Proceedings of the IEEE 1993 Data Compression Conference, Snowbird, Utah, pp. 195-203, April, 1993.

Full-Frame Compression of Tomographic Images Using the Discrete Fourier Transform

John D. Villasenor
Electrical Engineering Department
University of California, Los Angeles
405 Hilgard Avenue
Los Angeles, CA 90024-1594

Abstract

The unacceptability of block artifacts in medical image data compression has led to systems employing full-frame Discrete Cosine Transform (DCT) compression. Although the DCT is the optimum fast transform when block coding is used, we show here that it is outperformed by the Discrete Fourier Transform (DFT) and Discrete Hartley Transform (DHT) for images obtained using Positron Emission Tomograpy (PET) and Magnetic Resonance Imaging (MRI). This difference occurs because PET and MRI images are characterized by a roughly circular region D of non-zero intensity bounded by a region R in which the image intensity is essentially zero. Clipping R to its minimum extent can reduce the number of low-intensity pixels but the practical requirement that images be stored on a rectangular grid means that a significant region of zero intensity must remain an integral part of the image to be compressed. The DCT therefore loses its advantage over the DFT because neither transform introduces significant artificial discontinuities.

1. Introduction

The volume of medical image data is already very high, and will continue to grow as the resolution of current imaging modalities increases and the cost of using imaging techniques drops. Tomographic techniques such as PET, and MRI are of particular interest in compression because of their high dimensionality. Pressure to develop efficient compression algorithms will come not only from limitations in storage capacity, but from an increased need to transmit and provide remote access to medical images over existing telephone and other communications channels.

While image compression for non-medical applications has been the subject of intensive research in recent years, the resulting algorithms are often incompatible with the stringent image quality requirements of medical applications. For example, the JPEG [1] (Joint Photographic Experts Group) standard achieves compression by dividing the image into 8-pixel by 8-pixel blocks, with each block compressed using a discrete cosine transform (DCT). The block artifacts generated in the reconstructed

0-8186-3392-1/93 $3.00 © 1993 IEEE

image by this approach can render block coding unacceptable for medical applications because of the potential for misdiagnosis. Compression techniques are needed that meet the specific requirements of diagnostic accuracy and practicability for medical applications.

Current tomographic data volumes are still manageable (a typical PET study consists of approximately 15 image planes, each containing 128 by 128 pixels, all imaged over approximately 10 time steps), but are quickly growing. Whole-body PET involves as many as 500 planes, leading to an order of magnitude increase in data volume. While MRI has not traditionally involved the temporal variation inherent in many PET studies, it produces large amounts of data due to its higher resolution. MRI images contain 256 by 256 pixels and involve as many as 10 planes. Furthermore, new "fast" MRI techniques have introduced a temporal dimension to MRI imaging. Resolution of both PET and MRI has improved significantly in recent years and will continue to improve with advances in reconstruction algorithms. This coupled with increased clinical use will result in a need for practical, accurate tomographic compression techniques. There is growing interest in the feasibility of applying irreversible image compression to medical images [2, 3, 4]. Efficient irreversible compression techniques can be designed to yield reconstructed images which are of extremely high quality. Furthermore, studies have demonstrated that irreversible compression of digitized radiographs yields reconstructed images that are indistinguishable from the originals in terms of diagnostic accuracy [5].

The choice of the DCT for image compression is due to the optimal nature of the DCT for compression of most images. However, the assumptions used to determine this optimality no longer apply for tomographic images. We show here that when more appropriate image models are used, the DFT and DHT are superior to the DCT for compression of tomographic images. In contrast with most image classes, tomographic images have a consistent general structure consisting of an approximately circular region containing almost all of the image energy. While precise characterization of the size and form of this region is not possible, the nature of tomography as a method for cross-sectional imaging guarantees this basic structure and the associated low- or zero-intensity bounding region.

Transform compression, or coding, [7] relies on the use of image transforms to furnish an equivalent but alternative representation of the information in the image. If the transform is properly chosen, the energy in the image will be concentrated into a relatively small percentage of the transform coefficients; it is these coefficients which are quantized, stored, and then later used for generating the reconstructed image. The extent of distortion in the reconstructed image is a function of the manner in which the transform coefficients are quantized. Saving few coefficients yields an extremely high compression ratio but a noisy reconstructed image, while more coefficients gives a cleaner image but lower compression efficiency. Research on transform coding applied to medical images suggests that compression ratios of 10:1 to 20:1 can be obtained without sacrificing diagnostic accuracy when the transform coefficient quantization scheme is optimized [5].

Image compression has been most extensively studied in the context of video signals [8], where demands on the transmission channel are often severe and real-time performance is a requirement. These rigid bandwidth and speed constraints have forced tradeoffs in compression error and algorithm design, leading to the block transform approach which has become the standard for video compression [1, 9]. In block transform coding, the image is considered in blocks of 8 by 8 or 16 by 16 pixels, with a subset of the coefficients corresponding to each transformed block sent through the channel and used in reconstruction. The blocking artifacts visible in the reconstructed image render block coding unacceptable for medical applications. Instead, medical images have been compressed using full-frame techniques which treat the entire image as a "block," thereby avoiding the problem of blocking artifacts at the expense of a slightly increased computational burden. The great majority of computed tomography (CT) images possess low- or zero-intensity regions near image edges. This contrasts strongly with the image blocks used in video coding, in which edge pixels are no more or less active than those in the center of the block.

Section 2 discusses the energy compaction of the DCT, DFT, and DHT for a one-dimensional, first-order stationary Markov sequence. This leads to the well-known result that the DCT is the best fast transform when the interpixel correlation is high. The energy compaction properties are then treated under the assumption that the sequence is stationary Markov over its central portion and zero at both borders. This is a much more appropriate description of a typical tomographic image. In this second derivation, the DCT is no longer the optimal transform. Section 3 contains an extension of the argument to two dimensions, and Section 4 includes a comparison of the results of applying compression based on the the DCT and DFT to a PET image. The DHT, or Hartley transform [13, 14], is equivalent in terms of compression performance to the DFT, and in the case of tomographic images, superior to that of the DCT. The DHT is a real transform which has been shown to be essentially equivalent to the Fourier transform in all aspects, with the important difference that it produces real transforms from real signals. This allows the use of more elegant fast transform algorithms when the input data are real.

2. Comparison of Energy Compaction

The suitability of a given transform for image compression depends on the efficiency with which it can pack the energy of the image into a chosen number of transform coefficients. The expected variances of the transform coefficients can be expressed simply in terms of the second-order statistics of the input sequence.

Consider a random vector $u(n)$ of length N, where $0 \leq n \leq N-1$. The transform $v(k)$ can then be expressed $\mathbf{v} = \mathbf{T}\mathbf{u}$ where $\mathbf{v}$ and $\mathbf{u}$ are $N \times 1$ vectors and $\mathbf{T}$ is an $N \times N$ unitary transform matrix whose elements $T(k, n)$ are determined in accordance with the transform (i.e. the DCT, DFT, or DHT) to be implemented. Equations for

$T(k,n)$ is as follows:

$$\text{DCT: } T(k,n) = \begin{cases} 1/\sqrt{N} & k = 0, 0 \le n \le N-1 \\ \sqrt{2/N} \cos \frac{\pi(2n+1)k}{2N} & \text{other } k, n \end{cases} \tag{1}$$

$$\text{DFT: } T(k,n) = \frac{1}{\sqrt{N}} \exp(-j2\pi kn/N) \tag{2}$$

$$\text{DHT: } T(k,n) = \frac{1}{\sqrt{N}} [\cos(2\pi kn/N) + \sin(2\pi kn/N)] \tag{3}$$

Let $\mathbf{u}$ and $\mathbf{v}$ have means $\mathbf{m_u}$ and $\mathbf{m_v}$, and covariances $\mathbf{R_u}$ and $\mathbf{R_v}$ respectively. The vectors $\mathbf{m_i}$, $i = u, v$ have size $N \times 1$ and the covariance matrices $\mathbf{R_i}$ are $N \times N$. The covariance of the transform is then

$$\mathbf{R_v} = \mathbf{T R_u T^{*t}} \tag{4}$$

where $\mathbf{T^{*t}}$ is the conjugate of the (non-hermitian) transpose of $\mathbf{T}$. The variance σ_k^2 of transform coefficient $v(k)$ is the kth diagonal element of $\mathbf{R_v}$, and can be expressed from (4) as

$$\sigma_k^2 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} T(k,i) R_u(i,j) T^*(k,j) \tag{5}$$

If $\mathbf{u}$ is a Markov sequence with lag-1 correlation ρ, elements of $\mathbf{R_u}$ are given by $R_u(i,j) = \rho^{|i-j|}$. When the correlation ρ is high (typically taken to mean $\rho \ge 0.5$), substitution of this $\mathbf{R_u}$ into (4) shows that the DCT packs a higher portion of the signal energy in the first k coefficients than the DFT (or the DHT) and is therefore the best choice for compression [11]. It is this well-known result which has led to the choice of the DCT for block transforms in JPEG and MPEG.

Now consider a sequence $u(n)$ which is Markov for $m < n < N-1-m$, but whose first and last m values are known *a priori* to be zero. These (approximately) zero boundaries occur in almost all of the scan lines of many medical images. By explicitely considering the zero-intensity bounding regions, we are adopting a model which is fundamentally different from and more accurate than the stationary Markov model which is often used for medical images. We are still left with the assumption of stationarity over the non-zero regions of the image; we justify this in view of mathematical simplifications it allows, and because in contrast with the boundary regions, the nature of nonstationarities in the central region varies between image classes. The covariance matrix of this sequence will have zeros in its first and last m rows and columns, with a square region of non-zero entries in its center. Specifically,

$$R_u(i,j) = \begin{cases} 0 & 0 \le i,j \le m-1 \\ 0 & N-m \le i,j \le N-1 \\ \rho^{|i-j|} & \text{otherwise} \end{cases} \tag{6}$$

Use of this covariance matrix in (4) yields a set of variances for the transform coefficients corresponding to the truncated Markov sequence.

Figure 1a shows the energy compaction for the DFT and DCT for $N = 16$, $\rho = .85$, and $m = 0$. Theses curves were calculated using equation (5) with the covariance function $R_u(i,j) = \rho^{|i-j|}$. The figure gives the fraction of the total sequence energy preserved as a function of the number of coefficients k retained. It is assumed that the coefficients are arranged in order of decreasing energy. For the DFT, this involves separate consideration of the real and imaginary parts of the coefficients. When the redundancy of the DFT is exploited, there will be 16 independent values. The curve for the DHT would be identical to that for the DFT because the energy contained in origin-symmetric pairs of DFT and DHT coefficients is identical. Since $m = 0$, the sequence is stationary, and as expected the DCT has superior compression performance for any given k. Figure 1b results from replacing the covariance function used in Figure 1a with that given by equation (6), and shows the compression performance if zero boundary conditions are enforced by setting $m = 1$. As before, the length of the sequence is $N = 16$ and the correlation ρ is 0.85. The superior energy compaction efficiency of both the DFT and the DHT over the DCT is now evident. For any value of k, retaining k coefficients from the DFT or DHT will always give better compression than k coefficients from the DCT.

The assumption of a zero-intensity border region is not strictly true in practice. Consideration of low but non-zero pixel intensities along the borders can be modeled by appropriate modification of the covariance matrix in (6). As the energy in the border pixels is increased the advantage of the DFT over the DCT will gradually diminish. The exact threshold at which the DCT again becomes superior is a function of N and the assumed covariance matrix. For the example discussed above, the DFT remains superior even when a typical border pixel contains 20% as much energy as a pixel located in the image center. This is a comfortable margin since in typical PET images the relative strength of border pixels is much lower than 20%. An alternative solution involves simply replacing data along image borders with a smooth, low-intensity function which reaches zero near the image edge. This replacement can be restricted to areas outside the actual image target (e.g. the brain) to avoid any damage to data of interest.

3. Extension to Two Dimensions

The above arguments can be extended to two dimensions with analagous results. Assuming that the transform being used is separable (i.e. that the 2D transform can be obtained by taking successive transforms of the rows and then columns of the image), the image $u(m,n)$ and its transform $v(k,l)$ are related by

$$\mathbf{v} = \mathbf{T}\mathbf{u}\mathbf{T}, \tag{7}$$

where the matrix $\mathbf{T}$ is identical to that used in the one-dimensional case, and $\mathbf{u}$ and $\mathbf{v}$ are now $N \times N$ images. Non-square images can be treated in an analogous manner; we make the square assumption to simplify notation. The assumption of separability holds for the DFT and DCT but not for the DHT. A fast algorithm for the 2D DHT, however, does exist [13], and the results for the 2D DFT will apply to the 2D DHT in analogy with the 1D case. (Another proof of the energy equivalence of the Fourier and Hartley transforms in any number of dimensions can be made by considering the squares of sums of pairs of corresponding Fourier and Hartley transform coefficients).

The variance $\sigma_v^2(k,l)$ of the transform coefficients $v(k,l)$ can be written [11]

$$\sigma_v^2(k,l) = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1}\sum_{m'=0}^{N-1}\sum_{n'=0}^{N-1} T(k,m,)T(l,n)T^*(k,m')T^*(l,n')r(m,n,m',n') \quad (8)$$

where $r(m,n,m',n')$ is the covariance between $u(m,n)$ and $u(m',n')$ in the input image. No assumptions about separability of the covariance function are made. Now assume that the boundary of the non-zero region of the image is circular, and that it is inscribed in the square formed by the edges of $u(m,n)$. The covariance is assumed to be isotropic within D and zero when array elements external to D are involved. Letting d equal the distance between (m,n) and (m',n'), ie:

$$d = \sqrt{(m-m')^2 + (n-n')^2}, \quad (9)$$

the covariance function is

$$r(m,n,m',n') = \begin{cases} 0 & (m,n) \text{ or } (m',n') \text{ outside } D \\ \rho^d & \text{other } m,n,m',n' \end{cases} \quad (10)$$

Although no practical simplification of (8) exists for this covariance function, $\sigma_k^2(k,l)$ can easily be evaluated numerically. For example, for an image of size 16 by 16 with $\rho = .85$ and the covariance function of (10), the 13 lowest-frequency DCT coefficients contain (on average) 68.68% of the energy, while 13 real coefficients from the DFT contain 70.62%. If 22 of the 256 coefficients are saved the percentages rise to 75.93% and 76.66% respectively. As with the 1D case, the DFT offers superior energy compaction.

4. Discussion

The difference in the relative performance of the DCT and the DFT for the stationary versus nonstationary sequences can be understood by comparing the implicit periodicities of the DCT and DFT. The DFT enforces a periodicity of N on the input data. The resulting placement of the $n = 0$ element next to the $n = N-1$ element can introduce a discontinuity into the data which tends to spread the energy in the transform domain, thereby decreasing energy compaction efficiency. By contrast, the DCT can be obtained by a DFT of a length-$2N$ DFT sequence in which the values for $N \leq n \leq 2N-1$ are in effect a mirror image of the first N values. Periodic repetition of the length-$2N$ sequence is therefore free from artificial discontinuities.

Both the DFT and DCT compression were implemented on real tomographic images. In all cases, the energy retention and mean square error performance were superior with the DFT. Care was taken to ensure that the compression ratio for the DFT and DCT were equal after accounting for the complex nature of the DFT. For example, for 5:1 compression of a 128 × 128 PET image with transform $v(k,l)$, only the 3275 DCT elements $v(k,l)$ satisfying $\sqrt{k^2+l^2} < N/2$ were used. For the DFT, exactly the same number of real DFT coefficients were used (where complex coefficients furnish two real coefficients) in the reconstruction, and the hermitian redundancy of the DFT was exploited. An example of energy performance on a single PET image are as follows: For the DCT compression, the 3275 coefficients contained 99.889% of the image energy and gave a normalized mean square error (NMSE) of 0.1963%. For the DFT the 3275 coefficients contained 99.891% of the image energy, and gave an NMSE of 0.1926%. As expected from a transform coding scheme in which the highest-frequency transform components were discarded, the reconstruction error for both the DFT and DCT is high-frequency in nature. Trials using other images yielded similar results. It should be stressed that a more sophisticated bit allocation scheme would be used in an operational compression system (with an accompanying improvement in the compression efficiency for both the DCT and DFT), but that the fundamental superiority of the DFT would still hold.

5. Conclusion

The *a priori* knowledge regarding specific nonstationarities in tomographic images can be used in a non-stationary statistical image model. When this model is applied, the DFT and DHT replace the DCT as the optimum fast transform for full-frame image compression. As the percentages above show, in practice the use of the DFT instead of the DCT only results in an extremely small improvement in NMSE and energy compaction. What is significant, however, is that the DFT does offer better compression performance. Furthermore, the DFT and DHT are easier to implement than the DCT, involving (depending on the algorithm) simpler index mappings or fewer operations. Even if the energy compaction of the two transforms were equivalent, the DFT would still be a better choice than the DCT. Compression using the DHT would give a result identical to DFT compression. Such an approach would be more convenient because the purely real nature of the HT would simplify data handling in the transform domain, and the need to grapple with "real" FFT codes would be circumvented.

References

[1] G.K. Wallace, "The JPEG Still Picture Compression Standard," *Comm. ACM*, vol. 34, pp. 30-45, 1991.

[2] H.K Huang, S.C. Lo, B.K.T. Ho, and S.L. Lou, "Radiological image compression using error-free irreversible two-dimensional direct-cosine-transform coding techniques," *J. Opt. Soc. Am. A*, vol. 4, pp. 984-992, 1987.

[3] K.K. Chan, S.C. Lo and H.K. Huang, "Radiological image compression using full-frame cosine transform with adaptive bit allocation," *J. Comp. Med. Imaging Graphics*, vol. 13, pp. 153-159, 1989.

[4] B.K.T. Ho, J. Chao, C.S. Wu, and H.K. Huang, "Full-frame Cosine Transform Image Compression for Medical and Industrial Applications," *Machine Vis. Appl.*, vol. 3., pp. 89-96, 1991.

[5] J. Sayre, B.K.T. Ho, M.I. Boechat, T.R. Hall, and H.K. Huang, "The effect of full-frame image compression on diagnostic accuracy in hand radiographs with subperiosteal resorption," *Radiology*, Oct. 1992.

[6] H.K. Huang, *et al.* "Picture archiving and communication systems (PACS) for rdiological images: State of the art. CRC Critical Reviews in Medical Informatics," vol. 28, pp. 383-427,

[7] A.K. Jain, "Image Data Compression: A Review," *Proc. IEEE*, vol. 68, pp. 349-389, 1981

[8] D.E. Pearson and J.A. Robinson, "Visual communication at very low data rates," *Proc. IEEE*, vol. 69, pp. 542-551, 1985.

[9] D. Le Gall, "MPEG: A video compression standard for multimedia applications," *Comm. ACM*, vol. 34, pp. 46-58, 1991.

[10] V.R. Algazi and D.J. Sakrison, "On the optimality of the Karhunen-Loeve expansion," *IEEE Trans. Information Theory* pp. 319-321, 1969.

[11] A. K. Jain, *Fundamentals of Digital Image Processing*, Englewood Cliffs: Prentice-Hall, 1989, p. 140.

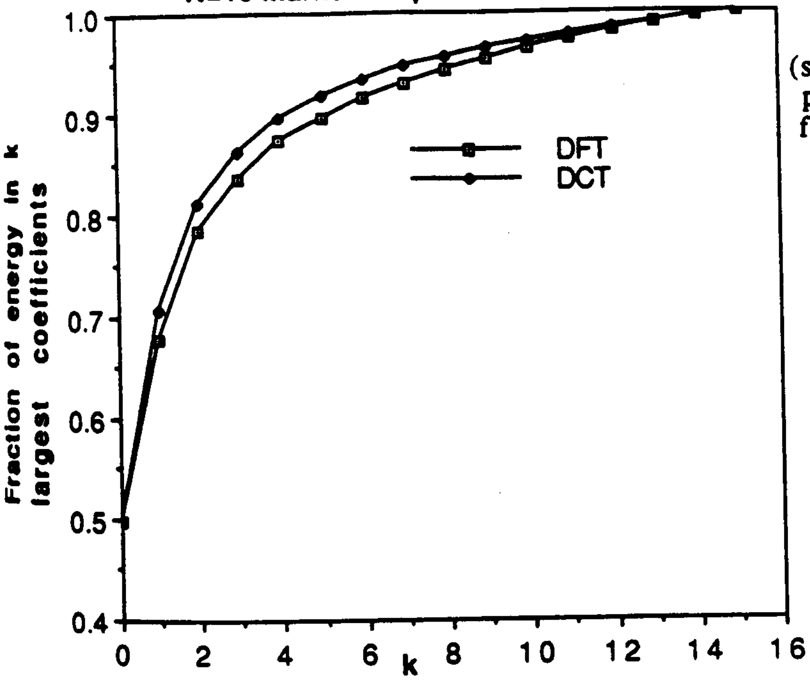
[12] E.A. Riskin, T. Lookabaugh, P.A. Chou, and R.M. Gray, "Variable rate vector quantization for medical image compression," *IEEE Trans. Med. Imaging*, vol. 9, pp. 290-298, 1990.

[13] R.N. Bracewell, "The fast Hartley transform," *Proc. IEEE*, vol. 72, pp. 1010-1018, 1984.

[14] R.N. Bracwell, *The Hartley transform*. New York, NY: Oxford Univ. Press, 1986.

Figure 1(following page): Energy compaction comparison for the DFT and DCT. An N=16 one-dimensional Markov sequence with correlation ρ=0.85 is used. In Figure 1a, the sequence is stationary and the DCT has superior energy compaction. Figure 1b shows the result if the two boundary elements (n=0 and n=15) are zero. In this non-stationary model, the DFT becomes superior.

Figure 1a: DCT/DFT Energy compaction
N=16 Markov sequence



(see previous page for figure caption)

Figure 1b: DCT/DFT Energy Compaction
N=16 Markov sequence, border pixels = 0