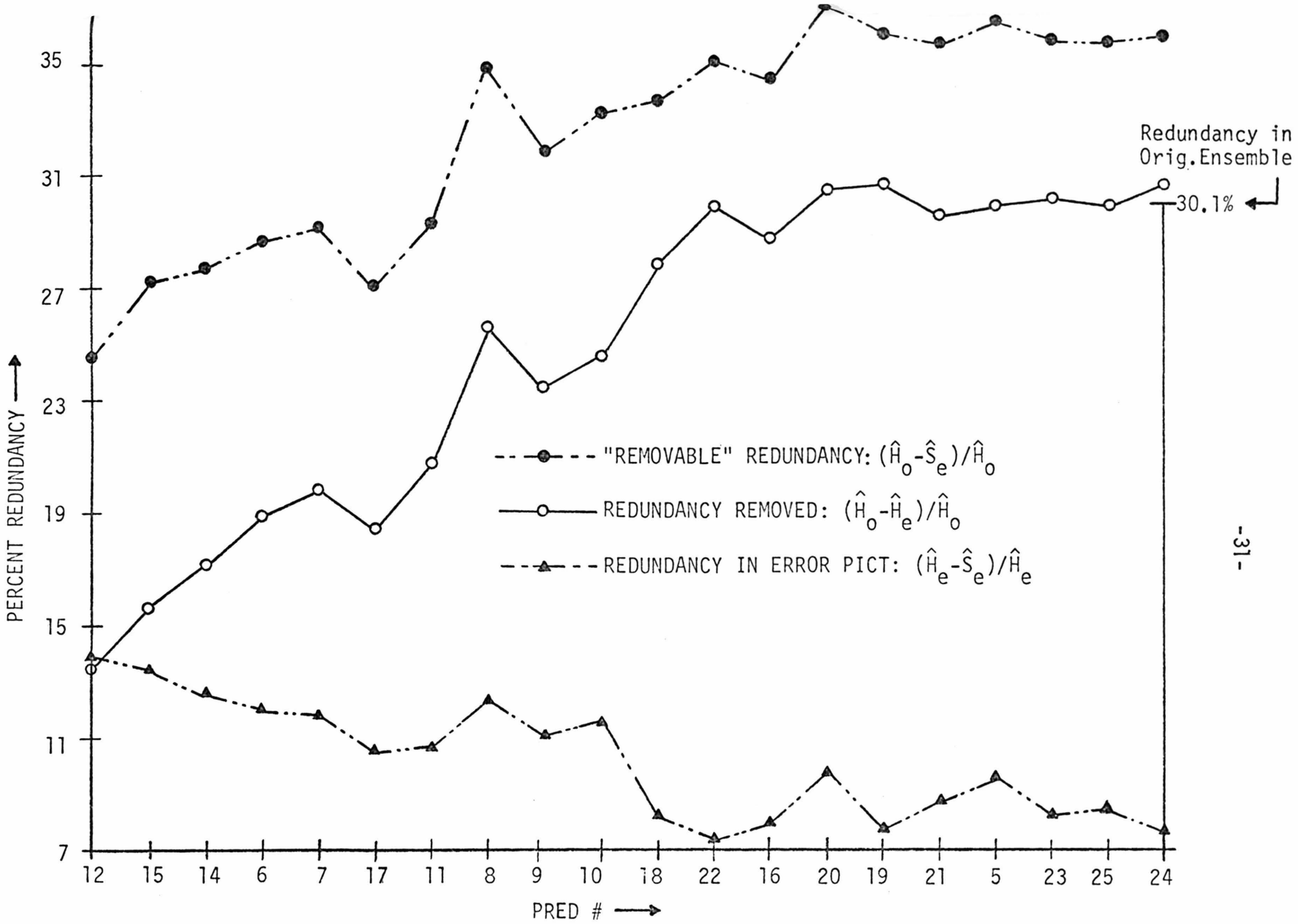


FIG. 2-8. FOR THE ENSEMBLE OF PICTURES PERCENTAGE OF REDUNDANCY REMOVED, REDUNDANCY IN THE ERROR PICTURE, AND REDUNDANCY "REMOVABLE" BY A GIVEN PREDICTOR versus PRED #

TABLE 2-4

THE MINIMUM VALUES OF H_e AND S_e, THE FIRST AND SECOND (w.r.t. PREV. PIXEL) ORDER ENTROPIES OF PREDICTION ERRORS

Picture	Min H_e	Min S_e	Min H_e Attained by Pred #	Min S_e Attained by Pred #
MOON	4.975	4.765	(20)*,21,5,23	(20),5,24,21
CHECK	4.322	3.927	(22),19,24,23	(22),20,8,19
KAREN	4.740	4.301	(22),24,25,23	(8),25,24,23
AERIAL	5.775	5.331	(24),20,19,23	(20),5,24,25
ENSEMBLE	5.075	4.623	(19),(24),20,25	(20),5,19,24

*Encircled Pred # attains the minimum value. Others follow closely in the given order. Difference between the first and second is about 0.03 in all cases except in H_e for CHECK where it is 0.12.

and the algorithm to construct such a picture.) in which all the triplets (formed by three consecutive pixels) are distinct, i.e., we expect the picture to be a fully "third order random" picture but measurements show its first and second order entropies are 8.000 and 7.994 respectively, while its third order entropy is 0 . Such an anomaly would not arise if we were to increase the size of the picture to 4096x4096, or else if we were to use 5 bits PCM pixels. Then we would be able to use the third order entropy in our redundancy calculations.

CHAPTER 3

QUANTIZATION

In this chapter we shall consider the design of quantizer to minimize "mean n^{th} power" of the quantization errors. We shall study the effects of "n" on the picture quality and the entropy of the quantized prediction errors. Finally, the second and third order entropies of the quantized errors and the design of Huffman codes will be considered.

3.1 The Quantization Problem

Let us denote the number of quantization levels by N and the levels themselves by $y_1, y_2, \cdots y_N$. Although in Fig. 1-1 the input variable to the quantizer is e_j, let us temporarily call it x. Also, let the input thresholds be $X_1, X_2, \cdots, X_{N+1}$ (Fig. 3-1) such that

$$X_1 \leq x \leq X_{N+1}$$

and

$$X_1 < X_2 < \cdots < X_{N+1} \quad .$$

Then we define the quantizer as a function Q that assumes the value y_k whenever the input variable x is between X_k and X_{k+1}, that is,

$$Q(x) = y_k \quad \text{for} \quad X_k < x \leq X_{k+1}$$

$$\text{and} \quad 1 \leq k \leq N \quad .$$

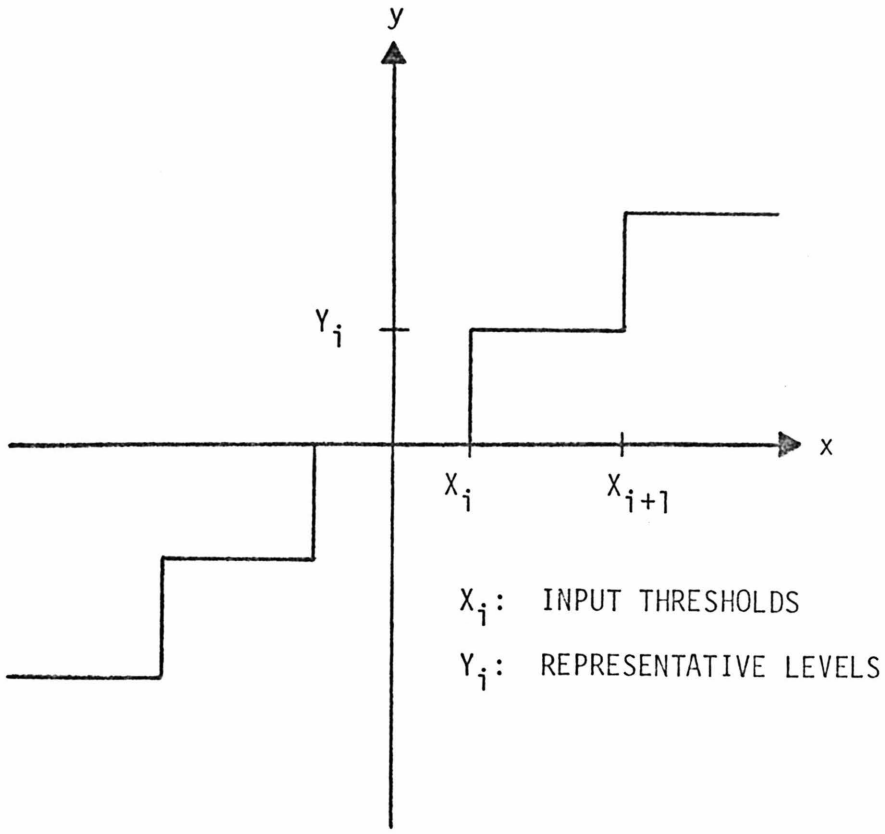


FIG. 3-1

Keeping in mind that x is a discrete variable and that $x - y_k$ is the quantization error for a typical input mapped to y_k, "mean n^{th} power" distortion for the k^{th} output level is given by

$$D_k = \sum_{X_k < x \leq X_{k+1}} |x - y_k|^n \, p(x)$$

where $p(x)$ is the probability of occurrence of x. The total distortion D, caused by the quantizer or the mean n^{th} power of the quantization errors becomes

$$D = \sum_{k=1}^{N} D_k$$

$$= \sum_{k=1}^{N} \sum_{X_k < x \leq X_{k+1}} |x - y_k|^n \, p(x) \quad .$$

Thus, given the probability distribution function $p(x)$ and the number of quantization levels N, the quantization problem consists of finding the input thresholds X_k's and the output representative levels y_k's such that the distortion D is minimized.

Here, since D is a function of the discrete variables X_k's and y_k's, it cannot be differentiated to find necessary conditions for local minimum. However, we can make use of the forward and the backward differences of D with respect to X_k's to characterize its local minimum. The necessary condition for this is given by

$$X_k = \left[\frac{y_{k-1} + y_k}{2} \right] , \quad 2 \leq k \leq N \quad , \qquad \text{(I)}$$

where the square brackets represent the largest integer contained in

the quantity inside them.

Derivation of Equation I is given in Appendix D. For the continuous case and $n = 2$ (i.e., mean square distortion case) the necessary conditions have been discussed by Max [17]. Our condition is a counterpart of one of his conditions. Extensions of his other condition to values of n other than 2 and for discrete case are also included in Appendix D, where we also give rationale behind not using them in our investigation. In our approach to design the quantizer, we shall use only the necessary condition given by Eq. I. Following is an account of the algorithm to carry out the design.

The Algorithm A1

In this iterative algorithm let us assume that we are given $p(x)$, the probability distribution function and $x_k^{(0)}$, $y_k^{(0)}$, $1 \leq k \leq N$, as an initial guess. Then we compute the next quantizer consisting of $x_k^{(1)}$'s and $y_k^{(1)}$'s in two steps.

In step 1 we fix $x_k^{(0)}$'s and find $y_k^{(1)}$'s one by one for $1 \leq k \leq N$. For a fixed k, $y_k^{(1)}$ is found so that

$$D_k = \sum_{x_k^{(0)} < x \leq x_{k+1}^{(0)}} |x - y_k^{(1)}|^n p(x)$$

is the minimum value of the distortion D_k in the k^{th} level. Since y_k can assume only finite number of discrete value, the value of $y_k^{(1)}$ can be found by a search procedure. Notice that $y_k^{(1)}$ is constrained by

$$x_k^{(0)} \leq y_k^{(1)} \leq x_{k+1}^{(0)} \quad ,$$

for if $y_k^{(1)}$ is outside the interval $(x_k^{(0)}, x_{k+1}^{(0)})$ we can make D_k smaller by bringing $y_k^{(1)}$ closer to the interval which contradicts the fact that $y_k^{(1)}$ was optimal. Thus our search for the optimal value $y_k^{(1)}$ is limited to a rather small set of values and could be made very fast in practical situations. In our implementation we have used binary search procedure. See Knuth [32] for other search algorithms. This procedure is repeated for all k from 1 to N.

In the second step we use these values of $y_k^{(1)}$ to compute $x_k^{(1)}$'s by enforcing the necessary condition

$$x_k^{(1)} = \left[\frac{y_{k-1}^{(1)} + y_k^{(1)}}{2} \right], \qquad 2 \leq k \leq N ,$$

which was presented as Eq. I above.

With the $x_k^{(1)}$'s and $y_k^{(1)}$'s we can do the second iteration, and so on. In each iteration we are only reducing the distortion D which is bounded below by 0. Hence our iterations must converge. In practice we may decide to stop iterating when the difference of the distortions obtained in two successive iterations becomes smaller than a predetermined quantity ε. As for the initial guess, we used uniform quantizer (i.e., x_k's and y_k's uniformly spaced) in our work. This completes description of the algorithm.

This algorithm can be used to design the quantizer if we are given the statistics of the prediction errors e_j. As portrayed in Fig. 1-1, in the final encoding scheme the prediction errors are based on the reconstructed picture, and hence they are dependent on the very

quantizer that we are seeking. We carry out design of such a quantizer by the following algorithm.

Complete Design of the Quantizer--Algorithm A2

To begin with, no encoding is done. As a first approximation we compute the statistics of the prediction errors based on the original. We obtain the probability density (p.d.) function of the prediction errors by assuming that their relative frequencies of occurrences closely resemble their respective probabilities. Then we use the algorithm A1 and this p.d. function to design our first quantizer. Using this in the next iteration of the complete design, we simulate the DPCM loop of Fig. 1-1 and, in the process, collect the statistics of e_j's which are now based on the newly reconstructed picture. This gives us a new p.d. function which, with the aid of algorithm A1 yields the second quantizer to be used in the DPCM loop the second time around, and so on. This way we expect the resulting quantizer to be matched with the p.d. function of the prediction errors to which it will give rise in the encoding process. This should in turn lead to a reduction in the resulting distortion.

The natural question that arises now is, when the iterations should be stopped. In the absence of precise analytic description of the picture, the answer to this question must remain empirical in nature. We simulated the algorithm A2 on the computer and tested it for a number of cases with varying number of levels N in the quantizer and the power "n" in the distortion criterion. Two representative cases are shown in Fig. 3-2. In this figure the vertical axis

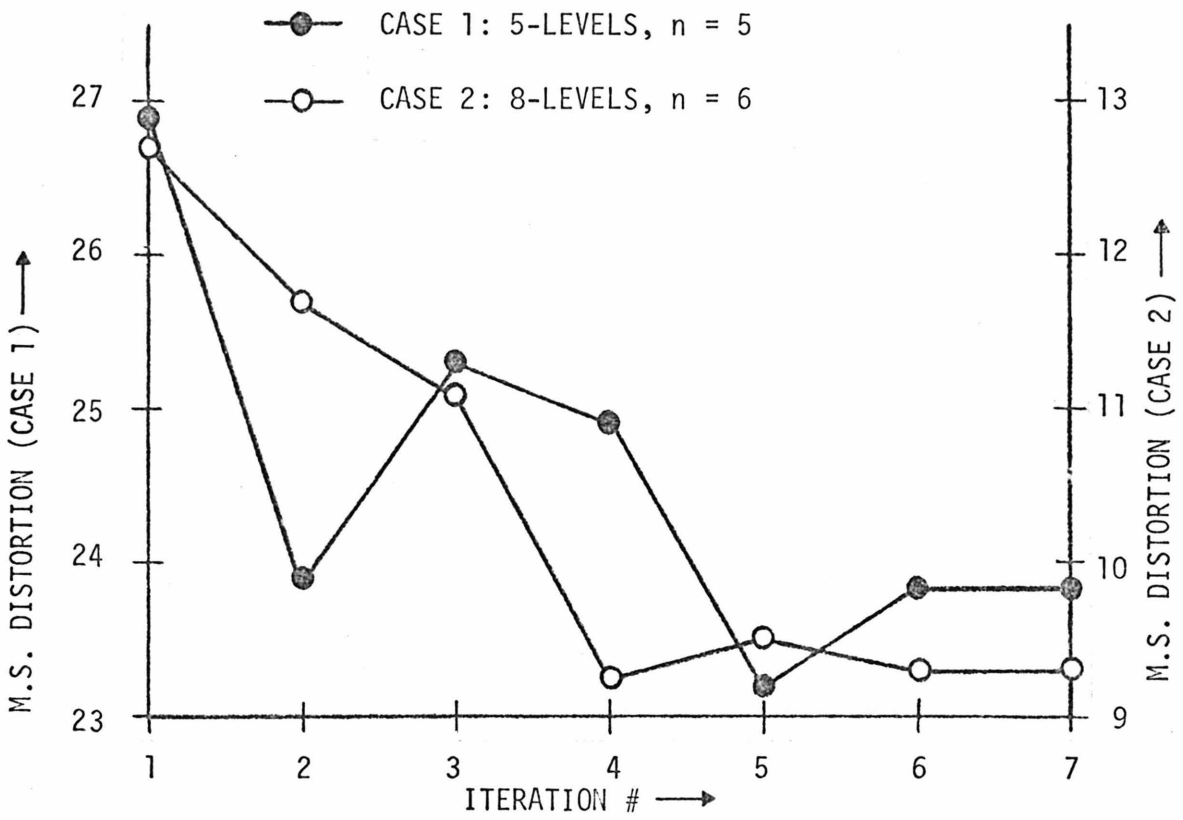


FIG. 3-2 CONVERGENCE OF LOOP ITERATIONS IN QUANTIZER DESIGN FOR KAREN (PRED #23)

represents M.S. distortion for ease of presentation, although $n = 5$ and 6 was used to design the quantizers. In case 2 with 8 levels, almost monotone convergence is obtained in six iterations, i.e., the quantizers obtained in the sixth and the seventh iterations are the same. In case 1, however, an oscillatory convergence is obtained in six iterations, but we pick the fifth quantizer since it has lower distortion value than the sixth one. In both the cases, as expected, there is marked reduction in the distortion, while the entropy of the quantized prediction errors, e_j's, increases by a negliglble amount (about .05 bits/pixel). This is accompanied with an improvement in the picture quality--the overall picture looks smoother. Case 1 is typical of quantizer design with the number of levels $N \leq 5$ while case 2 is typical of $N \geq 6$.

3.2 Discussion and Results

As "n" in the distortion criterion is increased, we observe significant effects on the quality of the encoded picture and its entropy. The main reason behind both these effects is, when we increase the value of "n", the large errors are being weighted much more than the small ones in evaluating the distortion D . As a result, the quantizer that minimizes D has representative levels that are spread out. This leads to the spreading out of the input thresholds also, because of Eq. I. This is shown in Fig. 3-3A with 9-level quantizers for KAREN using Pred #23. Similar trends are observed for the other originals also.

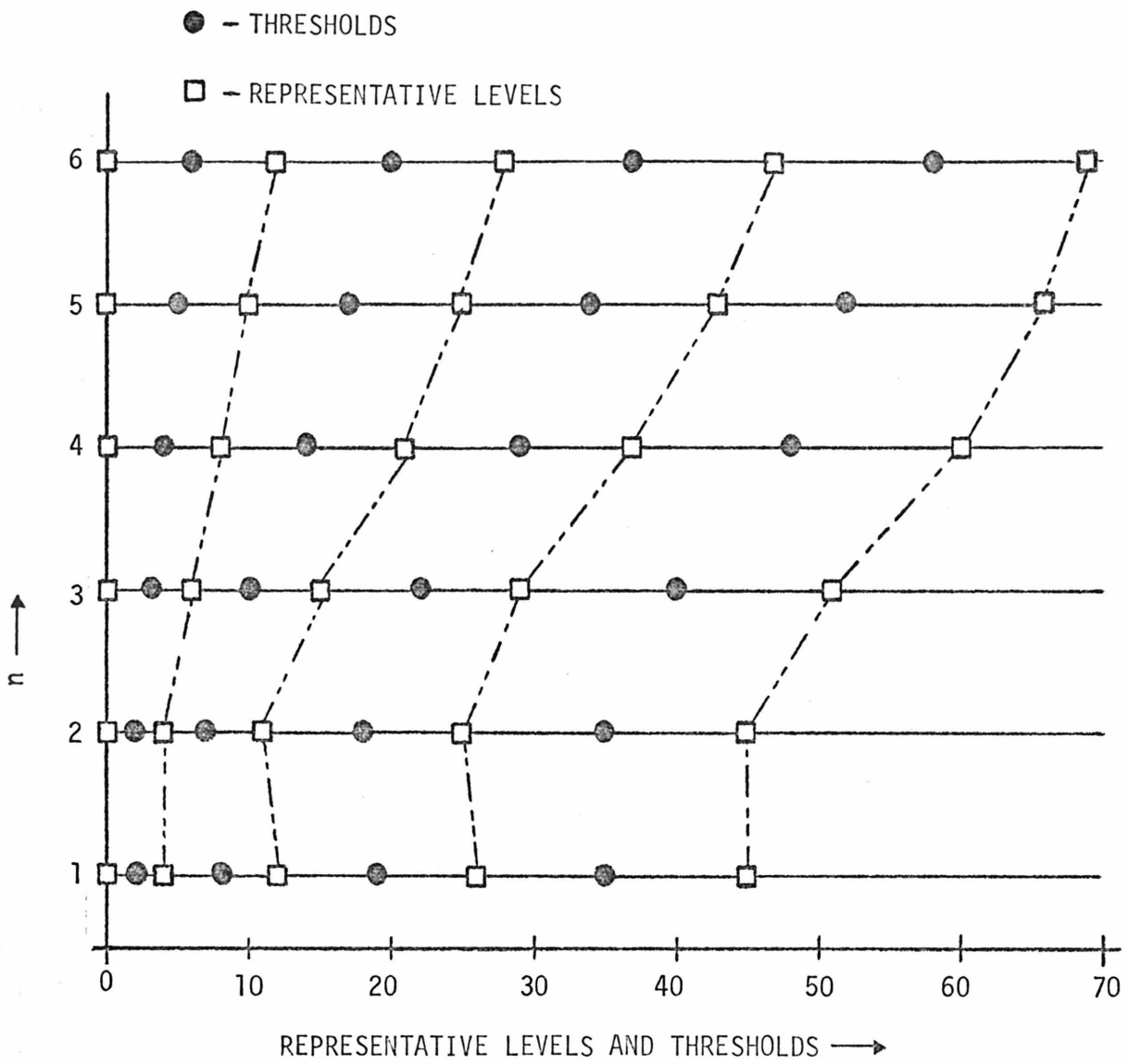


FIG. 3-3A 9-LEVEL QUANTIZERS FOR KAREN USING PRED #23 AND "MEAN n-TH POWER" DISTORTION CRITERIA

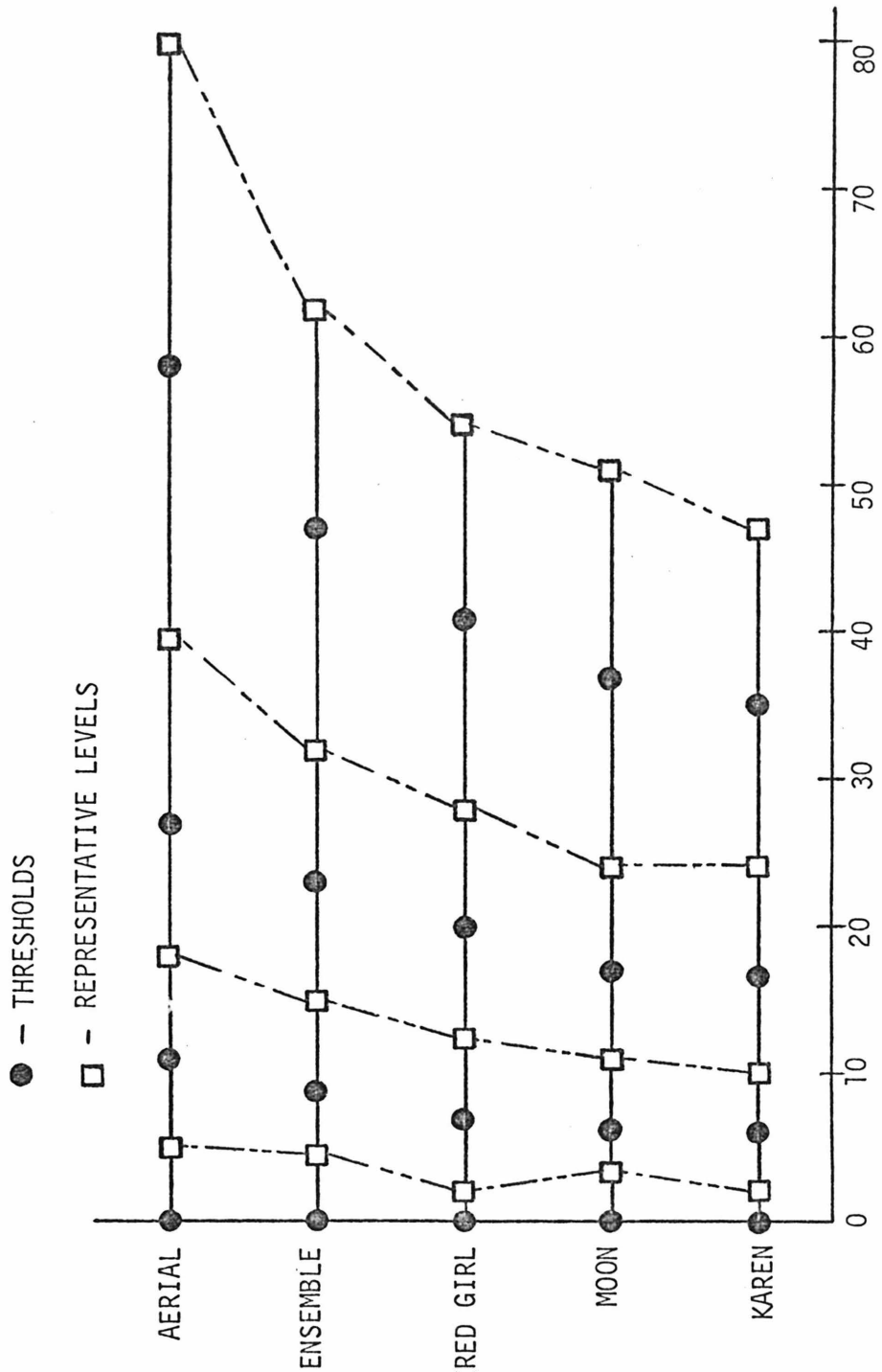


FIG. 3-3B. 8-LEVEL QUANTIZERS USING MEAN 3RD POWER DISTORTION CRITERION

The coarse quantization resulting from the use of large "n" has two major effects on the picture quality. Because of large outer levels of the quantizer, reproduction of relatively sharp edges is improved. On the other hand, because of large inner levels, artifacts begin to show up as granular noise in the smooth areas of the picture--the encoder is trying to track the slowly varying signal by taking large step jumps. We illustrate this by using 4-level quantizer and Pred #23 to encode KAREN. See PHOTO #16 and 17, also refer to the PHOTO-KEY. $n = 1$ was used in #16; sharp edges are poorly defined while little degradation is bserved in smooth areas like the hair region. In #17 we use $n = 4$. Note the remarkable improvement in edge reproduction, but the hair, cheek and the neck regions appear a little more granular than before. Overall, #17 is much more acceptable than #16.

Second effect of increased "n" is reduction in the entropy brought about by spreading out of the input thresholds. To see this, let us define

$$p_k = p(\ell_k) = \sum_{x_k < x \leq x_{k+1}} p(x) \quad ,$$

which is the probability of usage of the k^{th} level ℓ_k of the quantizer. See the shaded area in Fig. 3-4A. Then entropy of the encoded picture is

$$H = - \sum_{k=1}^{N} p_k \log p_k \quad .$$

Since the probability function $p(x)$ has a large peak near the origin,

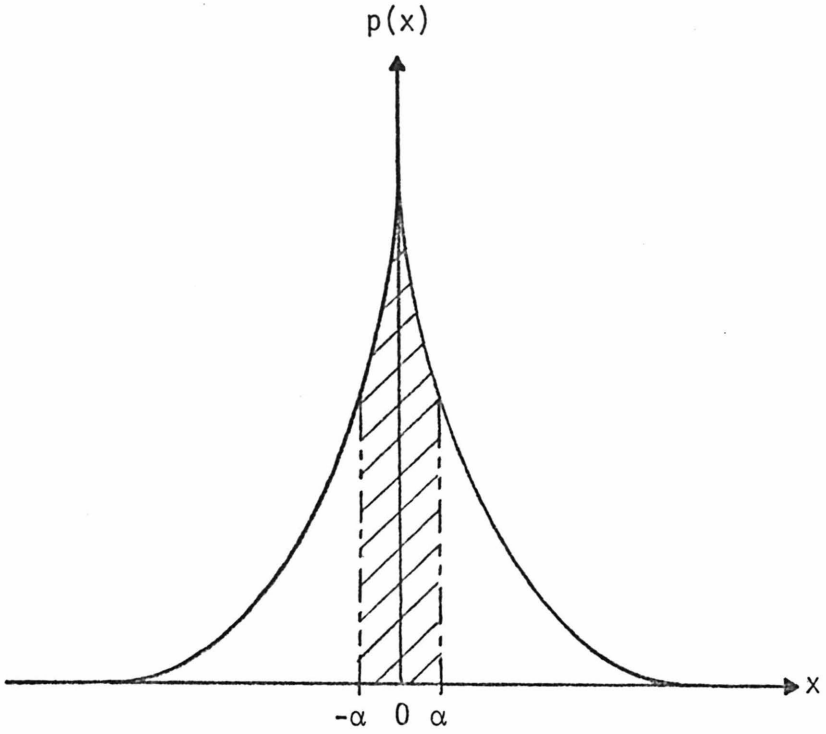


FIG. 3-4A

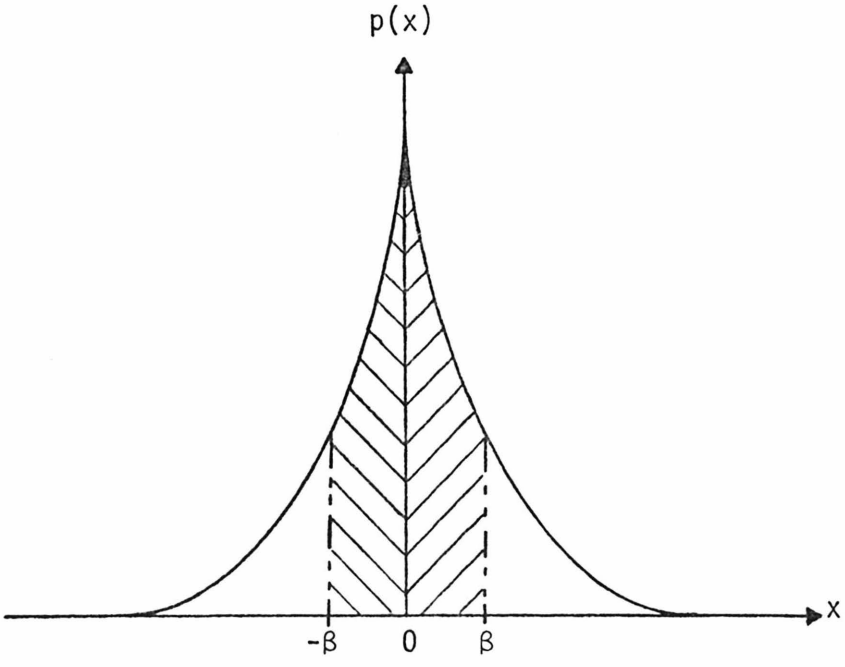


FIG. 3-4B

spreading out of the inner thresholds causes their corresponding p_k's to increase significantly. This makes the distribution $p(\ell_k)$ more peaked than before and thus reduces the value of H. This effect is observed for all the originals and is shown for KAREN in Fig. 3-5A. To study this figure pick some number of quantization levels and observe the thick dots moving down as "n" increases.

At this point we wish to draw the reader's attention to PHOTO #'s 9 and 11. In #9 we used $n = 1$, while in #11 we used $n = 5$. The entropy has gone down from 2.482 to 1.619, while there is only small degradation in the picture quality. Degradation in the hair area is the most prominent one. We rate PHOTO #'s 9 and 11 as 10 and 8 respectively on a scale from 0 to 10. Figure 3-5B shows this rating index for the other cases. For $n = 1$ deterioration in the picture quality is maximum as we reduce the number of levels in the quantizer, while for $n = 4$ this effect is minimum. This means that $n = 4$ is the most desirable value if we wish to reduce the entropy by dropping levels in the quantizer.

As we drop the levels we obtain quantizers with odd and even levels consecutively. What happens to the entropy then? Let us consider a symmetric quantizer with 2m+1 levels whose central threshold range $(-\alpha,\alpha)$ is shown in Fig. 3-4A, and another symmetric quantizer with 2m levels whose two central threshold ranges $(-\beta,0)$ and $(0,\beta)$ are shown in Fig. 3-4B. Due to large peak in $p(x)$ near the origin, the distribution $p(\ell_k)$ for the first quantizer will have one large peak in the middle corresponding to the shaded area in Fig. 3-4A. On the other hand, $p(\ell_k)$ for the second quantizer will have two large central

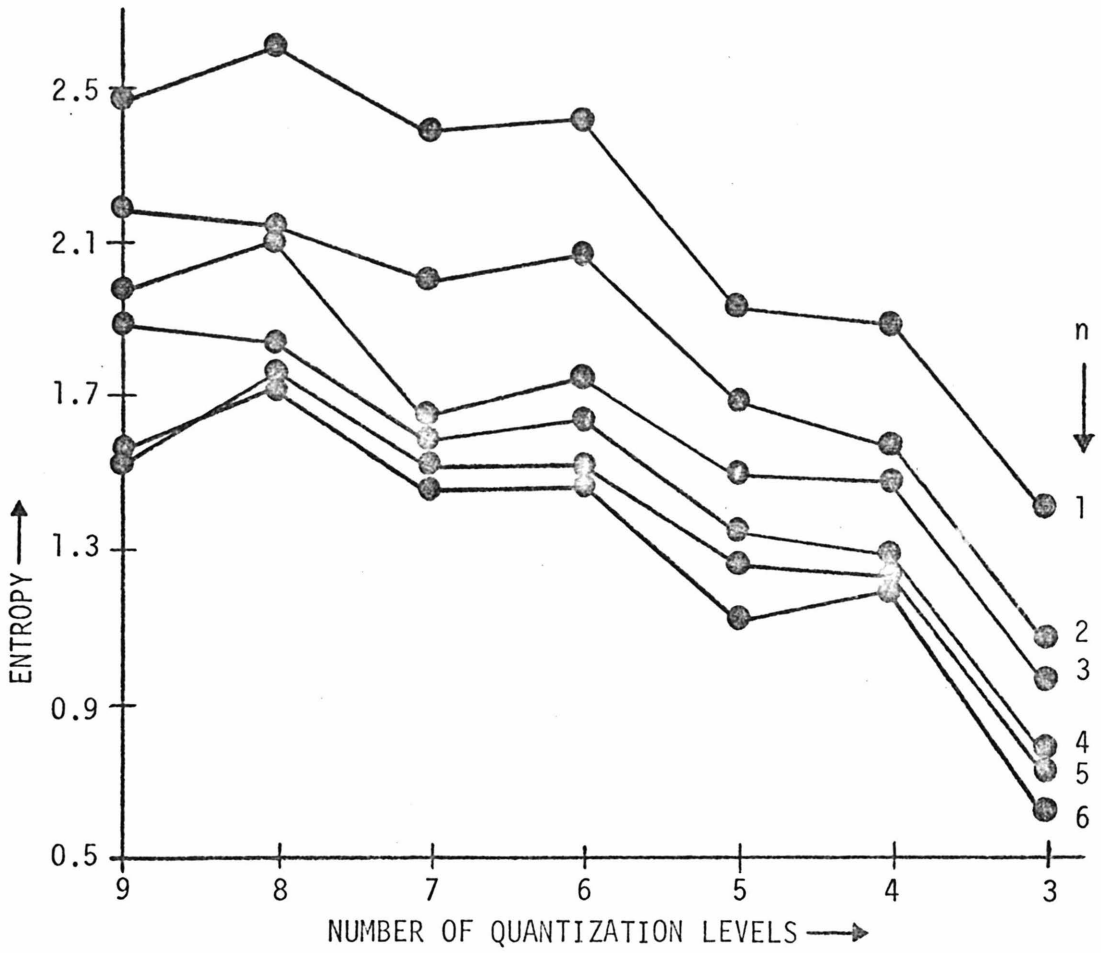
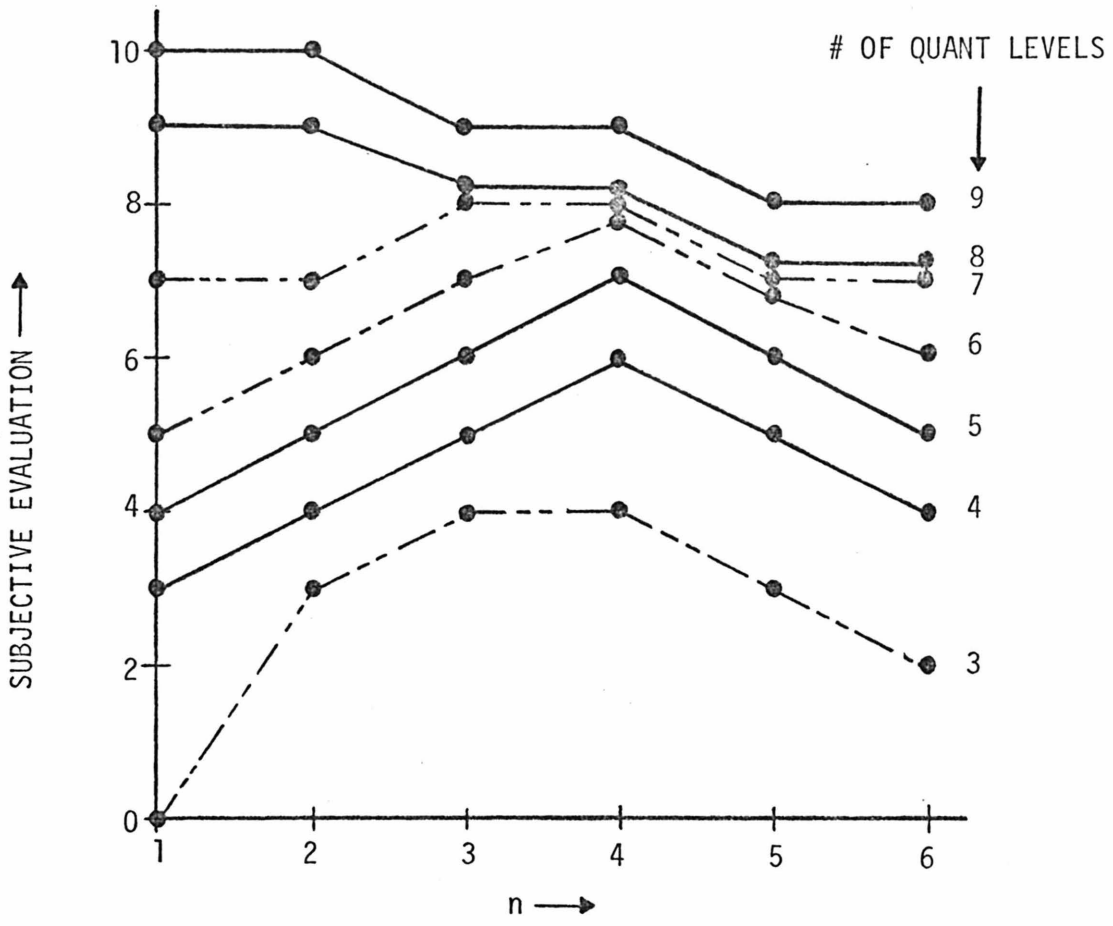


FIG. 3-5A. ENTROPY OF KAREN ENCODED WITH PRED #23 versus NUMBER OF QUANTIZATION LEVELS



NOTE: ORDINATES OF THICK DOTS TOUCHING EACH OTHER ARE THE NEAREST INTEGER VALUES

FIG. 3-5B. SUBJECTIVE EVALUATION OF KAREN ENCODED WITH PRED #23

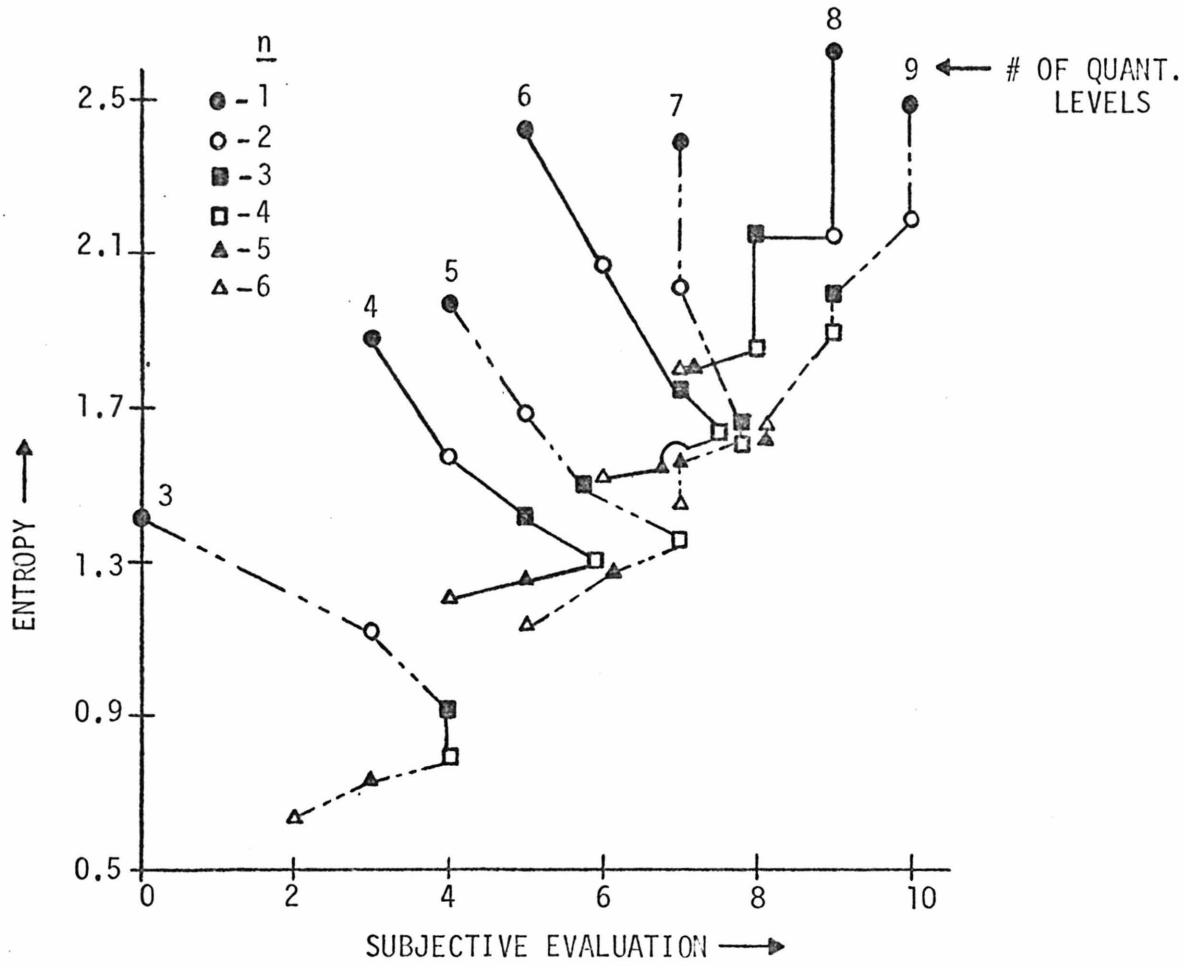


FIG. 3-5C. ENTROPY versus SUBJECTIVE QUALITY OF KAREN ENCODED WITH PRED #23

peaks of almost equal magnitude corresponding to the two shaded areas in Fig. 3-4B. This factor contributes toward making the second entropy higher than the first one. This trend is indeed observed in Fig. 3-5A, thus making the (2m+1)-level quantizer more desirable than the 2m-level one. This is not only because it gives lower entropy, but also because for fixed "n" the picture encoded with 2m+1 levels is of better quality than the one with 2m levels as indicated in Fig. 3-5B.

A moment ago we mentioned that the most desirable value of n is 4 from the point of view of dropping quantization levels to reduce the entropy. This value is the best from another point of view also, portrayed in Fig. 3-5C which is a plot of the entropy versus the subjective evaluation of the encoded pictures. For a fixed number of levels, if the value of n is made larger than 4 (shown by hollow squares), the resulting drop in entropy is associated with a sharp drop in subjective quality.

To summarize our discussion above, we find that $n = 4$ with odd number of quantization levels is the most suitable combination to encode KAREN. For other originals, we took some test cases and found that these conclusions are valid for them also. Figure 3-3B shows the quantizers for various originals using Pred #24. Notice the difference in the levels for different pictures. The ENSEMBLE quantizer was designed by taking an average of all the other four probability functions. The resulting pictures have been evaluated as in Table 3-1. Deterioration is most severe in KAREN and the encoded picture

TABLE 3-1

SUBJECTIVE EVALUATION OF PICTURES ENCODED WITH ENSEMBLE QUANTIZER OF FIG. 3-3B

PIX	ENSEMBLE QUANT. ENCODING	INDIVIDUAL QUANT. ENCODING
AERIAL	6	8
CHECK	7	9
MOON	5	8
KAREN	4	8

is subjectively unacceptable. This should be contrasted with the result obtained using adaptive encoding in Chapter 4.

3.3 Predictors Revisited

In this section we shall see what effects the predictors have on the encoded picture. Let us look at PHOTO #'s 12 and 13 which use Pred #10 and compare them with PHOTO #'s 9 and 11 which use Pred #23. Notice poor edge reproduction and general graininess in the former ones. In this case both the MAX and the VAR of Pred #23 are smaller than for #10. To see the usefulness of MAX criterion we pick Pred #22 and 24 for CHECK, because

$$\text{MAX (Pred \#24)} < \text{MAX (Pred \#22)}$$

while,

$$\text{VAR (Pred \#24)} > \text{VAR (Pred \#22)} \quad .$$

Results of encoding in the two cases with $n = 4$ are shown in PHOTO #'s 18 and 19 (also, see the PHOTO-KEY). The latter picture uses Pred #22 (which has low variance), has more contouring in the left-hand side background, and has more granular noise in the neck and the shirt regions. This is so because in the case of large MAX, spread of the probability distribution $p(x)$ is larger hence the quantizer levels get spread out as large errors receive more weight than the small ones for $n = 4$. Such is the case for all $n > 2$ as shown in Fig. 3-6. Similar results are obtained for KAREN with Pred #22 and 23, where

$$\text{MAX (Pred \#23)} < \text{MAX (Pred \#22)}$$

and

$$\text{VAR (Pred \#23)} \simeq \text{VAR (Pred \#22)} \ ,$$

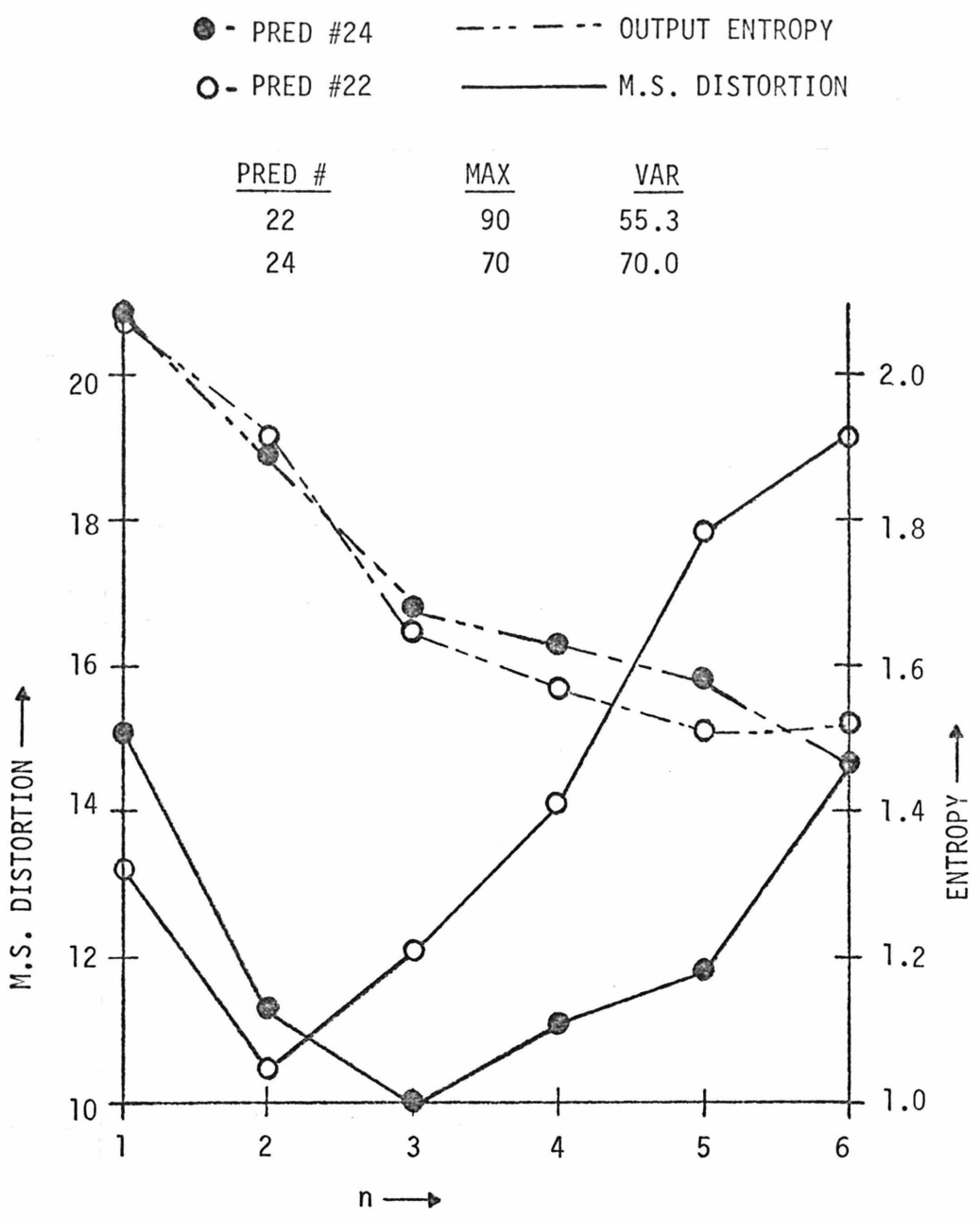


FIG. 3-6 ENCODING OF CHECKERED GIRL USING PRED #22 and 24 AND 6 LEVEL QUANTIZER

see Figs. 2-2 and 2-3. Thus we observe for more interesting values of n (i.e., the large ones), the MAX criterion is more useful as an aggregate measure of prediction errors than the VAR.

3.4 Huffman Codes and Higher Order Entropies

The first order entropy defined in Sect. 3.2 serves as only the lower bound on the information/pixel at which we can transmit the pictures. As a matter of practical interest we designed the variable length Huffman codes for the encoded pictures with the source alphabet consisting of ℓ_k , $1 \leq k \leq N$, the N levels of the quantizer. The average word length of this binary code comes to within 0.05 - 0.1 of the entropy per pixel for all the pictures.

Now we wish to consider the role of second order entropy as a means of further bit saving. For simplicity of the argument let us consider the last element predictor (i.e., #10) encoding. Let W and X be two nearest horizontal neighbors of Z , and let V and Y be its two nearest vertical neighbors as shown in Fig. 3-7. Also, let the pixels X, Y, etc. be encoded using the $\ell_x{}^{th}$, $\ell_y{}^{th}$ levels of the quantizer, etc. Then $H(\ell_z|\ell_x)$, the second order entropy of the encoded picture given the previous element in the same line as Z, and $H(\ell_z|\ell_y)$, the same quantity given the previous element in the same column are expected to bear the following relation

$$H(\ell_z|\ell_x) > H(\ell_z|\ell_y) \quad .$$

The reasoning behind this is outlined in the following steps.

PRED #10, n=1, 9-LEVEL QUANTIZER

Picture	$H(\ell_z)$*	$H(\ell_z\|\ell_y)$	$H(\ell_z\|\ell_x)$	$H(\ell_z\|\ell_x,\ell_y)$	$H(\ell_z\|\ell_x,\ell_w)$	$H(\ell_z\|\ell_y,\ell_v)$	$H(\ell_z\|\ell_x,\ell_u)$
KAREN	2.482	1.842	2.234	1.764	2.101	1.775	1.957
CHECK.GIRL	2.212	1.641	1.942	1.532	1.840	1.595	1.735
MOON	2.776	2.589	2.678	2.499	2.615	2.577	2.695
AERIAL	2.847	2.560	2.737	2.483	2.661	2.527	2.638

*ℓ_zth level of the quantizer was used to encode the pixel Z, etc.

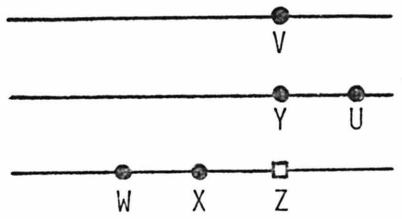


FIG. 3-7. FIRST, SECOND AND THIRD ORDER ENTROPIES OF ENCODED PICTURES

1. A second order entropy $H(\alpha|\beta)$ is small when the joint probability distribution $p(\alpha,\beta)$ has large peaks near the "main diagonal" (i.e., near $\alpha \simeq \beta$). This happens when the probability of the event $\alpha \simeq \beta$ is large.
2. $p(\ell_z|\ell_x)$ is expected to be less peaked near its main diagonal than $p(\ell_z|\ell_y)$ because
3. $\ell_z \simeq \ell_x$ implies that the picture is approximately a ramp through the pixels W, X, and Z, while
4. $\ell_z \simeq \ell_y$ only implies similarity between the slopes in the current and the preceding lines, and
5. the condition 3 is less likely to be met than the condition 4. Hence the relation between the two second-order entropies.

Similarly, we also expect

$$H(\ell_z|\ell_x,\ell_w) > H(\ell_z|\ell_y,\ell_v) > H(\ell_z|\ell_x,\ell_y) \quad .$$

Results of entropy measurements on various encoded pictures are shown in Fig. 3-7. The trend of all the results is the same as discussed above. Percentage reductions over $H(\ell_z)$, obtained by using $H(\ell_z|\ell_x)$ are 4, 4, 10, and 12% for MOON, AERIAL, KAREN, and CHECK respectively, while the corresponding figures for $H(\ell_z|\ell_y)$ are 7, 10, 26, and 26% respectively. Another observation of interest is that even $H(\ell_z|\ell_x,\ell_w)$ is larger than $H(\ell_z|\ell_y)$. $H(\ell_z|\ell_x,\ell_y)$ yields 12, 13, 30, 30% reduction in the four pictures in the same order as above. Similar trends are observed for other predictors also.

We also designed a Huffman code corresponding to $H(\ell_z|\ell_y)$. This was done by converting the two-dimensional source alphabet (ℓ_z,ℓ_y) corresponding to the probability distribution $p(\ell_z,\ell_y)$ to the one-dimensional source alphabet $(\ell_z-\ell_y)$ whose probability distribution $p(\ell_z-\ell_y)$ is used to design the code. The average word length of this code comes to within about 0.1 bits/sample of $H(\ell_z|\ell_y)$.

CHAPTER 4

ADAPTIVE ENCODING

In this chapter we consider schemes that make use of variations in the local structural properties of pictures, and thus encode them more efficiently. Such encoding can be done by dividing the picture in areas of low, medium and high details based upon the information available at the receiver end only. Encoding is done by both adaptive prediction and quantization based upon the three areas. Finally, we consider, using prediction error statistics in the three areas of all four originals, an ensemble quantizer that could be used to encode a class of pictures.

4.1 Segmentation

To predict the value of a pixel X depending on its local surroundings we first have to decide about the nature of these surroundings. This can be done by looking at the pixels that are in the vicinity of X in the picture. In order that we do not use additional bits to transmit information as to what area X belongs to, we must consider only the pixels occurring previous to X so that the same segmentation process can be carried out by the decoder also. Furthermore, in order to reduce the probability of getting the "segmenter" at the decoding end confused by spurious noise signals, we should use a number of neighboring pixels in the segmentation algorithm. Being led by these considerations, let us consider the neighborhood A(X) consisting of 22 of X's previous neighbors, as shown in Fig. 4-1. We wish to classify X in one of the three areas of low, medium and high details, called

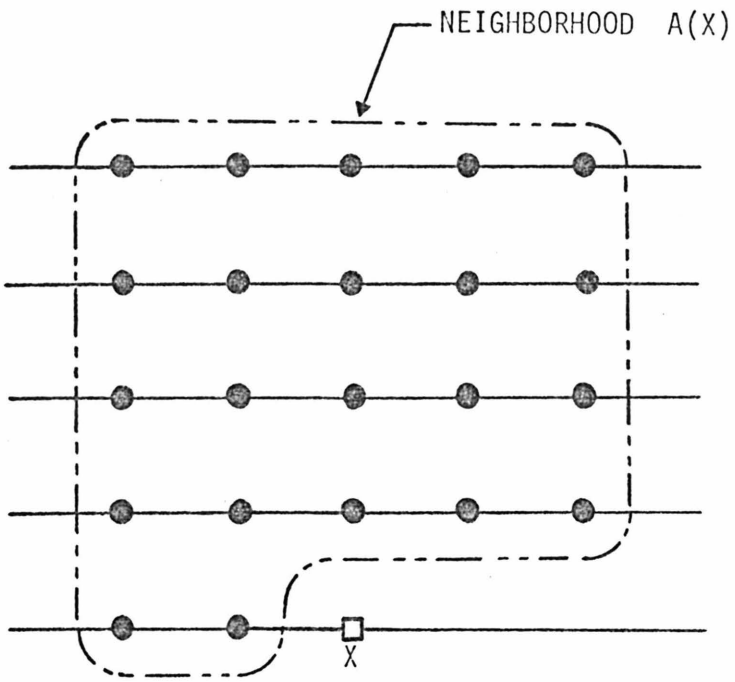


FIG. 4-1. ADAPTIVE PREDICTION OF THE PIXEL X

area 1, 2, and 3 respectively. For this, let us define $N(X)$ as the number of pixels in X's neighborhood $A(X)$, that differ from their immediately preceding vertical or horizontal neighbors by more than a given threshold T. For properly chosen T, this quantity $N(X)$ can serve as a measure of "activity" in the picture near the pixel X. Since there are only 22 pixels in $A(X)$ we observe

$$0 \leq N(X) \leq 22 \quad .$$

Next, we choose two integers n_1 and n_2 between 0 and 22, and classify X as belonging to area 1, 2, or 3 depending on whether the value of $N(X)$ lies in the interval $[0,n_1]$, $[n_1+1,n_2]$ or $[n_1+1,22]$, respectively.

By computer simulation of the above algorithm, we determined that the following set of values give satisfactory results:

$$\text{Threshold } T = 7 \quad ,$$
$$n_1 = 5 \quad ,$$
$$n_2 = 13 \quad .$$

The three areas in KAREN are shown in PHOTO #23 which has three brightness levels. White corresponds to the high details, while grey and black correspond to medium and low detail areas respectively. In Table 4-1 we show the fractional values of the three areas in various pictures. Note that MOON is predominantly a medium detail picture while AERIAL is a high detail one.

In the next step of our segmentation scheme we take area 3 which is predominant in high details, sharp edges, etc. and divide it in horizontal or vertical-edge areas. This is done by considering

TABLE 4-1

FRACTIONAL VALUES OF LOW, MEDIUM, AND HIGH DETAIL AREAS IN PICTURES

PIX	LOW DETAIL	MED DETAIL	HIGH DETAIL
MOON	0.04	0.71	0.25
CHECK	0.48	0.30	0.22
KAREN	0.30	0.40	0.30
AERIAL	0.06	0.33	0.61

$H(X)$ and $V(X)$, the sums of vertical and horizontal pixel differences in the neighborhood $A(X)$ of X. The pixel X is classified in

$$\text{Horizontal-edge-area} \quad \text{if} \quad H(X) > V(X)$$

or

$$\text{Vertical-edge-area} \quad \text{if} \quad H(X) \leq V(X) \ .$$

In PHOTO #24 we "paint" these areas over the original as white and black respectively.

Still further subdivision can be made by considering the pixels for which

$$\left|\frac{H(X) - V(X)}{H(X) + V(X)}\right| \leq 10\%$$

and grouping them in +45° or -45° slope areas. We present these in PHOTO #25 in dark black and white shades respectively, superimposed over the original. Since the total fraction of such pixels is very small, we shall not use such a fine subdivision in our encoding schemes.

4.2 Encoding

Similar to the nonadaptive case, the first step in designing an adaptive encoding scheme would be to investigate effectiveness of the various predictors in the three areas. In area 3, however, we have one more predictor available. In this predictor, referred to as (8,9), in the vertical edge area prediction is done by

$$\text{Pred \#8:} \quad X = (3E-B)/2 \ ,$$

while in the horizontal edge area we use

$$\text{Pred \#9:} \quad X = (3I-H)/2 \ .$$

As was done in Chapter 2, based on the original pictures we measure MAX and VAR of the prediction errors in the three areas. The average of these quantities, taken over the four originals has been plotted in Figs. 4-2, 4-3, and 4-4 for areas 1, 2, and 3 respectively. From these figures, as shown in Table 4-2, we determine Pred #'s 5, 24 and (8,9) to be the best for the three areas respectively. For these predictors the values of MAX and VAR for KAREN have been tabulated in Table 4-3. Notice a reduction in both the values as compared to the nonadaptive case (Figs. 2-2 and 2-3). The prediction error picture and the corresponding histograms for the three areas have been presented in PHOTO #'s 26-29. See also the PHOTO-KEY.

Next, for each area, we choose the value of the power "n" in the distortion criteria and ddesign the quantizer using prediction error statistics of that particular area using algorithm A1 of Chapter 3. Now the encoder has three predictors and three quantizers. Thus, to encode the pixel X, it uses the available reconstructed picture to decide which area X belongs to, and then it uses the appropriate predictor and the quantizer to encode the pixel. From here on we shall denote the adaptive predictor by Pred # $\alpha|\beta|\gamma$ where Pred #'s α, β and γ are used in areas 1, 2, and 3 respectively. We shall adopt similar notation for the number of quantizer levels and the value of "n" also.

By doing adaptive encoding, since we are making use of the local variations in the picture, we hope to improve the picture quality and we also hope to reduce the entropy by treating the three areas of the picture differently. To see how much improvement in the picture quality is achieved by quantizing the three error signals differently,

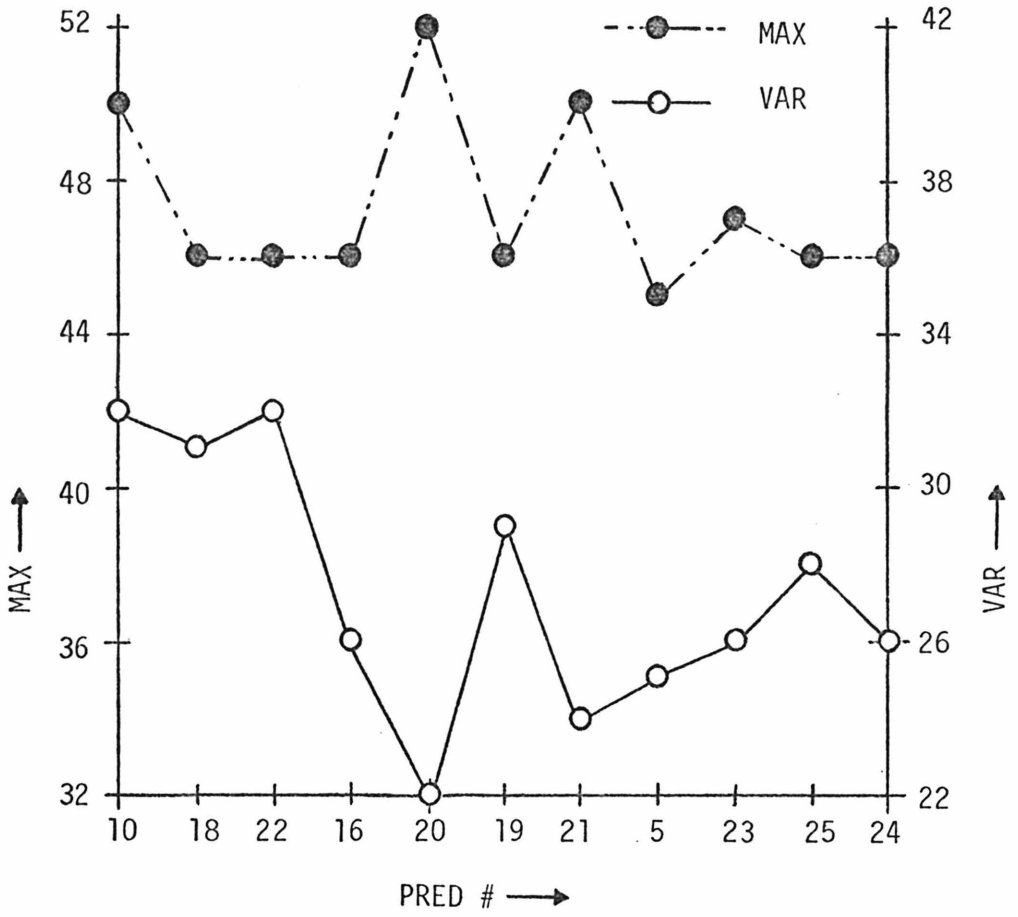


FIG. 4-2 MAX and VAR AVERAGED OVER FOUR ORIGINALS IN AREA 1 versus PRED #

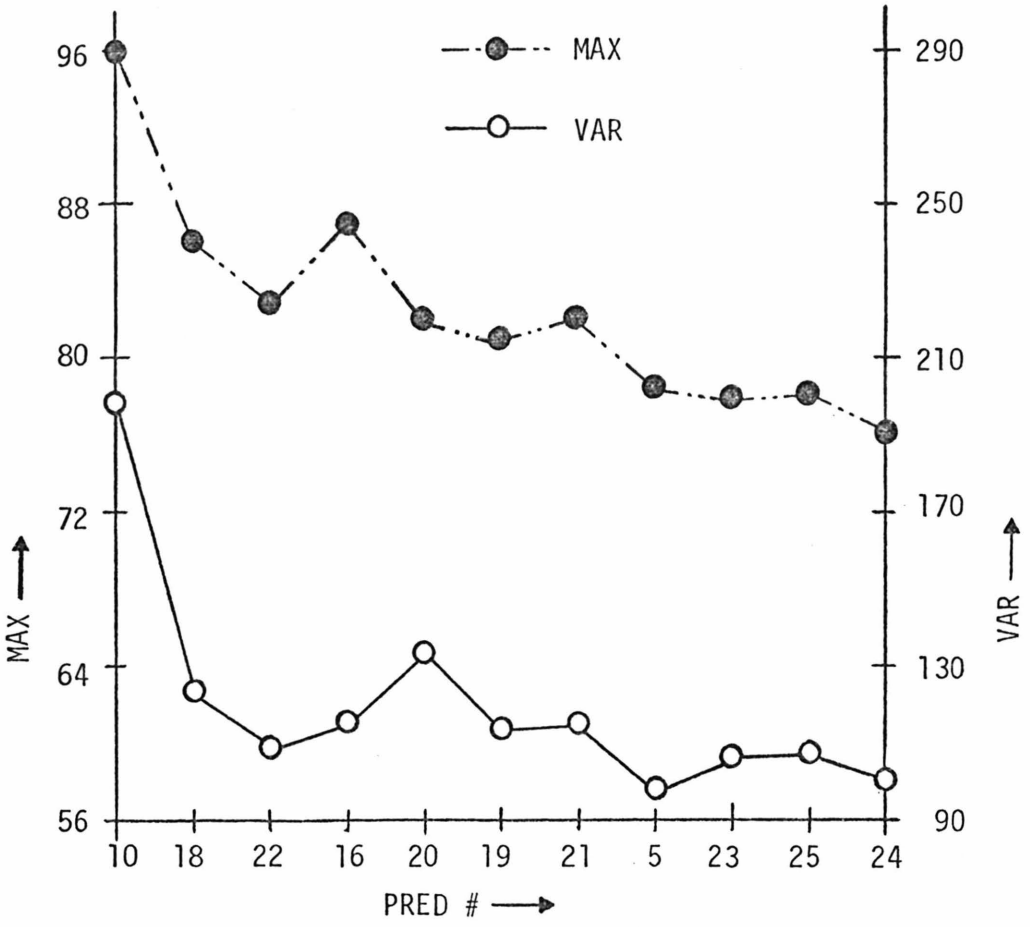


FIG. 4-3 MAX AND VAR AVERAGED OVER FOUR ORIGINALS IN AREA 2 versus PRED #

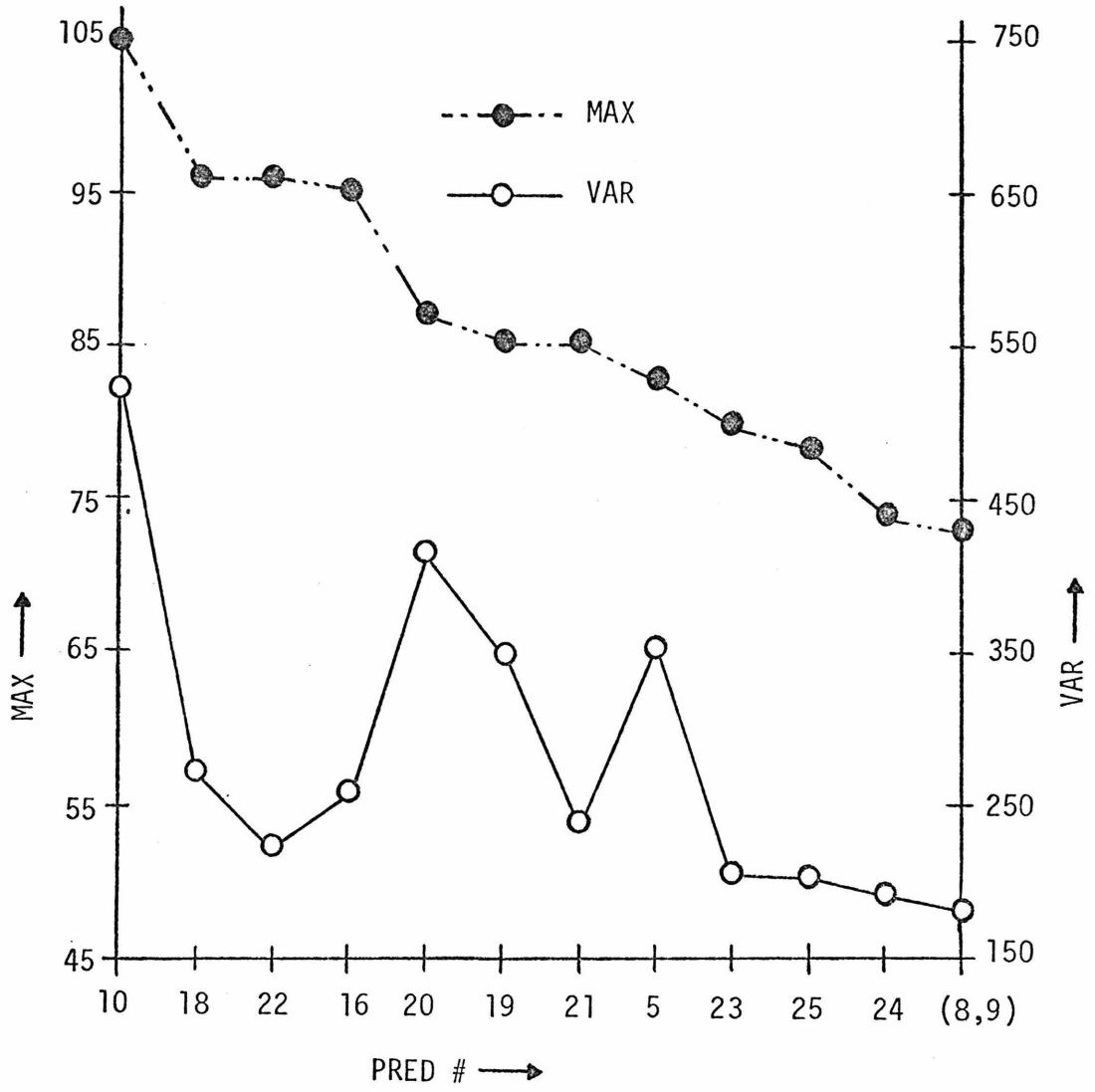


FIG. 4-4 MAX AND VAR AVERAGED OVER FOUR ORIGINALS IN AREA 3 versus PRED #

TABLE 4-2

CHOICE OF PREDICTORS BASED ON ENSEMBLE AVERAGES

FOR ADAPTIVE PREDICTION

AREA	PRED #'s
1	25,23,16,24,5*
2	21,25,5,23,24
3	21,23,25,24,(8,9)

*In order of increasing preference

TABLE 4-3

ADAPTIVE PREDICTION ON KAREN--MAX AND VAR OF PRED. ERRORS

AREA	PRED #	MAX	VAR
1	5	29	11
2	24	64	41
3	(8,9)	67	140
Total		67	60

we used Pred # 23|23|23 and 9|9|9 level quantizer with n = 6|6|6 to encode KAREN (See Tables 4-4 and 4-5). The quality of this picture (rated 9) is better than the one obtained by nonadaptive encoding with Pred #23, 9-levels and n = 6 (rated 8). The main difference is in smooth rendition of the hair region which was quite "grainy" in the nonadaptive case. However, the entropy has increased by about 0.19 bits/pixel. Now to reduce the entropy we have two alternatives available to us (1) we can reduce the number of levels in the three quantizers, or (2) we can use still higher values of n to design the quantizers. It is clear from the discussion in Chapter 3 that the second approach has already reached a state of diminishing returns as we are using "n" as high as 6. Therefore, we turn to the first alternative and use 9|9|7 level quantizers. The entropy does reduce to 1.751 but the picture looks grainy near the eye region due to coarse quantization. This forces us to retain more than 7 levels in area 3 but we can drop two levels each in areas 1 and 2, obtaining 7|7|9 level quantizers. Entropy goes down to 1.767 while there is no visible deterioration in the picture. Then we introduce the Pred # 5|24|24 followed by # 5|24|(8,9). With the latter one we obtain a 1.683 bit/pixel picture without losing any additional picture quality. This is due to low values of MAX and VAR with Pred # (8,9) in area 3. Now we reduce the number of quantizer levels to 4|7|8 and 4|6|8 to obtain 1.514 and 1.452 bits/pixel pictures, respectively. The quantizers are given in Table 4-4, while the encoded pictures are shown in PHOTO #'s 36 and 37. In both these pictures, the hair and the edge reproduction is quite good, while the cheek region is somewhat granular.

TABLE 4-4

Representative Levels* of Quantizers for Adaptive Encoding with Pred # 5/24/(8,9) and n = 6/6/6

A. Area 1

Pict.	# of Levels	Representative Levels
KAREN	4	3, 16
CHECK	4	4, 33
MOON	4	5, 23
AERIAL	4	6, 38
ENSEMBLE	4	4, 31

B. Area 2

KAREN	6	4, 25, 50
CHECK	6	6, 26, 58
MOON	6	6, 26, 57
AERIAL	6	9, 38, 80
ENSEMBLE	6	6, 29, 74

C. Area 3

KAREN	8	7, 29, 58, 96
CHECK	8	4, 16, 34, 59
MOON	8	6, 21, 45, 80
AERIAL	8	9, 33, 63, 103
ENSEMBLE	8	7, 26, 52, 92

*In view of Eq. I, Sect. 3.1, we omit the input thresholds. Also, because of symmetry, only positive half of the quantizer is presented.

TABLE 4-5

ENCODING KAREN

n = 6/6/6 IN AREAS 1/2/3

PRED #'s	# OF QUANT LEVELS	ENTROPY				SUBJECTIVE EVALUATION
		AREA 1	AREA 2	AREA 3	TOTAL	
23/23/23	9/9/9	1.667	1.714	2.208	1.845	9
"	9/9/7	1.659	1.721	1.887	1.751	6
"	7/7/9	1.639	1.546	2.202	1.767	9
5/24/24	"	1.632	1.540	2.200	1.760	9
5/24/(8,9)	"	1.640	1.544	1.917	1.683	9
"	4/7/8	1.173	1.562	1.799	1.514	8
"	4/6/8	1.183	1.400	1.801	1.452	8
"	4/6/8*	1.071	1.194	1.803	1.306	6

*Quantizer based on ensemble statistics

These pictures have been given a rating of 8.

Next, we take combined statistics of the prediction errors with Pred # 5|24|(8,9) from all the originals in the three areas. This is used to design the "ensemble quantizer" with 4|6|8 levels (Table 4-4). Such a quantizer could be used to encode any of the four pictures included in the ensemble. We used it to encode KAREN. The resulting picture as shown in PHOTO #38 has entropy of 1.3 bits/pixel. It appears more grainy in the face region than the previous two, while edge reproduction in it is quite good. Notice that in this case the subjective quality has dropped from 8 to 6 in going from the "individual" quantizer to the "ensemble" quantizer, while in the case of nonadaptive encoding the drop was from 8 to 4 (Table 3-1).

We summarize the above discussion by noting that the use of adaptive quantizer (with nonadaptive prediction) yields better picture quality than otherwise. Along with adaptive prediction, we can reduce the entropy of the encoded picture by reducing the number of levels in the quantizers without much deterioration in the picture quality. The number of levels could be reduced much more in the low detail areas than in the high detail ones. Furthermore, adaptive encoding schemes enable us to design "ensemble" quantizers that may be used to encode a given class of pictures such that the encoded pictures are subjectively acceptable.

REFERENCES

1. Oliver, B. N., Pierce, J. R., Shannon, C. E., "The Philosophy of Pulse Code Modulation", Proc. IRE 36, Nov. 1948, pp. 1324-1331.

2. Shannon, C. E., "A Mathematical Theory of Communication", BSTJ, July and Oct. 1948; Shannon and Weaver, The Mathematical Theory of Communication, University of Illinois Press, 1949.

3. Oliver, B. N., "Efficient Encoding", BSTJ 31, July 1952, pp. 724-750.

4. Graham, R. E., "Communication Theory Applied to TV Coding", Acta Electronica 2, 1957-58.

5. Schreiber, W. F., "The Measurement of Third Order Probability Distributions of TV Signals", Technicolor Corporation, Research Department, Burbank, California.

6. Cutler, C. C., "Differential Quantization of Communication Signals", Patent No. 2,605,361, July 29, 1952 (Applied for June 29, 1950).

7. Harrison, C. W., "Experiments with Linear Prediction in Television", BSTJ 31, July 1952, pp. 764-783.

8. Elias, P., "Predictive Coding", IRE Trans. Inform. Theory IT-1, March 1955, pp. 16-33.

9. Graham, R. E., "Predictive Quantizing of TV Signals," IRE Wescon Convention Record, Part 4, August 1958, pp. 147-156.

10. O'Neal, J. B., Jr., "Predictive Quantizing Systems (Differential Pulse Code Modulation) for the Transmission of Television Signals", BSTJ 45, No. 5 (May-June 1966), pp. 689.721.

11. Connor, D. J., Pease, R.F.W., and Scholes, W. G., "Television Coding Using Two-Dimensional Spatial Prediction", BSTJ 50, March 1971, pp. 1049-1061.

12. Kummerow, T., "Statistics for Efficient Linear and Non-Linear Picture Encoding", Proc. 1972 Intl. Telemetering Conf. 8, pp. 149-161.

13. Thoma, W., "Video Transmission Network with Intraframe DPCM and Optional Interframe Coding", Proc. 1972 Intl. Conf. on Communications 39, pp. 1-5.

14. Connor, D. J., "Techniques for Reducing the Visibility of Transmission Errors in Digitally Encoded Video Signals", IEEE Trans. Commun. Tech. COM-21, No. 3 (June 1973).

15. Kretzmer, F. R., "Statistics of Television Signals", BSTJ 31, July 1952, pp. 751-763.

16. Panter, P. F. and Dite, W., "Quantizing Distortion in PCM with Nonuniform Spacing of Levels", Proc. IEEE 39, Jan. 1951, pp. 44-48.

17. Max, J., "Quantizing for Minimum Distortion", IRE Trans. Inform. Theory IT-6, 1960, pp. 7-12.

18. Bruce, J. D., Optimum Quantization, MIT Research Lab. of Electronics, Tech. Rep. No. 429, March 1, 1965.

19. Bernard, S., "Instantaneous Companding of Quantized Signals", BSTJ 36, No. 3, May 1957, pp. 653-709.

20. Limb, J. O. and Mounts, F. W., "Digital Differential Quantizer for Television", BSTJ 48, No. 7, Sept. 1969, pp. 2583-2599.

21. Limb, J. O., "Source-Receiver Encoding of Television Signals, Proc. IEEE 55, April 1967, pp. 364-379.

22. Candy, J. C. and Bosworth, R. H., "Methods for Designing Differential Quantizers Based on Subjective Evaluation of Edge Busyness", BSTJ 51, No. 7, Sept. 1972, pp. 1495-1516.

23. Limb, J. O., "Picture Coding: The Use of a Viewer Model in Source Encoding", BSTJ 52, No. 8, Oct. 1973, pp. 1271-1301.

24. Limb, J. O., "Entropy of Quantized Television Signals", Proc. IEE 115, No. 1, Jan. 1968, pp. 16-20.

25. Golding, L. S. and Schultheiss, P. M., "Study of an Adaptive Quantizer", Proc. IEEE 55 (1967), pp. 293-297.

26. Wintz, P. A. and Tasto, M., "Picture Bandwidth Compression by Adaptive Block Quantization", Purdue University School of Electrical Eng., TR-EE 70-14, July 1970.

27. Brown, E. F. and Kaminsky, W., "An Edge-Adaptive Three-Bit Ten-Level Differential PCM Coder for Television", IEEE Trans. Commun. Tech. COM-19, No. 6, Dec. 1971, pp. 944-947.

28. Limb, J. O., "Adaptive Encoding of Picture Signals" in Picture Bandwidth Compression, Huang, T. S., and Tretiak, O. J., eds. Gordon and Breach, Science Publishers, 1972.

29. Cohen, F., "A Switched Quantizer for Nonlinear Coding of Video-Signals", Nachrichtentech Z. 25, 1972, pp. 554-559.

30. Kummerow, T., "DPCM Systems with Two-Dimensional Predoctors and Controlled Quantizer", NTG-Tagung Signalverarbeitung Erlangen, 1973, pp. 425-439.

31. Papoulis, A., Probability, Random Variables and Stochastic Processes, McGraw-Hill Book Company, Inc., New York, 1965.

32. Knuth, D. E., The Art of Computer Programming, Vol. 3, Addison-Wesley Publishing Company, Reading, Mass., 1973.

APPENDIX A

PHOTO-KEY

CAPTION	PHOTO #	DESCRIPTION
A. NON-ADAPTIVE SCHEMES		
Originals	1	MOON SHOT
	2	CHECKERED GIRL
	3	KAREN
	4	AERIAL VIEW

		PRED #	MAX	VAR
Prediction Error	5	10	109	271
Pictures - KAREN	6	24	70	79
Pred. Error Histograms				
Magnified for KAREN (Vert. Axis--Count of Occurrences; Horizontal Axis--Mag. of Error--32 per div.)	7	Pred #10 (Accompanies Photo #5)		
	8	Pred #24 (Accompanies Photo #6)		

		Pred #	n	ENT (bits/pixel)
	9	23	1	2.482
	10	23	3	2.130
KAREN Encoded	11	23	5	1.619
with 9-level quantizer	12	10	1	2.493
	13	10	3	2.077

CAPTION	PHOTO #	n	ENT	DESCRIPTION Pred #
CHECK GIRL Encoded	14	1	2.534	22
with 8-level quantizer	15	6	1.651	22

Effect of n on Edge Reproduction		n	ENT	Subjective Quality Index
KAREN Encoded with Pred #23 and	16	1	1.885	3
4-Level Quantizer	17	4	1.296	6 (Max value 10)

Effectiveness of MAX Criterion		Pred #	MAX	VAR	ENT	M.S. DIST
CHECK Encoded with n = 4 and 6-Level	18	24	70	70	1.631	11.091
Quantizer	19	22	90	55	1.577	14.193
KAREN Encoded with n = 4 and 6-Level	20	22	100	80	1.626	21.675
Quantizer	21	23	74	79	1.631	12.655
KAREN--Pred. Error Histogram for Pred #22	22	(Corresponds to PHOTO #20)				

B. ADAPTIVE SCHEMES

Segmentation in Low, Med. and High Detail Areas--KAREN — 23

Threshold = 7

Area Detail	Represented As	Count Range
Low	Black	0 - 5
Med.	Grey	6 - 13
High	White	14 - 22

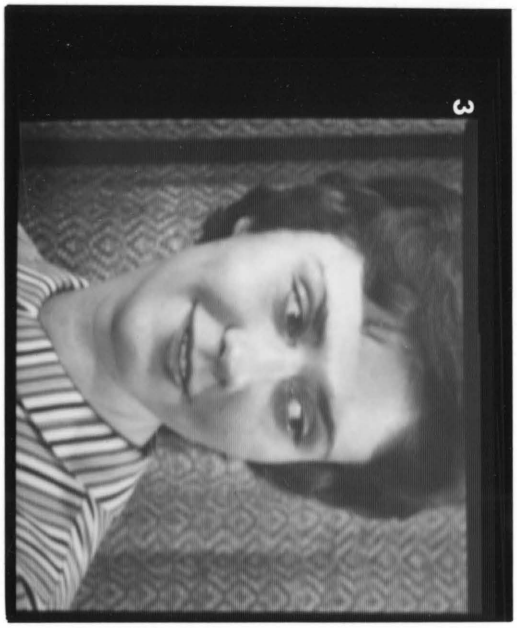
CAPTION	PHOTO #	DESCRIPTION
Horiz. and Vert. Edge Separation in High Detail Area of Photo #23	24	Horiz. edge area: White white Vert. edge area: Black black
Slant Edge Separation--in High Detail Area when Horiz. and Vert. Edge Contents are Equal in Photo #23	25	+45° Edge Area: Black black -45° Edge Area: White white

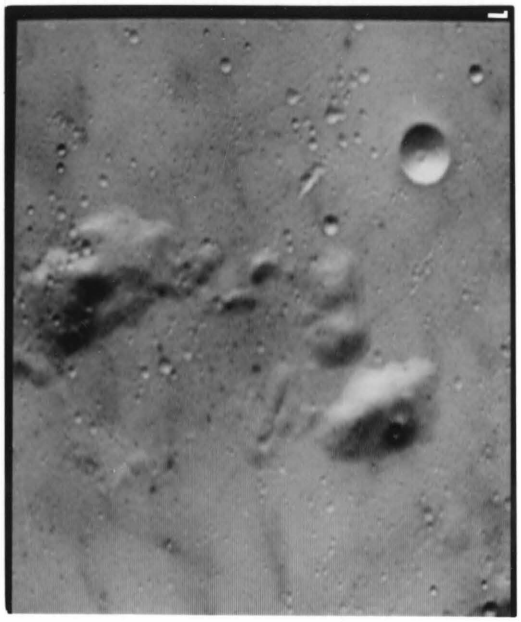
Pred. Error Picture with Adaptive Prediction--KAREN	26	Area Detail	Pred #	MAX	VAR
		Low	5	29	11
		Med.	24	64	41
		High (Horiz.Edge	9	67	140
		High (Vert. Edge	8		
		TOTAL		67	60

Pred. Error Histograms for		
Photo #26 (Vert. Axis--Count of Occurrences; Horiz. Axis--Mag. of Errors--32 per div.)	27	Low detail area
	28	Medium detail area
	29	High detail area

Segmentation--CHECK GIRL	30	Algorithm same as for KAREN
Horiz. and Vert. Edge Separation in High Detail Area of PHOTO #30	31	

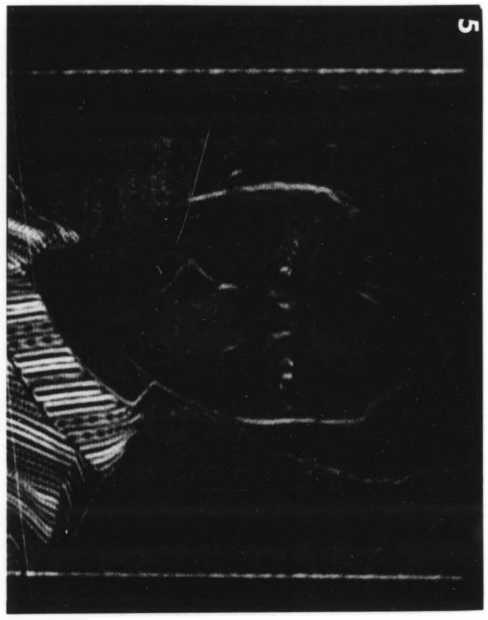
CAPTION	PHOTO #	DESCRIPTION		
Segmentation--MOON	32	Algorithm same as for KAREN		
Horiz. and Vert. Edge Separation in High Detail Area of PHOTO #32	33			
Segmentation--AERIAL	34	"	"	"
Horiz. and Vert. Edge Separation in High Detail Area of PHOTO #34	35	"	"	"
Adaptive Encoding of KAREN with Pred #'s 5/24/(8,9) and n = 6/6/6	36	4/7/8	1.514	8
	37	4/6/8	1.452	8
	38	4/6/8	1.336	6 (Ensemble Quantizer)
Demonstration Picture	39	First Order Ent = 8.000 Second Order Ent = 7.994 (w.r.t.prev. neighbor) Third Order Ent. = 0.000 (w.r.t.prev. two neighbors)		

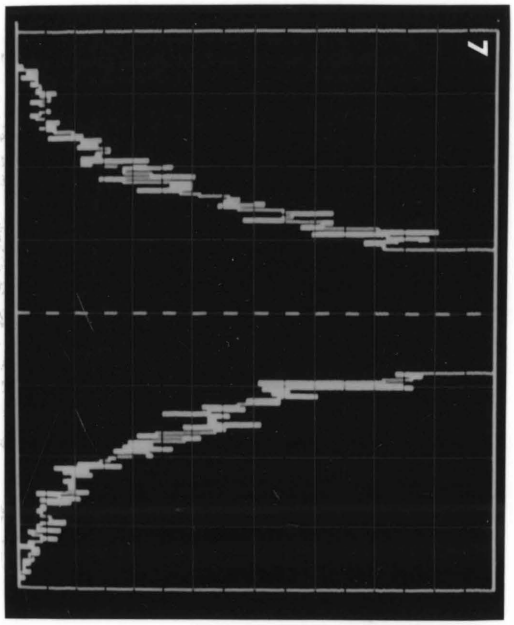

3

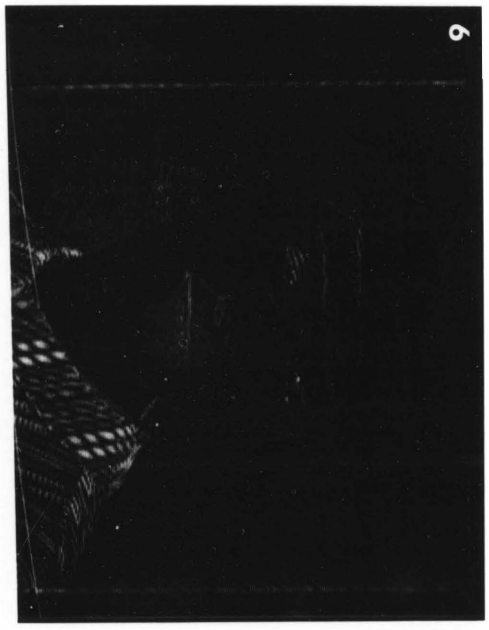

1

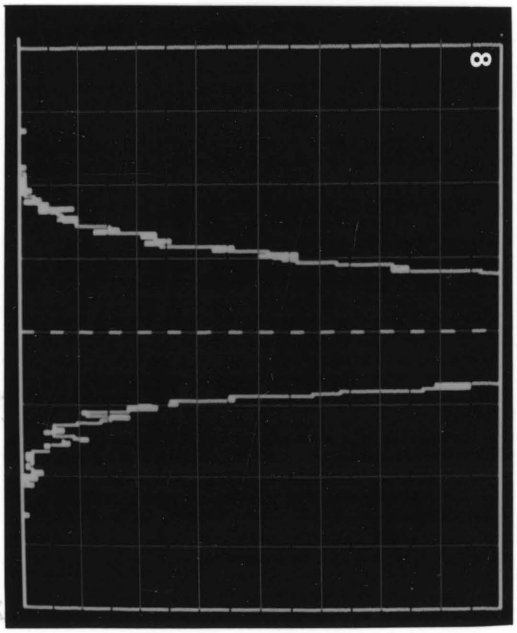

4

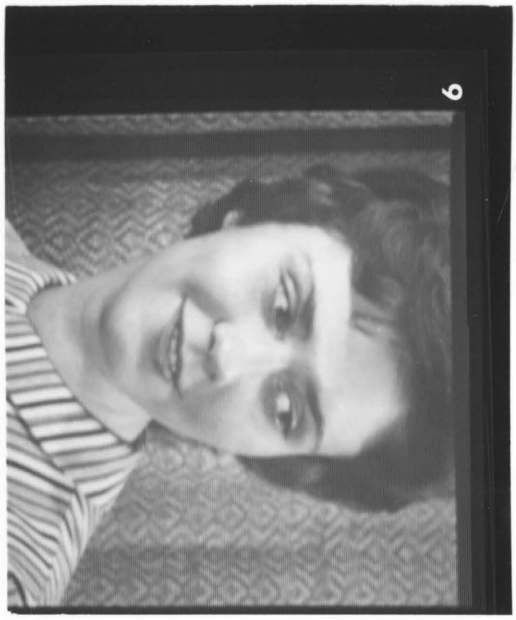

2

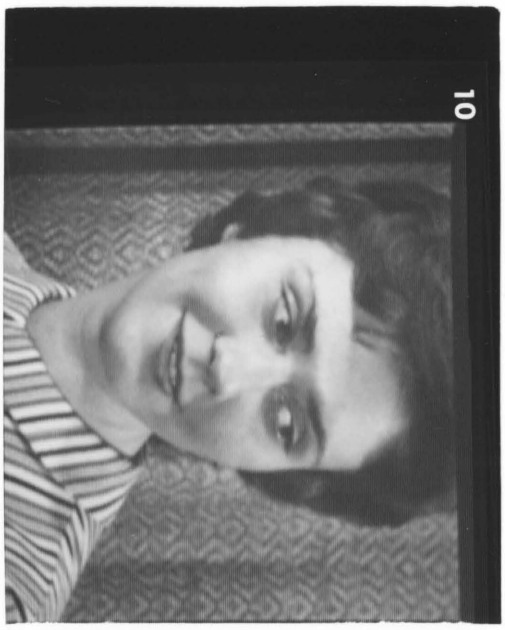

5


7


6


8

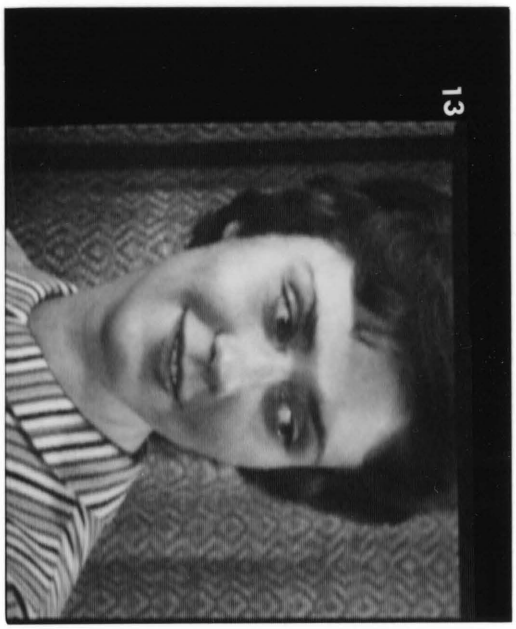

9


10


11


12

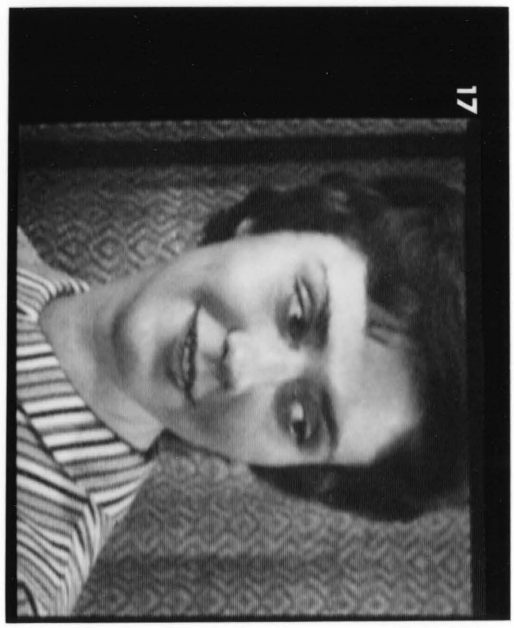

14


13


15

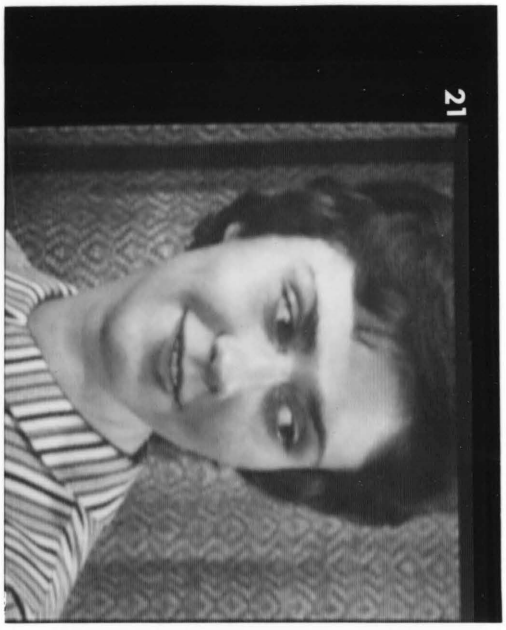

16

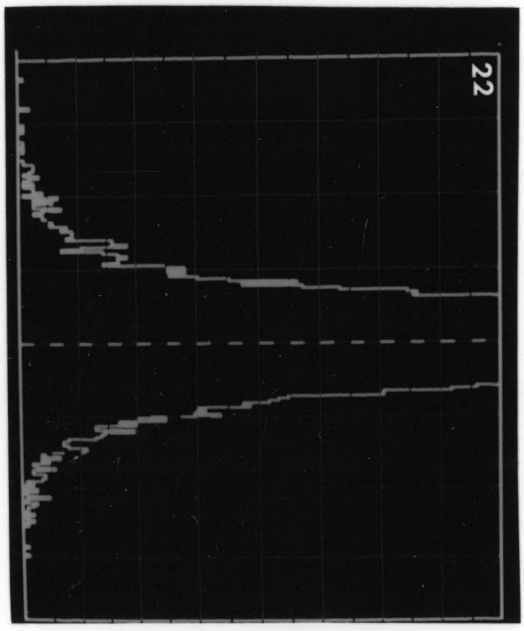

18


17

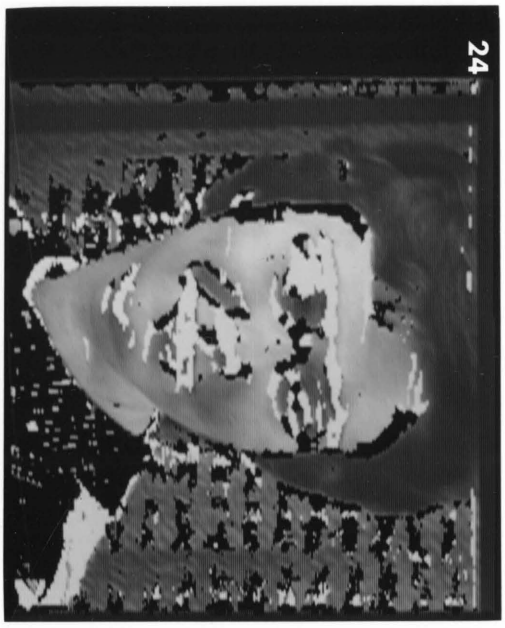

19

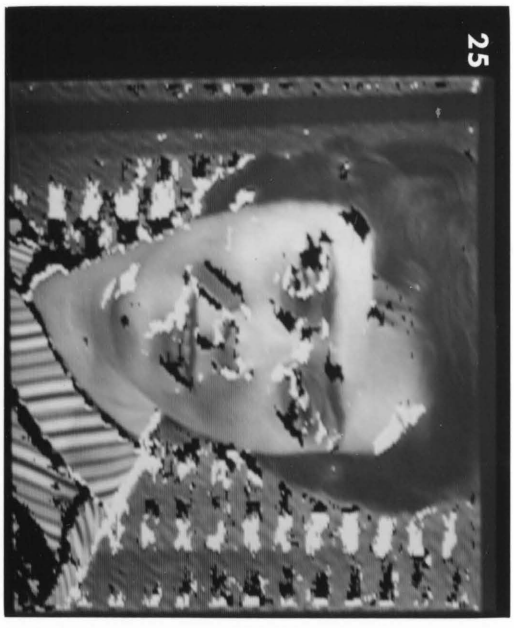

20

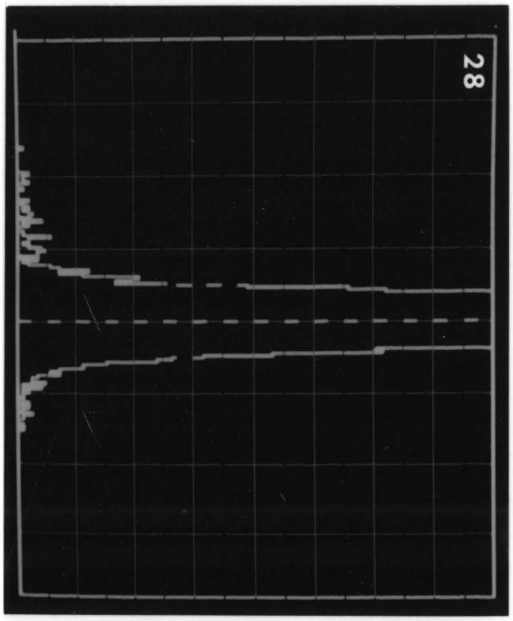

21

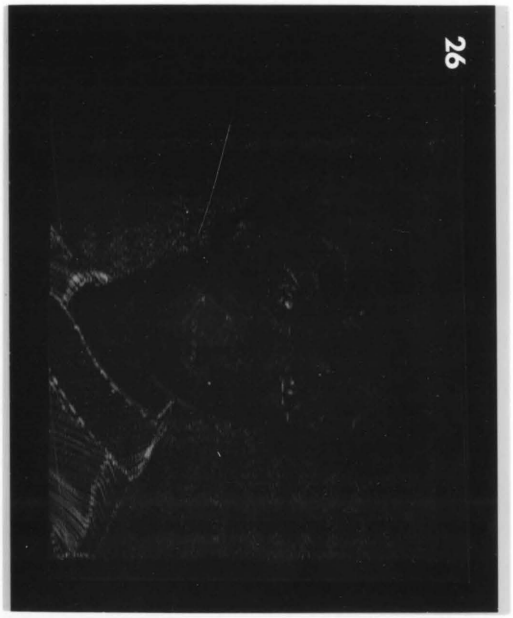

22

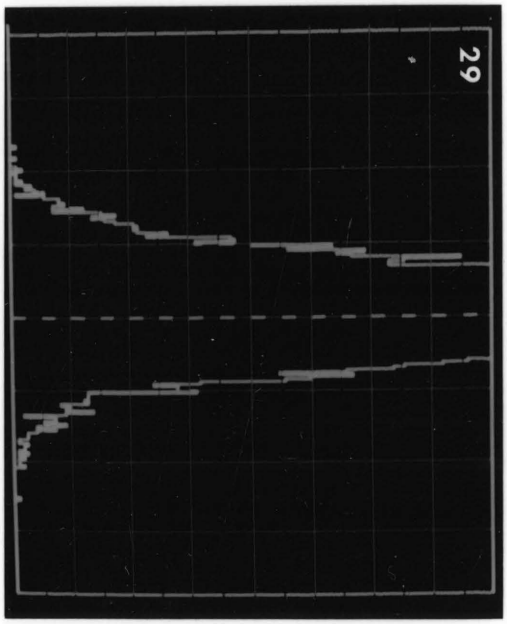

23

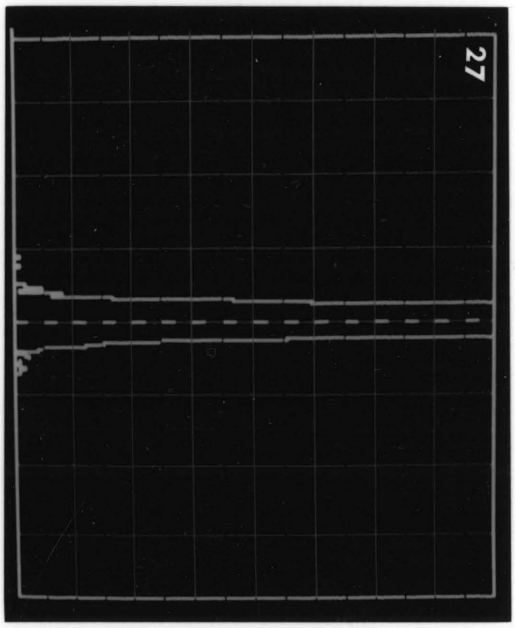

24

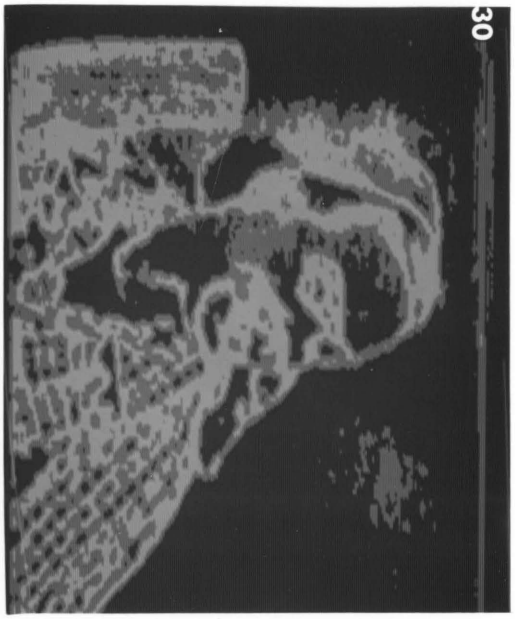

25

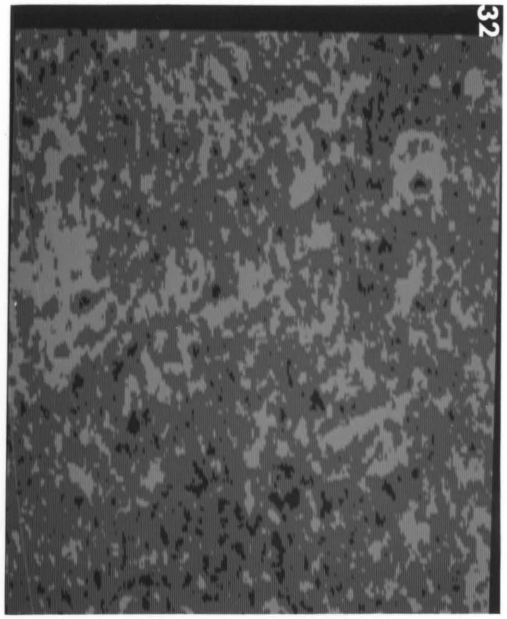

28


26

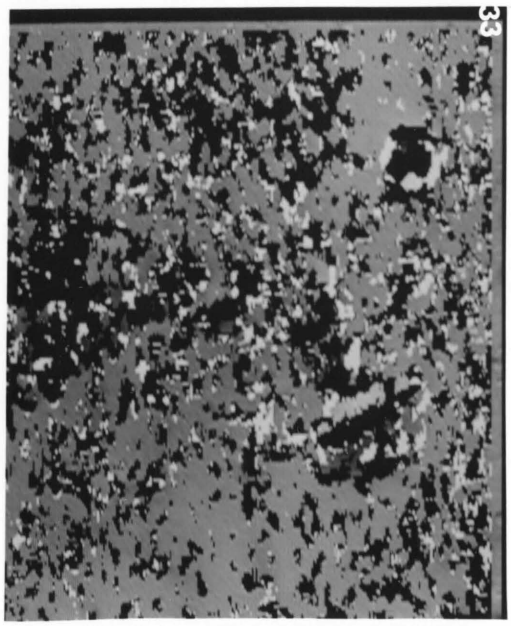

29

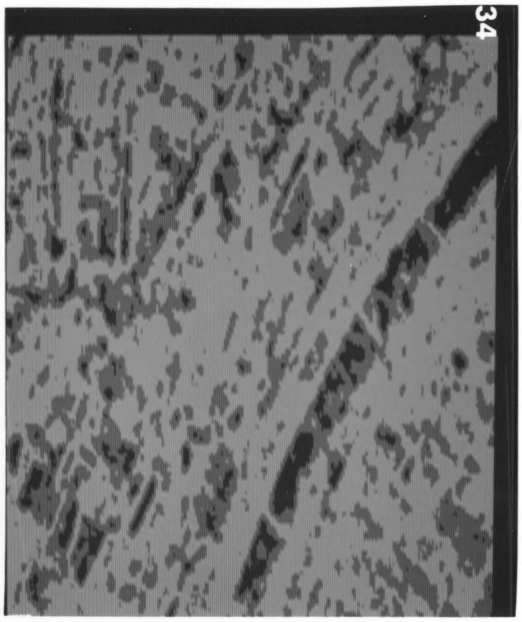

27

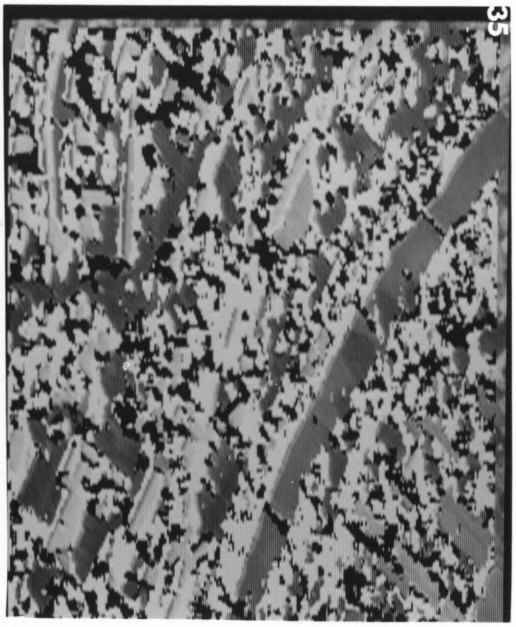

30

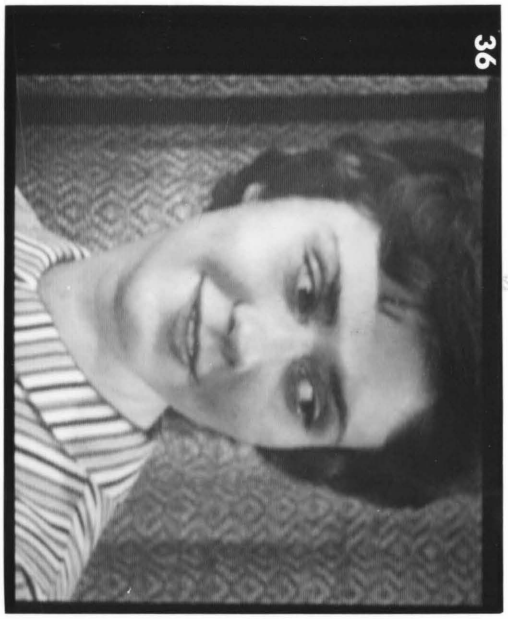

32


31

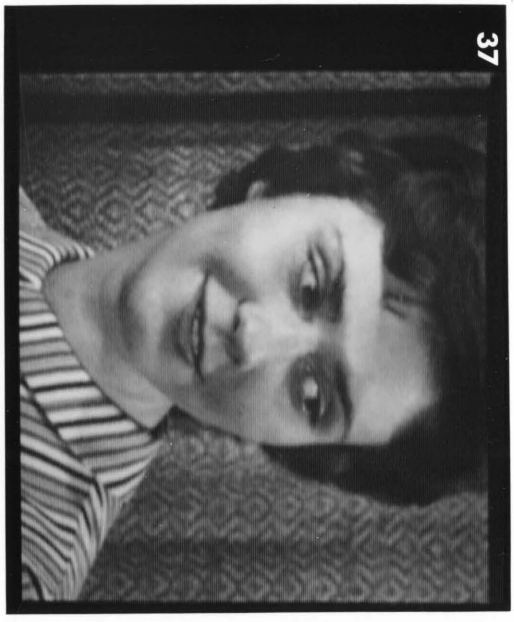

33

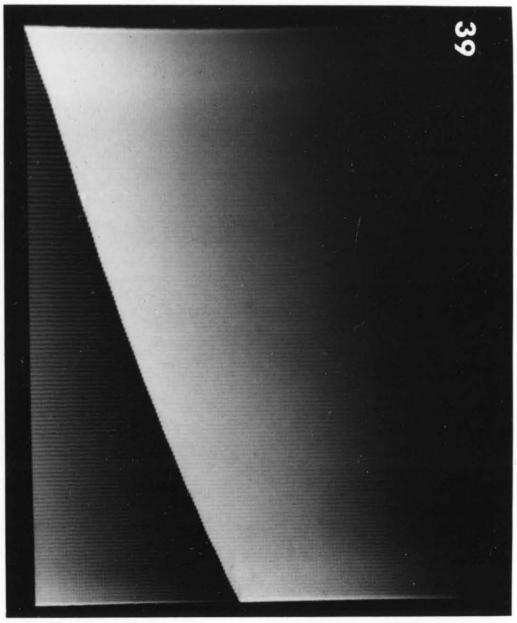

34


35


36


38

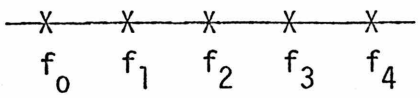

37


39

APPENDIX B

In this appendix we present proofs of the basic expressions behind extrapolative and interpolative type of predictors. The underlying concepts are Lagrange interpolation and Taylor series expansion which could be found in any standard text on Numerical Analysis.* For the sake of illustration we shall discuss third order predictors only as the first, second and fifth order ones can be derived in an an analogous fashion.

B.1 Third Order Polynomial Extrapolation and Interpolation



Suppose we are given that a function $f(x)$ assumes the values f_0, f_1, f_2 and f_3 at points x_0, x_1, x_2 and x_3 respectively. This allows us to find a unique third degree polynomial $p(x)$ passing through the four given points. Thus the third order extrapolated value of the function $f(x)$ to the point x_4 is given by $p(x_4)$.

Let us assume that x_i's are uniformly spaced with spacing h between them, i.e.,

$$x_i - x_{i-1} = h \ , \quad 1 \leq i \leq 4 \quad .$$

Then the third degree polynomial is given by

---

* See, for instance, Isaacson and Keller, Analysis of Numerical Methods (John Wiley, 1966).

$$p(x) = f_0\phi_0(x) + f_1\phi_1(x) + f_2\phi_2(x) + f_3\phi_3(x)$$

where $\phi_i(x)$ have the property that

$$\phi_i(x_j) = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \qquad \begin{matrix} 0 \leq i \leq 3 \\ 0 \leq j \leq 3 \end{matrix}$$

and they are given by

$$\phi_i(x) = \prod_{i \neq j} \frac{(x - x_j)}{(x_i - x_j)} \quad .$$

For example,

$$\phi_2(x) = \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)} \quad .$$

Note that

$$\phi_2(x_2) = 1 \quad ,$$

while

$$\phi_2(x_i) = 0 \quad \text{for } i = 0,1 \text{ and } 3.$$

Using this formulation we need $\phi_i(x_4)$ for $0 \leq i \leq 3$ to compute $p(x_4)$. For example,

$$\phi_2(x_4) = \frac{(4h)(3h)(h)}{(2h)(h)(-h)} = -6 \quad .$$

Similarly we obtain

$$\phi_0(x_4) = -1 \quad , \quad \phi_1(x_4) = \phi_3(x_4) = 4 \quad .$$

Therefore, we obtain the expression for Pred #13 as

$$p(x_4) = -f_0 + 4f_1 - 6f_2 + 4f_3 \quad ,$$

which has error of approximation $O(h^4)$ if $f^{(4)}(x)$ exists. For details on the error terms, see the reference cited in the beginning of this appendix.

Now to derive the basic expression for Pred #23 assume that f_o, f_1, f_3 and f_4 are given. Then the interpolated value of $f(x)$ at x_2 can be computed by forming another cubic $q(x)$ passing through x_o, x_1, x_3 and x_4 using Lagrange interpolation as discussed above. This yields the interpolated value as

$$q(x_2) = \frac{1}{6}(-f_o + 4f_1 + 4f_3 - f_4)$$

which is the same as Eq. I of Sect. 2.2. This approximation also has error terms of $O(h^4)$.

B.2 Deriving Expression for Pred #14

In this section we derive the expression for Pred #14. Suppose we know f_i, $i \leq 3$, then expand $f(x)$ in Taylor series about x_3 to obtain

$$f(x_4) \cong f(x_3) + h\, f'(x_3) + \frac{h^2}{2!} f''(x_3)$$

where we retain only up to second order terms. Now approximate the derivatives by finite differences of the function $f(x)$ as

$$f'(x_2) \cong \frac{1}{h}(f_2 - f_1) \quad ,$$

$$f'(x_3) \cong \frac{1}{h}(f_3 - f_2) \quad ,$$

$$f''(x_3) \cong \frac{1}{h}(f'(x_3) - f'(x_2))$$

$$\cong \frac{1}{h^2}(f_3 - 2f_2 + f_1) \quad .$$

Substituting these in the Taylor expansion we obtain

$$f(x_4) \tilde{=} \frac{1}{2}(5f_3 - 4f_2 + f_1) \quad .$$

The expression for Pred #15 can be derived in a similar way.

B.3

In Sect. 2.2 we gave a derivation of Pred #17. Here, we shall use the same nomenclature and derive expressions for Pred #'s 16 and 18 also. With reference to Fig. 2-1 we can write

$$X \tilde{=} (-H + 4I + 4\alpha - \beta)/6 \quad . \qquad (I)$$

In addition, let us also approximate α and β using Pred #21 as

$$\alpha \tilde{=} (X + K)/2 \qquad (II)$$

and

$$\beta \tilde{=} (\alpha + L)/2 \quad . \qquad (III)$$

Equations I, II and III are solved simultaneously to obtain Pred #16 as

$$\hat{X} = (-4H + 16I + 7K - 2L)/17 \quad .$$

To obtain Pred #18 we use expression for Pred #23 to approximate α and β as

$$\alpha \tilde{=} (-I + 4X + 4K - L)/6 \qquad (IV)$$

and

$$\beta \tilde{=} (-X + 4\alpha + 4L - M)/6 \quad , \qquad (V)$$

which are solved together with Eq. I, to obtain

$$\hat{X} = (-18H + 62I + 40K - 22L + 3M)/65 \quad .$$

APPENDIX C

In this appendix we discuss the third order entropy of digitized and sampled pictures. We also present the algorithm to make the demonstration picture (PHOTO #39) mentioned at the end of Sect. 2.4.

C.1

Let X, Y, Z denote the general triplet of three consecutive pixels in the picture. Then the third order entropy is

$$\begin{aligned} H(Z|X,Y) &= -\sum_{i,j,k} p(i,j,k) \log p(i,j,k) \\ &\quad + \sum_{i,j} p(i,j) \log p(i,j) \\ &= H(X,Y,Z) - H(X,Y) \end{aligned}$$

where p's are joint probability distributions and H's denote entropies.

In order to carry out the measurement of $H(Z|X,Y)$ on a given picture we count the number of occurrences $n(i,j,k)$ of the triplets (i,j,k) in which X,Y,Z assume values i,j,k respectively. In our case i, j and k can take on values between 0 and 255 for 8 bit PCM pictures. Then, letting N ($=256\times254 \cong 2^{16}$) be the total number of triplets encountered in the picture, we make the assumption

$$p(i,j,k) \cong \frac{n(i,j,k)}{N} \quad . \qquad (I)$$

Now, since i, j and k can assume 2^8 values each, there could be 2^{24} possible triplets. But out of these 2^{24} possible triplets we can observe at most about 2^{16}, which constitute a very small fraction

($\simeq \frac{1}{2^8}$) of the sample space. This implies that the probability assignments made according to Eq. I are necessarily at fault. This is so because in a probability experiment, for the relative frequencies to be close to their respective probabilities, it is necessary that the number of trials be at least as many as the total number of outcomes. (Consider the case of throwing a dice only 3 times and trying to assign probabilities based on the relative frequencies.)

To illustrate the point we have constructed the "random picture" PHOTO #39 in which

1. all the 256 pixel values occur equally frequently, i.e., $H(X) = 8$ bits/pixel,
2. all the 2^{16} doublets occur equally frequently, i.e., $H(X,Y) = 16$ bits/doublet, and
3. all the 2^{16} triplets in it are distinct, i.e., $H(X,Y,Z) = 16$ bits/triplet.

Hence

$$H(Y|X) = H(X,Y) - H(X) = 8$$

and

$$H(Z|X,Y) = H(X,Y,Z) - H(X,Y) = 0 \quad !$$

The measurements show that $H(Y|X)$ is 7.994 instead of 8. This is because ideally the picture should have had 256x258 pixels, but PHOTO #39 has only 256x256. Construction of such a picture has been described in the next section.

Such situations can be avoided if we make sure that the number of triplets ($\simeq M^2$ for M x M picture) in the picture is at least as large as the total number of possible triplets ($= 2^{3n}$ for n-bit/pixel picture). This would be the case if

$$M \cong 2^{3n/2} \quad .$$

We present this relation in tabular form below.

n (# of bits/pixel)	Smallest Picture Size for H(Z|X,Y) Measurement
8	4096 x 4096
7	2048 x 2048
6	512 x 512
5	256 x 256
4	64 x 64

According to this table, for 256 x 256 pictures we should not have more than 5 bits/pixel for the third order entropy measurement to be meaningful.

C.2

In this section we show how to make the demonstration picture PHOTO #39. Since intensities of pixels are represented by integers, we can use the terms "table of integers" and "picture" interchangeably.

Let us consider pictures with n bits/pixel, where n is even*, i.e., $n = 2m$.

*For odd n, the resulting table is not square, but the conclusions of Sect. C.1 are still valid.

Then,

$$\text{total number of triplets} = 2^{6m} ,$$

and $$\text{total number of doublets} = 2^{4m} .$$

Thus, using integers from 0 to $(2^{2m}- 1)$, we propose to make a table of numbers in which all the 2^{4m} possible doublets occur only once and no triplet is repeated. The size of such a table would be $2^{2m} \times (2^{2m}+ 1)$. Note that for an 8-bit picture, i.e., $m = 4$, this corresponds to 256×257.

The following observation will be useful.

Lemma 1: In the table of numbers mentioned above, if all the doublets are distinct, so are all the triplets.

Proof: Suppose there are two triplets (a_1,a_2,a_3) and (b_1,b_2,b_3) that are identical. Then the doublets (a_1,a_2) and (b_1,b_2) are also equal. This implies the result. Q.E.D.

Because of Lemma 1, we only have to form the table in which all doublets are distinct. We shall be working with the column of numbers given below and its cyclic permutations.

$$c_0 = (0,1,2,\cdots,2^{2m}- 1)^T$$

where "T" stands for transpose. We also call it column 0, meaning it is an upward cyclic shift of c_0 by 0 places. Similarly, c_j, or simply column j, would be obtained from c_0 by cyclically shifting it upwards by j places. For example,

$$c_2 = (2,3,4,\cdots,2^{2m}- 1, 0, 1)^T .$$

A column of pairs obtained by putting c_2 after c_0 will be called c_2c_0 or [2,0], i.e.,

$$[2,0] = (2,0;\ 3,1;\ 4,2;\ \cdots;\ 1,2^{2m}-1)^T .$$

Similarly for other pairs. Thus we construct the sets C and S as

$$C = \{c_i \mid 0 \leq i \leq 2^{2m}-1\}$$

$$= \text{set of all cyclic permutations of } c_0 ,$$

and $$S = \{\text{column pair}[i,j] \mid 0 \leq i \leq 2^{2m}-1 ,\ 0 \leq j \leq 2^{2m}-1\}$$

$$= \text{set of all column pairs } c_ic_j .$$

Now note that the column pair [2,0] has the same pairs in it as [3,1], [4,2], etc., as they have been obtained by cyclic shifts of [2,0]. This observation allows us to partition the set S in disjoint classes, such that column pairs belonging to the same class are cyclic shifts of each other. This is done by putting an equivalence relation on the set S, defined by

[i,j] and [k,ℓ] are related to each other iff,

1) they are cyclic shifts of each other,

or 2) $(i-j) \equiv (k-\ell) \pmod{2^{2m}}$.

Thus we obtain 2^{2m} equivalent classes named $e_0, e_1, \cdots, e_{2^{2m}-1}$. The name e_j of the j^{th} class is derived from the fact that

$$[0,j] \in e_j .$$

At this point we give an example for $m = 1$. We have

$$c_0 = (0,1,2,3)^T \quad ,$$

and S has 16 column pairs in it, which are divided into four equivalent classes as follows:

$\underline{e_0}$	$\underline{e_1}$	$\underline{e_2}$	$\underline{e_3}$
[0,0]	[0,1]	[0,2]	[0,3]
[1,1]	[1,2]	[1,3]	[1,0]
[2,2]	[2,3]	[2,0]	[2,1]
[3,3]	[3,0]	[3,1]	[3,2]

The rest of the discussion we shall present through this example for simplicity. We make a 4 x 5 table of numbers with the desired property in the following steps:

1. Pick any permutation of symbols e_0, e_1, e_2, e_3 ; say, we pick

$$e_1, e_2, e_3, e_0 \quad .$$

2. We shall translate this into a sequence of column numbers as follows. First replace e_1 by any one of its members, say by [2,3]. Then the sequence becomes

$$2,3,e_2,e_3,e_0 \quad .$$

Then choose the member of e_2 which has the form [3,*]. This will be a unique choice, which in this case is [3,1]. Replacing e_2 by the second column number which is 1, we get

$$2,3,1,e_3,e_0 \quad .$$

Next we replace e_3 by 0 to obtain

$$2,3,1,0,e_0 \quad ,$$

and finally

$$2,3,1,0,0 \quad (= c_2c_3c_1c_0c_0) \ .$$

Note that these numbers refer to column numbers. They translate into the table as follows:

c_2	c_3	c_1	c_0	c_0
2	3	1	0	0
3	0	2	1	1
0	1	3	2	2
1	2	0	3	3

In general we shall choose a sequence of e_j's of length 2^{2m} which will translate into a sequence of column names having length $2^{2m}+1$. This, in turn, will yield a table of size $2^{2m} \times (2^{2m}+1)$.

That there is no repetition of doublets in such a table follows from

$$e_i \cap e_j = \phi \quad \text{for} \quad i \neq 1 \quad .$$

In addition, the table has 2^{4m} doublets, each of which occurs only once, as there are only 2^{4m} total number of them. This completes the description of our method.

We give another example for $m = 2$, i.e., 4-bit/pixel picture, for which

$$c_0 = (0,1,2,3,4,5,6,7).$$

The sequence of e_j's used is

$$e_0, e_1, e_7, e_2, e_6, e_3, e_5, c_4 \ ,$$

which converts into the following sequence of column numbers

$$0,\ 0,\ 1,\ 0,\ 2,\ 0,\ 3,\ 0,\ 4 \ .$$

The tabular form of this is

c_0	c_0	c_1	c_0	c_2	c_0	c_3	c_0	c_4
0	0	1	0	2	0	3	0	4
1	1	2	1	3	1	4	1	5
2	2	3	2	4	2	5	2	6
3	3	4	3	5	3	6	3	7
4	4	5	4	6	4	7	4	0
5	5	6	5	7	5	0	5	1
6	6	7	6	0	6	1	6	2
7	7	0	7	1	7	2	7	3

For PHOTO #39 we used $m = 4$ and

$$e_0,\ e_1,\ e_{255},\ e_2,\ e_{254},\ e_3,\ e_{253},\ \cdots,\ e_{128} \ ,$$

whose column numbers are

$$0,\ 0,\ 1,\ 0,\ 2,\ 0,\ 3,\ 0,\ \cdots,\ 0,\ 128 \ .$$

APPENDIX D

In this appendix we derive Eq. I, Sect. 3.1, and then we discuss extension of Max's condition to the case of "mean n^{th} power" distortion criterion.

D.1

Recall from Sect. 3.1, that the discrete variable x, input to the quantizer Q, assumes only integral values, and that the input thresholds for the N-level quantizer are $X_1, X_2, \cdots, X_{N+1}$ such that

$$X_1 \leq x \leq X_{N+1} \quad ,$$

and

$$X_1 < X_2 \cdots < X_{N+1} \quad .$$

Then the representative levels y_k's are such that

$$Q(x) = y_k \quad \text{for} \quad X_k < x \leq X_{k+1}$$

$$\text{and} \quad 1 \leq k \leq N \ .$$

Let us, for our discussion here, write the distortion as

$$D = \sum_{k=1} \sum_{X_k < x \leq X_{k+1}} f(x - y_k)\, p(x)$$

where $f(\xi)$ is a strictly monotone increasing function for $\xi \geq 0$, and

$$f(\xi) = f(-\xi) \quad . \tag{I}$$

Note also that X_k's and y_k's are also integer valued variables. We prove the main result in two lemmas.

Lemma 1: For D to be minimum we must have

$$X_k < y_k \leq X_{k+1} \qquad \text{(II)}$$

for $1 \leq k \leq N$.

Proof: Note that D is the sum of N terms of the form

$$D_k = \sum_{X_k < x \leq X_{k+1}} f(x - y_k)\, p(x)$$

each of which is minimized only when Eq. II is satisfied for the particular value of k. This is so because if $y_k > X_{k+1}$ or $y_k \leq X_k$ then the value of D_k can be reduced by respectively reducing or increasing the value of y_k as f is an increasing function satisfying Eq. I.

Q.E.D.

Lemma 2: Given the values of y_k's, a necessary condition for D to be minimum is

$$X_k = \left[\frac{y_{k-1} + y_k}{2}\right] \qquad \text{(III)}$$

for $2 \leq k \leq N$, where the square brackets indicate the largest integer smaller than the quantity inside them.

Proof: We shall need the following property of local minimum of a function of discrete variable.

Let F be a function of a discrete variable ξ_n .

Define the forward difference of F at a point ξ_n as

$$\Delta(\xi_n) = F(\xi_{n+1}) - F(\xi_n) \quad .$$

Then ξ_n is a point of local minimum of F if the

following two conditions hold.

1. $\Delta(\xi_n) > 0$, and if $\Delta(\xi_n) = 0$, there exists a smallest positive integer N_1 such that $\Delta(\xi_{n+N_1}) > 0$.
2. $\Delta(\xi_{n-1}) < 0$, and if $\Delta(\xi_{n-1}) = 0$, there exists a smallest positive integer N_2 such that $\Delta(\xi_{n-N_2}) < 0$.

Here we consider D as a function of $N-1$ independent variables $X_2, X_3, \cdots, X_N$. Then its forward difference along the k^{th} "axis" is

$$\Delta^k(X_k) = D(X_2,\cdots,X_k+1,\cdots,X_N) - D(X_2,\cdots,X_k,\cdots,X_N)$$

where we have made use of the fact that X_j's assume integer values. Substituting for D we obtain

$$\Delta^k(X_k) = f(X_k+1 - y_{k-1})\ p(X_x+1) - f(X_k+1 - y_k)\ p(X_k+1)$$

and

$$\Delta^k(X_k-1) = f(X_k - y_{k-1})\ p(X_k) - f(X_k - y_k)\ p(X_k) \quad .$$

Now, under the additional assumption that $p(\xi) \neq 0$ for all ξ, we see that

$$\Delta^k(X_k) \neq 0$$

and

$$\Delta^k(X_k-1) \neq 0 \quad ,$$

as f is a strictly monotone function, and $y_{k-1} < y_k$. Therefore, applying properties 1 and 2 stated above we obtain

$$|x_k+1 - y_{k-1}| > |x_k+1 - y_k|$$

and

$$|x_k - y_k| < |x_k - y_k| \quad .$$

Then application of Lemma 1 yields

$$(x_k+1) - y_{k-1} > y_k - (x_k+1)$$

and

$$x_k - y_{k-1} < y_k - x_k \quad .$$

That is,

$$x_k + 1 > (y_{k-1} + y_k)/2$$

and

$$x_k < (y_{k-1} + y_k)/2 \quad ,$$

which is the same as

$$x_k = \left[\frac{y_{k-1} + y_k}{2}\right] \quad .$$

The condition that $p(\xi) \neq 0$ for all ξ is very easily relaxed by observing if we allow $p(\xi)$ to be zero for some ξ then there exist integers m and n such that

$$m \geq 0 \quad , \quad n \geq 0 \quad ,$$

$$p(x_k + m + 1) \neq 0 \quad , \quad \Delta^k(x_k + m) > 0$$

and $$p(x_k - n) \neq 0 \quad , \quad \Delta^k(x_k - n - 1) < 0 \quad .$$

Similar to the previous case, we obtain

$$x_k + m + 1 > (y_{k-1} + y_k)/2$$

and

$$x_k - n < (y_{k-1} + y_k)/2 \quad ,$$

which also yield the desired result.

Q.E.D.

D.2

In this section we present extension of Max's condition to the case of "mean n^{th} power" distortion criterion. Let us assume that x, x_k's and y_k's are all continuous variables in which case distortion is given by

$$D = \sum_{k=1}^{N} \int_{x_k}^{x_{k+1}} |x - y_k|^n \, p(x) \, dx \quad .$$

Then, one necessary condition which is also counterpart of Eq. I, Sect. 3.1, for D to be minimum is

$$x_k = (y_{k-1} + y_k)/2, \quad 2 \leq k \leq N \quad ,$$

as given by Max [17]. Proof of this condition is valid even if $|x - y_k|^n$ is replaced by $f(x - y_k)$ in the expression for D, where f is a monotone increasing function satisfying Eq. I. The second condition given by him was for the case $n = 2$ which is generalized in the following.

We shall need the following result in the main proof.

Lemma 3:

Suppose $f(x,y)$ is differentiable with respect to y, and that $f(x,y)$ and $f_y(x,y)$ both are continuous functions integrable with respect to x over all intervals of interest for all fixed values of y. Let

$$F(y) = \int_a^y f(x,y)\, dx \ .$$

Then F is differentiable with respect to y and

$$F'(y) = \int_a^y f_y(x,y)\, dx + f(y,y) \ .$$

Proof: We shall investigate the following as $h \to 0$.

$$\frac{F(y+h) - F(y)}{h} = \frac{1}{h} \int_a^{y+h} f(x,y+h)\, dx - \frac{1}{h} \int_a^y f(x,y)\, dx$$

$$= \int_a^y \frac{f(x,y+h) - f(x,y)}{h}\, dx + \int_y^{y+h} f(x,y+h)\, dx \ .$$

Then, by mean value theorem for integrals

$$\frac{1}{h} \int_y^{y+h} f(x,y+h)\, dx = f(\xi, y+h) \frac{(y+h) - y}{h} \quad \text{for} \quad y < \xi < y+h$$

which goes to $f(y,y)$ as $h \to 0$.

In addition to this, since $\frac{f(x,y+h) - f(x,y)}{h} \to f_y(x,y)$ as $h \to 0$ we see, for small enough h and $\varepsilon < 1$,

$$\left|\frac{f(x,y+h) - f(x,y)}{h}\right| - |f_y(x,y)| < \varepsilon < 1 .$$

That is, for fixed y ,

$$\left|\frac{f(x,y+h) - f(x,y)}{h}\right| \leq 1 + |f_y(x,y)|$$

and by dominated convergence theorem we get

$$\int_a^y \frac{f(x,y+h) - f(x,y)}{h}\, dx \to \int_a^y f_y(x,y)\, dx \quad \text{as} \quad h \to 0 .$$

Q.E.D.

Corollary 5: For $F(y) = \int_y^a f(x,y)\, dx$,

$$F'(y) = \int_y^a f_y(x,y)\, dx - f(y,y) .$$

Now we present the necessary condition.

Lemma 6:

A necessary condition for D to be minimum is

$$\int_{x_k}^{y_k} |x - y_k|^{n-1} p(x)\, dx = \int_{y_k}^{x_{k+1}} |x - y_k|^{n-1} p(x)\, dx$$

for $1 \leq k \leq N$, $n \geq 1$.

For even n , i.e., $n = 2m$, it could be rewritten as

$$\int_{x_k}^{x_{k+1}} (x - y_k)^{2m-1} p(x)\, dx = 0 \quad , \qquad m \geq 1 ,$$

which reduces to Max's condition for $m = 1$.

For odd n , i.e., $n = 2m+1$, it becomes

$$\int_{x_k}^{x_{k+1}} \frac{(x - y_k)^{2m+1}}{|x - y_k|} \, p(x) \, dx = 0 , \qquad m \geq 0 .$$

Proof:

For $n = 2m$ we can write D as

$$D = \sum_{k=1}^{N} \int_{x_k}^{x_{k+1}} (x - y_k)^{2m} \, p(x) \, dx .$$

Then

$$\frac{\partial D}{\partial y_k} = 2m \int_{x_k}^{x_{k+1}} (x - y_k)^{2m-1} \, p(x) \, dx = 0$$

yields the result.

For $n = 2m+1$ we write D_k as

$$D_k = \int_{x_k}^{x_{k+1}} |x - y_k| \, (x - y_k)^{2m} \, p(x) \, dx$$

$$= - \int_{x_k}^{y_k} (x - y_k)^{2m+1} \, p(x) \, dx + \int_{y_k}^{x_{k+1}} (x - y_k)^{2m+1} \, p(x) \, dx .$$

Then using Lemma 4 and Corollary 5 we obtain

$$\frac{\partial D}{\partial y_k} = \frac{\partial D_k}{\partial y_k} = -(2m+1)\int_{x_k}^{y_k} (x - y_k)^{2m}\, p(x)\, dx + (2m+1)\int_{y_k}^{x_{k+1}} (x - y_k)^{2m}\, p(x)\, dx = 0 \quad ,$$

which gives us

$$\int_{x_k}^{x_{k+1}} \frac{(x - y_k)^{2m+1}}{|x - y_k|}\, p(x)\, dx = 0 \qquad m \geq 0 \quad .$$

Q.E.D.

We feel that in Max's algorithm [17], to design the quantizer, enforcement of above conditions for general n will be quite cumbersome and therefore we resort to a different algorithm presented in Sec. 3.1. We make this algorithm efficient by making use of the discrete nature of the input variable x, as integer arithmetic on digital computer is much faster than floating point arithmetic.