PATIENT EXPOSURE OPTIMISATION THROUGH TASK-BASED ASSESSMENT OF A NEW MODEL-BASED ITERATIVE RECONSTRUCTION TECHNIQUE

Julien G. Ott[1],*, Alexandre Ba[1], Damien Racine[1], Nick Ryckx[1], François O. Bochud[1],
Hatem Alkadhi[2] and Francis R. Verdun[1]
[1]Institute of Radiation Physics, CHUV, Lausanne, Switzerland
[2]Institute of Diagnostic and Interventional Radiology, USZ, Zürich, Switzerland

*Corresponding author: julien.ott@chuv.ch

The goal of the present work was to report and investigate the performances of a new iterative reconstruction algorithm, using a model observer. For that, a dedicated low-contrast phantom containing different targets was scanned at four volume computed tomography dose index ($CTDI_{vol}$) levels on a Siemens SOMATOM Force computed tomography (CT). The acquired images were reconstructed using the ADMIRE algorithm and were then assessed by three human observers who performed alternative forced choice experiments. Next, a channelised hotelling observer model was applied on the same set of images. The comparison between the two was performed using the percentage correct as a figure of merit. The results indicated a strong agreement between human and model observer as well as an improvement in the low-contrast detection when switching from an ADMIRE strength of 1–3. Good results were also observed even in situations where the target was hard to detect, suggesting that patient dose could be further reduced and optimised.

INTRODUCTION

Over the past decade, the radiation dose delivered to patients via diagnostic X-ray imaging has continuously increased until today, where it reaches 25 % of the accumulated man-made and natural radiation contributions. Among that 25 %, computed tomography (CT) raises a particular concern, since this imaging modality represents in Switzerland for example 68 % of the collective dose, yet only 8 % of the number of examinations[1]. In this context, CT manufacturers have developed new strategies like iterative reconstruction (IR) algorithms in order to ensure that the benefits–risk ratio remains in favour of the patient. This new technology certainly improved the clinical practice[2], but it has also led to drastic changes in image perception. Thus, ensuring an adequate level of image quality while keeping patient's exposure as low as reasonably achievable constitutes a new challenge to be addressed. The use of task-based image quality assessment method could represent an efficient way to perform this optimisation scheme[3, 4]. Therefore, the goal of this study was to report and investigate the performances of a new IR technique using a model observer that mimics human detection of low-contrast targets: the channelised Hotelling observer (CHO) model.

MATERIALS AND METHODS

Data acquisition

A dedicated low-contrast phantom (QRM, Moehrendorf, Germany), mimicking the attenuation produced by a patient's chest, was used. The phantom could embed two different custom-made modules in its middle: a homogeneous modulus and another containing low-contrast spherical targets of 6 and 8 mm in diameter with contrast levels of 10 and 20 HU at 120 kVp.

Data acquisition was performed at the University Hospital Zurich on a third-generation dual-source 192-slice CT scanner (SOMATOM Force, Siemens Healthcare, Erlangen, Germany). A tube voltage of 120 kVp, a 300 mm display field of view (DFOV), a 512×512 matrix size and 2.0-mm-thick slices, which were reconstructed every 1.0 mm, were used. Acquisitions were performed in the helical mode with a pitch of 0.98. Four dose levels [1.0, 3.5, 8.0 and 15.0 mGy expressed in a volume computed tomography dose index ($CTDI_{vol}$) phantom of 32 cm in diameter] were investigated, using the procedure described in the IEC 60601-2-44[5] to measure the $CTDI_{vol}$. The phantom was scanned 20 times for each condition. Reconstructions were performed using the Siemens advanced model iterative reconstruction (ADMIRE) with strength levels 1 and 3. On the machine, users can choose ADMIRE levels ranging from 1 to 5, with level 1 being closest to the image impression of traditional, filtered back-projection, and level 5 showing the strongest noise reduction. In the end, 32 different categories were obtained (four dose levels, two ADMIRE levels, two contrast levels and two target sizes). From these sets of acquisitions, regions of interest (ROIs) of 41×41 pixels (0.59 mm pixel size) containing the centred targets were extracted. For each category, 100 images containing a signal

(20 scans × 5 targets with identical size and contrast in the phantom) and 1000 images with only noise were extracted. The ROIs that contained noise only were extracted from the homogeneous modulus, whereas the ROIs containing the signals came from the low-contrast modulus. This methodology enabled holding the same position on the (x,y) plane for both signal and noise ROIs, in turn enabling solving the noise stationarity problem.

Human observer

In the human observer experiment, three medical physicists took part in four alternative forced choice (4-AFC) experiments in order to yield a percentage of correct responses (PC) indicating how well they managed to detect the signals. The 4-AFC experiment consisted in selecting the signal-present image in a batch of three signal-absent images and one signal-present image, which were presented together in a randomised order. All observers were blinded to the CT acquisition and reconstruction conditions and began their test with a training session that was made of images acquired at high dose level. They were then asked to make decisions for all 32 categories acquired. The previously made acquisitions provided 100 signal-present ROIs and 1000 signal-absent ROIs for each category. ROIs among those data were selected randomly and used for the 4-AFC tests. For each observer and category, every answer to the 100 trials was stored and compared with the correct response, allowing the computation of the PC.

CHO model observer

A model observer enables to predict the detection of low-contrast signals by calculating a scalar response called the decision variable and denoted by λ_i. This parameter is given by

$$\lambda_i = w^T \times g_i, \tag{1}$$

where w is the template of the model observer, and g_i is the analysed ROI ($i = n$ or $i = s$ represents signal-absent or signal-present hypothesis, respectively).

The CHO model used in this study is an anthropomorphic model observer also including preprocessing of the image by a set of channels that enhance some spatial frequencies[6, 7]. The template w_{CHO} of this model is obtained as explained in the following part (extensive details can be found elsewhere[8–11]):

$$w_{CHO} = \left[\frac{1}{2}(K_{cs} + K_{cn})\right]^{-1}(\overline{g_{cs}} - \overline{g_{cn}}), \tag{2}$$

where

$$K_{cn} = U^T K_n U \quad \text{and} \quad K_{cs} = U^T K_s U. \tag{3}$$

In this equation, K_n and K_s are the covariance matrix calculated, respectively, from the signal-absent and signal-present data, and U is the matrix representation of the channel filters described more extensively below.

In Equation 2, $\overline{g_{cs}}$ and $\overline{g_{cn}}$ are the means of the channel outputs under a signal-present and signal-absent hypothesis and can be estimated according to

$$\overline{g_{cs}} = U^T \overline{g_s} \quad \text{and} \quad \overline{g_{cn}} = U^T \overline{g_n}. \tag{4}$$

The employed set of channels is called dense difference of Gaussian (DDoG) channels and was described by Abbey and Barrett in 2001[12]. It includes 10 channels for which the radial frequency profile of the jth channel is given by

$$U_j(\rho) = \exp\left[-\frac{1}{2}\left(\frac{\rho}{W\sigma_j}\right)^2\right] - \exp\left[-\frac{1}{2}\left(\frac{\rho}{\sigma_j}\right)^2\right], \tag{5}$$

where ρ is the spatial frequency, $W = 1.67$ defines the bandwidth of the channel, and σ_j is the standard deviation of the jth channel. Each σ_j value is defined by the equation $\sigma_j = \sigma_o \alpha^j$ with $\sigma_o = 0.005$ and $\alpha = 1.4$[12].

The decision variable can then be calculated by injecting the channel output of the ROI (denoted by g_{ci}) and the template w_{CHO} in Equation 1:

$$\lambda_i = w_{CHO}^T \times g_{ci}. \tag{6}$$

The PC was obtained using the CHO model to perform 4-AFC tests on the acquired images. The value of the decision variable was used to determine which of the four images contained the signal (the highest value of λ is supposed to be the signal-present image). Then, the results of the model observers were compared with the truth in order to enabling the computation of PCs.

Uncertainty estimation

The uncertainties of the models' results were estimated by performing bootstrap[13]. According to the bootstrap method and the set of 100 signal-present ROIs, the 4-AFC test was performed 150 times for each category, leading to 150 values of PC. Then, the mean and standard deviation of the 150 values obtained were computed in order to determine the final mean PC value as well as its standard deviation for each acquisition condition. This allowed to estimate a 95 % confidence interval. The uncertainties for the human observers were calculated using the results of the three different observers. For each category, the mean PC value of the observers and its

standard error were calculated in order to display a 95 % confidence interval.

RESULTS

In this section, the qualities of the images obtained using the ADMIRE algorithm with strength levels 1 and 3 are assessed through the performances of model (Figure 1) and human observers (Figure 2).

Results of the CHO model with the DDoG channels suggest that both strength levels exhibited PCs in the same range with an increase of a few per cent in the results when switching from strength level 1 to 3. This trend was, however, only observed for certain signals and dose range, namely the lowest contrast (10 HU) associated to low dose levels (1.0 and 3.5 mGy).

Human observers exhibited results very similar to the ones obtained with the CHO model in terms of PC values. The results also suggest that the use of higher ADMIRE strength level is useful to improve the detection of small-size and low-contrast signal under low CTDI$_{vol}$. On top of that, it was observed that the largest signal size (8 mm) with the highest contrast (20 HU) corresponded to a trivially easy task
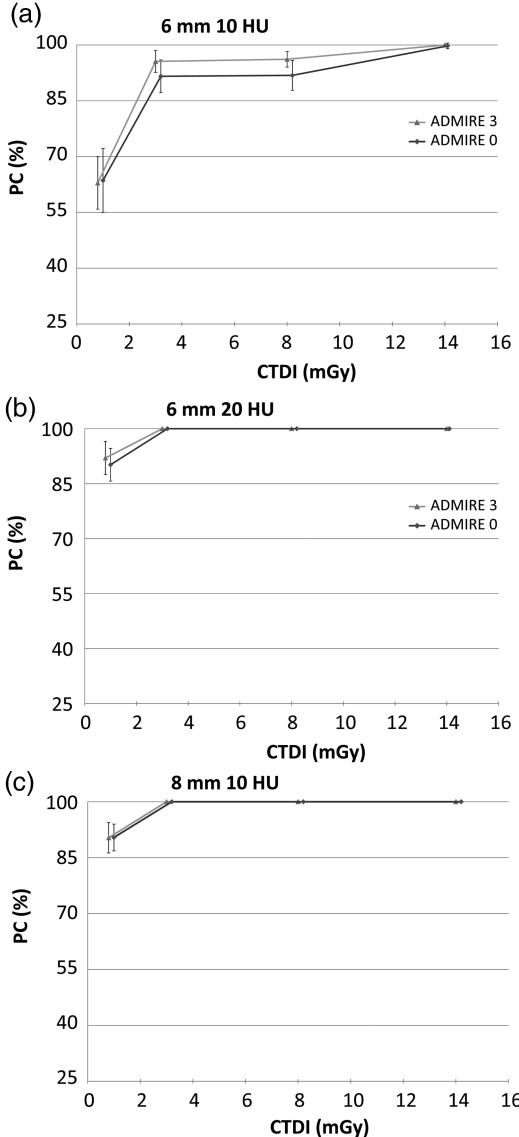

Figure 1. Results of the CHO model observer (in PC, obtained by performing 4-AFC tests).
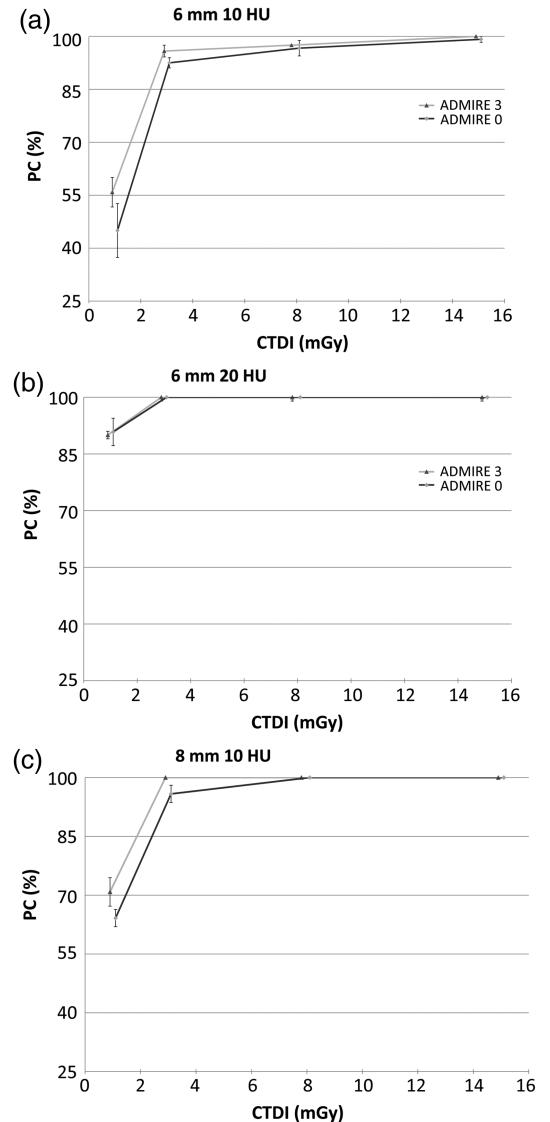

Figure 2. Results of the three human observers (in PC, obtained by performing 4-AFC tests).

with PC always equal to 100 % no matter which dose and ADMIRE level was employed. Those results remained true for both human and model observers and are therefore not presented in Figures 1 and 2.

Model and human observers exhibited a great adequacy in their results, both of them indicating that the use of a higher ADMIRE level enhances the detection when working under conditions where the signal is hard to detect. It was also witnessed that the PC increases with the given dose until it reaches an asymptote and that this asymptote is reached faster when using higher ADMIRE strength level. Furthermore, the results enlighten that no more dose increase is needed once an amount of ∼4 mGy is reached because the performance of the human observer is already almost maximal at this point.

DISCUSSION

The goal of the present investigation was to determine if the use of the ADMIRE algorithm at different strength could lead to a high detection performance, therefore allowing a further dose reduction in the clinical practice. The results showed that the CHO model with DDoG channels gave coherent results since it reproduced the behaviour of humans very well and under a wide range of conditions and signal contrasts. The model seems to be very efficient in low-contrast detection and even sometimes overestimates human results for low contrast and low doses. These results are coherent with recent studies from Leng and Yu[14] which showed the efficiency of the CHO model for the low-contrast detection. Also, both model and human observers reported a visible improvement in the low-contrast detection when increasing the ADMIRE strength. This trend was observed when working at low dose levels (<4 mGy) for all signal types. Indeed, when working at higher dose levels, the PCs always reached 100 %, letting no room for improvement. In the end, the use of ADMIRE makes it possible to diminish the dose without losing information in the image. Indeed, the PC results obtained in the study reach very high or perfect values for almost every acquisition condition and signal, indicating that a dose reduction without impacting the detection performance would be possible.

However, some limitations of the present study have to be underlined. First among them, the number of images acquired may be considered as low since not enough ROIs were disposed to separate them in two exploitable data sets. Usually, a first set is used for the computation of the covariance matrix in the determination of the template of the CHO model, while the second data set is used for the computation of the PCs. However, it is worth to underline that Barrett and Myers[15] who studied this problem concluded that using one single set of data remained a reliable way to proceed. Moreover, performing a limited number of

acquisitions (20 scans for each acquisition condition in the present case) allowed to reduce the operating time of the device, which is appreciable when working in a clinical environment. The second limitation faced is that the paradigm worked with (when the signal, location and background are exactly known) was simplified and therefore different from real anatomical conditions. The results could nevertheless be used in order to assess the performances of the tested IR algorithms, but it is worth to mention that there is room for a more complex study on the subject.

CONCLUSION

Nowadays, assessing CT image quality cannot be done with image space metrics anymore. Moreover, evidence indicates that frequency metrics should not be used either when working with IR. However, the task-based tool used in this investigation (CHO model observer associated to the DDoG channels) successfully demonstrated its ability to reproduce the human's response in a low-contrast detection task, thus establishing its reliability for image quality assessment.

The results obtained with this tool revealed that the ADMIRE algorithm led to high PCs even in situations where the target was harder to detect (i.e. low $CTDI_{vol}$ and contrast level). Also, using higher ADMIRE strength led to PC improvement, particularly in the low $CTDI_{vol}$ range. Therefore, using those benefits to keep the image quality equivalent to what was previously obtained would enable to spare some delivered dose.

All those elements suggest that the patient dose could be further optimised and reduced thanks to the use of the ADMIRE algorithm and this new CT unit.

REFERENCES

1. Samara, E. T., Aroua, A., Bochud, F. O., Ott, B., Theiler, T., Treier, R., Trueb, P. R., Vader, J. P. and Verdun, F. R. *Exposure of the Swiss population by medical x-rays: 2008 review.* Health Phys. **102**, 263–270 (2008).

2. Schindera, S. T., Odedra, D., Raza, S. A., Kim, T. K., Jang, H. J., Szucs-Farkas, Z. and Rogalla, P. *Iterative reconstruction algorithm for CT: can radiation dose be decreased while low-contrast detectability is preserved?* Radiology **269**, 511–518 (2013).

3. Brunner, C. C., Abboud, S. F., Hoeschen, C. and Kyprianou, I. S. *Signal detection and location-dependent noise in cone-beam computed tomography using the spatial definition of the hotelling SNR.* Med. Phys. **39**, 3214–3228 (2012).

4. Barrett, H. H., Myers, K. J., Hoeschen, C., Kupinski, M. A. and Little, M. P. *Task-based measures of image quality and their relation to radiation dose and patient risk.* Phys. Med. Biol. **60**, R1–R75 (2015).

5. International Electrotechnical Commission. International Standard IEC 60601-2-44. *Medical Electrical Equipment – Part 2-44: Particular Requirements for the Basic Safety and Essential Performance of X-Ray Equipment for Computed Tomography*, 3rd edn. IEC (2009).

6. Myers, K. J. and Barrett, H. H. *Addition of a channel mechanism to the ideal-observer model.* J. Opt. Soc. Am. **4**, 2447–2457 (1987).

7. Gallas, B. D. and Barrett, H. H. *Validating the use of channels to estimate the linear model observer.* J. Opt. Soc. Am. **20**, 1725–1738 (2003).

8. Beutel, J., Kundel, H. L. and Van Metter, R. L. *Handbook of Medical Imaging. Volume 1. Physics and Psychophysics.* SPIE Press (2000).

9. Castella, C., Abbey, C. K., Eckstein, M. P., Verdun, F. R., Kinkel, K. and Bochud, F. O. *Human linear template with mammographic backgrounds estimated with a genetic algorithm.* J. Opt. Soc. Am. **24**, B1–12 (2007).

10. Yu, L., Leng, S., Chen, L., Kofler, M., Carter, R. E. and McCollough, C. H. *Prediction of human observer performance in a 2-alternative forced choice low-contrast detection task using channelized Hotelling observer: impact of radiation dose and reconstruction algorithms.* Med. Phys. **40**, 041908 (2013).

11. Tseng, H. W., Fan, J., Kupinski, M. A., Sainath, P. and Hsieh, J. *Assessing image quality and dose reduction of a new x-ray computed tomography iterative reconstruction algorithm using model observers.* Med. Phys. **41**, 071910 (2014).

12. Abbey, C. K. and Barrett, H. H. *Human- and model-observer performance in ramp-spectrum noise: effects of regularization and object variability.* J. Opt. Soc. Am. **18**, 473–488 (2001).

13. Efron, B. and Tbishirani, R. J. *An Introduction to the Bootstrap.* CRC Press (1994).

14. Leng, S., Yu, L., Zhang, Y., Carter, R., Toledano, A. Y. and McCollough, C. H. *Correlation between model observer and human observer performance in CT imaging when lesion location is uncertain.* Med. Phys. **40**, 081908 (2013).

15. Barrett, H. H. and Myers, K. J. *Foundation of Image Science.* John Wiley & Sons (2004).