Journal of Digital Imaging

VOL 10, NO 3 AUGUST 1997

Evaluation of Irreversible Compression of Digitized Posterior-Anterior Chest Radiographs

Bradley J. Erickson, Armando Manduca, Kenneth R. Persons, Frank Earnest IV, Thomas E. Hartman, Gordon F. Harms, and Larry R. Brown

The purpose of this article is to assess lossy image compression of digitized chest radiographs using radiologist assessment of anatomic structures and numerical measurements of image accuracy. Forty posterior-anterior (PA) chest radiographs were digitized and compressed using an irreversible wavelet technique at 10, 20, 40, and 80:1. These were presented in a blinded fashion with an uncompressed image for A-B comparison of 11 anatomic structures as well as overall quality assessments. Mean error, root-mean square (RMS) error, maximum pixel error, and number of pixels within 1% of original value were also computed for compression ratios from 5:1 to 80:1. We found that at low compression (10:1) there was a slight preference for compressed images. There was no significant difference at 20:1 and 40:1. There was a slight preference on some structures for the original compared with 80:1 compressed images. Numerical measures showed high image faithfulness, both in terms of number of pixels that were within 1% of their original value, and by the average error for all pixels. Our findings suggest that lossy compression at 40:1 or more can be used without perceptible loss in the representation of anatomic structures. On this finding, we will do a receiver-operator characteristic (ROC) analysis of nodule detection in lossy compressed images using 40:1 compression.

KEY WORDS: data compression, picture archiving and communications system.

COMPRESSION OF medical images is of great interest because it can significantly reduce the time required to transmit images and the amount of space required to store them. Lossless techniques have the obvious appeal of no image degradation, but achieve only 2:1 to 3:1 reductions for medical images.[1] Lossy techniques reduce the amount of data transmitted to an arbitrarily small amount, depending on the degree of image degradation which is acceptable and the suitability of the compression technique for the type of image to be compressed. Many different techniques have been proposed including the JPEG (Joint Photographic Experts Group) algorithm,[2,3] other discrete cosine transform methods,[4] predictive coding,[5] and wavelet[6,7] methods.

Goldberg et al[6] evaluated wavelet compression and concluded that no clinically relevant degradation was present at compression ratios of up to 30:1. However, their conclusions are of limited value because only six chest radiographs were used (and each with a different pathology), cathode ray tubes (CRT) were used for display, and the results were combined with bone and abdomen films. It has not been shown that a technique appropriate for one type of image is appropriate for another.

This article focuses on the acceptability of lossy compression of digitized high quality PA chest radiograph films. Therefore, compression demands of one type of image were not combined with those of another. Furthermore, a substantially larger number of cases is used, increasing the power of the statistics. The current study also avoids the limited dynamic range of CRTs (which might not unmask more subtle degradation) and minimizes the effect that learning the computer image display software might have.

MATERIALS AND METHODS

Forty upright 14 × 17-in PA chest radiographs without significant pathological findings were digitized using a Lumisys 150 scanner at a resolution of 2,000 × 2,500 × 12 bits per pixel. The resulting uncompressed digital images were printed on a Kodak Model XLP (Kodak, Rochester, NY) laser printer. An intensity-correcting lookup table was created so that the printed

From the Department of Radiology, Mayo Clinic, Rochester, MN.

Address reprint requests to Bradley J. Erickson, MD, PhD, Dept of Radiology, Mayo Clinic, 200 First St SW, Rochester, MN 55905.

0897-1889/97/1003-0001$3.00/0

images matched the optical densities of the original films as determined by a step wedge phantom and densitometer. The maximum optical density of the laser-printed images was 3.2. The film from this uncompressed data served as the "original" in the A-B comparison (the original film was digitized and reprinted onto film before compression was performed). The digital image data were also compressed and decompressed using a lossy wavelet compression algorithm[7] at ratios of 10:1, 20:1, 40:1, and 80:1 and then printed using the same intensity correction. Compression ratios were calculated based on each pixel being stored as a 16-bit quantity. For numerical measurements, compression ratios from 5:1 to 100:1 at increments of 5 were generated.

Pairs of images (uncompressed versus one of the compressed images) were presented to the three raters (board-certified chest radiologists) in a blinded, random fashion using a two-alternative forced choice method. The evaluation consisted of an A-B comparison of the demonstration of 11 anatomical structures (see Table 1 for list of structures), and an overall quality and acceptability assessment; this method was chosen because it is the same method used for evaluating new screen/film combinations for chest radiography.[8] The scoring range for this system ranges from −3 to +3, in which negative scores meant B was significantly (−3), definitely (−2), or slightly (−1) better than A; positive scores meant A was significantly (3), definitely (2), or slightly (1) better than B; or 0 (A was the same as B). The three raters evaluated 80 pairs of images (each rater viewed a randomly selected group of 80 pairings out of the possible 200 pairings), for a total of 240 film comparisons from all three raters. A student's t-test was used for assessing differences among the ratings.

The wavelet compression method used here is described,[7] and is based on calculating the discrete wavelet transform (DWT) of the image to five levels of resolution with the 9-tap/7-tap biorthogonal filters of Antonini et al.[9] At the borders, the image is extended virtually by reflection. The wavelet coefficients are encoded with an embedded, quadtree-based technique termed "set partitioning in hierarchical trees" or SPIHT.[10]

Quantitative measures of image faithfulness were calculated for each image at compression ratios from 5:1 to 100:1 at an increment of 5:1, and included the mean pixel error,

$$\text{Mean Error} = \sum\nolimits_{x,y} |I_{x,y} - C_{x,y}| / \text{NumPixels}$$

RMS error,

$$\text{RMS Error} = \left[\left(\sum\nolimits_{x,y} |I_{x,y} - C_{x,y}|^2\right)\right]^{1/2} \Big/ [\text{NumPixels}]^{1/2}$$

maximum pixel error,

$$\text{Max Error} = |I_{x,y} - C_{x,y}|$$

and the percentage of pixels within 1% of their original value (based on the dynamic range of 0 to 4096).

Compression was performed on an IBM (IBM Corp, White Plains, NY) model 100 computer, with an Intel (Intel Corp, Santa Clara, CA) Pentium processor running at 90Mhz with 32 MB of RAM, using the OS/2 (IBM Corp, White Plains, NY) operating system.

RESULTS

A total of 240 images were compared, each with ratings of the 11 structures plus the overall assessment. Table 1 compares the different compression ratios and their ability to show the various anatomic structures as well as the overall acceptability for diagnostic use. Significantly, no images at any of the compression ratios were subjectively considered unacceptable for diagnostic use. We found no statistically significant difference (at the $P < .05$ level) between compressed and uncompressed images for any of the structures evaluated except for a slight preference for 10:1 compressed images when evaluating the spine ($P = .048$) and retrocardiac lung ($P = .042$), and a slight preference for original

Table 1. Results of Compression Comparison

	10:1		20:1		40:1		80:1	
Structure	Average	*P* Value	Average	*P* Value	Average	*P* Value	Average	*P* Value
T + B	0	0.500	0.016	0.161	−0.016	0.284	−0.024	0.285
PH	0	0.472	0.016	0.161	0	0.324	−0.020	0.162
LPE	0.027	0.160	0	0.500	−0.016	0.161	−0.024	0.162
VBI	0.041	0.130	0	0.500	−0.032	0.079	−0.143	0.006
DAE	0.027	0.079	0	0.448	0	0.285	−0.048	0.080
AER	0.055	0.051	0	0.323	0	0.500	0.024	0.285
Ribs	0.027	0.079	0	0.500	0	0.419	−0.048	0.080
VB	0.068	0.048	0.016	0.161	−0.016	0.161	−0.095	0.052
PVM	−0.027	0.079	−0.016	0.161	−0.032	0.079	−0.024	0.162
RDL	0.014	0.160	0	0.5	0	0.500	0.023	0.160
RCL	0.041	0.042	−0.016	0.161	−0.016	0.284	−0.024	0.162
Overall	0	0.500	0	0.500	0	0.500	−0.001	0.440

A score of 1 represents a slight preference for the compressed image, −1 a slight preference for the original, and 0 means no difference.

Abbreviations: T + B, trachea and bronchi; PH, pulmonary hila; LPE, left paraspinal edge; VBI, vertebral body interspaces; DAE, descending thoracic aortic edge; AER, azygoesophageal recess; VB, vertebral bodies; PVM, pulmonary vasculature and markings; RDL, retrodiaphragmatic lung; RCL, retrocardiac lung; Overall, overall appearance and esthetic quality.

images versus 80:1 when evaluating the vertebral body interspaces ($P < .006$). Approximately 82% of all assessments were 0, indicating that in the majority of cases there was no perceptible difference between compressed and uncompressed, and all other ratings were -1 or 1 indicating only a slight difference. Because all ratings were either 0 or ± 1, a student's t-test was deemed appropriate rather than a rank-based test. Only 1 of the 240 overall assessments indicated any preference, and that was 1 case of 80:1 where a slight preference was noted for the original. Figure 1 shows a region of interest from 1 of the test cases, showing both the original image and an 80:1 compressed image.

Figure 2 shows the results of mathematical measures (mean pixel error, RMS error, maximum error, and percentage of pixels within 1% of the original value). Generally, there was a fairly linear relation between these measures and the compression ratio. An exception to this was the maximum pixel error in an image. This curve had an irregular shape, which is of uncertain significance.

Compression and decompression times were similar for the same image size and compression ratio. Larger images or lower compression ratios resulted in longer times to compress or decompress, and ranged from 47 seconds to 267 seconds.

DISCUSSION

This study represents an early step in determining the amount of compression that can be used for transmitting and storing images used for primary interpretation. The motivation is clear: the cost of transmitting and storing images is inversely proportional to the compression ratio. It could be argued that this study does not show the ability to diagnose subtle pathologies. However, the method is the same as that used for evaluating new screen/film combinations, and evaluating the subtleties of normal anatomic structure is likely a good surrogate for detecting the effects of subtle pathology that might result in similarly subtle disruption of normal structures. And by evaluating many different types of structures (eg, bones as well as lung markings) we believe that it is likely that a range of pathologies should be equally preserved. One of the areas of concern was about loss of fine interstitial markings, but ratings did not reflect this. In fact, of the raters (L.R.B.) is a certified B reader who may be even more sensitive to interstitial patterns, but no difference in his ratings were noted. Nevertheless, this study does not prove that this wavelet compression technique produces no loss of diagnostic information. We aim to test this hypothesis in the next study, which will be an ROC analysis in diagnosing important and subtle pathologies on chest radiographs.

Previous studies have involved limited numbers of radiographs, most with varying pathologies, leading to inconsistency in how the technique was evaluated. By using an established method for evaluating image quality of chest radiographs,[8] we hope to be able to produce reliable, unbiased results. It is also recognized that even if interpretation of compressed images is proven to be diagnostically accurate, most radiologists prefer to evaluate images that are aesthetically pleasing, not simply diagnostically adequate. This study shows that compression at high levels does not have to degrade the appearance of images, and in fact, at lower levels, may make them more pleasing. It may seem surprising that there was a slight preference for 10:1 compressed images over the originals. However, at low compression ratios, the effect of running the compression algorithm is similar to performing wavelet denoising,[11] with the noise threshold dependent on the specific compression ratio. The noise, which tends to be high frequency and uncorrelated, is preferentially discarded. Therefore, wavelet compression increases the signal-to-noise ratio at sufficiently low compression rates (unless the image is noise-free). One might then expect that edge information would also be lost during compression. However, when compared with a simple low-pass filter, wavelet compression uses mixtures of basis functions to represent the image. We used basis functions that represent edges better than salt-and-pepper noise, and so it is not surprising that edges are maintained with this compression technique, while still suppressing noise.

We also recognize that this study uses digitized images, rather than directly acquired digital images. The digitizer almost certainly introduces noise into the digital image, but this is how teleradiology applications are often performed, and so we believe this does represent a useful evaluation.

The evaluation of compression techniques is quite labor intensive. For this type of study, it is necessary to have experienced radiologists review several sets of images several times. It would be

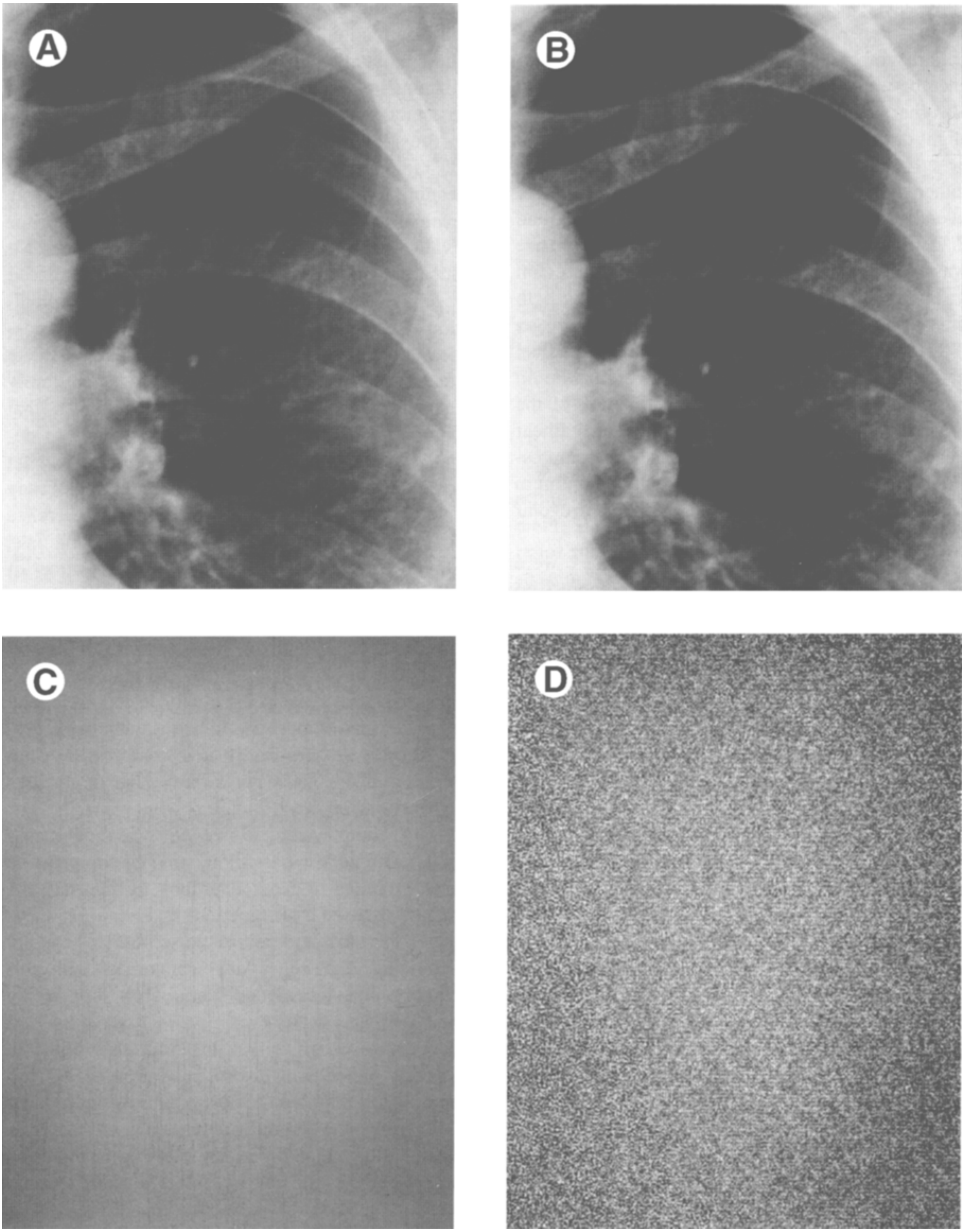

Fig 1. Comparison of a region of interest of the original image (A) and an 80:1 compressed image (B). A small area of sclerosis is noted in a left rib. Subtraction images printed at the same width (C) as (A) and (B), and printed at 1% of the width (D).

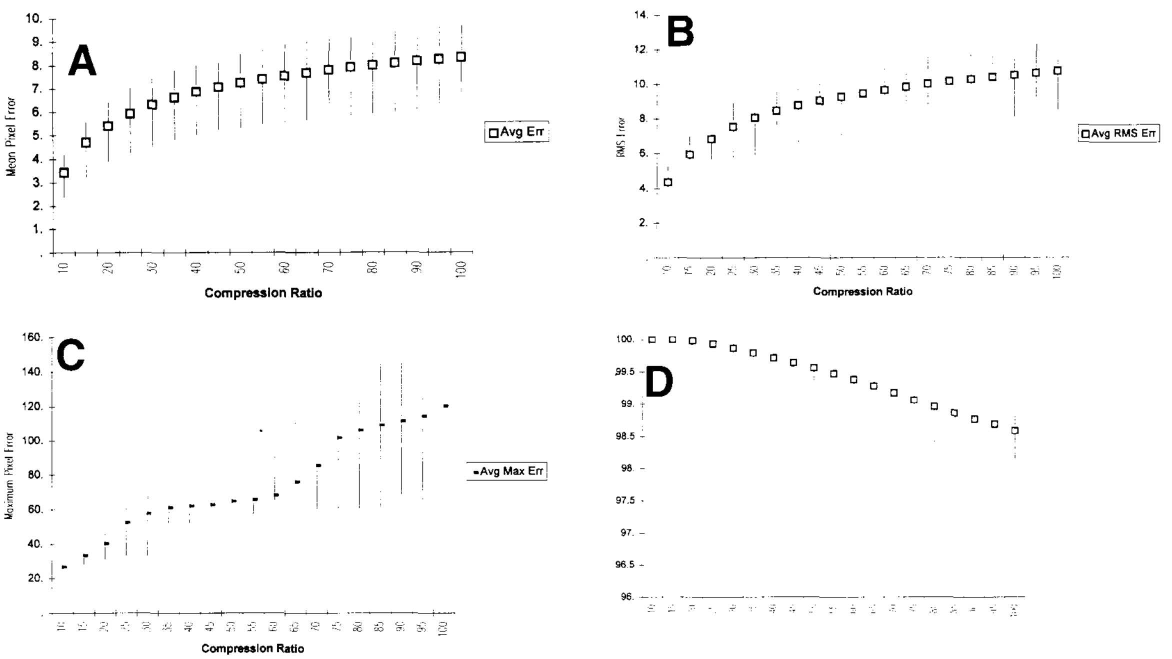

Fig 2. Numerical error measurements of images versus compression ratio. Fig 2A compares the mean pixel errors versus compression ratio showing the average for all 40 CXRs, as well as the best and worst case. Fig 2B compares the mean RMS error versus compression ratio for all cases, as well as the best and worst RMS error. Fig 2C displays the average of the maximum absolute pixel error versus compression ratio as well as the best case and worst case. Fig 2D displays the percentage of pixels that were within 1% of the original value versus compression ratio, as well as the best and worst case.

preferable if some analytical technique could measure diagnostic value and image quality (eg, aesthetic appearance). There are several commonly used measures of image faithfulness (such as RMS error and mean error), but it is difficult to relate these to perception of image degradation. Nevertheless, we believe it is valuable to document the mathematical accuracy of lossy compression techniques, even if they do not constitute the primary outcome measure. We have attempted to quantitate some of the basic measures of image fidelity, and although there were some trends, the correlation was not sufficiently clear that we would be comfortable replacing a radiologist's rating with an analytical measure. There has been some interesting recent work in this area,[12] but a suitable image quality measure does not currently exist.

CONCLUSION

Wavelet compression can be performed on digitized chest radiographs without perceptible loss in the ability to render a wide range of normal anatomic structures at ratios up to 40:1. Even at 80:1, degradation is surprisingly small. There is a slight preference for compressed images at low ratios. This study suggests—but does not prove—that compression ratios up to at least 40:1 may be applied to digitized chest radiographs without loss of diagnostic value. This will be tested in a future study.

REFERENCES

1. Aberle DR, Glesson F, Sayre JW, et al: The effect of irreversible image compression on diagnostic accuracy in thoracic imaging. Invest Radiol 28:398-403, 1993

2. Leger A, Omachi T, Wallace GK: JPEG still picture compression algorithm. Optical Eng 30:947-954, 1991

3. Wallace GK: The JPEG still picture compression standard. Commun Assoc Computing Machinery 34:30-44, 1991

4. Kostas TJ, Sullivan BJ, Ikeda M, et al: Clinical evaluation of irreversible image compression: analysis of chest imaging with computed radiography. Radiology 175:739-743, 1990

5. Mori T, Nakata H: Irreversible data compression in chest imaging using computed radiography: an evaluation. J Thorac Imaging 9:23-30, 1994

6. Goldberg MA, Pivovarov M, Mayo-Smith WW, et al: Application of wavelet compression to digital radiographs. AJR 163:463-468, 1994

7. Manduca A, Said A: Wavelet compression of medical images with set partitioning in hierarchical trees. Medical Imaging 1996: Image Display, Proc SPIE 2707 1996, pp 192-200

8. Swensen SJ, Gray JE, Brown LR, et al: A new asymmetric screen-film combination for conventional chest radiography: evaluation in 50 patients. AJR 160:483-486, 1993

9. Antonini M, Barlaud M, Mathieu P, et al: Image coding using wavelet transform. IEEE Trans Image Proc 1:205-220, 1992

10. Said A, Pearlman WA: A new fast and efficient image codec based on set partitioning in hierarchical trees. IEEE Trans Circuits and Systems for Video Tech 6:243-250, 1996

11. Donoho DI, Johnstone IM: Ideal denoising in an orthonormal basis chosen from a library of bases. Comptes Rendus Acad Sci 319:1317-1322, 1994

12. Lattner S, Good W, Maitz G: Visually weighted assessment of image degradation resulting from image compression. Proceedings of the SPIE, Medical Imaging, 1996, 11:59