AN ADAPTIVE NEAR INFORMATION PRESERVING IMAGE ENCODER

Saif S. Zahir
The University of British Columbia
Department of Electrical Engineering
Vancouver, B.C. V6T 1Z4
Tel. (604) 822-1375
Fax. (604) 822-5949
email: saifz@ee.ubc.ca

Morton Kanefsky
University of Pittsburgh
Department of Electrical Engineering
Pittsburgh, PA 15261
Tel. (412) 624-9677
Fax (412) 624-8002.
Mk1@vms.pitt.edu

ABSTRACT

A new adaptive near-information preserving image encoder that employs optimum predictive technique and that is insensitive to the type of image or the segmentation method employed, is presented. The encoder uses a non-symmetric half plan NSHP region of support (ROS) as well as a binary image that identifies the various regions of the segmented image as either stationary (i.e., homogenous) regions or non-stationary (i.e., transition) regions. Encoding is implemented via linear predictors whose coefficients, region of support, and prediction error quantization adapt depending on pixels location in the binary image. Reconstructed images are compared with those of segmentation based two-source coding algorithms and found to be objectively and subjectively significantly better.

1.0 INTRODUCTION

1.1 Non Adaptive Predictive Coding Methods

Let x(m, n) represents the pixel value of a two dimensional (2D) image at the position (m, n) and $\tilde{x}$ (m, n) represent its predicted value. To remove the redundancy, the predicted pixel value is subtracted from original pixel value to produce the prediction error, e(m, n) as follows: $e(m,n) = x(m,n) - \tilde{x}(m,n)$. The prediction error may be quantized e_q(m, n) for possible storage or transmission to the decoder via communication channel, and the reconstructed pixel $\hat{x}$(m, n), for both the encoder and the decoder is defined by the following : $\hat{x}(m,n) = \tilde{x}(m,n) + e_q(m,n)$. In order to obtain information-preserving compression, the prediction error is encoded in a Fano or Huffman code which recognizes that small errors are more likely than large errors. Often the errors are transmitted using crude differential pulse code modulation (DPCM) to avoid the complexity of Fano-Huffman type encoding. While this procedure is effective, it is not information preserving [1], [3].

In the two dimensional case, a linear AR prediction is defined as :

$$\hat{x}(m,n) = \sum_{(i,j)\in W}\sum b(i,j)\tilde{x}(m,n) + b_0 ,$$

where b(i, j) are the prediction coefficients, W is the region of support, and b_0 is a bias that satisfy the condition E[x(m, n) - $\tilde{x}$ (m, n)] and found to be equal to : b_0 = $[1 - \sum_{(i,j)\in W}\sum b(i,j)]\mu$ where μ is the mean value of x(m, n). Substituting for the bias b_0, the linear AR prediction becomes :

$$\hat{x}(m,n) = \sum_{(i,j)\in W}\sum b(i,j)x(m-i,n-j) + [1 - \sum_{(i,j)\in W}\sum b(i,j)]\mu$$

The linear prediction coefficients are determined by minimizing E{e^2(m, n)}, the mean squared value of the 2D prediction error, with respect to b(I, j) to produce the correlation function :

$$C_{k,l} = \sum_{(i,j)\in W}\sum b(i,j)C_{k-i,l-j}$$

where $C_{k,l}$ = E [x(m, n) x(m-k, n-l)] are the covariances to be measured. This enables the solution to the various covariances of the chosen window i.e., the 13 different covariances of our 6-pixels window (i.e., ROS) shown in Figure 1 below.

		0.11	
	0.22	0.20	0.11
0.11	0.25	X	

Figure 1. The 6-pixel ROS

2.0 THE PROPOSED ADAPTIVE NEAR-INFORMATION PRESERVING ALGORITHM

Assume that an image (I) can be segmented into a finite set of disjoint homogenous and transition regions, i.e., I = $U_{\forall(i,\ j)}$ x(i, j) = U_k R_{Hk} U_l R_{Tl} where Hk and Tl represent homogenous and transition regions respectively. Transition regions are treated alike whether or not they are merged. Further assume that the pixels in any homogenous region are independent of pixels in any other region. Thus E{ x_{Hk} (i, j)/$R_{Hk}{}^{C}$ } = μ_k, where c represent the complement and μ_k is the average value of the kth homogenous region. Under these conditions, the optimal causal predictor is given by x_{Hk} (i, j) = E{x_{Hk} (i, j) /R_{Hk} past}, where R_{Hk} past = I_{past} ∩ R_{Hk}. For Gaussian pixels, the optimal estimator is linear and can be written as :

$$\tilde{x}_{HK(i,\ j)} = \sum\sum_{WT} a(m,n)\tilde{x}(i-m,j-n)+[1-\sum\sum_{WT} a(m,n)\}\mu$$

where W_T is the intersection between some maximum region of support W_{max} and R_{Hk} past, and x(i, j) are the reconstructed image pixels.

The optimum prediction coefficients a(m, n) are determined from the Yule-Walker equations that are based on correlation measurements {C(m, n)} in each of the homogenous regions which need to be encoded.

Very small homogenous regions are best treated as transition regions. The number of coefficients, as well as the error variance ($\sigma_e{}^2$) depends on the size of the window W_T. The prediction error (e) is replaced by ±K for single bit quantization. The best K is given by K = E{ | e | } = [2 $\sigma_e{}^2$ / π]$^{1/2}$ for Gaussian errors. The error variance ranges from the value obtained for W = W_{max} up to the measured texture variance, C(0,0), if there are no pixels in the region of support.

We assume that texture details in nonstationary regions are relatively unimportant, and hence pixels in transition regions are predicted by local mean estimates over W_{max}. The optimal error quantization value in the transition regions is found by trial and error within a specified range, as mentioned above. This value is encoded along with average values and the correlation measurements of each homogenous region.

If the pixels can be considered Gaussian, then the optimal estimator is a linear estimator, or

$$\tilde{x}_{HK(i,\ j)} = \sum\sum_{WT} a(m,n)\tilde{x}(i-m,\ j-n)+b_0$$

where W_T is the inter-section between a maximum region of support W_{max} and the past of R_{Hk} i.e. W_T = W_{max} ∩ R_{HK} past. When encoding homogeneous regions, the optimal single bit quantization value was calculated. The optimal prediction coefficients are determined from Yule-Walker equations and the adaptive ROS mechanism.

Predictions can not be made in nonstationary transition regions. In the proposed algorithm, transition regions are predicted via local mean estimates of the pixel values in the region of support from :

$$\tilde{x}_{m,n} = \frac{1}{N}\sum\sum_{WT} x(n-i,m-j)$$

where N is the total number of pixels in the window, W_T.

The difference between the actual pixel value and predicted pixel value is the prediction error e_n = x_n - $\tilde{x}_n$. We replace e_n by ± K for single bit quantization. To obtain the optimal single bit quantization for the stationary homogeneous regions, we minimize the mean squared error given by E [(K - |e|)^2]. By taking the partial derivative with respect to K and then equating to zero which leads to K = E [|e|]. If f(e) = N (0, $\sigma_e{}^2$) where σ_e is determined for homogenous regions from Yule-Walker equations; then : K = σ_e √(2/π). If the maximum ROS is used (i.e., we are considering pixels in the interior of a homogeneous region). On the other hand, if there is no pixels in the region of support other than that to be predicted, then we set $\sigma_e{}^2$ = C(0, 0) where C(0, 0) is the measured texture variance. In general, for pixels near the edges of regions, K can assume values in the range given by :

$$\sigma^2 \sqrt{2/\pi} \le K \le \sqrt{2C(0,0)/\pi}$$

where C(0,0) is the measured covariance of the homogenous region.

The optimal quantization value for the transition regions is usually larger than in the homogeneous regions. An appropriate value for these regions can be found by a trial and error process. This value is encoded along with other statistical measurements for each of the homogeneous regions in the image.

3.0 SIMULATED TEST IMAGE

To assess the performance of the proposed algorithm, a special test image of 128 x 128 pixels was generated as shown in Figure 2 (Note : the results of the several test and real images are reported in a paper to the IEEE transaction on image processing). The chosen test image is nonstationary in the sense that it has a space-varying mean gray level value and space-varying texture. The test image consists of several homogeneous regions (objects and background) and two slowly changing nonstationary transition regions. Texture was generated that is a Gaussian and autoregressive process. Two different windows were employed for this purpose. The first is a 3 x 3 , quarter plane region of support predictor and the second is the 6-pixels window with a NSHP ROS shown in Figure 1. The weights of each window add up to unity for convenience. Also, to produce more textures for the different regions, two variances were chosen for the texture. Those were 4 and 16 respectively. For each region in the image a window and a variance value were chosen that are different from neighboring regions.

4.0 SIMULATION RESULTS

The mean squared error is the criteria used to test the fidelity of the reconstructed images of the four different methods. The MSE is given by :

$$MSE = \tfrac{1}{MN}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}[x(m,n) - \bar{x}(m,n)]^2$$

Table 1. depicts the MSE of reconstructed images of the four methods. The results show that the Proposed Method achieved a small MSE, less than one (0.98), which translates to a very high signal-to-noise ratio compared to a MSE of 155.5, 7.10, and 7.01 for the standard one source DPCM, TS-LMS, and TS-DPCM respectively. To show these results subjectively, we have displayed row 25 of the image for each method as shown in Figures 3, 4, and 5 respectively. A close look at these plots reveals that the one-source (standard) DPCM behaved poorly on and near edges, thus causing quality degradation and generated high noise in homogenous regions. As for the two source coding TS-DPCM and adaptive TS-LMS methods, they failed to track variation in texture in homogenous regions as well as different texture in different regions. These results are obtained with almost perfect separation of the underlying structure from the textural image which is very difficult to obtain under normal processing. The sudden change of texture in the different regions caused the TS-LMS method generated a non-stable encoder which necessitated recalculation of the stability factor and re-implementation of the encoder.

Table 1. The MSE of the four coding techniques

TECHNIQUE	**MSE**
Standard DPCM	155.5
TS-LMS	7.10
TS-DPCM	7.01
THE PROPOSED ALGORITHM	**0.98**

5.0 CONCLUSIONS

We have presented a new adaptive near-information preserving predictive encoder for images with arbitrarily shaped image segments. The proposed algorithm targets a wide range applications where information preserving is of primary concern. This algorithm has several advantages over predictive coding techniques, DPCM and adaptive LMS methods. First, the proposed encoder significantly outperforms the idealized two-source DPCM and LMS methods by achieving nearly 7.1 times less MSE than either of the two methods and 155.5 times less than the standard DPCM method while maintaining almost a perfect tracking of texture variation in homogenous regions and from one region to another. Second, edges were near optimally preserved without additional cost in MSE or tracking texture variations. Based on objective and subjective results obtained, we can claim that the proposed encoder is a near information preserving algorithm. The binary image and statistical measurements require approximately 0.03 bits/pixel.

6.0 BIBLIOGRAPHY

1. Jain, A. K., *Image Data Compression : A Review*, Proc. IEEE, Vol. 69, No. 3, 1981, pp. 349-389.
2. Kayhan, A. S., *An Algorithm for Two Source Image Coding*, M. Sc. Thesis, University of Pittsburgh, 1988.
3. Gilge, M., Engelhardt, T. and Mehlan, R., *Coding of Arbitrarily Shaped Image Segments Based on a Generalized Orthogonal Transform*, Signal Processing : Image Comm., Vol.1, 1989, pp. 153-180.
4. Chaparro, L. F., Kanefsky, M., and Popoutsis, *The Use of Multiple Source for the Modeling and Coding of Nonstationary Images*, IEEE Trans. Systems, Man, and cyber., Vol. SMC-16, pp. 482-486, 1986.
5. Zahir, S., Lee, S. and Kanefsky, M., *A Segmentation Based Predictive Coding*, Proc., Data Compression Conference, Utah, March, 1994.
6. Chaparro, L. F., Kanefsky, M., and Boudaoud, M., *Two-Dimensional Linear Prediction Covariance Method and its Recursive Solution*, IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-17, No. 4, pp. 617-621, 1987.
7. Leou, F-C and Chen, Y-C, *A Contour-Based Image Coding Technique With Its Texture Information Reconstructed by Polyline Representation*, Signal Processing, Vol. 25, 1991, pp. 81-89.
8. Beggar, M.J., Morris, O. J. and Constantinides, A.G., *Segmented-Image Coding: Performance Comparison With the Discrete Cosine Transform*, IEE Proc. Vol.135 Pt. F, No.2, April 1988, pp. 121-132.
9. Strobach, P., Linear Prediction Theory : A Mathematical Basis for Adaptive Systems, Springer-Verlag, 1990.
10. Maragos, P. A., Schafer, R.W., and Mersereau, R.M., *Two Dimensional Linear Prediction and its Application to Adaptive Coding of Images*, IEEE Trans. on ASSP, Vol. ASSP - 32, No.6, pp. 1213-1229, 1984.
11. Kocher, M. and Leonardi, R., *Adaptive Region Growing Technique Using Polynomial Functions for Image Approximation*, Signal Processing, Vol. 11, No. 1, 1986, pp. 47-56.
12. Nasiopoulos, P., Ward, R., and Morse, D., *Adaptive Compression Coding*, IEEE Trans. on Comm., Vol. 39, No.8, 1991.
13. Young-Sik Chung and Morton Kanefsky, *DPCM Using A Backward Predictor For The Compression Of Nonnegative Images*, in Ph.D. Thesis of University of Pittsburgh, 1988.
14. Moghaddamzadeh, A and Bourbakis, N., *A Fuzzy Technique for Image Segmentation of Color Images*, Third IEEE International Conference on Fuzzy Logic, pp. 83-88, 1994.
15. Hadhoud, M. and Thomas, D., *The Two-Dimensional Adaptive LMS(TSLMS) Algorithm*, IEEE Trans. on Circuits and Systems, Vol.15, No. 5, 1988.

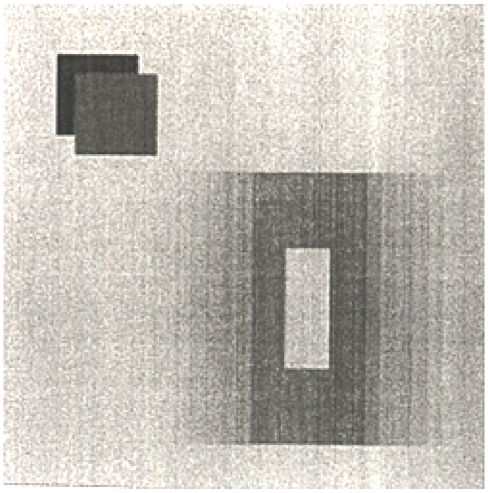

Figure 2. Test Image

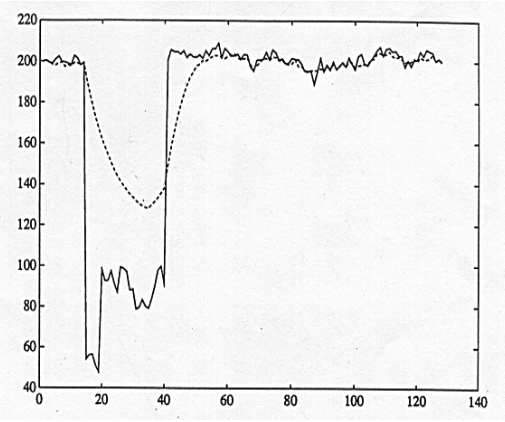

Figure 3. Standard DPCM

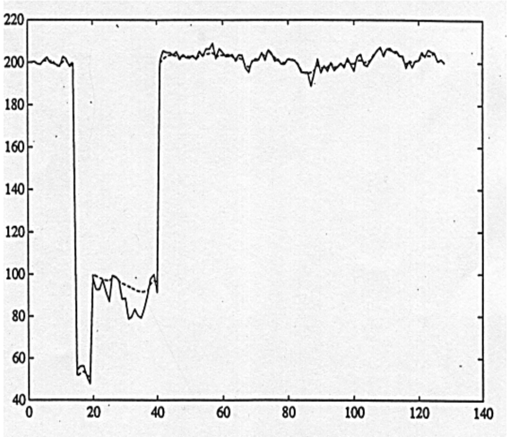

Figure 4.   TS-LMS method

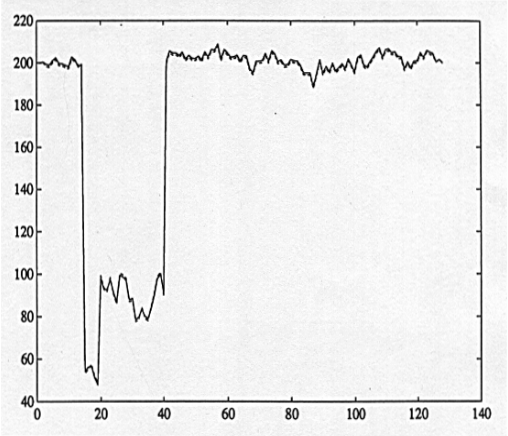

Figure 5. The Propose Algorithm