Medical image compression using a new subband coding method

Faouzi Kossentini and Mark J. T. Smith

School of Electrical & Computer Engineering
Georgia Institute of Technology
Atlanta GA 30332-0250

Allen Scales

Nichols Research Corporation
Huntsville AL 35815-1502

Doug Tucker

University of Alabama at Birmingham
Birmingham AL 35233

ABSTRACT

A recently introduced iterative complexity- and entropy-constrained subband quantization design algorithm is generalized and applied to medical image compression. In particular, the corresponding subband coder is used to encode computed tomography (CT) axial slice head images, where statistical dependencies between neighboring image subbands are exploited. Inter-slice conditioning is also employed for further improvements in compression performance. The subband coder features many advantages such as relatively low complexity and operation over a very wide range of bit rates. Experimental results demonstrate that the performance of the new subband coder is relatively good, both objectively and subjectively.

1 INTRODUCTION

Subband image coding of 8-bit/pixel natural images has been studied extensively in the literature.[1] Common to all subband image coding systems is the decomposition of the input image into subband images using a two-dimensional, mostly separable, filter bank. The resulting subband images are then quantized and entropy coded separately. Since the subband images typically have different statistical properties, a bit allocation algorithm is usually used to distribute bits among the subbands.

The subband image coder proposed in references[2,3] is different in that the design algorithm optimizes the subband quantizers and associated entropy coders jointly within and across the subbands in a complexity- and entropy-constrained framework. Advantages of the design algorithm are that it provides much greater control on the complexity-performance tradeoffs by using multistage residual vector quantizers,[4,5] and that no bit allocation

algorithm is required. This coder works very well for quasi-stationary signals such as most natural images. It can be designed to match the global statistics of a class of images by using a representative training sequence, and can be adapted to local statistics that are specific to individual images through the use of adaptive arithmetic coding.[6,7]

Like many natural images, medical images that are acquired from the same anatomical section using the same imaging modality are also quasi-stationary. A specific class of images, such as computed tomography (CT) axial slice head images, features similar global structural appearances due to the similarity in anatomical and tissue structures among different patients. On the other hand, anomalies such as pathologies or image artifacts, different density tissues, and different imaging conditions, produce image patterns that are not part of the training sequence. In medical imaging, some of these local statistics represent very critical information, and failing to reproduce such unique patterns can significantly impair the usefulness of the compressed medical image.

The problem of subband coding medical images using the coder proposed in references[2,3] is addressed in this paper. Although there are many similarities between natural and medical images, the problem of subband coding medical images is very different. Medical images are obtained from a variety of devices, and the images produced have different characteristics (e.g. dynamic range, spatial resolution) as well as distinct statistical dependencies. Performance can be improved by designing the filters, decomposition structures, quantizers, and entropy coders differently. For example, medical images contain a significant amount of both high and low frequency information. Thus, uniform decompositions fair better than the octave-band (or wavelet) decomposition frequently used in natural image subband coding. Moreover, a higher degree of fidelity is required in the compressed-decompressed images. Experimental work[8] shows that the choice of filters and filter design parameters has little or no effect on the reproduction quality in the low bit rate range. However, as will be discussed in this paper, filters do affect the subband coder's performance both objectively and subjectively in the high bit rate (high fidelity) range. Another problem associated with high fidelity subband coding is the large complexity usually required by the quantizers and corresponding entropy coders. Fortunately, since the proposed subband coder employs multistage residual vector quantizers, the complexity associated with both quantization and entropy coding is still relatively low.

The subband coder described in references[2,3] exploits both statistical intra-band and inter-band dependencies within an image simultaneously, mainly through complexity-constrained high order conditional entropy coding. In this work, inter-band dependencies both within a slice image and between slice images are exploited, resulting in a 5-10 % improvement in compression-complexity performance for the same reproduction quality. Next, we provide a brief description of the coder's components. This is followed by a discussion of design and complexity issues. This paper concludes with a discussion of the application of the subband coder to medical images and a presentation of some CT Head image coding experimental results.

2 THE SUBBAND CODER

Figure 1 shows the block diagram of the subband encoder used in this work. As is the case in conventional subband coding, the input image is first decomposed into M subband images using an analysis transformation. In this work, we employ a uniform tree-structured decomposition which is based on 2-band exact reconstruction filter banks. Each subband image is then encoded using a sequence of P_m $(1 \leq m \leq M)$ residual vector quantization (RVQ) fixed length encoders. Multistage RVQ is instrumental in drastically reducing the complexity of encoding/decoding as well as entropy coding, while still maintaining good rate-distortion performance. Advantages of multistage RVQs will be described in the following sections. Although encoding optimality can generally be achieved through exhaustive searching of the RVQ stage codebooks in all subbands (i.e. embedding the synthesis transformation in the encoding procedure), experiments have shown that dynamic M-search[9] of the stage codebooks in each subband separately usually leads to the best complexity/performance tradeoffs.

The output symbol of each of the stage vector quantizers is fed into an entropy coder driven by a high order

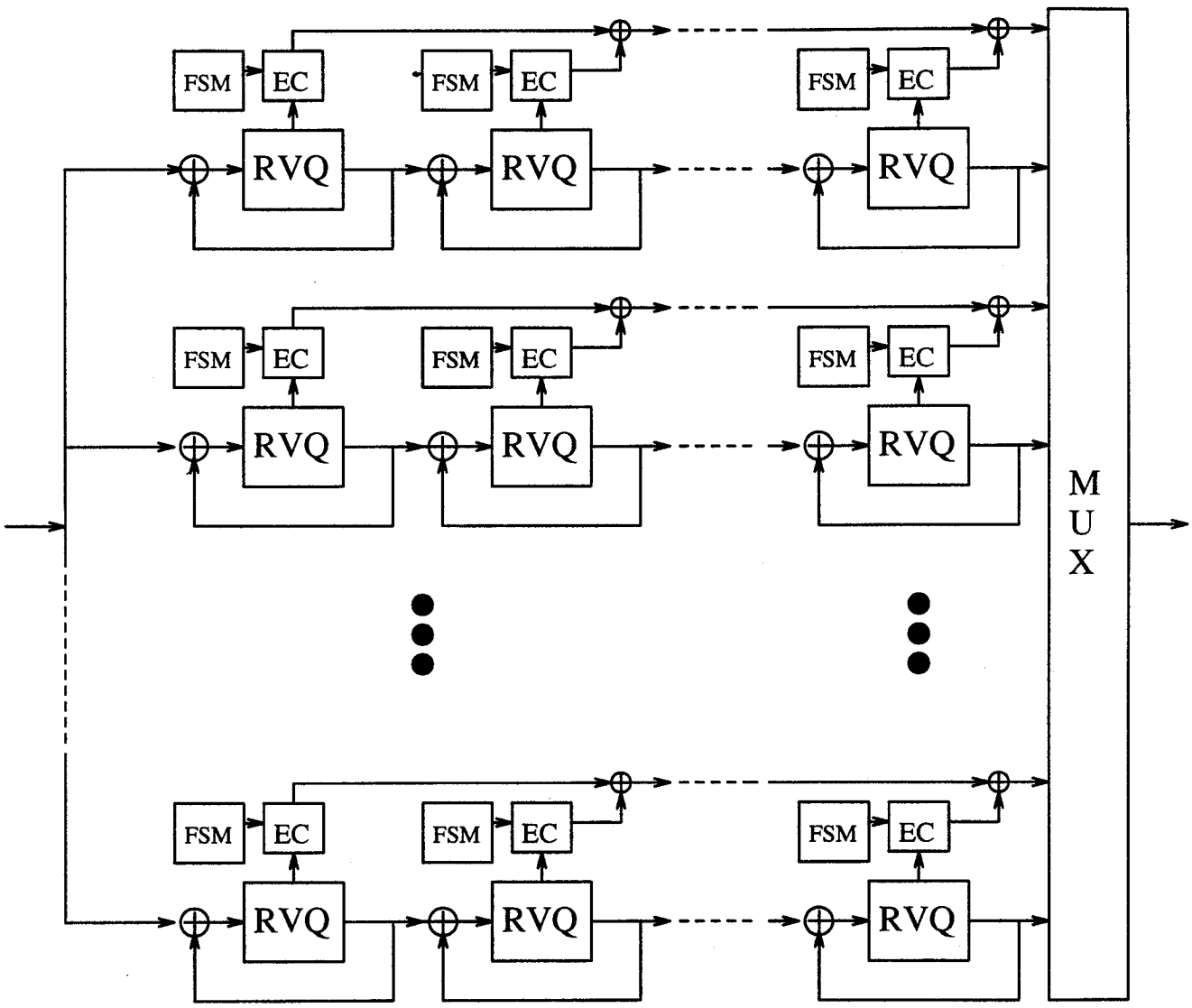


Figure 1: Basic block diagram of the subband encoder.

stage statistical model that is governed by a finite state machine (FSM). The FSM allows the statistical model to utilize information about previously coded stage vectors. A nonlinear function F given by $u = F(s_1, s_2, \ldots, s_n)$, where $s_1, s_2, \ldots, s_n$ are n conditioning symbols, or previous outputs of particular fixed-length RVQ stage encoders, is used here to determine the conditioning state u. As will be described in the next section, F is a many-to-one function that is represented by a table mapping each combination of realizations of conditioning random variables into a conditioning state. Since only previously coded symbols are used by the FSM, no side information is necessary and the decoder can track the state of the encoder by only storing the same table. Finally, the output bits of the entropy coders are combined together and sent to the channel.

3 DESIGN AND IMPLEMENTATION ISSUES

The algorithm used to design the subband coder minimizes iteratively the expected distortion subject to a constraint on the complexity-constrained high order conditional entropy of the stage vector quantizers (VQs). The popular squared error measure is used here as the distortion measure. This design algorithm is based on a Lagrangian minimization, and is a generalization of entropy-constrained algorithms described in.[10,11,5] Details of the *joint* optimality conditions used in the development of the algorithm and convergence issues are discussed elsewhere.[3]

Given a Lagrangian parameter, λ, which is chosen based on the overall rate and distortion of the subband system (i.e. no bit allocation algorithm is required), the entropy-constrained joint subband quantization algorithm consists of three optimization steps. The encoder optimization step involves exhaustively searching all RVQ stage codebooks, a task which requires a huge computational load. A large reduction in complexity can be achieved by using dynamic M-search. This results in only a small loss of performance. The decoder optimization step consists of using the Gauss-Seidel algorithm[5] to minimize iteratively the average distortion between the input and the synthesized reproduction of all stage codebooks in all subbands. The complexity can be drastically reduced by, for example, grouping neighboring stage codebooks in neighboring subbands and jointly optimizing each group independently. This typically results in less than a 0.10 dB loss in signal-to-noise (SNR) performance.

Since actual entropy coders are not used explicitly in the design process, the entropy coder optimization step is equivalent to a high order statistical modeling procedure. In terms of complexity (i.e. computational load and memory requirements), high order statistical modeling is potentially the most demanding task of the design algorithm. However, using the multistage residual structure not only substantially reduces the large complexity demands, usually associated with high order conditional entropy coding, but also makes exploiting high order statistical dependencies much easier by producing multiresolution approximations of the input subband images. However, there are still many issues to be addressed. Multistage RVQs reduce the complexity because the output alphabet of the stage quantizers is typically very small (e.g., 2, 3, or 4), but the complexity of a stage entropy coder is still exponentially dependent on its order (number of conditioning symbols or random variables) and/or the output alphabet sizes of the stage quantizers. Moreover, multistage RVQs also introduce another dimension to the statistical modeling problem, which significantly increases the number of possible combinations of conditioning symbols. Finally, many of the frequencies of combinations of conditioning symbols, gathered during the training process and used as estimates for probabilities, have zero values, producing empty states. This complicates the encoding stage because a combination of conditioning symbols corresponding to an empty state may occur. This is the so-called empty state problem, a problem usually associated with finite state machines.

In reference,[2] a complexity-constrained statistical modeling algorithm is proposed that attempts to simultaneously solve the above problems. To help illustrate the algorithm, Figure 2 shows the inter-stage, inter-band, and intra-band conditioning scheme employed in this work. Each image shown in the figure is a multistage approximation of a particular slice image. Note that statistical dependencies both within and across slice images can be exploited. For each stage (m, p) in each subband m, a 5-dimensional initial region of support $\mathcal{R}_{m,p}$ containing a sufficiently large number $R_{m,p}$ of conditioning symbols is first chosen. Then, the $n_{m,p}$, $n_{m,p} << R_{m,p}$, con-

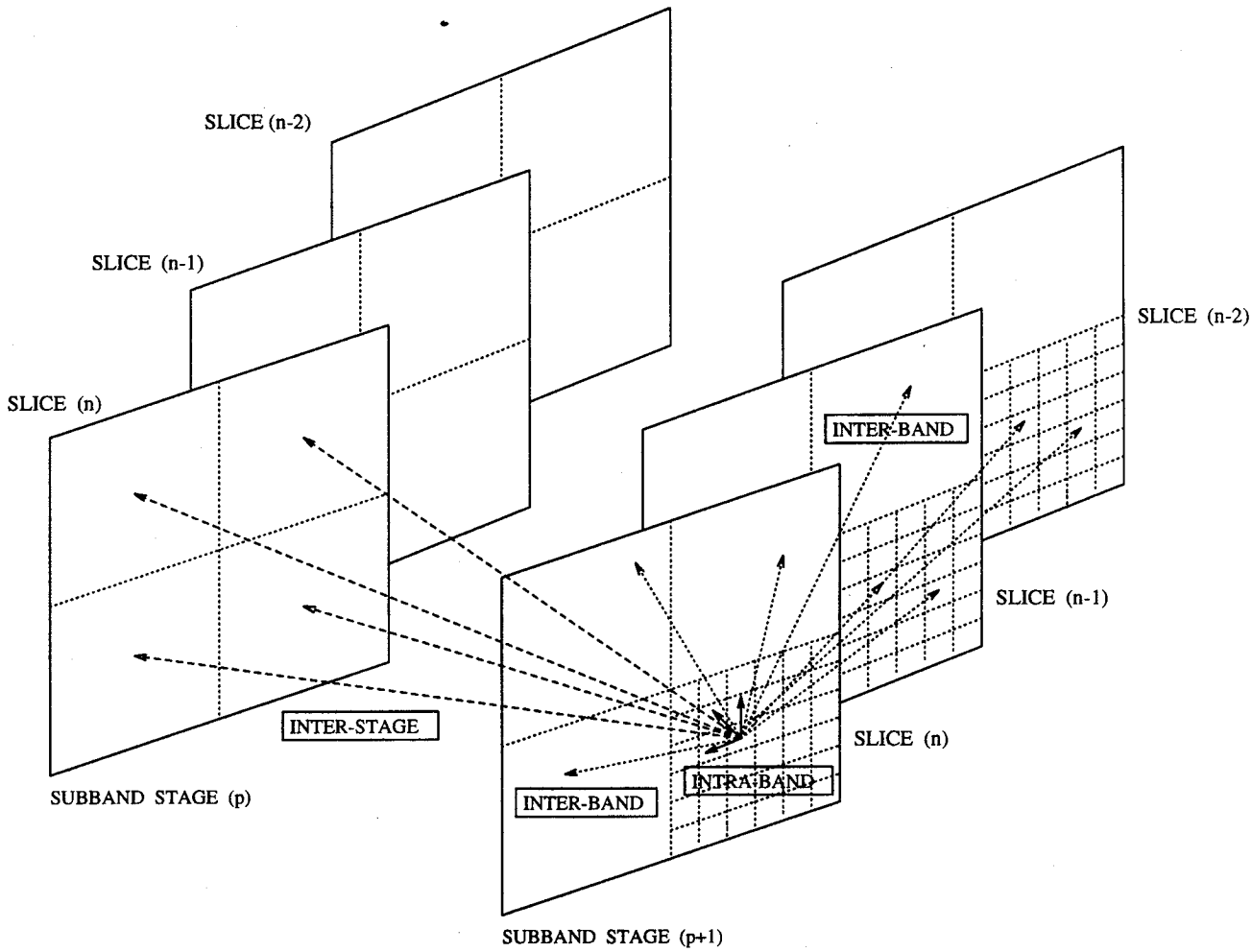


Figure 2: Inter-stage, inter-band, and intra-band conditioning scheme within an image sequence.

ditioning symbols $s^1, \ldots, s^{n_{m,p}}$ that lead to the smallest $n_{m,p}$th order conditional entropy $H(J_{m,p}|s^1, \ldots, s^{n_{m,p}})$ are located by building a special tree and using the dynamic M-search algorithm.

The next step of the algorithm is to find orders of all stage statistical models such that the average entropy in all subbands given a fixed level of complexity, expressed here in terms of total number of probabilities to be computed/stored, is minimized. The process described above is repeated for each stage (m, p) and many values of $n_{m,p}$, producing, let's say, $L_{m,p}$ complexity-entropy pairs per stage. For each complexity-entropy pair, complexity is given by $\mathcal{N}_{m,p} = S_{m,p} N_{m,p}$, where $S_{m,p}$ is the number of all combinations of realizations of the conditioning symbols and $N_{m,p}$ is the output alphabet size of stage p in band m. Once all complexity-entropy pairs are obtained, a tree with $\sum_{m=1}^{M} P_m$ branches, where P_m is the number of stage codebooks in the mth subband, can be built. Each branch of the tree is a unary tree of length $L_{m,p}$, and each node represents a complexity-entropy pair. The generalized BFOS algorithm[12] is then used to minimize the average entropy subject to a constraint $\mathcal{N}_1$ on the total number of conditional probabilities.

The FSM statistical model for each stage (m, p) employs a mapping F to determine the state given $n_{m,p}$ available conditioning symbols. This mapping F is one-to-one and is actually given by a table that contains the numbers $0, 1, \ldots, S_{m,p} - 1$, representing each of the possible combinations. Up to this point, the number of conditioning states for each of the stages is typically large. Many of the corresponding tables of probabilities may still be empty after the design is completed, thereby occupying memory which can usually be more efficiently used. Moreover, as mentioned earlier, these empty tables may be visited during actual encoding even though they were never visited during the design process. Therefore, the last step of the algorithm is to further reduce the number of conditioning states through quantization. In this work, the PNN algorithm[13] has been shown to be successful in reducing the number of states by orders of magnitude while still bounding the loss in entropy performance to about 1%. The PNN algorithm first merges all of the empty states with the least probable state into *one* conditioning state, thereby completely removing empty states. Then, the two conditioning states resulting in the lowest increase in entropy (when merged) are combined into one conditioning state, and so on until only one state, which represents one table of first order probabilities, is obtained. Since the objective is to minimize the complexity-constrained average entropy, the BFOS algorithm is again used, where a much smaller complexity value $\mathcal{N}_2$ is the constraint or criterion.

In the context of medical image compression, quantization of the conditioning states has two important advantages. First, the stage statistical model orders can be allowed to grow to relatively large numbers, which generally results in significantly lower average entropy because most medical images feature high order global statistical dependencies. This also incurs a small additional enoding/decoding complexity, since only larger mapping tables have to be stored/accessed. Second, the merging process improves the robustness of the subband coder because only global statistics are carried through, and the possibility of a strong mismatch between individual medical images and the subband coder is less likely.

4 EXPERIMENTAL RESULTS

A total of 90 axial slice CT head images with no abnormal findings were used for training and testing. Images were selected retrospectively from studies of 10 patients undergoing scans as a part of their clinical care. A General Electric (Waukasha, Wi) Hi-Lite Advantage CT scanner was used to produce all images which were either 3 mm (posterior fossa, 120 kVp, 320 mAs) or 5 mm (mid-brain, 120kVp, 240 mAs) thick slices. All images were of size 512×512 with 12 bit/pixel amplitude resolution. No special image processing or reconstruction algorithms were applied. Each image was extracted from the CT scanner's proprietary database using software tools supplied by the manufacturer. Subsequently, the proprietary header information was removed and the raw images were stored with 16-bit amplitude precision. A set of 12 slice images was kept for testing, and was not used as part of the training sequence.

Two experiments were performed. The first investigated the performance of the spatial subband coder, while the second considered exploiting both spatial and inter-slice dependencies within the CT image sequence. In both experiments, each slice image was first fed into a 2-level balanced tree structured filter bank, producing 16 image subbands. The A11 allpass polyphase exact reconstruction IIR filters[14] were used in our simulations. Many other filters, such as the Johnston 16-tap and 32-tap QMFs[15] and the Daubechies 32-tap wavelet filters,[16] were tested and were found to be inappropriate. The SNR reconstruction performance of these filters for the test CT images did not exceed 52 dB even when quantization was not performed. This is unsatisfactory in light of the fact that the medical community demands a SNR reconstruction performance that is usually 50 dB or higher.

In this work, we employ a vector size of 1×1 (scalar quantizer). Although k-dimensional vector quantizers are potentially better than scalar quantizers, their complexity is very large. Thus we found scalar quantizers to be more appropriate, particularly considering the high rates of operation. To initialize the design algorithm, a multistage residual scalar quantizer (RSQ) is obtained for each subband image, as described in reference.[3] The number of scalars in each RSQ stage codebook is set to 3. Non-uniform stage codebook sizes were considered, but no significant improvement in rate-distortion performance was obtained. Furthermore, choosing a uniform stage codebook size for all RSQs in all subbands simplifies both quantization and arithmetic encoding/decoding. We have also tried stage codebooks of sizes $2, 3, 4, 5, \ldots$ and have determined that 3-scalar stage codebooks provide the best complexity-performance tradeoffs for the training CT sequence.

In both experiments, dynamic M-search with a fixed threshold of 10 was used in the encoder optimization. Moreover, a joint decoder optimization between stages only is used in both cases. In other words, no joint optimization between subband decoders is performed. During the statistical modeling procedure, the value of $\mathcal{N}_1$ was set to 8192, and the value of $\mathcal{N}_2$ was set to 1024. For each state of the FSM model at stage (m, p), only two probabilities, quantized to values between 1 and 256, are needed by each adaptive arithmetic coder. Since the probabilities are constrained to be powers of 2, no multiplications are necessary in the implementation of the arithmetic encoders/decoders. Dynamic adaptation[17] was performed to further lower the bit rate. Although good performance high rate coders typically require a large design complexity, such is not the case in the first experiment. About 12 CPU hours on a Sparc 10 Sun Station were required to design subband coders operating at rates between 0.80 and 2.0 bpp. However, the design complexity in the second experiment is relatively large. More specifically, more than two days in CPU time were required to design the same number of codebooks and corresponding entropy coders. This is due to the fact that inter-slice conditioning requires that a much larger region of support be used, which complicates statistical modeling.

The encoding/decoding complexity and memory of the CT image subband coder are relatively small. The memory required to store all codebooks for each rate-distortion point is only 1152 bytes, while that required to store the conditional probabilities is approximately 1024 bytes. Furthermore, the average number of operations (multiplies/adds) required for encoding is 14.64 per input sample. Decoding requires 10 multiplies/adds. By placing some constraints on the coder, encoding/decoding can also be implemented without multiplications. However, such constraints also affect the rate-distortion performance. Full evaluation of a multiplication-free implementation of this subband coder for medical image compression is the subject of further research.

The objective quality of the reconstructed CT slice images is very good. Table 1 shows rates and SNRs for all 12 test CT slice images for average rates of 2.00, 1.50, 1.0 and 0.80 bpp, corresponding to compression ratios of $6:1$, $8:1$, $12:1$, and $15:1$, respectively. The SNR is defined by

$$\mathrm{SNR} = -10 \log_{10} \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} (x(i,j) - \hat{x}(i,j))^2}{\sum_{i=1}^{N} \sum_{j=1}^{M} (x(i,j) - \mu)^2} \tag{1}$$

where $N \times M$ is the number of samples in the image, $x(i,j)$ and $\hat{x}(i,j)$ represent the original and the coded value (respectively) of the (i,j)th sample, and μ is the mean of $x(i,j)$. Figure 3(a) shows the original slice image # 11. Figure 3(b) shows the residual image formed by taking the absolute difference between the original image and the reconstructed one at a bit rate of 0.73 bpp. Note that the intensities of the residual image have been magnified by a factor of 16. Finally, Table 2 compares the bit rates and SNRs of the first and second experiments for the

	6:1		8:1		12:1		15:1	
	BR	SNR	BR	SNR	BR	SNR	BR	SNR
SLICE #1	2.16	56.83	1.75	52.89	1.13	48.75	0.89	46.35
SLICE #2	2.12	57.06	1.61	52.99	1.12	48.64	0.92	46.50
SLICE #3	2.09	57.11	1.67	53.05	1.16	48.77	0.89	46.41
SLICE #4	2.11	56.97	1.59	52.89	1.05	48.61	0.83	46.24
SLICE #5	1.99	57.27	1.42	53.00	0.98	48.90	0.79	46.62
SLICE #6	2.01	57.13	1.46	52.92	1.01	48.83	0.80	46.59
SLICE #7	2.03	57.18	1.48	52.94	1.02	48.82	0.82	46.54
SLICE #8	2.01	57.31	1.45	53.03	0.95	48.89	0.80	46.62
SLICE #9	1.94	57.35	1.39	53.10	1.02	48.86	0.77	46.61
SLICE #10	1.89	57.33	1.42	53.03	0.98	48.84	0.74	46.58
SLICE #11	1.84	57.36	1.38	53.06	0.93	48.89	0.73	46.65
SLICE #12	1.84	57.13	1.33	52.88	0.90	48.93	0.71	46.

Table 1: Bit rate (BR) in bits per pixel (bpp) and signal-to-noise ratio (SNR) in decibels (dB) for the 13 slice images used in the first experiment at compression ratios of 6:1, 8:1, 12:1, and 15:1.

	6:1		15:1	
	BR	SNR	BR	SNR
Non-inter-slice	1.84	57.36	0.73	46.65
Inter-slice	1.71	57.29	0.68	46.68

Table 2: Bit rate (BR) in bits per pixel (bpp) and signal-to-noise ratio (SNR) in decibels (dB) for the slice image #11 at compression ratios of 6:1 and 15:1.

slice image # 11 at the two 6 : 1 and 15 : 1 compression ratios. Looking at Table 2, one can see that inter-slice conditioning resulted in a 7 % decrease in bit rate roughly for approximately the same objective quality.

Compressed and reconstructed images of the non-inter-slice conditioning experiment were also viewed by an experienced radiologist for his impressions. Viewing was performed in a low-light environment. Images were displayed on an Image Systems M21PMAX 1280 x 1024 display using a Dome MD2kEISA display controller on a DELL Omniplex Pentium personal computer running MS-DOS 6.21 and an image viewing software customized from DOME's software library. The same series of 12 images, each compressed at 6:1, 8:1 and 12:1, were used. For each image viewing, the original image and a single compressed-reconstructed image were displayed together. All images compressed at 6:1 were viewed first, followed by the 8:1 and the 12:1 images, respectively. The radiologist was allowed to adjust window and level settings and no time constraints were imposed. The radiologist's impression was solicited. The radiologist reported no noticeable difference between the original image and the 6:1 or 8:1 compressed-reconstructed images. For two of the twelve 12:1 compressed-reconstructed images the observer noted slight enhancement of the high frequency component (noise) of the compressed-reconstructed image.

5 CONCLUSIONS

The results of the preliminary viewing of the compressed-reconstructed images by a radiologist were encouraging. We are currently conducting more rigorous observer performance tests to determine objectively the performance of radiologists using the compressed-reconstructed images. The computational complexity and memory requirements make this coder a suitable candidate for implementation in real-time hardware.

6 ACKNOWLEDGMENTS

The authors would like to acknowledge J. Kevin Smith, M.D., Ph.D. for his assistance. The first author would also like to acknowledge Wilson C. Chung for supplying software implementations for the various filters studied in this paper.

7 REFERENCES

[1] J. W. Woods, ed., *Subband Image Coding.* Norwell, MA: Kluwer Academic Publishers, 1991.

[2] F. Kossentini, W. Chung, and M. Smith, "Subband image coding with jointly optimized quantizers," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing*, (Detroit, MI, USA), Apr. 1995.

[3] F. Kossentini, W. Chung, and M. Smith, "A jointly optimized subband coder," *Submitted to Transactions on Image Processing*, July 1994.

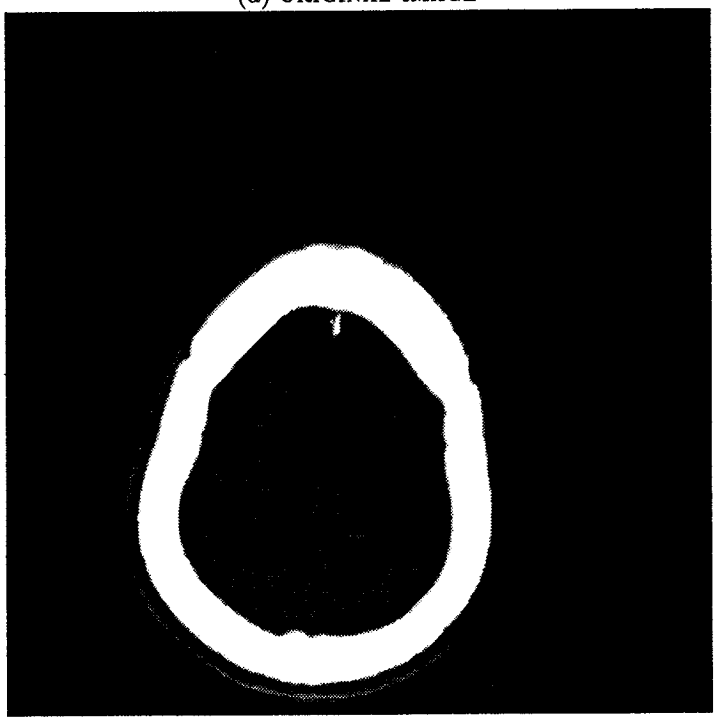
[4] B. H. Juang and A. H. Gray, "Multiple stage vector quantization for speech coding," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 1, pp. 597–600, April 1982.

[5] F. Kossentini, M. Smith, and C. Barnes, "Necessary conditions for the optimality of variable rate residual vector quantizers," *Submitted to Transactions on Information Theory in June 1993. Revised in May 1994.*

[6] G. Langdon, "An introduction to arithmetic coding," *IBM J. Res. Dev.*, vol. 28, pp. 135–149, Mar. 1984.

[7] A. C. Popat, *Scalar Quantization with Arithmetic Coding.* PhD thesis, M.I.T., Cambridge, MA, 1986.

(a) ORIGINAL IMAGE



(b) RESIDUAL IMAGE

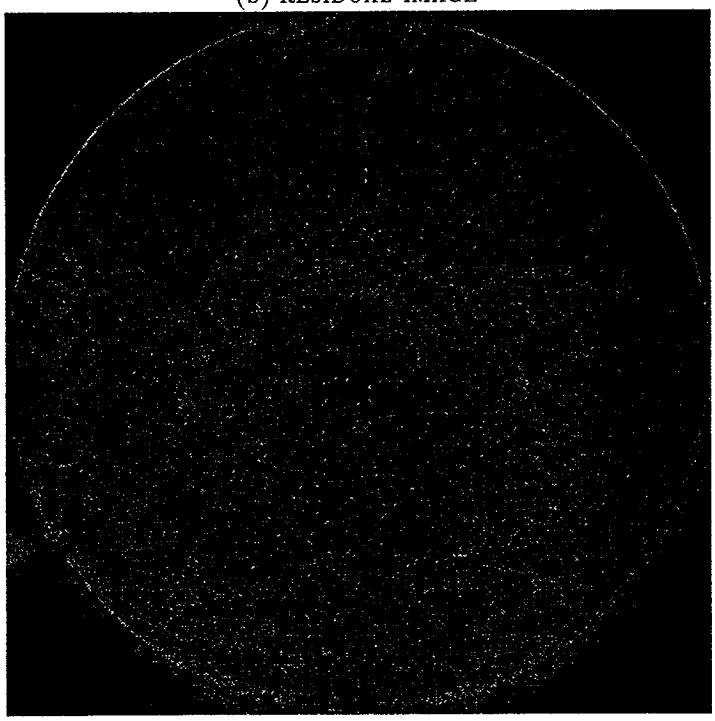


Figure 3: The original slice image #11 and the corresponding residual image at a bit rate of 0.73 bpp. Intensities of the residual image have been magnified by a factor of 16.

[8] A. Docef, F. Kossentini, W. Chung, and M. Smith, "Multiplication-free subband coding of color images," in *Data Compression Conference*, (Snowbird, UT, USA), Mar. 1995.

[9] F. Kossentini and M. Smith, "A fast searching technique for residual vector quantizers," *Signal Processing Letters*, vol. 1, pp. 114–116, July 1994.

[10] P. A. Chou, T. Lookabaugh, and R. M. Gray, "Entropy-constrained vector quantization," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-37(1), pp. 31–42, Jan. 1989.

[11] F. Kossentini, M. Smith, and C. Barnes, "Entropy-constrained residual vector quantization," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing*, vol. V, (Minneapolis, MN, USA), pp. 598–601, Apr. 1993.

[12] E. A. Riskin, "Optimal bit allocation via the generated BFOS algorithm," *IEEE Trans. on Information Theory*, vol. 37, pp. 400–402, Mar. 1991.

[13] W. H. Equitz, "New vector quantization clustering algorithm," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 1568–1575, Oct. 1989.

[14] M. Smith and S. Eddins, "Analysis/synthesis techniques for subband image coding," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 38, pp. 1446–1456, Aug. 1991.

[15] J. Johnston, "A filter family designed for use in quadrature mirror filter banks," *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing*, pp. 291–294, April 1980.

[16] I. Daubechies, *Ten Lectures on Wavelets.* Philadelphia, Pennsylvania: SIAM, 1992.

[17] G. G. Langdon and J. Rissanen, "Compression of black-white images with arithmetic coding," *IEEE Transactions on Communications*, vol. 29, no. 6, pp. 858–867, 1981.