Wavelet-Based Vector Quantization for High-Fidelity Compression and Fast Transmission of Medical Images

Sunanda Mitra, Shuyu Yang, and Vadim Kustov

Compression of medical images has always been viewed with skepticism, since the loss of information involved is thought to affect diagnostic information. However, recent research indicates that some wavelet-based compression techniques may not effectively reduce the image quality, even when subjected to compression ratios up to 30:1. The performance of a recently designed wavelet-based adaptive vector quantization is compared with a well-known wavelet-based scalar quantization technique to demonstrate the superiority of the former technique at compression ratios higher than 30:1. The use of higher compression with high fidelity of the reconstructed images allows fast transmission of images over the Internet for prompt inspection by radiologists at remote locations in an emergency situation, while higher quality images follow in a progressive manner if desired. Such fast and progressive transmission can also be used for downloading large data sets such as the Visible Human at a quality desired by the users for research or education. This new adaptive vector quantization uses a neural networks-based clustering technique for efficient quantization of the wavelet-decomposed subimages, yielding minimal distortion in the reconstructed images undergoing high compression. Results of compression up to 100:1 are shown for 24-bit color and 8-bit monochrome medical images.

Copyright © 1998 by W.B. Saunders Company

COMPRESSION OF MEDICAL IMAGES for fast transmission is not accepted universally because of possibility of losing diagnostic information. Recently, however, it has been reported¹ that up to 30:1 compression is acceptable using wavelet-based compression techniques, rather than JPEG compression, by which more than 10:1 compression is not considered acceptable by most radiologists.

Wavelet coding^{2,3} is a class of subband coding that allows selective coding of subbands significant to visual perception. The encoding process actually

determines the distortion introduced by quantization. It is well known from rate distortion theory that vector quantization yields optimal distortion. However, implementation of vector quantization specifically the encoding process is quite involved. On the other hand, the decoding process in vector quantization is relatively simple and can be achieved from look-up tables.

Generally, a simple statistical clustering technique is used for generating the code vectors in vector quantization.^{4,5} We have developed novel clustering techniques^{6,7} for on-line generation of code vectors^{8,9} and applied them to quantization of wavelet-transformed subimages. The selection of the number of clusters generated is decided by the quantization level to be applied, ie, whether coarse or fine quantization is needed. The code vectors representing the clusters are generated here by an adaptive vector quantization technique, namely AFLC-VQ (adaptive fuzzy leader clustering-vector quantization).¹⁰ AFLC-VQ is based on an integration of self-organizing neural networks with fuzzy membership values of data samples for updating the centroid values of the clusters as the samples are presented.

The filter length of the wavelet transformation matrix plays an important role on the accuracy of the encoding scheme. Based on the distribution characteristics of the wavelet coefficients, appropriate filter lengths for specific wavelets are chosen before encoding.¹¹

We have used this multiresolution encoding and decoding scheme for large color images such as NASA Landsat and the Visible Human color images, as well as standard color images using color transformation from RGB to YIQ and L*a*b* planes.¹² Different color planes were used to determine the effectiveness of color planes on adaptive arithmetic coding for lossless coding.

For color or monochrome images, AFLC-VQ yields excellent reconstructed images even at very low bit rates corresponding to compression ratios greater than 100:1 with acceptable mean square error (MSE) and peak signal-to-noise ratio (PSNR). The advantage of being able to use such high compression ratio without introducing sufficient

From the Department of Electrical Engineering, Texas Tech University, Lubbock, TX.

Supported in part by Contract No. P NLM 97-062/VMS from the National Library of Medicine, National Institutes of Health, and by the Advanced Research Program (ARP) from the state of Texas (Grant No. 003644-176-ARP).

Address reprint requests to Sunanda Mitra, PhD, Department of Electrical Engineering, Texas Tech University, Lubbock, TX 79409-3102.

*Copyright © 1998 by W.B. Saunders Company
0897-9118/98/1104-2003\$8.00/0*

Fig 1. Schematic diagram of progressive transmission of coded images.

distortion in the reconstructed image is that, under emergency situation, the compressed image can be sent to a remote location quickly over the Internet to a radiologist for first evaluation while images with better quality can follow in a progressive manner if desired.

In this report, we briefly describe the principle of vector quantization, our adaptive clustering techniques, and compare the results of our compression technique with a well-known wavelet-based scalar quantization technique, namely EZW¹³ and JPEG,¹⁴ for medical images.

METHODOLOGY

Image Coding

Figure 1 shows a general scheme of radiographic image coding. This scheme also provides a completely lossless image coding with a compression ratio ranging from 7:1 to 10:1 depending on the image to background area ratio, image content, and the use of an adaptive arithmetic coding.¹² If exact reproduction of the original image at the decoding stage is not essential, then a lossy compression scheme using either scalar or vector quantization is used.

Quantization

A. Scalar quantization. In scalar quantization the gray level value of each pixel in an original image or transformed image (as in JPEG or wavelet-based coding) is mapped to a lower range of values thus reducing the bits per pixel (bpp) values. In JPEG, discrete cosine transform is performed on each 8×8 block of an image, and the transformed image is then quantized by a scaled normalization matrix before applying other encoding steps. Due to this small block divisions designed to minimize the statistical variance within the block, an artifact arising from the Gibb's phenomenon or commonly known as the blocking artifact yields significant distortion in the decoded image at high compression ratios. In fact, many radiologists do not find the

image quality of the decoded images acceptable when compressed over 10:1 compression ratios by the JPEG scheme.

B. Wavelet-based scalar quantization. Since multiresolution wavelet decomposition allows subband coding that can be easily tailored to visual perception and yields natural subsampling to smaller sizes without causing severe blocking artifacts observed in JPEG compression scheme, wavelet-based compression techniques have been chosen in recent years specifically when compression ratios higher than 10:1 are desired. Most popular compression techniques employ scalar quantization because of its simplicity of implementation, despite suboptimal

Fig 2. Comparison of reconstructed image from AFLC-VQ, JPEG, and EZW compression on cervical spine image (spine.raw, greylevel, $1,024 \times 1,024$).

Table 1. Comparison Between AFLC-VQ, JPEG, EZW (spine.raw)

Compression Scheme	CR	MSE	PSNR
AFLC-VQ	43.81	4.13	41.98
	58.94	4.32	41.77
	99.90	22.26	34.66
JPEG	43	1.69	45.86
	58	2.30	44.51
	99	53.05	30.88
EZW	40	315.97	23.13
	60	278.17	23.69
	100	335.50	22.87

Abbreviation: CR, compression ratio.

performance. The most well-known wavelet-based compression algorithm with successive approximation scalar quantization involving adaptive arithmetic coding is known as EZW. We have implemented a modified version of EZW so that the performance of our adaptive vector quantization technique can be compared with it.

C. Vector quantization. Vector quantization is the process of grouping similar vectors into the same cluster so that the centroid vector of each cluster is able to reconstruct the entire set of vectors in that cluster with a minimum distortion dictated by the similarity and optimization criteria employed. Therefore, the centroid vectors or the code vectors form the codebook that consists of the addresses of all code vectors, thus reducing the file size of the original data set tremendously while introducing minimal distortion. The distortion introduced depends on the complexity of the similarity criterion, optimization criterion, and the level of quantization used. In most vector quantization techniques, Euclidean distance measure is used as a similarity measure. We have used an integrated neuro-fuzzy clustering

technique⁶ where a modified adaptive resonance theory (ART)-type¹⁵ neural networks initially clusters the vectors into similar groups and then the cluster centroids are updated by an optimization constraint, including fuzzy membership values of the data samples following the fuzzy C means (FCM)^{16,17} nonlinear equations. A later modification of this algorithm also included a new similarity measure, enabling formation of clusters of any shape and a Kohonen type upgrading rule for the centroids.¹⁸ These modifications allow better allocation of the vectors to the right clusters by introducing an intracluster membership function while upgrading the centroids. It also eliminates some of the problems with ART-type networks due to normalization of the sample vectors. Although the achievability of minimum distortion at a specific bit rate by vector quantization has been theoretically proven from rate distortion theory almost half a century ago,¹⁹ practical implementation of vector quantization for small sizes and classes of images has been achieved relatively recently.⁵ Many of the earlier algorithms using simple statistical clustering suffer from a number of problems namely lack of convergence, getting trapped in local minima, and inability to handle large data sets. More advanced vector quantization algorithms^{8,20} have eliminated some of these problems. However, vector quantization of large data sets as encountered in many medical images still remains a challenging problem.

We present an adaptive vector quantization technique that is capable of encoding large size as well as color images with minimum distortion in the decoded images even at a compression ratios greater than 100:1. The success of this new technique depends on an adaptive clustering technique with efficient optimization criteria in combination with multiresolution wavelet decomposition.

For each image class, a suitable vigilance parameter, τ , needs to be selected for a range of clusters that could be formed to

Fig 3. Original vm1480 image and wavelet-decomposed image.

Table 2. Comparison Between AFLC-VQ, JPEG, EZW (vm1480.raw)

Compression Scheme	CR	MSE	PSNR
AFLC-VQ	42.56	23.65	34.39
	74.14	31.00	33.22
	132.13	49.71	31.27
JPEG	42	14.36	36.56
	74	25.92	33.99
	99	343.09	22.79
EZW	40	32.06	33.07
	70	42.97	31.80
	130	55.52	30.69

generate multiresolution codebooks for wavelet-decomposed subimages. Once the parameter τ is established, the optimization process involves updating of the cluster centroids, initially formed by the modified self-organizing neural network architecture, by using the well-known FCM equations.

D. FCM algorithm. In nonfuzzy or crisp clustering, one pattern is assigned to exactly one cluster. On the contrary, fuzzy clustering provides partitioning results with additional information supplied by the cluster membership values of the data samples indicating different degrees of belongingness. Usually an optimization criterion involving an *objective function* that acts as a performance index of clustering is used.

A general form of the objective function is:

$$J(u_{ij}, v_k) = \sum_{i=1}^c \sum_{j=1}^n \sum_{k=1}^c g[w(x_i), u_{ij}] d(x_j, v_k) \quad (1)$$

where $w(x_i)$ is the a priori weight for each x_i and $d(x_j, v_k)$ is the degree of dissimilarity between the data x_j and the supplementary element v_k , which can be considered the central vector of the k th cluster.

One of the widely used clustering methods based on equation 1 is the FCM algorithm developed by Bezdek.^{16,17} This objective function of the FCM algorithm takes the form of

$$J(u_{ij}, v_k) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m \|x_j - v_k\|^2, \quad m > 1, \quad (2)$$

where m is called the exponential weight that influences the degree of fuzziness of the membership (partition) matrix

resulting in the following nonlinear equations for the centroids and the membership function:

$$v_i = \frac{1}{\sum_{j=1}^n (u_{ij})^m} \sum_{j=1}^n (u_{ij})^m x_j, \quad i = 1, 2, \dots, c \quad (3)$$

$$u_{ij} = \frac{(1/\|x_j - v_i\|^2)^{1/m-1}}{\sum_{k=1}^c (1/\|x_j - v_k\|^2)^{1/m-1}}, \quad (4)$$

$i = 1, 2, \dots, c; \quad j = 1, 2, \dots, n$

The system described by equations 3 and 4 cannot be solved analytically. However, the FCM algorithm provides an iterative approach to approximating the minimum of the objective function starting from a given position.

E. AFLC. AFLC is a hybrid neurofuzzy system that can be used to learn cluster structure embedded in complex data sets, in a self-organizing, stable manner. A choice of Euclidean metric is made in developing the AFLC system while keeping a simple control structure adapted from ART-1. The AFLC algorithm initially starts with the number of clusters, c , set to zero. The system is initialized with the input of the first feature vector X_j . As is in leader clustering, this first input is said to be the prototype of the first cluster. The next normalized input feature vector is then applied to the bottom-up weights in a simple competitive learning scheme, or dot product. The node that receives the largest input activation Y_i is chosen as the prototype vector. As in the original ART-1:

$$Y_i = \max_j \sum_{k=1}^p X_{jk} b_{ki}, \quad 1 \leq j \leq N \quad (5)$$

After a node is selected as a winner, the comparison between the cluster prototype associated with Y_i (the centroid v_i) and the current input vector X_j is performed as follows:

$$R = \frac{\|X_j - v_i\|}{\frac{1}{N_i} \sum_{k=1}^{N_i} \|X_k - v_i\|} < \tau \quad (6)$$

where N_i is the number of samples in class i .

If the ratio, R , is less than a user-specified threshold, τ , then the input is found to belong to the cluster originally activated by

Table 3. Vector Quantization for Cervical Spine Image (spine1024.raw, 1,024 × 1,024)

EX	CR	Level 4						Level 3					
		V		H		D		V		H		D	
		τ	C	τ	C	τ	C	τ	C	τ	C	τ	C
1	43.81	1.0	76	1.0	81	1.0	46	1.0	141	1.0	149	1.0	69
2	58.94	1.0	76	1.0	81	1.0	46	1.0	141	1.1	136	1.2	52
3	99.90	1.0	76	1.0	81	1.0	46	1.0	141	1.1	136	1.2	52
EX	CR	Level 2						Level 1					
		V		H		D		V		H		D	
		τ	C	τ	C	τ	C	τ	C	τ	C	τ	C
1	43.81	1.1	169	1.0	195	1.0	114	1.5	148	1.2	404	1.1	501
2	58.94	1.2	150	1.0	195	NA	NA	1.5	148	1.3	178	NA	NA
3	99.90	1.2	150	1.0	195	1.0	114	NA	NA	NA	NA	NA	NA

the simple competition. The choice of the value of τ is determined empirically after a number of initial runs.

When the input is classified as belonging to an existing cluster, it is necessary to update the prototype (centroid) and the bottom-up weights associated with that cluster. First, the degree of membership (μ_{ij}) of X_j to the winning cluster (v_i) is calculated using equation 4.

RESULTS

For large-size images, parameters for controlling the compression ratios in an efficient manner by vector quantization is a difficult task. In the current design, a four-level wavelet decomposition is followed by quantization with 8 bits for vector indices

in the first level and 9 bits for the other levels for the lowest compression used. To achieve higher compression, a combination of quantization of the higher two levels and selective retention of subimages of the first two levels are used.

The decoder has the identical codebook for regenerating all the vectors and an inverse wavelet transform module to reconstruct the original image. The fidelity of the reconstructed image is evaluated by computing the MSE, PSNR.

For color images, first, the image is traditionally split into the RGB three-color planes. The process

Table 4. Vector Quantization for Visual Human Image (vm1480, 2,048 × 1,024, red plane)

EX	Level 4						Level 3					
	V		H		D		V		H		D	
	τ	C	τ	C	τ	C	τ	C	τ	C	τ	C
1	1.0	326	1.0	304	1.0	242	1.7	501	1.5	483	1.2	499
2	1.0	326	1.0	304	1.0	242	1.7	501	1.5	483	1.2	499
3	1.6	251	1.2	234	1	242	1.85	221	1.78	216	1.45	201
EX	Level 2						Level 1					
	V		H		D		V		H		D	
	τ	C	τ	C	τ	C	τ	C	τ	C	τ	C
1	1.9	490	1.9	299	1.47	469	1.75	381	1.7	385	1.36	456
2	N/A	N/A	1.9	299	N/A	N/A	N/A	N/A	1.7	385	N/A	N/A
3	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
EX	Level 4						Level 3					
	V		H		D		V		H		D	
	τ	C	τ	C	τ	C	τ	C	τ	C	τ	C
1	1.0	339	1.0	326	1.0	253	1.64	479	1.68	315	1.35	470
2	1.0	339	1.0	326	1.0	253	1.64	479	1.68	315	1.35	470
3	1.7	240	1.25	254	1.2	217	2.1	235	1.7	256	1.6	253
EX	Level 2						Level 1					
	V		H		D		V		H		D	
	τ	C	τ	C	τ	C	τ	C	τ	C	τ	C
1	2.15	498	2.0	345	1.6	332	1.8	364	2.0	286	1.4	432
2	N/A	N/A	2.0	345	N/A	N/A	N/A	N/A	2.0	286	N/A	N/A
3	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
EX	Level 4						Level 3					
	V		H		D		V		H		D	
	τ	C	τ	C	τ	C	τ	C	τ	C	τ	C
1	1.0	306	1.0	288	1.0	208	1.55	475	1.45	499	1.2	458
2	1.0	306	1.0	288	1.0	208	1.55	475	1.45	499	1.2	458
3	1.4	244	1.15	245	1.0	208	2.0	245	1.9	92	1.45	218
EX	Level 2						Level 1					
	V		H		D		V		H		D	
	τ	C	τ	C	τ	C	τ	C	τ	C	τ	C
1	1.93	493	1.90	435	1.53	509	1.8	499	1.8	456	1.7	75
2	N/A	N/A	1.90	435	N/A	N/A	N/A	N/A	1.8	456	N/A	N/A
3	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Abbreviation: N/A, not applicable.

is then repeated for each color plane and a code-book is generated for each color plane. At the decoder, the reconstructed color planes are synthesized to form the original image.

We used DAUB12 for wavelet decomposition based on the statistical analysis of the wavelet coefficients in the subimages. This allowed us to vary the control parameters in the clustering algorithm smoothly to achieve a desired bit rate.

Figure 2 shows the original and reconstructed images of a typical cervical spine image at compression ratio 99.90:1 using AFLC-VQ and at similar compression ratios (100:1) by EZW. In the EZW algorithm, the target bit rate can be exactly achieved but in AFLC-VQ, the achievable bit rates are approximate values of the target rates set. The JPEG compressed and reconstructed image at 99.1 with obvious distortion is also shown for comparison

in Fig 2. Table 1 shows the MSE and PSNR values for different compression ratios by using AFLC-VQ, JPEG, and EZW for a 8-bit cervical spine image.

Figure 3 shows the original vm1480 image (one slice of the male Visible Human data set) and its wavelet-decomposed image.

Table 2 shows comparative MSE and PSNR values for the above image for different compression ratios by the algorithms used. EZW-reconstructed images result in much higher errors as reflected by the MSE values.

Table 3 shows the selected τ values and the corresponding clusters at each level of subimages of the cervical spine image using AFLC-VQ.

Table 4 shows the selected τ values and the corresponding clusters at each level of subimages of the vm1480 image using AFLC-VQ for three-

Fig 4. Comparison of reconstructed image from AFLC-VQ, JPEG, and EZW compression on visible human image (vm1480.raw, color, 2,048 × 1,024).

color planes using the size $2,048 \times 1,024$. The original size is $2,048 \times 1,216$. The reduction in size is done to make each side as a multiple of two as required by the algorithm; however, this reduction has been done carefully so as not to exclude any part of the relevant image.

Figure 4 shows sections of vm1480 after reconstruction from the image compressed at 42:1, 74:1, and 132:1 with some visible distortions at the highest compression ratios for all three algorithms, but the distortion introduced by JPEG at 99:1 is the worst.

CONCLUSION

A well-designed vector quantization algorithm with proper choice of wavelet and the wavelet filter lengths based on the probability distribution of the coefficients results in an encoding scheme with very low bit rate output at minimum distortion.

REFERENCES

1. Goldberg MA, et al: Application of wavelet compression to digitized radiographs. *AJR* 163:463-468, 1994
2. Antonini M, Barlaud M, Mathieu P, et al: Image coding using wavelet transform *IEEE Transact* 1:205-220, 1992
3. Mallet SG: A theory of multiresolution signal decomposition: The wavelet representation. *IEEE Transact* 11:674-693, 1989
4. Gersho A, Gray RM: *Vector Quantization and Signal Compression*. Boston, MA, Kluwer Academic, 1995
5. Linde Y, Buzo A, Gray RM: An algorithm for vector quantizer design. *IEEE Transact* 28:84-95, 1980
6. Newton SC, Pemmaraju S, Mitra S: Adaptive fuzzy leader clustering of complex datasets in pattern recognition. *IEEE Transact* 3:784-800, 1992
7. Kim YS, Mitra S: An adaptive integrated fuzzy clustering model for pattern recognition. *Fuzzy Sets Syst* 65:297-310, 1994
8. Mitra S, Pemmaraju S: Adaptive vector quantization using an ART-based neuro-fuzzy clustering algorithm. Presented at the International Conference on Neural Networks, Washington, DC June 3-6, 1996
9. Mitra S, Pemmaraju S: Optimized vector quantization with fuzzy distortion measure. *SPIE 10th International Symposium on Aerospace Sensing*, Orlando, FL, April 8-12, 1996
10. Mitra S, Pemmaraju S: Efficient color image compression using integrated fuzzy neural networks for vector quantization. Presented at the SMC Symposium, Orlando, FL, October 12-17, 1997
11. Mitra S, Wilson M, Kompella S: Performance of multiresolution pattern classifiers in medical image encoding from wavelet coefficient distributions. *SPIE Proceedings* 3328 on Medical Imaging, San Diego, CA February 1998, pp 256-263
12. Zamora G, Mitra S: Lossless coding of color images using color space transformations. *11th IEEE Symposium on Computer Based Medical Systems*, 1998, pp 13-18
13. Shapiro JM: Embedded image coding using zerotrees of wavelet coefficients. *IEEE Transact* 41:3445-3462, 1993
14. Pennebaker WB, Mitchell JL: *JPEG Still Image Data Compression Standard*. New York, NY, Reinhold, 1993
15. Carpenter GA, Grossberg S: A massively parallel architecture for a self-organizing neural pattern recognition machine. *Comput Vision Graphics Image Process* 37:54-115, 1987
16. Bezdek J: *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York, NY, Plenum, 1981
17. Mitra S, Castellanos R: Comparative evaluation of pattern recognition algorithms: Statistical, neural, fuzzy and neuro-fuzzy techniques. Presented at *SPIE Symposium Proceedings* 3455, San Diego, CA, July, 1998
18. Kohonen T: *Self-Organizing and Associative Memory*. New York, NY, Springer, 1988
19. Shannon CE: A mathematical theory of communications. *Bell Syst Tech J* 27:379-423, 1948
20. Special Issue on Vector Quantization. *IEEE Transact* vol 5, no 2, 1996
21. The Second Visible Human Conference. National Institutes of Health, Bethesda, MD, October 1-2, 1998