Radiologist Evaluation of a Multispectral Image Compression Algorithm for Magnetic Resonance Images

Patrick T. Cahill, Thomas Vullo, Jian-Hong Hu, Yao Wang, Michael D.F. Deck, Rene Manzo,
Karen Weingarten, and John A. Markisz

With the advent of teleradiology and picture archiving and communication systems (PACS), the expense and time required for image transmission and long term image archiving become important. The use of validated image compression algorithms can greatly reduce these costs. A lossy, multispectral image compression scheme at compression ratios (CR) of 25:1 and 32:1 was used for a set of 26 different patient MR exams. The original and compressed/decompressed (CD) image sets were evaluated in a blinded fashion by four radiologists in two phases. The main objective was to determine whether radiologic interpretation would vary between the two types of CD image sets and the corresponding originals. In general, the compression algorithm caused a slight decrease in image quality; however, the interpretation of pathology did not change between the original and CD image sets. In only one case at the maximum CR = 32 did one of four radiologists change the interpretation of pathology after CD. In this study, lossy multispectral image compression of MR images at CR = 25 maintained diagnostic integrity. This could play a significant role in image storage and communications.

Copyright © 1998 by W.B. Saunders Company

KEY WORDS: Image compression, magnetic resonance imaging, telemedicine, image quality

WITH THE EXPANDING use of telemedicine, teleradiology, and PACS,¹ the cost and time for communications and image archiving become increasingly important. An average digital radiologic examination contains 15 to 20 megabytes and requires at least 69 minutes of transmission time using standard phone lines operating at 28.8 Kbytes/second. T1 lines, at a minimum cost of \$500 per month, would reduce transmission time to approximately two minutes. Digital mammography, having higher resolution and over 100 Mbytes per exam, would require even longer transmission times.

From the Departments of Electrical Engineering and Physics, Polytechnic University, Brooklyn, NY; and the Department of Radiology, The New York Hospital—Cornell University Medical College, New York, NY.

This research was supported in part by a biomedical engineering research grant from the Whitaker Foundation.

Address reprint requests to Patrick T. Cahill, PhD, 530 East 72nd Street, New York, N.Y. 10021.

*Copyright © 1998 by W.B. Saunders Company
0897-1889/98/1103-0004\$8.00/0*

A 1000 bed hospital may require an image archive of four terabytes per year. In any case, an important consideration is how will archival and transmission costs be paid, especially in the face of projections of decreased insurance reimbursement. This will require increased efficiency and throughput.

Image compression using at least a 20:1 ratio would make a significant impact in reducing the size and cost of telemedicine and image archival systems. Numerous image compression algorithms have been developed to help address these problems.^{2,3} Although many compression schemes use wavelet transforms,⁴ MR images, which are acquired with numerous pulse sequences such as gradient echo, and T1, T2 and proton weighted acquisitions, are more suitable for the use of multispectral image compression methods.⁵

Lossy image compression, which is characteristic of algorithms having high CRs, can add noise and occasionally image artifacts. Because of these considerations, the main concern is whether radiologic interpretation of normalcy or pathology would differ between the CD image set versus the original non-compressed set.

To evaluate image compression algorithms for diagnostic integrity, several aspects must be considered. If two images, an original and its highly CD version, are compared side by side, differences can be easily identified. When viewed alone, the determination of whether a given patient exam is original, or has undergone CD can be very difficult since even the original images can have significant variations in image quality. Degradation can be due to poor hardware tuning and instabilities, radiofrequency noise interference, and patient related factors such as dental fillings, implants and motion artifacts (respiratory, cardiac, tremor, discomfit, and so on). These effects produce a spectrum of image artifacts, and signal-to-noise ratio and contrast-to-noise ratio reductions.

To evaluate the utility of a lossy compression scheme, a valid study should involve some or all of the following: several radiologists, each of whom performs independent blinded interpretations; include a sufficient number of image sets; have several phases using image sets with a variety of

subtle and obvious pathologies; employ a spectrum of exams having a broad range of image quality; and consider intra- and inter-variability in individual radiologist interpretations.

This study focused on the use and comparative radiologic evaluation of a lossy multispectral image compression method⁵ for MR studies. A total of 26 neuroradiology cases, which included a range of pathologies normally present in most positive brain MR exams, underwent CD at CRs of 25:1 and 32:1. These image sets were interpreted by four radiologists. The ultimate purpose was to determine whether the underlying radiologic interpretation would change when comparing both CD image sets to the originals.

METHODS

Twenty four MRI patient studies containing a range of typical neuropathologies (Table 1) were selected from the neuroradiology wet-reading log book. In addition, two normal studies were included for a total of 26 image sets. From each complete MR exam, which averaged about 140 images, 30 images consisting of ten contiguous T1 (TR/TE = 600/25), and ten each of the corresponding dual echo T2 (TR/TE = 2100/25, 90) weighted transverse slices (slice thickness = 3 mm with 1 mm interslice gaps) were selected which covered the areas of pathology.

Multispectral image compression was performed according to the methods of Hu et al.⁵ Briefly, the 12 bit gray values for the three corresponding images (one T1 weighted and two dual echo T2) obtained at the same pixel location were treated as a "spectral" vector. The vector for each pixel was linearly predicted from the vectors at adjacent previous pixel locations. A third order linear predictor was used which predicted a new pixel from three previous pixels (left, top, and top-left). The

coefficients were then determined using a forward adaptive method.

An excitation vector was then determined through an analysis-by-synthesis procedure. To produce the excitation vector, a candidate codeword was used to excite each linear predictor (known as the excitation codeword), and the one yielding the closest match to each original image value was chosen. This method is an extension of the code-excited linear prediction (CELP) method commonly used for compression of one dimensional speech signals.⁶ To preserve image quality, the synthesized image using the selected codewords and the linear predictors was further compared to the original image, and the error was coded using a vector quantizer. By varying the number of excitation and error codewords, the CR could be adjusted. Two levels of compression were evaluated, with average CRs of 25:1 (CD25) and 32:1 (CD32). These two CRs were selected since 25:1 was initially determined by non-radiologists to demonstrate no difference in image quality, and for 32:1 moderate differences were seen. For CR = 25, the number of codewords for excitation and error compensation codebooks were 1024 and 512, respectively. For CR = 32:1, 256 and 512 codewords were used for the excitation and error compensation codebooks, respectively.

Both Phase I and Phase II interpretations were based on hard copy presentation. The 30 images for each original, CD25 and CD32 image sets were presented on two sheets of 14 × 17 inch transparency film photographed in a matrix of 5 rows by 3 columns. Each row consisted of an individual slice with the corresponding T1 and dual echo T2 images in each of the columns. The sheets of film were coded so as not to identify which type of study they represented. Each original and its two corresponding CD images were photographed at the same window and level.

The original and CD image sets were evaluated in a blinded fashion in two phases by three board certified attendings (two neuroradiologists, one general radiologist) collectively having over 38 years experience with MRI, and a neuroradiology fellow. Each radiologist performed interpretations individually and without knowledge of her/his or another radiologist's previous results. The age and medical history of the patients were not provided.

Phase I consisted of three blinded reading sessions each separated by a minimum of three weeks. For each session, approximately one third of the 26 original image sets were presented with one third each of the CD25 and CD32 image sets, arranged in random order. For this phase, the image sets were shown one at a time and in no case was an original set and any of its corresponding CD sets shown in the same reading session. Thus all 26 original, CD25 and CD32 image sets were interpreted independently by each radiologist after three such sessions. For each Phase I session, the radiologists were asked specific questions (Fig 1) for each study with answers recorded by an impartial observer. The image quality was ranked numerically from zero (excellent) to two (poor). Similarly, the presence of artifacts was rated from zero (none) to two (excessive). The number of slices containing pathology and the pathology type were recorded. Also, confidence of interpretation, ease of detection of pathology and the radiologist's impression of image type, ie, original, CD or can't tell (CT) were recorded.

Phase II involved two blinded reading sessions, each at least

Table 1. Types of Pathology Found in the MRI Studies

Categories	Subcategories
I. Normal	
II. Hemorrhage:	
Parenchymal	A = Acute
Subarachnoid	B = Subacute
Subdural	C = Old/Chronic
III. Mass Effect:	
Focal lesion	
Enhancing mass	
Midline shift	
Caudal shift	
IV. Infarct:	
Cortical	A = Edema
White Matter	B = Chronic
V. Hydrocephalus	
Diffuse White Matter Changes	
VI. Others:	
Congenital/Structural Abnormality	
Vascular Abnormality	
Aneurysm	
Infection	

Radiologist Code _____ Case Number _____

1. Image Quality: 0 = Excellent 1 = Normal 2 = Poor

2. Artifacts: 0 = None 1 = Moderate 2 = Excessive

3. Study Impression: Normal _____ Pathology _____

Pathology Type _____

Number of slices observed _____

Confidence of Interpretation _____

1 = Questionable 2 = Likely 3 = Definite

Detectability of Pathology _____

1 = Difficult 2 = Easy

4. Image Type: Original _____ Compressed _____ Cant Tell _____

Fig 1. Phase I questions asked all four radiologists for each original, CD25 and CD32 image sets.

two weeks apart for the four radiologists. In the first session, each radiologist was presented with the 26 original together with their corresponding CD25 sets, arranged in random order. In the second session, the original and CD32 sets were compared. Each radiologist was once again asked a series of questions (Fig 2) for each study whose answers were recorded by an impartial observer. First, the radiologist was asked to select the image which was felt to be the original. Four sets of parameters were evaluated: image quality for original and CD image set and which was better overall; differences in five features of normal anatomy (white/gray matter contrast, visibility of the internal auditory canal, ventricles, optic nerves, and overall impression); differences in four aspects of pathology (size/extent, texture, boundary and overall appearance); and most importantly, would interpretation of pathology be different for the original compared to its corresponding CD set. Appearance of normal anatomy and pathology features were rated on a scale from zero to three with zero representing no difference between the original and CD set, and three being unacceptable.

For one neuroradiologist, in a separate reading session, the originals were presented together with both CD25 and CD32 image sets in a random fashion. The radiologist ranked the image quality using a score of A (best), B (second best) and C (third best).

RESULTS

In Fig 3, representative original T1 and dual echo T2 images for a typical slice are shown in addition to their corresponding CD25 and CD32 images. A slight increase in image noise is seen in the CD

image sets; however, the anatomy is clearly defined and there is minimal image degradation.

For one of the 26 CD image sets, a noticeable "boxy" artifact was introduced into three of 10 T1 weighted images (Fig 4). This artifact was only slightly apparent on the corresponding dual echo T2 images. This image set was maintained in the study to measure whether the radiologic interpretation would change; however, it was not considered in the data analysis for Phase I and Phase II, leaving a total of 25 image sets.

Consistent with most lossy compression schemes, quantization caused an increase in image noise. This was insignificant in most cases; however, visible degradations (Fig 5) occurred in 3 of 26 image sets.

Phase I

Phase I results for the four radiologists are summarized in Table 2 and Figs 6 and 7. Image quality of the CD sets was consistently rated somewhat poorer than the originals. Three of four radiologists ranked the original set as having the same degree of artifacts on average than corresponding CD25 sets. This is due in part to the smoothing effect on anatomic borders of the compression

Radiologist Code _____ Case Number _____

1. Pick the set you consider as having better quality.

Better Image Set Number _____

Rate the Image Quality of both #1 _____ #2 _____

0 = Excellent 1 = Normal 2 = Poor

2. Compare original and CD sets for the following features:

0 = No change 1 = Slightly Different

2 = Noticeably Different 3 = Unacceptable

Normal Image Features:

White / Gray matter contrast _____

Internal Auditory Canal _____

Ventricles _____

Optic Nerves _____

Overall Image Quality _____

Pathology Appearance / Interpretation

Check if Normal Study _____

Size / Extent _____

Texture _____

Boundary _____

Overall Impression _____

Fig 2. Phase II questions asked all four radiologists for each original, CD25 and CD32 image sets.

3. Would you change your interpretation of pathology? _____

algorithm. For CD32 image sets, the presence of artifacts was rated slightly higher and is due to the increase in quantization noise which overwhelmed the smoothing effect seen in CD25 images. The average number of slices in which pathology was observed increased slightly with CD25 and CD32 for radiologists 2 and 4, decreased for CD25 and was effectively the same for CD32 for radiologist 1 and increased for CD25 and decreased for CD32 for radiologist 3.

When deciding whether the image sets were original or CD, neuroradiologist 2 differentiated between the two primarily from the slight increase in background noise in the latter type. The other three based their interpretation solely on neuroanatomy and did not consider the background appearance. None of these three radiologists identi-

fied accurately greater than 50% of the CD images sets (Fig 6).

All four radiologists had a high confidence of interpretation. With 3.0 by definition being definite, the average score for all CD sets was 2.90, and for the originals, 2.96. This is consistent with the radiologist's opinion that pathology was easy to detect. In this case, with 2.0 by definition representing a ranking of easy, the CD average ranking was 1.98 and the original average was 2.0.

The two normal studies were detected by all four radiologists in the original and CD25 images. One radiologist interpreted the CD32 of a normal case as having pathology. Another radiologist interpreted the original and both CD sets of two additional studies as having normal age-related changes. For these latter studies, one had diffuse

Fig 3. Comparative T1 and dual echo T2 images of the original (top), CD25 (middle) and CD32 (bottom) sets for a single slice of a representative case.

white matter changes and the second had a small focal lesion in a single slice. Since the patient age or medical history were not provided, it was felt that these two cases could represent normal age

related changes. For the radiology fellow, two studies were identified as normal for the original and CD25 sets, and as having pathology on the CD32 set. In addition, pathology was diagnosed in

Fig 4. T1 image of the CD32 set showing the "boxy" artifact.

two of the original and CD32 sets, but not on the CD25 set. All these cases were interpreted as having pathology by the two neuroradiology attendings for both the original and CD image sets. The "boxy" CD artifact seen in a single study (Fig 4) was identified by all four radiologists; however, only one changed the interpretation of pathology when comparing the original with the CD sets in Phase II.

Table 2. Phase I Results Including the Average Rankings of Image Quality (0 = Excellent, 1 = Normal, 2 = Poor) and Presence of Artifacts (0 = None, 1 = Moderate, 2 = Excessive), and the Sum of the Number of Slices Where Pathology was Identified for 25 Original (ORIG), CD25 and CD32 Sets

Radiologist Number	Image Type	Image Quality	Presence of Artifacts	# Slices Pathology
1	ORIG	0.76	0.40	135
	CD25	0.84	0.40	125
	CD32	1.28	0.60	138
2	ORIG	0.76	0.16	128
	CD25	1.08	0.36	132
	CD32	1.16	0.40	141
3	ORIG	0.24	0.12	108
	CD25	0.64	0.12	116
	CD32	0.72	0.16	99
4	ORIG	0.52	0.24	113
	CD25	1.08	0.24	125
	CD32	1.12	0.28	129
Average	ORIG	0.57	0.23	121
	CD25	0.91	0.28	125
	CD32	1.07	0.36	127

In Fig 7 are shown the percent of studies where either four, three or two radiologists agreed on their determination of normalcy or type of pathology. Although the percentages for the original, CD25 and CD32 image sets were consistent, all radiologists agreed in their interpretations in less than 50% of cases.

In Tables 3 and 4, the intra- and interobserver kappa values are presented for the four radiologists in the interpretation of pathology type. For a kappa value of 0, any agreement is based upon pure

Fig 5. T1 images of the original (left), CD25 (middle) and CD32 (right) sets showing the effects of salt and pepper noise particularly noticeable here in the lateral ventricles.

Fig 6. Percent of studies which were interpreted by four radiologists as the original, can't tell or compressed in Phase I.

chance whereas a kappa of 1 represents true agreement without any element of chance. The intraobserver variation when the original set was compared to their CD25 and CD32 sets, and the CD25 set compared to CD32, had an average kappa of 0.69 for the four radiologists. However, if radiologist 1 was excluded, the average intraobserver kappa was 0.77 for the determination of pathology type. The interobserver kappa value for characterizing the same type of pathology was 0.58 when averaged over all radiologists, and 0.68 for radiologists 2, 3 and 4.

Phase II

In this phase, the original and corresponding CD image sets were presented side by side in random

order. The rankings for image quality and the average differences between original and CD types are summarized in Table 5 and Fig 8. When shown side by side, all radiologists ranked the original set as being better than the corresponding CD image set, except in one case where the CD25 was considered better by radiologist 3. The average differences in image quality ranking between the original and the CD25 and CD32 image sets were 0.71 and 0.77, respectively, where 1.0 by definition represents slightly different. If the results of radiologist 1 are not considered, the average differences in image quality between the original and CD25 and CD32 were 0.62 and 0.63, respectively. For the 25 exams (Fig 8), the four radiologists rated on average 34% and 33% of the CD25 and CD32

Fig 7. Percent of studies where either four, three or two radiologists agreed on their determination of normalcy or type of pathology in Phase I for the original, CD25 and CD32 image sets.

Table 3. Intra-observer Kappa Values for Identifying Pathology Types by Four Radiologists (R) During Phase I

	R 1	R 2	R 3	R 4	Average
ORIG/CD25	0.27	0.89	0.80	0.61	0.64
ORIG/CD32	0.51	1.00	0.75	0.64	0.73
CD25/CD32	0.56	0.90	0.85	0.51	0.71
Average	0.45	0.93	0.80	0.59	0.69

images, respectively, as having no change from the original. In addition, 61% and 58% were rated as having slight differences (one unit), while 5% and 9% had noticeable differences (two units).

Radiologist 1 compared the originals simultaneously with both CD versions in a separate reading session. The results are summarized in Fig 9. The original sets were ranked best for 84% of the cases, CD25 was ranked second best for 96% of the cases, and CD32 was ranked third best for 84% of the image sets.

In Tables 6 and 7, comparisons between original and CD image sets for normal anatomy and pathology features, respectively, are given for each radiologist where a lower score represents a smaller difference. For the CD25 image sets, the average change in the overall impression of normal features compared to the original for four radiologists was 0.81, where 1.0 represents an impression of slightly different. For the CD32 images, the average change was 0.97. For ranking of pathology, the overall changes between the original, and CD25 and CD32 image sets were 0.43 and 0.57, respectively, where 1.0 represents an impression of slightly different.

The final and most important question asked the four radiologists was whether their interpretation of pathology would change after CD. For the CD25 sets, the four radiologists unanimously said NO for all 25 cases. For the one case with CD "boxy" artifacts (Fig 4), the CD32 set was scored by all radiologists as having noticeably different texture and boundary features; however, only two radiologist would change their interpretation. For one of the three studies with quantization noise artifacts (Fig 5), only radiologist 1 would change the interpretation of pathology for CD32. Here, the

Table 4. Interobserver Kappa Values for Identifying Pathology Types by Four Radiologists (R) from Phase I

	R 1/2	R 1/3	R 1/4	R 2/3	R 2/4	R 3/4	Average
Original	0.47	0.43	0.43	0.75	0.70	0.59	0.56
CD25	0.56	0.41	0.44	0.75	0.76	0.66	0.60
CD32	0.65	0.40	0.55	0.65	0.75	0.59	0.60
Average	0.56	0.41	0.47	0.72	0.74	0.61	0.59

Table 5. Phase II Comparison by Four Radiologists of Image Quality Ratings (0 = Excellent; 1 = Normal; 2 = Poor) for Original (ORIG) and CD Image Sets, and the Average Image Quality Rating Difference (AVG DIF) Between Original and CD Image Sets

Radiologist	Image Types	% Original Rated Best	Avg. Quality		Avg. Dif.
			ORIG	CD	
1	CD25	100	0.12	1.12	1.00
	CD32	100	0.25	1.46	1.21
2	CD25	100	0.24	1.04	0.80
	CD32	100	0.12	1.08	0.96
3	CD25	96	0.44	1.00	0.56
	CD32	100	0.48	1.04	0.56
4	CD25	100	0.80	1.24	0.44
	CD32	100	0.88	1.24	0.36
Average	CD25	99	0.40	1.10	0.70
	CD32	100	0.43	1.20	0.77

right lateral ventricle appeared to have some degree of intraventricular hemorrhage (Fig 5).

DISCUSSION

This study involved a radiologic evaluation by four radiologists of a multispectral image compression algorithm in two phases. Phase I included three blinded sessions which simulated the typical radiology reading room conditions and compared image quality and interpretation of pathology for the original and CD image sets. Intra- and interobserver variability in radiologist's interpretations was also assessed. In Phase II, where originals were compared directly against CD25 and CD32 image sets in a random arrangement, the ultimate goal was to determine whether the radiologist's overall interpretation would change for CD compared to the original.

For Phase I, significant variability between radiologists was observed. For example (Table 2), for the original image sets there was an average variation between radiologists from 0.24 to 0.76 relative units (>200%) in rating of image quality, a 0.12 to 0.40 variation (>200%) in the assessment of artifacts, and a 25% variation in the number of slices (108 to 135 for the original sets) in which pathology was identified. All radiologists were unanimous that identification of pathology was easy in the original and both CD image sets. The radiologists were self consistent in their interpretations, however, and well within the norms of variance.⁷

In the determination of pathology, interobserver agreement for at least three radiologists for the three Phase I sessions was approximately 90%.

Fig 8. Percent of cases in Phase II where the ranking of image quality for original and CD sets, interpreted in pairs by four radiologists, differed by 0, 1 or 2 units.

Typical variability has been reported to be approximately 20% in spine MR exams.⁷ Part of the variability here resulted from the radiologist's lack of knowledge of patient age and medical history. For example, hydrocephalus or small focal infarcts were determined in some cases but considered age-related changes in others.

For Phase II, interobserver variability in the interpretation of specific features was also present. For example when rating image quality for the selected original vs. the CD25 image sets (Table 5), the four radiologists reported no differences in 3, 5, 11 and 15 of the 25 studies, respectively. Rankings of five descriptors of normal features (Table 6) and the four descriptors of pathology (Table 7), varied significantly between radiologists for both CRs.

For example, the average rating for the appearance of white/gray matter ranged on average from 0.28 to 1.00 for CD25. Also, the average difference in overall impression of pathology between original and CD image sets varied between 0.00 and 0.91 for CD25 and CD32, respectively.

In some cases, a smaller difference between the original and CD32 image sets were reported compared to the original and CD25 set. Variations between the original and CD image sets for the same radiologist, and between different radiologists in part reflects variations in their interpretations from one reading session to another. This is particularly true since radiologists are not normally asked to quantify image characteristics.

Regardless of the variability in Phase II interpre-

Fig 9. Interpretation for whether the original, CD25 or CD32 image sets were considered best, second best or third best in image quality when viewed together by one radiologist in Phase II.

Table 6. Phase II Ranking of Differences Between Original and CD Image Sets (0 = No Change; 1 = Slightly Different; 2 = Noticeably Different; 3 = Unacceptable) in Normal Features for White/Gray Matter Contrast (W/G), Internal Auditory Canal (IAC), Ventricles (VEN), Optic Nerve (OPT) and Overall Impression (OVL)

Radiologist	Type	W/G	IAC	VEN	OPT	OVL
1	CD25	0.48	0.00	0.00	0.00	0.60
	CD32	0.33	0.00	0.00	0.00	0.92
2	CD25	0.88	0.00	0.04	0.00	1.04
	CD32	0.16	0.00	0.08	0.09	1.12
3	CD25	0.28	0.22	0.56	0.43	0.68
	CD32	0.28	0.23	0.68	0.35	0.88
4	CD25	1.00	0.33	0.48	0.91	0.92
	CD32	0.96	0.78	0.80	0.91	0.96
Average	CD25	0.66	0.14	0.27	0.34	0.81
	CD32	0.43	0.25	0.39	0.34	0.97

tations of specific features for both normal and pathology, identification of those exams which were normal, and most importantly, whether interpretation of pathology would change, demonstrated nearly perfect agreement. Three radiologists did not change their interpretation in all cases. For radiologist 1, interpretation of pathology changed in one of 25 cases at CD32 (Fig 5). The other three commented in this latter case that there were changes in the textural characteristics of the pathology after CD compared to the original, but their interpretation of pathology did not change. Once again, some variability in radiologic interpretations is expected.⁷

In the operation of this lossy multispectral compression algorithm, one study was observed as having an unacceptable "boxy" artifact on three of the T1 weighted images (Fig 4). This was due to the insufficient number of codewords used for representing the excitation vectors and errors. This case was

Table 7. Phase II Ranking of Differences Between Original and CD Image Sets (0 = No Change; 1 = Slightly Different; 2 = Noticeably Different; 3 = Unacceptable) in Pathology Features for Size/Extent (SIZ), Texture (TEX), Boundary (BOU) and Overall Impression (OVL), and the Number of Cases Where a Change of Interpretation (NCI) was Reported

Radiologist	Type	SIZ	TEX	BOU	OVL	NCI
1	CD25	0.09	0.26	0.22	0.13	0
	CD32	0.23	0.36	0.32	0.48	1
2	CD25	0.04	0.65	0.09	0.74	0
	CD32	0.09	0.18	0.29	0.91	0
3	CD25	0.00	0.10	0.00	0.00	0
	CD32	0.00	0.05	0.00	0.00	0
4	CD25	0.61	0.61	0.78	0.87	0
	CD32	0.87	0.87	0.87	0.87	0
Average	CD25	0.19	0.41	0.27	0.43	0
	CD32	0.30	0.37	0.37	0.57	0.25

included in the reading sessions in order to observe whether these artifacts would change the radiologist's interpretations, but was excluded in the tabulated results. Salt and pepper noise, which is typical of lossy compression algorithms, was also present in some images (Fig 5). This noise is due to quantization errors and is most noticeable in homogeneous regions of an image.

The complete set of MR images for each pulse sequence, the patient's age, and medical history were not supplied to the radiologists. To perform CD of all patient images, which averaged 140 per study, and to have them interpreted by four radiologists, would have been overly expensive and time consuming. The lack of age and medical history contributed in part to the differences in Phase I interpretations in three cases where diagnosis was considered that of normal age related changes. Nevertheless, the final interpretations of four radiologists were remarkably self consistent, where only radiologist 1 reported a difference in interpretation of pathology for one case. Thus, the radiologic interpretation between the original and CD image sets varied in only one out of 200 total cases (25 image sets $\times$ 4 radiologists $\times$ 2 CRs) or 0.5%. An ROC analysis was not performed for the following reasons. First, as has been noted by Cosman et al,⁸ ROC analysis can only evaluate the effect of CD on the binary diagnostic decision (normal or with pathology). On the other hand, it is desirable to evaluate the effects of CD on various characteristics including contrast, texture and shape of normal and pathologic features, which is a multivariable problem. Secondly, the majority of the test images used for this study contained obvious pathology which would render an ROC analysis difficult.

There is a need to develop a set of MR exams that are difficult to interpret and which should be examined by radiologists in many imaging centers for continued evaluation of CD algorithms. Similar studies need to be performed for computed tomography and digital mammography. In addition, the features used to evaluate changes in image appearance (normal and pathology) need to be carefully examined. Image compression techniques will become extremely useful and cost effective for both the transmission and storage of medical images once sufficient comparative studies of original to CD exams covering the full spectrum of radiologic imaging modalities have been performed and appropriate CRs are determined.

REFERENCES

1. Cahill PT, Vullo T: Picture archiving and communications systems: Status, problems, and needs. *Information Display* 12:36-40, 1996
2. Aberle D, Gleeson F, Sayre JW, et al: The effect of irreversible image compression on diagnostic accuracy in thoracic imaging. *Invest Radiol* 28:398-403, 1993
3. Wong S, Zaremba L, Gooden D, et al: Radiologic Image Compression: A Review. *Proc IEEE* 83:194-219, 1995
4. Goldberg MA, Pivovarov M, Mayo-Smith WW, et al: Applications of wavelet compression to digitized radiographs. *AJR Am J Roentgenol* 163:463-468, 1994
5. Hu JH, Wang Y, Cahill PT: Multispectral code excited linear prediction coding and its applications in magnetic resonance images. *IEEE Trans Image Processing* 6:1555-1566, 1997
6. Kleijn WB, Paliwal KK (eds): *Speech Coding and Synthesis*. New York, NY, Elsevier, 1995
7. Jensen MC, Brant-Zawadzki MN, Obuchowski N, et al: Magnetic resonance imaging of the lumbar spine in people without back pain. *N Engl J Med* 331:69-73, 1994
8. Cosman PC, Tseng C, Gray RM, et al: Tree-structured vector quantization of CT, chest scans: Image quality and diagnostic accuracy. *IEEE Trans Medical Imaging* 12:703-709, 1993