Ultrasound Grayscale Image Compression With JPEG and Wavelet Techniques

Kenneth R. Persons, Patrice M. Palisson, Armando Manduca, William J. Charboneau,
E. Meredith James, Nicholas T. Charboneau, Nicholas J. Hangiandreou, and Bradley J. Erickson

The purpose of the study was to evaluate the effects of lossy compression on grayscale ultrasound images to determine how much compression can be applied while still maintaining images that are acceptable for diagnostic purposes. The study considered how the acquisition technique (video frame-grabber versus directly acquired in digital form) influences how much compression can be applied. For directly acquired digital images, the study considered how text (that is burned into the image) affects the compressibility of the image. The lossy compression techniques that were considered include JPEG and a Wavelet algorithm using set partitioning in hierarchical trees (SPIHT).

KEY WORDS: lossy compression, JPEG, wavelet compression, ultrasonography, Medical Images, picture archiving and communication systems, teleradiology.

BACKGROUND

Lossy Compression Techniques

Lossy compression techniques, as the name implies, are not "bit-preserving." Some of the original information is lost in favor of effectively compressing the image for purposes of more efficient transmission or storage. There are a variety of lossy compression techniques available, and this ultrasound (US) compression study focused on 2 of them, JPEG and a wavelet-based technique using set partitioning in hierarchical trees (SPIHT).[1]

JPEG[2] is an ISO (International Standards Organization) algorithm developed by the Joint Photographic Experts Group in the 1980s. Because it is a standard, it has achieved universal acceptance and has been adopted for use in a variety of applications, including teleradiology and PACS. The JPEG lossy compression algorithm is a 3-step process that operates on the image in 8 × 8 blocks.

- *Step 1* applies a discrete cosine transform (DCT) to the 8 × 8 block, creating a new 8 × 8 block in the spectral domain. This is a lossless step.
- *Step 2* is where lossy compression occurs. The 8 × 8 spectral block is divided by an 8 × 8 quantization table, thus generating an integral approximation of the spectral 8 × 8 block, where important coefficients are closely approximated, and less important coefficients are roughly approximated or zeroed.
- *Step 3*, called encoding, compactly represents the remaining integral coefficients in the quantized block. Encoding is a lossless step.

The particular quantization table used in Step 2 determines the amount of compression achieved. The quantization tables defined by JPEG were tuned to the human visual system to bias the algorithm toward preserving information to which the eye is most sensitive. Tables that result in a more rough approximation of the 8 × 8 spectral block achieve greater amounts of compression. For this study we used the JPEG compressor provided by the Independent JPEG Group.

Wavelets[3,4] are mathematical functions that, like sinusoids, can form a basis for analyzing arbitrary data sets. Unlike sinusoids, wavelets are spatially localized, and thus better suited for representing local features in an image, such as boundaries. An application for which wavelets are well suited is compression.

Wavelet compression,[5,6] like JPEG, is a 3-step process.

- *Step 1* is a lossless step that uses a pair of wavelets (low and high pass filters) to transform the image from a spatial to a spectral or wavelet representation. This step is called a discrete wavelet transform.
- *Step 2* is a lossy quantization step in which the coefficients are approximated. The quantization technique used by this wavelet algorithm is called SPIHT. In simple terms, the technique approximates the most important coefficients first and, through a series of passes, refines those approximations. The resultant data stream can be truncated at any point to yield an approximate representation of the coefficients, which can be transformed back into an approximate representation of the

---

From the Mayo Foundation, Rochester, MN.

Address reprint requests to Kenneth R. Persons, Mayo Foundation, 200 First St SW, Rochester, MN 55905.

0897-1889/00/1301-0005$10.00/0

image. The shorter the data stream (greater compression), the more rough the approximation.

- *Step 3* is a lossless step in which additional compression is achieved by more compactly representing the coefficient stream through techniques such as arithmetic, binary, and/or run length encoding.

Unlike JPEG, a wavelet compression standard has not been established (as of the time of this writing). Different wavelet pairs can be used to perform the discrete wavelet transform in Step 1, and a variety of approaches (such as SPIHT) can be used to achieve quantization in Step 2.

OBJECTIVES

The primary objective of this study was to determine to what extent grayscale US images can be compressed (using JPEG and wavelet technology) and still remain of "diagnostic quality." "Diagnostic quality" indicates that the radiologist considered the image to be of high enough quality that the ability to make a diagnosis was not compromised.

Secondary objectives included:

- Comparing the effectiveness of JPEG and SPIHT wavelet compressors.
- Determining what influence the acquisition technique (video frame-grabbed versus acquired directly in digital form) has on compressibility of the images.
- Examining how much the text embedded into the image affects the compressibility of the image.

MATERIALS AND METHODS

Acquisition and Selection of Images for the Study

Images were acquired using state-of-the-art US scanners from 2 vendors (designated A and B). Both scanners were capable of directly generating digital US images and exporting them to a DICOM receiver. A picture archiving and communication system (PACS) was attached to the scanners so the same images could be acquired via the video frame-grabber. Each scanner was used in clinical practice for 1 week to acquire a typical mix of US studies. Every image generated during the week was acquired into the PACS (via the video frame-grabber), printed to film (from the PACS), and acquired directly from the scanner to a DICOM receiver station in the compression research laboratory. This library of US images was used as a pool from which a representative mix of images from a variety of anatomic regions were selected, including the liver, kidney, thyroid, abdomen, gall bladder, and aorta.

To determine the range of compression levels to use, a pre-evaluation was performed (in which the radiologists were looking for acceptability for diagnostic purposes) using images compressed with JPEG and SPIHT wavelet at 8:1, 10:1, 12:1, 14:1, 16:1, 18:1, 20:1, and 40:1. The results of this pre-evaluation indicated that 10:1 was always acceptable; however, at ratios between 12:1 and 18:1, some images had observable differences (from the original), which made the radiologists uncomfortable reading them. It was decided that the appropriate compression levels for the study would be 10:1, 12:1, 15:1, 17:1, 20:1, and 25:1, and that more images should be processed at ratios of 12:1, 15:1, and 17:1 because this appeared to be the range at which some of the images became of questionable diagnostic value. Table 1 shows how many images were selected for each acquisition method, compression type, and level. To facilitate a comparison of the wavelet and JPEG results, the specific JPEG quality factor that yielded the desired compression ratio was calculated and applied for each image.

Processing of Selected Images

A naming convention was established to be able to identify how and from which scanner the images were acquired, what type of compression was applied and at what level, and whether or not text was removed before compression. The images were then processed as indicated in Table 1. Each processed image was paired with its original (randomly, so sometimes the original would appear on the left, and sometimes on the right in a viewing pair). Twelve random image pairs were grouped for a particular patient, and a DICOM header added (or modified) to make all 12 pairs appear under one patient name, patient ID, and study on a DICOM viewstation.

Table 1. Number of Images for US Compression Study

Acquisition Method	Compression Ratio	Wavelet Images	JPEG Images	Repeat Images	Total Images
Frame-grabbed	10:1	6	6		12
	12:1	16	16		32
	15:1	16	16		32
	17:1	16	16		32
	20:1	6	6		12
	25:1	6	6		12
Directly acquired	10:1	6	6		12
	12:1	16	16		32
	15:1	16	16		32
	17:1	16	16		32
	20:1	6	6		12
	25:1	6	6		12
No text	10:1	0	0		0
	12:1	6	6		12
	15:1	6	6		12
	17:1	6	6		12
	20:1	6	6		12
	25:1	6	6		12
Repeats	1:1			48	48
Total		162	162	48	372

NOTE. Half the images were from vendor A and half were from vendor B ultrasound scanners.

Evaluation Protocol

Thirty-one patient series (of 12 image pairs each) were generated and sent via a DICOM network connection to a DICOM viewstation for use in the evaluation. The viewstation was located in a US reading area where standard viewing conditions existed, and to which the 3 radiologists who would be assessing the image quality could have easy access.

The images were read in pairs as they appeared on the viewstation (without the use of zoom, window/level, or brightness/contrast tools) in a blinded-paired fashion. Each compressed/decompressed image was shown next to the original. The radiologist knew that one or both of the images were originals, but did not know which image was original, and which (if either) was compressed.

For each pair, the radiologist was asked to identify whether:

- Both images are of *diagnostic quality,* and if so:
  - both images are of the *same* diagnostic quality *or*
  - they prefer the image on the left (a) or right (b)
- Only the left (a) or right (b) image is of *diagnostic quality,* the other was not.

Other Measures to Assess Image Quality and the Effects of Compression

Objective measures (maximum pixel error, average pixel error, average root mean square [RMS] pixel error, and signal-to-noise ratio [SNR]) were calculated for every image to compare how well the 2 compressors performed with respect to mathematical accuracy. These measures give an indication of how similar the pixel values of the compressed/decompressed images are to those of the original.

In the compression lab, observations were made to note any visible effects of compression and to determine the maximum compression level where there were no observable differences between the compressed and original image. The methods used to do this were (1) rapidly toggling back and forth between the original and compressed image looking for any subtle changes and (2) simultaneously (side by side) zooming regions of interest (ROIs) within the image, looking for any artifacts or differences.

RESULTS

Image Quality: JPEG Versus SPIHT Wavelet Subjective Results

Figures 1 and 2 summarize the radiologist readings for all acquired images (also see Table 1).

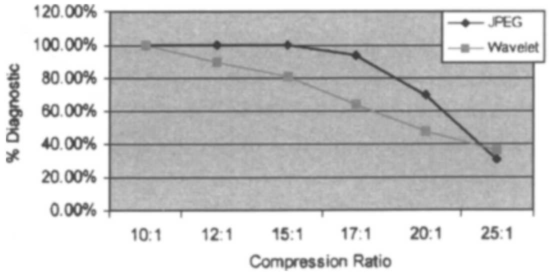


Fig 1. Percentage of images considered of "diagnostic quality."

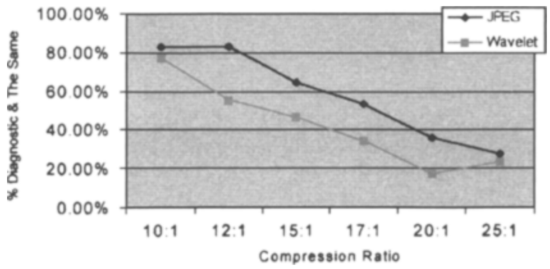


Fig 2. Percentage of images of "diagnostic quality" and "the same."

Figure 1 shows, for JPEG and wavelet, as a function of compression ratio, the percentage of image pairs in which both images were considered of diagnostic quality. For JPEG, 100% of the images were considered of diagnostic quality at compression ratios of up to 15:1. For SPIHT wavelet, 100% of the images were considered of diagnostic quality at a compression ratio of 10:1. From a subjective viewpoint, at low to moderate levels of compression, JPEG clearly produced more diagnostically acceptable images than SPIHT wavelet.

Figure 2 shows, for the pairs in which both images were considered of diagnostic quality, the percentage of pairs where the radiologists considered both images to be "the same" (ie, they did not prefer one image over the other). For cases in which both images were of diagnostic quality, the JPEG compressed images were more likely to be considered the same. For most cases in which the images were not considered the same, the original was preferred over the compressed image; however, for low levels of JPEG compression (10:1), there was a slight preference for the compressed image.

Figure 3 contrasts the subjective results from Fig 1 (percent of images considered to be of diagnostic quality) with one of the objective measures (aver-

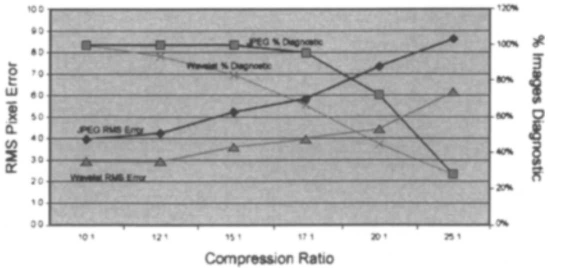


Fig 3. Percentage of Images: Diagnostic quality versus average RMS pixel error as a function of compression ratio.

age RMS pixel error). Objective measures indicate how similar the pixel values of the compressed/decompressed image are to those of the original image. For each objective measure (maximum pixel error, average pixel error, average RMS pixel error, and signal-to-noise ratio), SPIHT wavelet consistently produced a more "mathematically accurate" image than JPEG at all compression levels, especially at higher levels. See the Discussion section for additional comments.

Effects of Acquisition Method and Matrix Size

The image itself (the content of the image, the method by which it was acquired, and the matrix size, ie, number of bytes used to store the image information) influences the extent to which an image can be compressed and still be considered diagnostically acceptable.[1] For example, vendor A's scanner produced a DICOM digital image with a matrix size of 640 × 486 × 8 (311 KBytes). Vendor B's scanner produced a similar image in a matrix size of 888 × 666 × 8 (591 KBytes). As is shown in Fig 4 and 5, the vendor B DICOM image was more tolerant to compression, but because of the larger size, this does not necessarily result in a smaller image file size.

Figure 4 shows the JPEG results from Fig 1 according to acquisition method. The frame-grabbed images (matrix size, 640 × 486 × 8) were acquired from both vendor A and vendor B US scanners via a PACS video capture station. The DICOM vendor A (matrix size, 640 × 480 × 8) and vendor B (matrix size, 888 × 666 × 8) images were acquired directly from the scanner in DICOM digital format. The frame-grabbed images and DICOM vendor A images had a similar matrix size, and were considered diagnostically acceptable at JPEG compression levels of up to 15:1. The larger vendor B images were considered diagnostically acceptable at JPEG compression levels of up to 20:1.

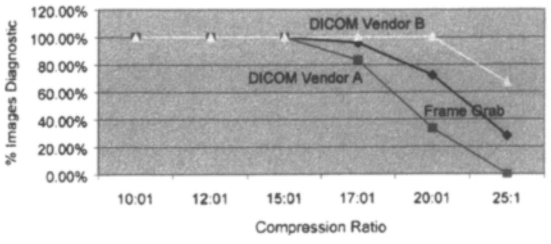


Fig 4. **JPEG results for 3 acquisition methods.**

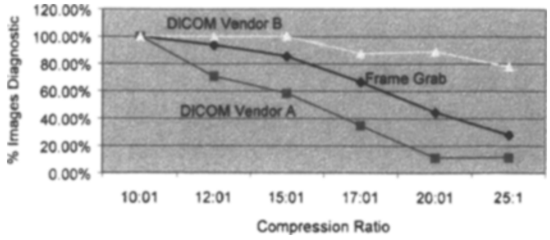


Fig 5. **Wavelet results for 3 acquisition methods.**

Figure 5 shows the same measurements as Fig 4 for wavelet compression. The video frame-grabbed images and DICOM vendor A images (with a similar matrix size) were considered diagnostically acceptable at SPIHT wavelet compression levels of 10:1. Video frame-grabbed images were more likely to be considered of diagnostic quality than directly acquired vendor A images at compression levels greater than 10:1. The larger vendor B images were considered diagnostically acceptable at SPIHT wavelet compression levels of up to 15:1.

Effects of Text Removal

Text adds substantial high-frequency content to an image, which makes it harder to compress. To understand the influence of text on the compressibility of DICOM images, the text was removed before compression, and added back in after compression. It was possible to do this because the text was added to the digitally generated files using a particular grayscale value (eg, 254).

Figure 6 shows the effects of removing text on the diagnostic acceptability of the images for both JPEG and wavelet compressed images. The results showed consistently, for both compression techniques, that the percentage of images considered of diagnostic quality is higher for images with text removed, especially at higher compression levels. See the Discussion section for more information.

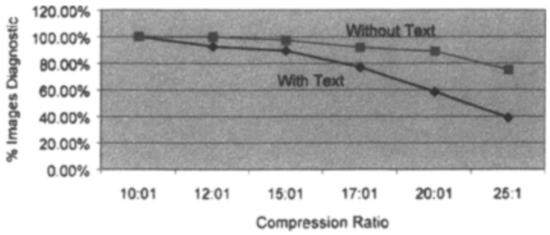


Fig 6. **Effects of text removal on compressibility of US images.**

LAB OBSERVATIONS

Using techniques such as rapidly toggling back and forth between the original and compressed image, or zooming ROIs in the original and compressed images side by side, the maximum compression level where there were no observable differences was approximately 5:1 for both JPEG and SPIHT wavelet. At compression levels higher than 5:1, we observed subtle differences between the 2 images. The following observations were noted.

JPEG Blocking Artifact

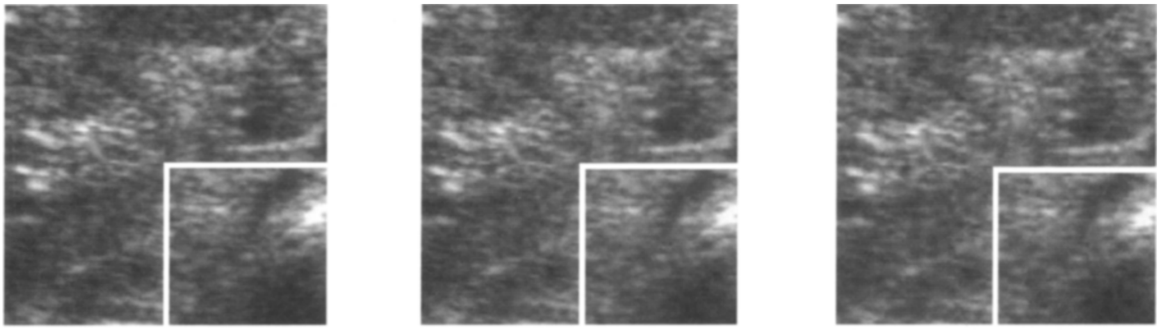
Subtle blocking artifacts were barely visible in some JPEG images compressed at 10:1, and were somewhat more visible in JPEG images compressed at 15:1. These artifacts were much more pronounced when the images were zoomed by a factor of 2 (Fig 7). In the subjective study, radiologists did not use zooming or brightness/contrast tools to manipulate the images, and despite the appearance of these very subtle blocking artifacts, all JPEG images (compressed at 10:1, 12:1, or 15:1) were still considered to be of diagnostic quality. At 12:1 and 15:1, however, radiologists indicated a preference for the original image over the compressed, and this may have been a factor that influenced their decision. If zooming, brightness/contrast, or other image enhancement tools are used to read US images, such artifacts might be a concern, especially at ratios of 12:1 and 15:1.

JPEG Text Distortion

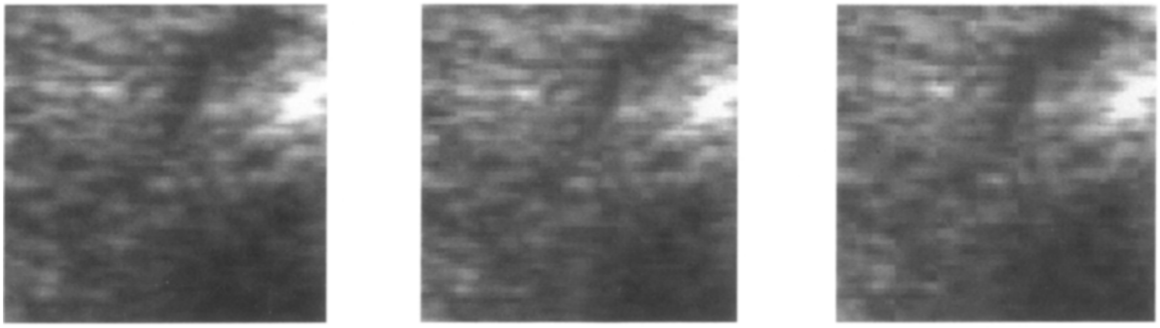
The sharp edges of text and fiduciary markers in the image cause a ghosting or ringing artifact along the text or marker borders in JPEG compressed images (Fig 8). This ghosting artifact did not affect the diagnostic acceptability of JPEG images (compressed up to 15:1). However, it may have been a factor that influenced radiologists to show a preference for the original image over the same image JPEG compressed image at 12:1 and 15:1.

Wavelet Blurring and Artifact

The wavelet compressed images showed a visible softening or blurring of image structure at lower compression levels than JPEG, especially in low-contrast areas of the image (Fig 9). Blurring and/or subtle wavelet "rice grain" artifact some-

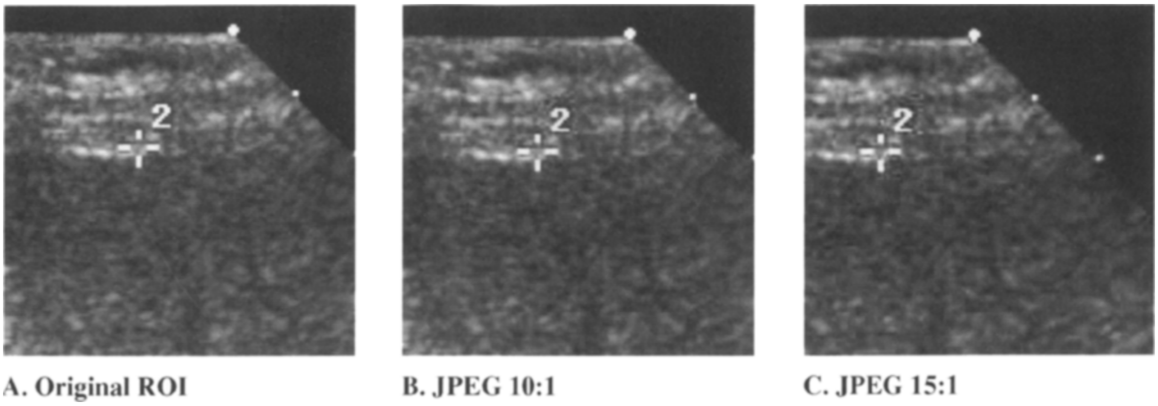


A. Original US ROI B. JPEG 10:1 C. JPEG 15:1



D. Original X2 E. JPEG 10:1 X2 F. JPEG 15:1 X2

Fig 7. JPEG blocking artifact and effects of zooming an ROI.



A. Original ROI B. JPEG 10:1 C. JPEG 15:1

Fig 8. JPEG ghosting around text and fiduciary markers.

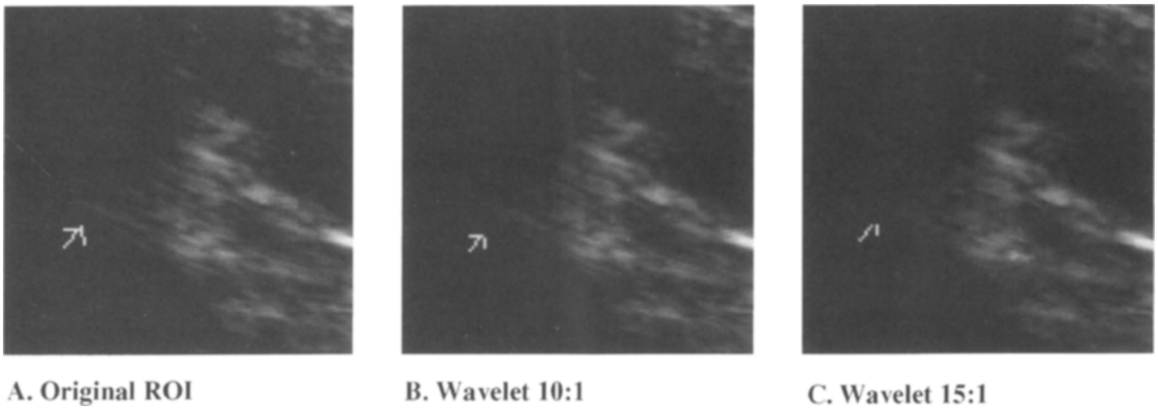
times were visible in SPIHT wavelet images compressed at 10:1, and usually were apparent when compressed at 15:1 (Fig 9). At 10:1, the subtle blurring and artifact were not significant enough to affect the diagnostically acceptability of the images, but the more significant blurring and artifact apparent at 15:1 caused these images to be considered nondiagnostic in many cases. As with the JPEG blocking artifact (Fig 7), the wavelet blurring and artifact were more apparent if the images were zoomed.

DISCUSSION

Subjective and Objective Results

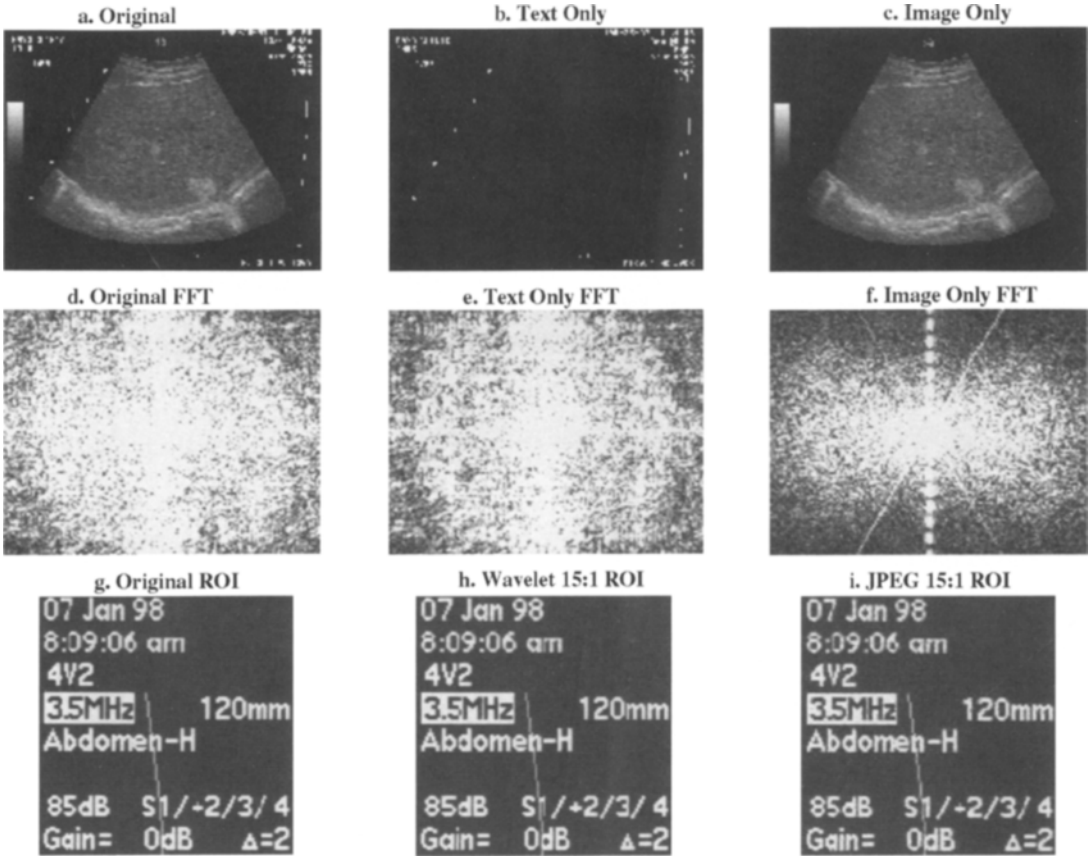
From a subjective viewpoint at low to moderate levels of compression, JPEG produced a more diagnostically acceptable image than SPIHT wavelet, regardless of the method of acquisition. From an analytical viewpoint, SPIHT wavelet produced a more mathematically accurate image than JPEG by every objective measure (Fig 3). We suggest that the reason for this lies in the nature of US images, and the very different quantization processes used by the 2 compression algorithms. The JPEG quantization process is tuned to the human visual system; it attempts to preserve frequencies to which the eye is most sensitive, regardless of image content. The SPIHT quantization process is tuned to produce the most mathematically accurate approximation of the image, and is sensitive to the image content, giving priority to preserving frequencies of significant magnitude over those of lesser magnitude.

Consider the nature of grayscale US images. They contain significant amounts of embedded text, which generates high-frequency coefficients of substantial magnitude in the spectral or wavelet domain. Figure 10 shows fast fourier transforms (FFTs) with spectral representations of the image with text, text only, and image only. Because the SPIHT quantization process is sensitive to the magnitude of the coefficients, it does a better job of preserving the text than JPEG, but at the expense of



A. Original ROI B. Wavelet 10:1 C. Wavelet 15:1

Fig 9. Wavelet artifact and blurring.



a. Original b. Text Only c. Image Only

d. Original FFT e. Text Only FFT f. Image Only FFT

g. Original ROI h. Wavelet 15:1 ROI i. JPEG 15:1 ROI

Fig 10. Influence of text on compressibility of the image. Figures a-c show spatial and d-f show spectral representations of the image with text, text only, and image only. Figures 10 g-i show a text ROI as an original ROI (10g), after SPIHT wavelet compression at 15:1 (10h), and after JPEG compression at 15:1 (10i).

using bytes that could have been used to better approximate image content. At 15:1, the text in the wavelet compressed image was still crisp, whereas that in the JPEG image was blurred and showed ringing along the edges (Figs 10h and 10i). Text, however, is not the entire cause. Even after removal of the text, JPEG produced more diagnostically acceptable images.

Grayscale US images have a speckle background that generates many decorrelated medium- to high-frequency coefficients of small to moderate magnitude in the spectral or wavelet domain. These smaller coefficients in the higher frequency wavelet subbands are the most vulnerable to being attenuated by the SPIHT quantization process, which results in a subtle blurring or "softening" of the speckle structure. JPEG, on the other hand, produces images (in the diagnostically acceptable range) that appear slightly sharper with more contrast, even though there can be subtle changes in the speckle structure.[7] At low to moderate compression levels, the slightly sharper JPEG compressed images are visually preferable to the subtly softened SPIHT wavelet images.

CONCLUSION

Several important conclusions can be drawn from the results of this study.

1. Using the JPEG standard algorithm at 10:1, grayscale US images (frame-grabbed or digitally acquired through a DICOM interface on the scanner) were considered of diagnostic quality, with a slight preference for the compressed image over the original. This slight preference for the compressed image at low levels of JPEG compression probably is related to the removal of noise.[1,5] JPEG 10:1 compression corresponds to a quality factor of approximately 70 to 75 for grayscale US images.

2. Using the JPEG standard algorithm at 12:1 and 15:1, grayscale US images (frame-grabbed or digitally acquired through a DICOM interface on the scanner) were considered of diagnostic quality, with a preference for the original image over the compressed image.
3. Using the JPEG standard algorithm at compression levels greater than 15:1, images were not necessarily of diagnostic quality (except for vendor B DICOM images, which had a larger matrix size; see item 9 below).
4. A slight blocking artifact was visible on images compressed with the JPEG algorithm at 10:1, and this was more apparent at 12:1 and 15:1. The artifact was enhanced by manipulating the image (zooming, applying filters, adjusting the brightness/contrast).
5. Using the SPIHT wavelet algorithm at 10:1, grayscale US images (frame-grabbed or digitally acquired through a DICOM interface on the scanner) were considered of diagnostic quality, with a preference for the original over the compressed image.
6. Using the SPIHT wavelet algorithm at compression levels greater than 10:1, images were not necessarily of diagnostic quality.
7. Although all video frame-grabbed US images and directly acquired vendor A DICOM images (with a similar matrix size) were considered of diagnostic quality up to 15:1 (JPEG) and 10:1 (SPIHT wavelet), the frame-grabbed images were more likely of diagnostic quality than the directly acquired vendor A DICOM images at higher levels of compression.
8. Objective results indicate that SPIHT wavelet compression produced a more mathematically accurate image (lower average error per pixel, lower average RMS error, higher SNR) than JPEG compression at all compression levels, especially higher levels. Despite this, the JPEG compressed image was preferred to the SPIHT wavelet compressed image at most compression levels. This confirms that there is not a direct correlation between objective and subjective results.
9. One factor that affects the compressibility of images is the matrix size of the image. The vendor B DICOM US scanner uses a much larger matrix size for each image (888 × 666 × 8 versus 640 × 480 × 8 for vendor A), and as a result was more compressible. This is not necessarily a benefit, however, because the resultant compressed file was larger for a similar level of compression acceptance. The vendor B DICOM image would have to be compressed at 28:1 to match the size of the same frame-grabbed or vendor A DICOM file compressed at 15:1.
10. Text burned into an image significantly reduces the compressibility of the image.

REFERENCES

1. Persons KR, Palisson P, Manduca A, et al: An analytical look at the effects of compression on medical images. J Digit Imaging 10:60-66, 1997 (suppl 1)

2. Wallace GK: The JPEG still picture compression standard. Comm ACM 34:30-44, 1991

3. Rioul O, Vetterli M: Wavelets and signal processing: IEEE Signal Process Mag 5:14-38, 1991

4. Villasenor J, Belzer B, Liao J: Wavelet filter evaluation for image compression. IEEE Trans Image Proc 4:1053-1060, 1995

5. Erickson BJ, Manduca A, Palisson P, et al: Wavelet compression of medical images. Radiology 206:599-607, 1998

6. Manduca A: Compressing images with wavelet/subband coding. IEEE Engineer Med Biol 0739-5175:639-646, Sept/Oct 1995

7. Persons KR, Palisson PM, Manduca A, et al: JPEG/wavelet ultrasound compression study. SPIE Med Imaging 3662-33:232-247, 1999