IEEE International Conference on Image Processing, Quebec City, Canada, Sep. 2015.

DATA RATE AND DYNAMIC RANGE COMPRESSION OF MEDICAL IMAGES: WHICH ONE GOES FIRST?

Shahrukh Athar, Hojatollah Yeganeh and Zhou Wang

Dept. of Electrical & Computer Engineering, University of Waterloo, Waterloo, ON, Canada
Email: {shahrukh.athar, hyeganeh, zhou.wang}@uwaterloo.ca

ABSTRACT

Advances in the field of medical imaging have led to an immense increase in the volume of images being acquired. A fast growing application is to enable physicians to access image data remotely from any viewing device. This casts new challenges for data rate compression. Meanwhile medical images typically have High Dynamic Range (HDR), which needs to be transformed to Low Dynamic Range (LDR) through a so-called "windowing" operation in order for them to be viewed on standard displays to best visualize specific types of content such as tissues or bone structures. This leads to a basic question: Should data compression be performed before windowing or vice versa? Answering this question needs domain knowledge and also requires comparing HDR and LDR images in terms of objective measures, which has only recently become possible. In this paper, we compare the two alternative schemes by using a recently proposed structural fidelity measure. Our study suggests that data compression followed by windowing delivers better performance than the other alternative.

***Index Terms*—** Medical image compression, high dynamic range, windowing, tone mapping, image quality assessment, structural fidelity

1. INTRODUCTION

Technological advances in medical imaging have led to a rapid increase in the resolution of medical images and in the number of images being acquired every day. An increasingly important application is to enable physicians to access image data remotely from any viewing device, including mobile devices. To achieve this goal, reducing the data rate by lossy image compression is desirable as long as it does not affect the diagnostic quality [1, 2]. An additional complication is that medical images often have higher dynamic range (HDR) than that of the remote display devices. For example, images stored using the Digital Imaging and Communications in Medicine (DICOM) standard [3] typically have a precision of 12 to 16 bits per pixel, while most image display devices today are made for Low Dynamic Range (LDR) images which have lower precision (8 to 10 bits per pixel). Tone mapping operators (TMOs) are necessary to compress HDR images to LDR [4, 5]. To reduce the dynamic range of DICOM images, a tone mapping process known as "windowing" is applied that usually linearly maps the intensity region of interest (typically larger than the dynamic range of the displays) from HDR image to the dynamic range of the LDR display. In summary, HDR medical images need to go through both lossy data rate compression and dynamic range compression (windowing) before they are displayed on remote devices. Both operations cause information loss.

An important question thus arises: Should data rate compression be performed before windowing or vice versa? Surprisingly, previous effort to compare these two alternative schemes is missing in the literature. Subjective evaluation is highly valuable but is time consuming, expensive and cannot be embedded in optimization processes. Objective Image Quality Assessment (IQA) is a much more convenient approach to compare the final LDR image with its HDR reference. In [6], objective IQA metrics including PSNR, SSIM [7], MS-SSIM [8] and IW-SSIM [9] are employed to evaluate a final LDR image after windowing the HDR reference image to an intermediate LDR reference. However, the windowing process itself results in loss of information and hence the LDR reference image is no longer pristine, which may lead to unreliable comparison.

In this work, we shall compare the two possible schemes of compression followed by windowing (CW) and windowing followed by compression (WC) in a series of experiments by using an objective structural fidelity measure that has recently been proposed [10, 11]. This is one of the first method that allows the comparison of two images of different dynamic ranges, thus avoiding the need of relying on an LDR reference image. Our results suggest that data rate compression followed by windowing is generally a better option.

2. WINDOWING AND CROSS DYNAMIC RANGE QUALITY MODEL

Windowing is a mapping function that maps an interval of interest of the high dynamic range image to a lower dynamic range. This interval, which is a subset of the high dynamic range, is typically called a *Window* and is defined by two parameters: (1) *Window Width* (W), the range of the interval and (2) *Window Center* (C), the center of the interval. For an HDR image, let l_l and l_u define the lower and upper bounds of the *window* range. They can be expressed in terms of *W* and *C* as: $l_l = C - \frac{W}{2}$ and $l_u = C + \frac{W}{2}$. The windowing function maps the HDR interval $[l_l, l_u]$ to the typical LDR interval $[0, 255]$. The values of *W* and *C* are adjusted to create different windows that are suitable to view particular regions of the human body (bones, tissue, etc). For example, the LDR image in Fig. 1(a) was obtained by applying (*W* = 350 and *C* = 30) whereas the one in Fig. 1(b) was obtained by applying (*W* = 1500 and *C* = -600) on the same HDR image. In medical imaging, windowing is usually performed linearly. However, it was shown that this does not lead to optimal perceptual image quality and that adaptive windowing based on a family of piecewise linear and sine bases functions delivers better performance [10, 12], where the optimal windowing solution is determined by an objective IQA model derived from the Tone-Mapped image Quality Index (TMQI) [11].

TMQI allows the comparison of HDR images with their LDR counterparts. TMQI consists of two components: (1) a structural fidelity measure and (2) a statistical naturalness measure. Since medical images are not natural images, the statistical naturalness measure is not necessary and the structural fidelity (SF) measure is adopted here.

The *SF* quality measure is developed based on a modified contrast comparison component and the structure comparison component of the Structural Similarity Index (SSIM) [7] and is first applied to the images locally by

$$SF_{local}(x, y) = \frac{2\acute{\sigma}_x \acute{\sigma}_y + C_1}{\acute{\sigma}_x^{\,2} + \acute{\sigma}_y^{\,2} + C_1} . \frac{\sigma_{xy} + C_2}{\sigma_x \sigma_y + C_2} \qquad (1)$$

Here x and y are local image patches in the HDR and LDR images respectively. The second term in Eq. (1) is the structure comparison component and is the same as in SSIM, where σ_x, σ_y and σ_{xy} are the standard deviations and cross correlation between x and y, and C_1 and C_2 are stability constants. The first term in Eq. (1) is a modified contrast comparison component. It penalizes the contrast differences between the HDR and LDR image patches when the contrast is significant in one patch but insignificant in the other. This is achieved by passing the local standard deviations through a non-linear cumulative Gaussian mapping function, resulting in the modified standard deviation values $\acute{\sigma}$ [10, 11].

The local *SF* quality measure is applied to the LDR medical image and its HDR reference by using a sliding window that runs across the images. This results in a *Quality Map* that shows how structural fidelity varies across the image. Sample images and *SF* quality maps are given in Fig. 1, where brighter regions in the *SF* maps suggest that the information is well preserved, while darker regions indicate there is significant information loss. Since we cannot see HDR images on standard LDR displays, quality maps are powerful tools

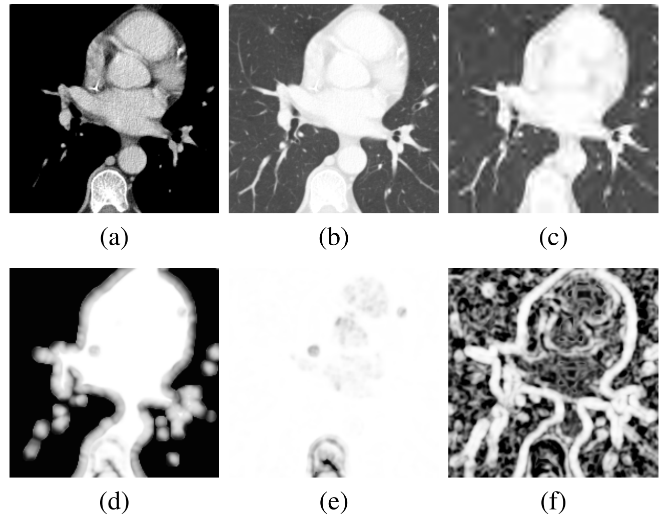


Fig. 1: (a) LDR Image (W = 350, C=30), (b) LDR Image (W = 1500, C = -600), (c) JPEG2000 compressed version of (b) (Compression Ratio: 50); (d)-(f): SF quality maps of (a)-(c).

that allow us to observe information loss as a result of operations such as windowing and data compression. For instance, JPEG2000 compression causes blurriness and loss of structural details of the central region in Fig. 1(c), which is well predicted by the *SF* map in Fig. 1(f). Finally, the quality map is averaged to provide an overall structural fidelity quality score:

$$SF = \frac{1}{N} \sum_{i=1}^{N} SF_{local}(x_i, y_i) \qquad (2)$$

where N is the total number of patches and x_i, y_i are the ith patches in the HDR and LDR medical images respectively. The parameters are given by $C_1 = 0.01$, $C_2 = 10$, and a Gaussian sliding window of size 11×11 and standard deviation of 1.5 was employed.

3. EXPERIMENTS AND RESULTS

We now define the two alternative schemes of medical image compression followed by results and discussions. In both cases, we use JPEG2000 to compress the images because it allows both 8-bit and 16-bit compression. We use linear, *SF* optimized piecewise linear, and *SF* optimized sine functions as the windowing operators.

In the first scheme, the LDR image is obtained first by applying windowing to the HDR DICOM image. Then 8-bit JPEG2000 compression is applied on this LDR image to obtain the compressed LDR image. We call this scheme "*Windowing-Compression*" (WC). The compressed LDR image can then be transmitted to the end users. The benefit of this scheme is to maximally reduce the transmission band-

Table 1: Experiment 1: *SF* Quality scores for the cases of *NC*, *WC* and *CW*. (CR: Compression Ratio)

Image	CR	Window	SF_{NC}	SF_{WC}	SF_{CW}
General	10	Linear	0.9989	0.913	**0.9686**
		PW Linear	0.999	0.9135	**0.9684**
		Sine	0.9993	0.9102	**0.9658**
Knee	15	Linear	0.9844	0.9061	**0.9605**
		PW Linear	0.9978	0.9172	**0.9689**
		Sine	0.9838	0.8966	**0.9597**
CAP	15	Linear	0.9794	0.8216	**0.9252**
		PW Linear	0.9794	0.8228	**0.9249**
		Sine	0.9789	0.8175	**0.9251**

Table 2: Experiment 1: Image sizes for the cases of *NC*, *WC* and *CW*. (CR: Compression Ratio)

Image	CR	$Size_{NC}$	$Size_{WC}$	$Size_{CW}$
General	10	512 KB	25.1 KB	50.7 KB
Knee	15	345 KB	11 KB	22.5 KB
CAP	15	512 KB	16.6 KB	33.6 KB

Table 3: Experiment 2: Compression Ratios (CR) for *WC* and *CW* that result in similar SF Quality scores and image sizes.

Image	Window-Compression			Compression-Window		
	CR	SF_{WC}	$Size_{WC}$	**CR**	SF_{CW}	$Size_{CW}$
General	**10**	0.9135	25.1 KB	**19**	0.9158	26.3 KB
Knee	**15**	0.9172	11 KB	**30**	0.9174	11 KB
CAP	**15**	0.8228	16.6 KB	**29**	0.8248	17.1 KB

width, but it suffers from two drawbacks: (1) lossy image compression is being applied to an already altered image (as a result of windowing); (2) since the end users receive a LDR image, further windowing is not possible and other body parts, that require different windows, cannot be retrieved. This can lead to retransmission requests and the requirement to store the original uncompressed HDR images at the source, which takes more storage space.

In the second scheme, image compression of the original HDR DICOM image is done first by using 16-bit JPEG2000 compression. This results in a compressed HDR image transmitted to the remote user. At the user side, the compressed HDR image is decompressed and windowing is applied to obtain a LDR image. We call this scheme "*Compression-Windowing*" (CW). The apparent advantages of this approach are: (1) different types of windows can be applied at the user end, and thus different body parts can be viewed without retransmission of the images; (2) compressed HDR images, instead of the original HDR images, can be stored at the source which leads to reduced storage requirement. The drawbacks are: (1) more bandwidth maybe required to transmit the image; (2) typical browsers may not be able to decode and show images compressed at 16 bits, and thus dedicated software may need to be installed at the user end.

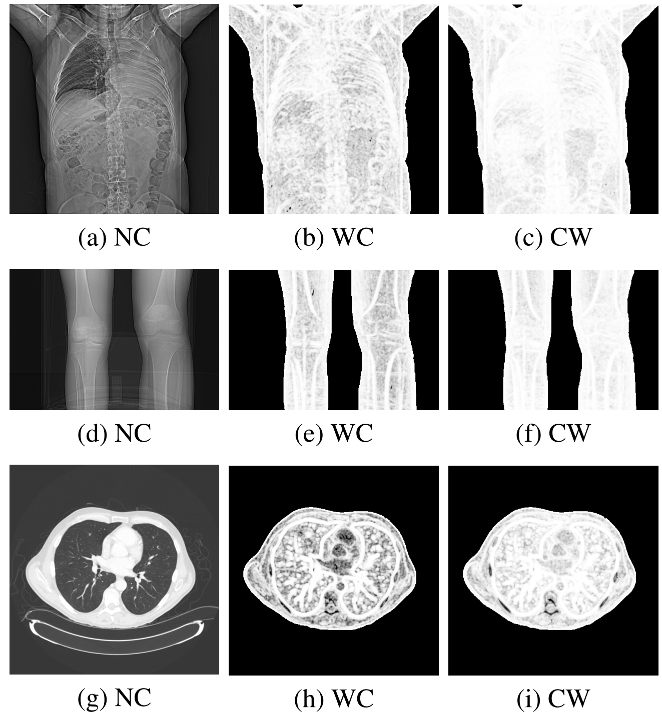


Fig. 2: Experiment 1: Only windowed images (*NC*): (a) General, (d) Knee, (g) CAP; Quality Maps for *WC*: (b) General, (e) Knee, (h) CAP; Quality Maps for *CW*: (c) General, (f) Knee, (i) CAP

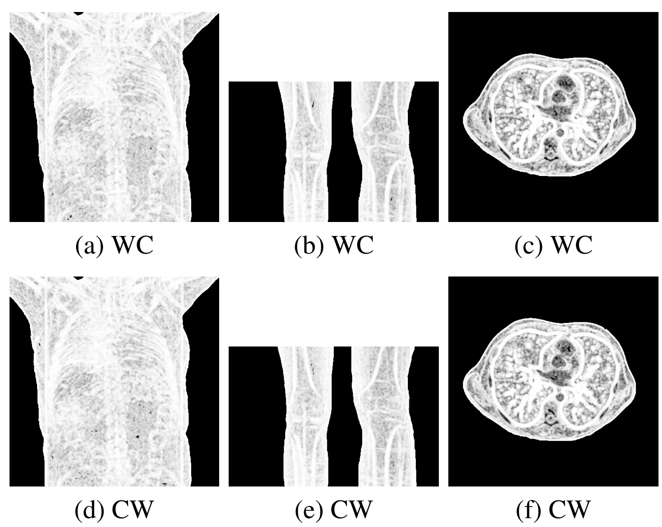
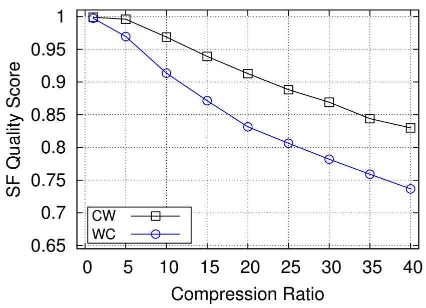
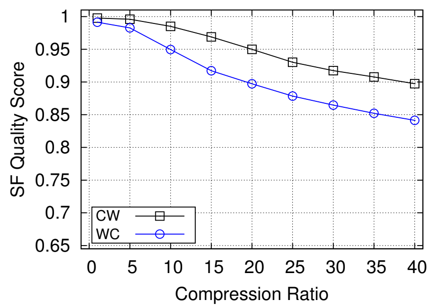


Fig. 3: Experiment 2: Quality Maps for *WC*: (a) General, (b) Knee, (c) CAP; Quality Maps for *CW*: (d) General, (e) Knee, (f) CAP

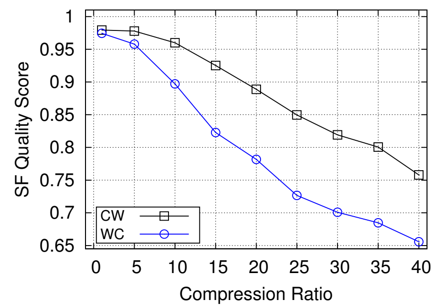
Three experiments were carried out on three types of CT images that cover the thorax and abdomen regions of the human body, the knee, and the Chest, Abdomen and Pelvis En-



(a) General



(b) Knee



(c) CAP

Fig. 4: Experiment 3: SF Quality Score VS Compression Ratio.

hanced Body, respectively. The three types of images named *General*, *Knee* and *CAP* fall in the categories of Body, Musculoskeletal system (MSK) and Chest for which Koff et al. [2] recommended JPEG2000 compression ratios of 10, 10-15 and 10-15, respectively.

In ***Experiment 1***, compression ratios of 10, 15 and 15 were set for *General*, *Knee* and *CAP* images respectively and *SF* quality scores for the cases of *No Compression* (NC), *Windowing-Compression* (WC) and *Compression-Windowing* (CW) were obtained along with the quality maps. Table 1 shows the *SF* quality scores obtained for various windowing functions. The *NC* case can be regarded as the maximum possible quality score for each image and window type. It can be observed that the *CW* scheme performs better than *WC* in terms of *SF* quality scores for all images and window types. In case of *NC* and *WC*, windows that use piecewise linear and sine bases perform better than linear windows. However, this is not always the case in the *CW* scheme. Further investigations are desirable to design adaptive approaches for joint data compression and windowing.

Fig. 2 shows the *NC* LDR versions of the three test images along with the quality maps obtained for the best performing window for the cases of *WC* and *CW*. The superior performance of *CW* in all three cases can be visually observed from its quality maps. At the same time it is evident from Table 2 that for the same compression ratio, *CW* images have almost twice the file sizes of the *WC* images. Therefore, it can be concluded that for the compression ratios recommended in [2], the *CW* scheme performs better than *WC* in terms of image quality but it also takes more storage space and transmission bandwidth. However, it is also evident from Table 2 that the *CW* scheme still results in substantial savings in terms of size when compared to the *NC* case.

In ***Experiment 2***, we used the compression ratios recommended in [2] for the *WC* scheme and then varied the compression ratios for *CW* in order to find compression ratios for which the *SF* quality scores of *CW* were nearly the same as the corresponding scores for *WC*. The results are presented in Table 3, where it is evident that both schemes have nearly the same *SF* quality scores and image file sizes when the compression ratios used in *CW* were almost twice of those used in *WC*. The quality maps of the *WC* and *CW* images corresponding to the results of Table 3 are shown in Fig. 3. Visual inspection conforms that the quality maps of the two image compression schemes in this case look similar to each other.

In ***Experiment 3***, we determined *SF* Quality scores and Image sizes at nine different compression ratios (1, 5, 10, 15, 20, 25, 30, 35, 40). This was done in order to observe if the findings of Experiments 1 and 2 hold for a wide range of compression ratios. Fig. 4 shows the *SF* Quality score VS Compression ratio plots and it can be observed that for all compression ratios tested and for all images, the *CW* scheme performs better than *WC*. Similarly, it can also be observed that for any given compression ratio in case of *WC*, the *CW* scheme has a similar *SF* quality score at almost twice that compression ratio.

4. CONCLUSION

In this work, we compared two different schemes for data rate and dynamic range compression of medical images by using a recent objective *SF* measure. We found that the *CW* scheme performs better than the *WC* scheme in terms of the *SF* measure. This was also visually validated by observing the quality maps. For a given compression ratio, we found that the *CW* scheme results in images with larger file size as compared to *WC*. However, it still leads to substantial file size reduction when compared to the *NC* case. We also found that when the *CW* scheme uses almost twice the compression ratio of the *WC* scheme, it has quality scores and image file sizes similar to *WC*. This is a significant advantage of the *CW* scheme because it means that images of similar size and quality can be obtained as compared to *WC* with an added advantage of applying any windowing scheme at the user end. A distinction needs to be made between 8-bit and 16-bit compression regarding dedicated compression ratios that are acceptable to radiologists. More work also needs to be done on adaptive joint data compression and windowing.

5. REFERENCES

[1] D. A. Koff and H. Shulman, "An overview of digital compression of medical images: can we use lossy image compression in radiology?," *Canadian Association of Radiologists Journal*, vol. 57, no. 4, pp. 211–217, 2006.

[2] D. Koff, P. Bak, P. Brownrigg, D. Hosseinzadeh, A. Khademi, A. Kiss, L. Lepanto, T. Michalak, H. Shulman and A. Volkening, "Pan-Canadian evaluation of irreversible compression ratios (lossy compression) for development of national guidelines," *Journal of digital imaging*, vol. 22, no. 6, pp. 569–578, 2009.

[3] O. S. Pianykh, *Digital imaging and communications in medicine (DICOM)*, Springer, 2012.

[4] F. Banterle, A. Artusi, K. Debattista, and A. Chalmers, *Advanced high dynamic range imaging: theory and practice*, CRC Press, 2011.

[5] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, *High dynamic range imaging: acquisition, display, and image-based lighting*, Morgan Kaufmann, 2010.

[6] I. Kowalik-Urbaniak, E. R. Vrscay, Z. Wang, C. Cavaro-Menard, D. Koff, B. Wallace, and B. Obara, "The impact of skull bone intensity on the quality of compressed CT neuro images," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2012, pp. 83190L1–83190L9.

[7] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, 2004.

[8] Z. Wang, E. P. Simoncelli and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on*. IEEE, 2003, vol. 2, pp. 1398–1402.

[9] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *Image Processing, IEEE Transactions on*, vol. 20, no. 5, pp. 1185–1198, 2011.

[10] H. Yeganeh, Z. Wang and E. R. Vrscay, "Adaptive windowing for optimal visualization of medical images based on a structural fidelity measure," in *International Conference on Image Analysis and Recognition*, pp. 321–330. Springer, 2012.

[11] H. Yeganeh and Z. Wang, "Objective quality assessment of tone-mapped images," *Image Processing, IEEE Transactions on*, vol. 22, no. 2, pp. 657–667, 2013.

[12] N. Nikvand, H. Yeganeh, and Z. Wang, "Adaptive windowing for optimal visualization of medical images based on normalized information distance," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 1200–1204.