Copyright © 1971 American Telephone and Telegraph Company
THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 3, March, 1971
Printed in U.S.A.

Television Coding Using Two-Dimensional Spatial Prediction

By D. J. CONNOR, R. F. W. PEASE and W. G. SCHOLES

(Manuscript received September 30, 1970)

Three classes of differential coders have been built which exploit the two-dimensional spatial correlation present in television pictures. The three classes are distinguished according to the technique used to predict the value of the picture element (pel) to be coded. In the first class, the prediction is the value of the previous element along the scan line averaged with one or more elements in the previous line (of the same field). In the second class, the prediction is the sum of the value of the previous pel with the local element difference in the previous line. In the third class, the transmitter and receiver choose the prediction from either the previous element or some combination of the previous element with the elements in the previous line. The choice is based on differences measured between elements already coded. These are available both at the transmitter and receiver, so no extra information need be transmitted. In the absence of transmission errors, the pictures resulting from all three classes of coder are markedly improved over those in which only the previous element is used as a prediction. In particular, vertical and sloping edges are now well defined. The effect of a transmission error on a single frame of the received picture is scarcely visible in the first class of coder but is much more visible in the second and third classes.

I. INTRODUCTION

Intensity correlation in television pictures extends both horizontally and vertically.[1-3] For a number of good reasons, most practical predictive encoders have exploited only the horizontal correlation:[3]

(*i*) In a noninterlace system, access to the vertically adjacent picture element (pel) requires storage of one line of information (about 8,000 bits for broadcast TV or 2,000 bits for *Picture-*

phone® Service). In a 2 : 1 interlaced system, a field plus a line must be stored (10^6 bits for broadcast or 250,000 for *Picturephone* Service). If access to the vertically adjacent pel in the same field is desired, a line of storage is required.

(*ii*) Most television systems, including *Picturephone* Service, use an interlaced scan. Because of the storage problem, access to the vertically adjacent pel in the same field is most practical. However, this pel is about twice as far away (spatially) from the current pel as is the previous pel in the same scan line. The vertical correlation of intensity within a field is thus less than the horizontal correlation.

(*iii*) Statistical studies of television pictures suggested that for linear predictive encoding very little extra redundancy could be removed by exploiting the vertical correlation as well as the horizontal correlation.[2,3]

Given these reasons, why look at the use of vertically adjacent pels in various predictive encoding schemes? First, in recent years the cost of storage has fallen dramatically and hence the extra memory required for storing one line is not a serious objection. At present, a 2,000-bit line store using MOS shift registers costs about $600.00 and fits on one printed circuit board containing 40 dual-in-line packages. Future costs should be much lower (less than $100.00).

To counter the second and third points above, let us consider for a moment the operation of a differential quantizer. A differential quantizer predicts the value of the current pel by using information that has already been transmitted. The discrepancy between the uncoded value and the prediction is quantized and added to the prediction to form the coded value of the current pel.

Most practical differential quantizers which make use of the horizontal correlation of intensity produce pictures of generally acceptable quality, except for sharp vertical or diagonal edges. These are blurred and have an annoying "busyness." The problem is in the prediction. The differential quantizer has no way of "anticipating" a vertical edge. Consequently, the prediction it makes is grossly in error at these edges.

By making use of information from the previous line, it is possible to anticipate the vertical edges and most of the diagonal ones. To study this possibility, we have built and/or simulated on a computer a number of differential quantizers. These coders can be divided into three classes which are distinguished by the way in which the information from the previous line is used to "anticipate" or predict the value of the current pel. In Fig. 1 we give a labelled diagram of the pels near the current pel.

In the first class of coder, the prediction is the average of the value of the previous pel, A, and the value of one or more of the vertically adjacent pels, B, C, and D. We call this "averaged prediction." In the second class of coder, the prediction is the sum of A and the scaled horizontal element difference or slope between elements on the previous line, e.g., $C - B$ or $\frac{1}{2}(D - B)$. This has been called "planar prediction."[4]

These first two classes of coders use linear prediction since the prediction is a linear sum of previously coded intensities. The third class of coder uses a nonlinear prediction in that the choice of prediction mode is signal-dependent. The prediction can be either A or some linear combination of A, B, C and/or D. The choice is dependent on the magnitude of intensity differences between the pels in the previously coded neighborhood of the current pel, X. This class of coder can be called "optional prediction." (R. E. Graham has proposed and simulated some coders of this type.[5])

In the following sections we describe the particular coders we have built, and give some examples of processed pictures. We also include a discussion of the effect of transmission errors for the various coding strategies.

II. DESCRIPTION

The scan format is similar to that used in *Picturephone* Service: there are 271 scan lines with a 2:1 interlace, 248 elements per line and the frame rate is 30 Hz.

A schematic of a general differential quantizer that uses information from the preceding line is given in Fig. 2. The quantizer characteristics are shown on Table I. The particular coders to be described below differ only in the implementation of the predictor. Note that with the use of sign prediction[6] these are three-bit coders.

For comparison, we chose as a reference an element-coder (i.e., the

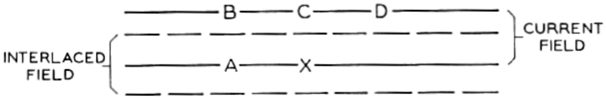


Fig. 1—Diagram of picture elements (pels) near the current pel.

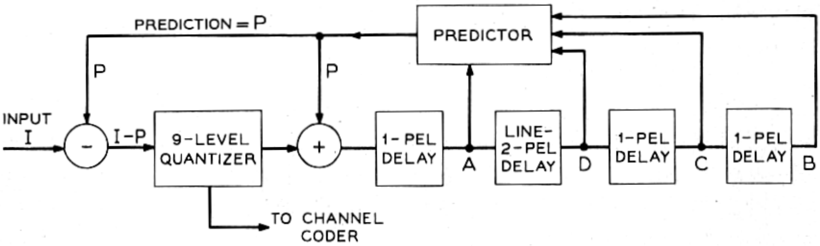


Fig. 2—Diagram of differential quantizer using information in previous line to aid prediction of current pel. The letters *A, B, C* and *D* refer to the pels whose positions are shown in Fig. 1.

value of pel *A* is used as a prediction) having the same quantizer characteristics as above. Although it might be argued that it would make more sense to optimize the quantizing scale for each type of prediction, we found that the subjectively optimum quantizing scale varied as much with picture content as with prediction strategy. For most experiments, therefore, we used one quantizing scale; a few results on varying the quantizing scale are, however, also described in the next section.

In the subjective tests, the 5½ inch by 5 inch raster was viewed through a polarizing screen from a distance of 36 inches in a room having average office illumination (70 foot candles). Two scenes were used: (*i*) a model head, "Penelope," with a draped background (Fig. 3a); and (*ii*) a photographic slide, "Karen," (Fig. 3b).

III. AVERAGED PREDICTION

In this type of coder, the prediction is formed by adding the value of pel *A* to the value of some combination of the pels *B, C* and *D*. Three coders were built and tested; the predictions used were: (*i*) $(A + D)/2$, (*ii*) $(A + C)/2$, (*iii*) $[A + (C + D)/2]/2$. Figure 4a shows a diagram of the arrangement used to test coder (*i*).

In the absence of channel errors, the displayed picture at the trans-

Table I—Encoder Characteristics

Input Levels	±0–1	2–7	8–17	18–33	34–255
Output Levels	±0	4	11	25	42

mitter is the same as at the receiver[6] and hence only the transmitter loop was built for the real time tests.

In comparison with the reference coder, all three of the coders gave significantly improved rendition of the vertical and most of the sloping edges. Horizontal edges and those sloping slightly upward appear blurred in comparison with the same edges encoded by the element coder. However, it is a static blurring which is subjectively much less annoying than the "edge busyness" which occurs on some edges encoded by the element coder. The granular noise in pictures encoded with the averaged prediction coders is associated more with high vertical spatial frequencies than with high horizontal spatial frequencies. This effect is to be expected because the relatively large vertical step size will result in more frequent use of the larger (and less accurate) quantized difference signals in regions containing vertical detail. In picture areas with no directional properties, the noise level is about the same as with the element coder. Some observers, however, mentioned that it appeared to be a higher frequency noise.

In comparisons among the "averaged prediction" coders, (*i*) and (*iii*) appeared preferable to (*ii*) since they were able to adequately encode edges with smaller positive slopes than was (*ii*). This is understandable since the presence of pel D in their prediction means that such edges are anticipated. There was very little to choose between coders (*i*) and (*iii*) and so we carried out subjective tests in which eight unskilled and eight skilled observers were asked to state a preference between pictures obtained using coder (*i*) and the reference coder. With "Penelope," all eight of the skilled and seven of the eight unskilled observers preferred the picture given by coder (*i*). With

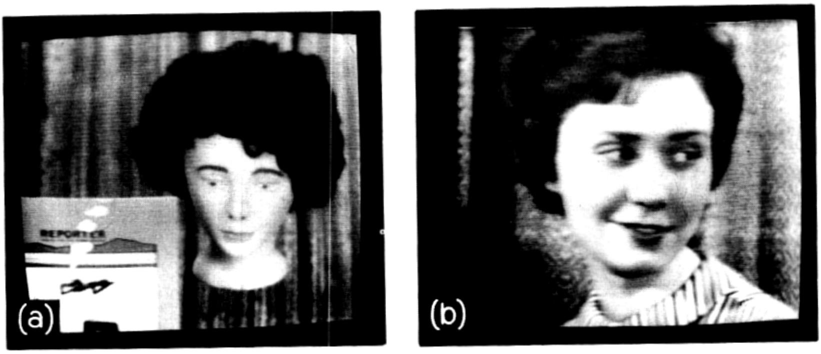
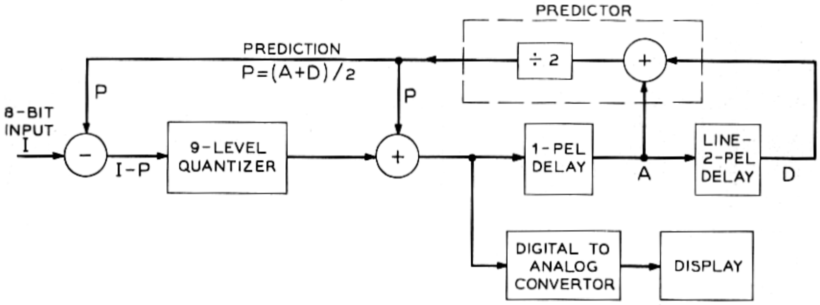
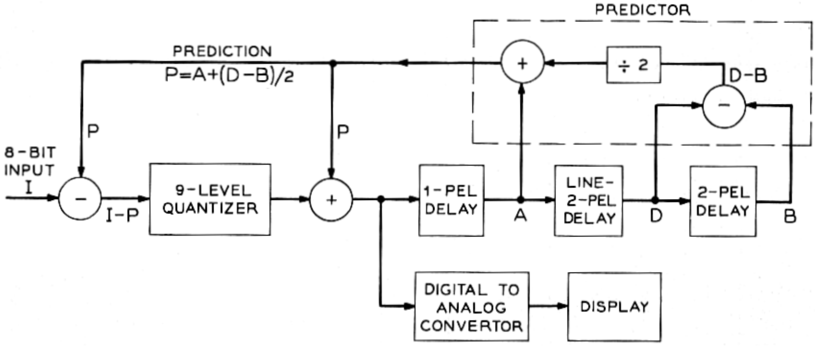


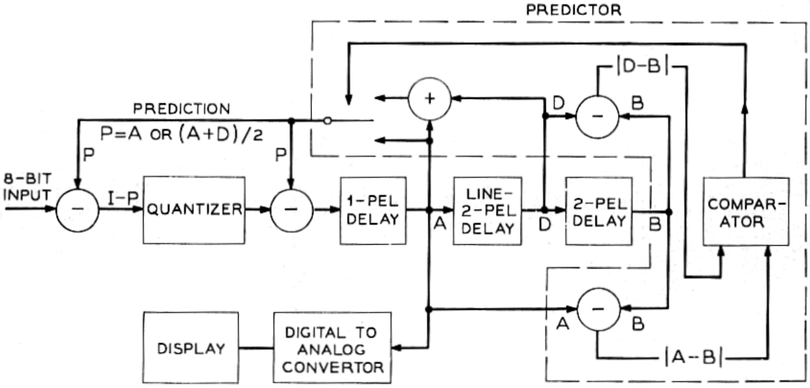
Fig. 3—Scenes viewed for evaluating coders, (a) Penelope and (b) Karen.



(a)



(b)



(c)

Fig. 4a—Diagram of apparatus used to evaluate coder (*i*) in which the value $(A + D)/2$ is used as a prediction of the current picture element. The displayed picture is equivalent to the received picture providing there are no channel errors.

Fig. 4b—Apparatus used to evaluate coder (*v*) in which $A + (D - B)/2$ is used as a prediction of the current pel. The position of pels A, B and D is shown in Fig. 1.

Fig. 4c—Apparatus used to test the optional prediction encoder. If $|A - B| > |D - B|$, A is used as the prediction. If $|A - B| \leqq |D - B|$, $(A + D)/2$ is used. The positions of A, B and D are shown on Fig. 1.

"Karen," all subjects preferred the picture given by coder (*i*). (The skilled observers all remarked on the better edge rendition of the vertical stripes in Karen's blouse.) As the quantizing scale is made coarser, there is improved rendition of vertical edges in the picture from the reference coder but the granular noise also increases. In this situation, the comparison between the two coders is closer, but coder (*i*) was still preferable.

The element coder and coders (*i*), (*ii*) and (*iii*) were simulated on a DDP-224 computer system. Only a single frame was processed so it was not possible to judge frame to frame noise. In the first set of experiments the magnitude of the difference between the original picture (coded as eight-bit PCM) and pictures from the above coders was displayed. Figures 5a and 5b show the discrepancy between the PCM coded picture and pictures coded with the element coder and coder (*i*) respectively. (Figures 5c and d are for coders discussed

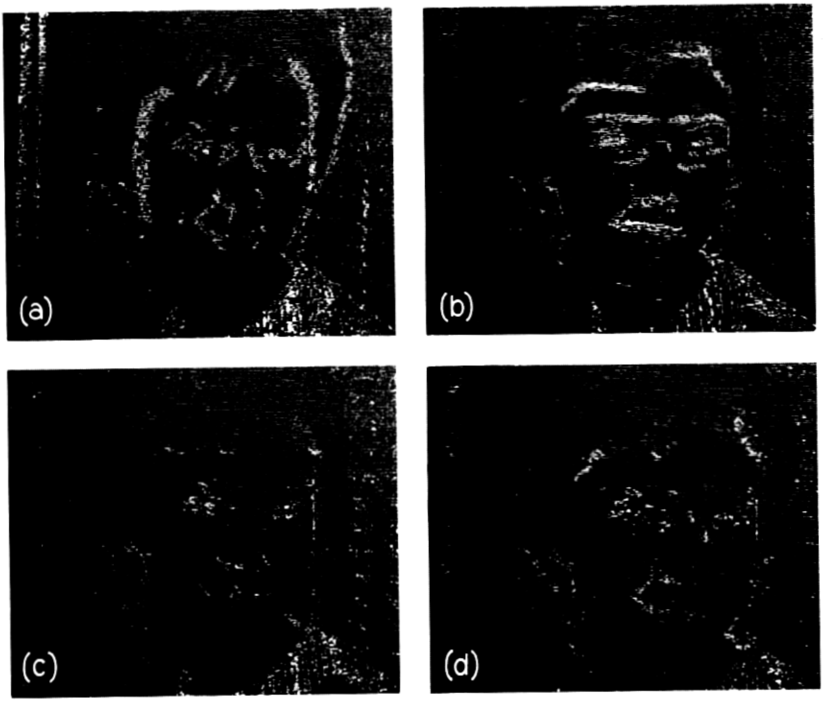


Fig. 5—Magnitude of the difference between a picture coded as 8-bit PCM and: (a) the reference coder; (b) coder (*i*) ($[A + D]/2$ is prediction); (c) coder (*v*) ($[A + (D - B)/2]$ is prediction); (d) optional prediction coder of Fig. 4c.

below.) As expected, the element coder showed differences along vertical and diagonal edges. The average prediction coders, on the other hand, gave differences chiefly along horizontal edges. Coder (*iii*) seemed to show the least discrepancy from the eight-bit coded picture. No subjective improvement was realized by choosing coder (*iii*) over coder (*i*) in the experimental model.

We also investigated on the computer the effect of transmission errors on a single frame of the received picture. It has been suggested that because the average prediction coders made use of information from the previous line as well as from previous pels on the current line, the effect of an error on subsequent lines would be serious. However, because the prediction is an average, the effect of the error decays spatially. For example, consider the effects of an error of 64 quantum levels in coder (*i*). This error results in the array of intensity errors shown in Table II. Note that the effect of the error decays rapidly both horizontally and diagonally, but more slowly vertically. Also, the spatial derivative of the error pattern is small except near the original error. This fact should help reduce the visibility of the error pattern. In fact, the pictures generated on the computer showed that even serious errors (e.g., a change of sign on the outer level) resulted in a pattern that was scarcely visible.

Contrast the visibility of the errors in Figs. 6a and b. The picture in Fig. 6a was coded with the reference coder. The upper, dark line was caused by a positive-to-negative sign change in an outer quantizer level. The lower light line results from a negative-to-positive sign change in the third quantizer level. The identical type of errors are present in the picture in Fig. 6b which was coded with coder (*i*). The error in the outer quantizer level is visible as a small black dot at the outside corner of the left-hand eye. The error in the third quantizer level appears as a white streak toward the right end of the upper lip. In any event, the errors are far less visible than the equivalent errors in Fig. 6a.

IV. PLANAR PREDICTION ENCODERS

Two coders were built and tested. The predictions used were (*iv*) $A + (C - B)$, and (*v*) $A + (D - B)/2$. Figure 4(b) shows a schematic for coder (*v*).

Both coders showed improvement over the reference coder, but not over the coders in the previous section. In particular, the planar prediction encoders gave rise to "edge-busyness" on edges sloping upward

TABLE II—INTENSITY ERRORS FOR AVERAGE PREDICTION ENCODER*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Horizontal Distance in Pels →x	
1								64	32	16	8	4	2	1			
2																	
3							32	32	24	16	10	6	3	1			
4																	
5						16	24	24	20	15	10	7	4	2	1		
6																	
7					8	16	20	20	18	15	11	8	5	3	1		
8																	
9				4	10	15	18	18	16	14	11	8	6	3	2	1	
10																	
11			1	5	10	14	16	16	15	13	11	8	5	3	2	1	
12																	
13			3	7	10	14	15	15	14	12	10	7	5	3	2	1	
14																	
15		2	4	7	11	13	14	14	13	11	9	7	5	4	2	1	
16																	
17	1	2	5	8	10	12	13	13	12	10	9	7	6	4	3	2	1
Vertical 18																	
Distance in Lines 19	1	3	6	8	10	12	12	12	11	10	8	7	6	4	3	2	1

y ↓

---

* The pattern of intensity errors that results from an error of 64 levels at pel 8,1 ($x = 8$, $y = 1$) when $(A + D)/2$ is used as a prediction of the current pel. The intensity of the error decays spatially and, perhaps more importantly, there is no abrupt large spatial difference in intensity except near pel 8, 1.

Fig. 6—Decoded pictures showing the effect of intensity errors resulting from channel errors. Figures (a) and (b) each contain two error patterns resulting from errors of 84 and 50 levels (out of 255 levels). (a) Derived using the reference coder (differential quantizer using A as the prediction); (b) results from the use of coder (*i*) ($[A + D]/2$ is prediction); (c) results from the use of coder (*v*) ($[A + (D - B)/2]$ is prediction) when there is a single intensity error of 84 levels; (d) results from the use of the optional predictor (see Fig. 4c) when there is a single intensity error of 50 levels. Copies of the original prints may be obtained from the authors.

to the left. In addition, coder (*iv*) gave rise to the same problem on edges sloping upward to the right.

Consider an edge passing to the left of B in Fig. 1 and between A and X so that $(C - B) = (D - B) = 0$. The values of $A + (C - B)$ and $A + (D - B)/2$ will both be poor predictions of X (the same as the element coder but worse than $(A + D)/2$.) Also, if an edge passes between A and X and between D and C, then $A + (C - B)$ will be a poor prediction of X.

The discrepancy between the picture coded as eight-bit PCM and that coded with coder (*v*) is shown in Fig. 5c. Although on the average, the picture shows less discrepancies than the equivalent

TABLE III—INTENSITY ERRORS FOR PLANER PREDICTIVE ENCODER*

Horizontal Distance in Pels →

Start of Blanked Region ↓

							210	211	212	213	214	215	216	217	218	219	220	221	222	223
1							0	0	64	64	64	64	64	64	64	64	64	64	64	0
3							0	32	64	64	64	64	64	64	64	64	64	64	32	0
5						0	16	48	64	64	64	64	64	64	64	64	64	48	16	0
7					0	8	32	56	64	64	64	64	64	64	64	64	56	32	8	0
9					4	20	44	60	64	64	64	64	64	64	64	60	44	20	4	0
11				2	12	32	52	62	64	64	64	64	64	64	62	52	32	12	2	0
13			1	7	22	42	57	63	64	64	64	64	64	63	57	42	22	7	1	0
15		0	4	14	32	50	61	64	64	64	64	64	64	61	50	32	14	4	0	0
17	0	2	9	23	41	56	63	64	64	64	64	64	63	56	41	23	9	2	0	0
19	1	6	17	33	49	60	64	64	64	64	64	64	60	49	33	17	6	1	0	0
21	3	11	25	4	54	62	64	64	64	64	64	62	54	41	25	11	3	0	0	0

y ↓ Vertical Distance in Lines
(unlike Table II, the interlaced lines are not shown)

---

* Pattern of intensity errors resulting from an error of 64 levels. At pel 212, 1 when A + (D − B)/2 is used as a prediction of the current pel. Note that the intensity error is constant to the right of pel 212, 1 until the start of the blanked portion of the line (pel 223, 1) when the accumulated value is forced to zero. On successive lines this (correct) reset value is partially used in the prediction near the right-hand side and so the error is reduced.

picture for coder (i), the discrepancies that do occur lie on vertical or diagonal edges sloping upward to the left. It is on just these edges that the "edge busyness" occurs.

The effect of a transmission error is an area of uniformly altered brightness as shown in Fig. 6c. How this pattern arises is illustrated in Table III where we have assumed an error of +64 levels in the element in the top row of numbers.

In view of the increased visibility of error patterns and the absence of a significant improvement in picture quality over the averaged prediction encoder, we saw no overall advantage in the planar prediction encoders for fixed length codes.

V. OPTIONAL PREDICTION ENCODERS

A number of varieties of optional prediction encoders were investigated. The one examined most closely uses either A or $(A + D)/2$ as a prediction. If the difference $|A - B|$ is greater than the difference $|D - B|$, A is used as the prediction, otherwise $(A + D)/2$ is used (Fig. 4c). These differences are made up of pels already coded and available (in the absence of transmission errors) at both the receiver and transmitter. Thus, no extra information need be sent to the receiver to indicate the prediction used by the transmitter.

The pictures obtained showed less granular noise than those from any of the other encoders. They showed little deviation (Fig. 5d) from the pictures coded as eight-bit PCM. Most edges are reproduced well, but those sloping at about 45 degrees upward to the left showed some local slope overload due to a mistaken choice of prediction. Unfortunately, a mistaken choice on one line tended to lead to similar mistakes on the next line and hence such clusters of mistakes tended to be visible. Different strategies for making the choice altered the slope of the edges on which the trouble occurred. So far only simple strategies have been tried; more complicated ones may well eliminate the problem. Obviously, there is a wide variety of techniques to be explored. For instance, there can be more than two options. In view of the low granular noise of the encoded pictures, such exploration is well worthwhile.

The effect of a transmission error on the decoded picture can be catastrophic. The receiver can lose track of the prediction option selected by the transmitter. Indeed, the effect of a large error tends to make the receiver choose incorrectly. This is illustrated by supposing the pel D has an error of +64 levels. This generates a large spatial change; $|D - B|$ is large and hence $(A + D)/2$ will probably be used as the prediction. This propagates the error to the next line. If A should have been used as a prediction instead of $(A + D)/2$, then the differences are added to the wrong prediction at the receiver. The effect on the decoded picture can be catastrophic as is illustrated in Fig. 6d.

VI. DISCUSSION

Because of the improved picture quality and the insensitivity to transmission errors compared with a standard differential quantizer, coders (*i*) and (*iii*) seem at present to be the most practical for intrafield coding of pictures for transmission. This is especially true when fixed length codes are to be used. We have still to determine the best quantizing characteristics and the best weighting of pel A with pels on the previous line to form a prediction.

Except for a few poorly predicted pels, the optional prediction encoders give a better rendering of the scenes described than do the fixed prediction encoders [e.g., coders (*i*) through (*v*)]. Hence we are continuing to study ways of removing the last few flaws in the coded picture. The sensitivity of this type of coder to transmission errors may be a serious problem. We are considering a number of methods for reducing this sensitivity.

REFERENCES

1. Kretzmer, E. R., "Statistics of Television Signals," B.S.T.J., *31,* No. 4 (July 1952), pp. 751–763.
2. O'Neal, J. B., Jr., "Predictive Quantizing Systems (Differential Pulse Code Modulation) for the Transmission of Television Signals," B.S.T.J., *45,* No. 5 (May–June 1966), pp. 689–721.
3. Limb, J. O., "Entropy of Quantized Television Signals," Proc. IEE, *115,* No. 1 (January 1968), pp. 16–20.
4. Harrison, C. W., "Experiments with Linear Prediction in Television," B.S.T.J., *31,* No. 4 (July 1952), pp. 764–784.
5. Graham, R. E., "Predictive Quantizing of Television Signals," IRE WESCON Record, *2,* Part 4, (1958), pp. 147–157.
6. Limb, J. O., and F. W. Mounts, "A Digital Differential Quantizer for Television," B.S.T.J., *48,* No. 7 (September 1969), pp. 2583–2599.