

JPEG2000 Still Image Coding Quality

Tzong-Jer Chen · Sheng-Chieh Lin · You-Chen Lin ·
Ren-Gui Cheng · Li-Hui Lin · Wei Wu

Published online: 16 April 2013
© Society for Imaging Informatics in Medicine 2013

Abstract This work demonstrates the image qualities between two popular JPEG2000 programs. Two medical image compression algorithms are both coded using JPEG2000, but they are different regarding the interface, convenience, speed of computation, and their characteristic options influenced by the encoder, quantization, tiling, etc. The differences in image quality and compression ratio are also affected by the modality and compression algorithm implementation. Do they provide the same quality? The qualities of compressed medical images from two image compression programs named Apollo and JJ2000 were evaluated extensively using objective metrics. These algorithms were applied to three medical image modalities at various compression ratios ranging from 10:1 to 100:1. Following that, the quality of the reconstructed images was evaluated using five objective metrics. The Spearman rank correlation coefficients were measured under every metric in the two programs. We found that JJ2000 and Apollo exhibited indistinguishable image quality for all images evaluated using the above five metrics ($r > 0.98$, $p < 0.001$). It can be concluded that the image quality of the JJ2000 and Apollo algorithms is statistically equivalent for medical image compression.

Keywords Image compression · JPEG2000 · Image quality

Introduction

Recent technological advances have made digital radiology a practical alternative to the film-based system [1]. The fast retrieval and ease of transmission of digital data make this alternative particularly attractive. The image information flow is the essential idea behind digital radiology, i.e., fast retrieval, transfer, display, and archiving. Digitized images must have high quality and resolution and efficiently manage a very large data volume in general [2]. The amount of data generated by imaging devices of all digital radiological modalities is massive and steadily increasing [3]. However, both the speed of data transmission and the space storage requirements depend on the amount of data. The data increase poses technical challenges in regard to effectively archiving, transferring, and interpreting these data [4]. Data compression techniques substantially reduce the image data volume generated and thus increase information flow efficiency. The amount of image volume affects picture archiving and communications systems (PACS) effectiveness, as well as telemedicine networks and radiology information systems.

Data compression methods can be reversible (lossless) or irreversible (lossy) [5]. A reversible scheme will allow exact recovery of the original image from the compressed version but can only achieve a maximum compression ratio (CR) (i.e., $CR = \text{original image volume} / \text{compressed image volume}$) of about 2.5 [3]. A lossy scheme will not allow exact recovery after compression but can achieve greater compression factor, but a higher CR has an image quality trade-off.

The wavelet transform-based image compression algorithms are recognized as a better method to compress, archive, and communicate medical images [6]. This algorithm is now available to an extensive medical system user base with the approval of JPEG2000 (ISO 15444-1) as a supported method within the DICOM standard (Digital Imaging and Communications in Medicine).

T.-J. Chen (✉) · R.-G. Cheng · L.-H. Lin · W. Wu
Department of Mathematics & Computer Science, Wuyi
University, Wuyishan, Fujian, China 354300
e-mail: d838502@alumni.nthu.edu.tw

S.-C. Lin
Department of Dental Laboratory Technology, Shu-Zen College
of Medicine and Management, Kaohsiung, Taiwan 82144

Y.-C. Lin
Department of Radiology, Taichung Veterans General Hospital,
Taichung, Taiwan 40705

There are many JPEG2000-based popular image compression programs, such as the Aware JPEG2000 (Aware, Inc., Bedford, MA), the Pegasus PICS Tools JPEG2000 (Apollo; Pegasus Imaging Co., Tampa, FL) [7], and the JJ2000 (JJ2000 version 4.1, available on the Internet at <http://jj2000.epfl.ch>). The JPEG2000 standard is only defined for the stream syntax decoder. These independently developed JPEG2000 compression software coding algorithms have different encoders. These differences directly affect the coding efficiency, speed, and implementation complexity [8]. Zhang et al. used the performance of a model observer to optimize the JPEG2000 encoder options through a genetic algorithm procedure [2]. Recently, Kim et al. reported that the achieved CR at a given nominal CR was not always comparable between the Aware and Apollo systems, two popular commercial JPEG2000 compression programs [7]. They also found that some of the definitions are inconsistent between the programs, CR for example, and the details are not provided in the user manuals. The quality of compressed images may be different because of these options.

Two image compression programs are JPEG2000 encoded, the Apollo and the JJ2000, developed by different organizations, respectively. Apollo has a window mode interface and is easy to use, but time-consuming. The JJ2000 has Microsoft

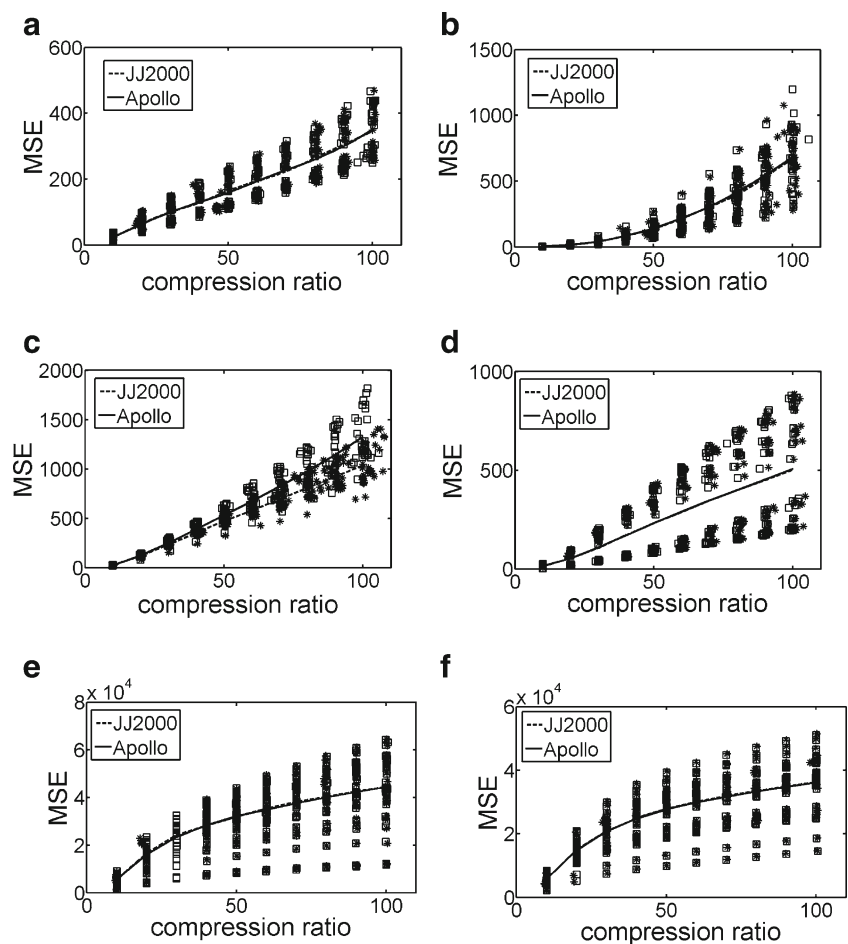
prompt mode, and you have to prepare a program before using it. The above factors will be considered by hospitals when they acquire compression algorithms for their PACS system. Besides, is there a difference in quality between these two programs?

For some time now, considerable effort has been made to evaluate digital image compression techniques to permit the image quality required for medical images [9–20]. These evaluations include subjective and objective metrics. The subjective image quality measurement requires extreme caution and is costly. The objective metrics are relatively easy and efficient.

Some objective metrics have been widely used for image quality evaluation [6, 13, 21–23]. These metrics can be categorized into three groups: (1) pixel-based metrics, peak signal-to-noise ratio (PSNR) and mean square error (MSE) [13, 23]; (2) window-based metrics, which include the Q index [23], Moran peak ratio (MPR) [6, 21, 22]; and (3) a human visual system (HVS)-based metrics [13].

This study applied these algorithms to three medical modality images first. The images were compressed at ten different CRs. Following that, the qualities of the reconstructed images were measured using five objective metrics. Images, only original, were chosen from a PACS system from a general

Fig. 1 Comparison of JJ2000 (broken lines, asterisk) and Apollo (straight line, square) image quality for various CRs: **a** CT abdominal, **b** CT head, **c** MR abdominal, **d** MR head, **e** DR abdominal, **f** DR chest using MSE



hospital in Central Taiwan. These modalities are computer tomography (CT), magnetic resonance (MR) images, and digital radiography (DR) images. Sixty images were chosen for each modality, totaling 180 images.

Kim et al. [4] suggested that the Spearman rank correlation coefficients can be used to present how an objective metric correlates to visual analysis or to the other objective metrics. They found that the PSNR results, HDR-VDP (high-dynamic range visual difference predictor), and MS-SSIM (multi-scale structural similarity) correlated well with the results from five responding radiologists [4]. In this report, the Spearman rank correlation coefficients were measured under every metric between two programs.

Methods and Materials

A and B were used to represent the original and processed images in the following equations. J_A and J_B were used to denote the pixel values, respectively.

The Pixel-Based Metric

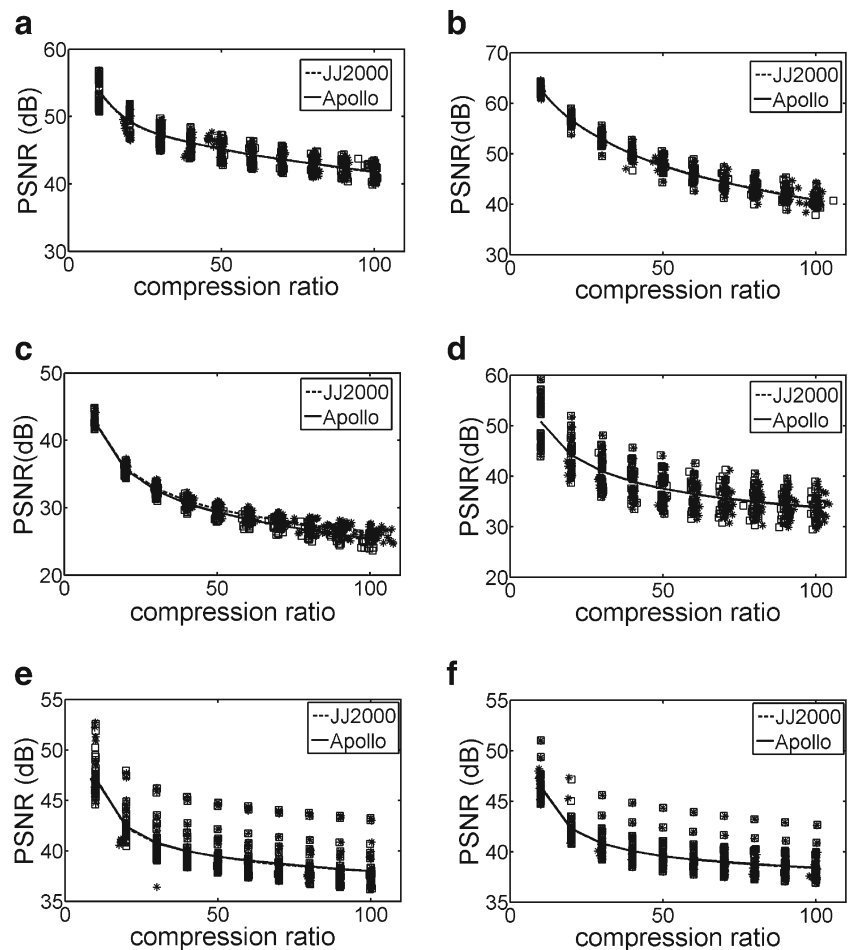
The pixel-based metrics generally measure image quality degradation using the differences between all corresponding pixels. In this study, the PSNR and MSE were used as pixel-based metrics. These metrics are widely used as image quality indicators to indicate how “close” one image is to another. The error measure-based metrics are listed below.

$$MSE = \frac{1}{N} \sum_i [J_A(i) - J_B(i)]^2 \quad (1)$$

$$PSNR \text{ (dB)} = 10 \log_{10} \left[\frac{(2^n - 1)^2}{\sqrt{\frac{1}{N} \sum_i [J_A(i) - J_B(i)]^2}} \right] \quad (2)$$

where n is the depth of the bits in a pixel, and N is the total number of pixels in the image. A lower MSE and higher PSNR correspond to better image quality.

Fig. 2 Comparison of JJ2000 (broken lines, asterisk) and Apollo (straight line, square) image quality for various CRs: **a** CT abdominal, **b** CT head, **c** MR abdominal, **d** MR head, **e** DR abdominal, **f** DR chest using PSNR



The Window-Based Metrics

These metrics evaluate image quality from a local region instead of using pixels. The quality of images in window-based metrics uses the mean, variance, or covariance calculation between two windows. These windows in corresponding positions in the image have equal size. Two window-based image quality metrics were used in this study: a Q index proposed by Wang et al. [23] and the MPR by Chen et al. [21, 22]. Both metrics estimate image spatial information from a local region with an 8×8 window size instead of a single pixel [7, 23]. The Q index is estimated as follows:

$$Q = \frac{4\sigma_{AB}M_A M_B}{(\sigma_A^2 + \sigma_B^2)(M_A^2 + M_B^2)} \tag{3}$$

where M and σ^2 are the mean and variance of the pixel values inside the window, and

$$\sigma_{AB} = \frac{1}{N-1} \sum_{i=1}^N [W_A(i) - M_A][W_B(i) - M_B] \tag{4}$$

is the covariance between windows in an image (N =number of pixels in the window). The $W_A(i)$ and $W_B(i)$ are the gray levels of the pixels in the windows.

The dynamic range for Q should be in $[0, 1]$. The value 1 is achieved when J_A and J_B are identical, which means the

$$\sigma^2 = \frac{N[(N^2 - 3N + 3)S_1 - NS_2 + 3S_0^2] - K[N(N-1)S_1 - 2NS_2 + 6S_0^2]}{(N-1)(N-2)(N-3)S_0^2} - a^2 \tag{7}$$

where K is defined by

$$K = N \sum (x_i - \bar{x})^4 / \left[\sum (x_i - \bar{x})^2 \right]^2 \tag{8}$$

where $S_1=2S_0$ and $S_1=8(8mn-7m-7n+4)$. The standardized normal statistic

$$Z = \frac{C - a}{\sigma} \tag{9}$$

can be used to determine the structural information in an image [7, 21, 22].

Collecting all Z values in an image and then sorting them into bins can produce a Z histogram. The spatial correlation increases with the amount of image blurring and accompanies the increase in Z value. This Z value will increase in

best quality. The Q index is calculated using a sliding window approach [23].

The Moran coefficient C and standard score Z for pixels in an $m \times n$ window are calculated as follows [24]:

$$C = \frac{N \sum_{j=1}^{m \times n} \sum_{i=1}^{m \times n} \delta_{ij}(x_i - \bar{x})(x_j - \bar{x})}{S_0 \sum_{i=1}^{m \times n} (x_i - \bar{x})^2} \tag{5}$$

where x_i is the gray level of pixel i ; \bar{x} is the mean gray level inside the window; $\delta_{ij}=1$ if pixel i and j are adjacent, and 0 otherwise; $S_0=2(2mn-m-n)$, where m and n are the number of rows and columns in the window, and N is the total number of pixels in the window. The numerator is a measure of the covariance, and the denominator is a measure of the variance among the pixels. For a larger C value, there is a greater correlation between pixels, and the image is blurred. When the size of N is large enough (i.e., >25), the variable approximately follows a normal distribution with the mean and variance given by [24]

$$a = -1/(N - 1) \tag{6}$$

and

certain areas to form a peak [7, 22]. The MPR is a peak ratio of the Z value between manipulated and original images. It has been proven to correspond well to the image variation in spatial properties. The higher the Q and the lower the MPR correlate well with better image quality.

The HVS-Based Metric

Because a human observer is the end user of image quality measurement, an image quality model based on HVS seems to be more appropriate for user perception. In order to obtain a closer relation with the assessment by the human visual system, both the original and compressed images can be preprocessed via filters that simulate the HVS. The models for the human visual system are, in general, given as a band-pass filter with a transfer function in polar coordinates [13].

$$K(\rho) = \begin{cases} 0.05e^{\rho^{0.054}} & \rho < 7 \\ e^{-9[|\log_{10}\rho - \log_{10}9|]^{2.3}} & \rho \geq 7 \end{cases} \tag{10}$$

where $\rho=(u^2+v^2)^{1/2}$. An image was first transformed using 2-D discrete cosine transform (DCT) to frequency domain as $\Phi(u, v)$. Following that, it was processed through a spectral filter and then inverse DCT transformed. This process can be expressed as the $T\{\cdot\}$ operator, i.e.,

$$T\{C(i, j)\} = \text{DCT}^{-1}\left\{K\left(\sqrt{u^2 + v^2}\right)\Phi(u, v)\right\} \quad (11)$$

DCT^{-1} is the 2-D inverse DCT. The measure of multi-spectral images for this study is shown below

$$H = \frac{\sum |T\{A(i, j)\} - T\{B(i, j)\}|}{\sum |T\{A(i, j)\}|} \quad (12)$$

where $A(i, j)$ and $B(i, j)$ represent the original and manipulated images, respectively. It is obvious that the lower the H values are, the better is the image quality.

Images

For extensive evaluation, we applied the above algorithms to various medical image modalities: CT (image size of 512×512 and 12 bits deep), MR (image size of 512×512 for head

and 256×160 for abdomen, all images are 12 bits deep), and DR images (image size of $2,000 \times 2,000$ and 15 bits deep). Sixty images were chosen for each modality, totaling 180 images. For CT and MR images, 30 head images and 30 abdominal images were used. For DR, 30 chests with 30 abdominal images were used. All of these images were randomly chosen from a PACS system from a general hospital in Central Taiwan.

Results

The images were first compressed at ten different CRs (10~100). Following that, the quality of the reconstructed images was evaluated using the above five metrics.

The Pixel-Based metrics

The differences between all corresponding image pixels are the basis for the quality of images in the pixel-based metrics. The MSE and PSNR were used in this study as pixel-based image quality indicators. The lower the MSE and the higher the PSNR values correspond to images that are close with better image quality.

Fig. 3 Comparison of JJ2000 (broken lines, asterisk) and Apollo (straight line, square) image quality for various CRs: **a** CT abdominal, **b** CT head, **c** MR abdominal, **d** MR head, **e** DR abdominal, **f** DR chest using Q index

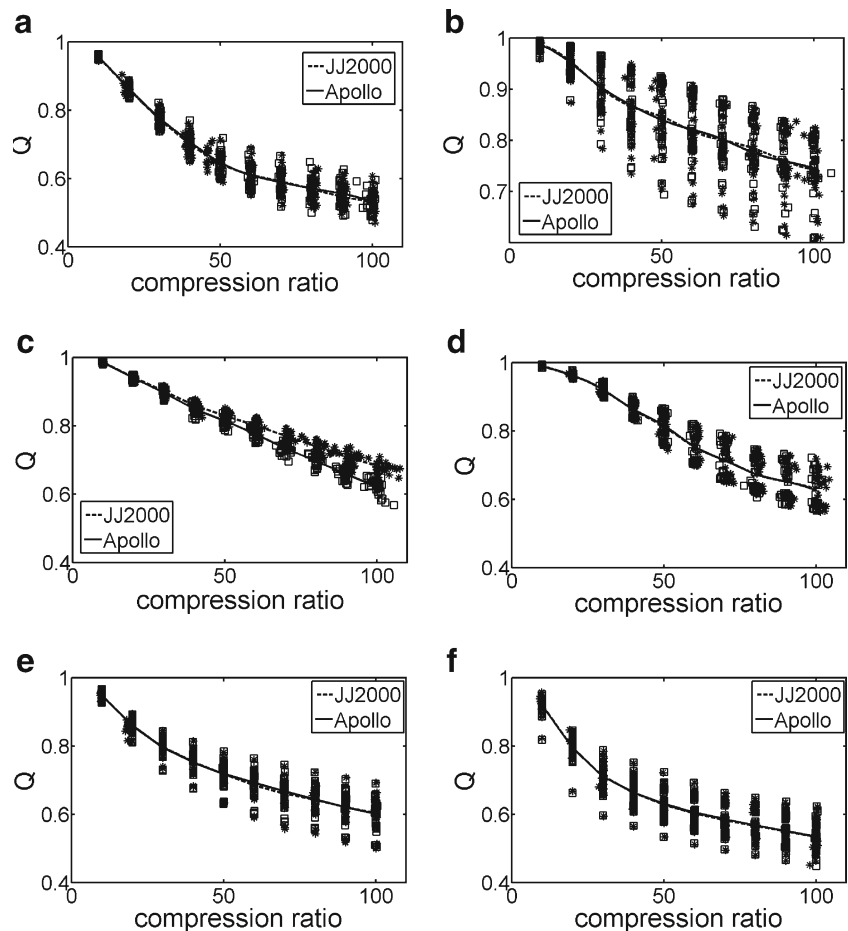


Fig. 4 Comparison of JJ2000 (broken lines, asterisk) and Apollo (straight line, square) image quality for various CRs: **a** CT abdominal, **b** CT head, **c** MR abdominal, **d** MR head, **e** DR abdominal, **f** DR chest using MPR

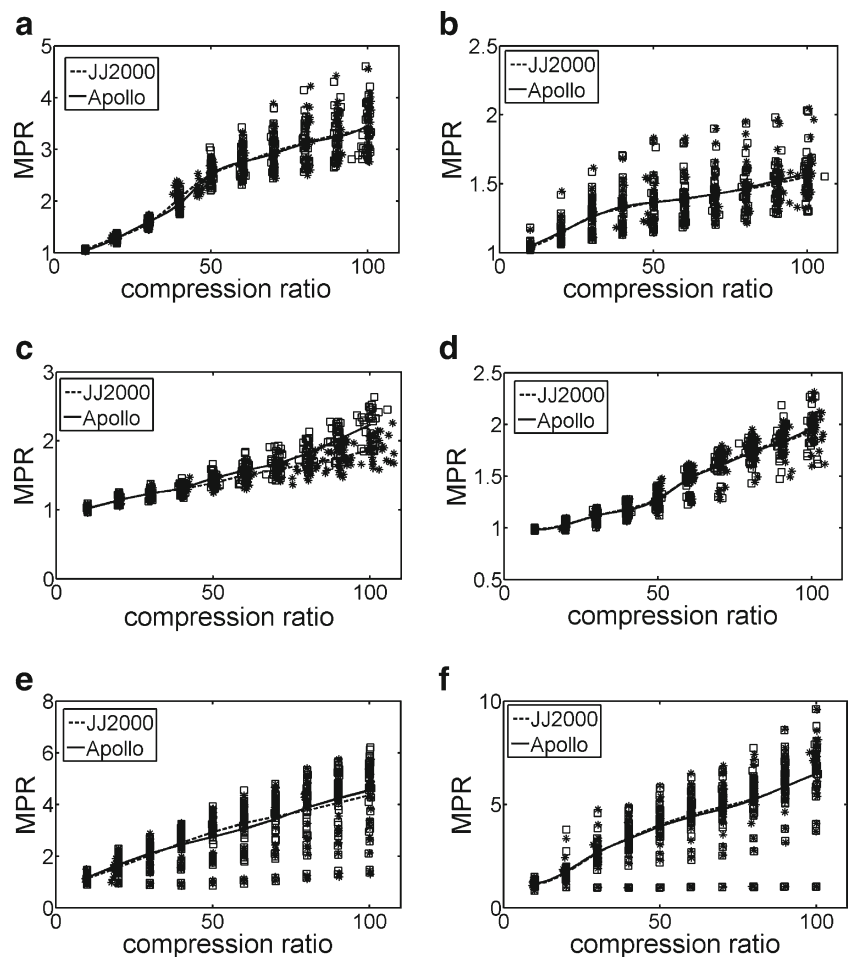


Figure 1 shows the MSE results for three modalities. The lines in this figure show the average MSE values vs. CRs, respectively. The results of CT abdominal and head images are shown in Fig. 1a and b; MR abdominal and head images, in Fig. 1c and d; and DR abdomen and chest images, in Fig. 1e and f, respectively. The line trends are obviously the same for intra-modalities and different for inter-modalities. The CT gets a lowest MSE than MR, and DR images are the highest, as a CR value is set around 50. The qualities of CT images are better than MR and DR. There is no difference in the image quality between JJ2000 and Apollo using MSE except for the MR abdominal images.

Figure 2 shows the PSNR results for all three modalities. The trends are obviously the same for these three modalities. The CT obtained the highest PSNR followed by the MR head and DR images. The MR abdominal images have the lowest PSNR as CR value around 50. The qualities of compressed images using PSNR are equal to those using MSE. Except for the

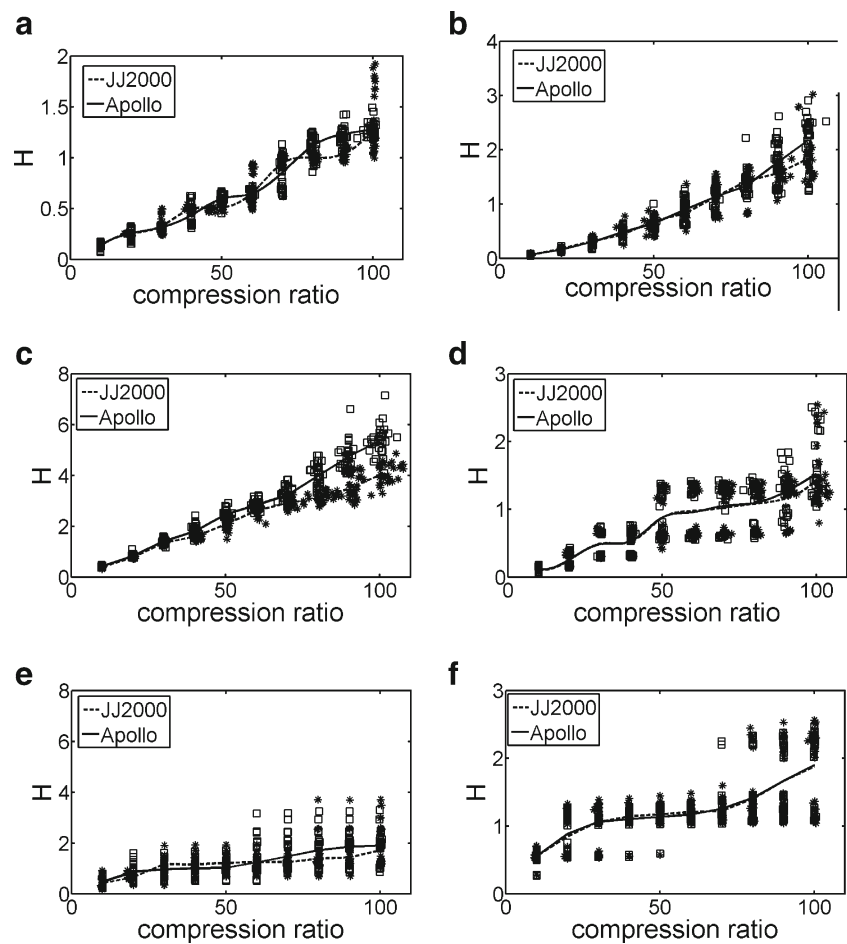
MR abdominal images, the image qualities of between JJ2000 and Apollo are the same.

The Window-Based Metrics

The window-based metrics evaluate image quality from a local region instead of using pixels because it was reported that the window approach is more like human vision than pixels in an image [23]. The *Q* and MPR were used in this study as image quality indicators to indicate how two image correlate. The higher the *Q* and the lower the MPR correlate well with better image quality.

Figures 3 and 4 show the *Q* and MPR results for three modalities. The CT head and MR images have higher *Q* value as shown in Fig. 3. This means that these images have higher qualities when compressed at some CR values (e.g., CR=50). These results are equal to those using pixel-based metrics. The MPR results coincide with those of the *Q* evaluation. The CT head and MR images have lower MPR values in Figs. 3 and

Fig. 5 Comparison of JJ2000 (broken lines, asterisk) and Apollo (straight line, square) image quality for various CRs: **a** CT abdominal, **b** CT head, **c** MR abdominal, **d** MR head, **e** DR abdominal, **f** DR chest using HVS



4. Only little differences in image quality could be found between JJ2000 and Apollo by Q and MPR.

The HVS-Based Metric

Figure 5 shows the image quality evaluation results for the HVS-based metric. The trends are obviously the same for all three modalities. The MR abdominal images get the highest H values when the CR value is set around 50. The compressed CT image quality is the best. The average H value lines of JJ2000 and Apollo obviously interlace. This means that the metric is not consistent in responding to the CR variation.

Statistical Analysis

Spearman's rank correlation coefficient is used as a measure of linear relationship between two sets of ranked data. The measurements of five metrics on two programs for 180 images are shown in Figs. 1, 2, 3, 4 and 5. The trends of averaged line looked the same but are distinguishable between these two programs. The

minimum Spearman's rank correlation coefficient (r) is 0.98, and the maximum p value is 0.001 for five metrics and three modalities for 30 measurements (six measurements for each figure). The high correlation coefficients mean that the qualities are indistinguishable between JJ2000 and Apollo.

Discussion

Image Size

No major differences were found statistically in image quality measurements between JJ2000 and Apollo using the above five metrics. The line trends are almost the same for all 120 DR and CT images and 30 MR head images, as seen in Figs. 1, 2, 3, 4 and 5. These figures show that the mean quality values for these two algorithms are almost matched. However, the MR abdominal images using JJ2000 was always slightly superior to those images compressed using the Apollo algorithm for all five metrics, as indicated in Fig. 6.

The size of the MR abdomen image is just 256×160 , which is much smaller than the other medical images. This factor may result in differences between the two compression algorithms. To verify, a MR head image was chosen randomly with an area of 256×160 cropped from this 512×512 image. Following that, this cropped image was compressed at ten different CRs before using Apollo and JJ2000, respectively. The image qualities of these compressed images were evaluated using the five metrics. The results show that JJ2000 is slightly superior to Apollo for all the metrics. This was consistent with the MR abdominal image results. The image size is a factor that affects the quality evaluation results for different image compression algorithms and may be studied in the future.

The Pixel and Window-Based Metrics

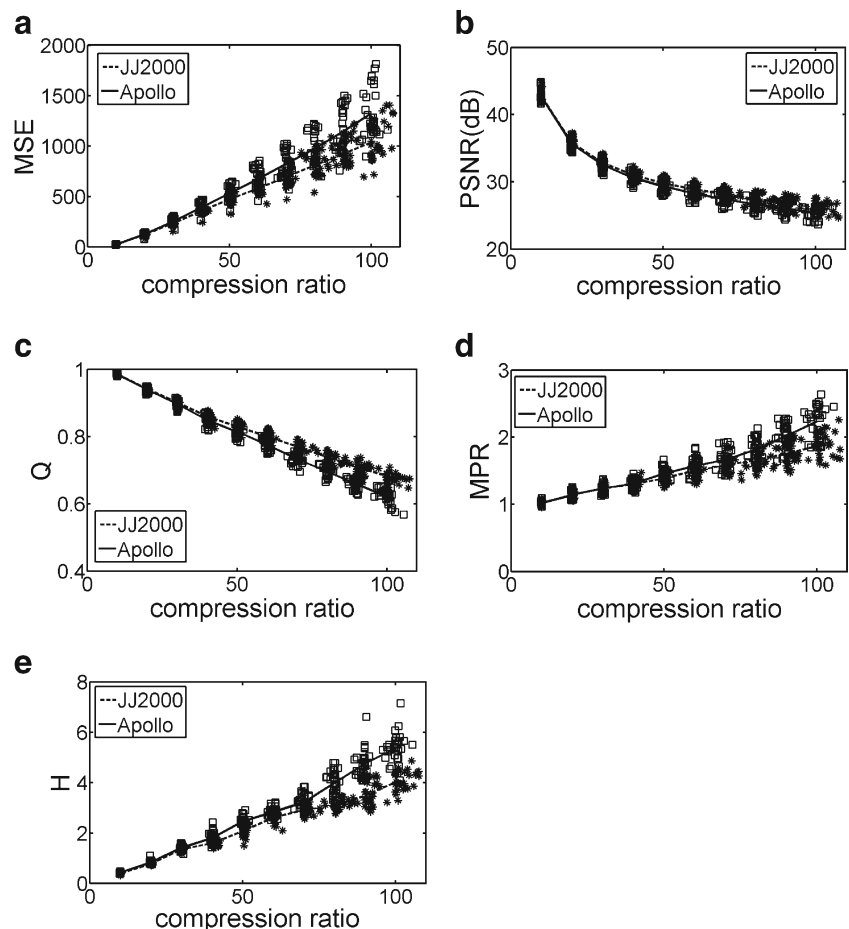
For the metrics, the line trends obviously go the same for intra-modalities. However, they are different for inter-modalities, as noted in Figs. 1, 2, 3 and 4. The line trend directions for pixel-based metrics, MSE and

PSNR, are different, but they have same pattern, as shown in Figs. 1 and 2. This effect occurs also for window-based metrics, Q and MPR, as indicated in Figs. 3 and 4. Images with equal modality show the same trends and metrics with equal bases showing the same patterns. Both MSE and PSNR measure closeness, and both Q and MPR estimate correlation. The trend directions are opposite, but the pattern is the same. Using either MSE or PSNR is equally for pixel-based metrics. For window-based metrics, Q and MPR both perform the same.

Modalities

CT, MR, and DR medical images are formed using various imaging means. Both CT and DR are produced using X-rays. However, a CT image is created by reconstruction, filtering multifarious processes to produce a tomographic image, and DR is only a projection of the total attenuation coefficient along the line of radiation. The pixel value of a MR image has only relative intensity. For each modality, the average quality evaluation lines

Fig. 6 Comparison of image quality for JJ2000 (broken lines, asterisk) with Apollo (straight line, square) for MR abdominal images in various CRs using **a** MSE, **b** PSNR, **c** Q index, **d** MPR, **e** HVS



obviously follow the same trends corresponding to each metric, respectively. This dependency is more apparent for pixel-based metrics, as noted in Figs. 1 and 2, and moderate for window-based metrics, as shown in Figs. 3 and 4. Histograms of pixel values can be produced if all of the image gray levels are arranged into bins. We chose six images from these modalities, one for each, and sorted these values to produce histograms. There are three groups that can be found apparently in these gray level histograms for modalities. Most CT pixel values are around 900–1,500. MR values are lower than 500, and DR spreads from approximately 4,000 to 15,000. The trend may result from these different gray level distributions.

Conclusions

These two JPEG2000-coded medical image compression algorithms are different in interface, convenience, speed of computation, and a number of coding options. However, the image qualities of Apollo and JJ2000 are statistically indistinguishable for medical image compression.

The differences in image quality under a metric between two programs may also be calculated using Spearman rank correlation coefficients. We found that the image quality is indistinguishable between JJ2000 and Apollo for all images evaluated using the above five metrics ($r > 0.98$, $p < 0.001$). It can be concluded that the quality is statistically equivalent for the JJ2000 and Apollo medical image compression algorithms coded in JPEG2000 with options.

Acknowledgments The authors would like to thank Prof. KS Chuang of the National Tsing-Hua University for his help in editing this paper.

References

- Huang HK: PACS—Picture Archiving and Communication Systems in Biomedical Imaging. VCH Publishers, New York, USA, 1996
- Zhang Y, Pham BT, Eckstein MP: Automated Optimization of JPEG 2000 Encoder Options Based on Model Observer Performance for Detecting Variable Signals in X-Ray Coronary Angiograms. *IEEE Trans Med Imaging* 23:459–474, 2004
- Erickson BJ: Irreversible compression of medical images. *J Digit Imaging* 15:5–14, 2002
- Kim KJ, Kim B, Mantiuk R, Richter T, Lee H, Kang HS, Seo J, Lee KH: A comparison of three image fidelity metrics of different computational principles for JPEG2000 compressed abdomen CT images. *IEEE Trans Med Imaging* 8:1496–1503, 2010
- MacMahon H, et al: Data Compression: Effect on diagnostic accuracy in digital chest radiography. *Radiology* 178:175–179, 1991
- Shiao YH, Chen TJ, Chuang K, Lin CH, Chuang CC: Quality of Compressed Medical Images. *J Digit Imaging* 20:149–159, 2007
- Kim KJ, Kim B, Choi SW, Kim YH, Hahn S, Kim TJ, Cha SJ, Bajpai V, Lee KH: Definition of Compression Ratio: Difference Between Two Commercial JPEG2000 Program Libraries. *Telemed e-Health* 14:350–354, 2008
- Rabbani M, Joshi R: Signal Processing: *Image Commun* 17:3–48, 2002
- Brennecke R, Burgel U, Rippin G, Post F, Rupprecht HJ, Meyer J: Comparison of image compression viability for lossy and lossless JPEG and Wavelet data reduction in coronary angiography. *Int J Cardiovasc Imaging* 17:1–12, 2001
- Ebrahimi F, Chamik M, Winkler S: JPEG vs. JPEG2000: an objective comparison of image encoding quality. Applications of Digital Image Processing XXVII. *Proc SPIE* 5558:300–308, 2004
- Eikelboom RH, Yogesan K, Barry CJ, Constable IJ, Tay-Kearney ML, Jitskaia L, House PH: Methods and limits of digital image compression of retinal images for telemedicine. *Investig Ophthalmol Vis Sci* 41:1916–1924, 2000
- Hui OT, Besar R: Medical image compression using JPEG2000 and JPEG: a comparison study. *J Mech Med Biol* 2:313–328, 2002
- Ismail A, Bulent S, Khalid S: Statistical evaluation of image quality measures. *J Electron Imaging* 11:206–223, 2002
- Iyriboz TA, Zukoski MJ, Hopper KD, Stagg PL: A comparison of wavelet and Joint Photographic Experts Group lossy compression methods applied to medical images. *J Digit Imaging* 12(2 Suppl 1):14–17, 1999
- Janhom A, Stelt PF, Sanderink GCH: A comparison of two compression algorithms and the detection of caries. *Dentomaxillofac Radiol* 31:257–263, 2000
- Kalyanpur A, Neklesa VP, Taylor CR, Daftary AR, Brink AR: Evaluation of JPEG and wavelet compression of body CT images for direct digital teleradiologic transmission. *Radiology* 217:772–779, 2000
- Li F, Sone S, Takashima S, Kiyono K, Yang ZG, Hasegawa M, Kawakami S, Saito A, Hanamura K, Asakura K: Effects of JPEG and wavelet compression of spiral low-dose CT images on detection of small lung cancers. *Acta Radiol* 42:156–160, 2001
- Przelaskowski A: Hybrid Vector Measures of Compressed Medical Images. *SPIE Symposium Medical Imaging: Image Perception and Performance*. 2001. http://www.ire.pw.edu.pl/~arturp/Publikacje/ap_HVM.pdf.
- Ricke J, Maass P, Lopez HE, Liebig T, Amthauer H, Stroszczyński C, Schauer W, Boskamp T, Wolf M: Wavelet versus JPEG (Joint Photographic Expert Group) and fractal compression. Impact on the detection of low-contrast details in computed radiographs. *Investig Radiol* 33:56–463, 1998
- Slone RM, Foos DH, Whiting BR, Muka E, Rubin DA, Pilgram TK, Kohm KS, Young SS, Ho P, Hendrickson DD: Assessment of visually lossless irreversible image compression: comparison of three methods by using an image-comparison workstation. *Radiology* 215:543–553, 2000
- Chen TJ, Chuang KS, Chang JH, Shiao YH, Chuang CC: A blurring index for medical images. *J Digit Imaging* 19:118–125, 2006
- Chen TJ, Chuang KS, Jay W, Chen SC, Hwang IM, Jan ML: A novel image quality index using Moran I statistics. *Phys Med Biol* 48:131–137, 2003
- Wang Z, Bovik AC: A universal image quality index. *IEEE Signal Process Lett* 9:81–84, 2002
- Cliff AD, Ord JK: *Spatial process: models and applications*. Pion, London, UK, 1981