Assessing computed tomography image quality for combined detection and estimation tasks

Hsin-Wu Tseng
Jiahua Fan
Matthew A. Kupinski

Assessing computed tomography image quality for combined detection and estimation tasks

Hsin-Wu Tseng, a,b Jiahua Fan, b and Matthew A. Kupinski a,*

aUniversity of Arizona, College of Optical Sciences, Tucson, Arizona, United States
bGE Healthcare, CT Systems Engineering, Waukesha, Wisconsin, United States

Abstract. Maintaining or even improving image quality while lowering patient dose is always the desire in clinical computed tomography (CT) imaging. Iterative reconstruction (IR) algorithms have been designed to allow for a reduced dose while maintaining or even improving an image. However, we have previously shown that the dose-saving capabilities allowed with IR are different for different clinical tasks. The channelized scanning linear observer (CSLO) was applied to study clinical tasks that combine detection and estimation when assessing CT image data. The purpose of this work is to illustrate the importance of task complexity when assessing dose savings and to move toward more realistic tasks when performing these types of studies. Human-observer validation of these methods will take place in a future publication. Low-contrast objects embedded in body-size phantoms were imaged multiple times and reconstructed by filtered back projection (FBP) and an IR algorithm. The task was to detect, localize, and estimate the size and contrast of low-contrast objects in the phantom. Independent signal-present and signal-absent regions of interest cropped from images were channelized by the dense-difference of Gauss channels for CSLO training and testing. Estimation receiver operating characteristic (EROC) curves and the areas under EROC curves (EAUC) were calculated by CSLO as the figure of merit. The one-shot method was used to compute the variance of the EAUC values. Results suggest that the IR algorithm studied in this work could efficiently reduce the dose by ~50% while maintaining an image quality comparable to conventional FBP reconstruction warranting further investigation using real patient data.

Keywords: computed tomography; iterative reconstruction; channelized scanning linear observer; detection; estimation; estimation receiver operating characteristic; EROC curves.

Paper 17162PR received Jun. 20, 2017; accepted for publication Oct. 27, 2017; published online Nov. 21, 2017.

1 Introduction

The advent of multiple detector computed tomography has led to various approaches in medical computed tomography (CT) imaging. Its rapid volume acquisition makes CT angiography and CT colonography routine. However, the wide and increasing usage of CT in modern medicine has increased patient dose, which has become a concern in the CT community. Discussions regarding the associated estimated risk for radiation-induced cancer have been well publicized. Regardless of the merits of these various arguments, it is always desirable to limit the dose to the patient. Consequently, different hardware and algorithm solutions have been developed to maintain or even improve the image quality while the radiation dose is reduced. The traditional CT reconstruction algorithm filtered back projection (FBP) is well known for its speed and robust image quality. However, FBP images suffer from noise and artifact contaminations especially in low radiation dose conditions. To lower the radiation dose without sacrificing image quality, several iterative reconstruction (IR) algorithms were developed and introduced commercially in the past several years. The challenge of reducing dose is to maintain a clinically acceptable image quality while decreasing exposure level.

To evaluate the image qualities of different image systems, state-of-the-art medical image quality assessment methods that extract desired information from the images for clinically interesting tasks are now known to be good choices. A clinically interesting and relevant task might be a detection task that requires distinguishing normal cases from diseased cases. Another example of a clinically relevant task is an estimation task where the observer must provide information regarding, for example, the size and contrast of the lesions. Image quality in medical systems should be measured by an observer performing a task or tasks of clinical interest. However, there is not a clear choice for this task in CT imaging. Our previous studies have shown different dose-savings depending upon the choice of task. In this study, we focused on the application of the channelized scanning linear observer (CSLO) on CT images to quantitatively evaluate the performance of different reconstruction algorithms under tasks that include signal detection, localization, and estimation of size and contrast in an attempt to obtain a more complete picture of dose savings. It should be noted that this paper will focus on a mathematical observer model and that validation of this observer model against human performance will not be discussed in this paper but will be addressed in a future publication. The outline of this paper is as follows: Sec. describes the materials and methods used. More specifically, Sec. illustrates the idea of the scanning linear observer (SLO); Sec. depicts the selection of channels and the concept of the channelization mechanism applied on the SLO; Sec. depicts...
outlines the training dataset, testing dataset, and the estimation
method of variance for observer study; Sec. 2.4 provides the
brief ideas of IR algorithm; and Sec. 2.5 details the phantoms
used in this study and the process of data generation and acquis-
tion. Results of dose reduction based on the comparison between algorithms are then shown in Sec. 4 followed by con-
cclusions and discussions in Sec. 4.

2 Materials and Methods

2.1 Scanning Linear Observer

In this study, we focus on the application of the CSLO on CT
imaging systems. The concept of scanning linear estimation is
to approximate the mode of the posterior density and perform a
pseudo maximum a posteriori (MAP) estimation. This requires a
scan of parameter space to compare solutions and find the maxi-
mum. The general formula of MAP estimation is

$$\hat{\theta}_{\text{MAP}} = \arg \max_{\theta} \left[ \frac{p(\theta|g)}{p(g)} \right] = \arg \max_{\theta} \frac{p(g|\theta)p(\theta)}{p(g)}.$$  (1)

where $\theta$ and $\hat{\theta}$ are the parameters we are interested in and
the estimated parameter, respectively. The probability density function $p(g|\theta)$ is the likelihood of data conditioned on the parameters to be estimated. To easily optimize this function, we will consider the Gaussian likelihood. Note this approximation does not imply that joint pdf $p(g, \theta)$ is also Gaussian. Based on this approximation, the conditional likelihood $p(g|\theta)$ can be described by

$$p(g|\theta) = \frac{1}{\sqrt{2\pi^d \det(K_{\theta\theta})}} \exp \left\{ -\frac{1}{2} \| g - \bar{g}(\theta) \|^2 K_{\theta\theta}^{-1} \| g - \bar{g}(\theta) \| \right\}.$$  (2)

where $\bar{g}(\theta)$ is the mean image averaged over the parameters $\theta$, $K_{\theta\theta}$ is the sample covariance matrix conditioned on the parameters, and $\det(\cdot)$ is the determinant of the matrix. Instead of evaluating the covariance matrix and its inverse for every parameter, our second approximation is to use the mean of $K_{\theta\theta}$ averaged over all $\theta$ as was done in Ref. 3. To avoid the exponential term in Eq. (2), the common strategy is to operate the natural logarithm on both sides and ignore the term independent of the parameters $\theta$, which leads to

$$\ln[p(g|\theta)] \approx -\frac{1}{2} \| g - \bar{g}(\theta) \|^2 K_{\theta\theta}^{-1} \| g - \bar{g}(\theta) \| + \ln[p(\theta)].$$  (3)

Thus, the scanning linear estimator that maximizes the pos-
terior density under these approximations is equivalent to

$$\hat{\theta}_{\text{SL}}(g) = \arg \max_{\theta} \{ g(\theta)^T K_{\theta\theta}^{-1} g(\theta) - \frac{1}{2} \| g(\theta) \|^2 K_{\theta\theta}^{-1} \| g(\theta) \| + \ln[p(\theta)] \}.$$  (4)

The first term is a linear operation applied on the testing image data by $g(\theta)^T K_{\theta\theta}^{-1}$ from the training dataset. The second term is a shifted term due to the different parameters $\theta$. The third term $\ln[p(\theta)]$ is assumed to be a flat prior, where all values of $\theta$ are equally likely. Thus, this third term is a constant, independent of $\theta$, and ignored. So, the final equation becomes

$$\hat{\theta}_{\text{SL}}(g) = \arg \max_{\theta} \{ g(\theta)^T K_{\theta\theta}^{-1} g(\theta) - \frac{1}{2} \| g(\theta) \|^2 K_{\theta\theta}^{-1} \| g(\theta) \| \}.$$  (5)

This observer operates on the data linearly, even though, in
general, the linear template is a nonlinear function of $\theta$. In the
estimation process, the observer seeks the value of $\theta$ that will
maximize this linear operation and give the estimated parameter
$\hat{\theta}_{\text{SL, est}}$.

2.2 Channel Selection and Channelization

As mentioned earlier, we chose to use the CSLO to perform our
combined detection/estimation tasks. This observer model has the
benefit of being computationally practical and also relevant
in terms of approximating human-observer performance. Many
observer models, including the CSLO, require the estimation
and inversion of a covariance matrix. The size of the regions
of interest (ROI) patches in our study is $M = 100 \times 100$ pixels.
The minimum number of sample images to achieve an invertible
estimate of an $M \times M$ covariance matrix is $M + 1$. To make
the inverse of a covariance matrix practical, we reduce the dimen-
sion of dataset through channelization. The channels selected in
our study are 10 dense-difference of Gauss (DDOG) chan-
nels. The 10 DDOG channels employed in this study have been demonstrated to match the human-observer behavior in
signal-detection tasks for radially symmetric signals. In addi-
tion to reducing the dimension of the covariance matrix,
DDOG channels also have been proved to have a property of mimicking human visual system in the pure detection
task. These channels have not been verified to match human-
observer performance when used in a scanning observer as we are using them in this work. However, we chose to use
this observer model, because the channels are based on the
human visual system, even if the scanning mechanism is not.

The channelization process maps the image onto the chan-
nels. This process can be expressed as

$$g^c = Tg,$$  (6)

where the superscript $c$ refers to channelized data and trans-
spanese, respectively, and $T$ is the channels. The dimension of $g$ is $M$, which, for this study, is $100^3$. The dimension of $g^c$ is 10 for the 10 DDOG channels. Thus, the minimum number of training images to achieve an invertible estimate of the covariance matrix is 11. The channelized image data $g^c$ is used as input in
Eq. (5), and the final equation for CSLO in this study becomes

$$\hat{\theta}_{\text{CSL}}(g) = \arg \max_{\theta} \{ g_{\text{ch,training}}^c(\theta)^T K_{\theta\theta,\text{training}}^{-1} g_{\text{ch,training}}^c(\theta) - \frac{1}{2} g_{\text{ch,training}}^c(\theta)^T K_{\theta\theta,\text{training}}^{-1} g_{\text{ch,training}}^c(\theta) + \ln[p(\theta)] \}.$$  (7)

where $g_{\text{ch,training}}^c(\theta)$ is the average of signal-present image from training image dataset [Eq. (6)], $K_{\theta\theta,\text{training}}$ is the covariance matrix calculated from $g_{\text{ch,training}}^c(\theta)$ and averaged over different signal parameter vectors $\theta$ [Eqs. (5) and (6)], and $g_{\text{ch,testing}}$ is either a signal-present or signal-absent image. The second term in Eq. (5) is independent of testing data, but it has to be retained since it is $\theta$ dependent.
2.3 Tasks and Test Statistics

The CSLO was designed for pure estimation tasks but can be readily extended to include detection by including the value of the maximum in Eq. (7) as the test statistic. Thus, the test statistic that is used to decide whether the signal will be called present or not is the maximum of Eq. (7) instead of the argument maximum. This is similar to the scanning Hotelling observer. Note here that if \( g_{\text{ch,testing}} \) in Eq. (7) is a signal-present image, then the test statistic generates a random variable as \( t_{\text{CSL,2}} \); on the other hand, if \( g_{\text{ch,testing}} \) is from signal-absent image pool, then the test statistic generates a random variable \( t_{\text{CSL,1}} \). A decision that the signal is present is made when the test statistic is greater than the test-statistics threshold; otherwise, the observer decides that the signal is absent. For combination tasks that include detection and estimation, the CSLO scans parameters in the parameter domain to search for \( \theta \) that maximizes Eq. (7). \( \theta \) is a parameter vector including locations, sizes, and contrasts of signals. When the estimated parameters \( \theta_{\text{CSL}} \), the output of Eq. (7), are equal to the true parameters \( \theta_{\text{true}} \), the estimation is counted as being correct.

Note that the implementation of Eq. (7) is potentially non-linear in \( \theta \) but is linear in the image data. Thus, we refer to the implementation of Eq. (7) as an EROC template where the template itself depends on the parameters to be estimated (e.g., location, size, and contrast). To implement location estimation, we crop the EROC templates down (Fig. 1) to exclude the mostly zero portions of the template. This increases computational efficiency substantially. The cropped EROC template contains all of the information of original image template but with smaller size. With this approximation, a cropped EROC template was scanned at every possible location and for every possible value of size and contrast on every testing image. There were \( 41 \times 41 \) possible locations, 5 possible sizes, and 4 possible contrasts in the signal-present images. A statistics map (Fig. 2) was generated after the cropped EROC template was scanned over a single testing image. If the test statistic is above threshold, the observer model then chooses the location of signal based on the highest pixel value in the statistics map and estimates the size and contrast for the given signal in the testing image based on the value of \( \theta \). For signal-present images, if the difference between the estimated location and the true location (the center) was less than location threshold (the radius of the smallest signal) and the estimated size and contrast were equal to the true values, this combined detection and estimation was considered successful. According to the test-statistics thresholds, two test statistic distributions (\( t_{\text{CSL,2}}, t_{\text{CSL,1}} \)) and the results of estimation, an estimation receiver operating characteristic (EROC) curve can be generated. The area under the EROC curve (EAUC) value is the figure of merit used in this study. The higher the EAUC value, the better the image quality of the system under the combination of detection and estimation tasks.

2.4 Training, Testing, and Variance Estimation

The training data are used to estimate the parameters that define the model observers, i.e., the means and the covariance matrices. As mentioned before, although the minimum required training
samples’ size is 11, to have more accurate estimate statistics, 200 samples were used for training and 300 samples were used for testing. For the CSLO, the training dataset is used to estimate $\bar{g}_{2,\text{training}}(\theta)$, and $K_{g_{2,\text{training}}}$, and the equations are given by

$$g_{2,\text{training}}^\text{ch}(\theta) = \frac{1}{N} \sum_{i=1}^{N} g_{2,i}^\text{ch,training}(\theta), \quad (8)$$

$$K_{g_{2,\text{training}}}(\theta) = \frac{1}{N} \sum_{i=1}^{N} [g_{2,i}^\text{ch,training}(\theta) - g_{2}^\text{ch,training}(\theta)]^t \times [g_{2,i}^\text{ch,training}(\theta) - g_{2}^\text{ch,training}(\theta)]. \quad (9)$$

$$\bar{K}_{g_{2,\text{training}}} = \frac{1}{\text{number of } \theta} \sum_{\theta} K_{g_{2,\text{training}}}(\theta). \quad (10)$$

Once these parameters were estimated, the testing images were substituted into Eq. (7) to estimate the EAUC values.

There are several ways to estimate the variance of the EAUC: bootstrap, jackknife, and shuffle methods. They all have been studied previously. In this work, the variance of EAUC values is estimated by a completely nonparametric and unbiased approach, referred to as the Barrett–Clarkson–Kupinski method or the one-shot method. This variance-estimation method is a variant of the original Dorfman–Berbaum–Metz technique but does not rely on resampling techniques.

### 2.5 Computed Tomography Iterative Reconstruction Approach Studied in This Work

To achieve good image quality at low flux CT imaging conditions and maintain reconstruction speeds comparable to FBP, we have designed an IR approach that de-emphasizes the system optics modeling. Imaging-chain statistical noise modeling, physics modeling, and object modeling have been included in this IR algorithm. The detailed description of this algorithm will be presented in a future publication. The focus in this paper is to apply the model observer approach designed above to quantitatively evaluate the dose saving or image quality improvement capability of this IR algorithm developed in house.

### 2.6 Phantoms, Data Acquisition, and Generation

In this work, low contrast (LC) objects embedded in the Medical Imaging Technology Alliance (MITA) CCT 189 phantom (Medical Imaging and Technology Alliance constructed by the Phantom Laboratory) were imaged on a GE Discovery CT750 HD CT system. Examples of signal-present images are shown in Fig. 3. In this study, the MITA phantom was used for signal detection and signal size and contrast estimation task. The properties of LC objects in this phantom are listed in Table 1. Only axial scans were used in this work. The large body bowtie was used for the MITA CT IQ LCD phantom with the body ring attached. The slice thickness was 0.625 mm, and the collimator aperture used was 20 mm.

In this study, the x-ray tube current was varied to achieve different radiation dose levels. For each dose level, 50 identical images were acquired.

*Fig. 3* Examples of images used for combination of detection, size and contrast estimation task. All of the images shown are the averaged results of 500 FBP images. The various signal parameters are: (a) 3 mm 14 HU, (b) 5 mm 7 HU, (c) 7 mm 5 HU, (d) 10 mm 3 HU, (e) 15 mm 14 HU, (f) 15 mm 7 HU, (g) 15 mm 5 HU, and (h) 15 mm 3 HU.
scans were acquired. A total of 10 signal-present and signal-absent ROI pairs were extracted from different longitudinal locations (along the CT system table direction) from each scan. The 50 scans and 10 extracted ROI pairs per scan resulted in 500 individual ROI pairs for each LC object at every dose level. Images were reconstructed at a field of view 180 mm with a matrix size of $512 \times 512$ image pixels. A random order of 500 images was split into training and testing image datasets. The same randomized sequence was used for every study. The center of each signal was determined by analyzing the mean image. The signal was always in the center of the ROI, and the signal-absent ROIs were extracted from regions distant from the signals to avoid any overlap between signal-present and signal-absent ROIs.

For each object and dose level selected, there were 500 independent signal-present and signal-absent image pairs. These image pairs were randomly split into 200 pairs for training and 300 pairs for testing in the observer study. EROC curves for the combined detection and estimation task were generated, and the corresponding areas under the curves were calculated. As mentioned above, the variances of EAUC values were estimated via the one-shot method.29

Table 1 MITA body phantom signal parameters.

<table>
<thead>
<tr>
<th>Object</th>
<th>Diameter (mm)</th>
<th>Contrast relative to background (HU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>3</td>
</tr>
</tbody>
</table>

3 Results

FBP and IR were compared for different tasks. Quality-dose characteristic (QDC) curves33 (Fig. 4) were generated by plotting different dose levels and their associated EAUC values for both FBP and IR algorithms. For each task, 50 scans were made for each dose level and there were 10 different dose levels as shown in Fig. 4. The results show that the QDC curve of IR is higher than that of FBP at all dose levels. It suggests that the IR is better than FBP in terms of image quality at the same dose level. A comparison could also be made across dose levels for an equal EAUC value. This comparison indicates that the required dose level is lower for IR than FBP to achieve the same image quality. Based on these QDC curves, achievable dose reduction can be easily derived. All objects in the MITA phantom were scanned axially for all experiments using the parameters mentioned above, and signal-location-unknown detection, size, and contrast estimation were performed.

Because the location of signal is unknown for the observer, the observer had to localize the signal in the provided testing image first and followed by estimating the size and the contrast of the signal. As mentioned in Sec. 2, the location threshold is about the radius of the smallest signal. The total number of possible locations for signals is $41 \times 41$. Thus, similar to signal location known study, the guessing value of EAUC is given by

$$EQ_{\text{TG}} = \frac{1}{2\pi r^2} \times \text{num}_{\text{obj}} \times \text{num}_{\text{loc}}$$

which, for this problem, is 0.0021. The result indicates that a 50% dose reduction could be achieved using the IR algorithm. EROC curves are shown in Fig. 5. The significant difference

![Combination task of signal location unknown detection, size and contrast estimation](image)

**Fig. 4** An example of QDC curve for the task of signal detection, localization, and estimation of size and contrast.
The improvement of image quality of IR is substantial. This improvement not only helps observers to determine if tumors exist and where the tumors located relatively correctly but also helps observers to distinguish the physical properties such as the size and the contrast of the targeted lesion easily.

4 Discussions and Conclusions

Different IR algorithms have been designed on CT systems to assist the patient radiation dose reduction. The nonlinearity nature of both the CT system and an IR algorithm makes the assumptions of a linear and shift-invariant system far from valid. Thus, conventional image-quality metrics such as modulation transfer function, noise power spectrum, and detective quantum efficiency while well-suited for characterizing detector performance, are not well-suited for overall CT image quality evaluation. In view of this, mathematical model observers have been selected as surrogates to study the task-based image quality performance of CT imaging systems. In this work, we used CSLO to measure the image quality under a more realistic task, the detection, and estimation, at different dose levels for two different reconstruction methods. Radiation dose levels required for LC object detection and estimation in body phantoms were examined and evaluated quantitatively. This work provides an objective way to assess the image quality of CT imaging systems for dose saving evaluation. However, unlike pure detection tasks, currently there is no proper way of adding internal noise in the combination of detection and estimation tasks so internal noise was not used in this study. Our results provide the upper bounds of observer studies.

In this study, different combinations of clinically relevant tasks were considered and evaluated by CSLO. Our results suggest that CSLO could be used as an image quality assessment tool for performance evaluation of CT imaging systems. Future works include the addition of internal noise, correlation with human-observer performance, and head mode phantom evaluations. In addition, we note that conclusions drawn from this study must be tempered by the fact that the phantom is uniform. Thus, adding anatomical variability, possibly through the use of real patient data, will be a possible topic in the future. Additionally, having more signals of different sizes and contrasts in the new design of phantom is in the future plan as well. On the other hand, this method currently is still time-consuming and requires numerous images. To make this method more practical, further design of the CSLO using fewer images is needed and work has progressed in this area. Much like the differences in dose savings we observed when the tasks were made more complex, we expect that there will be differences in the performance when anatomical variability and more complicated signal models are used.

In this study, we evaluated the performance of FBP and an IR algorithm. Our results indicate that the IR algorithm can make a significant dose reduction without compromising image quality. According to our results, the performance of the IR algorithm shows that the noise has been decreased greatly compared to the traditional FBP algorithm for all of the combinations of detection and estimation tasks. Approximately 50% dose reduction can be achieved using this IR algorithm for tasks that combine signal-location-unknown detection and estimation of signal size and contrast. Thus, to achieve the same image quality, the IR algorithm requires less x-ray radiation exposure to patients. To provide this benefit of lower radiation dose to patients, our results support the use of IR in clinical CT. Further study is needed with anatomical variability, real data, and even more complicated tasks to further validate this dose-reduction claim.

Disclosures

Dr. Jiahua Fan is an employee of GE Healthcare, Waukesha, Wisconsin, which is the organization that funded this research project.

Acknowledgments

This work was supported by GE Healthcare, Waukesha, Wisconsin. The authors appreciate the technical discussions and support from members of Advanced Sciences and Engineering Team in GE Healthcare.

References


Hsin-Wu Tseng received his PhD from the College of Optical Sciences, University of Arizona, Tucson, Arizona, USA, in 2015. Currently, he is a postdoctoral scientist in biomedical imaging innovation and clinical translation next-gen CT Research Lab, Department of Medical Imaging, College of Medicine, University of Arizona. His current research interests and expertise include image quality, design, development, and clinical translation of novel x-ray imaging systems.

Jiahua Fan received his PhD in electrical and computer engineering from University of Arizona, Tucson, Arizona. After working on medical-display and medical image processing as an assistant research scientist at University of Arizona, he joined GE Healthcare, Waukesha, Wisconsin. Currently, he is a principal scientist and works on image quality, physics, calibration, and reconstruction topics of CT systems. He has over 70 scientific publications and 16 patents issued.

Matthew A. Kupinski is a professor at the University of Arizona with appointments in the College of Optical Sciences, Department of Medical Imaging, and the Program in Applied Mathematics. He received his BS degree in physics from Trinity University in San Antonio, Texas, in 1995, and received his PhD in 2000 from the University of Chicago. He has worked in diverse areas of imaging including x-ray, gamma-ray, diffuse optical, magnetic resonance, and neutron imaging.