Perceived Fidelity of Compressed and Reconstructed Radiological Images: A Preliminary Exploration of Compression, Luminance, and Viewing Distance

Thomas K. Pilgram, Richard M. Slone, Edward Muka, Jerome R. Cox, and G. James Blaine

The authors' goal was to explore the impact of image compression algorithm and ratio, image luminance, and viewing distance on radiologists' perception of reconstructed image fidelity. Five radiologists viewed 16 sets of four hard-copy chest radiographs prepared for secondary interpretation. Each set included one uncompressed, and three compressed and reconstructed images prepared using three different algorithms but the same compression ratio. The sets were prepared using two subjects, four compression ratios (10:1, 20:1, 30:1, 40:1), and two luminance levels (2,400 cd/m², standard lightbox illumination, and 200 cd/m², simulating a typical CRT display). Readers ranked image quality and evaluated obviousness and clinical importance of differences. Viewing distances for image screening, inspection, and comparison were recorded. At 10:1 compression, the compressed and uncompressed images were nearly indistinguishable; the three algorithms were very similar, and differences were rated "not obvious" and "not important." At higher compression, readers consistently preferred uncompressed images, with notable differences between algorithms. The obviousness and clinical importance of differences were rated higher at lightbox luminance. Viewing distances appeared to be idiosyncratic.

Copyright © 1998 by W.B. Saunders Company

KEY WORDS: thorax, radiography, images, compression, fidelity.

AS ELECTRONIC imaging becomes more common in medical practice, the problems of storage and transmission of image data become more important.¹ Image compression techniques are designed to reduce redundancy by appropriate image coding and, as a result, can significantly reduce effective image data size. Before presentation, the coded representation is reconstructed to form an exact or approximate reproduction of the original image. Currently available image compres-

sion techniques can reduce the storage and transmission size of images by more than an order of magnitude, but at the cost of some degradation in image fidelity. If the degradation in image fidelity were small enough to be clinically unimportant, the use of compressed images could dramatically reduce the financial and time costs of electronic storage and transmission.

Several studies have attempted to demonstrate the feasibility of using recently developed compression algorithms for diagnostic imaging, particularly chest radiography,²⁻⁵ which is the most common type of diagnostic imaging.⁶ These studies have shown that the algorithms examined produce mild image degradation, without loss of diagnostic performance at low compression ratios.

Within the context of enterprise-wide distribution of images intended for secondary interpretation, original images, whether screen-film or digitally captured, will likely need to be reduced to pixels of about 1-k by 1-k and converted into an appropriate signal space before transmission on the network. In keeping with this premise, we used a suitably prepared, distributable form of the original image in its uncompressed state rather than the "true original" as the reference image. This scenario is not unlike commonly practiced teleradiology wherein a preliminary reading is made by a radiologist at a distant site using a soft-copy display that cannot render the "true original."

Although studies exist that compare compression algorithms,⁷⁻⁹ we are not aware of any study of the role of reading distance, image luminance, or reduced-image size for the purpose of secondary image review. Herein we report a study designed to explore the roles of the compression algorithm and ratio, image luminance, and viewing distance on the observability of reconstruction artifacts. Previous studies have shown that readers notice subtle differences during image comparisons, even when these findings did not affect diagnostic performance.¹⁰ Based on this principle, we employed rank ordering of simultaneously viewed images,

---

From the Mallinckrodt Institute of Radiology, Washington University School of Medicine, St Louis, MO.

Address reprint requests to Thomas K. Pilgram, PhD, Mallinckrodt Institute of Radiology, Washington University School of Medicine, 510 S Kingshighway Blvd, Box 8131, St Louis, MO 63110.

Copyright © 1998 by W.B. Saunders Company
0897-1889/98/1104-0002\$8.00/0

which is sensitive and does not have the sample-size demands of ROC analysis.¹¹

MATERIALS AND METHODS

Study Design

The broad design of the study was to evaluate readers' image preferences while they viewed multiple versions of the same image side-by-side, thus allowing careful scrutiny of subtle differences. The chest radiograph provides a broad range of structure and tone characteristics, providing the opportunity for the reader to note degradation (artifacts) in areas of uniform opacity, along the interface of differing tissue density and within the intricate trabecular and pulmonary parenchymal detail. The effects of compression algorithm and ratio, image luminance, and reading distance were studied using multiple replications of a single image, but we selected two to allow for the possibility of detecting individual differences. Clinical diagnosis was not an issue; although the patients in the study had pulmonary abnormalities, neither detection nor description was included in the image evaluation.

Images

The cases selected included a 43-year-old woman with a small right pneumothorax and a 15-year-old boy with mediastinal adenopathy and a 14-mm nodule in the right lung base, visible through the right hemidiaphragm. Posteroanterior radiographs were obtained in a dedicated chest radiography room using a medium-resolution screen-film system (Quanta Fast Detail-T, Dupont, Wilmington, DE). They were phototimed at 180 cm using a 10:1, 40-line-per-centimeter fixed grid and 115-kVp technique.

The screen-film images were digitized at high resolution on a research digitizer¹² and reduced, after low-pass filtering, to 1,024 pixels by 1,024 lines to simulate a practical format for secondary interpretation. The 12-bit density data were converted nonlinearly into an 8-bit DAC space that can be directly displayed by low-cost video displays likely to be found in an enterprise-wide, image distribution system (PACS). The images were produced on a laser printer (Kodak Ektascan blue base film, KELP Model 100 laser printer with digital interface, Interpolation method: cubic convolution; Eastman Kodak Co, Rochester, NY). The image size was 17 by 17 cm, to approximate the size of a desktop CRT display.

Sixteen sets of four images were prepared, which included an original and three compressed/reconstructed images. Each compressed/reconstructed image in a set was made using a different compression algorithm but the same compression ratio (10:1, 20:1, 30:1, or 40:1). Two of the algorithms, embedded zero-tree wavelet (EZW) and wavelet scalar quantization (WSQ) were based on the discrete wavelet transform, and the third algorithm was the commonly available block-DCT (Joint Photographic Experts Group (JPEG)¹³⁻¹⁵ (Fig 1). A previous report describes the complete image-processing process.¹⁶

The four images in each set were arranged in random order for viewing by the reader. There were eight image sets (two patients and four compression ratios), which were duplicated (except for image position) for viewing at the two luminance levels. Eight image sets were presented at typical lightbox luminance (2,400 cd/m²), and another eight were presented at a

lower level of luminance consistent with typical, gray-scale CRT displays (200 cd/m²) by placing a clear base, optical attenuation film behind the image on the lightbox. Luminance is used because it is the physical correlate of retinal illuminance.

Observers

All four readers were board-certified radiologists specializing in chest radiography, with an average of 8 years' experience (range, 2 to 20 years). They viewed the sets one at a time, and (1) ranked the images for image quality (fidelity, degradation, artifact) after considering all available structural and tone information of the image, (2) rated the obviousness of the image quality difference between the best and second-best image (not obvious, moderately obvious, very obvious), and (3) rated the clinical importance of the image quality difference between the best and second-best image (not important, moderately important, very important).

Three viewing distances were recorded for each image set, by asking the readers to position themselves at the distance at which they: (1) would scan the best image for initial diagnostic screening (screening distance), (2) would inspect the best image to confirm or reject a suspected abnormality (interpretation), and (3) could confidently discern the difference in image quality between the best and second-best image (inspection). The CRT luminance level images were presented first, and low-compression ratio images were presented before high-compression images.

Analysis of the data took the form of descriptive statistics, in graphic form. Bar graphs and scatter plots were constructed and examined for trends. In some cases, means were calculated to present trends more concisely.

RESULTS

There was no perceived difference between the uncompressed reference and compressed image at the 10:1 compression ratio, but there were clear differences at 40:1. The differences were greater at lightbox luminance than at CRT luminance (Fig 2). At low-compression ratios (10:1 and 20:1), there was only a slight tendency to rank the uncompressed image above the compressed and reconstructed image. At high-compression ratios (30:1 and 40:1), all readers consistently ranked the uncompressed image best. All three compression algorithms performed equally at low compression levels. At higher compression levels and from an image quality perspective, EZW was ranked best, followed by WSQ and JPEG.

Readers generally felt that the differences between the best and second-best images were "not obvious" or only "moderately obvious" at low-compression ratios, and were "very obvious" at high-compression ratios (Fig 3). Generally the differences appeared more obvious at lightbox level than at CRT luminance. These differences

Fig 1. Uncompressed reference and compressed/reconstructed images. Select region of case one comparing: (A) Original reference image, (B) JPEG 20:1, (C) JPEG 40:1, (D) Wavelet based EZW 20:1, and (E) EZW 40:1. Note the block artifact in the 40:1 JPEG image is not appreciable at 20:1. Degradation of the EZW image takes the form of "blurring" similar to low-pass filtering, which is not noticeable at 20:1.

Fig 2. Perceived image quality. Mean rank of images (1 = best, 4 = worst) for four readers, by image compression algorithm and image compression ratio, for low- and high-luminance presentations. (A) Results for images presented at CRT luminance. (B) Results for images presented at lightbox luminance. Deviation from 2.5 implies preference, as random assignment of ranks would yield a mean score of 2.5.

generally were “not important” clinically at low-compression ratios at either luminance level, but the differences were more important at lightbox luminance (Fig 4).

The readers, as a group, positioned themselves

more closely to interpret than to screen images (44 cm v 59 cm, respectively), but there was considerable individual variation in the difference between the distances. Luminance did not affect the distance at which readers positioned themselves for

Fig 3. Obviousness of image degradation. Cumulative ratings of four readers for CRT and lightbox luminance presentations, evaluating the obviousness of the differences between the best and second-best images, by compression ratio. (A) Results for images presented at CRT luminance. (B) Results for images presented at lightbox luminance.

screening, but they tended to inspect the light-box images at a slightly closer distance than the CRT luminance images (42 cm v 46 cm, respectively).

As the compression level increased, readers were better able to confidently discern the differences between the best and second-best images at greater distances. The mean distance at which

Fig 4. Clinical importance of differences. Cumulative ratings of four readers, evaluating the clinical importance of the differences between the best and second-best images, by compression ratio, for CRT and lightbox luminance presentations. (A) Results for images presented at CRT luminance. (B) Results for images presented at lightbox luminance.

readers positioned themselves for this task was 34 cm for 10:1 compressed images, 40 cm for 20:1 and 30:1 images, and 45 cm for 40:1 images. The mean inspection distance was 40 cm for all com-

pression levels; thus, the readers were forced to go closer than their diagnostic inspection distance to discern differences in image quality at 10:1 compression.

DISCUSSION

The small number of cases and small image size introduce opposite effects. Small sample size allows for recall, making readers more sensitive to subtle differences in image quality, whereas small image size makes subtle differences in images harder to detect.

A number of the findings are well-supported, and can form an empirical foundation for designing a conclusive study. First, it is clear that at low-compression ratios readers are relatively insensitive to differences between uncompressed and compressed/reconstructed images, or between different compression algorithms. The inability of readers to consistently rank the uncompressed image better than the 10:1 compressed image, and the consistent rating of "not obvious" for describing the difference and "not important" for the clinical importance between the first- and second-ranked images at 10:1 compression, suggest that low compression ratios will be suitable for images intended for electronic distribution and teleradiology.

As the level of image compression increases, readers begin to prefer the uncompressed image and to develop clear and consistent preferences among the compression algorithms. However, even when they can consistently rank images, readers tend to consider the differences as neither obvious nor clinically important. The result suggests that quality differences introduced by compression are less noticeable at low-luminance levels typical of CRT soft-copy displays, which is expected given the contrast sensitivity of the human visual system. Of particular interest is the readers' unanimous

opinion that at CRT luminance, despite a detectable difference in image quality, the difference introduced by 20:1 compression is not clinically important, which suggests that higher compression ratios will be more acceptable for soft-copy (CRT) than for hard-copy display (lightbox).

Readers were able to rank images in a consistent fashion,⁹ and presenting several images dramatically reduces the number of image sets that must be read, compared with pairwise presentation. To collect our data using pairwise image presentation would have required six times as many image sets, with associated increased time commitments for readers. This is specifically relevant to the design of more comprehensive studies that investigate image compression in a medical setting.

Reader positioning can be a problem in viewer preference studies. The most common solution, positioning readers at a set viewing distance, does not take into account individual differences in visual acuity or the possibility, confirmed by our results, that some readers may position themselves differently for different diagnostic tasks. When reading distance is uncontrolled, readers position themselves more closely to spot differences in image quality than they do to inspect a suspicious finding noticed when screening a patient image. Allowing the readers to select their own reading distances, and recording those distances, solves these problems.

ACKNOWLEDGMENTS

The authors are grateful to Drs David S. Gierada, R. Gilbert Jost, and Pamela K. Woodard, for reading the images in this study.

REFERENCES

1. Bramble JM, Huang HK, Murphey MD: Image data compression. *Invest Radiol* 23:707-712, 1988
2. Ishigaki T, Sakuma S, Ikeda M, et al: Clinical evaluation of irreversible image compression: Analysis of chest imaging with computed radiography. *Radiology* 175:739-743, 1990
3. MacMahon H, Doi K, Sanada S, et al: Data compression: Effect on diagnostic accuracy in digital chest radiography. *Radiology* 178:175-179, 1991
4. Aberle DR, Gleeson F, Sayre JW, et al: The effect of irreversible image compression on diagnostic accuracy in thoracic imaging. *Invest Radiol* 28:398-403, 1993
5. Goldberg MA, Pivovarov M, Mayo-Smith WW, et al: Application of wavelet compression to digitized radiographs. *AJR* 163:463-468, 1994
6. Bunge RE, Ducca AT, Properzio WS: Exposure in chest radiology—NEXT results in optimization of chest radiography, US Dept. HHS Publication (FDA) 80-8124, 172-181, 1980
7. Jones PW, Daly S, Gaborski RS, et al: Comparative study of wavelet and DCT decompositions with equivalent quantizations and encoding strategies for medical images. *Proc SPIE* 2431:571-582, 1995
8. Eckert MP: Lossy compression using wavelets, block DCT, and lapped orthogonal transforms optimized with a perceptual model. *Proc SPIE* 3031:339-350, 1997
9. Persons K, Palisson P, Manduca A, et al: An analytical look at the effects of compression on medical images. *J Digit Imaging* 10: pp 60-66, 1997 (suppl 1)
10. Mori T, Nakata H: Irreversible data compression in chest imaging using computed radiography: An evaluation. *J Thorac Imaging* 9:23-30, 1994

11. Holbert JM, Staiger M, Chang TS, et al: Selection of algorithms for digital image compression: A rank-order study. *Acad Radiol* 2:273-276, 1995
12. Kocher TE, Whiting B: Design considerations for a high-resolution film scanner for teleradiology applications. *Proc SPIE* 1446:459-464, 1991
13. Shapiro J: Embedded image coding using zerotrees of wavelet coefficients. *IEEE Trans Signal Processing* 41:3445-3462, 1993
14. US Department of Justice, Federal Bureau of Investigation. WSQ gray-scale fingerprint image compression specification. IAFIS-IC-0110v2, 1993
15. Pennebaker WB, Mitchell JL: *JPEG Still Image Data Compression Standard*. New York, NY, Van Nostrand Reinhold, 1993
16. Cox JE, Muka E, Wang X, et al: Factors affecting the selection of compression algorithms for projection radiography. *Proc SPIE* 3031:256-264, 1997