JOURNAL OF VISUAL COMMUNICATION AND IMAGE REPRESENTATION
Vol. 8, No. 1, March, pp. 21–26, 1997
ARTICLE NO. VC970323

Lossless Image Compression Using the Discrete Cosine Transform

Giridhar Mandyam,*,† Nasir Ahmed,‡ and Neeraj Magotra‡

†*Texas Instruments Inc., P.O. Box 660199, Dallas, Texas 75266-0199,* ‡*Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, New Mexico 87131*

Received March 6, 1995; accepted March 14, 1996

In this paper, a new method to achieve lossless compression of two-dimensional images based on the discrete cosine transform (DCT) is proposed. This method quantizes the high-energy DCT coefficients in each block, finds an inverse DCT from only these quantized coefficients, and forms an error residual sequence to be coded. The number of coefficients used in this scheme is determined by using a performance metric for compression. Furthermore, a simple differencing scheme is performed on the coefficients that exploits correlation between high energy DCT coefficients in neighboring blocks of an image. The resulting sequence is compressed by using an entropy coder, and simulations show the results to be comparable to the different modes of the lossless JPEG standard.

1. INTRODUCTION

Compression of images is of great interest in applications where efficiency with respect to data storage or transmission bandwidth is sought. Traditional transform-based methods for compression, while effective, are lossy. In certain applications, even slight compression losses can have enormous impact. Biomedical images or synthetic-aperture radar images are examples of imagery in which compression losses can be serious.

The discrete cosine transform (DCT) has been applied extensively to the area of image compression. It has excellent energy-compaction properties, and as a result has been chosen as the basis for the Joint Photography Experts' Group (JPEG) still-picture compression standard. However, losses usually result from the quantization of DCT coefficients, where this quantization is necessary to achieve compression. In this paper, an alternative lossless method is proposed which takes advantage of not only the energy compaction properties of the DCT, but also the correlation that exists between high-energy coefficients in neighboring transformed blocks of data.

* Supported by NASA Grant NAGW 3293 obtained through the Microelectronics Research Center, The University of New Mexico.

2. THE ALGORITHM

The DCT has long been used as a method for image coding and has now become the standard for video coding [2]. Its energy compaction capability makes it ideal for efficient representation of images. Given a square image block $\mathbf{F}$ of size m by m, an m by m matrix $\mathbf{C}$ is defined by the equation:

$$\begin{aligned} C_{ij} &= \frac{1}{\sqrt{m}} \quad i = 0 \quad j = 0 \ldots m-1 \\ &= \sqrt{\frac{2}{n}} \cos \frac{(2i+1)j\pi}{2m} \quad i = 1 \ldots m-1 \quad j = 0 \ldots m-1. \end{aligned}$$

Thus the DCT of $\mathbf{F}$ is defined as

$$\mathbf{f} = \mathbf{C}\mathbf{F}\mathbf{C}^{\mathbf{T}} \tag{1}$$

The DCT is a unitary transform, meaning that the inversion can be accomplished by

$$\mathbf{F} = \mathbf{C}^{\mathbf{T}}\mathbf{f}\mathbf{C}. \tag{2}$$

Unfortunately, the DCT coefficients, i.e., the entries in $\mathbf{f}$, are evaluated to infinite precision. In traditional coding methods based on the DCT, all compression and all losses are determined by quantization of the DCT coefficients. Even for lossless image compression, this problem cannot be avoided, because storing the coefficients to their full precision (which is determined by the machine one is using) would not yield any compression. What is proposed is to evaluate all entries of the DCT matrix out to only B digits past the decimal point. This means that the DCT coefficients will have precision out to $2B$ digits past the decimal point. A major consequence of this action is that the resulting DCT matrix is no longer unitary, and the inverses of the DCT matrix and its transpose must be evaluated explicitly, i.e.,

$$\mathbf{F} = \mathbf{C}^{-1}\mathbf{f}(\mathbf{C}^{\mathbf{T}})^{-1}. \tag{3}$$

1047-3203/97 $25.00

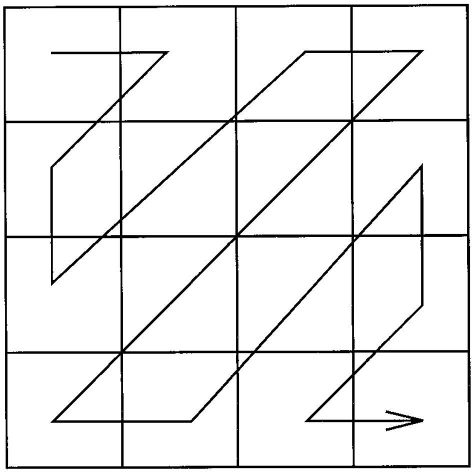

FIG. 1. Highest to lowest energy coefficients.

Of course, one must choose B such that these inverses exist. For instance, in the case where B is 1, i.e., the entries of $\mathbf{C}$ are evaluated to only one place beyond the decimal point, the inverse matrix does not exist in the 8 by 8 pixel block case. Since the entire operation involves matrix multiplications, a total of $2m^3$ multiplications and $2m^2(m-1)$ additions are required to evaluate all the entries of $\mathbf{F}$ in (3).

Once the DCT coefficients have been computed, we retain only w high energy coefficients to be used for the calculation of an approximation to the original data matrix $\mathbf{F}$. One needs to choose w such that a desired amount of energy compaction is obtained. The high-energy coefficients in general will appear in the same locations inside $\mathbf{f}$; e.g. the three highest-energy coefficients always appear in f_{00}, f_{01}, and f_{11}, the upper-left corner of $\mathbf{f}$ (for a detailed discussion, see [3]). In Fig. 1, the high to low energy coefficients are scanned for an 8 by 8 DCT block. The remaining DCT coefficients are assumed to be zero. Then the inverse-DCT is calculated (without assuming unitarity) and a resulting matrix $\mathbf{Fn}$ results. From this, an error residual matrix $\mathbf{E}$ can be defined:

$$E_{ij} = F_{ij} - Fn_{ij}. \tag{4}$$

By retaining $\mathbf{E}$ and the w quantized DCT coefficients, perfect reconstruction of $\mathbf{F}$ can be achieved.

After selecting the high energy DCT coefficients, we perform linear prediction on the nonzero DCT coefficients by using a simple differencing scheme. Between neighboring blocks there exists some correlation between the corresponding high energy coefficients. After specifying w of these high energy coefficients, each of these coefficients can be encoded as the error residual resulting from subtracting the corresponding DCT coefficient from a neighboring block; in this paper, the block immediately to the left is chosen, except for any of the leftmost blocks in the image, which use the block immediately above. As a result, the high energy coefficients are decorrelated between blocks, and the overall *first-order* entropy of the resulting data is decreased. In addition, the entropy of the entries of $\mathbf{E}$ are of lower entropy than the original data matrix; therefore, a suitable entropy coder can be used, such as an adaptive Huffman coder or an arithmetic coder. Moreover, since we are transmitting error residuals, the choice of the parameter B, which determines the precision of the transmitted DCT coefficients, becomes less crucial, as increasing B will result in little decrease of $\mathbf{E}$ in many cases. For instance, in the 8 by 8 case, experimentation showed that $B = 2$ was adequate to achieve maximal compression (as stated before, for $B = 1$ the DCT matrix inverses do not exist in the 8 by 8 case). The minimum choice of B is case-dependent.

2.1. *Determination of w*

As one increases the number of high-energy coefficients retained, the first-order entropy of the entries of $\mathbf{E}$ steadily decreases. Unfortunately, a tradeoff exists in that the memory required to store the high-energy coefficients increases, since, even with differencing, these coefficients are still of relatively high entropy. So one must find a middle point, and to do so, the following performance metric called the *potential compression statistic, p*, is proposed:

$$\begin{aligned} p(w) = \sum_i \{&\text{No. of bits needed to store } w \\ &\text{coefficients at block } i \; + \qquad (5) \\ &\text{First Order entropy of } \mathbf{E} \text{ at block } i\}. \qquad (6) \end{aligned}$$

As w increases from 1 to m^2 (each data block being m by m pixels), $p(w)$ will reach a global minimum. The value of w at that minimal $p(w)$ is the value used for that particular image. The main reason that p would vary from one image to another is that a particular scan (Fig. 1) of the high to low energy coefficients in each block has been chosen, and this scan may not correspond to the actual ordering of these coefficients in a particular block [6]. However, the scan that has been chosen has been shown to be optimal under the assumption that the data follows a first-order Markov process [7].

3. EXPERIMENTS

The algorithm was tested on three standard grayscale 256 by 256 pixel images (from the University of Southern California's image processing database), each quantized

FIG. 2. Cameraman.

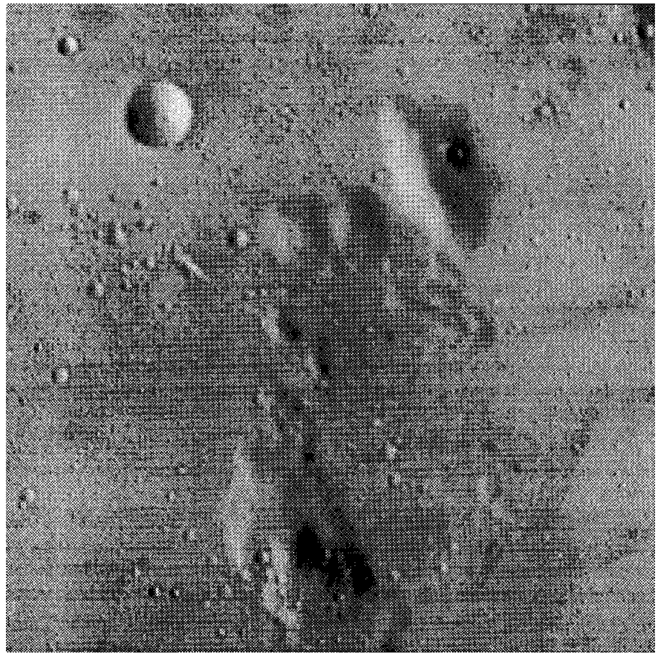

FIG. 4. Moon.

to 8 bits per pixel. The original images are shown in Figs. 2–4. An 8 by 8 block was used to obtain the DCT coefficients. By using the *potential compression statistic,* it was found that the optimal value for w was 3 for all of the test images. The value for B was two; moreover, in all experiments it was found that after deriving the DCT coefficients for each block, the precision of these coefficients could be reduced by a factor of 1000 (i.e. three decimal digits) without affecting the first-order entropy of the error residuals; the advantage of this is a reduction in first-order entropy of the DCT coefficients. The proposed method was compared to seven fixed filters for the present lossless JPEG standard (given in Table 1) [10], with the first-order entropy of the error residuals for both methods given Table 2, where the proposed method is designated by "MDCT," denoting modified DCT. There are three overhead values per block for the MDCT method; for a 256 by 256 image, this implies that there are 3072 overhead values assuming a block size of 8 by 8. For the JPEG lossless filters, the overhead is practically nothing, except for a handful of

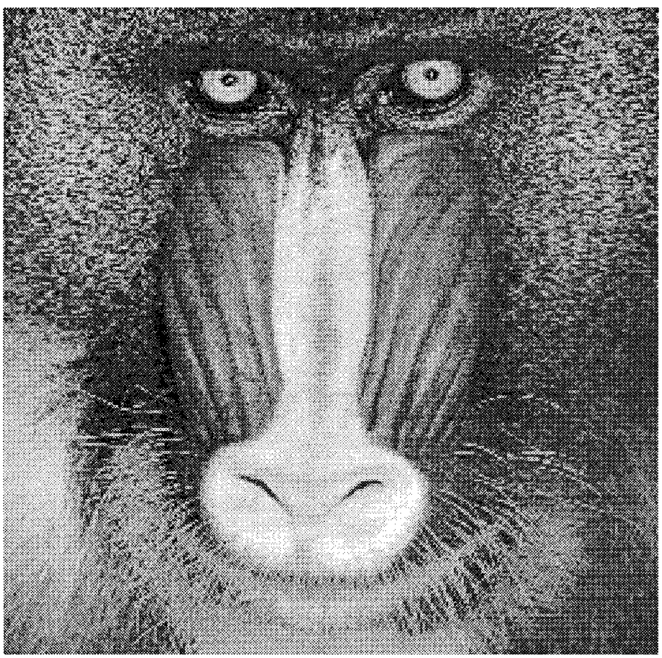

FIG. 3. Baboon.

TABLE 1
Prediction Modes for Lossless JPEG (a Is a Left-Neighboring Pixel, b Is an Upper-Neighboring Pixel, and c Is an Upper-Left-Neighboring Pixel)

Lossless JPEG filter	Method of prediction
JPEG 1	a
JPEG 2	b
JPEG 3	c
JPEG 4	$a + b - c$
JPEG 5	$a + (b - c)/2$
JPEG 6	$b + (a - c)/2$
JPEG 7	$(a + b)/2$

TABLE 2
Entropy Results for Test Images (bits/value)

	Cameraman	Baboon	Moon
Original image	6.90	6.78	6.81
MDCT error residuals	5.71	6.57	5.17
MDCT overhead	4.89	5.45	4.48
JPEG 1	5.44	5.78	5.50
JPEG 2	5.31	5.93	5.23
JPEG 3	5.66	6.02	5.52
JPEG 4	5.53	6.18	5.58
JPEG 5	5.38	6.86	5.35
JPEG 6	5.31	6.90	5.27
JPEG 7	5.22	6.62	5.05

TABLE 4
Compression Results for Rice Coding

	Cameraman	Baboon	Moon
MDCT	32.19%	16.35%	32.43%
JPEG 1	31.21%	15.32%	30.42%
JPEG 2	34.02%	14.43%	33.95%
JPEG 3	29.13%	12.72%	30.10%
JPEG 4	30.04%	11.70%	29.39%
JPEG 5	32.00%	15.00%	32.17%
JPEG 6	32.94%	14.63%	33.26%
JPEG 7	34.18%	17.67%	35.99%

startup values for prediction. Knowing this, *projected* compression performance was evaluated using the entropy values of Table 2; the results are given in Table 3, where the performance criterion used is the percentage size reduction (with respect to the number of bytes) of the original file, given by

$$\frac{\text{Original File Size} - \text{Compressed File Size}}{\text{Original File Size}} \times 100\%. \quad (7)$$

It can be seen from the results in Table 3 that the new algorithm can be *expected* to perform approximately as well as the seven lossless modes of JPEG. However, upon application of a standard entropy coder, compression results could end up different than what was predicted. As an example, a Rice coder ([11, 12]) was used for entropy coding the error residuals for both the proposed method and the lossless JPEG modes. The advantage of the Rice coder is that it is an adaptive coder which conforms appropriately to different data blocks *without using codebooks;* this also yields an efficient hardware implementation [12]. The data block size used was 64; this was to take advantage of the possible differences in entropy between different DCT error residual blocks. This same block size was applied to the seven JPEG modes. An adaptive Huffman coder on the other hand was used to code the overhead associated with the proposed DCT-based method; this was due to the fact that the Rice coder chooses a "winner" out of several fixed entropy coders to compress a sequence, and for each 8 by 8 image data block, the three low frequency coefficients retained for reconstruction are uncorrelated and do not in general follow a distribution corresponding to these fixed entropy coders. The compression results are given in Table 4, from which it can be seen that the proposed method provides very close performance to most of the JPEG lossless modes. It must be noted here that if a Rice coder was to be run in conjunction with any of the lossless JPEG modes, one would most likely not use such a large block size, as what is gained in reduction in overhead is lost by a reduction in *adaptability*. Therefore, in accordance with the recommendations of the JPEG committee, an adaptive Huffman coder was used to do all the coding for each method, with the results given in Table 5. Although the performance of the proposed method worsens under this kind of entropy coding and the lossless JPEG modes' performance improves when compared to the previous Rice coding example, the proposed method still outperforms several lossless JPEG modes for each of the three images.

TABLE 3
Projected Compression Results for Test Images

	Cameraman	Baboon	Moon
MDCT	26.87%	16.12%	33.62%
JPEG 1	32.00%	27.75%	31.25%
JPEG 2	33.62%	25.87%	34.62%
JPEG 3	29.25%	24.75%	31.00%
JPEG 4	30.87%	22.75%	30.25%
JPEG 5	32.75%	14.25%	33.12%
JPEG 6	33.62%	13.75%	34.12%
JPEG 7	34.75%	17.25%	36.87%

TABLE 5
Compression Results for Adaptive Huffman Coding

	Cameraman	Baboon	Moon
MDCT	24.41%	13.66%	31.98%
JPEG 1	26.59%	26.75%	30.37%
JPEG 2	31.09%	24.14%	33.91%
JPEG 3	22.65%	22.78%	30.21%
JPEG 4	28.32%	18.55%	29.57%
JPEG 5	30.63%	12.35%	32.29%
JPEG 6	31.72%	11.71%	33.30%
JPEG 7	33.12%	16.32%	36.30%

4. CONCLUSIONS

A new lossless image compression scheme based on the DCT was developed. This method caused a significant reduction in entropy, thus making it possible to achieve compression using a traditional entropy coder. The method performed well when compared to the popular lossless JPEG method. Future work will focus on finding an efficient hardware implementation, possibly taking advantage of commonality between the new method and the existing DCT-based lossy JPEG method.

ACKNOWLEDGMENTS

The authors thank Mr. Jack Venbrux of the NASA Microelectronics Research Center at the University of New Mexico for his help with the Rice coding algorithm. The authors also thank the anonymous reviewers for their constructive comments and suggestions.

REFERENCES

1. G. Mandyam, N. Ahmed, and N. Magotra, A DCT-based scheme for lossless image compression, *IS&T/SPIE Electronic Imaging Conference,* San Jose, CA, February, 1995.
2. N. Ahmed, T. Natarajan, and K. R. Rao, Discrete Cosine Transform, *IEEE Trans. Comput.* **C23,** January 1974, 90–93.
3. K. R. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages, and Applications.* Academic Press, San Diego, 1990.
4. A. K. Jain, *Fundamentals of Digital Image Processing,* Prentice–Hall, Englewood Cliffs, NJ, 1989.
5. M. F. Barnsley, and L. Hurd, *Fractal Image Compression,* A. K. Peters, Wellesley, MA, 1993.
6. W. H. Chen and C. H. Smith, Adaptive coding of monochrome and color images. *IEEE Trans. Commun.* **25**(11), November 1977, 1285–1292.
7. P. Yip and K. R. Rao, Energy packing efficiency for the generalized discrete transforms, *IEEE Trans. Commun.* **26**(8), August 1978, 1257–1262.
8. I. H. Witten, R. M. Neal, and J. G. Cleary, Arithmetic coding for data compression, *Commun. ACM.* **30**(6), June 1987, 520–540.
9. R. E. Blahut, *Principles and Practice of Information Theory,* Addison–Wesley, Menlo Park, CA, 1990.
10. P. E. Tischer, R. T. Worley, and A. J. Maeder, Context based lossless image compression, *Comput. J.* **36**(1), January 1993, 68–77.
11. R. F. Rice, P. S. Yeh, and W. H. Miller, Algorithms for a very high speed university noiseless coding module, *JPL Publication 91-1,* Pasadena, CA, JPL Publication Office, February 15, 1991.
12. J. Venbrux, P. S. Yeh, and M. N. Liu, A VLSI chip set for high-speed lossless data compression. *IEEE Trans. Circuits Syst. Video Technol.* **2**(4), December 1992, 381–391.
13. R. B. Arps and T. K. Truong, Comparison of international standards for lossless still image compression, *Proceedings of the IEEE.* Vol. 82. No. 6. June, 1994. pp. 889–899.
14. N. D. Memon and K. Sayood, Lossless image compression: A comparative study, *IS&T/SPIE Electronic Imaging Conference.* San Jose, CA, February, 1995.

GIRIDHAR MANDYAM was born in Dallas, Texas on October 15, 1967. He received the B.S.E.E. degree Magna Cum Laude from Southern Methodist University in 1989. From 1989 to 1991 he was employed with Rockwell International in Dallas. From 1991 to 1993 he was a teaching assistant in the Signal and Image Processing Institute (SIPI) at the University of Southern California, where he obtained the M.S.E.E. degree in May 1993. In May 1993, he joined Qualcomm, Inc., in San Diego, CA as an engineer in the Systems Group. In January 1994, he took a leave of absence to pursue doctoral studies at the University of New Mexico under a fellowship provided by the NASA Microelectronics Research Center (MRC) at UNM. He defended his dissertation in December, 1995. He has organized and chaired sessions at the IEEE ISE '95 Conference in Albuquerque, New Mexico; the IEEE Asilomar '95 Conference in Asilomar, California; and the IEEE ISCAS '96 Conference in Atlanta, Georgia. His interests are image coding, image restoration, adaptive algorithms, spectral estimation, orthogonal transforms, systolic architectures, and spread spectrum communications. He joined the Wireless R & D group at Texas Instruments in Dallas, Texas in April, 1996.

NASIR AHMED (Fellow, IEEE 1985), was born in Bangalore, India, in 1940. He received the B.S. degree in electrical engineering from the University of Mysore, India, in 1961, and the M.S. and Ph.D. degrees from the University of New Mexico in 1963 and 1966, respectively. From 1966 to 1968 he worked as a Principal Research Engineer in the area of information processing at the Systems and Research Center, Honeywell, Inc., St. Paul, Minnesota. He was at Kansas State University, Manhattan, from 1968 to 1983. Since 1983 he has been a Professor of Electrical and Computer Engineering (EECE) at the University of New Mexico, Albuquerque. In August 1985 he was awarded one of twelve Presidential Professorships at the University of New Mexico, and was also elected Fellow of the Institute of Electrical and Electronic Engineers for his contributions to digital signal processing and engineering education. He became the Chairman of the EECE Department in July 1989. He is the leading author *Orthogonal Transforms for Digital Signal Processing* (Springer Verlag, 1975), and *Discrete-Time Signals and Systems* (Reston, 1983), and co-author of *Computer Science Fundamentals* (Merrill, 1979). Dr. Ahmed is also the author of numerous technical papers in the area of digital signal processing. He was an Associate Editor for the *IEEE Transactions on Acoustics, Speech, and Signal Processing* (1982–1984) and is currently an Associate Editor for the *IEEE Transactions on Electromagnetic Compatibility* (Walsh Functions Applications). Dr. Ahmed served as the Interim Dean of Engineering at the University of New Mexico from October 1994 to January 1996. He is now the Interim Associate Provost for Research and Dean of Graduate Studies at the University of New Mexico.

NEERAJ MAGOTRA was born in Jamshedpur, India on December 5, 1958. He obtained his B.Tech. in Electrical Engineering from the Indian Institute of Technology (Bombay, India) in 1980, his M.S. in Electrical Engineering from Kansas State University in 1982 and his Ph.D. in Electrical Engineering from the University of New Mexico in 1986. From 1987 until 1990 he held a joint appointment with Sandia National Laboratories and the Department of Electrical and Computer Engineering (ECE) at the University of New Mexico, Albuquerque, New Mexico. He is currently employed as an Associate Professor in the ECE Department at the University of New Mexico. He has been involved in signal/image processing research in the areas of seismic, speech, and radar (synthetic array radar imaging) signal processing for the past fourteen years. The research involved the theoretical design of algorithms as well as their real time implementation on DSP chips. He has served as past associate editor of the IEEE transactions on Signal Processing and served on the organizing committee of ICASSP 1990. He has also helped organize special sessions for the IEEE Asilomar Conference on Signals, Systems, and Computers, the Midwest symposium on Circuits and Systems, and the International Symposium on Circuits and Systems. He is a member of Sigma Xi, Phi Kappa Phi, Tau Beta Pi, the Institute of Electrical and Electronics Engineers (IEEE), and the Seismological Society of America. He has authored/co-authored over 70 technical articles including journal papers, conference papers, and technical reports.