Authors' accepted manuscript of the article published in
2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)
http://dx.doi.org/10.1109/EMBC.2013.6610265

Compressibility variations of JPEG2000 compressed computed tomography

Jean-François Pambrun and Rita Noumeir

***Abstract*— Compression is increasingly used in medical applications to enable efficient and universally accessible electronic health records. However, lossy compression introduces artifacts that can alter diagnostic accuracy, interfere with image processing algorithms and cause liability issues in cases of diagnostic errors. Compression guidelines were introduced to mitigate these issues and foster the use of modern compression algorithms with diagnostic imaging. However, these guidelines are usually defined as maximum compression ratios for each imaging protocol and do not take compressibility variations due to image content into account. In this paper we have evaluated the compressibility of thousands of computed tomography slices of an anthropomorphic thoracic phantom acquired with different parameters. We have shown that exposure, slice thickness and reconstruction filters have a significant impact on compressibility suggesting that guidelines based solely on compression ratios may be inadequate.**

I. Introduction

Image compression is increasingly used in the medical domain to enable instantaneous and universally available electronic health records (EHR) and therefore increase the quality of care. Medical images can be losslessly compressed by up to two-thirds. Better compression 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.

Estimating the impacts of these distortions is very difficult. Images with seemingly similar characteristics compressed using identical compression parameters can result in very different reproduction fidelity; some images could preserve all diagnostic qualities while others may become completely unusable. Because of liability issues raised the possibility of diagnostic errors caused by lossy compression, radiologists are not generally inclined to use compression techniques that would produce visually lossy results[1]. Compression guidelines were introduced to mitigate those issues, but variations [2][3] in image compressibility suggest that broad guidelines allows only for conservative and suboptimal compression.

In this paper, we will investigate the relation between CT acquisition parameters and image fidelity with JPEG2000 compression. In this context, image fidelity represents the faithfulness of reproduction and should not be confused with image quality which is related to subjective perception of an observer and its ability to perform the diagnostic task.

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

Jean-François Pambrun and Rita Noumeir are with the Department of Electrical Engineering, École de technologie supérieure (ÉTS), 1100 Notre-Dame Ouest, Montréal, Québec, Canada, jf.pambrun@gmail.com

II. Methodology

Our objective is to expose compressibility variations with computed tomography images that are caused by different acquisition parameters. To achieve this, multiple series of a single subject acquired with similar equipment is needed. We have restricted our study to computed tomography because it is known to be poorly compressible and represents an increasing amount of data. Acquisition parameters are available from each image DICOM header. However, these parameters are inconsistently reported between imaging devices. Exposure, for instance, can have a sightly different definition while filter type may not have any direct equivalence with different hardware configurations.

The National Cancer Institute has made many diagnostic image collections publicly available to encourage and support cancer research through their Cancer Imaging Archive project. One of these, labeled Phantom FDA[4], 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 applications. To meet their requirements, these researchers repeatedly scanned an anthropomorphic thoracic phantom with synthetic lung nodules acquired with different acquisition parameters. These parameters are presented in table I. They include five slice thickness varying from 0.8 mm to 5 mm, three effective exposures varying from 25 mAs to 200 mAs, two slice collimation configurations, two pitch configurations and two types of reconstruction filters. Two nodule layouts were made publicly available through the Cancer Imaging Archive. For this experiment, we have selected every parameter permutations for one of these layout. That is 23,767 images across 72 series. Slice thickness depends on slice collimation : 0.8 mm and 1.5 mm slices are acquired with a collimation of

TABLE I

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)	16x0.75, 16x1.5
Pitch	0.9, 1.2

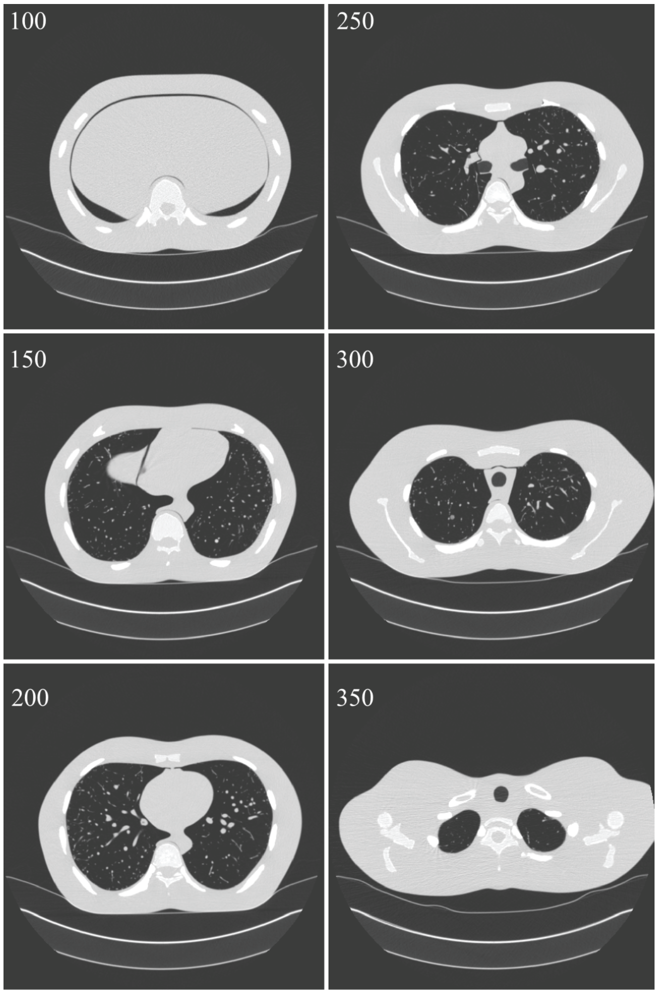

Fig. 1. Image content relative to slice location indicated in the top left corners

16×0.75 mm, 2 mm and 5 mm slices are acquired with a collimation of $16{\times}1.5$ mm and 3 mm slices are acquired with both configurations. All 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 across all acquisitions. Figure 1 shows six images of the phantom with their slice location in the top left corner. The scanned area spans about 30 cm with the slice location ranging from 90 mm to 389 mm. The thinnest acquisition (0.8 mm) had inter-slice spacing of 0.4 mm and contained 750 images while the thickest (5 mm) had a inter-slice spacing of 2.5 mm and contained only 120 slices.

We have compressed each image with a JPEG2000 coder using multiple compression ratios including : lossless, 4:1, 5:1, 6:1, 8:1, 10:1, 15:1 and finally 30:1 which is twice the recommendation for computed tomography published by the Canadian Association of Radiologists (CAR)[5]. This should provide a wide enough range of compression distortions. The codec used in our experiment was an open source JPEG2000 implementation[6]. R[7] was used for statistical analysis.

Compression ratios are computed with respect to actual file size including DICOM and JPEG2000 headers. The fidelity of every compressed image is evaluated using Mean Squared Error (MSE) and Peak-signal-to-noise ratio (PSNR). MSE and PSNR are equivalent, but PSNR can be easily interpreted as the ratio of signal range over noise in decibels. MSE is computed with

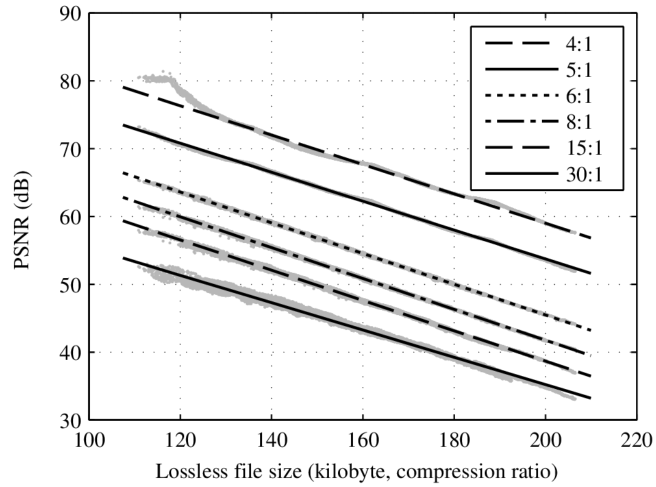

Fig. 2. PSNR of lossy compressed image plotted against lossless file size. Each point represents the PSNR of an image compressed at a specific lossy compression ratio. This PSNR is plotted against the lossless size of that image. The figure also shows various lines fitted on points corresponding to that compression ratio. The PSNR is directly correlated to the lossless compression image size.

$$\mathrm{MSE} = \frac{1}{\mathrm{MN}} \sum\nolimits_{\mathrm{i}} \sum\nolimits_{\mathrm{j}} \left[\mathrm{I_o}\left(\mathrm{i},\mathrm{j}\right) - \mathrm{I_c}\left(\mathrm{i},\mathrm{j}\right)\right]^2 \tag{1}$$

where I_0 is the original image and I_c is the compressed image. PSNR is computed with

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

where I_{range} is the range of the signal. We have calculated the range of the signal in all image and found it to be 1600. Although bit allocated were 16 and bits stored were 12 suggesting a dynamic range of 4096, we have used 1600 for I_{range} to compute PSNR values. PSNR was chosen because it is used in JPEG2000 rate allocation algorithms and it will be used to estimate image compressibility. Fitted models are evaluated with the coefficient of determination (R^2), Root mean squared error of the prediction (RMSE) and Pearson correlation coefficient (CC). The coefficient of determination can be interpreted as the fraction of variance explained by the model.

III. Results

The impact of acquisition parameters on compressibility can be observed by comparing:

1) file sizes of both images after lossless compression
2) image fidelity of both image compressed using the same compression ratio.

Most JPEG2000 coder algorithms are designed to minimize MSE (maximize PSNR) when a target compression ratio

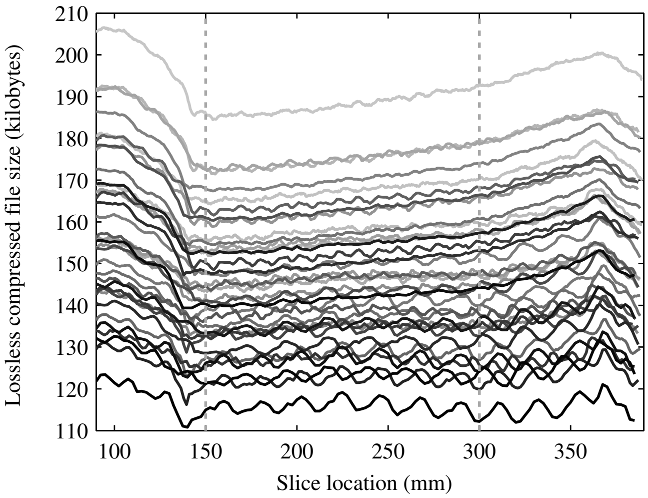

Fig. 3. Lossless file size shown with respect to slice location. Each curve represent one series. Two consecutive images from the same series have very similar compressibility. For a specific series (or curve), compressibility varies with slice location.

is specified. As a result, lossless file size and lossy image fidelity in PSNR for a given ratio are virtually equivalent. This is illustrated in Figure 2 where the PSNR of all 23,767 images compressed a 4:1, 5:1, 8:1, 15:1 and 30:1 plotted against their respective lossless file sizes. The relation is almost perfectly linear with very little variations and fidelity decreases with increased lossless file size. With images compressed at 8:1, a commonly used compression ratio with computed tomography, the linear regression is extremely well fitted with a R^2 of 0.99, a RMSE of 0.13dB and a Person correlation coefficient of 0.99. We can also note from Figure 2 that fidelity varies widely, more than 20dB, within any compression ratio while the average fidelity difference between each compression ratios is about 3dB. This suggest that image content and acquisition parameters have a much more significant impact on image fidelity than the targeted compression ratio itself. Furthermore, the figure shows that images with lossless file sizes smaller than 150 kilobytes are less affected by 30:1 compression than images with a lossless file sizes larger than 185 kilobytes but compressed with a target compression ratio of 8:1. In fact, the fidelity of 15% of all images compressed at 15:1, the maximum ratio allowed by the CAR guidelines, and 4% of those compressed at 8:1 are below the median fidelity of those compressed at 30:1. Nonlinearities at the lowest compression ratio for the smallest lossless file sizes appeared because these image could have been losslessly compressed if the integer filter bank of the codec were used instead of the floating point alternative reserved for lossy applications. Using float filters result in floating point coefficients that generate rounding errors. This explains why the PSNR values saturates around 80 dB for images smaller than 120kB compressed at 4:1.

Figure 3 shows the size of each of the 23,767 losslessly compressed plotted against slice location. Each series is displayed in varying shades of gray.

Compressibility variations between series are immediately obvious The average lossless file size of the best case was 116 kB while the worst case was 193 kB, a 66% difference that is the result of varying five acquisitions parameters. Compressibility variation along the subject's z axis are also apparent. Every series exhibits a similar behavior with respect to slice location and, as expected, consecutive images from each series show very similar compressibility.

We have showed that the image content as well as acquisition parameters have a significant impact on compressibility without identifying which is more significant. In the dataset that was used, five parameters were varied between each acquisition and it would be very valuable to identify exactly to what degree each one of these parameters are responsible for those variations. Figure 4a to 4e shows boxplots of 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 effects. In fact, t-testing indicates that pitch doesn't have any statistically significant impact on compressibility.

Using the simple quadratic regression with exposure, thickness, filter type and collimation :

$$\begin{aligned} \mathrm{PSNR} \sim\ & \beta_i + \beta_f \mathrm{Filter} + \beta_{\mathrm{c}} \mathrm{Collimation} \\ & + \beta_e \mathrm{Exposure} + \beta_{\mathrm{t}} \mathrm{Thick} \\ & + \beta_{e^2} \mathrm{Exposure}^2 + \beta_{\mathrm{t}^2} \mathrm{Thick}^2 \end{aligned} \tag{3}$$

we can obtain a well fitted model with a coefficient of determination (R^2) of .94, a prediction error of 4.4kB in terms of lossless file size (1.05dB in terms of 8:1 quality) and a Pearson correlation coefficient of .97. Table II shows standardized coefficients where each predictor and observed variables are transformed using $(X - E[X])/\sigma(X)$ to provide more uniform interpretation of the results.

Standardized coefficients can provide an estimation of the relative importance of each parameter suggesting that exposure and thickness have the most significant impacts on compressibility followed by filter type and finally slice collimation. Because covariance can introduce bias[8] with standardized coefficient other more robust techniques were developed to identify the contribution of each predictors to R^2. proportional marginal variance decomposition (PMVD) [9] is one such metric and according to this indicator, 53%

TABLE II

STANDARDIZED COEFFICIENTS FOR PREDICTING PSNR WHEN COMPRESSED AT 8:1

	PSNR (8:1)
Intercept (β_i)	0.44
Exposure (β_e)	0.73
Slice Thickness (β_t)	0.68
Filter Type (β_f)	-0.34
Slice Collimation (β_c)	-0.05
Exposure^2 (β_{e^2})	-0.31
Slice Thickness^2 (β_{t^2})	-0.13

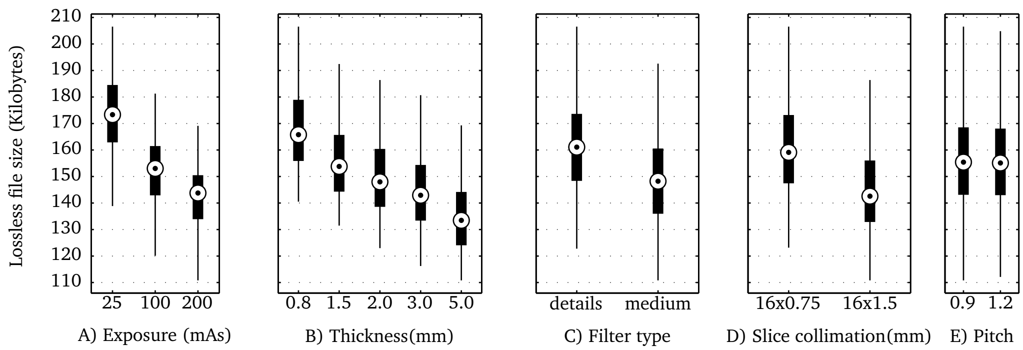

Fig. 4. Boxplots using all 23,676 images. Box is located at median, extends from the 25th to the 75th percentiles and the whiskers extend to the most extreme data points not considered outliers)

of the variation can be explained by exposure, 34% by slice thickness and 13% by filter type. PMVD reveals that slice collimation has no effect on compressibility on its own. Slice collimation is in fact entirely covariant with slice thickness because the 16×0.75 mm configuration can only be used with slice thicknesses of .8, 1.5 and 3 mm while 16×1.50 mm is used with thicknesses 2, 3 and 5 mm.

Finally, adding compression ratios typically used in CT applications, 6:1, 8:1, 10:1 and 15:1, to our predictor of equation 3 yields a new model with R^2 of 0.93, a RMSE of 1.3dB and a Person correlation coefficient of 0.96. Analyzing each predictors relative importance with PMVD reveals that compression ratio can only explain 28% of the PSNR variation while exposure represents 38%, slice thickness 25% and filter type 9%. Therefore, with this dataset, acquisition parameters affects the resulting fidelity more than the compression ratio itself.

Acquisition noise is inherently hard to compress and it is directly related to the number of photons generated by the X-Ray source which in turn is linked to both exposure and slice thickness. Increasing exposure or slice thickness increases the number of photons received at the detector reducing acquisition noise and increasing compressibility. On the other hand, images acquired using "detail" filters are less compressible because of the increased High frequency fine details while noise is also slightly attenuated by the "medium" filter.

IV. Conclusion

Producing compression guidelines for medical applications is not an easy task. Many factors affect compressibility and, consequently, the overall fidelity of compressed images. Coding algorithms and compression ratios are obviously important factors but other parameters can also have a significant impact. Our study suggest that exposure is the most significant parameter and it was linked to more than 50% of the compressibility variations in our experiment, followed by slice thickness with about 30% and filter type with 10%. Such wide variation within a single modality and a single subject suggest that guidelines based on compression ratios may be inadequate.

References

[1] E. Seeram, "Using irreversible compression in digital radiology: a preliminary study of the opinions of radiologists," *Proceedings of SPIE*, 2006.

[2] H. Ringl, R. E. Schernthaner, A. A. Bankier, M. Weber, M. Prokop, C. J. Herold, and C. Schaefer-Prokop, "JPEG2000 compression of thin-section CT images of the lung: effect of compression ratio on image quality." *Radiology*, vol. 240, no. 3, pp. 869–77, Sep. 2006.

[3] T. J. Kim, K. H. Lee, B. Kim, K. J. Kim, E. J. Chun, V. Bajpai, Y. H. Kim, S. Hahn, and K. W. Lee, "Regional variance of visually lossless threshold in compressed chest CT images: lung versus mediastinum and chest wall." *European journal of radiology*, vol. 69, no. 3, pp. 483–8, Mar. 2009.

[4] M. a. Gavrielides, L. M. Kinnard, K. J. Myers, J. Peregoy, W. F. Pritchard, R. Zeng, J. Esparza, J. Karanian, and N. Petrick, "A resource for the assessment of lung nodule size estimation methods: database of thoracic CT scans of an anthropomorphic phantom." *Optics express*, vol. 18, no. 14, Jul. 2010.

[5] CAR, "CAR Standards for Irreversible Compression in Digital Diagnostic Imaging within Radiology," pp. 1–11, 2008.

[6] F. Auli-Llinas, "boi software," 2011. [Online]. Available: http://www.gici.uab.es/BOI/

[7] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2012. [Online]. Available: http://www.r-project.org

[8] A. Kraha, H. Turner, K. Nimon, L. R. Zientek, and R. K. Henson, "Tools to support interpreting multiple regression in the face of multicollinearity." *Frontiers in psychology*, vol. 3, no. March, p. 44, Jan. 2012.

[9] B. Feldman, *Relative Importance and Value Manuscript (Version 1.1, March 19 2005)*, 2005.