International Symposium on Signal Processing and its Applications, ISSPA, Gold Coast, Australia, 25-30 August, 1996. Organized by the Signal Processing Research Centre, QUT, Brisbane, Australia.

A HYBRID CODER FOR CODE-EXCITED LINEAR PREDICTIVE CODING OF IMAGES

Farshid Golchin and Kuldip K. Paliwal*

School of Microelectronic Engineering, Griffith University
Brisbane, QLD 4111, Australia
K.Paliwal@me.gu.edu.au, F.Golchin@me.gu.edu.au

ABSTRACT

In this paper, we propose a hybrid image coding system for encoding images at bit-rates below 0.5 bpp. The hybrid coder is a combination of conventional discrete cosine transform (DCT) based coding and code-excited linear predictive coding (CELP). The DCT-based component allows the efficient coding of edge areas and strong textures in the image which are the areas where CELP based coders fail. On the other hand, CELP coding is used to code smooth areas which constitute the majority of a typical image, to be encoded efficiently and relatively free of artifacts.

1. INTRODUCTION

Most images consist of large smooth areas and textures which are separated by edges. This characteristic of images is exploited in many image coding schemes in use today. Although there is no underlying source model for images, the smooth areas and textures in images can be well modelled as the output of a 2-dimensional autoregressive (AR) process. The use of this well known property in image coding dates back to the earliest DPCM and transform coding techniques [1]. It is this property of images that allows a DCT to perform almost as well as the optimal Karhunen-Loeve Transform (KLT).

Code-Excited Linear Predictive (CELP) coding [2] has demonstrated excellent results for the encoding of AR sources and is the basis for many speech coding methods in use today. So the question arises why CELP coding has not found wider application in image coding. In the past, there have been attempts, with varying degrees of success to apply the principles of CELP coding to images [3], [4] and [5]. Gimeno et. al. [6] have successfully applied a similar technique to the coding of textures.

CELP coding's lack of success in the coding of images is largely due to its inability to encode edge areas. Edge areas are the areas in an image where we find the AR model to be least appropriate. As a result, edges are badly degraded and may even cause instability in the system. However, while the edge areas are degraded, we find that smooth areas and textures are very well preserved, even at low bit-rates. This led us to investigate combining CELP coding a with another technique, such as DCT, which is better suited to encoding edge areas. The resultant hybrid coder is able to use CELP for efficiently encoding the smoother regions while the DCT-based coder is used to encode the edge areas. In this paper, we describe the hybrid coder and examine its results when applied to a typical monochrome image.

2. CELP CODING OF IMAGES

In CELP coders, source signals are approximated by the output of synthesis filter (IIR) when an excitation (or code) vector is applied to its input. Essentially, the encoding process consists of finding the appropriate excitation vector and synthesis filter parameters to approximate the desired source signal. Provided that the excitation vector and the synthesis filter can be represented efficiently (i.e. using few bits), compression can be achieved.

2.1. Image Synthesis using CELP

For our CELP coder, we essentially follow the implementation described in [3]. Since CELP coding is a form of analysis-by-synthesis coding, it is probably best described by its decoder. The block diagram of a CELP decoder is given in Fig. 1. In this figure, y(i,j) denotes the output the decoder which is the intensity of the pixel at location (i,j). In order for images to follow the AR model, we need to remove the local means in the image (in this case for 8x8 blocks of pixels). The local means are subtracted from the image data at the encoder and then quantized and transmitted as side information to the decoder. At the CELP decoder, the local mean for each block, denoted in Fig, 1 by M(m,n) is added back to the synthesized, zero-mean image, s(i,j). The indices (m,n) denote the location of the b x b pixel blocks within the image.

The decoded image is synthesized block by block by selecting an optimum codevector from a codebook, multiplying it by a suitable gain factor G(m,n), before passing it through the

* F. Golchin is supported in part by a CSIRO Division of Radiophysics postgraduate scholarship.

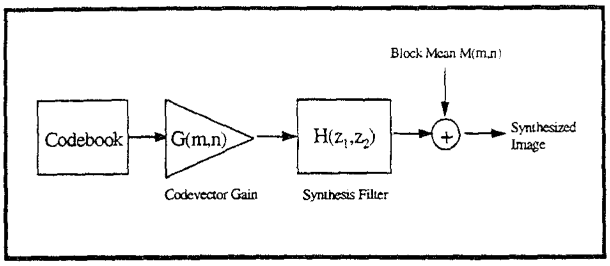

Fig. 1. The CELP decoder

synthesis filter $H(z_1,z_2)$. $H(z_1,z_2)$ describes an all-pole filter where $H(z_1,z_2) = 1 / A(z_1,z_2)$. The coefficients of the all-zero filter $A(z_1,z_2)$ are the 2-dimensional LP coefficients of the image, that is:

$$A(z1,z2) = 1 - \sum_k \sum_l a(k,l)\, z_1^{-1} z_2^{-1} \quad , (k,l) \in \mathbf{R}.$$

Here, $\{a(k,l)\}$ denote the LP coefficients of the image and **R** is the set which defines the region of support for the predictor. In our coder we choose the three pixels at the locations (x-1,y), (x,y-1) and (x-1,y-1) to predict the pixel at location (x,y) and hence the region of support for the predictor is defined by:

$$\mathbf{R} = \{ (-1,0), (0,-1), (-1,-1) \}.$$

The LP coefficients $\{a(k,l)\}$ are found using the well known covariance method. Three coefficients are calculated for the entire image and must be sent along to the decoder as side information.

The image synthesis can be accurately described by the following difference equation:

$$y(i,j) = \sum_k \sum_l a(k,l)\, s(i-k,j-l) + M(m,n) + u(i,j), \quad (k,l) \in \mathbf{R}.$$

where y(i,j) is the reconstructed image and s(i,j) is the synthesized image before the mean is added to it. The excitation signal u(i,j) is found by multiplying the optimum codevector by the gain factor G(m,n).

2.2 The CELP Encoder

A diagram of the CELP encoder is provided in Fig. 2. The encoding process in CELP is equivalent to estimating the parameters of the synthesis system and transmitting them in quantized form. The process can be divided into the following steps:

1. Calculate LP coefficients for the image.
2. Divide image into b x b pixel blocks.
3. Calculate and remove the mean for each block.

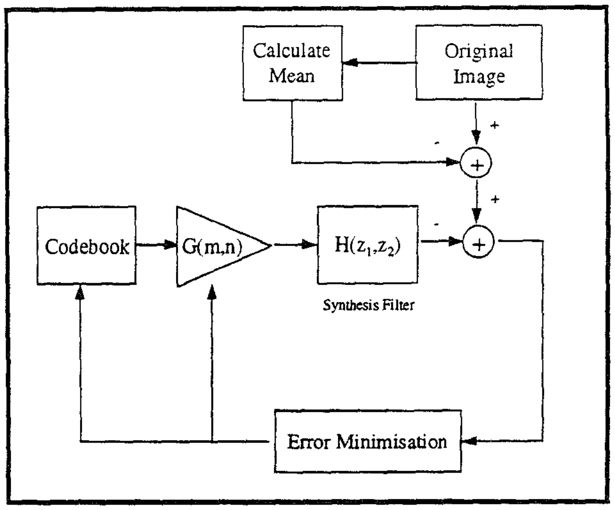

Fig. 2. The CELP encoder

4. Find the optimum codevector and gain value by passing all codevectors through the synthesis filter to find which results in the least reconstruction error.

5. Transmit the index of the codevector and the value of the gain factor (in quantized form).

Rather than quantizing G(m,n) directly, a logarithmic value of G(m,n) was quantized using a Lloyd-Max quantizer with 4 bits. The side information that needs to be sent along to the decoder consists of the local means of the image blocks and the LP coefficients. The local means are DPCM encoded to remove the correlation among neighbouring block means. The DPCM scheme utilised in this case, uses a 4-6 bit Lloyd-Max quantizer. The LP coefficients are quantized with very high accuracy (up 32 bits for each coefficient).

The codebook used, is a collection of 256, 64-dimensional (for 8x8 image blocks), unity norm vectors of gaussian random numbers. Each vector in the codebook can be addressed by an 8-bit index value.

Testing the CELP coding system described here demonstrated poor results. Even when leaving the mean and gain values unquantized and extending the codebook to 4096 vectors, we found the coded images to be severely degraded, specially at the edge areas. As mentioned before, edges do not follow the source model which CELP coding relies upon. Therefore, when an edge is encountered the reconstruction error becomes quite large. Due to the predictive nature of the encoder, the reconstruction error at the edges spreads to all neighbouring pixels (even non-edge pixels).

One way to overcome this problem would be to limit the use of CELP coding to only encode the smoother areas in images and use a different coding scheme for other blocks. Since CELP coding operates on blocks of pixels, it would be helpful to use another block based scheme for encoding the edges. For this purpose we chose a simple DCT-based coder which is described in the following section.

3. THE HYBRID CELP CODER

The Hybrid CELP coder consists of two simpler coders. The first is the CELP coder described in the previous section and the second is a simple DCT-based coder. Each image block is coded using CELP or DCT, depending on which method will produce the least distortion in the reconstruction. The block size chosen in this case is 8x8 pixels.

Since the encoder is using one of the two coding techniques for each blocks, the decoder also needs to be informed of this decision. This means an additional cost of 1 bit per block to specify the coding method being used.

The DCT coding follows along the same lines as [7], except we only make use of a single class of blocks. In the DCT coder, the 8x8 DCT of all image blocks are taken and bit budget for a block is allocated among the DCT coefficients based on an estimate of the coefficient's variance (over all image blocks). Since the local means have already been removed, there is no need to allocate any bits to the (0,0) coefficient, commonly known as the DC coefficient of the DCT blocks.

The CELP coder is capable of encoding the shade and smooth texture blocks quite efficiently and hence requires fewer bits than the DCT coder. In the optimal case, an optimisation procedure would be used to decide on the proportion of the blocks encoded using CELP and DCT and the bits used for each type of block. However, a sub-optimal but much simpler solution is found in setting the DCT bit budget (bits per pixel for DCT coded blocks). This will place an upper limit on the overall bit-rate. If any blocks at all are encoded using CELP then the overall bit-rate is reduced. The number of blocks that are encoded using CELP at low bit-rates, is usually around half the total number blocks. Also, by only setting the DCT bit-budget, we place an upper limit on the coding error. The error will be at most equal to that of a DCT coder operating at the designated bit-rate.

4. RESULTS

The Hybrid CELP coder was applied to encoding 512x512 pixel monochrome images. The CELP codebook was designed using a simple random number generator to generate the random vectors which were normalised to unity norm and placed into the codebook. A codebook training procedure such as the LBG algorithm is not required in this case. An LBG codebook was also experimented with, but did not offer any significant advantage. The codebook was designed to contain 256, 64-dimensional vectors (to use for 8x8 image blocks).

We found that 4 bits were sufficient for the DPCM coder to quantize the means of the image blocks. The logarithmic values of the gain were quantized using 3 bits. One bit was used to specify whether the block is encoded using CELP or DCT coding. This bit-allocation scheme results in an effective bit-rate of 0.25 bpp (bits per pixel) within the CELP coded blocks. The contribution of the LP coefficients to the bit-rate is negligible ($\sim 4\times10^{-4}$ bpp).

The DCT coder was set to operate at 0.4 bpp. Taking into account the mean (4 bits) and classification information (1 bit) this results in an effective bit-rate of 0.48 bpp within the DCT coded blocks.

The original of the image "Bank" is provided in Fig. 3. This image was encoded first using only the DCT coder described above (i.e. without CELP). The quality of the reconstructed image was quite poor. A portion of the reconstructed image has been magnified and shown in Fig. 4. It is evident that mosquito noise has degraded all the smooth areas and textures within the image.

The Hybrid CELP coder was then used to encode the same image. We found that the CELP component of the Hybrid CELP coder was used to encode 1994 of the total 4096 image blocks. CELP has reduced the bit-rate of the image from 0.48 bpp in the previous case down to 0.36 bpp while guaranteeing the same or better quality of reconstruction. The reconstructed image from the hybrid coder can be seen in Fig. 5. Again, a portion of the reconstructed image is magnified in Fig. 6. Comparison with the DCT only image in Fig. 4, clearly highlights the improvement in the quality of the smooth regions.

5. CONCLUSION

The principles of CELP coding were applied to the encoding of images. It was found that CELP coding is not capable of encoding the edge areas in images with any acceptable level of quality. However, its good performance over the smooth regions of images can still be taken advantage of, if it is combined with another coding technique.

A hybrid coder, which used a simple DCT coding as an alternative to CELP was implemented. Comparison with the DCT coder alone, shows that the hybrid coder is able to reduce the overall bit-rate while improving the quality of the reconstructed image.

The results in this paper suggest that CELP provides a viable option for encoding the smooth parts in images. The inclusion of CELP coding into existing block-based techniques such as JPEG may improve the overall reconstruction quality or reduce the bit-rate.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Geoff T. Poulton of the CSIRO Division of Radiophysics for his encouragement and helpful discussions.

REFERENCES

[1] A.K. Jain, "Image Data Compression: A Review," *Proc. IEEE*, Vol. 69, pp. 349-389, March 1981.

[2] M.R. Schroeder and B.S. Atal, "Code-Excited Linear Prediction (CELP): High quality speech at very low bit-rates," *Proc. ICASSP*, pp. 937-940, 1985.

[3] K.K. Paliwal, "Low Bit-rate Image Coding Using 2-D Linear Prediction and 2-D Stochastic Excitation," *Proc. ICASSP*, pp. 800-803, 1988.

[4] L. Torres and E. Arias, "Stochastic Vector Quantization of Images," *Proc. ICASSP*, pp. 385-388,1992.

[5] L. Torres and E. Arias, "Improvements on Stochastic Vector Quantisation of Images," *Proc. ICASSP*, pp. 237-240, 1993.

[6] D. Gimeno, L. Torres and J.R. Casas, "A New Approach to Texture Coding Using Stochastic Vector Quantization," *Proc. ICIP*, pp. 119-123, 1994.

[7] W. Chen and C.H. Smith, "Adaptive Coding of Monochrome and Color Images," *IEEE Trans. Commun.*, Vol. COM-25, pp. 1285-1292, Nov. 1977.

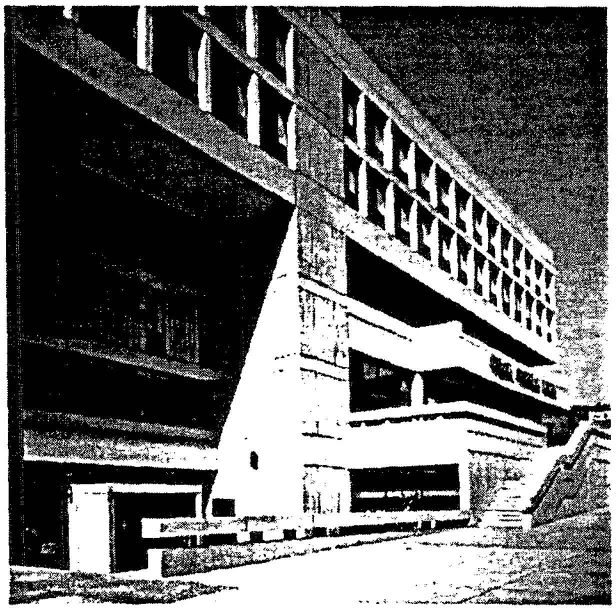

Fig. 3. The Original "Bank" image

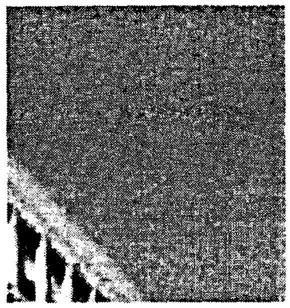

Fig. 4. A portion of the DCT coded image

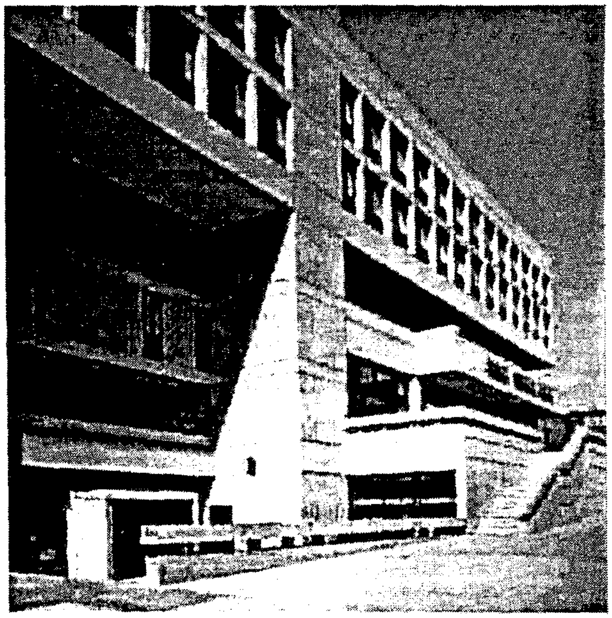

Fig. 5. Reconstructed image using the hybrid coder

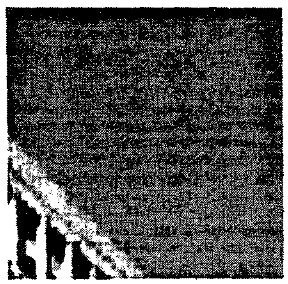

Fig. 6. A portion of the image encoded using the hybrid coder