J Digit Imaging (2015) 28:636–645
DOI 10.1007/s10278-015-9791-7

Computed Tomography Image Compressibility and Limitations of Compression Ratio-Based Guidelines

Jean-François Pambrun [1] · **Rita Noumeir** [1]

Published online: 25 March 2015

Abstract Finding optimal compression levels for diagnostic imaging is not an easy task. Significant compressibility variations exist between modalities, but little is known about compressibility variations within modalities. Moreover, compressibility is affected by acquisition parameters. In this study, we evaluate the compressibility of thousands of computed tomography (CT) slices acquired with different slice thicknesses, exposures, reconstruction filters, slice collimations, and pitches. We demonstrate that exposure, slice thickness, and reconstruction filters have a significant impact on image compressibility due to an increased high frequency content and a lower acquisition signal-to-noise ratio. We also show that compression ratio is not a good fidelity measure. Therefore, guidelines based on compression ratio should ideally be replaced with other compression measures better correlated with image fidelity. Value-of-interest (VOI) transformations also affect the perception of quality. We have studied the effect of value-of-interest transformation and found significant masking of artifacts when window is widened.

Keywords Image compression · Image artifact · Image quality · Image visualization · JPEG2000 · Computed tomography · Exposure · Slice thickness · Filter type

✉ Jean-François Pambrun
jf.pambrun@gmail.com

1 Department of Electrical Engineering, École de Technologie Supérieure (ETS), 1100 Notre-Dame Ouest, Montréal, Québec, Canada

Introduction

We reasonably expect instant access to a wealth of information. With the Internet and cloud computing, we are also used to very efficient collaboration mobile applications. However, health-care information exchange is very slowly following this trend. Patients' records are still commonly handled manually and spread across multiple institutions. As a result, records are not readily available or are incomplete; patients may be required to repeat exams, which causes treatment delays and reduces productivity.

Being aware of the financial and health implications, many authorities around the world started laying groundwork for an electronic health record (EHR) that will be universally accessible and readily available and contains information relevant to all aspects of patient care: demographics, contact information, allergies, intolerances, laboratory results, diagnostic imaging, pharmacological and immunological profiles, etc. Achieving this will require tremendous resources. In Canada, for instance, the cost of providing a pan-Canadian Electronic Health Record for each one of its 35 million citizens is expected to be over $3.5 billion [1].

The implementation of high-capacity redundant data centers and the deployment of robust network infrastructures are some of the factors that contribute to such high costs. This is mostly due the vast amount of data produced every day by modern diagnostic imaging devices. For instance, computed tomography (CT) can generate image stacks containing thousands of slices that can weigh more than a gigabyte. Moreover, these images need to be archived for a very long time, usually until the patient's death, and remain readily available throughout his/her life.

This issue can be mitigated with the use of data compression. Images can be losslessly compressed by up to two thirds. Compressing to a greater extent is desirable to further reduce bandwidth and storage requirements, but lossy compression introduces artifacts and distortions that, depending on their levels, can alter diagnostic accuracy and may interfere with image processing techniques used in computer-aided diagnostic applications [2].

Estimating the impacts of these distortions is very difficult. Images with seemingly similar characteristics that are compressed using identical compression parameters can result in very different reconstruction fidelity; some can preserve all their diagnostic qualities, while others may become completely unusable. Because of liability issues raised by possible diagnostic errors caused by lossy compression, radiologists generally are not inclined to use compression techniques that would produce visually lossy results [3]. Compression guidelines were introduced to enable the use of lossy compression, but variations [4, 5] in image compressibility suggest that broad guidelines allow only for conservative and suboptimal compression.

The term fidelity is used throughout this paper to quantify the accuracy of the reconstruction. On the other hand, image quality depends on the subjective perception of an observer and his/her ability to perform a specific task. Image quality can also be improved with image processing techniques.

Most research on this topic was aimed at finding the maximum safe compression ratios for a given modality or organ in order to propose guidelines for practitioners. Conversely, our objective with this paper is to identify and raise awareness on the limitations inherent to the reliance on compression ratios to characterize image fidelity. To achieve this, we will study the impact of image content and five acquisition parameters on the compressibility of the computed tomography of a lung phantom and we will show that those factors are more closely related to fidelity than the compression ratio itself. We will also investigate the relation between CT acquisition parameters and noise in addition to analyzing how they affect fidelity after compression. Finally, we will examine the effects of different value-of-interest (VOI) transformations commonly used to adapt the high dynamic range of medical images to the limited range of most displays.

Previous Work

In 2006, a survey of radiologists' opinions on compression [3] revealed that lossy compression was already used in the USA for both primary readings and clinical reviews, while Canadian institutions remained much more cautious about irreversible compression. In the USA, two radiologists out of five reported using lossy compression before primary reading and all reported using lossy compression for clinical reviews. The compression ratios that they used ranged between 2.5:1 and 10:1 for CT and up to 20:1 for computed radiography. Comparatively, only three Canadian radiologists out of six reported using lossy compression at all and only one reported using compression before primary reading. Furthermore, two of them declared using compression ratios between 2.5:1 and 4:1, which are effectively, or very close to, lossless levels. Almost every radiologist expressed concerns regarding litigations that could emerge from incorrect diagnostic based on lossy compressed images, and all was aware that images from different modalities require different compression ratios. In view of that, most radiology departments from the USA had conducted their own tests to establish visually lossless compression ratios with the assumption that imperceptible distortions cannot impair diagnostic accuracy in any way.

This task of finding visually lossless compression thresholds is usually done by asking trained radiologists whether pairs of unaltered and compressed images are identical. These studies are structured as a two-alternative forced choice (2AFC) experiments where observers can either examine both images side-by-side or alternate between both images to determine if the distortion is perceivable or not. This exercise is repeated with many images compressed at different compression ratios to find a visually lossless threshold for a given modality and/or organ system. Images compressed with compression ratios below this threshold are then assumed diagnostically lossless. Interestingly, while performing these experiments, some researchers [6–8] noticed that when radiologists could perceive differences between both images they often preferred the compressed version. A possible explanation for this is that when compression ratios are increased beyond visually lossless thresholds, acquisition noise is significantly attenuated before the signal itself. This is supported in [8, 9] by the absence of structures in difference images from visually lossy image pairs indicating that noise is likely lost before any diagnostically important information. This suggests that it might be desirable to compress diagnostic images beyond visually lossless levels.

The impacts of compressibility differences between different modalities on image fidelity are widely known [2, 3, 9], but as variations within the same modalities are not as widely acknowledged, guidelines are often only defined on a modality basis. However, these variations can be fairly significant. As an example, tomographic images of the chest walls are far less tolerant to compression than those of the lung [5] and thinner slice thicknesses are known to have adverse effects on compressibility [10]. Because of this, recommendations from the Canadian Association of Radiologists (CAR) specify different compression ratios for different anatomical regions and CT scans are divided in six subtypes (angiography, body, chest, musculoskeletal, neuroradiology, and pediatric) each with different compression ratios. However, these recommendations ignore key acquisition parameters that may have substantial impacts on compression such as reconstruction kernel [9] and slice thickness [11] that are known to reduce compressibility. Researchers in [9] observed a relationship between compressibility and the relative importance of the energy of the lower subbands in the wavelet domain. Because acquisition parameters are linked to lower subbands' energy levels, they concluded that compression ratio recommendations should not be developed on a modality or organ system basis. Compressibility variations within images have also been observed in [12] with a regional difference between the lungs, chest wall, and mediastinum. Interestingly, they noted that while the lung had lower peak signal-to-noise ratio (PSNR), it had higher perceptual rating. More recently, they have tried [13] to predict the perceived image quality using only parameters extracted from Digital Imaging and Communications in Medicine (DICOM) headers and found that compression ratio and slice thickness are the two best predictors. Unfortunately, they have limited their model to these two variables even if other parameters are known to be correlated to compressibility.

In an effort to foster the use of image compression in diagnostic imaging applications, researchers conducted a large-scale pan-Canadian study [14] on irreversible compression for medical applications. It involved 100 staff radiologists analyzing images from several modalities. Images were compressed using multiple compression ratios, and each pair was rated using a six-point scale. Diagnostic accuracy was also evaluated by requiring radiologists to examine images of known pathologies. As a result, guidelines based on compression ratio were proposed for computed radiography, computed tomography, ultrasound, and magnetic resonance, but the effects of acquisition parameters were ignored and slice thickness for CT scans was restricted to 2.5 mm and higher. This work resulted in the recommendations on irreversible compression [15] that have been published by the CAR and are used today in Canada.

Recommendations like those from well-established organizations are essential, but compression ratio, on which these ratios are usually based, is poorly correlated with image fidelity [3] because deterioration levels depend highly on image information [16] and noise [17]. The CAR acknowledged this by providing different guidelines for different scenarios, but it is still only a coarse approximation and image fidelity cannot be guaranteed for a given compression ratio. Therefore, fidelity metrics should be used instead of compression rate in medical application [16]. Furthermore, differences in coder implementations [5] can produce different results even when using identical target compression ratio. Most JPEG2000 coders use the mean-squared error to regulate compression, but this is not a requirement of the standard, which is completely open to other implementations that could produce completely different outcomes [18]. Moreover, different codec vendors use different compression ratio definitions, either based on stored or allocated bits, resulting in 25 % differences [19]. The CAR does not specify which definition should be used with its recommendations, and even if they did, radiologists would probably be unaware of the implementation used by their software. Because of all these factors, standardization of image quality or fidelity measurement and compression parameters for clinical applications is desirable [20].

Another issue specific to medical imaging is the high dynamic range. Diagnostic images usually have more than 255 (8 bits) gray levels. Visualization systems cannot display such a wide range, but these images can be dynamically adapted using VOI transformations [21] that can be manipulated by the clinician in order to explore a different grayscale window. Papers such as [4, 11, 22, 23] on diagnostic image quality have used fixed values of interest for their evaluation. However, diagnosis may require unpredictable settings [3] and a narrower window can make distortions more apparent, while a wide widow can mask them. Observers should be allowed to freely modify [14] the value-of-interest setting as they would in their practice; otherwise, their observation may be skewed. Another option is to allow customization within a reasonable range as in [24, 25]. Imposing a lower limit on window width eliminates the case where a single and, otherwise, invisible artifact is amplified and becomes obvious because the window width is narrower than the display range.

Methodology

As stated above, the existence of significant compressibility differences between imaging modalities and the variations due to noise levels is well known. However, to our knowledge, the impact of CT acquisition parameters, analyzed individually, has never been thoroughly studied. This is the gap that this experiment seeks to demonstrate.

Data

Our objective is to study two challenges related to diagnostic image compression: the compressibility variations in computed tomography caused by different acquisition parameters and the impact of window width on the perception of compression artifacts. We have restricted our study to computed tomography because it is known to be poorly compressible and generates an increasing amount of data. To achieve that objective, multiple series of the same region of the same subject, acquired with different acquisition parameters, are needed.

Acquisition parameters are available from each image DICOM header. However, different implementations inconsistently report these parameters. Exposure, for instance, is not consistently reported across different devices and reconstruction filters may not have any direct equivalent for different hardware configurations. Moreover, the field of view and subject size, when not kept constant, make comparative evaluation very difficult. For example, when the field of view is increased for a specific subject, the easily compressible black background fills a larger image region resulting in increased compressibility. For these reasons, we used multiple series of the same subject acquired with the same equipment to ensure that acquisition parameters are consistently reported.

Fortunately, the National Cancer Institute made many diagnostic image collections publicly available to encourage and support cancer research through their Cancer Imaging Archive project. One of these collections, labeled *phantom FDA* [26], perfectly fits the requirements of our experiment. It was developed in an effort to evaluate the effects of acquisition parameters on the accuracy of automated lung nodule size estimation algorithms used in computer-aided diagnostic solutions. To meet their requirements, the researchers repeatedly scanned an anthropomorphic thoracic phantom with synthetic lung nodules using different acquisition parameters. These parameters are presented in Table 1. Parameters include five slice thicknesses varying from 0.8 to 5 mm, three effective exposures from 25 to 200 mAs, two slice collimation configurations, two different pitches, and two types of reconstruction filter. Two different nodule layouts were made available through The Cancer Imaging Archive, and we have selected all series, each with a different parameter combination, of the nodule layout labeled *2*. That is 23,767 individual images across 72 series. Slice thickness depends on slice collimation, and only 3-mm-thick slices can be acquired with both collimator configurations. All the series were acquired using a Philips 16-row scanner (Mx8000 IDT; Philips Healthcare, Andover, MA), and precautions were taken to preserve a constant positioning of the phantom between acquisitions. Figure 1 shows six images of the phantom with their slice locations displayed in the upper left corner. The scanned area spans about 30 cm with the slice location ranging from 90 to 389 mm. Slices were acquired with a slice overlap of 50 %; the thinnest acquisitions (0.8 mm) had a spacing of 0.4 mm and contained 750 images, while the thickest (5 mm) series had a slice spacing of 2.5 mm and contained only 120 slices.

Compression

We have compressed each image with JPEG2000 using multiple compression ratios including lossless, 4:1, 5:1, 6:1, 8:1, 10:1, 15:1, and 30:1. The wide range of compression ratios covers ratios that are normally used, except for 30:1 which is twice the CAR-recommended ratio for CT. Our compression ratio is calculated using the allocated file size including headers; the codec used is an open source JPEG2000 implementation [27]. The software was able to compress high-dynamic-range images.

Table 1 Acquisition parameters

Parameter	Values
Slice thickness (mm)	0.8, 1.5, 2, 3, 5
Effective dose (mAs)	25, 100, 200
Filter type	Detail, medium
Slice collimation (mm)	16×0.75, 16×1.50
Pitch (mm)	0.9, 1.2

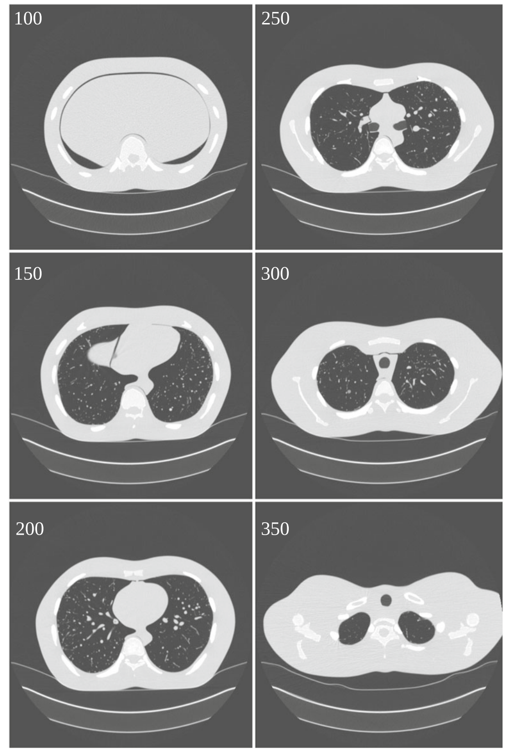

Fig. 1 Image content relative to slice location. *The number displayed in the upper left corner of the image* indicates the slice location

Fidelity Evaluation

The fidelity of every compressed image is evaluated: using maximum absolute difference, (1) mean-squared error (MSE), and (2) PSNR. Maximum absolute difference is the absolute difference of the most altered pixel using the compression process.

MSE is computed with

$$\mathrm{MSE} = 1/(MN)\Sigma_i \Sigma_j [I_0(i,j) - I_c(i,j)]^2 \quad (1)$$

where I_0 is the original image and I_c is the compressed image.

PSNR is computed with

$$\mathrm{PSNR} = 20\log\left(I_{\mathrm{range}}/\sqrt{\mathrm{MSE}}\right) \quad (2)$$

where I_{range} is the range of the signal; therefore, PSNR is the signal-to-noise ratio in decibels. We have calculated the range of the signal in all images and found it to be 1600. Although the bit allocated was 16 and bits stored were 12, suggesting a dynamic range of 4096, we have used 1600 for I_{range} to compute PSNR values.

Compressibility Evaluation

Compressibility can be evaluated by

1. Observing the file size after lossless compression compared to the uncompressed file size
2. Comparing the relative image fidelity of the two different compressed images with the same compression ratio.

With the first measure, if one image has a smaller file size, we can conclude that it is more compressible. With the second measure, if one image has a higher PSNR or conversely a lower MSE than another one, we can also conclude that it is more compressible.

Most JPEG2000 coder are designed to minimize MSE (maximize PSNR) for a specified target compression ratio. As a result, both proposed measures are equivalent. This is illustrated in Fig. 2: the PSNR of all 23,767 images compressed at 4:1, 5:1, 8:1, 15:1, and 30:1 plotted against their respective lossless file size. The relation is linear except for images compressed at 4:1 with lossless file size below 128 kB because these images could have been compressed losslessly using reversible filter banks. Naturally, fidelity decreases for a given ratio when the lossless file size increases.

Statistical Analysis

In order to evaluate the impact of each acquisition parameter on compressibility, we have used the R software [28] to perform statistical analysis and to fit models. Fitted models are evaluated with the coefficient of determination (R^2), root mean-squared prediction error (PE), and Pearson correlation coefficient (CC).

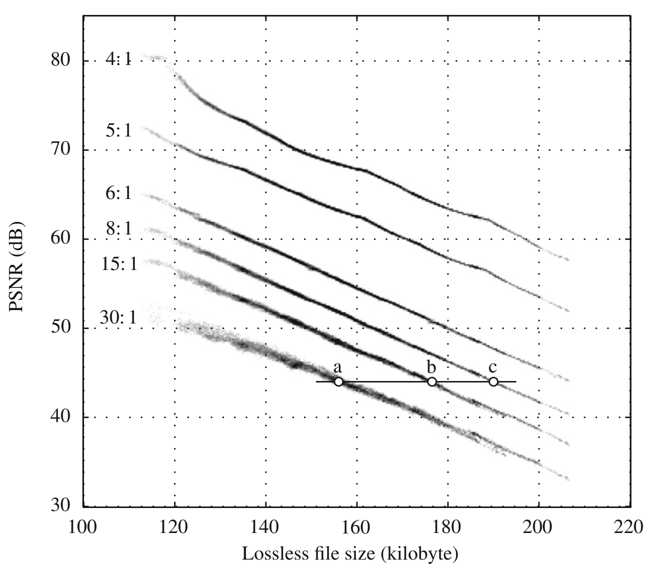

Fig. 2 PSNR of lossy compressed image plotted against lossless file size. *Each point* represents the PSNR of an image compressed at a specific lossy compression ratio. This PSNR is plotted against the lossless size of that image. PSNR is directly correlated to the lossless compression image size

A linear regression was performed between PSNR of images compressed at 8:1 and their corresponding lossless file sizes. The model is extremely well fitted (R^2=0.99, PE=0.13 dB, CC=0.99), indicating that both the PSNR at a fixed compression ratio and the lossless file size can be used interchangeably to estimate compressibility.

Results

Impacts of Image Content

From Fig. 2, we note that

1. For a specific compression ratio, compressibility varies for more than 20 dB for different images, suggesting that this variation is due to image content.
2. For a given image, the fidelity decreases by only 3 dB when a compression ratio is decreased from 6:1 to 8:1 or from 8:1 to 15:1 or from 15:1 to 30:1.

This suggests that image content, defined by slice location and acquisition parameters, has a more significant impact on image fidelity than compression ratios.

Figure 2 shows that 15 % of images compressed at 15:1 (point b) have a fidelity lower than the median of those compressed at 30:1 (point a); likewise, 4 % of those compressed at 8:1 (point c) have a lower fidelity than the median of those compressed. In other words, images with lossless file sizes smaller than 155 kB, compressed at 30:1, are less degraded than images with lossless file sizes larger than 190 kB, compressed at 8:1.

Figure 3 shows the size of each losslessly compressed image, plotted against slice location, for all 23,767 images. Each series is displayed using a curve with different gray levels. Series are acquired with different acquisition parameters. Compressibility variations between series are very important. For the same slice location and subject, the average lossless file size was 116 kB in the best case and 193 kB in the worst case, a 66 % difference.

Compressibility variations along the subject are also apparent. Every series exhibits a similar behavior with respect to slice location, and adjacent images from the same series have similar compressibility.

The maximum absolute difference between the original and compressed images for the most damaged pixel is displayed in Fig. 4. With images compressed at 15:1, the maximum absolute error varies by a factor of 10; 10 % (above point b) of the images compressed at 15:1 and 4 % of the images compressed at 8:1 (not shown) have higher maximum absolute error than the median of those compressed at 30:1 (point a).

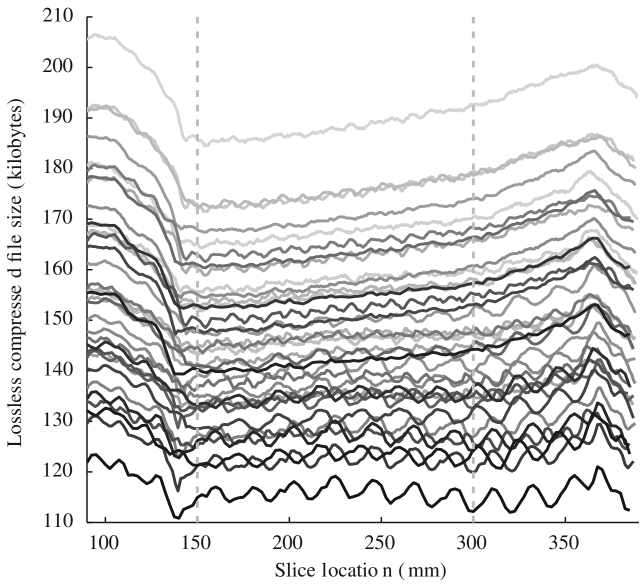

Fig. 3 Lossless file size shown with respect to slice location. *Each curve* represents one series. Two consecutive images from the same series have very similar compressibility. For a specific series/curve, compressibility varies with slice location. Between locations 150 and 300, compression is best because noise is less

Impacts of Acquisition Parameters

Impacts on Prediction

We have shown that the image content as well as the acquisition parameters have a significant impact on compressibility without identifying which one of the acquisition parameters has the most significant impact. In the dataset that was used, five parameters were varied between each acquisition; our objective here is to study the impact of each acquisition parameter, such as exposure. We have grouped the images in subsets of equal exposure. In our case, we have three groups of images: (1) acquired with 25 mAs, (2) acquired with 100 mAs, and (3) acquired with 200 mAs. For each group of images, we have measured the file size. We show in Fig. 5 where the boxes are centered at the mean and extend between the 25th and 75th percentiles.

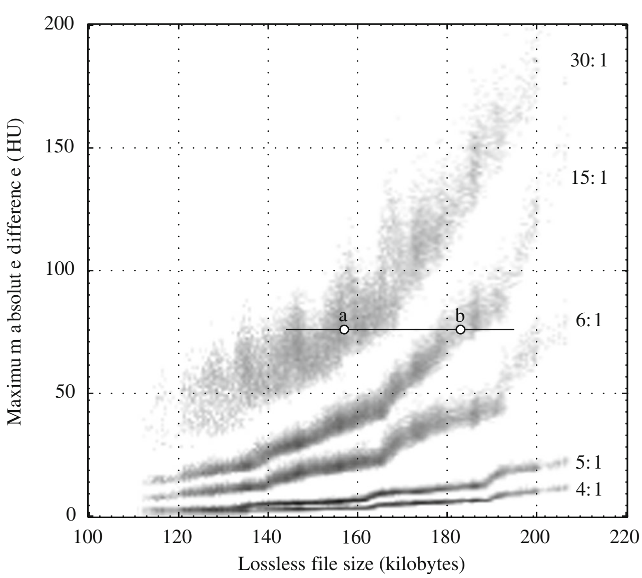

Fig. 4 Maximum absolute difference of lossy compressed image plotted against lossless file size

When exposure increases from 25 to 100 mAs, the boxes do not overlap. This suggests that exposure has a definitive impact on compressibility. When exposure increases, the file size decreases and compressibility increases.

Images have been grouped into subsets of equal thicknesses; box plots for thicknesses of 0.8, 1.5, 2, 3, and 5 mm are shown in Fig. 5b. It is clear than when thickness increases, so does compressibility.

Images where divided into two groups according to filter type: medium and detail. Images processed with the *medium* filter contain less noise but have lower spatial resolution. It is clear from Fig. 5c that compressibility is increased with the use of the medium filter.

Images are separated in two subsets according to slice collimation: 16×0.75 and 16×1.5 mm. Figure 5d suggests that when slice collimation is decreased, compressibility increases. Finally, images are separated into two groups according to pitch: 0.9 and 1.2 mm. Figure 5e suggests that pitch has no effect on compressibility.

Figures 5a, e shows box plots on the impact of each one of these five parameters on lossless file size. These plots clearly indicate that there is a link between exposure, thickness, filter type, slice collimation, and compressibility. Pitch, on the other hand, seems to have little effect. In fact, *z* testing indicates that the means of both groups are statistically identical and that pitch does not have any statistically significant impact on compressibility. This may appear counterintuitive, and it will be discussed later.

Impacts on Fidelity

Figure 6 shows histograms of PSNR differences between images at the same location taken from two series acquired while varying one single parameter. The reference series was acquired with an exposure of 200 mAs, a slice thickness of 5 mm, and a medium filter. This series corresponds to the best possible compressibility in our dataset. Figure 6a shows the impact on compressibility when reducing exposure from 200 to 100 mAs (dark gray) and from 200 to 25 mAs (light gray). Figure 6b shows the impact on compressibility when reducing the thickness from 5 to 3 mm, from 5 to 2 mm, from 5 to 1 mm, and from 5 to 0.8 mm. Figure 6c shows the impact on compressibility when changing the filter from detail to medium.

With this dataset, we observe the following:

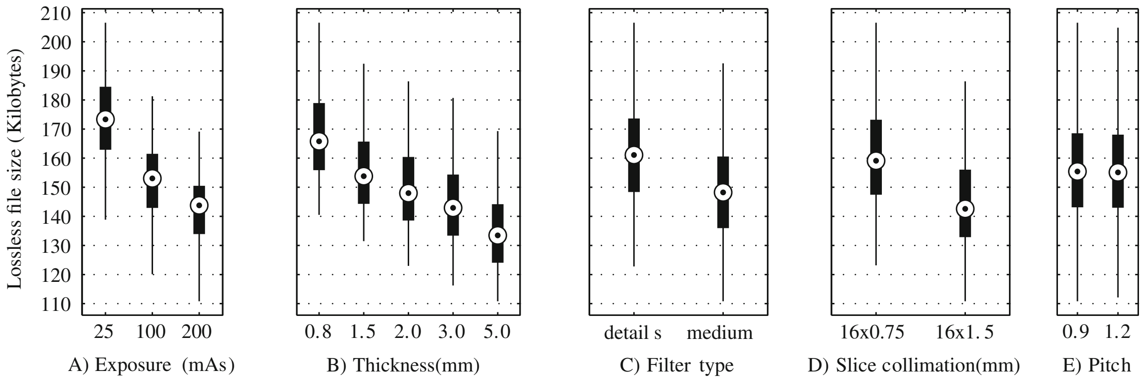

Fig. 5 Box plots using all 23,676 images. *Boxes* are located at median and extend from the 25th to the 75th percentiles, and whiskers extend to the most extreme data points that were not considered outliers

1. A 7 dB reduction in fidelity when exposure is reduced from 200 to 25 mAs and a 2 dB reduction when exposure is reduced from 200 to 100 mAs

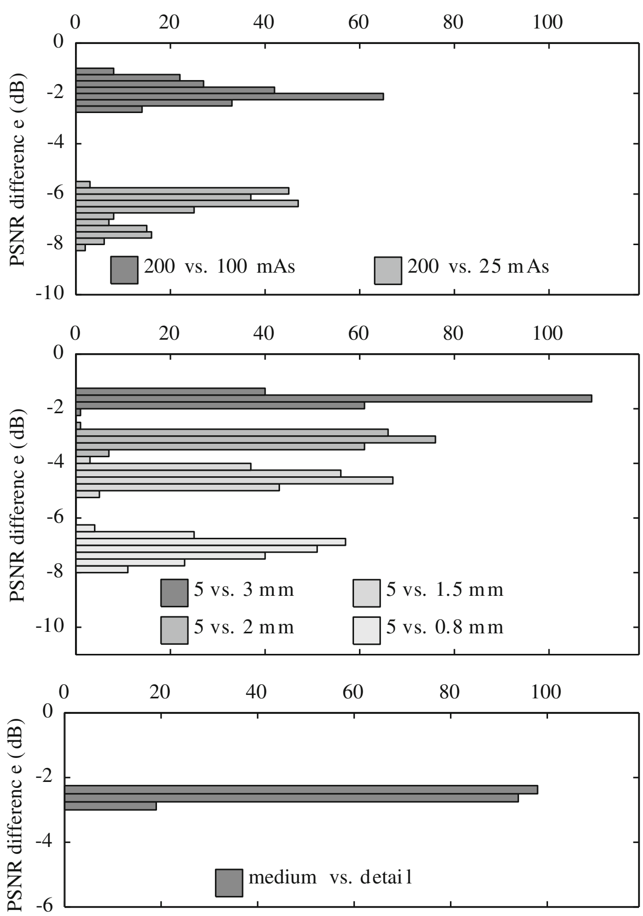

Fig. 6 The reference series was acquired with an exposure of 200 mAs, a slice thickness of 5 mm, and a medium filter. This series corresponds to the best possible compressibility in our dataset. **a** The impact on compressibility when reducing exposure from 200 to 100 mAs (*dark gray*) and from 200 to 25 mAs (*light gray*) on compressibility. **b** The impact on compressibility when reducing thickness from 5 to 3 mm, from 5 to 2 mm, from 5 to 1 mm, and from 5 to 8 mm. **c** The impact on compressibility when changing from detail to medium filter

2. A 7 dB reduction in fidelity when slice thickness is reduced from 5 to 0.8 mm, a 3 dB reduction when thickness is reduced from 5 to 2 mm, and a 5 dB reduction when thickness is reduced from 5 to 1.5 mm
3. A 2.5 dB reduction in fidelity when the detail filter is used instead of the medium filter.

Relative Importance of Each Parameter

To evaluate the relative importance of each acquisition parameter on compressibility, we have fitted a quadratic model to predict the PSNR of images compressed at 8:1 using following equation (3):

$$\begin{aligned} \mathrm{PSNR} \sim B_i + B_f \times \mathrm{Filter} + B_c \times \mathrm{Collimation} + B_e \times \mathrm{Exposure} \\ + B_t{*}\mathrm{Thick} + {B_e}^2 \times \mathrm{Exposure}^2 + {B_t}^2 \times \mathrm{Thick}^2 \end{aligned} \tag{3}$$

Because each acquisition parameter does not have the same distribution in terms of average and standard deviation, we have normalized each beta variable in Eq. 3 by subtracting its mean and dividing by its standard deviation. Using a normalized predictor, the quadratic model can be represented according to the beta coefficients shown in Table 2. The model is a well-fitted model with a coefficient of determination (R^2) of 0.94, a prediction error of 1.05 dB, and a Pearson correlation coefficient of 0.97. Beta values provide an estimation of the relative importance of each parameter. Moreover, when

Table 2 Beta coefficient for predicting PSNR when compressed at 8:1

	Beta coefficient values
Intercept (B_i)	0.44
Exposure (B_e)	0.73
Slice thickness (B_t)	0.68
Filter type (B_f)	−0.34
Slice collimation (B_c)	−0.05
Exposure^2 (${B_e}^2$)	−0.31
Slice thickness^2 (${B_t}^2$)	−0.13

considering only images located between slice locations at 150 and 300 mm, the quadratic model is even better, cutting prediction errors by half.

B_e and B_t, being larger, suggest that exposure and thickness have the most significant impact on compressibility, followed by filter type and slice collimation. Because of the bias introduced by the covariance between predictors [29], other methods were developed to evaluate the contribution of each predictor to R^2. By using the proportional marginal variance decomposition (PMVD) [30], we have found 53 % of the prediction provided by exposure, 34 % from slice thickness, and 13 % from filter type. We have also found that slice collimation has no effect on compressibility by itself. The covariance between collimation and slice thickness is high because collimation of 16×0.75 mm has been used to acquire series with slice thicknesses of 0.8, 1.5, and 3 mm; likewise, collimation of 16×1.5 mm has been used only with thicknesses of 2, 3, and 5 mm.

We have fitted another model that includes a compression ratio as a predictor, in order to compare the impact of compression ratios with other acquisition parameters. To fit the model, we have considered compression ratios that are usually used with CT images: 6:1, 8:1, 10:1, and 15:1. We added two terms for a compression ratio to Eq. 3: one linear and one quadratic.

Analyzing each predictors' relative importance with PMVD, we found that compression ratio can only explain 28 % of the PSNR variations while exposure explains 38 %, slice thickness 25 %, and filter type 9 %. Therefore, with this dataset, acquisition parameters affect the compression fidelity significantly, more so than compression ratio.

Impacts of Noise

Exposure and slice thickness are directly related to noise in computed tomography. Noise is a key factor in image compression. Noisy images are hard to compress because they produce many small uncorrelated coefficients in high-frequency wavelet subbands.

In our experiment, noise was estimated for each series by calculating the variance within a uniform region of the first slice. This uniform region of 208 by 94 pixels represents an area of the phantom molded in urethane with a constant Hounsfield unit value. Noise alone is a fair predictor (R^2= 0.85, PE=7.3 kB) of image compressibility, and it is much more accurate than any other single predictor. Using exposure, thickness, filter type, and slice collimation to predict noise yields a good fit (R^2=0.90, CC=0.95). PMVD reveals that exposure explains 67 % of noise in our highly controlled model, slice thickness 27 %, and filter type 6 %.

Noise was added as a predictor to the quadratic model in Eq. 3. The quality of the model was not significantly improved because noise and the other predictors are highly correlated. Commonality analysis [29] is used to identify the unique contribution of every single parameter and the common contribution of every possible combination of parameters to R^2. It provides separate measures for the explained variances of each individual parameter as well as measures for the shared variance of all combinations of parameters. It is mostly useful when the regression contains significant multicollinearity and suppressions as is the case with this model. Commonality analysis measures always sums to R^2 which is 0.92 in this case. Table 3 shows each contribution in percentage of R^2. Exposure, slice thickness, filter type, and noise each uniquely accounts for less than 5 % of the compressibility variance, while noise and exposure commonly account for 45 %, noise and slice thickness for 26 %, and noise and filter type for 8 %. Slice collimation has no effect on compressibility but is highly co-dependent on slice thickness. Noise, slice collimation, and slice thickness together account for 15 % of the total variance.

Impacts of Window/Level Transform on Image Fidelity

Image visualization requires a "window and level" transformation in order to select parts of the pixel dynamic range to display. Standard ranges of values of interest (VOIs) are defined for specific tasks and anatomical regions. CT values are shifted and scaled to create presentation values (p values) that fit the dynamic range of the display. These different VOI settings affect the image fidelity by masking coding artifacts. To illustrate this phenomenon, three common VOI settings were used to transform CT values into p values: (1) abdomen, centered on 60 Hounsfield units (HU) with a window width of 400 HU; (2) lung, centered on −500 HU and spanning at 1500 HU; (3) bone, centered on 750 HU and spanning at 3500 HU. Figure 7 shows the PSNR computed on the p values

Table 3 Commonality analysis

	Total
Unique to exposure	3.36
Unique to thickness	3.63
Unique to collimation	0.27
Unique to filter type	4.81
Unique to noise	5.37
Thickness and exposure	−2.23
Thickness and collimation	2.26
Exposure and filter type	−2.18
Noise and exposure	45.46
Noise and thickness	11.25
Noise and filter type	8.27
Noise, thickness, and collimation	15.02
Noise, exposure, and thickness	1.36
Exposure, thickness, and filter type	2.27
Exposure, filter type, and noise	2.20
Total	100.00

Entries with small contribution (<1 %) were removed

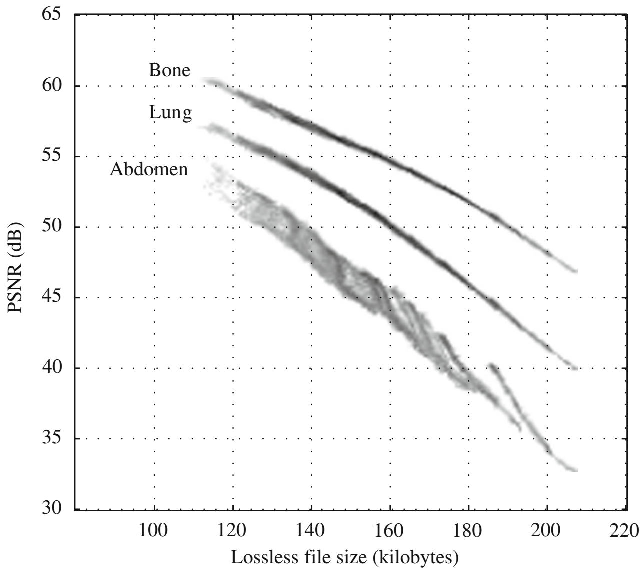

Fig. 7 *Each point* represents the PSNR computed on presentation values obtained after applying the grayscale window transform against the lossless file size for a specific image. The image displayed with the abdomen window shows lower fidelity, while those presented with the bone VOI appears to have higher fidelity

plotted against lossless file size. The display range (I_{range}) considered was 256. When the abdomen is displayed with 256 gray levels, distortions are attenuated by a factor of 1.5. In that case, a CT value difference of 3 would show up as a p value difference of 2. On the other hand, distortions that occur outside this range, where HU values are clamped to either 0 or 255, would become completely invisible. The window width used to visualize lung is large, more than six times the display range. Only large distortions can be noticed. The bone window is even larger, 14 times the display range. Consequently, distortions are significantly masked. Therefore, narrow windows can accentuate distortions, while wide windows can significantly underestimate them. This should be carefully considered when designing metrics or observer-based fidelity studies.

Discussion

Our results have shown that noise is a key factor in image compressibility. The quantum noise found in computed tomography images is governed by the Poisson statistic law, and the signal-to-noise ratio (SNR) is proportional to the squared root of N, the number of photons [5]. With all other acquisition parameters kept constant, the number of photons generated by the X-ray source is directly proportional to current and time product, in milliampere second, called exposure. Increasing this parameter by a factor of 2 causes a 41 % increase in SNR. This relation holds for slice thickness as well since the number of photons reaching the X-ray detectors is directly proportional to the detector size. Because of noise, compressibility is increased with exposure and slice thickness. Moreover, high-frequency details in the image are harder to compress and are attenuated by the averaging over a larger region along the z-axis, which increases with slice thickness.

In multi-slice CT scanners, the pitch is defined as the table feed for each complete revolution of the X-ray detectors and source. A pitch of 1 indicates a table feed equivalent to the width of the detector array. If speed or coverage is needed, images can be reconstructed with less than a full rotation, resulting in pitches higher than 1. Conversely, slices reconstructed with pitches lower than 1 are reconstructed with more than one revolution [21], resulting in increased exposure. Consequently, with all other parameters kept constant, the number of photons emitted per slice is inversely proportional to the pitch. Therefore, increasing the pitch introduces more noise and reduces compressibility. However, multi-slice scanner manufacturers usually use an alternative definition of exposure that takes pitch into account—effective exposure. Effective exposure, or mAs per slice, allows radiologists to estimate acquisition signal-to-noise ratio with fewer parameters. As a result, to keep effective current constant in our experiment, the X-ray tube current was increased by 30 % when the pitch was increased from 0.9 to 1.2.

Reconstruction filters are not standard across manufacturers. However, detail filters usually accentuate high frequencies while increasing noise. CT images of bones have high contrast and can benefit from sharper fine details without suffering from significant increase in noise levels. On the other hand, soft tissues have lower contrast and it is preferable to attenuate noise using a medium filter in spite of lower spatial resolution [21]. Therefore, detail and medium filters are commonly called bone and soft tissue filters. Images acquired using detail filters are less compressible because of increased high-frequency details.

Conclusion

Producing compression guidelines for medical applications is not an easy task. Many factors affect the overall fidelity of compressed images. Coding algorithms and compression ratios are obviously important factors, but other parameters can also have significant impacts on image fidelity and, consequently, diagnostic quality. Our study showed that image content as well as acquisition parameters significantly affect image compressibility of computed tomography.

Exposure appears to be the most significant parameter as it accounted for about half of the compressibility variations, followed by slice thickness and filter type. Noise is known to be poorly compressible, and all the three parameters are directly related to noise levels of the acquired image. Smaller slice thicknesses and detail filter type are also associated with higher spatial resolution and higher frequency content; they therefore present additional challenges for image compression. Slice collimation and pitch did not have any effect on compressibility. Pitch did not impact noise levels and therefore compressibility because it was taken into account in the effective exposure parameter.

Visualization transformations such as window and level scaling can significantly alter the perception of quality. Great care is needed while choosing VOI parameters during comparative study on image quality. Moreover, compression metrics that take into account noise and grayscale transformations would be more suitable for medical image compression.

Finally, in light of on the body of literature, the experiment, and the discussion presented in this paper, we recommend that rate-based guidelines be phased out in favor to quality-based guidelines. Future work includes proposing fidelity metrics other than global PSNR to control the quantification step during lossy compression.

Acknowledgments This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

1. Infoway (2010) Annual Report 2009–2010: Reporting to Canadians
2. European Society of Radiology (ESR): Usability of irreversible image compression in radiological imaging. A Position Paper by the European Society of Radiology (ESR). Insights Imaging 2:103–115, 2011
3. Seeram E, Horii SC, Ratib OM: Using irreversible compression in digital radiology: a preliminary study of the opinions of radiologists. Medical Imaging 2006: PACS and Imaging Informatics. SPIE, pp 614510–614510–22, 2006
4. Ringl H, Schernthaner RE, Bankier AA, Weber M, Prokop M, Herold CJ, Schaefer-Prokop C: JPEG2000 compression of thin-section CT images of the lung: effect of compression ratio on image quality. Radiology 240:869–877, 2006
5. Kim TJ, Lee KH, Kim B, Kim KJ, Chun EJ, Bajpai V, Kim YH, Hahn S, Lee KW: Regional variance of visually lossless threshold in compressed chest CT images: lung versus mediastinum and chest wall. Eur J Radiol 69:483–488, 2009
6. Persons K, Palisson P, Manduca A, Erickson BJ, Savcenko V: An analytical look at the effects of compression on medical images. J Digit Imaging 10:60–66, 1997
7. Ponomarenko N, Krivenko S, Lukin V, Egiazarian K, Astola J: Lossy Compression of Noisy Images Based on Visual Quality: A Comprehensive Study. EURASIP J Adv Signal Process 2010: 976436, 2010
8. Erickson BJ, Manduca A, Persons KR, Earnest 4th, F, Hartman TE, Harms GF, Brown LR: Evaluation of irreversible compression of digitized posterior-anterior chest radiographs. J Digit Imaging 10: 97–102, 1997
9. Erickson BJ, Manduca A, Palisson P, Persons KR, Earnest 4th, F, Savcenko V, Hangiandreou NJ: Wavelet compression of medical images. Radiology 206:599–607, 1998
10. Woo HS, Kim KJ, Kim TJ, Hahn S, Kim B, Kim YH, Lee KH: JPEG 2000 compression of abdominal CT: difference in tolerance between thin- and thick-section images. AJR Am J Roentgenol 189: 535–541, 2007
11. Bajpai V, Lee KH, Kim B, Kim KJ, Kim TJ, Kim YH, Kang HS: Differences in compression artifacts on thin- and thick-section lung CT images. AJR Am J Roentgenol 191:W38–43, 2008
12. Kim KJ, Kim B, Lee KH, Kim TJ, Mantiuk R, Kang H-S, Kim YH: Regional difference in compression artifacts in low-dose chest CT images: effects of mathematical and perceptual factors. AJR Am J Roentgenol 191:W30–7, 2008
13. Kim KJ, Kim B, Lee H, Choi H, Jeon J-J, Ahn J-H, Lee KH: Predicting the fidelity of JPEG2000 compressed CT images using DICOM header information. Med Phys 38:6449–6457, 2011
14. Koff D, Bak P, Brownrigg P, Hosseinzadeh D, Khademi A, Kiss A, Lepanto L, Michalak T, Shulman H, Volkening A: Pan-Canadian evaluation of irreversible compression ratios ("lossy" compression) for development of national guidelines. J Digit Imaging 22:569–578, 2009
15. Canadian Association of Radiologists: CAR Standards for Irreversible Compression in Digital Diagnostic Imaging within Radiology. 1–11, 2008
16. Fidler A, Skaleric U, Likar B: The impact of image information on compressibility and degradation in medical image compression. Med Phys 33:2832–2838, 2006
17. Janhom A, van der Stelt PF, van Ginkel FC, Geraets WG: Effect of noise on the compressibility and diagnostic accuracy for caries detection of digital bitewing radiographs. Dentomaxillofac Radiol 28: 6–12, 1999
18. André T, Antonini M, Barlaud M, Gray R: Entropy-Based Distortion Measure and Bit Allocation for Wavelet Image Compression. IEEE Trans. on Image Processing 16(12):3058–64, 2007
19. 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 J E Health 14:350–354, 2008
20. Fidler A, Skaleric U, Likar B: The effect of image content on detail preservation and file size reduction in lossy compression. Dentomaxillofac Radiol 36:387–392, 2007
21. Shipping F, Bushberg JT, Anthony Seibert J, Leidholdt EM Jr, Boone JM, Sellers T: The Essential Physics of Medical Imaging (2nd Edition), 2.00 edition. 933, 2001
22. Kalyanpur A, Neklesa VP, Taylor CR, Daftary AR, Brink JA: Evaluation of JPEG and wavelet compression of body CT images for direct digital teleradiologic transmission. Radiology 217:772–779, 2000
23. Flint AC: Determining optimal medical image compression: psychometric and image distortion analysis. BMC Med Imaging 12:24, 2012
24. Ringl H, Schernthaner R, Sala E, El-Rabadi K, Weber M, Schima W, Herold CJ, Dixon AK: Lossy 3D JPEG2000 compression of abdominal CT images in patients with acute abdominal complaints: effect of compression ratio on diagnostic confidence and accuracy. Radiology 248:476–484, 2008
25. Kim B, Lee KH, Kim KJ, Mantiuk R, Kim H-R, Kim YH: Artifacts in slab average-intensity-projection images reformatted from JPEG 2000 compressed thin-section abdominal CT data sets. AJR Am J Roentgenol 190:W342–50, 2008
26. Gavrielides MA, Kinnard LM, Myers KJ, Peregoy J, Pritchard WF, Zeng R, Esparza J, Karanian J, Petrick N: A resource for the assessment of lung nodule size estimation methods: database of thoracic CT scans of an anthropomorphic phantom. Opt Express 18:15244–15255, 2010
27. Auli-Llinas F: Boi JPEG2000 Codec, 2013
28. R Core Team: R: A Language and Environment for Statistical Computing. ISBN 3–900051–07–0, 2012
29. Kraha A, Turner H, Nimon K, Zientek LR, Henson RK: Tools to support interpreting multiple regression in the face of multicollinearity. Front Psychol 3:44, 2012
30. Grömping U: Estimators of Relative Importance in Linear Regression Based on Variance Decomposition. Am Stat 61:139–147, 2007