IP-1 | 897

Military Technical College
CAIRO - EGYPT

PERFORMANCE EVALUATION OF DPCM FOR AERIAL IMAGES

BY

Brig.A.A.SHAHIN(Ph.D)
Brig.M.S.GHONIEMY(Ph.D)
Brig.D.M.I.ELRAYESS(Ph.D)
Col.A.A.ELBATTA

ABSTRACT

Image coding(compression) is useful in image transmission and storage where the aim is to minimize the bandwidth for transmission and memory for storage.This paper presents a review of techniques used for coding of digital images.A short review of transform coding,pulse code modulation,and hybrid coding techniq- ues is introduced. Predictive coding is presented,analysed,simu- lated,and tested for the image of Minea area (Egypt) and the results are displayed.

INTRODUCTION

Image coding is concerned with the minimization of the number of information carrying units used to represent an image. Image compression is useful in image transmission and storage where the aim is to minimize the bandwidth for transmission and memory for storage. The efficiency of any coding algorithm is measured by:its data compressing ability, the resulting distorsion of decoded image, and its implementation complexity. Image coding techniques comprise:Transform coding,pulse code modulation(PCM), predictive coding(also known as DPCM), hybrid coding, and miscellaneous techniques that do not fall into any of the above categories.

This paper contains six sections. Section 1 is a short review of transform coding techniques. Section 2 introduces pulse code modulation(PCM) techniques. Section 3 is a detailed analysis of DPCM technique optimized to achieve the best reconstruction of decoded image. Section 4 presents a short review of hybrid coding. Section 5 displays the simulation and results of coding the picture of Minea area using DPCM technique with different number of bits per picture element. The paper is terminated by a conclusion with a short reference to contour coding of images as an introduction to the second generation image coding techniques which achieve high compression ratio.

1.TRANSFORM CODING

In transform coding, the picture is divided into sub-pictures and then each of these sub-pictures is transformed into a set of more independent coefficients.The coefficients are then quantized and decoded for transmission. At the receiver, the received bits are decoded,and inverse transformed to recover intensities of picture elements. Achieved compression results from dropping smaller coefficients and coarsely quantizing the others as required by the picture quality.Performance of a transform coder is determined by the

shape and size of the sub-pictures, type of transformation used, selection of the coefficients to be transmitted and quantization of them, and the bit assignor which assigns a binary word for each of the quantizer outputs. The optimum transform must achieve smallest mean-square reconstruction error, maximum energy compaction in the smallest number of coefficients and result in uncorrelated transform coefficients. The transform that achieves these requirements is the Karhunen-Loeve transform(KLT). The optimum transform is computed from the covariance function of the pel vector U(in one dimensional case).

$$Cu=E[(U-E(U)).(U-E(U)^T] \quad (1)$$

where E is the statistical expectation and T denotes transpose. Rows of the optimum transform matrix are normalized eigenvectors of the matrix Cu,i.e.they are solutions of the equation:

$$Cu.U=\lambda_i.U \quad (2)$$

where λ_i's are the corresponding eigenvalues.

Some of the problems encountered in this case is that the covariance function is not stationary, and one must choose different covariance matrices matched with different regions of the image. Another approach is to use an average covariance matrix but in this case the performance is highly degraded. The second problem is the computation of eigenvectors of Cu which can be singular. The third problem is the quantity of mathematical computations required to compute the transformation (N^2 complex operations are required to compute one transform coefficient, where N is number of pixels per block). These problems have led to the use of sub-optimal transforms satisfying both good performance and mathematical tractability. Usually these transformations are unitary. Unitary transforms preserve image energy and compact it in a few number of coefficients (usually low frequency).

From these transforms we distiguish:

FOURIER TRANSFORM

$$\Phi(u,v)= \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} u(j,k).\exp[\frac{2\pi i}{N} (uj+vk)] \qquad (3)$$

where u(j,k) is the (j,k) element of U.

COSINE TRANSFORM

$$\Phi(u,v)= \frac{2}{N} c(u).c(v). \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} u(j,k).\cos(\frac{\pi}{N} u(j+.5)) .\cos(\frac{\pi}{N} v(k+.5)) \qquad (4)$$

where

$$c(u)= \frac{1}{\sqrt{2}} \quad \text{for } u=0$$

$$= 1 \quad \text{otherwise}$$

$$c(v)= \frac{1}{\sqrt{2}} \quad \text{for } v=0$$

$$= 1 \quad \text{otherwise}$$

SINE TRANSFORM

$$\Phi(u,v)=\frac{2}{(N+1)} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} u(j,k).\sin[\frac{(j+1)(u+1)\pi}{(N+1)}].\sin[\frac{(k+1)(u+1)\pi}{(N+1)}] \qquad (5)$$

HADAMARD TRANSFORM

$$\Phi(u,v)= \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} u(j,k).(-1)^{p(j,k,u,v)} \qquad (6\text{-a})$$

where

$$p(j,k,u,v)=\sum_{i=0}^{n-1} (u_i.j_i + v_i.k_i) \qquad (6\text{-b})$$

$n=\log_2(N)$, and u_i , v_i, j_i , and k_i are the bit states of the binary representation of u,v,j, and k, respectively.

Obtaining the coefficients,a sample selection criterion is used to choose the coefficients to be transmitted. There are two criteria for sample selection:zonal sampling and threshold sampling. In zonal sampling, some region is chosen and all coefficients in this region are transmitted. The best zone chosen is that based on the variance of the coefficients where the coefficients with larger variances are selected [1]. In threshold sampling, samples greater than a specified threshold are chosen. The chosen samples are then quantized separately. The number of bits allocated to each coefficient is chosen in proportion to its variance such that the total number of bits per sub-picture is constant. Then output levels of the quantizer are coded and transmitted. Transform coding has good performance in the presence of channel errors. If a coefficient is decoded erroneously at the receiver due to a transmission error, on inverse transformation only pels in the same block are affected. As the block size decreases the averaging property decreases, and the transmission errors appear as blotches in the reconstructed picture.Fig.1 shows a simplified block diagram of transform coding.

2.PULSE CODE MODULATION

Image coding by PCM is nothing more than a time discrete and amplitude discrete representation of the signal. Time discretization is obtained by sampling the picture. Amplitude discretization can be obtained by quantizing each sample using N levels. Each level is represented by a binary word of B bits, where $B = \log_2 N$. In the decoder , these binary words are converted to discrete amplitude levels, and then the time sequence of the amplitude levels is low-pass filtered. Basic PCM encoder affords an uncommon simplicity compared to other encoders, but suffers from inefficiency since it does not use redundancy present in the picture signal.

For good quality original pictures,as the number of quantization levels is decreased, quantization errors are seen

as false contours. Visibility of this noise can be decreased by adding some high frequency noise to the original signal before quantization. This noise causes the coded signal to oscillate between the quantizing levels, thereby increasing the frequency content of output noise. PCM requires 128-256 quantizing levels for good picture quality.

3.PREDICTIVE CODING

In basic predictive coding,a prediction of the sample to be encoded is made from previously coded information that has been encoded (Fig.2). The error resulting from the subtraction of the prediction from the actual value of the sample is quantized into a set of discrete amplitude levels either of fixed or variable word length and sent to the channel for transmission. Thus the encoder has three components: Predictor,quantizer,and code assignor. Optimization of these components for picture coding will be discussed. Predictors can be classified as linear or nonlinear depending upon whether the prediction is a linear function of the previously transmitted samples or not. Also they can be classified according to the location of previous elements used: One-dimensional predictors use previous elements in the same line, two-dimensional predictors use elements in the previous line as well, whereas interframe predictors use picture elements from previously transmitted frame(in case of TV). Predictors can be fixed or adaptive. Fixed predictors maintain the same characteristics independent of the data, whereas adaptive predictors change their characteristics as a function of data.

Let $[u_{i,j}]$ be a set of picture elements indexed according to their spatial location as shown in Fig.3. Let R(i,j,p,q) be the autocorrelation function of the real random field to which the picture belongs :

$$R(i,j,p,q)=E[u(i,j)u(p,q)] \qquad (7)$$

If the three nearest neighboring elements that have already

been transmitted are used to form the linear estimate for a picture element u(i,j).

$$\hat{u}(i,j)=a_1u(i-1,j)+a_2u(i-1,j-1)+a_3u(i,j-1) \qquad (8)$$

where the unknowns a_1, a_2 and a_3 are chosen such that the mean square estimation error

$$e^2 = E[(u(i,j)-\hat{u}(i,j))^2] \qquad (9)$$

is minimized.

Substituting (8) in (9), differentiating the resulting expression with respect to a_1, a_2 and a_3, separately, equating each derivative to zero and using (7) we obtain the following equations :

$$a_1.R(i-1,j;i-1,j)+a_2.R(i-1,j-1;i-1,j)+a_3.R(i,j-1;i-1,j) = R(i,j;i-1,j) \qquad (10\text{-a})$$

$$a_1.R(i-1,j;i-1,j-1)+a_2.R(i-1,j-1;i-1,j-1)+a3.R(i,j-1;i-1,j-1) = R(i,j;i-1,j-1) \qquad (10\text{-b})$$

$$a_1.R(i-1,j;i,j-1)+a_2.R(i-1,j-1;i-1,j-1)+a_3.R(i,j-1;i-1,j-1) = R(i,j;i,j-1) \qquad (10\text{-c})$$

If the random field is assumed to be homogeneous, with zero mean and autocorrelation function given by :

$$R(x,y)=R(0,0).\exp[-c_1|x|-c_2|y|] \qquad (11)$$

then (10) reduces to :

$$a_1.R(0,0)+a_2.R(0,1)+a_1.R(1,1) = R(1,0) \qquad (12\text{-a})$$

$$a_1.R(0,1)+a_2.R(0,0)+a_3.R(1,0) = R(1,1) \qquad (12\text{-b})$$

$$a_1.R(1,1)+a_2.R(1,0)+a_3.R(0,0) = R(0,1) \qquad (12\text{-c})$$

From (11) , $R(1,1)=R(1,0)R(0,1)/R(0,0)$. Using this

relationship and solving (12) for a_1 , a_2 , a_3 we obtain

$$a_1 = R(1,0)/R(0,0) \qquad (13\text{-a})$$

$$a_2 = -R(1,1)/R(0,0) \qquad (13\text{-b})$$

$$a_3 = R(0,1)/R(0,0) \qquad (13\text{-c})$$

The error signal e(i,j) at each picture element is given by

$$e(i,j) = u(i,j)-\hat{u}(i,j)$$

$$= u(i,j)-[a_1.u(i-1,j)+a_2.u(i-1,j-1)+a_3.u(i,j-1)] \qquad (14)$$

It is clear that if u has zero mean, then e has zero mean also, and the variance of the error signal is given by

$$E[e^2(i,j)]=E\{[u(i,j)-(a_1.u(i-1,j)+a_2.u(i-1,j-1)+a_3.u(i,j-1))]^2\} \qquad (15)$$

Expanding the right hand side of (15), using (7),(13) and the assumption of homogenity we get

$$\sigma^2 = E[e2(i,j)]$$

$$= R(0,0)-[R^2(1,0)/R(0,0)+R^2(0,1)/R(0,0)-R^2(1,1)/R(0,0)] \qquad (16)$$

It is clear from (16) that the variance of the error signal is smaller than the variance of the picture elements given by R(0,0).

Returning to (8) and assuming that the prediction is made from all the previous elements, substituting into (9) and differentiating, we obtain

$$E[(u(i,j)-\hat{u}(i,j)).u(p,q)]=0 \qquad (17)$$

for all u(p,q) used in the calculation of u(i,j).

$$E[e(i,j).e(p,q)]=E[(u(i,j)-\hat{u}(i,j)).(u(p,q)-\hat{u}(p,q))]$$

$$=E[(u(i,j)-\hat{u}(i,j)).u(p,q)]-E[(u(i,j)-\hat{u}(i,j)).\hat{u}(p,q)]$$

$$= 0 \qquad (18)$$

From the orthogonality principle of linear estimation the first term in the right hand side is zero, the second term is also zero because if u(p,q) is used in estimation of u(i,j), then all points used in estimation of u(p,q) will also be used in estimation of u(i,j) and the estimation error will also be orthogonal to u(p,q). In the special case of representing the picture by a wide sense Markov process, then

$$E[e(i,j).e(p,q)]=0 \qquad (19)$$

if only three neighboring elements are used in the estimation. This implies that the estimation errors are uncorrelated, which is the first step needed for image compression.

In DPCM coding systems it has been found that the prediction error can be well modeled by laplacian density of the form

$$p(e) = \frac{1}{\sqrt{2}\sigma_e} \exp\left(-\frac{\sqrt{2}|e|}{\sigma_e}\right) \qquad (20)$$

where σ_e is the standard deviation of the prediction error [6].

The MSE has been computed by Habibi,[2], for different numbers of picture elements within a frame. His results,show that if the predictor coefficients are matched to statistics of a picture, then for that picture the MSE decreases significantly by using up to three picture elements, further decreases are small. This is shown in Fig.4. If the coefficients are not exactly matched, the decrease in MSE is not significant. The improvement in MSE by using two-dimensional prediction is small, but subjectively the rendition of vertical edges due to it is significantly improved [3].

Because of the nonstationarity of picture statistics, it will

be advantageous to change the prediction based on the local properties of the picture signal. For the intraframe case, the popular method is that of Graham [1] where the horizontal and vertical correlations are computed and the predictor is chosen in the direction of largest correlation.This is shown in Fig.5.

From point of view of transmission errors, it is advantageous in one-dimensional case to use optimum predictor due to its stability, so propagation of transmission errors will decay rapidly. Previous element predictor is usually unstable,[4], and to avoid bad effects of transmission errors, a leak factor is used in the predictor.In case of two-dimensional predictor,the optimum predictor is usually unstable for most correlation functions used [4]. In this case the simplest method to overcome bad effects of transmission errors is to substitute the erroneous line by the previous line. In case of adaptive predictor, transmission errors has two effects:use of wrong value for prediction, and the other is the wrong selection of predictor. The best solution in this case is to use error detecting and correcting codes [1].

DPCM schemes achieve compression, mainly due to non quantizing the prediction error as finely as the original image. Three types of degradations can be seen due to improper design of DPCM quantizer. These are referred to as granular noise, slope overload and edge busyness. Granular noise is caused by coarse quantization of flat areas in the picture. Slope overload is a result of small dynamic range of the quantizer(largest quantizer step) which prevents the quantizer from rapid response to high contrast edges. For edges whose contrast changes gradually, the quantizer output oscillates around the signal from line to line giving the appearance of a " busy edge". This is shown in Fig.6. To avoid some of these degradations,optimum quantizer is used. Parameters of this quantizer are chosen to minimize the average mean square quantization error. Considering Fig.7, if u is the input to the quantizer with probability density p(u), then

$$u_j = (y_j+y_{j-1})/2 \tag{21}$$

and

$$y_j = \int_{u_j}^{u_{j+1}} u.p(u).du \; / \; \int_{u_j}^{u_{j+1}} p(u).du \tag{22}$$

where $u_1<u_2<\ldots<u_{N+1}$ and $y_1<y_1<\ldots<y_N$ are the decision and reconstruction levels of the quantizer. Solution of (21) and (22) is performed recursively by chosing y_1 somewhat larger than u_1, calculating u_2 from (22) and substituting u into (21)to compute y_2, and so on until all u_j's and y_j's are determined.

The mean-square quantization error q can be expressed in terms of the probability density of the prediction error signal as

$$\sigma^2_q = \sigma^2_e - \sum_{j=1}^{N} y^2_j \int_{u_j}^{u_{j+1}} p(u).du \tag{23}$$

where σ_e is the variance of error signal as specified in (16).

To adapt the quantizer to image data, the innovation $\delta(i,j)$ given by

$$\delta(i,j)=u(i,j)-\hat{u}(i,j) \tag{24}$$

is used as a training data for the quantizer Fig.8. The standard deviation and the mean of the innovation sequence δ are used to normalize the innovation sequence to a zero mean unity variance sequence in order to use Table.6-1 that gives reconstruction and decision levels of optimum quantizer [6] .

4.HYPRID CODING

Since transform coding has good performance in the presence of channel errors but its implementation is complex, while DPCM has the advantage of simple implementation but lower

performance in the presence of channel errors, it will be advantageous to combine both systems in a hybrid one making use of both advantages simultaneously. In this case the picture is divided into small sub-blocks(either one or two dimensional).These small blocks are used in evaluating transform coefficients, and performing DPCM of the coefficients using the previously transmitted blocks as predictions. A block diagram of hybrid coder is shown in Fig.8.

5. SIMULATION AND COMPUTER RESULTS

In this paper both the mean-square reconstruction error and subjective quality performance of 3 types of adaptive DPCM predictors with optimum quantizer are compared. These are : Optimum 1-d system, Optimum 2-d system, and previous element system.

The decoded pictures resulting from these three systems are compared subjectively with that resulting from a previous element system that uses an uniform quantizer. The quantizer performance is compared with the rate distorsion function of the prediction error.

The system used is the remote sensing image processing system RIPS which consists of cromemco personal computer, a terminal, a digitizer, and a 256 grey level video monitor .

The picture used is that of Minea area in Egypt taken by LANDSAT satellite. The picture is digitized into 240 x 256 pixels. Each pixel of this array is represented by 8 bits .

The algorithm used to implement this task consists of the following steps :

1. Compute the horizontal and vertical correlation factors of the picture in use by using

$$\rho_h = \sum_{i=1}^{240} \sum_{j=1}^{255} [u(i,j).u(i,j+1)/u^2(i,j)]/61200$$

$$\rho_v = \sum_{i=1}^{239} \sum_{j=1}^{256} [u(i,j).u(i+1,j)/u^2]/61184$$

For the picture used $\rho_h = \rho_v = 0.4845$

2. Compute the innovation sequence δ using (24) . The estimate u is calculated from (8) by using

$$a_1 = \rho_h$$

$$a_2 = -\rho_h \rho_v$$

$$a_3 = \rho_v$$

3. Compute the mean m and the variance σ^2 of the sequence by using

$$m = \sum_{i=1}^{N} \sum_{j=1}^{M} \delta(i,j)/NM$$

$$\sigma^2 = \sum_{i=1}^{N} \sum_{j=1}^{M} [\ (i,j)-m]2/NM$$

4. Use m and σ^2 to normalize the sequence δ. The formula used is

$$\bar{\delta}(i,j)=(\delta(i,j)-m)/\sigma$$

5. The quantizer decision and reconstruction levels are taken from Table.6-1 according to the number of bits used in the quantizer.

6. Compute the prediction error sequence using

$$e(i,j)=u(i,j)-\hat{u}(i,j)$$

where $\hat{u}(i,j)$ is the estimation from previously transmitted samples. Quantize this error, assign a code word to it, and transmit .

The reproduced value $\dot{u}(i,j)$ is computed from

$$\dot{u}(i,j) = \hat{u}(i,j) + \dot{e}(i,j)$$

where $\dot{e}(i,j)$ is the quantized value of $e(i,j)$. This value will be used in calculation of the predicted value of the next sample(s) using (8) .

7. Calculate the mean-square reconstruction error (MSRE) using

$$MSRE = \sum_{i=1}^{N} \sum_{j=1}^{M} (u(i,j) - \dot{u}(i,j))^2 / NM$$

8. Compute the variance of the prediction error using (16), and the variance of the quantization error using (23)

9. Compute the rate distortion function R(D) of the prediction error sequence (it's elements are uncorrelated) using [7]

$$R(D) = \frac{1}{2} \log_2 (\sigma^2_e / \sigma^2_q)$$

6. CONCLUSION

A review of image coding techniques has been presented. Computer simulation of 4 different DPCM systems is performed. The results indicate better performance in using adaptive optimum quantizer in the feedback loop of the encoder .

The results of computer simulation indicated the following

1. The optimum quantizer performed approximately 1 bit behind the rate distortion function .

2. There are no significant differences in the subjective quality of the decoded picture between the 3 predictors which contain an optimum quantizer. This is a result of the nature of the used picture which contains large areas of slowly varying grey levels, requiring only 1 bit for coding when an optimum quantizer is used in the feedback loop of the coder. From point of view of mean-square reconstruction error, the 2-d optimum encoder has the smallest one while the previous element encoder has the largest one .

3. The subjective quality of decoded picture using optimum

quantizer in the feedback loop is very much better than those resulting when non optimum quantizer is used. The quantizer degradations are very clear in the last case. Image segmentation and contour coding is under investigation as high compression coding techniques. In this technique, image is segmented into adjacent regions and the signal in each region is approximated by two dimensional polynomial function. Polynomial coefficients are chosen to minimize the mean square reconstruction error between the actual pixels in a region and their approximation. Then contours are coded using chain coding and polynomial coefficients are coded as binary values to the largest possible degree of quantization precision [5].

REFERENCES

1. Netrali and Limb,"Picture coding:A review",Proc. IEEE, vol. 68, NO.3. March 1980,pp.366-406.
2. A.Habibi,"Comparison of N-th-order DPCM encoder with linear transformation and block quantization techniques",IEEE Trans. commun. Technol.,vol. Com-19,pp.948-956. Dec. 1971.
3. J.B.Oneal,"Predictive quantizing differential pulse code modulation for the transmission of television signals",Bell Syst. Tech.J.,vol.45,pp. 689-722.May-June 1966.
4. J.Woods, " Stability of DPCM coders for television", IEEE. Trans. Commun.,vol.Com-23,pp. 845-846, Aug. 1975.
5. M.Kocher and R.Lionardi,"Adaptive region growing technique using polynomial functions for image approximation",Signal processing, April 1986, pp. 47-60.
6. W.K.Pratt,"Digital Image Processing",Prentice-Hall, Englewood Cliffs,N.J. 1978 .
7. T.Berger," Rate Distortion Theory", Prentice-Hall, Englewood Cliffs,N.J. 1971 .

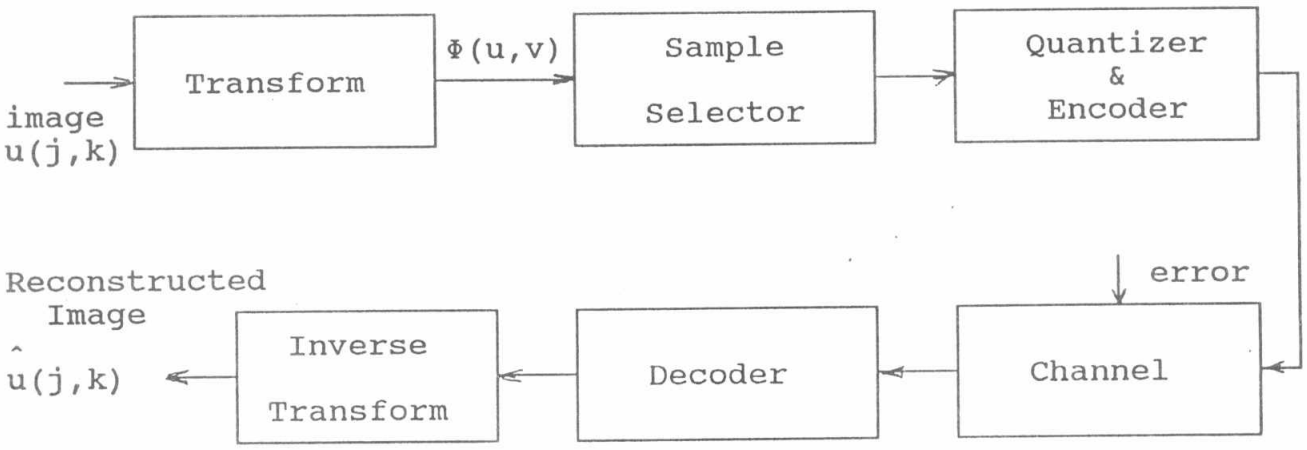


Fig.1 Block diagram of transform image coding

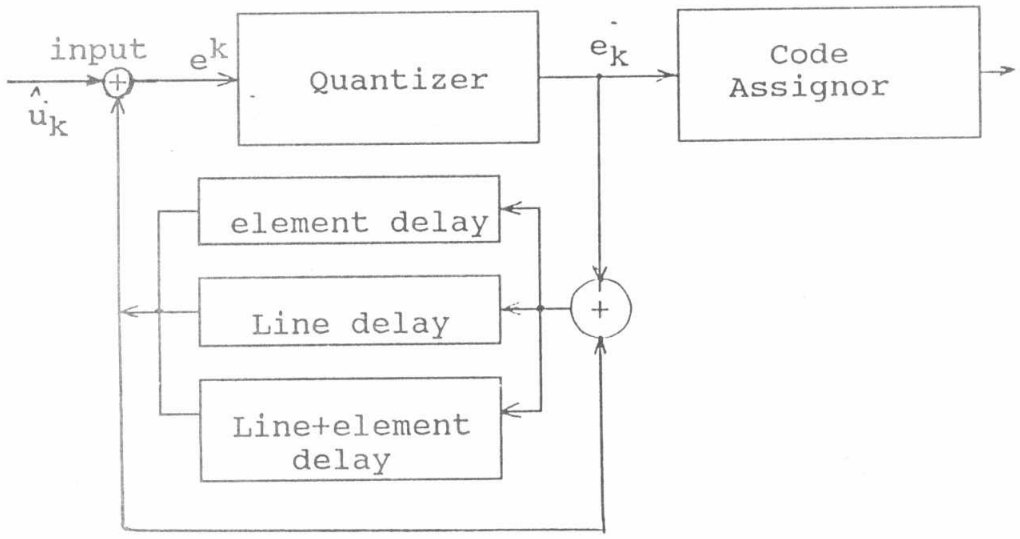


Fig.2.a Block diagram of 2-D DPCM system transmitter.

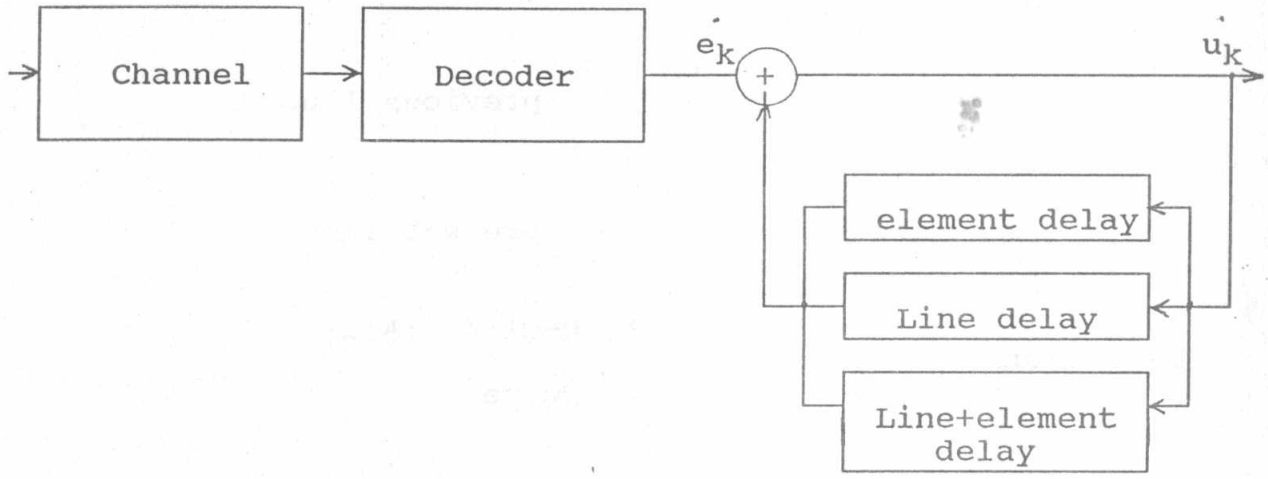


Fig.2.b Block diagram of 2-D DPCM system receiver.

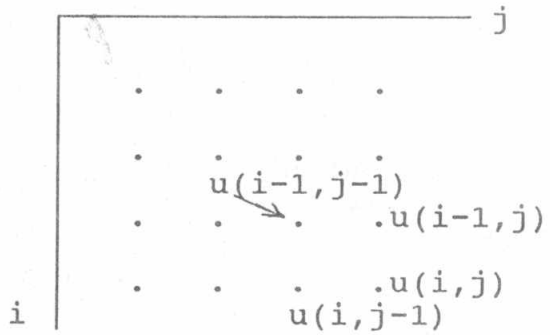


Fig.3 Calculation of prediction from previously transmitted samples

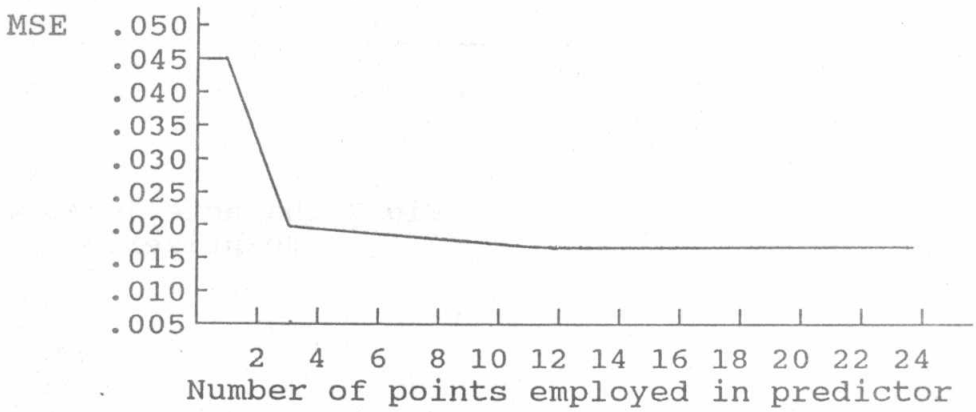
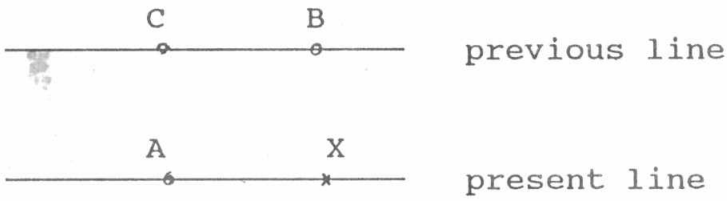


Fig.4 MSE as a function of the order of the predictor



$$\hat{X}(\text{prediction of } X) = \begin{cases} A & \text{if } |B-C| < |A-C| \\ B & \text{otherwise} \end{cases}$$

Fig.5 Graham's rule for adaptive prediction

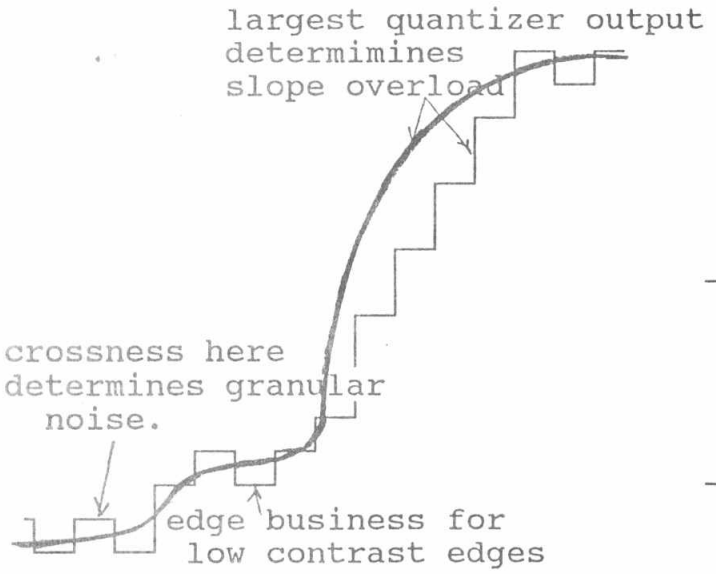


Fig.6 Quantizer degradation in DPCM coding.

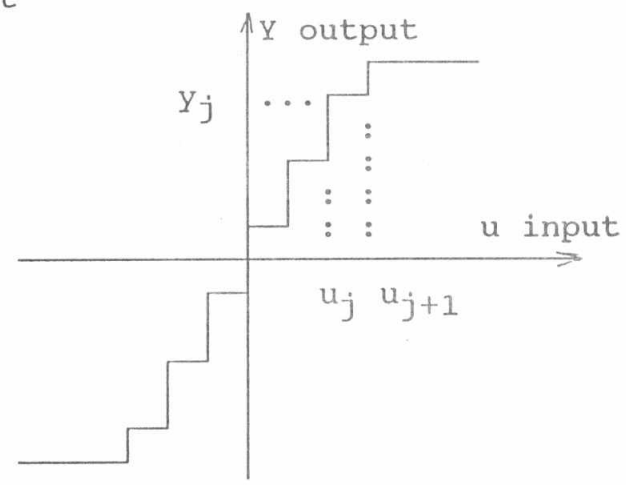


Fig.7 Characteristic of a quantizer.

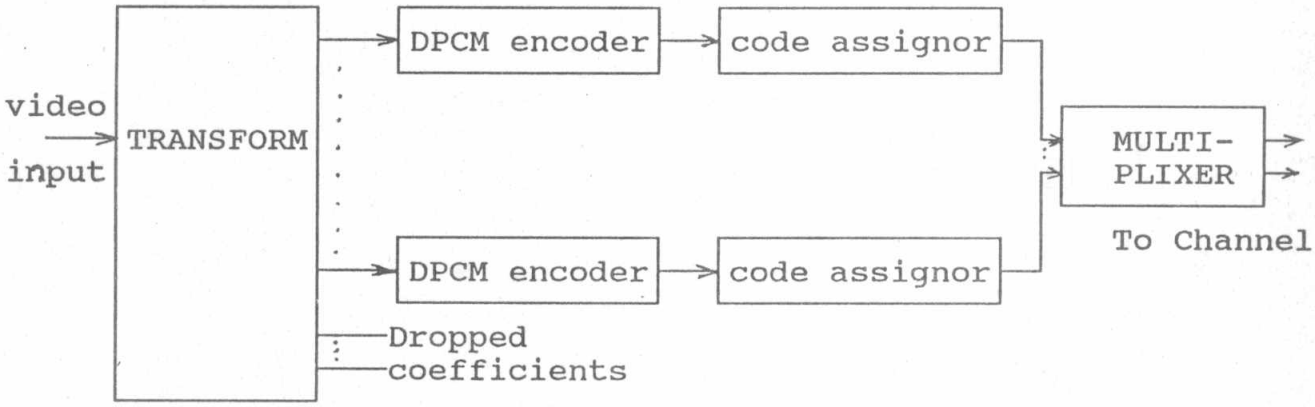


Fig.8 Block diagram of hyprid coder

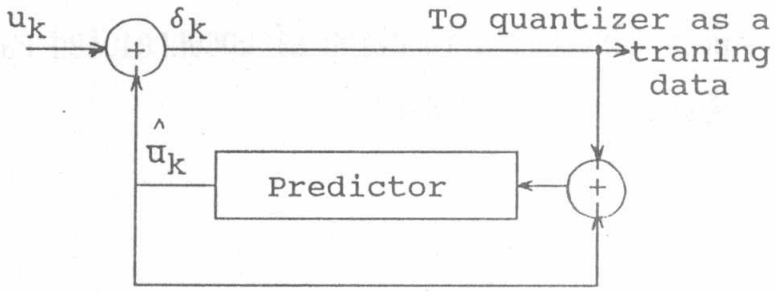


Fig.9 Calculation of innovation

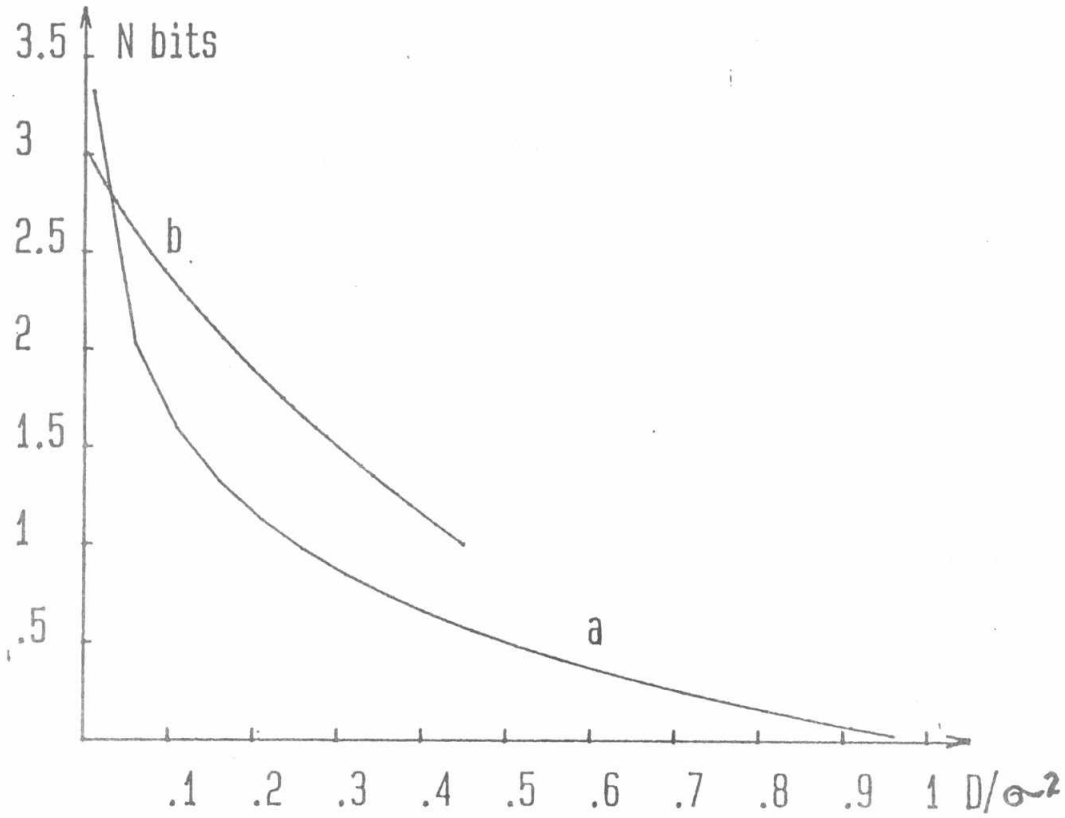


Fig.10.a-Rate distorsion function of uncorrelated random variables
b-Quantizer performance

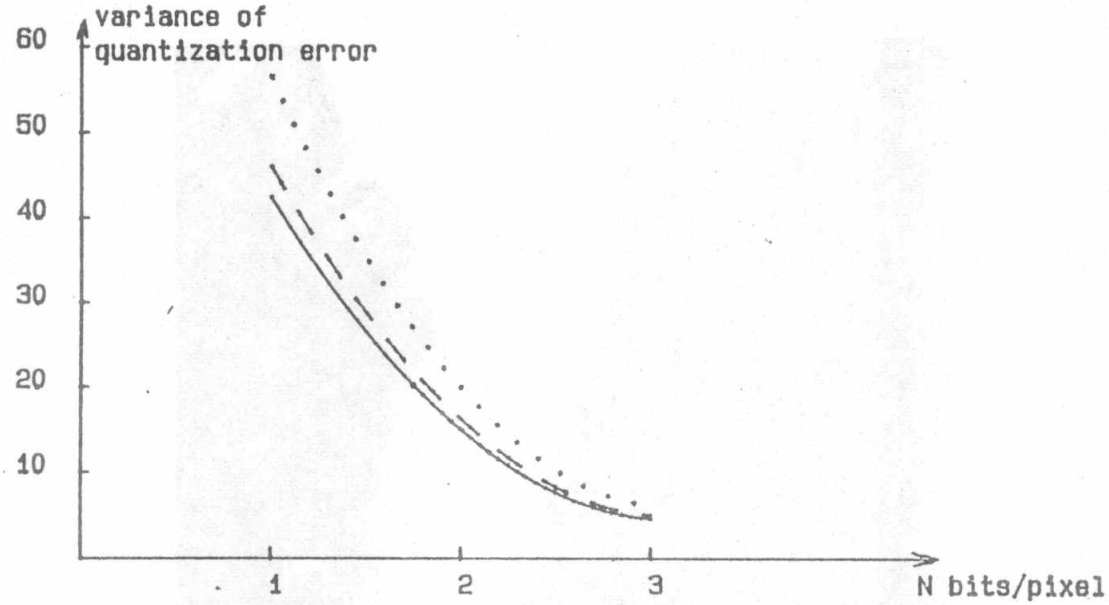


Fig.11.a Mean square quantization error versus number of bits/pixel
a) _____ optimum 2-d DPCM b) __ __ optimum 1-d DPCM
c). . . . previous element DPCM

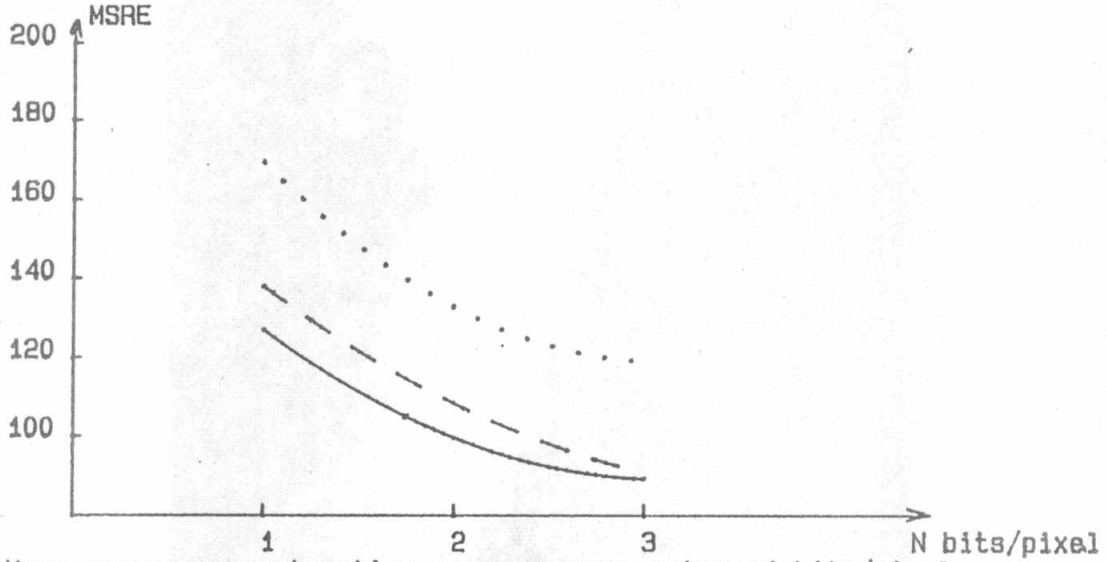


Fig.11.b Mean square reconstruction error versus number of bits/pixel
a) _____ optimum 2-d DPCM b) __ __ optimum 1-d DPCM
c). . . . previous element DPCM

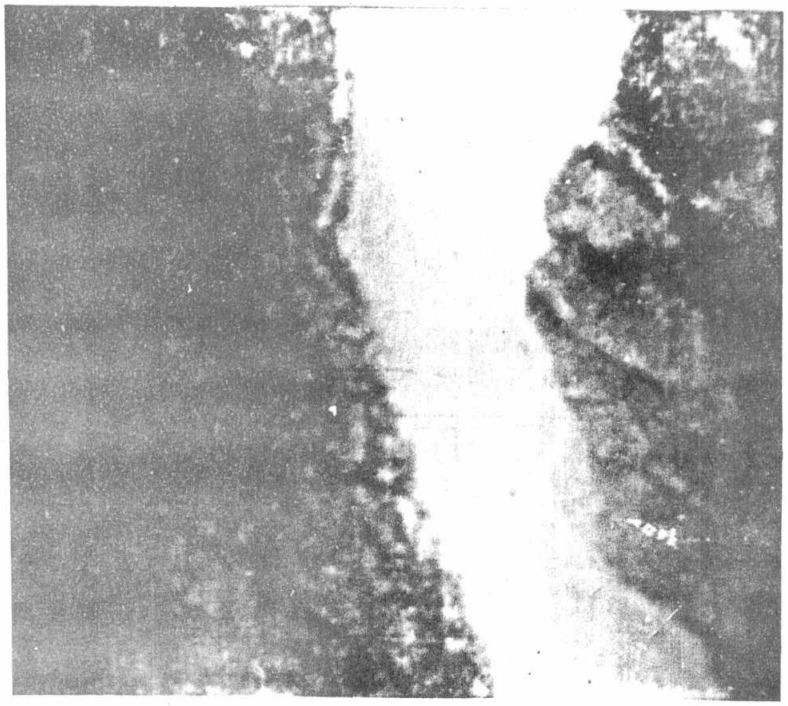


Fig.12.a Original picture of Minea area (8 bits/pixel)

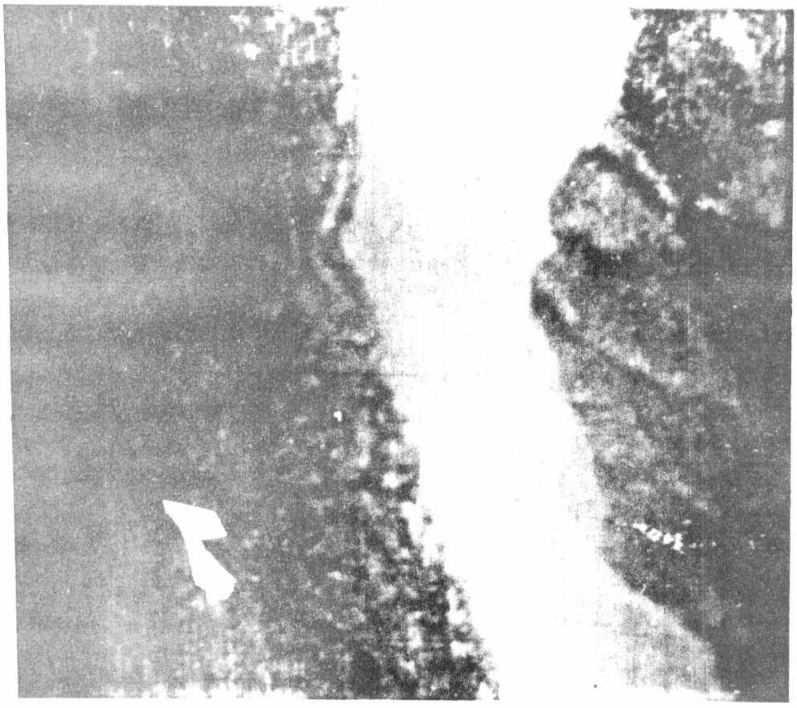


Fig.12.b Decoded picture of optimum 1-d DPCM system with optimum quantizer (1 bit/pixel)

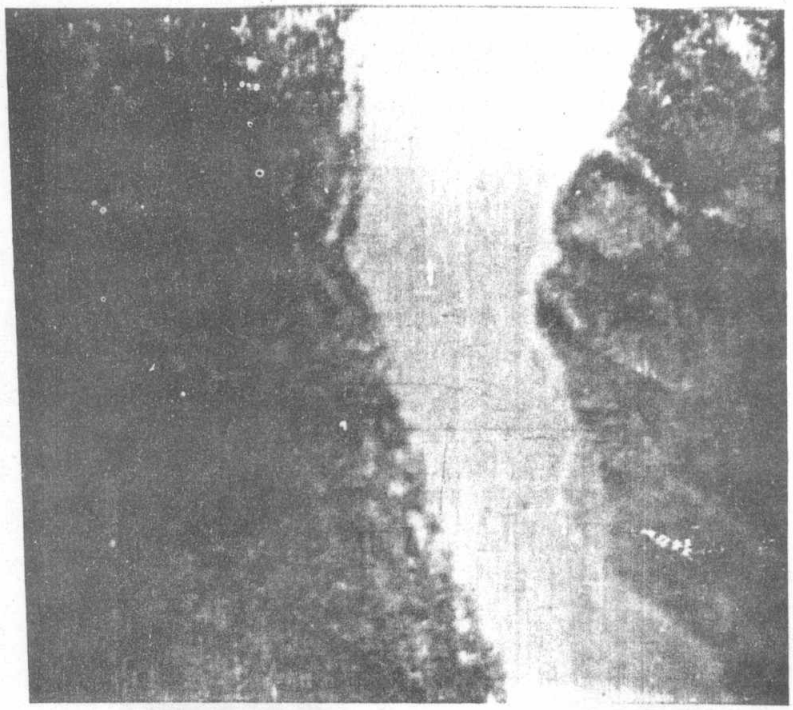


Fig.12.c Decoded picture of optimum 2-d DPCM system with optimum quantizer (1 bit/pixel)

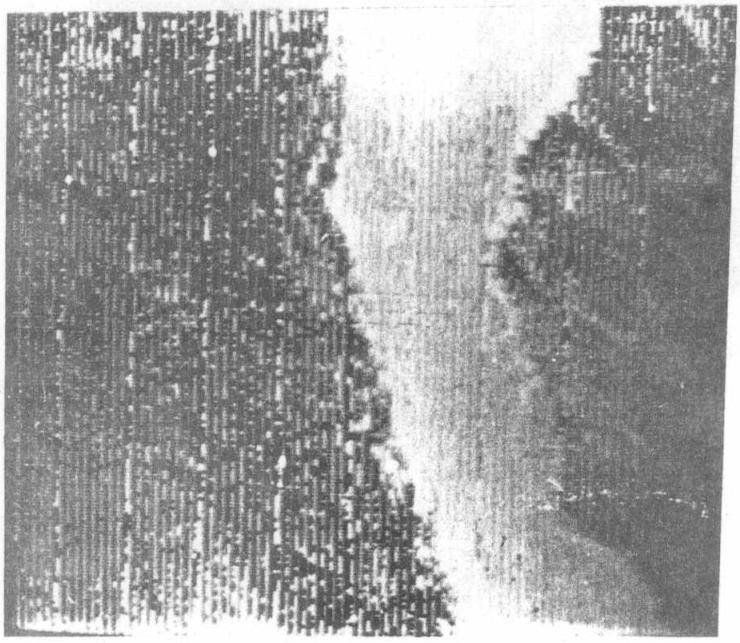


Fig.12.d Decoded picture of previous element DPCM system with optimum quantizer (1 bit/pixel)

Military Technical College
CAIRO - EGYPT

Fig.12.e Decoded picture of previous element DPCM system with uniform quantizer (2 bits/pixel)