26th IEEE Southeastern Symposium on System Theory (SSST'94), Athens, OH, pp. 457-461,March 20-22, 1994.

Image Data Compression Using Multiple Bases Representation

John F. Tilki and A.A. (Louis) Beex

The Bradley Department of Electrical Engineering
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061-0111

Abstract

Digitized images contain huge amounts of information which strain, or exceed, the capacity for their real-time processing, storage, and retrieval. Various compression techniques have been developed to reduce the amount of data necessary for representation. We report on a hybrid image data compression procedure based on a multiple bases representation.

The multiple bases representation technique described herein utilizes advantages of transform coding, vector quantization, and predictive coding, while aiming to circumvent the associated disadvantages of each. Preliminary results indicate that this procedure can outperform conventional compression methods, and yield high compression ratios while avoiding prohibitive computational complexity.

1. Introduction

Digitized images contain huge amounts of information which strain, or exceed, the capacity for their real-time processing, transmission, storage, and retrieval. Various compression techniques exist which reduce the amount of data necessary for representation. We are developing a hybrid image data compression procedure based on a multiple bases representation.

Conventional data compression schemes typically tradeoff between computational complexity and compression yield. Predictive coders are fast and simple, but they usually provide low compression rates. Similarly, transform coding methods trade some compression yield for computational efficiency. In contrast, vector quantization algorithms achieve high compression rates, but are computationally intense.

The hybrid method under development combines advantages of transform coding, vector quantization, and predictive coding, while aiming to circumvent the associated disadvantages of each. Called Multiple Bases Representation (MBR) [1], this technique uses several fast transformations to map the image data into multiple orthogonal basis sets.

2. Conventional image compression methods

In digital images, picture elements or pixels are usually highly correlated with their neighbors. This results in a redundancy of information. From information theory we know that the measure of information content is based on the amount of surprise that a 'bit' of information provides. A pixel value that is the same as the previous pixel value contains no 'surprise' and hence little information. Therefore, since the information is actually contained in the changes from pixel to pixel, a picture can be represented more efficiently by coding these changes.

2.1 Predictive coders

Predictive coders directly utilize the fact that picture elements in local regions are highly correlated with one another. Differential Pulse Code Modulation (DPCM) is the simplest predictive coding scheme, in which the coder predicts that the next pixel has the same value as the previous pixel. The predicted value is subtracted from the actual value, and the difference is then coded. Higher order predictors use more than a single past pixel in the prediction.

2.2 Transform coding

The fact that pixels are highly correlated with their neighbors also suggests that pictures contain an important low-pass component. Transform coding methods utilize this fact by transforming the image into the frequency domain, followed by efficient coding of the resulting coefficients. Since most of the higher frequency components lack energy, they can be coded with fewer bits or eliminated completely. Various fast algorithms are available to perform the necessary transformations onto the corresponding orthogonal basis vectors.

2.3 Hybrid methods

Hybrid techniques combine transform and predictive coding. Fast transformations are performed on a block of data in one dimension, and the results used to predict in the other dimension to further decorrelate the image data.

2.4 Vector quantization

Vector quantization is an extension of transform coding in that the transform domain is expanded into a general vector space. Since the basis vectors are no longer required to be orthogonal, they can be chosen to better conform to the signal structure. Each vector of data in the image is mapped to its nearest neighbor in the vector space. The resulting codewords are stored or transmitted, and the image is reconstructed by mapping the codewords back into reconstruction vectors. Since the vectors in the vector space no longer correspond to those of orthogonal transformations, an exhaustive search of the entire vector space must be made to find the nearest neighbor match to the data. This full search is generally computationally intensive and time consuming.

3. Multiple bases representation

We have developed an image data compression technique which utilizes advantages of transform coding, vector quantization, and predictive coding, while circumventing many of the associated disadvantages of each. This Multiple Bases Representation (MBR) technique uses several fast transformations to map the image data into multiple orthogonal basis sets. The coefficient with the most energy is found and stored, and the result is then subtracted from the original data forming a residual signal. The residual signals are then recursively operated on in a similar manner until specified stopping criteria are met. This process is called Recursive Residual Projection (RRP). A common stopping criterion is mean-square-error (MSE), and the recursive residual projection is halted after the energy of the residual signal falls below a given threshold. Other possible stopping criteria include maximum absolute value of the residual signal or a maximum number of coefficients.

As in vector quantization, the vectors in all of the MBR basis sets are not orthogonal, so they can better conform to the image data. Within each transformation however, the vectors are orthogonal, and fast algorithms exist which perform the appropriate mapping. This combination provides a compression scheme which uses multiple fast transforms in parallel, and which has high energy packing characteristics like vector quantization.

The three transforms used in MBR to date were chosen for their fast implementations and other special properties. The Discrete Cosine Transform (DCT) [2] is able to efficiently represent texture information and has excellent energy compaction for images. The Fast Haar Transform (FHT) [2] is very useful for feature extraction and the representation of edges. Finally the Identity Transform (IDT) developed for this project can efficiently reduce localized errors. The basis vectors of the IDT are simply the natural Euclidean basis vectors which have been normalized to have zero mean.

4. Statistical modeling

Once the entire image has been decomposed, the coefficients must be quantized, and these quantized coefficients along with the vector positions must be mapped into a finite symbol set for transmission or storage. Statistical modeling is used to parameterize the coefficient and vector probability density functions (pdfs). These models are then used to entropy code all of the symbols which compose the image. By assuming a statistical structure *a priori* and parameterizing, only a few parameters are needed to represent the coefficient and vector position densities, and the densities can readily be recreated in the reconstruction process.

After viewing the histograms for the DCT and FHT transform vector positions (indices) from several test images, it was decided that a gamma density function would be an appropriate model (see Figure 1) due to the higher probability of the low-frequency vectors

$$p(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}} \qquad (1)$$

where $x > 0$, a > 0, and b > 0. The gamma density takes on a variety of forms depending on the values of the a and b parameters. The a parameter determines how quickly the function rises initially, since for small x the x^{a-1} term dominates the expression. Similarly b controls how quickly the function decays for large x. Furthermore, when a equals one the density degenerates to the exponential pdf, and when b equals two it degenerates to the chi-square density [3].

Unlike the DCT and FHT basis vectors, the positions of the IDT vectors are not frequency ordered. Hence the lower position vectors do not have a higher probability of occurrence than the higher position ones. In fact, all of the IDT basis vectors are equally probable, since a very large or very small single pixel can occur at all positions with equal probability. Hence a uniform pdf is assumed for the IDT vector position density (see Figure 2).

In order to model the histograms of the AC transform coefficients, a double-sided version of the gamma density was developed (see Figure 3).

$$p(x) = \frac{1}{2\Gamma(\alpha)\beta^{\alpha}} \, |x|^{\alpha-1} \, e^{-\frac{|x|}{\beta}} \qquad (2)$$

where $\alpha > 0$ and $\beta > 0$.

The DC coefficient histogram is best matched by a double-sided exponential density (see Figure 4). Since the double-sided gamma function becomes the double-sided exponential when α equals one, the same model structure can be used here as well.

The method of moments [3] is used to estimate the α and β parameters for the densities based on the first, second, and fourth moments of the coefficient values and vector positions.

5. Optimal quantization of coefficients

After the pdfs of the coefficients have been modeled, the coefficient values need to be quantized. For this an optimal Lloyd-Max [4, 5] quantizer for minimum mean square error is employed, based upon the parameterized pdf models.

6. Entropy coding

Once all of the coefficients have been quantized, the entire image can be represented by a finite symbol set. However, in order to efficiently assign codewords to the symbols, the probability of each symbol needs to be calculated. The probabilities for all coefficient reconstruction levels are computed based upon the parameterized coefficient pdfs. Likewise, the probability of occurrence of each vector is calculated by integration of its position pdf model.

The symbols representing the coefficients and vectors are then assigned codewords using a modified Huffman [6] coder. Since it is possible to have a large number of symbols, a regular Huffman coder could not be implemented. This is because the codeword lengths become prohibitively long for a large number of symbols [2]. The modified Huffman coder was developed such that the longest codeword length is 256 bits. If there are more than 256 symbols to be coded, successive codebooks are used, where the last symbol in each codebook points to the next codebook. The modified Huffman codebook efficiencies consistently exceeded 98% in experiments [7].

7. Image Reconstruction

The first task in image reconstruction is to recreate the vector position and coefficient pdfs from the α and β parameters which were transmitted or stored in full 32-bit real precision. Once the pdfs are available, the probability of each coefficient and vector position can be computed. Based on these probabilities, the Huffman codewords can be inverse mapped into the appropriate symbols. Each block of the image is then reconstructed by multiplying all of the specified basis vectors by their coefficients, and then summing all of the components. Once a block is reconstructed, the results are used to predict the same block on the next line, since it is actually the difference signal that is coded.

8. Experimental Results

The MBR procedure was applied to the "LENA" image, using 32 sample blocks and a stopping MSE threshold of 4 percent. Examples of the resulting coefficient and position pdfs, together with their models, are shown in Figures 1 through 4. The model pdfs are then used for quantization, where 16, 8, 8, and 8 quantization levels were used for the DC, DCT, FHT, and IDT coefficients, respectively. The actual MSE achieved in the reconstruction was 0.90 percent, and the reconstructed image was represented with 0.78 bits per pixel. The original and reconstructed images are shown in Figures 5 and 6 respectively.

9. Conclusions

The image data compression technique presented in this paper combines the advantages of several conventional compression methods, yet it avoids many of the associated disadvantages. By using an expanded set of representation vectors, excellent energy compaction is achieved. In addition, prohibitive computational complexity is avoided by choosing component basis sets for which fast orthogonal transformations are available. An experimental result for "LENA" indicates that the MBR technique can produce a high quality reconstruction at well under 1 bit per pixel.

REFERENCES

[1] F. G. Safar, Signal Compression and Reconstruction Using Multiple Bases Representation, M. S. Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA, June 1988.

[2] A. K. Jain, Fundamentals of Digital Image Processing, (Englewood Cliffs, NJ: Prentice Hall, 1989).

[3] J. S. Milton and J. C. Arnold, Probability and Statistics in the Engineering and Computing Sciences, (New York: McGraw-Hill, Inc., 1986).

[4] S. P. Lloyd, "Least Square Quantization in PCM," Bell Telephone Laboratories Memo., Murray Hill, NJ, 1957.

[5] J. Max, "Quantizing for Minimum Distortion," *IRE Transactions on Information Theory*, Vol. IT-6, pp. 7-12, March 1960.

[6] D. A. Huffman, "A Method for the Construction of Minimum-Redundancy Codes," *IEEE Proceedings of the IRE*, Vol. 40, No. 9, pp. 1098-1101, 1952.

[7] R. Khanna, Image Data Compression Using Multiple Bases Representation, M. S. Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA, September 1990.

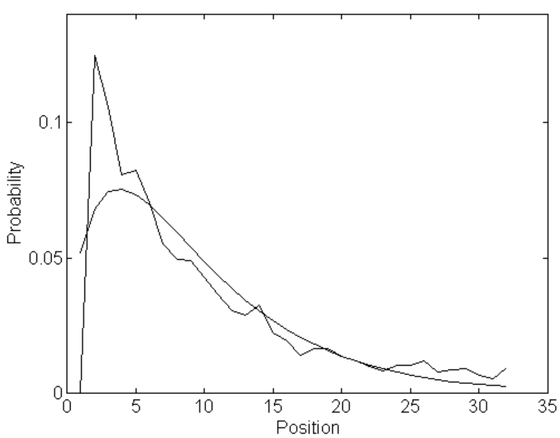

Figure 1: DCT Position PDF

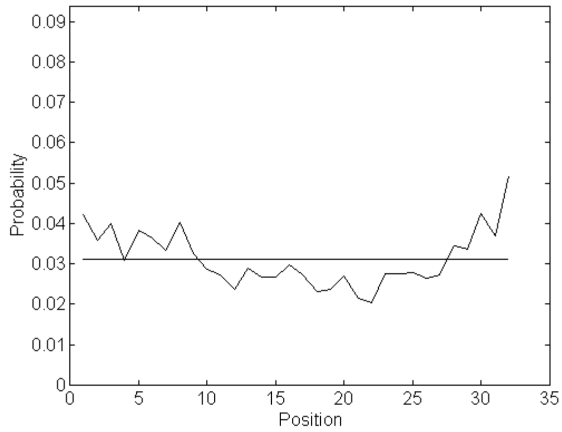

Figure 2: IDT Position PDF

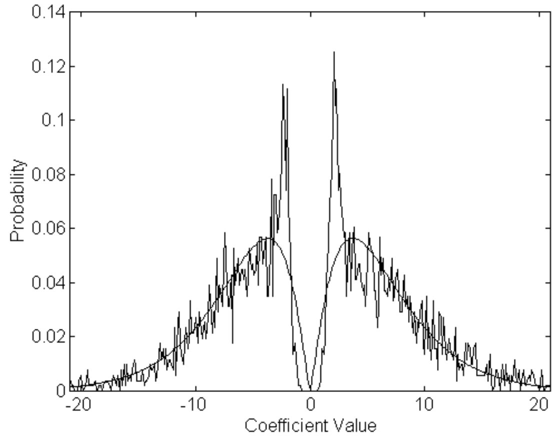

Figure 3: FHT Coefficient PDF

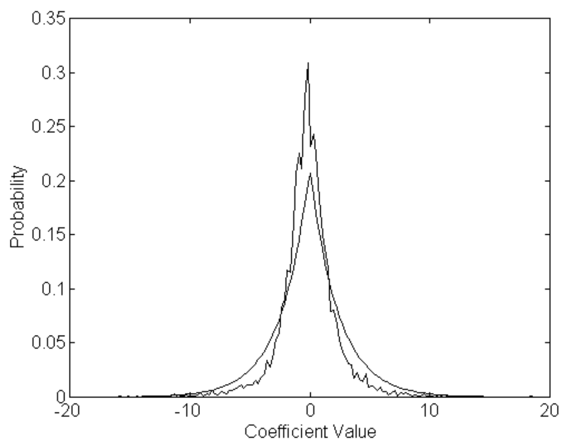

Figure 4: DC Coefficient PDF