Report Number: WUCS-86-12

1986-06-01

Progressive Coding and Transmission of Digital Diagnostic Pictures

Sharaf E. Elnahas, Kou-Hu Tsou, Jerome R. Cox, Rexford L. Hill, and R. Gilbert Jost

In radiology, as a result if the increased utilization of digital imaging modalities, such as computed tomography (CT) and magnetic resonance imaging (MRI), over a third of the images produced in a typical radiology department are currently in digital form, and this percentage is steadily increasing, Image compression provides a means for the economical storage and efficient transmission of these diagnostic pictures. The level of coding distortion than can be accepted for clinical diagnosis purposes is not yet well-defined. In this paper we introduce some constraints on the design of existing transform codes in order to achieve progressive image...
Read complete abstract on page 2.

Progressive Coding and Transmission of Digital Diagnostic Pictures

Sharaf E. Elnahas, Kou-Hu Tsou, Jerome R. Cox, Rexford L. Hill, and R. Gilbert Jost

Complete Abstract:

In radiology, as a result if the increased utilization of digital imaging modalities, such as computed tomography (CT) and magnetic resonance imaging (MRI), over a third of the images produced in a typical radiology department are currently in digital form, and this percentage is steadily increasing, Image compression provides a means for the economical storage and efficient transmission of these diagnostic pictures. The level of coding distortion than can be accepted for clinical diagnosis purposes is not yet well-defined. In this paper we introduce some constraints on the design of existing transform codes in order to achieve progressive image transmission efficiently. The design constraints allow the image quality to be asymptotically improved such that the proper clinical diagnoses are always possible. The modified transform code outperforms simple spatial-domain codes by providing higher quality of the intermediately reconstructed images. The improvement is 10 dB for a compression factor of 256:1. and it is as high as 17.5 dB for a factor of 8:1. A novel progressive quantization scheme is developed for optimal progressive transmission of transformed diagnostic images. Combined with a discrete cosine transform, the new approach delivers intermediately reconstructed images of comparable quality twice as fast as the more usual zig-zag sampled approach. The quantization procedures is suitable for hardware implementation.

PROGRESSIVE CODING AND TRANSMISSION OF DIGITAL DIAGNOSTIC PICTURES

Sharaf E. Elnahas, Kou-Hu Tzou, Jerome R. Cox, Jr., Rexford L. Hill and R. Gilbert Jost

WUCS-86-12

June 1986

Department of Computer Science
Washington University
Campus Box 1045
One Brookings Drive
Saint Louis, MO 63130-4899

This work was presented in part at SPIE's 29th Annual International Technical Symposium, San Diego, California, August 1985.

Progressive Coding and Transmission of Digital Diagnostic Pictures

SHARAF E. ELNAHAS, KOU-HU TZOU, SENIOR MEMBER, IEEE, JEROME R. COX, JR., FELLOW, IEEE, REXFORD L. HILL, AND R. GILBERT JOST

Abstract—**In radiology, as a result of the increased utilization of digital imaging modalities, such as computed tomography (CT) and magnetic resonance imaging (MRI), over a third of the images produced in a typical radiology department are currently in digital form, and this percentage is steadily increasing. Image compression provides a means for the economical storage and efficient transmission of these diagnostic pictures. The level of coding distortion that can be accepted for clinical diagnosis purposes is not yet well-defined.**

In this paper we introduce some constraints on the design of existing transform codes in order to achieve progressive image transmission efficiently. The design constraints allow the image quality to be asymptotically improved such that the proper clinical diagnoses are always possible. The modified transform code outperforms simple spatial-domain codes by providing higher quality of the intermediately reconstructed images. The improvement is 10 dB for a compression factor of 256:1, and it is as high as 17.5 dB for a factor of 8:1.

A novel progressive quantization scheme is developed for optimal progressive transmission of transformed diagnostic images. Combined with a discrete cosine transform, the new approach delivers intermediately reconstructed images of comparable quality twice as fast as the more usual zig-zag sampled approach. The quantization procedure is suitable for hardware implementation.

I. INTRODUCTION

DIGITAL diagnostic pictures are now used for an increasing number of imaging modalities including CT, MRI, nuclear medicine, ultrasound, and digital vascular imaging. In the near future it is likely that many standard radiographic images will also be available in digital form. Considerable interest in picture archiving and communication systems (PACS) was a result of the ever increasing number of digital diagnostic images. The quantities of image data are large [1]–[3]. Data compression techniques can be used to minimize, where possible, the burdens of storage and transmission. However, the data compression requirements for PACS are varied, and different solutions are likely to be required for different aspects of the problem. In general, algorithms for archival purposes, which require a minimum of time for decompression, are favored. In most situations it is not important how long it takes to compress an image for storage in the archive, but clinical image retrieval demands will often require prompt decompression. The new developments in optical and magnetic storage technology [4], [5] are anticipated to help alleviate digital storage problems. These developments will reduce the storage cost per bit significantly. However, it is likely that, for the foreseeable future, significant cost savings will be realized by reducing the storage requirements through data compression. Data compression techniques can play an important role in digital image transmission, reducing the time required to move an image from one location to another, but the algorithms suitable for data transmission may be entirely different from those algorithms that are designed for archival purposes. Dunham *et al.* [6] reviewed the major classes of data compression with respect to their suitability for digital images in radiology.

In a PACS environment, noninvertible compression provides high compression factors (20:1, 30:1); but the distortion introduced by noninvertible compression has yet to be evaluated clinically for different classes of radiology images and diagnostic modalities. Noninvertible algorithms might seem unsuited to situations in which fine image detail is required for diagnosis and to situations in which the images become the object of later automatic computer processing. On the other hand, invertible compression provides perfect image reconstruction with low compression factors (2:1, 4:1). One may combine invertible and noninvertible procedures into one system, wherein the advantages of both high compression factors and perfect image reconstruction can be realized as follows. For archiving purposes, two types of data are generated. The first type is noninvertible compressed data. The second type is the mathematical difference between the imperfectly reconstructed image, from the first type of data, and the original image data before compression was applied. Elnahas [7] has shown that the sum of these two types of data results in compression factors ranging from 2:1 to 4:1 for CT and MRI pictures. For transmission purposes, the first type of data can be used to achieve high compression factors provided that the introduced distortion is acceptable. Should the distortion be unacceptable, the second type of data is transmitted such that perfect image reconstruction is obtained at the price of low compression factors. Therefore, efficient techniques for both invertible and noninvertible compression are needed

Manuscript received October 21, 1985; revised February 19, 1986. This work was presented in part at SPIE's 29th Annual International Technical Symposium, San Diego, CA, August 1985.

S. E. Elnahas and K.-H. Tzou are with GTE Laboratories Inc., Waltham, MA 02254.

J. R. Cox, Jr. is with the Department of Computer Science, Washington University, St. Louis, MO 63130.

R. L. Hill and R. G. Jost are with the Mallinckrodt Institute of Radiology, St. Louis, MO 63110.

IEEE Log Number 8608299.

for PACS. This is particularly desirable when the picture network is large and includes low-bandwidth channels for long-distance switched communication. For local environments, where the workstations are very close to the picture archive and direct communication using large-bandwidth channels is affordable, noninvertible compression might not be necessary.

Progressive transmission of digital pictures permits the initial reconstruction of an approximate picture followed by a gradual improvement in the quality of image reconstruction [8]–[13]. The technique is useful for the transmission of pictures over low-bandwidth channels such as telephone lines. Teleconferencing and telebrowsing are good examples of prospective applications. The concept of progressive transmission is of particular importance in an electronic radiology environment. A radiologist browsing through many remotely stored pictures of a patient may need to quickly abort transmission of one or more unwanted pictures as soon as they are recognized at the local image display unit. Here we assume that the radiologist had already browsed on a distributed file system and selected a patient who has several diagnostic studies and several images per study. We further assume that the local workstation is not close to the picture archive and that direct picture transmission, over a large-bandwidth channel, is not economical. Once the desired picture is identified during early steps of the progression, more information can be transmitted, thereby improving the subjective quality of the received picture until a clinical diagnosis is possible.

In order to evaluate the performance of progressive transmission techniques, we assume that the image has already been digitized into an $N \times N$ array of picture elements with L bits/pixel. The performance criterion is as follows.

> For X bits transmitted, such that $X < N^2L$, the best progressive transmission system will give the best quality of intermediate image approximations within some complexity limit.

It is worth noting that, by the above criterion, we evaluate the performance of progressive image transmission from the standpoint of data compression. The goal is to use the transmitted information in the most efficient way independent of bit rate. Knowlton [9] has proposed a simple encoding scheme, spatial-domain encoding, for progressive transmission of gray-scale pictures. His approach has the advantage of simple implementation with no coding distortion in the final reconstructed image. Due to the nature of successive picture division introduced in Knowlton's method, the accumulated number of transmitted bits to achieve the ith intermediate image reconstruction is proportional to 2^i.

In transform coding of digital pictures [14], the input image is divided into subblocks, $u_{n,m}$, $n, m = 1, 2, \cdots, N/M$, where a subblock is an array of $M \times M$ picture elements. Each subblock $u_{n,m}$ is transformed into $U_{n,m}$ by the linear transformation T, as shown in Fig. 1. In the

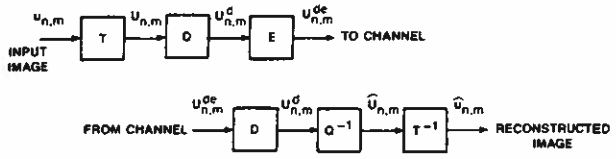


Fig. 1. Transform coding.

transform domain, $U_{n,m}$ is quantized into $U^d_{n,m}$ by the quantizer Q. Then, the quantized subblock $U^d_{n,m}$ is encoded into $U^{de}_{n,m}$, by the entropy encoder E, and transmitted over the channel. At the receiver end, $U^{de}_{n,m}$ is entropy decoded into $U^d_{n,m}$, dequantized into $\hat{U}_{n,m}$, and inverse-transformed into the reconstruction $\hat{u}_{n,m}$. Transform coding is suitable for progressive transmission of pictures in the sense that an initial subset of transform domain samples yields an approximate picture, while the latter ones add detail. Takikawa [11] focused on a fast algorithm for progressive reconstruction of transformed images rather than on the compression efficiency. In his method the $M \times M$ transform coefficients are decomposed into $(\log_2 M + 1)$ complementary matrices; each of them can be inverse transformed by $1 \times 1, 2 \times 2, 4 \times 4, \cdots, M \times M$ fast transformations.

Takikawa's approach provides an efficient computation of progressive reconstruction and at the same time reduces the accumulation of roundoff errors due to the progression. However, like Knowlton's approach, the accumulated number of transmitted bits grows exponentially with the order of an intermediate step in the progression. In this paper we propose more efficient schemes for progressive transmission based on discrete cosine transform coding of digital pictures. In Section II we formalize the concept of progressive transmission of transformed images and introduce some design constraints that make the system desirable from the standpoint of electronic radiology. Using the design constraints, we modify an existing transform code and compare its performance to that of spatial-domain coding. In order to achieve optimal progressive transmission of transformed images, progressive quantization schemes are needed. A novel quantization technique is developed in Section III. Application of the concept of progressive transmission to electronic radiology is introduced in Section IV, and simulation results for panels of digital diagnostic images are presented. The paper describes software simulation results and the fundamental algorithms that may be used to design a working system. Hardware based on the presented methods has not been built yet. Conclusions are drawn in Section V.

II. Progressive Transmission

For the purpose of progressive transmission, we define $S_i[U^{de}_{n,m}]$ to be the size, in bits, of the portion of $U^{de}_{n,m}$ transmitted to produce the ith intermediate reconstruction. We note that

$$S_i[U^{de}_{n,m}] = S_{i-1}[U^{de}_{n,m}] + d_{n,m,i} \tag{1}$$

where $d_{n,m,i}$ is the number of bits transmitted during the ith step in the progression, $i = 1, 2, 3, \cdots$, and

$S_0[U^{de}_{n,m}] = 0$. At the decoding end, the recently received $d_{n,m,i}$ bits are combined with the previously received $S_{i-1}[U^{de}_{n,m}]$ bits to achieve the ith intermediate reconstruction of $U_{n,m}$. Now, with X_i denoting the accumulated number of transmitted bits, to achieve the ith intermediate reconstruction of the entire image, we have

$$X_i = \sum_{n=1}^{N/M} \sum_{m=1}^{N/M} S_i[U^{de}_{n,m}] \leq N^2 L \tag{2}$$

where equality occurs when all of the transform-domain samples are quantized at L bits/sample and transmitted over the channel.

The question now is how to choose the subsets of transform-domain samples and $d_{n,m,i}$ for all n, m, and i. Two points are important here. First, the order in which the transform-domain samples are transmitted should be known by both the encoder and decoder such that the decoder can properly accumulate the progressively transmitted data to obtain the intermediate picture approximations. Second, some quality measure for picture reconstruction should be directly or indirectly involved in choosing the progressive subsets of transform-domain samples. Let us consider possible approaches for using transform codes to achieve progressive image transmission. For example, one might ideally transmit the transform samples $U_{n,m}(i, j)$ in the order of their magnitudes for all $1 \leq n, m \leq N/M$ and $1 \leq i, j \leq M$. However, this order is not known by the decoder. Transmitting the magnitude order would require a very large amount of side information. Ngan [12] has suggested another approach in which the transform samples are transmitted in the order of a zig-zag sampled pattern. This pattern determines the order of transmitting elements from a given subblock. The order of subblocks is not well defined for achieving reasonable gradual improvements in the quality of progressively reconstructed images. The bit-assignment matrices of the adaptive system of Chen and Smith [15] can provide a practical solution of this problem. The remainder of this section is devoted to the details of this solution when it is applied to digital diagnostic images.

A. Procedure

First, we introduce some design constraints that will make the solution desirable from the standpoint of clinical diagnosis. In radiology, the level of coding distortion that can be accepted for diagnosis purposes is not yet well defined. What a radiologist actually needs is an image compression system from which he can control the quality of reconstructed images by reducing the level of distortion so that the proper clinical diagnoses can be made. Our design constraints are given by

$$\sum_{n=1}^{N/M} \sum_{m=1}^{N/M} d_{n,m,i} = C \tag{3}$$

where C is a constant for all i. That is, we assume a fixed-length transmission. Furthermore, we choose $C = NL$, which leads to $X_N = N^2 L$. Therefore, N transform elements, quantized at L bits/element, are transmitted at each intermediate step. The final step is the Nth step, and the accumulated number of transmitted bits X_N to achieve the final image reconstruction is $N^2 L$. We emphasize that we progressively transmit the transform samples quantized at the full resolution of L bits/sample. The bit-assignment matrices of Chen and Smith [15], where more bits are assigned to subblocks of the image high-activity regions and fewer bits are assigned to subblocks of the image low-activity regions, are used only to tell in what order we transmit the transform samples.

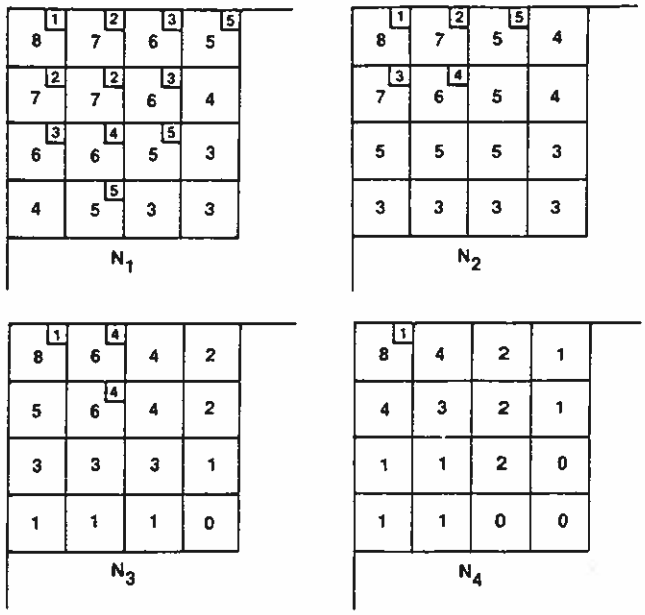


Fig. 2. Bit-assignment matrices. (Numbers at the upper-right corners of some locations indicate the order of transmission for Steps 1–5 in the progression.)

The mechanism of using the bit-assignment matrices for determining the order of transmission is best illustrated by an example. Consider the 4×4 low-frequency portions of the $M \times M$ bit-assignment matrices (see Fig. 2). Here, N_1 corresponds to the highest activity level, and N_4 corresponds to the lowest activity level. The full resolution L is assumed to be 8 bits/pixel. In the first step of the progression, all of the dc samples at location (1, 1) are transmitted. In the second step, all elements at locations (1, 2), (2, 1), and (2, 2) are transmitted from the highest activity level, and all elements at location (1, 2) are transmitted from the second activity level. In the third step, all elements at location (2, 1) are transmitted from the second level, and all elements at locations (1, 3), (2, 3), and (3, 1) are transmitted from the highest activity level. In the fourth step, all elements at location (3, 2) from the highest activity level, all elements at location (2, 2) from the second level, and all elements at locations (1, 2) and (2, 2) from the third level are transmitted. In the fifth step, all elements at locations (1, 4), (3, 3), and (4, 2) from the highest activity level and all elements at location (1, 3) from the second level are transmitted. The process is continued in the obvious manner for the remainder of the steps in the progression.

Therefore, for $N = 256$ and $M = 16$, the four bit-assignment matrices can be used to control the progressive transmission as follows. For each intermediate step, a combination of four entries is chosen from the bit-assignment matrices. Each entry will determine the transmission of 64 transform-domain samples from one of the four classification groups of transform subblocks. Entries from the bit-assignment matrices are chosen as the largest entries at the ith step for all i. A tie-breaking rule is needed in the case of equality of two (or more) largest entries from two (or more) different matrices. As an example, consider the different possibilities of choosing a combination of four entries for the third step from the matrices given in Fig. 2. The entry at location (2, 1) of N_2 has the largest value of "7," and, therefore, it is one of the four needed entries. For the remaining three entries, we have different possibilities since there are seven largest entries at this point—namely, entries at locations (1,3), (2,3), (3,1), and (3, 2) from N_1; an entry at location (2, 2) from N_2; and entries at locations (1, 2) and (2, 2) from N_3. All of these entries have the largest value of "6." Our tie-breaking rule was to choose entries at locations (1, 3), (2, 3), and (3, 1) from N_1 as the remaining three entries for the third step in the progression. In other words, we gave the priority of transmission to elements from high-activity regions. By this tie-breaking rule, the quality of reconstruction of the high-activity regions will be improved in the early steps of the progression. Details of the low-activity regions will be added in later steps.

B. Simulation

Fig. 3 demonstrates the simulation results of the above progressive transmission scheme when applied to the chest CT image of Fig. 4(d). The first 48 picture approximations are shown in Fig. 3 with the first reconstruction at the upper-left corner. The quality is gradually improved from top to bottom in the first column of picture approximations. The progression is continued at the bottom picture of the second column and is improved from bottom to top in this case. The cycle is repeated every two columns of picture approximations in a cosine-like form. The relative quality of intermediate image approximations can be compared, and recognized as superior, to that of Knowlton [9], Burt and Adelson [10], or Takikawa [11]. The compared image reconstructions should be at the same compression factor. For example, the relative quality of Knowlton's ninth approximation (compression factor of 64:1) should be compared to that of the bottom approximation of the first column in Fig. 3, Knowlton's tenth approximation (32:1) should be compared to the top picture approximation of the second column, and so on.

C. Performance

Quantization of the transform coefficients prevents the perfect reconstruction of the transformed images. However, since the transform-domain samples are quantized at L bits/sample, the same intensity resolution as that of the input image, we can expect the sequence of intermediate image approximations to converge into an almost perfect reconstruction of the input image. In other words, we claim that the subjective quality of the high-order image approximations will be the same as that of the input image, as shown in Fig. 4. Quantitatively, Fig. 5 shows how the signal-to-noise ratio (SNR) converges reasonably into a high value of 56.5 dB as a function of the order of image approximation. The SNR for the ith intermediate reconstruction is defined as follows:

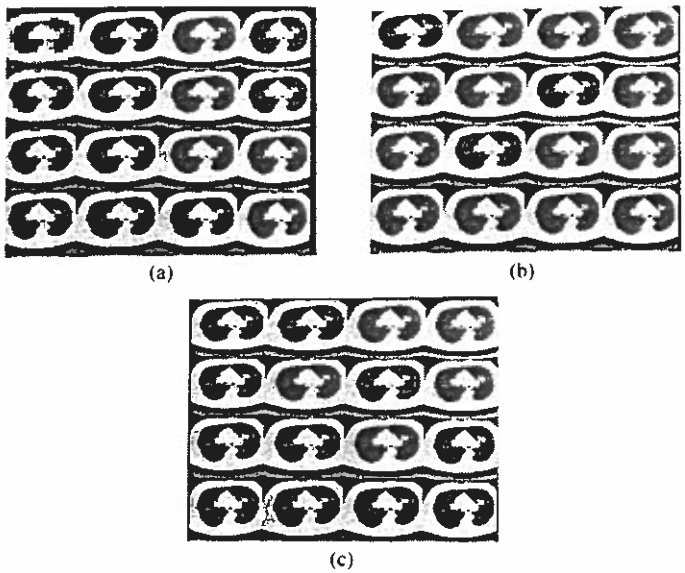


Fig. 3. Low-order approximations: (a) 1st through 16th, (b) 17th through 32nd, and (c) 33rd through 48th.

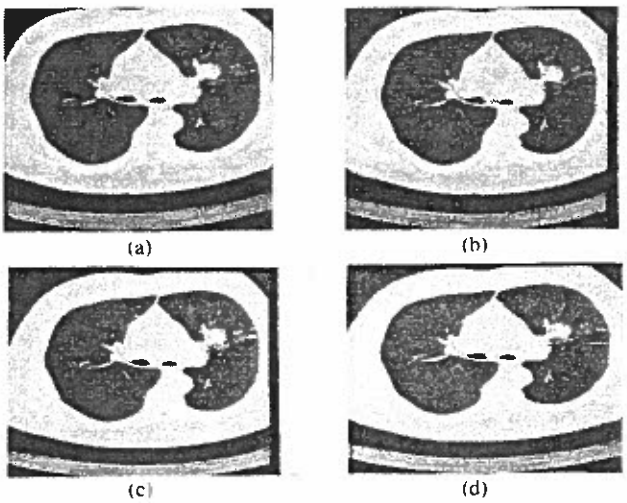


Fig. 4. High-order approximations: (a) 50th, (b) 125th, (c) 225th, and (d) original.

$$\text{SNR} = 10 \log_{10} \frac{(\text{peak-to-peak value})^2}{\text{mean square error}} = 10 \log_{10} \frac{(2^L)^2}{1/N^2 \sum_{x=1}^{N} \sum_{y=1}^{N} [u(x, y) - \hat{u}_i(x, y)]^2} \quad (4)$$

where $u(x, y)$ are elements of the input image, and $\hat{u}_i(x, y)$ are elements of the ith intermediate reconstruction. In

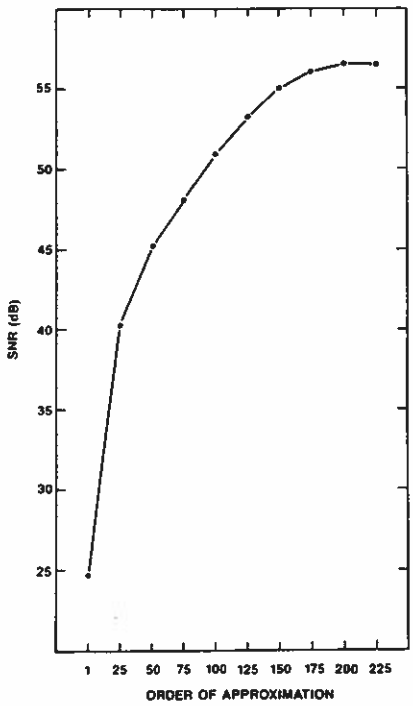


Fig. 5. Objective quality of reconstructed images.

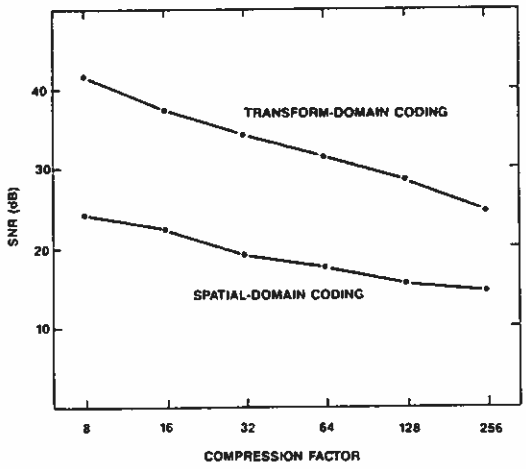


Fig. 6. Knowlton's approximations versus transform approximations.

Fig. 6 we compare the SNR values of our intermediate image approximations to those of Knowlton's approximations [9] that are obtained after the gray-scale values of all picture elements have been formatted into a hierarchical structure of picture subdivisions of successive sizes from entire image down to basic element values. An improvement of about 10 dB is achieved for a compression factor of 256:1. The improvement is as high as 17.5 dB for a compression factor of 8:1. We point out that no quality measure has yet been directly involved in choosing the progressive subsets of transform-domain samples. In the following section we include the mean squared error as a quality measure and derive an optimal progressive quantization scheme. The ability to optimize the different modules of the system is actually why transform coding techniques have promise for achieving more efficient systems for the progressive transmission of pictures.

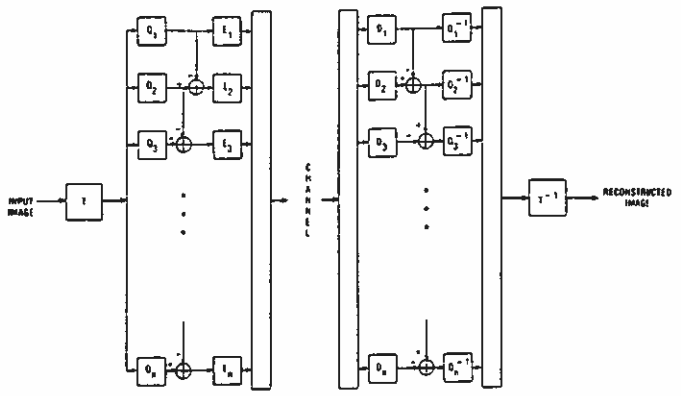


Fig. 7. Transform progressive coding: "T" is a 2-D linear transformation such as the discrete cosine transform. For $i = 1, 2, 3, \cdots$, "Q_i" is a 2-D quantization scheme for optimal 2-D bit-assignment matrices that are designed for a coding rate of R_i bits/pixel such that $R_{i+1} > R_i$; "E_i" is an entropy encoder; "D_i" is an entropy decoder; and "Q_i^{-1}" is a 2-D inverse-quantization scheme. "T^{-1}" is a 2-D inverse transformation.

III. Progressive Coding

The general transform coding procedure of Fig. 1 can be extended to the case of transform progressive coding as shown in Fig. 7. Bit-assignment matrices are designed for coding rates of R_1, R_2, $\cdots$, and R_n bits/pixel, and the transform samples are quantized by Q_1, Q_2, $\cdots$, and Q_n, respectively. The coding rates are chosen such that $R_{i+1} > R_i$ for any intermediate step i. In the first step of the progression, transform coefficients quantized at the lowest coding rate, R_1 bits/pixel, are entropy encoded by E_1 and transmitted over the channel. In the second step, the R_1-quantized samples are subtracted from the R_2-quantized samples. The difference is entropy encoded by E_2 and transmitted over the channel. At the receiver end, the recently received difference is entropy decoded by D_2 and added to the previously decoded data. The result is then dequantized by Q_2^{-1} and inverse transformed to obtain a reconstructed image. The process is repeated for higher code rates. If the quantizers are optimized in the sense of minimum mean squared error (MMSE), the intermediate reconstructions will be optimized in the same sense. In Fig. 8 we compare the subjective quality of intermediate image reconstructions to that of the previous section. The photo shown in Fig. 8(a) is from the scheme of the previous section, while that of Fig. 8(b) is from the optimal progressive coding scheme, both at a compression factor of 16:1. Now, even though the intermediate quantization and reconstruction are optimized in the MMSE sense, the intermediate transmission of data is not well-optimized. As an example, consider the second step in the progression with $R_1 = 1$ bit/pixel and $R_2 = 2$ bits/pixel. The simple difference between the R_2-quantized samples and the R_1-quantized samples will, in general, require more than 1 bit/pixel for transmission. In order to achieve optimal progressive transmission, the quantization scheme must provide an embedded code, that is, the transmitted bits corresponding to a lower bit rate must be contained within the code of a higher bit rate. We have examined both the uniform and nonuniform Max [16]

(a) (b)

Fig. 8. Subjective quality at a compression factor of 16 : 1 from: (a) previous algorithm: and (b) optimal algorithm.

quantizers and found that neither of them has the embedding property. The remainder of this section is devoted to the design of an optimal embedded code for the optimal progressive transmission of digital pictures.

A. Embedded Code

For the convenience of this discussion, we describe the embedded code for the case of nonadaptive transform coding with a zonal bit-assignment scheme where the number of bits allocated to each transform subblock is the same. Certainly, the proposed technique can be easily extended to adaptive transform coding. We further assume the segments of progressive data to be of fixed length $d_{n.m}$ for all the intermediate steps. The incremental bit rate for each step is $d_{n.m}/M^2$ bits/intermediate pixel. Let $b_i(k, l)$ and $b_{i+1}(k, l)$ be the number of bits allocated to the transform coefficient at location (k, l) for the intermediate steps of order i and $i + 1$, respectively. The number of incremental bits $\Delta b_{i+1}(k, l)$ to be transmitted at the intermediate step of order $i + 1$ is given by $\Delta b_{i+1}(k, l) = b_{i+1}(k, l) - b_i(k, l)$. The bit-assignment rule based upon the rate-distortion theory [17] guarantees $\Delta b_{i+1}(k, l)$ to be always nonnegative. The maximum number of bits allocated to each coefficient is set to L in order to match the resolution of the input image. An example of incremental bit-assignment maps designed for a block size of 4×4 with $L = 8$ bits/pixel and incremental rate of 1 bit/intermediate pixel is shown in Fig. 9. The sequence of bits allocated to each coefficient can be read from the map at its corresponding location. For example, the bit-assignment sequence for the dc term, location (1, 1), is "41111000" while the sequence for the coefficient at location (1, 2) is "31111100." The proposed scheme can be viewed as slicing the full bit-assignment map, 8 bits at all locations, into layers of incremental bit-assignment maps and transmitting a slice of bits at each intermediate step. When the incremental bit rate is reduced, the total number of incremental bit-assignment maps is increased. The overhead information corresponding to these maps may be reasonably large if the incremental bit rate is very small. Nevertheless, there is no need to send the incremental maps as side information. Both the transmitter and receiver can design the same maps based upon the standard deviations of transform coefficients that can be transmitted beforehand, a modest amount of side information.

The bit-sliced progressive quantization scheme requires the quantizers to be able to make progressively finer reconstructions based upon the already received information and the additional information received at each intermediate step. In the example shown in Fig. 9, 4 bits are allocated to the dc component in the first step and an additional bit in each of the next four steps. Therefore, the dc term is initially quantized by a 4-bit quantizer; the quantization is refined into 5 bits in the second step, into 6 bits in the third step, and so on. Let y_i and y_{i+1} be the binary representations of b_i- and b_{i+1}-bit quantizer outputs, corresponding to an input x, respectively. Then, y_{i+1} should be able to be represented as $(y_i, \Delta y_{i+1})$ so that the previous output code is embedded within the current output code. Consequently, a finer reconstructed value x_{i+1} can be obtained from the already received y_i and the additional information Δy_{i+1} in the intermediate step of order $i + 1$. Neither the uniform nor the nonuniform Max quantizer has the embedding property. Fig. 10 shows the thresholds of 1- to 4-bit nonuniform Max quantizers, where only the positive portions are shown since a symmetric probability density function has been assumed. The quantization procedure can be clearly explained by the binary tree representation of thresholds for a logarithmic search as shown in Fig. 11. For each quantizer, the procedure starts at the root of the tree and advances one level down the tree for each additional bit assigned. If the quantizer input is greater than the threshold, a "1" is

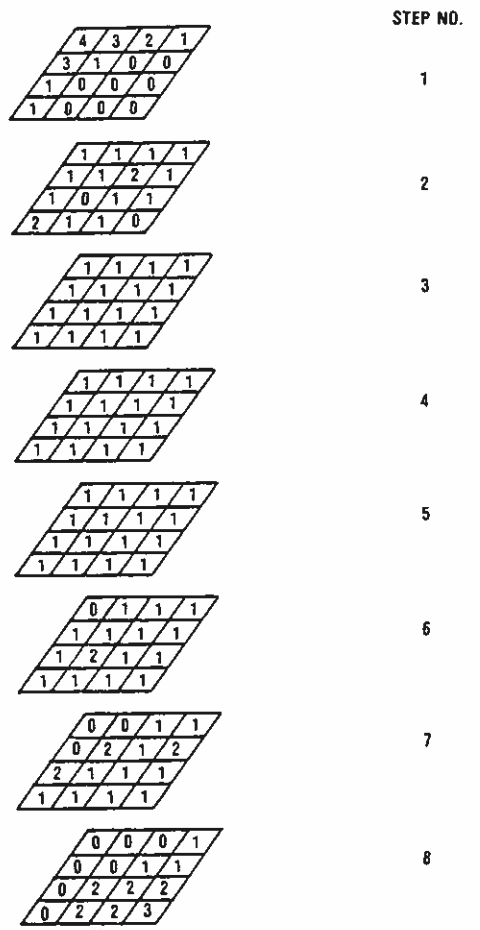


Fig. 9. Incremental bit-assignment maps for a block size of 4×4 and incremental bit rate of 1 bit/pixel.

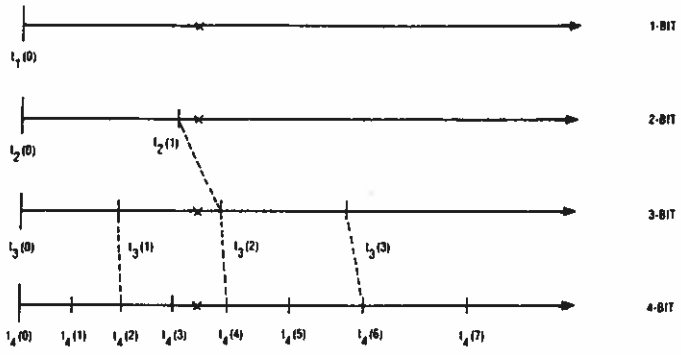


Fig. 10. Thresholds of 1- to 4-bit nonuniform Max quantizers.

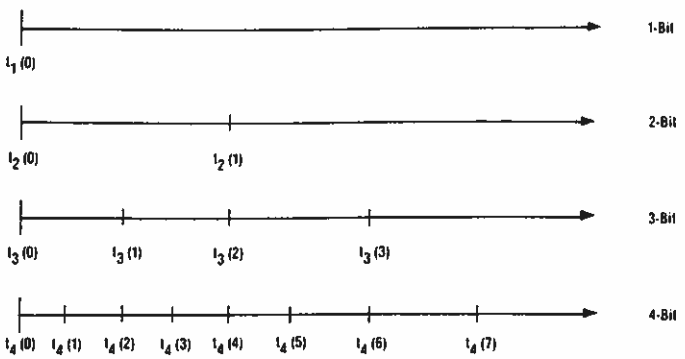


Fig. 12. Threshold-aligned quantizers. (Only the positive portions are shown.)

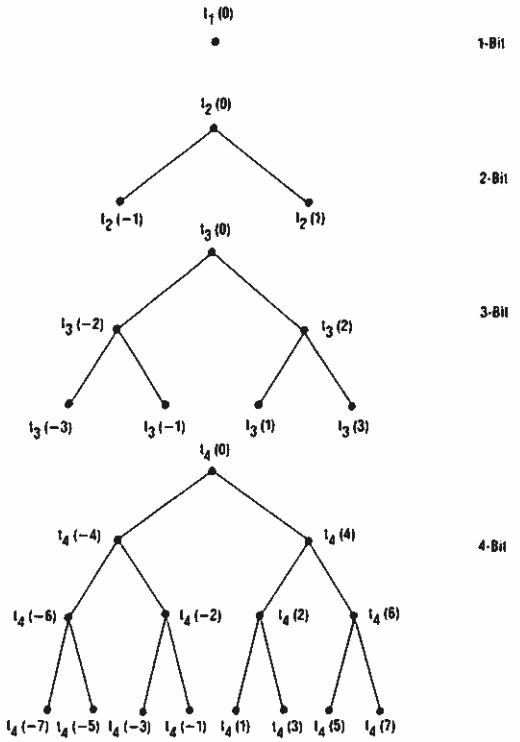


Fig. 11. Binary tree representations of the quantizer thresholds corresponding to Fig. 10.

transmitted, and the right branch of the node is followed to reach the next threshold to be compared. If the input is smaller than the threshold, opposite actions are taken. Now assume that the input marked as an "x" in Fig. 10 is to be quantized sequentially by the quantizers. Referring to Figs. 10 and 11, we can find that the output codes for the 1-, 2-, 3-, and 4-bit quantizers are "1," "11," "101," and "1011," respectively. Therefore, the nonuniform Max quantizer is not embedded since "11" is not included in "101." Similarly, we can show that the uniform Max quantizer is also not embedded.

B. Threshold-Aligned Quantizers

However, we can modify the quantizers by aligning their thresholds, linked with dashes in Fig. 10, to the same levels to achieve the embedding property. For example, we can fix the thresholds for the 3-bit quantizer and align $t_2(1)$ and $t_4(4)$ with $t_3(2)$, $t_4(2)$ with $t_3(1)$, and $t_4(6)$ with $t_3(3)$. Fig. 12 shows the threshold-aligned quantizers. The effect of aligning the thresholds in Fig. 10 is equivalent to that of overlaying the binary trees in Fig. 11. There are two major issues that we have to deal with while designing the threshold-aligned quantizers. First, we have to choose a quantizer as a reference. Examining the optimal quantizers designed for Gaussian and Laplacian sources [16], [18], [19], we find that the magnitudes of thresholds to be aligned are increasing with the number of output levels of the quantizers. If we choose the 2-bit quantizer as the reference, the threshold-aligned 8-bit quantizer will suffer significant degradation. The reverse is also true. A compromise solution is to choose the 4- or 5-bit quantizer as the reference so that the 2- and 8-bit quantizers will suffer mild degradations.

Second, we have to find the optimal finer thresholds and reconstructed values for quantizers with higher resolution than the reference and the optimal reconstructed values for quantizers with lower resolution than the reference. For example, if the 3-bit quantizer is chosen as the reference, we should align $t_2(1)$ with $t_3(2)$ and find an optimal reconstructed value for each interval partitioned by the aligned thresholds for the 2-bit quantizer. For the 4-bit quantizer, we should align $t_4(2)$ with $t_3(1)$, $t_4(4)$ with $t_3(2)$, and $t_4(6)$ with $t_3(3)$. Then we have to find the optimal finer thresholds $t_4(1)$, $t_4(3)$, $t_4(5)$, and $t_4(7)$ and an optimal reconstructed value for each interval specified by a pair of thresholds. In other words, for those quantizers with higher resolution than the reference, we have to solve for an optimal finer threshold t_m and a pair of optimal reconstructed values $\hat{x}_l$ and $\hat{x}_u$ for each pair of aligned thresholds (t_l, t_u) such that the mean squared quantization error (MSQE)

$$\text{MSQE} = \int_{t_l}^{t_m} (x - \hat{x}_l)^2 \, p(x) \, dx + \int_{t_m}^{t_u} (x - \hat{x}_u)^2 \, p(x) \, dx \tag{5}$$

is minimized, where $p(x)$ is the probability density function (pdf) of x. Following the same derivations used by Max [16], we obtain a set of equations as sufficient conditions to minimize the MSQE

TABLE I
COMPARISON OF MSQE AMONG THE UNIFORM MAX, NONUNIFORM MAX, AND THRESHOLD-ALIGNED QUANTIZERS FOR GAUSSIAN SOURCES

No of Bits	MSQE		
	Uniform Max Quantizer	Nonuniform Max Quantizer	Embedded Max Quantizer
	Gaussian		
1	.3634	.3634	.3634
2	.1188	.1175	.1203
3	.03744	.03455	.03485
4	.01154	.009501	.0009501
5	.003495	.002505	.002524
6	.001041	.0006442	.0006596
7	.0003035	.0001635	.0001705
8	.00008714	.00004117	.00004347

TABLE II
COMPARISON OF MSQE AMONG THE UNIFORM MAX, NONUNIFORM MAX, AND THRESHOLD-ALIGNED QUANTIZERS FOR LAPLACIAN SOURCES

No of Bits	MSQE		
	Uniform Max Quantizer	Nonuniform Max Quantizer	Embedded Max Quantizer
	Laplacian		
1	.5	.5	.5
2	.1963	.1762	.1826
3	.07175	.05448	.05516
4	.02535	.01537	.01537
5	.008713	.004102	.004146
6	.002913	.001061	.001097
7	.0009486	.0002699	.0002863
8	.0003014	.00006806	.00007403

$$\int_{t_l}^{t_m} (x - \hat{x}_l)\, p(x)\, dx = 0 \tag{6}$$

$$t_m = (\hat{x}_l + \hat{x}_u)/2 \tag{7}$$

$$\int_{t_m}^{t_u} (x - \hat{x}_u)\, p(x)\, dx = 0. \tag{8}$$

For most pdf's, it is not easy to solve (6)–(8) simultaneously. We used the numerical method suggested by Max [16] to solve the set of equations. For those quantizers with lower resolution than the reference, we only need to find an optimal reconstructed value $\hat{x}$ for each given pair of aligned thresholds (t_l, t_u) such that

$$\text{MSQE} = \int_{t_l}^{t_u} (x - \hat{x})^2\, p(x)\, dx \tag{9}$$

is minimized. The corresponding sufficient condition is

$$\int_{t_l}^{t_u} (x - \hat{x})\, p(x)\, dx = 0 \tag{10}$$

and it is also solved numerically. Using the 4-bit nonuniform Max quantizer as the reference, we designed the threshold-aligned quantizers for both Gaussian and Laplacian sources. Comparisons are made among the uniform Max, nonuniform Max, and threshold-aligned quantizers; the results are shown in Table I for a Guassian source and in Table II for a Laplacian source. For both sources, the threshold-aligned quantizers suffer little degradation due to their suboptimal thresholds, and their performance is extremely close to that of the nonuniform Max quantizers.

C. Implementation

Simple quantization is very desirable for practical transform coding systems. The complexity of the threshold-aligned quantizer is suitable for hardware implementation. The quantization procedure can be easily implemented as successive comparisons, and only one comparison is required for each allocated bit. When the quantization is refined from i bits to $(i + 1)$ bits, the finer threshold of the $(i + 1)$-bit quantizer to be compared to an input data point can be solely determined from the output code of the i-bit quantizer. Therefore, we can store the thresholds in a look-up table and address the table by the previous output code to obtain finer quantization. The size of the look-up table is approximately 256 storage words, which is reasonably small for hardware implementation in a ROM. Dequantization is even simpler since only table look-up is involved. The size of the look-up table for dequantization is approximately 512 storage words, which is also small. We like to point out that no accumulation of round-off errors due to progression will occur in the proposed progressive quantization scheme since the reconstructed coefficients can be directly read from look-up tables and since no arithmetic operations are involved throughout the progression.

By software simulation, we have applied the proposed progressive quantization scheme to a digital image of 512 × 512 pixels with 8-bit resolution. At $\frac{1}{32}$ bit per intermediate pixel, 8 kbits are transmitted at each intermediate step. For the purpose of comparison, the zig-zag sampled approach of Ngan [12], where the transmission starts with the lower frequency coefficient and continues on to the high-order frequency by following a zig-zag pattern, was also simulated for block sizes of 16 × 16 and 32 × 32. The 8-bit nonuniform Max quantizer is used in the simulation for the zig-zag sampled scheme. The SNR's of the two schemes are shown in Figs. 13 and 14. It is clearly shown in the figures that the bit-sliced approach is superior to the zig-zag sampled approach through almost the entire course of progression. The efficiency of the bit-sliced approach in terms of the bit rate required to achieve a certain SNR value is about twice as good as the zig-zag sampled approach for bit rates less than 4 bits per pixel. When the accumulated bit rate reaches 8 bits per pixel, both systems assign a full 8 bits to each coefficient. In this case, the zig-zag sampled approach should outperform the bit-sliced approach since the former uses an optimal quantizer while the latter uses a suboptimal quantizer. As we pointed out in Tables I and II, the degradation due to the suboptimal quantization is negligible. Simulation results showed that the SNR's are almost identical at a block size of 32 × 32, and the degradation is less than 0.2 dB at a block size of 16 × 16.

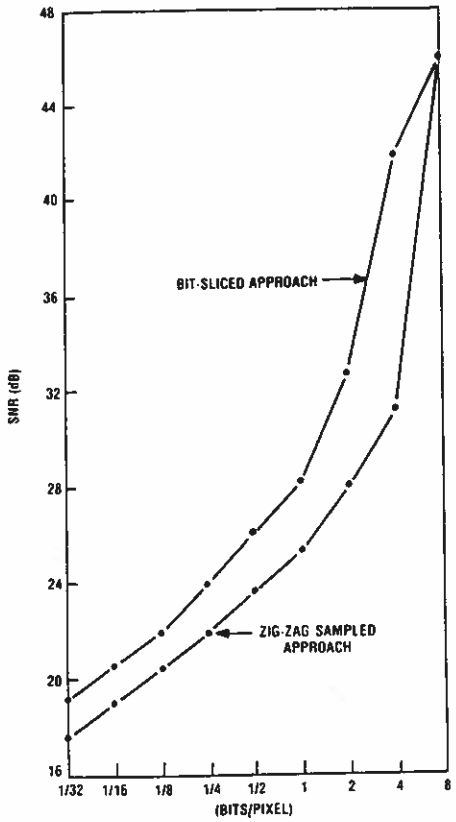


Fig. 13. Comparison of performance between the bit-sliced and the zig-zag sampled approaches at a block size of 32 × 32.

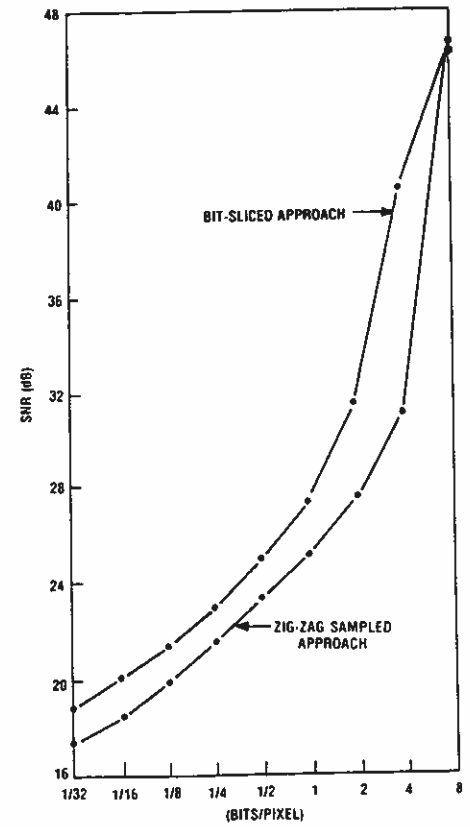


Fig. 14. Comparison of performance between the bit-sliced and the zig-zag sampled approaches at a block size of 16 × 16.

IV. Application to Electronic Radiology

In radiology, as a result of the increased utilization of digital imaging modalities, such as CT and MRI, over a third of the images produced in a typical radiology department are currently in digital form, and this percentage is steadily increasing [3]. New commercial offerings make it possible to routinely digitize film images for clinical use [20], and radiology equipment manufacturers are developing products that will produce "standard" examinations, such as chest and bone images, in digital form [2]. This infusion of digital image sources is occurring at a time when significant new technical developments in the field of digital storage and transmission are close at hand, and this has stimulated planning for the development of picture archiving and communication systems that are capable of transmitting, storing, processing, and displaying radiologic image data [21], [22]. There are important economic and medical reasons for this trend. Studies show that significant cost benefits can result from electronic storage of medical images [2], and it is anticipated that the instantaneous, reliable electronic distribution of radiology images to the appropriate clinical decision-making area will expedite the delivery of quality medical care.

The problem of storage cost has impeded the development of a comprehensive electronic image archive. New developments in storage technology, both optical [4] and magnetic [5], promise a potential solution. It has been estimated [1] that a large radiology department requires an archive capacity of no more than 0.5 to 2 × 10^{13} bits per year, a requirement that could be reduced through data compression [7].

Image presentation is an important aspect in electronic radiology. Traditionally, films for critical care patients are kept in the radiology department on one of several multiviewers with rotating panels. Electronic multiviewers have been developed [23] to electronically simulate traditional multiviewing environments. Progressive transmission of digital pictures can play a key role in developing efficient schemes of image transmission and presentation. A radiologist will have a means to quickly browse through many remotely stored picture panels. As soon as the desired picture panel is recognized, more panel detail can be progressively transmitted until a particular image within the panel is determined. Next, more details of this particular image can be progressively transmitted such that a clinical diagnosis can be made. We have applied the progressive transmission technique, based on adaptive transform coding without progressive quantization, to the panel of 16 CT images shown in Fig. 15. The first three approximations are depicted in Fig. 16. At this level of picture detail, it could be possible to make a decision on which particular images within the panel should be transmitted with sufficient fidelity to make diagnosis possible. Anticipatory paging of image data may improve, where possible, the performance of image transmission at high resolution. Utilizing *a priori* information, the local station can start prefetching compressed images

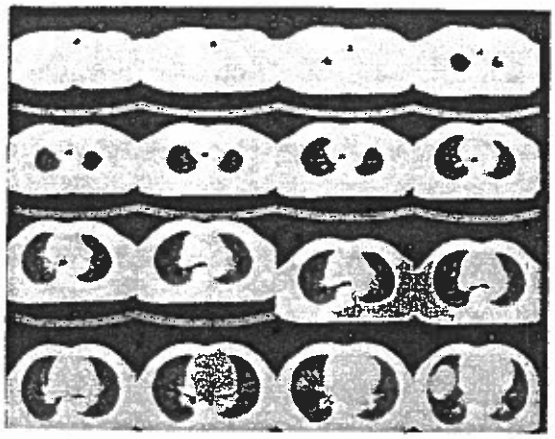
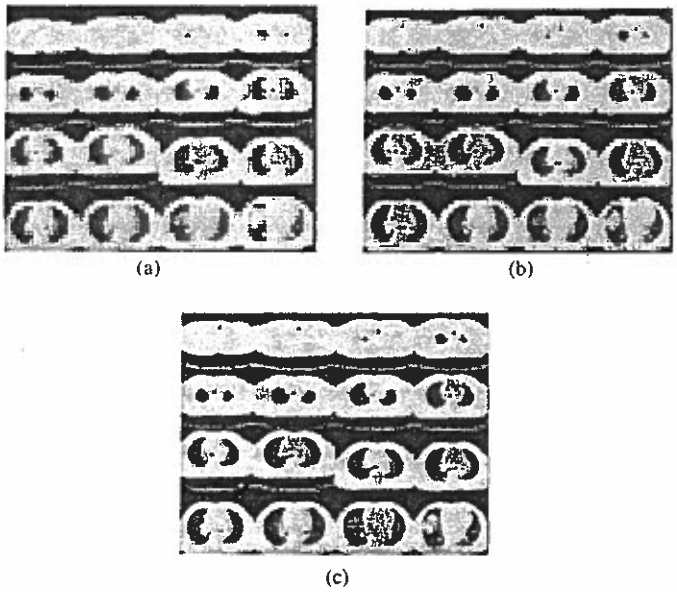


Fig. 15. Original picture panel of 16 CT images.



(a) (b) (c)

Fig. 16. Progressive approximations of picture panels: (a) first, (b) second, and (c) third approximations.

to the local disk whenever the traffic on the picture network allows such a prefetching transmission of data. In this case, the compression techniques discussed herein will improve the latency; the time required to access the data from the local disk.

V. Conclusions

The concept of progressive transmission presents a technical solution to the clinical suitability of imperfect reconstruction from encoded digital diagnostic pictures. It allows the subjective quality to be progressively improved so that proper diagnoses are possible. The greatest value will be in situations where the picture panels are displayed at remote workstations after transmission over low-bandwidth channels such as telephone lines. Transform coding techniques provide an efficient means for achieving progressive transmission of digital pictures. The relative quality of intermediate image approximations from transform-domain coding is superior to that of spatial-domain coding. An improvement of about 10 dB is achieved for a compression factor of 256:1. The improvement is as high as 17.5 dB for a compression factor of 8:1. The SNR converges reasonably into a high value of 56.5 dB as a function of the order of image approximation. This assures that the sequence of intermediate image approximations converges into an almost perfect reconstruction of the input image.

The ability to optimize the different modules of a coding system is actually why transform coding techniques have a potential promise in achieving more efficient schemes for the progressive transmission of pictures. The bit-sliced progressive transmission has been shown to be more efficient in delivering image quality than the zig-zag sampled approach. The SNR's shown in Figs. 13 and 14 indicate that the bit-sliced approach is about twice as efficient as the zig-zag sampled approach. We conjecture that incorporating a human visual model [24], [25] with the progressive coding system is a way to match the objective quality with the subjective quality and will achieve more efficient coding subjectively. Complexity of the threshold-aligned quantizers is suitable for hardware implementation. Other points for future investigation include optimal entropy coding and algorithms for efficient inverse transform of the sparse matrices corresponding to the progressive reconstructions of pictures.

Acknowledgment

The authors wish to thank D. Gray, Manager of the Signal Processing Department of GTE Laboratories, for his helpful discussions and suggestions during the progress of this work. The patience and help of M. Layton of GTE Laboratories in preparing the manuscript are acknowledged.

References

[1] J. R. Cox, G. J. Blaine, R. L. Hill, and R. G. Jost, "Study of a distributed picture archiving and communication system for radiology," in *Proc. 1982 PACS Conf., SPIE*, 1982, vol. 318, pp. 133-142.

[2] S. J. Dwyer, III, *et al.*, "Cost of managing digital diagnostic images for a 614 bed hospital," in *Proc. 1982 PACS Conf., SPIE*, 1982, vol. 318, pp. 3-8.

[3] R. G. Jost, "Image management system for radiology—A current perspective," in *Proc. 4th World Conf. Medical Inform.*, Amsterdam, The Netherlands, Aug. 1983, pp. 365-368.

[4] G. M. Blom, "Optical disc storage," in *Proc. 1982 PACS Conf., SPIE*, 1982, vol. 318, pp. 43-47.

[5] R. L. Benton, "Storage technology for the 1980's," in *Proc. 1982 PACS Conf., SPIE*, 1982, vol. 318, pp. 60-64.

[6] J. G. Dunham, R. L. Hill, G. J. Blaine, D. L. Snyder, and R. G. Jost, "Compression for picture archiving and communication in radiology," in *Proc. 1983 PACS Conf., SPIE*, 1983, vol. 418, pp. 201-208.

[7] S. E. Elnahas, "Data compression with applications to digital radiology," D. Sc. dissertation, Washington Univ. St. Louis, MO, 1984.

[8] K. R. Sloan, Jr. and S. L. Tanimoto, "Progressive refinement of raster scan images," *IEEE Trans. Comput.*, vol. C-28, pp. 871-874, Nov. 1979.

[9] K. Knowlton, "Progressive transmission of grey-scale and binary pictures by simple, efficient, and lossless encoding schemes," *Proc. IEEE*, vol. 68, pp. 885-896, July 1980.

[10] P. J. Burt, and E. H. Adelson, "The laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. COM-31, pp. 532-540, Apr. 1983.

[11] K. Takikawa, "Fast progressive reconstruction of transformed images," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 111-117, Jan. 1984.
[12] K. N. Ngan, "Image display techniques using the discrete cosine transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 173-177, Feb. 1984.
[13] S. E. Elnahas, R. G. Jost, J. R. Cox, and R. L. Hill, "Progressive transmission of digital diagnostic images," *Applications of Digital Image Processing VIII, Proc. SPIE*, 1985, vol. 575, pp. 48-54.
[14] A. K. Jain, "Image data compression: A review," *Proc. IEEE*, vol. 69, pp. 348-389, Mar. 1981.
[15] W. H. Chen and C. H. Smith, "Adaptive coding of monochrome and color images," *IEEE Trans. Commun.*, vol. COM-25, pp. 1285-1292, Nov. 1977.
[16] J. Max, "Quantization for minimum distortion," *IEEE Trans. Inform. Theory*, vol. IT-6, pp. 7-12, Mar. 1960.
[17] L. D. Davisson, "Rate distortion theory and applications," *IEEE Proc.*, vol. 60, pp. 800-808, 1972.
[18] M. D. Paez and T. H. Glisson, "Minimum mean-squared-error quantization in speech PCM and DPCM systems," *IEEE Trans. Commun.*, vol. COM-20, pp. 225-230, Apr. 1972.
[19] K. N. Ngan, "Adaptive transform coding of video signals," *IEE Proc.*, part F, vol. 129, pp. 28-40, Feb. 1982.
[20] R. G. Fraser, E. Breetnach, and G. T. Barnes, "Digital radiography of the chest: Clinical experience with a prototype unit," *Radiol.*, vol. 148, pp. 1-5, July 1983.
[21] J. R. Cox, G. J. Blaine, R. L. Hill, R. G. Jost, and C. D. Shum, "Some design considerations for picture archiving and communication systems," *IEEE Computer Mag.*, pp. 39-49, Aug. 1983.
[22] G. J. Blaine, R. L. Hill, J. R. Cox, and R. G. Jost, "A PACS workbench at MIR," in *Proc. 1983 PACS Conf., SPIE*, 1983, vol. 418, pp. 80-86.
[23] R. G. Jost, R. L. Hill, S. S. Rodewald, and A. P. Rueter, "An electronic multiviewer," in *Proc. ACR 8th Conf. Comput. Appl. Radiol.*, St. Louis, MO, May 1984, pp. 303-311.
[24] K. Tzou, T. R. Hsing, and J. G. Dunham, "Applications of physiological human visual system model to image compression," *Appl. Digital Image Processing VII, Proc. SPIE*, 1984, vol. 504, pp. 419-424.
[25] K. Tzou, "A physiologically based human visual system model for threshold vision and image compression," D.Sc. dissertation, Washington University, St. Louis, MO, 1983.