PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

A visibility overshoot index for interventional x-ray image quality assessment

Asli Kumcu, Ljiljana Platisa, Bart Goossens, Amber Gislason-Lee, Andrew Davies, et al.

Asli Kumcu, Ljiljana Platisa, Bart Goossens, Amber J. Gislason-Lee, Andrew G. Davies, Gerard Schouten, Dimitri Buytaert, Klaus Bacher, Wilfried Philips, "A visibility overshoot index for interventional x-ray image quality assessment," Proc. SPIE 12467, Medical Imaging 2023: Image Perception, Observer Performance, and Technology Assessment, 124670T (3 April 2023); doi: 10.1117/12.2652548

SPIE.

Event: SPIE Medical Imaging, 2023, San Diego, California, United States

A visibility overshoot index for interventional X-ray image quality assessment

Asli Kumcu[a], Ljiljana Platiša[a], Bart Goossens[a], Amber J. Gislason-Lee[b], Andrew G. Davies[b], Gerard Schouten[c,d], Dimitri Buytaert[e], Klaus Bacher[e], and Wilfried Philips[a]

[a]TELIN-IPI-imec, Ghent University, Ghent, Belgium
[b]Division of Medical Physics, University of Leeds, Leeds, UK
[c]Fontys Hogescholen, Eindhoven, the Netherlands
[d]Philips Healthcare, Best, the Netherlands
[e]Department of Human Structure and Repair, Ghent University, Ghent, Belgium

ABSTRACT

Dose reduction remains an important goal in interventional X-ray. We propose an image quality (IQ) measure called the Visibility Overshoot Index. Given a patient image and a specified clinical task, the index quantifies the maximum acceptable dose reduction. The dose control system can then use this information to deliver the minimum dose necessary for detection of clinical signals, reducing unnecessary radiation exposure. We developed an experimental visual model to estimate signal detectability as a function of image features such as noise and signal contrast. The model is used to find a feature's threshold – the maximum change in noise or signal contrast where signal detectability remains possible. An automated algorithm measures the magnitudes of these features on a frame. Visibility Overshoot is expressed in terms of the image features: the Noise Overshoot and Contrast Overshoot indices are the ratio of the threshold to measured noise/contrast. The indices demonstrate good agreement with detector dose, Channelized Hotelling Observer results, and clinicians' judgments. In our study of a cylindrical object phantom acquired at seven dose levels, we found that the Noise Overshoot index is linearly related to the square root of detector dose and the CHO detectability index, with Pearson correlation 0.995–1.0 for signals 1-4 mm diameter. For interventional cardiology and neurology sequences acquired at standard and 25–30% dose, the index and clinicians rank IQ similarly. Results on the phantom suggest at least 15% dose reduction could be achieved in fluoroscopy mode. Our patient-specific IQ approach could bring additional dose savings to clinical practice.

Keywords: image quality, video quality, patient-specific, task-based, fluoroscopy, angiography, dose reduction, dose optimization

1. INTRODUCTION

We propose a novel method to measure patient-specific image quality (IQ) in interventional X-ray images, with the goal of improving dose reduction for a given clinical task. Interventional X-ray systems are used to conduct procedures where imaging and intervention occur simultaneously, such as navigating a catheter through blood vessels, diagnosing a blocked coronary artery, and treatment with stent placement. In angiography procedures, which is the focus of this work, patients are exposed to X-rays while a contrast medium is injected into blood vessels. This shows the vessels as dark, low intensity objects on a monitor. Depending on the procedure, X-ray pulses are typically delivered at 2 to 15 frames per second (FPS). The doctor – an interventionalist – is able to dynamically view anatomy in real-time. Total X-ray exposure time to both the patient and staff can range from several minutes of fluoroscopy and 10 to 15 short angiography runs during a standard procedure, such as the ones in this study,[1,2] to over one hour of fluoroscopy for more complex procedures.[3] Fluoroscopy is a lower-dose acquisition used for navigation, and the higher-dose angiography run is for diagnostic tasks. Although modern dose reduction techniques have significantly reduced X-ray exposure for standard procedures,[4–6] they have in

Send correspondence to A.K.: asli.kumcu@ugent.be

Medical Imaging 2023: Image Perception, Observer Performance, and Technology Assessment, edited by Claudia R. Mello-Thoms, Yan Chen, Proc. of SPIE Vol. 12467, 124670T
 · 1605-7422 · doi: 10.1117/12.2652548

turn facilitated longer procedures of increasing complexity,[3,7] effectively limiting the overall savings to both patients and staff. Therefore, dose reduction remains an important goal in interventional X-ray imaging.

A fundamental goal of radiation protection in X-ray imaging is to deliver *as low as reasonably achievable* dose of ionizing radiation, in which its benefits outweigh its harms (ALARA principle[8]). In angiography, the 'benefits' of X-ray exposure can be expressed in various ways, such as enabling the visibility of clinical objects during fluoroscopy (e.g. guidewires, stents, and needles) and the accurate performance of a diagnostic task during angiography.[9] Both are a *task-based* approach to measuring *image quality* – the preferred approach to optimizing the performance of medical imaging systems. That is, a system's IQ is defined as the efficacy of an image for a given detection task, rather than the image's technical characteristics such as noise and contrast.[10] Efficacy is measured as the average performance of a human observer for conducting a detection task.[10] For this reason, automated measurement of IQ in medical applications is challenging.

Model observers (MO) are a class of algorithmic observers that predict human detection performance and therefore directly quantify IQ. Ideal and quasi-ideal observers and the Channelized Hotelling Observer (CHO) are several examples. MOs are regularly used for offline IQ/dose optimization and have a long history in interventional X-ray – for example to optimize dose-pulse/frame rate frequency,[11,12] motion blur,[13] contrast enhancement[14] and noise reduction [15,16] algorithms, X-ray detectors,[17] and general IQ.[18–20] However, MOs such as the CHO require carefully selected datasets and their detection capabilities do not generalize well beyond the datasets used for training. Others are only suitable for signal known exactly / background known exactly (SKE/BKE) tasks. Thus, the suitability of most MOs for evaluating the IQ of patient images is limited.

For on-line task-based dose optimization, the latest automatic dose control (ADC) systems used by several interventional X-ray manufacturers[21,22] simulate the IQ of an object embedded in a patient. IQ is measured as the (quasi) ideal observer's spatial frequency corrected signal difference to noise ratio (SDNR).[23] The approach incorporates physical modeling of the properties of the patient (based on estimated thickness), the object (e.g. guidewire), and the acquisition parameters to simulate the object's SDNR. The system optimizes and maintains the target IQ level. As far as we are aware, this approach does not measure IQ directly on the acquired image, after simulation. In other modalities, the 'in-vivo detectability index' is a patient-specific method recently developed for computed tomography (CT).[24,25] This approach has shown agreement with human detection accuracy for clinical CT but has not been tested in interventional X-ray imaging, in which sequences are comprised of high-contrast signals viewed in dynamic mode. Other dose optimization approaches include calibrated patient thickness-dose curves[26] and post-processing to enhance low-dose acquisitions,[5] such as denoising or contrast enhancement. However, these approaches specific to interventional X-ray do not quantify the IQ of the *acquired* image, and thus do not take advantage of the possibility that a detection task could potentially be accurately conducted at a lower dose level.

We propose a task-based method in which the task is defined as detection of contrast-enhanced vessels, and ALARA as the minimum dose that achieves this. We hypothesize that greater dose savings can be achieved by exclusively targeting the minimum dose at which the detection task can be accurately performed. The underlying assumption is that beyond a certain point, a reduction in image noise or an increase in signal contrast does not help the doctor execute the task more accurately. Once the interventionalist detects the clinical object, higher dose levels may produce a more pleasing image, but it does not affect the task – vessel detection has already been realized at the lower dose. Therefore the dose control system should deliver the minimum dose necessary to ensure the detection of clinical objects, and no more. We refer to any excess dose above this point as *overshoot*. Determining this minimum point is a more straightforward task in interventional X-ray, where we know in advance that our clinical signal of interest should be present – e.g. contrast-enhanced vessels – and must always be visible to the interventionalist. This is unlike diagnostic and screening radiology – for example mammography – where lesions may be absent in images acquired from healthy patients.

The minimum dose level can be found by identifying the minimum IQ level at which clinical objects become reliably visible to a human observer – the 'detection task'. This is accomplished with an experimental visual model we developed that represents the probability of detecting a signal as a function of the magnitude of *image features*. These features are signal contrast ratio, image noise, and background intensity. Adjusting a feature's magnitude can affect signal detectability, but we are interested in the range of values that do *not* impair detectability. The visual model can directly estimate the acceptable range for a given feature. The target

value is that which results in the lowest dose output. For example, we can determine how much X-ray tube current can be reduced (resulting in higher image noise) before the signal drops below the noise level and can no longer be detected by the clinician. Assume we detect a clinical signal in the image and measure the noise standard deviation in the background as 2%. Using the model, we estimate that noise can be increased up to maximally 5% without affecting the detectability of that signal. We conclude that an *overshoot* has occurred. We call this the *Noise Overshoot index* (o'_σ), defined as the ratio of the two values: $o'_\sigma = 5/2 = 2.5$. Thus, the system can increase noise by a factor 2.5 and the vessel will continue to be visible, if its contrast ratio and background intensity remain constant. We can apply the same procedure to estimate the maximum acceptable loss of contrast: the *Contrast Overshoot index* (o'_c).

We refer to this class of indices collectively as the *Visibility Overshoot index*, shortened to *Overshoot index* or *Overshoot*. The amount of dose reduction is directly calculated from one of the Overshoot indices, depending on which acquisition parameters should be adjusted; we focus primarily on tube current, which affects noise. We found that for a phantom dataset, the Noise Overshoot index is a function of the square root of the detector dose. For this dataset, detector dose is predominantly a function of tube current and pulse duration. When dose is reduced, the amount of Noise Overshoot is proportionally lower. The dose control system can therefore map the desired reduction in Overshoot to a reduction in tube current and/or pulse duration. Using the previous example, if Noise Overshoot is 2.5 and we target an Overshoot of 2 (above the threshold of detectability of Overshoot = 1, as a safety margin), the difference of 0.5^2 is proportion to the acceptable dose reduction.

We envision a workflow that operates in real-time during a patient's procedure. In the first frame, the system delivers its reference dose level, for example based on the patient's weight. Our proposed method measures image features in the frame and estimates the amount of Noise Overshoot. If Overshoot occurs, the equivalent reduction in tube current and pulse duration are computed from the index and the desired Overshoot target. Subsequent frames are imaged at the lower dose level, as long as the field of view remains unchanged. Results on the phantom dataset show that Noise Overshoot can help achieve at least 15% dose reduction in fluoroscopy mode, when the target Overshoot is set to 2. The index also shows high correlation with a CHO for a cylindrical object phantom and corresponds with clinicians' image quality judgments for two clinical datasets. The results demonstrate that automatic dose reduction can be achieved using the Noise Overshoot index.

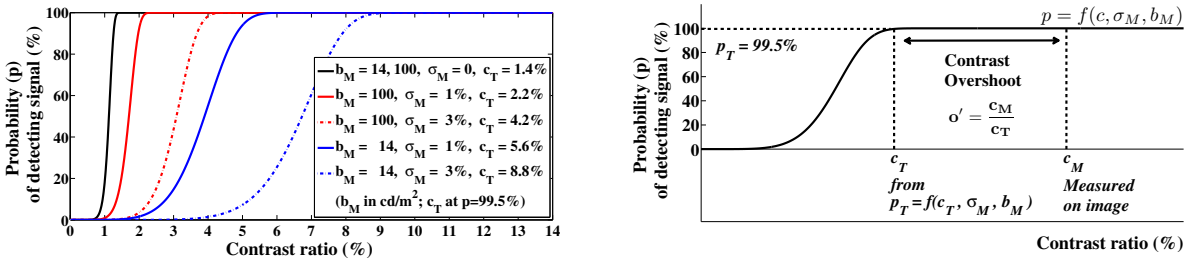
This work contains several novelties. First, the *Overshoot* appears to be a new IQ FOM, as far as we are aware. Existing measures such as CNR and detectability FOMs do not explicitly quantify the limits of dose reduction. Second, computation of the Overshoot relies on an unconventional use of the well-known psychovisual model of contrast detection, and contains several differences compared to typical implementations. IQ algorithms that employ visual models, such as the contrast sensitivity function (CSF), often target non-medical applications and are commonly used to assess whether image artifacts (e.g. noise) are below the threshold of detectability, often compared to a reference image. However, we employ a visual model to quantify the IQ of signals above threshold detection ('suprathreshold' range), rather than estimating whether distortions are visible. In addition, our model can operate on single images; a reference image is not required. Unlike the CSF, we do not consider the spatial extent of the signal. However, we do incorporate the effect of background intensity, which has a significant influence on threshold contrast. This also differentiates our model from CNR, which does not account for the impact of background intensity on signal visibility. The third novelty of this work is the fully-automated, patient-specific algorithm. The pipeline automatically extracts clinically relevant signals, estimates image features in the luminance domain, and computes a local and global Overshoot index around clinical signals using the visual model.

In Section 2, we describe the visual model and the Visibility Overshoot index, the automated algorithm, and the datasets and analysis used in this study. The Results are shown in Section 3, in which the Overshoot index is compared to dose, the detectability index from a CHO, and clinicians' judgments. The Discussion is in Section 4 and we end with the Conclusions in Section 5.

2. METHODS

The derivation of the Overshoot index is described in Section 2.1. The fully-automated algorithm that estimates image features and the Overshoot index on patient images is described in Section 2.2. The datasets used in this study are described in Section 2.3, and an explanation of the statistical analysis is in Section 2.4.

2.1 Derivation of Overshoot FOM



(a) Example of visual model output

(b) Derivation of the Contrast Overshoot index

Figure 1. The visual model is used to estimate a signal's threshold contrast c_T and its corresponding Overshoot. In (a), c_T is a function of the background luminance b_M and dynamic noise σ_M around the signal, shown for the experimentally derived model explained in Section 2.1. We define threshold contrast at 99.5% detection probability. Higher noise levels and lower background intensity flatten the response curve. This reflects the observer's decision uncertainty, and shifts c_T to higher thresholds. Values for b_M and σ_M represent realistic image characteristics as measured on the PhantomChest dataset. In (b), the Contrast Overshoot index is derived from the relationship between the estimated c_T and the signal's measured contrast c_M. The Noise Overshoot index is derived using the same principle.

The *Visibility Overshoot Index* (o'), shortened to *Overshoot*, represents the factor by which either image noise or contrast can be adjusted without affecting the clinician's ability to detect a clinical signal. For example, if we the dose control system would reduce X-ray tube current, image noise will increase, while signal amplitude will remain approximately constant. The 'threshold noise' is the maximum noise standard deviation in which the signal remains detectable. We can estimate this value and compare it to the actual noise standard deviation measured on the image. We define the *Noise Overshoot* index with respect to noise standard deviation (σ) as

$$o'_\sigma = \frac{\sigma_T}{\sigma_M}\,, \tag{1}$$

where σ_T is threshold noise and σ_M is the image's current noise level. If we seek maximum dose reduction, Overshoot should be minimized for a given detected signal. Thus the system should deliver the minimum dose necessary for a clinician to reliably visualize the signal, and no more.

We use the same approach to estimate the limit of contrast reduction. The *Contrast Overshoot* for signal contrast ratio (c) is

$$o'_c = \frac{c_M}{c_T}\,, \tag{2}$$

where c_M is the contrast ratio measured for the signal and c_T is 'threshold contrast' – the minimum contrast at which the signal remains detectable. The algorithm that measures σ_M and c_M on images is explained in Section 2.2. Estimates of σ_T and c_T are made from the visual model explained next.

The visual model $p = f(c, \sigma, b)$ expresses the probability p that an observer detects a signal as a function of the image features, where b is the background intensity. The model is comprised of a series of psychometric functions we derive experimentally, represented by Weibull cumulative distribution functions (CDF). When σ and b are held constant, the function is monotonic with contrast c, as shown in Fig. 1. Thus, an observer is more likely to detect a signal as contrast increases. The function undergoes a shift when σ and b are varied, as shown in Fig. 1a for 6 scenarios: dynamic noise with spatial standard deviation 0, 1 and 3%, and background intensities 14 and 100 cd/m^2, selected to represent realistic image characteristics as explained later in this section. The legend shows the corresponding threshold contrast c_T when detection probability reaches 99.5% for that condition. As seen from the figure, more contrast is required to detect signals in noisier and darker backgrounds; c_T is higher. However, detection probability is limited to $0 \leq p \leq 1$; above a certain contrast, p is always 1. Therefore, we seek the lowest contrast where detection remains possible: the threshold contrast c_T when $p = 1$. Threshold contrast is estimated from the inverse CDF; we solve for c_T at a detection threshold of $p_T = 99.5\%$, holding σ_M

and b_M constant. After c_T is estimated, calculation of the Contrast Overshoot is straightforward: $o'_c = c_M/c_T$. Likewise, we solve for σ_T to estimate the Noise Overshoot: $o'_\sigma = \sigma_T/\sigma_M$.

We have evaluated three formulations of the Overshoot and found the ratio of threshold-to-measured value the most meaningful. The two other formulations are the absolute difference between threshold and measured σ or c, and the relative percentage. The absolute difference is not proportion to the change in dose; one unit difference will have a different effect size depending on the feature's magnitude. Relative percentage suffers from the same problem, with absolute difference in the numerator. Moreover, one unit difference will be amplified when the denominator is smaller. The ratio is independent of these effects.

Psychometric functions are derived from Yes/No staircase experiments. Experimental data are fit to the function $p = \gamma + (1 - \gamma - \lambda) \cdot F(x)$, where $F(x)$ is the Weibull CDF: $F(x) = 1 - \exp(-x/\alpha)^\beta$ and x is contrast. Terms γ and λ are the guessing and lapse rates; they constrain the lower and upper asymptotes of the function. Factors α and β define the shape of the CDF for a given test condition. The psychometric functions for intermediate noise and background levels are derived by interpolating the α and β values corresponding to the tested conditions. The factor α is linearly interpolated between two tested background or noise levels; β is logarithmic. Threshold contrast is computed from the inverse CDF for a test condition with known $\alpha, \beta, \gamma, \lambda$, given by: $c_T = \alpha(-\log(1 - \hat{p}_T))^{1/\beta}$ where $\hat{p} = (p_T - \gamma)/(1 - \gamma - \lambda)$ and $p_T = 0.995$. Threshold noise is found by interpolating α, β corresponding to c_M, b_M.

Our choice of $p_T = 99.5\%$ detection probability is unusual. Contrast detection experiments are often designed such that observers achieve 75% correct responses, as this provides the most accurate estimate of the visual response function.[27] Threshold contrast is defined with respect to 75% detection probability. However, reducing the signal's amplitude to this level within a dose control system would render the signal too subtle for diagnostic purposes. Hence we have defined the lower boundary for Overshoot as the threshold contrast at which a signal is correctly detected 99.5% of the time; threshold contrast is determined analytically, as explained above. Our staircase experiments employ the standard 75% target, as detailed below.

Fig. 1a illustrates our motivation for selecting a higher operating point. As seen on the left-most curve, the visual response for images without noise approaches a step function. In noisier and darker backgrounds, the function undergoes two changes. First, the midpoint of the function shifts to the right – more contrast is required to detect the signal, as would be expected. Furthermore, the function becomes flatter. As the task difficulty increases, the observer is unable to consistently detect the signal over a wider range of contrasts. Thus, the observer's decision uncertainty is higher. For example, on a dark, noisy background ($b_M = 14$, $\sigma_M = 3\%$), the observer's range of uncertainty extends to 8.8% signal contrast before reaching maximum detectability. If we would instead select 75% probability, the corresponding contrast threshold is lower, at 7.4%. If the dose control system would target the 7.4% threshold, vessels may be rendered too subtle for diagnostic use. Our choice of $p_T = 99.5\%$ avoids this risk and acts as a safety margin.

Staircase experiments were conducted with a single-interval adjustment-matrix (SIAM) procedure.[28] SIAM adaptively adjusts the signal contrast based on the observer's response for signal positive (SP) and signal negative (SN) stimuli. Contrast thresholds were determined for a range of test conditions: additive white Gaussian noise in the range of 0–8.4% (luminance units) and background intensities of 3, 8, 26, and 170 cd/m^2. These values were selected to reflect realistic image characteristics, based on measurements from the PhantomChest dataset described in Section 2.3. Experiments were conducted on a medical display with dynamic range 0.57–202.7 cd/m^2. Signals were Sloan letters embedded in the test conditions, adapted from Ref. 29. Signals were hypointense, consistent with clinical signals in interventional X-ray. Stimuli were displayed dynamically at 15 FPS. Two interleaved staircases targeted 50% and 75% correct responses. Although the targets in the staircase experiments are lower than $p_T = 99.5\%$, their purpose is to derive the full psychometric function. Per test condition, the staircase responses were fit to psychometric functions using psignifit 3.0.[30] Function parameters were estimated with constrained maximum likelihood estimation.

2.2 Fully automated algorithm

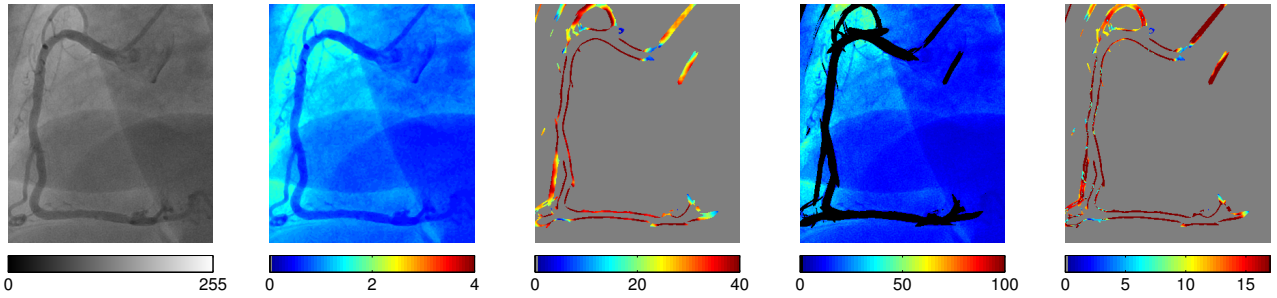
The fully automated algorithm we developed segments contrast-filled vessels from a single frame in the patient image and measures image features in luminance: contrast ratio of vessel edges (c_M), noise standard deviation

(σ_M), and background intensity (b_M), as illustrated in Fig. 2. These features serve as the inputs to the visual model described in Section 2.1, allowing for the calculation of the Overshoot index for the frame. Algorithm parameters were tuned using ground truth extracted from the PhantomChest dataset (see Section 2.3).

Pixel grey levels are converted to luminance (cd/m^2) using the DICOM grayscale display function (GSDF)[31,32] conversion for a generic display. A pixel-to-luminance look-up table (LUT) is generated for a sequence based on the image bit depth and target display characteristics. Display bit depth and minimum/maximum display luminance are based on the staircase experiment described in Section 2.1. All subsequent processing is conducted in luminance domain units, referred to here as *intensity* (i).

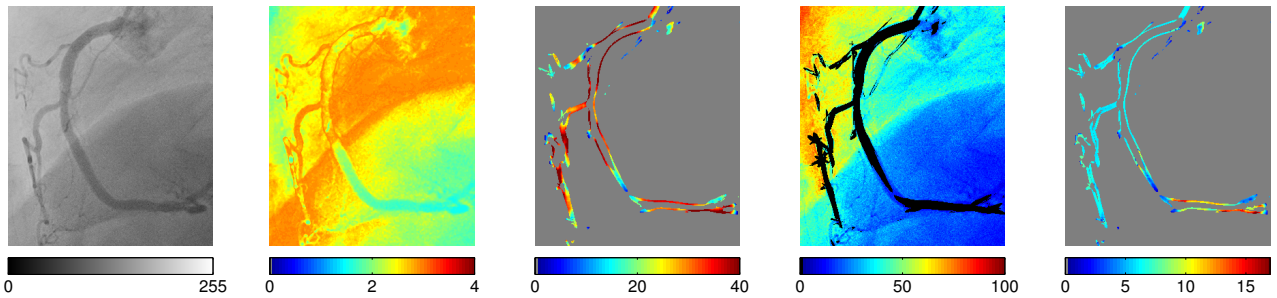
We derive an image noise map σ_M for each frame using an adaptation of Ref. 33. Our implementation generates an intensity-dependent noise model, as typically found in fluoroscopy images, but it also accounts for the effect of post-processing, which distorts the relationship. This adaptation is based on a method originally proposed by Ref. 34, in which the intensity gradients in a uniform image corrupted by Gaussian noise can be described with a Rayleigh distribution, with the peak of the probability density function (PDF) equal to the noise standard deviation σ. In non-uniform images, the peak of the PDF can shift away from the 'true' noise value σ. Ref. 33 makes an adaption for this shift and derives spatio-temporal gradients from first order horizontal and vertical wavelets; one σ value is estimated for the entire image. In our implementation, we compute σ within small intensity bins and employ additional adaptations to remove the shift in the peak. Noise is computed only in the spatial domain.

The adapted noise algorithm employs *Symlets 2* discrete stationary wavelet decomposition to compute the



Input: Xper | $\tilde{\sigma}_M = 0.9$ cd/m^2 *Noise* | $\tilde{c}_M = 25.4\%$ *Contrast ratio* | $\tilde{b}_M = 14.3$ cd/m^2 *Background level* | *Output*: $\tilde{o}'_\sigma = 13.8$ *Noise Overshoot*

(a) Xper



Input: Clarity30 | $\tilde{\sigma}_M = 2.5$ cd/m^2 *Noise* | $\tilde{c}_M = 23.8\%$ *Contrast ratio* | $\tilde{b}_M = 35.4$ cd/m^2 *Background level* | *Output*: $\tilde{o}'_\sigma = 6.0$ *Noise Overshoot*

(b) Clarity30

Figure 2. Examples of the fully automated algorithm's intermediate steps and its output, the Overshoot map. Data from the Cardiac dataset: (a) the standard-dose Xper system, (b) the low-dose Clarity30 system. Overshoot ($\tilde{o}'_\sigma$) is lower in (b) than in (a), indicating lower image quality – the vessels in (b) are closer to the threshold of visibility. This is due to the higher noise level in (b), which is not compensated by the brighter background. The median values are shown.

image gradient magnitude $G = \sqrt{(g_h^2 + g_v^2)}$, where g_h and g_v are the first-order horizontal and vertical Symlet coefficients $CD_1^{(h)}$ and $CD_1^{(v)}$. The gradients G are binned by their corresponding image intensity. Within a bin n with mean intensity i_n and vector of gradients G_n, gradient coefficients are iteratively shrunk to eliminate edge responses, such as anatomical and other structural boundaries in the image. Any wavelet coefficients greater than the shrink factor, computed as the value of the Donoho and Johnstone VisuShrink[35] estimator, are set to zero. The remaining coefficients $\hat{G}_n$ are fit to the Rayleigh distribution with maximum-likelihood estimation; the distribution's scale parameter (PDF peak) is the initial σ_n estimate. After applying a linear shift calibrated to the training data, similar to Ref. 33, a model of the shifted $\hat{\sigma}_n$ values are fit to the intensities i_n over all n bins. In unprocessed fluoroscopy data, intensity-dependent X-ray noise can be modeled with various functions, e.g. as discussed in Ref. 36. However, clinical systems apply non-linear processing to raw acquisitions, e.g. contrast enhancement, noise suppression, or other post-processing, which tends to distort this relationship. Therefore, we construct a linear spline between $(i_j, \hat{\sigma}_j)$ and $(i_{j+1}, \hat{\sigma}_{j+1})$ for each interval $j = 0, 1, .., n-1$, giving $\hat{\sigma}$ as a function of i. We rename $\hat{\sigma}$ to σ_M, to refer to the *measured* noise standard deviation, giving $\sigma_M = f(i)$. Finally, to generate the noise map, for a pixel at coordinate (x, y), the median luminance $(\tilde{i})$ in a 3x3 region of interest (RIO) is used to look up the local noise level $\sigma_M(x, y) = f(\tilde{i}(x, y))$. A new noise map σ_M is generated for each frame.

Contrast ratio c_M and background intensity b_M maps for each frame are computing on image intensities after extracting a vessel mask. Vessel edges are segmented using an implementation of the ProbShrink estimator extended to nonstationary noise.[37,38] This approach estimates the probability of a wavelet coefficient representing a signal embedded in noise. A map of the nonstationary image noise (σ_M) is used to estimate the 'true' correlation between coefficients in a local neighborhood, giving a better estimate of the probability that a coefficient represents a 'true' signal. Our approach employs shearlets[39] – directionally sensitive wavelets that are well suited for rotationally asymmetric ('sheared') signals such as contrast-enhanced vessels. The first shearlet scale with 4 orientations are used for the clinical datasets; the first scale with 2 orientations are used for the PhantomCHO dataset, which contains rotationally-symmetric signals.

Once the vessel mask is extracted, the intensities of the vessels (i_v) and background (i_b) are taken as the median value within small ROIs. The Weber contrast between vessel edges and background is measured as $c_M = (i_b - i_v)/i_b$ within the ROIs. We rename i_b to b_M for notational ease and consistency with the other image features. The noise and contrast Overshoot indices, o'_σ and o'_c, can now be computed by taking the image features c_M, σ_M, b_M as inputs to the visual model from Section 2.1. Indices are computed as the median values within a small sliding window.

2.3 Datasets

This study includes one training dataset (*PhantomChest*) and three testing datasets (*PhantomCHO, Cardiac, NeuroDSA*). An overview of their image characteristics is given in Table 1.

The training dataset *PhantomChest*,[40] generated by the authors (AGL, AD, AK), was used to tune the extraction of image features described in Section 2.2. Acquisitions of a static anthropomorphic chest phantom containing contrast-filled coronary arteries (Radiology Support Devices Alderson Phantoms, Long Beach, USA) were made on an Allura interventional X-ray system (Philips Healthcare, Best, The Netherlands) with and without 10 cm polymethyl methacrylate (PMMA) to simulate standard and large chest thickness, respectively. Sequences of 2 seconds were acquired at six dose levels per phantom size at 15 frames per second (FPS). Image are in raw format without any post-processing or proprietary image processing, except for logarithmic normalization of pixel intensities. Frames are 14-bit, 890 x 890 pixels. Entrance skin dose was calculated as 0.308, 0.643, 0.978, 1.302, 1.973, 2.613 mGy/sec and 2.475, 3.320, 4.140, 4.980, 6.700, 9.330 mGy/sec for the standard and large phantoms, respectively. Ground truth data were generated from manual segmentation of vessels on the highest dose sequences.

The *PhantomCHO*[19] dataset was used to evaluate the Overshoot indices on controlled dose levels and to benchmark Overshoot performance against a state-of-the-art IQ measure – a CHO with Gabor channels and bias correction to estimate the detectability index ($d'_{\infty,o}$ in Ref. 19, which we shorten to d' in this work). The cylindrical object phantom is comprised of a 2.5 mm thick PMMA phantom containing 7 rotationally symmetric objects of 0.5–4.0 mm diameter. Sequences were acquired at 7 dose level for 1204 frames at 30 FPS on a Siemens

Table 1. Image characteristics of interventional X-ray sequences

Dataset	Type	Dose levels	Seq/ dose	Frames/seq	FPS	Bit depth (bpp)	Frame size
PhantomChest[40]	training phantom	12	1	30	15	14	890 x 890
PhantomCHO[19]	phantom	7	1	301	30	8	512 x 512
Cardiac[2]	clinical	2	72 ; 69	61 [34–147]	12.5 ; 15	8	512 x 512
NeuroDSA[1]	clinical	2	40	20 [15–23]	2	14	1000 x 1000

Seq = Sequence. Frames/seq given as: median [minimum–maximum]. FPS = frames per second. bpp = bits per pixel. Frame size in pixels.

Artis Zee VC21 system (Siemens Medical Solutions, Germany). Frames are 8-bit, 512 x 512 pixels. This study analyzed 301 frames. Detector target doses (DTD) are 6, 10, 18, 36, 65, 120, 240 nGy/frame generated from the following acquisition settings: peak tube potential from 93.6 to 96.2 kilovoltage peak (kVp), pulse duration from 3.5 to 9.2 milliseconds (ms), and tube current from 9 to 246 milliamps (mA).

The *Cardiac*[2] dataset is comprised of cardiology procedures acquired in a clinical setting at two dose levels: a standard dose Philips *Xper* system and a *Clarity30* mode – 30% of standard dose and ClarityIQ denoising technology. Seventy-two patients were divided over the two modes, with two views acquired per patient, for a total of 72 sequences for *Xper* and 69 for *Clarity30*. Sequences were acquired at 12.5 FPS (*Xper*) and 15 FPS (*Clarity30*) for a median of 61 frames (minimum/maximum 34–147 frames) per sequence, depending on room, system, and anatomy. Images are 8-bit, 512 x 512 pixels. Kerma air products for the total procedure, including multiple angiographic runs, were measured as 47.4 (33.6–66.5) and 12.0 (7.59–20.1) Gy·cm^2 (median and interquartile range) for *Xper* and *Clarity30*, respectively. Four interventional cardiologists evaluated 3 image attributes (resolution, image contrast, noise) and perceived IQ on a 0–5 scale (Very poor to Excellent).

The *NeuroDSA*[1] dataset is comprised of digital subtraction angiography (DSA) neuroradiology procedures acquired in a clinical setting at two dose levels: a standard dose mode Philips *Xper* FD20/20 biplane and a *Clarity25* mode – 25% of standard dose and ClarityIQ. Twenty patients underwent interventions under both modes, with two views per patient, for a total of 40 sequences per system. Sequences were acquired at 2 FPS for 20 (15–23) frames. Images are 14-bit, 1000 x 1000. Dose Area Products for sequences were measured as 553 (479–682) and 142 (118–170) Gy·cm^2 (median and interquartile range) for *Xper* and *Clarity25*, respectively. Three neuroradiologists evaluated perceived IQ on 3 bolus phases using a 1–5 scale (Unsatisfactory for diagnosis to Superior visualization of vasculature); scores were averaged to generate an overall average perceived IQ score.

2.4 Analysis

For the PhantomCHO dataset, we fit a regression between noise Overshoot o'_σ and detector target dose (DTD), and between o'_σ and the CHO detectability index d'. We calculate the regression coefficients and the Pearson correlation coefficient (PCC) between terms. We derive one Overshoot index per signal diameter per frame (o'_σ) by taking the intra-frame median within an ROI around each signal.

For the Cardiac and NeuroDSA datasets, the IQ of the standard-dose and low-dose systems are compared. ANOVA is used to conduct statistical significance testing between scores for the two systems. ANOVA is Analysis of Variance, a form of linear regression where predictors are categorical (the two systems under test). Clinicians' IQ ratings and the Overshoot indices are analyzed separately, to evaluate global trends. We derive one Overshoot index for a sequence ($\tilde{o}'_\sigma$, $\tilde{o}'_c$) by pooling pixel values: first by taking the intra-frame median, then the inter-frame median. Frames with too few detected pixels are automatically identified and excluded (e.g. pre- or post-bolus frames).

While the previous analysis is useful for offline optimization, one Overshoot index per frame may not be sufficiently informative to drive a real-time dose control system. Patient frames will contain various vessel

diameters and background anatomy, and relying on a global index risks over- or under-estimating the margin for dose reduction. Therefore, we construct a heatmap of Overshoot distributions to examine how its behavior changes by frame and dose. First, the relative frequencies of all o' values for one frame are calculated. Data are grouped into bins of [1 o' unit x 10%] (PhantomCHO) or [1 o' unit x 1% relative frequency] (clinical datasets). For example, the first Overshoot bin is comprised of $1 < o' \leq 2$. The distributions from all frames for all sequences are superimposed. This is plotted as a heatmap (shown in Fig. 5, with color-coding representing the number of frames falling within a bin. The plots are limited to $o' \leq 17$ for easier visualization; many images contain larger Overshoot values corresponding to highly-visible main vessels.

3. RESULTS

Fig. 2 shows an example of the algorithm's intermediate steps and its final output, the Noise Overshoot map. Examples are taken from the Cardiac dataset. Fig. 2a is a frame from the standard-dose Xper system. Fig. 2b is the low-dose Clarity30 system. The Overshoot map for (b) Clarity30 suggests lower image quality – lower median $\tilde{o}'_\sigma$ scores signify that on average, the vessels are closer to the threshold of visibility. As seen from the intermediate outputs, this is due to the higher noise level. Moreover, the brighter background is unable to compensate for the increased noise. The slightly lower contrast level also contributes to lower IQ.

In the following sections, we show the relationship of the Overshoot index to dose and CHO d' in Section 3.1 and Overshoot versus clinical judgments in Section 3.2. We examine the distribution of Overshoot scores within a frame in Section 3.3 and describe how this could potentially be used to set a maximum boundary for dose reduction in a dose control system.

3.1 Overshoot versus dose and d'

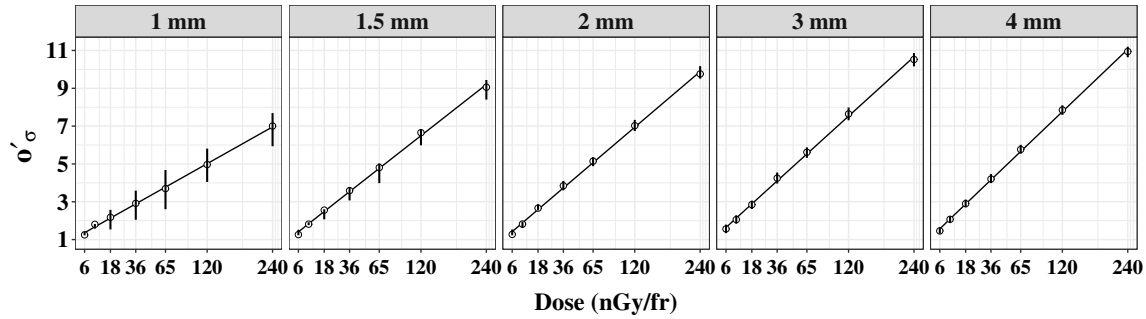
Fig. 3 shows Noise Overshoot o'_σ is positively correlated with both dose and d'. The Pearson correlation coefficients (PCC) between the square root of the dose and Overshoot is between 0.999 and 1, as seen in Fig. 3a. Similarly, d' has a linear relationship to Overshoot as shown in Fig. 3b, with PCC between 0.995 and 1.

As seen on both figures, o'_σ is a function of signal diameter. This effect arises because smaller signals were measured to have lower contrast, leading to smaller Overshoot, even at the same dose level (determined from the intermediate outputs, not shown). Therefore, the regression line between dose and Overshoot is flatter for smaller signals (Fig. 3a). However, the fit with d' changes in the opposite direction, becoming steeper for smaller signals (Fig. 3b). The slope increases to 0.9 as signal diameter decreases to 1 mm. This is because d' is also a function of signal size but increases faster than o'_σ. For example, at 240 nGy/fr, d' for the 4 mm signal is a factor 4.8 greater than d' of the 1 mm signal;[19] for o'_σ this factor is approximately 1.6.

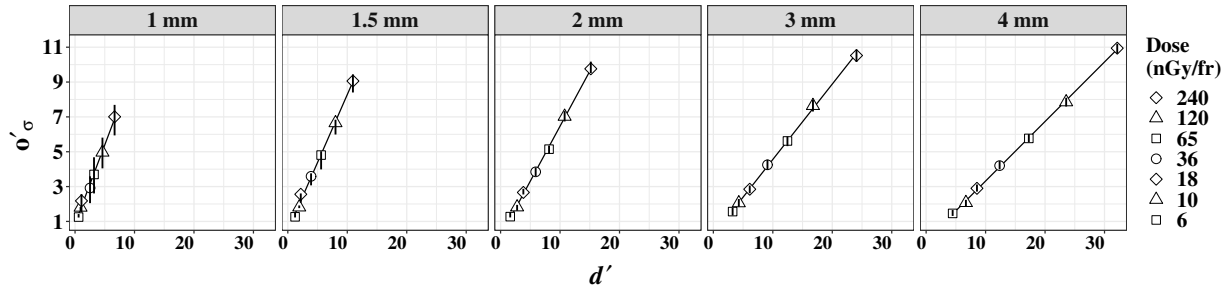
Notably, the algorithm is able to quantify signals near the threshold of visibility ($o'_\sigma \sim 1$) on images acquired at the lowest doses. For example, for the 1 mm signal at 6 nGy/fr, median o'_σ is 1.25. However, the algorithm could not compute Overshoot for the smaller signals (0.5 and 0.75 mm), due to unreliable segmentation at these diameters. Signals smaller than 1 mm diameter correspond to the septal segments branching off the Left anterior descending (LAD), based on coronary artery measurements taken from angiography sequences.[41] Signals 1 mm and larger were correctly segmented, corresponding to all 12 segments of the three main coronary arteries (Right coronary artery, Left main, Left circumflex), 11 of 12 segments of the LAD, and nearly all remaining branch arteries. In clinical practice, the entire length of the vessel would typically be imaged, whereas this phantom represents a cross-section of a vessel. In addition, our automated algorithm is optimized to extract vessel-like objects and not circular targets. This is also the reason for the larger o'_σ error bars on the 1 mm signal. The algorithm does automatically extract signals 1 mm or larger for at least 70% of frames.

3.2 Overshoot versus Clinical judgments

For clinical datasets, o'_σ and o'_c are both in concordance with clinicians' judgments, as seen in Fig. 4. For the Cardiac dataset (Fig. 4a), clinicians judged the noise content of Xper as better (Q_Noise plot, red points), in agreement with the indices. However, clinicians on average rate the two systems as equivalent for Overall IQ (Q_IQ). Evidently, their Overall IQ judgment was not affected by noise, leading to the conclusion reached by Ref. 2 that dose reduction up to 75% can be achieve while maintaining diagnostic image quality. As neither

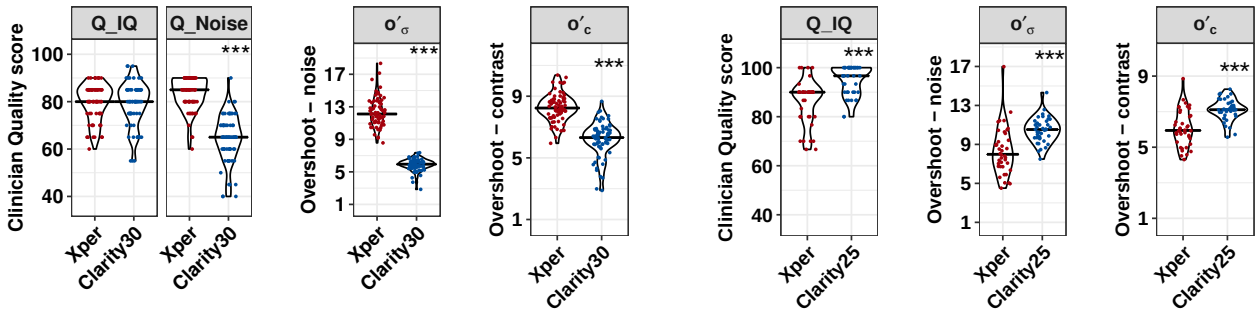


(a) Noise Overshoot versus Detector Target Dose



(b) Noise Overshoot versus CHO d'

Figure 3. Noise Overshoot on the PhantomCHO dataset for the five largest signal diameters: (a) o'_σ versus dose, showing a linear relationship with dose after square-root transformation. Pearson correlation coefficient ranges from 0.999 to 1.0. (b) o'_σ versus d'. Pearson correlation ranges from 0.995 to 1.0. For both plots, the linear regression line is shown. Error bars indicate the 25%-75% quartile range across frames.



(a) Cardiac

(b) NeuroDSA

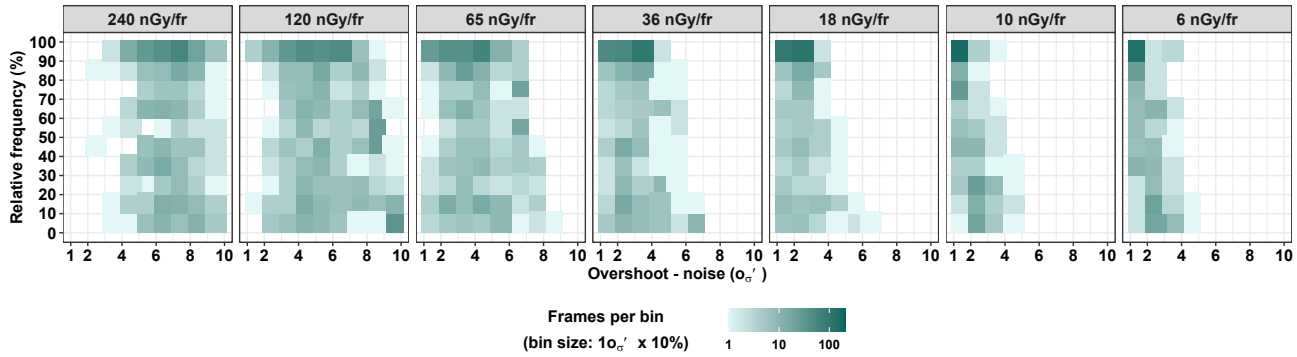
Figure 4. For the two clinical datasets, violin plots showing agreement between clinician's quality judgments and the two visibility Overshoot indices. For all y-axes, higher scores indicate better IQ scores. Individual points reflect the median score for one sequence; points have been horizontally jittered for better visualization. The width of the violin indicates the density of points at that coordinate. Stars indicate level of statistical differences between the two systems. Q_IQ is an overall IQ rating, Q_Noise is the rating of IQ with respect to noise.

index has $\tilde{o}'$ values near 1, our results support this conclusion. For the NeuroDSA dataset, shown in Fig. 4b, clinicians judged the Clarity25 system significantly better (Q_IQ, blue points). The indices are in agreement, indicating that the IQ of the Clarity25 system is on average better.

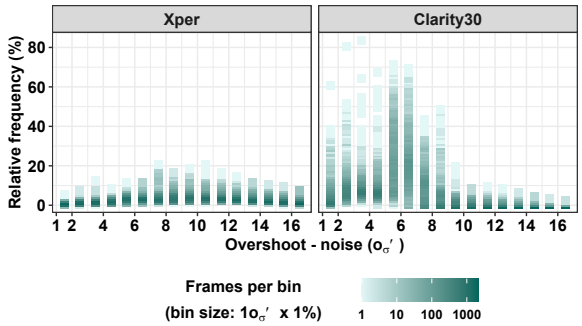
3.3 Overshoot distribution

A dose control system operating in clinical practice requires setting a stopping point, especially if combined with an approach that aims to reduce dose as low as practically possible. Although we could define $o' = 1$ plus an additional safety as a reasonable stopping point, the use of one global FOM for an entire frame will be inadequate. As shown for the PhantomCHO dataset in Section 3.1, Overshoot values for larger signals are up to a factor 1.6 greater than for smaller signals. Therefore, it would be imprudent to define one 'average' Overshoot target for clinical images comprised of an uncontrolled number and diameters of vessels being imaged. Dose reduction tuned to the average Overshoot could reduce the visibility of smaller vessels below the clinician's detection threshold. We investigate an alternative – the distribution of Overshoot values within a frame. The heatmaps in Fig. 5 represents superimposed histograms, generated from one histogram per frame for all sequences/datasets. Color-coding represents the number of frames falling into a bin. Additional explanation is given in Section 2.4.

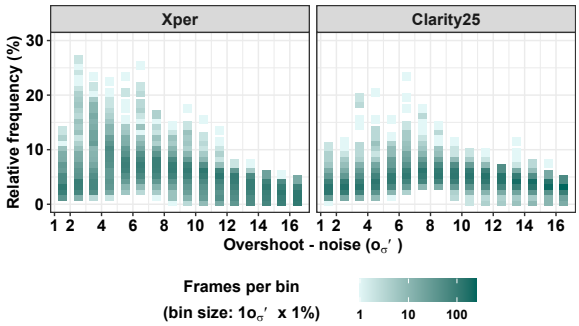
The trends are similar for all three datasets: sequences with higher IQ show a more uniform distribution across all Overshoot values. Conversely, sequences with lower IQ have a distribution that peaks at lower Overshoot



(a) PhantomCHO



(b) Cardiac



(c) NeuroDSA

Figure 5. Heatmap of Overshoot distributions, showing a shift towards lower overshoot values at lower image quality levels. Heatmaps are generated by calculating the Overshoot histogram within a frame and superimposing the distributions for all frames and sequences at that dose level; a detailed explanation is in Section 2.4. On (a) PhantomCHO: at 240 nGy/fr, values are distributed at higher quality levels and no frames contain values in bin $1 \leq o'_\sigma < 2$. With decreasing dose and IQ, the distribution is concentrated at smaller o'_σ values with greater frequency. By 6 nGy/fr, most o'_σ values are at the threshold of visibility. On the clinical datasets, the lower IQ sequences exhibit the same shift to lower Overshoot values: (b) Cardiac–Clarity30 and (c) NeuroDSA–Xper.

values. This trend is most easily seen on the PhantomCHO dataset, shown in Fig. 5a. At the highest dose level, 240 nGy/fr, the distribution is concentrated equally at higher Overshoot values. Moreover, no frames contain Overshoot values at the limit of visibility (bin $1 \leq o'_\sigma < 2$). Dose reduction shifts the distribution to the left, concentrating at smaller o'_σ values with greater frequency. By 6 nGy/fr, most o'_σ values are at the threshold of visibility for almost all frames. Fig. 5b shows the same trend for the Cardiac dataset, where the Clarity30 system was rated as lower IQ than Xper. The Clarity30 distribution peaks at lower Overshoot values, while the standard-dose Xper has a uniform distribution with relatively low Overshoot frequencies across frames. For the NeuroDSA dataset, the Xper system was rated as lower IQ then Clarity25, but the differences between the two systems are smaller than those for the Cardiac dataset. Therefore, the two distributions in Fig. 5c appear similar. Nevertheless, the Xper distribution is shifted left and peaks at lower Overshoot values.

To implement this in clinical practice, we would first define an Overshoot threshold; for example, $o'_\sigma = 2$, which includes a margin of safety above the threshold of visibility. The dose control system would then adjust the dose while tracking the frequency of Overshoot values falling below the threshold. When the number of pixels in this range would exceed a certain frequency, for example when more than 10% of pixels have $o'_\sigma < 2$, the system would cease dose reduction. Based on the heatmaps of the clinical datasets in Fig. 5, many of the frames in the low-IQ sequences are concentrated around 10% frequency for $1 \leq o'_\sigma < 2$. Additional experiments will help determine whether such an approach will match human judgments in detecting the lowest acceptable dose level, and whether the same frequency target would generalize across different image content.

4. DISCUSSION

This work presents an IQ index and an automated algorithm to measure patient-specific IQ for interventional X-ray imaging. The Overshoot Index reflects the 'excess' IQ in the image, beyond what is required to reliably visualize contrast-enhanced vessels – 'the clinical task'. The dose control system will reduce dose until the target Overshoot is reached. For example, at a target Overshoot of 1, signals will be at the threshold of detectability. A higher target Overshoot may be defined, as an added safety margin. We found a linear relationship between the Noise Overshoot index and the square root of the detector dose. Thus the the maximum allowable dose reduction can be calculated directly and is equivalent to the difference between the measured and target Overshoot values.

The dose control system will operate in two steps. The first step is identical to existing clinical systems. The system will acquire one frame at its reference dose level, for example based on the patient's weight. The proposed method will take over in the second step. The automated algorithm will measure image features in that frame and estimate the amount of Noise Overshoot. If Overshoot occurs, the equivalent reduction in tube current and pulse duration will be computed from the maximum allowable reduction in Overshoot. As long as the field of view remains unchanged, subsequent frames will be imaged at the lower dose level.

We hypothesized that by seeking the minimum dose that facilitates the clinical task, additional dose reduction beyond a system's default configuration could be achieved. Our study suggests this is feasible: results on the PhantomCHO dataset suggest a dose reduction of 15–58% depending on the the desired safety margin. The authors of Ref. 19, who have shared the PhantomCHO with us, state that fluoroscopy is routinely acquired at 18 nGy in their clinic. Noise overshoot at 18 nGy/fr is estimated as $o'_\sigma = 2.14$ for the 1 mm signal. If the dose control system is configured to target $o'_\sigma = 2$, the equivalent dose reduction would be approximately 15%, with a target DTD of 15.4 nGy/fr. For a more aggressive target of $o'_\sigma = 1.5$, dose savings could reach nearly 58% at DTD 7.6 nGy/fr. Either target translates into a significant reduction in dose exposure for both patients and staff in the operating theater.

The Overshoot index may also function as a suitable substitute for the well-known CHO in real-time patient-specific dose optimization, given several caveats. Our results show that Noise overshoot is linearly correlated with d' from a CHO optimized for interventional X-ray. Models such as the CHO mimic detection performance of humans using a signal detection theory (SDT) approach. The model estimates statistical differences between signal present (SP) and signal negative (SN) stimuli and computes the probability that a given stimulus belongs to one class. SDT explicitly incorporates the phenomenon that observers – models and humans – may incorrectly identify a SP stimulus as SN and vice versa, depending on the amplitude of the signal and image characteristics. Thus the detectability of a signal is given as a probability, not a binary visible/not-visible class. In our approach,

however, the segmentation algorithm assigns a binary class to the signal. The vessel detector currently works by classifying a candidate pixel as SP or SN based on a hard probability threshold. This threshold – i.e. the ROC operating point of the detector – was optimized for the PhantomChest ground truth segmentation, and does not guarantee the same operating point holds for other datasets. In future work, we will investigate an adaptive detector that adjusts the classification threshold based on a fixed ROC operating point. However, the visual model implicitly handles the phenomenon of decision uncertainty. Our model is derived from staircase experiments incorporating a SDT protocol. Threshold contrast is computed from the visual model where the probability of detection is greatest, i.e. the least decision uncertainty. Our approach does not require extensive training, unlike the CHO, but it is dependent on accurate measurement of the visual model across a range of conditions, such as different types of noise correlation – a current limitation of our method.

For the clinical datasets, the Overshoot indices display the same trends as clinicians' judgments. However, the two IQ ratings exhibit low correlation and the data are highly dispersed. This may be due to the experimental setup. Clinicians were asked to evaluate perceived IQ; the experiment did not involve a vessel detection task. Ref. 42 found a similar discrepancy in a digital pathology application. Six veterinary pathologists rated the perceived IQ of altered digital pathology slides under two conditions: while free viewing and while conducting a diagnostic task. The introduction of a clinical task led to different judgments of perceived IQ, possibly by inducing a shift in the participants' attention and perceived IQ criteria.

The proposed work brings several benefits over existing IQ measurement and dose optimization approaches. First, unlike the CNR metric, the Overshoot explicitly incorporates the effect of background intensity, which has a significant impact on threshold contrast. This effect is easily seen from the examples in Fig. 1a. Comparing the two backgrounds 14 versus 100 cd/m^2 with 3% noise standard deviation, reflecting realistic interventional X-ray values, the equivalent CNR at *threshold contrast* using $\mathrm{CNR} = c_T/\sigma_M$ is $\mathrm{CNR}_{14} = 2.9$ and $\mathrm{CNR}_{100} = 1.4$. Thus, while we would conclude that the first signal has twice the CNR of the second, in fact an observer's ability to detect the two signals on a detection task are equivalent: detection probability of 99.5%. From the perspective of performing the clinical task, which we defined as the ability to detect contrast-enhanced vessels, the two stimuli have equivalent IQ. Second, the formulation of the Overshoot index explicitly incorporates the lower bound of a clinical signal's visibility threshold. This allows the dose control system to target maximum dose reduction while ensuring the interventionalist's ability to carry out their clinical task is not affected. Third, the proposed method does not require prior training as with some MOs. Post-processing algorithms, such as denoising and contrast enhancement, distort spatio-temporal noise correlation as well as the intensity-noise model. A CHO would require continuous retraining of the image statistics as different post-processing parameters are explored, whereas our approach can be applied directly. Finally, the Overshoot index is independent of the underlying algorithm used to quantify image features.

The proposed approach contains several limitations, in addition to those already mentioned. The algorithm relies on accurate segmentation of contrast-enhanced vessels and cannot currently estimate the visibility of other important clinical objects such as guidewires. In the future, deep learning algorithms trained on angiography sequences could be investigated; several implementations from the literature are publicly available. The index could also estimate Overshoot from the virtual guidewire approach used by several manufacturers, explained in Section 1. Second, our visual model does not consider aspects such as the spatial extent of the signal, noise correlation, object motion, presentation frame rate, and display LCD response rate, among others. To improve the model's accuracy, we are exploring adapting existing visual models to this application. In addition, the algorithm could be adapted to estimate Overshoot with respect to perceived luminance at the eye, rather than display luminance. Finally, the relationship found between Noise Overshoot and dose has not been verified for other exposure parameters, such as tube voltage, filtration, and focal spot extent, nor for other vendors' systems. We consider this a proof-of-concept approach; despite the aforementioned approximations in the model, the method performs well on the tested datasets.

5. CONCLUSIONS

We propose a novel index and a fully automated algorithm to measure patient-specific image quality (IQ) in interventional X-ray images, with the goal of improving dose reduction for a given clinical task. We define the task as detection of contrast-enhanced vessels. The method determines the maximum increase in image noise

or reduction in signal contrast that does not affect task performance, which we call the Noise and Contrast Overshoot indices, respectively. We refer to this class of indices as the Visibility Overshoot.

The Noise Overshoot index is found to have high correlation with d' from a Channelized Hotelling Observer on a phantom dataset. Both indices show good agreement with clinicians' subjective judgments for clinical datasets. When either index is coupled to acquisition parameters, the dose control system will deliver a lower dose to the patient, proportional to the amount of Overshoot in a given acquisition frame. For the phantom dataset, the Noise Overshoot index varies linearly with the square root of detector dose, where dose is predominantly a function of tube current and pulse width. The index predicts a dose reduction of 15–58% may be feasible in fluoroscopy mode, depending on the desired safety margin. The proposed approach shows promise for greater dose reduction compared to existing dose optimization methods. This reduction will benefit patients undergoing interventional X-ray procedures and staff in the operating theater.

ACKNOWLEDGMENTS

This research was funded by the ECSEL PANORAMA project (co-funded by grants from Belgium, Italy, France, the Netherlands, the United Kingdom, and the ENIAC Joint Undertaking). Phantom data and CHO scores kindly provided by Kenneth Fetterly, Mayo Clinic, Rochester, Minnesota. NeuroDSA dataset kindly provided by Philips Healthcare, the Netherlands. We thank Robert Hofsink from Philips Healthcare and Dr. Yves Taeymans from Ghent University Hospital for their assistance with this work.

REFERENCES

[1] Söderman, M., Holmin, S., Andersson, T., Palmgren, C., Babić, D., and Hoornaert, B., "Image noise reduction algorithm for digital subtraction angiography: Clinical results," *Radiology* **121262** (June 2013).

[2] Eloot, L., Thierens, H., Taeymans, Y., Drieghe, B., De Pooter, J., Van Peteghem, S., Buytaert, D., Gijs, T., Lapere, R., and Bacher, K., "Novel X-ray imaging technology enables significant patient dose reduction in interventional cardiology while maintaining diagnostic image quality," *Catheterization and Cardiovascular Interventions* **86**(5), E205–E212 (2015).

[3] Osei, B., Xu, L., Johnston, A., Darko, S., Darko, J., and Osei, E., "Retrospective study of patients radiation dose during cardiac catheterization procedures," *The British Journal of Radiology* **92**, 20181021 (July 2019).

[4] Kastrati, M., Langenbrink, L., Piatkowski, M., Michaelsen, J., Reimann, D., and Hoffmann, R., "Reducing radiation dose in coronary angiography and angioplasty using image noise reduction technology," *The American Journal of Cardiology* **118**, 353–356 (Aug. 2016).

[5] Gislason-Lee, A. J., Keeble, C., Egleston, D., Bexon, J., Kengyelics, S. M., and Davies, A. G., "Comprehensive assessment of patient image quality and radiation dose in latest generation cardiac x-ray equipment for percutaneous coronary interventions," *Journal of Medical Imaging* **4**, 025501 (May 2017).

[6] Pizano, A., Khurram, A., Chamseddin, K., Timaran, C. H., Baig, S., Shih, M., Xi, Y., Guild, J., and Kirkwood, M. L., "New imaging technology system reduces patient radiation dose during peripheral arterial endovascular interventions," *Journal of Vascular Surgery* **76**, 500–504 (Aug. 2022).

[7] Kaatsch, H. L., Schneider, J., Brockmann, C., Brockmann, M. A., Overhoff, D., Becker, B. V., and Waldeck, S., "Radiation exposure during angiographic interventions in interventional radiology – risk and fate of advanced procedures," *International Journal of Radiation Biology* **98**, 865–872 (Jan. 2022).

[8] International Commission on Radiological Protection (ICRP), "Recommendations of the ICRP. ICRP Publication 26," *Annals of the ICRP* (1977).

[9] Justino, H., "The ALARA concept in pediatric cardiac catheterization: techniques and tactics for managing radiation dose," *Pediatric Radiology* **36** (July 2006).

[10] Barrett, H. H. and Myers, K. J., [*Foundations of Image Science*], John Wiley & Sons, Inc., Hoboken, NJ (2013).

[11] Aufrichtig, R., Thomas, C. W., Xue, P., and Wilson, D. L., "Model for perception of pulsed fluoroscopy image sequences," *Journal of the Optical Society of America A* **11**, 3167–3176 (Dec. 1994).

[12] Zhang, S., Abbey, C. K., Teymoorian, A., Da, X., Whiting, J. S., and Eckstein, M. P., "Model observers for complex discrimination tasks: assessments of multiple coronary stent placements," *Proc. SPIE* **7627**, 76270W–76270W–10 (2010).

[13] Tao, A. and Fetterly, K., "Integration of high velocity test object motion into a channelized hotelling observer for the assessment of x-ray angiography systems," *Physics in Medicine & Biology* **64**, 185011 (Sept. 2019).

[14] Jabri, K. N. and Wilson, D. L., "Quantitative assessment of image quality enhancement due to unsharp-mask processing in x-ray fluoroscopy," *Journal of the Optical Society of America A, Optics and Image Science* **19**, 1297–1307 (July 2002).

[15] Wilson, D., Jabri, K., and Aufrichtig, R., "Perception of temporally filtered X-ray fluoroscopy images," *IEEE Transactions on Medical Imaging* **18**(1), 22–31 (1999).

[16] Ortenzia, O., Trojani, V., Bertolini, M., Nitrosi, A., Iori, M., and Ghetti, C., "Radiation dose reduction and static image quality assessment using a channelized hotelling observer on an angiography system upgraded with clarity IQ," *Biomedical Physics & Engineering Express* **6**, 025008 (Feb. 2020).

[17] Fetterly, K. A., "Performance assessment of active vs passive pixel x-ray angiography detector systems using a bias-corrected channelized hotelling observer and adult patient-equivalent experimental conditions," *Medical Physics* **45**, 4888–4896 (Oct. 2018).

[18] Tapiovaara, M. J., "SNR and noise measurements for medical imaging. II. application to fluoroscopic x-ray equipment," *Physics in Medicine and Biology* **38**, 1761–1788 (Dec. 1993).

[19] Fetterly, K. A. and Favazza, C. P., "Direct estimation and correction of bias from temporally variable non-stationary noise in a channelized hotelling model observer," *Physics in Medicine and Biology* **61**, 5606–5620 (July 2016).

[20] Villa, R., Paruccini, N., Baglivi, A., Signoriello, M., Velasquez, R. A. M., Morzenti, S., Ponti, E. D., and Crespi, A., "Model observers for low contrast detectability evaluation in dynamic angiography: A feasible approach," *Physica Medica* **64**, 89–97 (aug 2019).

[21] Desponds, L., "AutoRight(TM) - Intelligence Inside," Tech. Rep. white paper JB70418XX, GE Healthcare (2019).

[22] Tashenov, S., "OPTIQ – A new approach to image quality and dose in minimally invasive procedures," Tech. Rep. white paper Version 1.0 VE20 online pdf 7666 0620, Siemens Healthcare GmbH (2020).

[23] Dehairs, M., Bosmans, H., and Marshall, N. W., "Implementation of a spatio-temporal figure of merit for new automatic dose rate control regimes in dynamic x-ray imaging," *Physics in Medicine & Biology* **64**, 045001 (Feb. 2019).

[24] Smith, T. B., Solomon, J., and Samei, E., "Estimating detectability index in vivo: development and validation of an automated methodology," *Journal of Medical Imaging* **5**, 031403 (Dec. 2017).

[25] Smith, T. B., Abadi, E., Solomon, J., and Samei, E., "Development, validation, and relevance of in vivo low-contrast task transfer function to estimate detectability in clinical CT images," *Medical Physics* **48**, 7698–7711 (Nov. 2021).

[26] Gislason, A. J., Hoornaert, B., Davies, A. G., and Cowen, A. R., "Allura Xper cardiac system implementation of automatic dose rate control," Tech. Rep. 4522 962 71201, Philips (Aug. 2011).

[27] Kingdom, F. and Prins, N., [*Psychophysics: A Practical Introduction*], Academic Press, London NW1 7BY, UK (Nov. 2009).

[28] Kaernbach, C., "A single-interval adjustment-matrix (SIAM) procedure for unbiased adaptive testing," *J. Acoust. Soc. Am.* **88**, 2645–2655 (Dec. 1990).

[29] Kumcu, A., Platiša, L., and Philips, W., "Effects of static and dynamic image noise and background luminance on letter contrast threshold," in [*7th International Workshop on Quality of Multimedia Experience (QoMEX 2015)*], 1–2 (26–29 May 2015).

[30] Fründ, I., Haenel, N. V., and Wichmann, F. A., "Inference for psychometric functions in the presence of nonstationary behavior," *Journal of Vision* **11**(6), 16 (2011).

[31] Fetterly, K. A., Blume, H. R., Flynn, M. J., and Samei, E., "Introduction to grayscale calibration and related aspects of medical imaging grade liquid crystal displays," *Journal of Digital Imaging* **21**(2), 193–207 (2008).

[32] National Electrical Manufacturers Association (NEMA), "Digital imaging and communications in medicine (DICOM), part 14: Grayscale standard display function," (2015).

[33] Zlokolica, V., Pižurica, A., and Philips, W., "Noise estimation for video processing based on spatial-temporal gradient histograms," *IEEE Signal Processing Letters* **13**, 337–340 (June 2006).

[34] Bracho, R. and Sanderson, A. C., "Segmentation of images based on intensity gradient information," in [*Proceedings of CVPR-85 Conference on Computer Vision and Pattern Recognition, San Francisco*], 341–347 (1985).

[35] Donoho, D. L. and Johnstone, I. M., "Ideal spatial adaptation by wavelet shrinkage," *Biometrika* **81**, 425–455 (Sept. 1994).

[36] Cesarelli, M., Bifulco, P., Cerciello, T., Romano, M., and Paura, L., "X-ray fluoroscopy noise modeling for filter design," *International Journal of Computer Assisted Radiology and Surgery* **8**(2), 269–278 (2013).

[37] Goossens, B., Pizurica, A., and Philips, W., "EM-based estimation of spatially variant correlated image noise," in [*2008 15th IEEE International Conference on Image Processing*], 1744–1747, IEEE (2008).

[38] Goossens, B., Pižurica, A., and Philips, W., "Removal of correlated noise by modeling the signal of interest in the wavelet domain," *IEEE Transactions on Image Processing* **18**, 1153–1165 (June 2009).

[39] Goossens, B., Aelterman, J., Luong, H., Pižurica, A., and Philips, W., "Efficient design of a low redundant discrete shearlet transform," in [*Local and Non-Local Approximation in Image Processing, 2009. LNLA 2009. International Workshop on*], 112–124 (Aug 2009).

[40] Kumcu, A., Platiša, L., Chen, H., Gislason-Lee, A. J., Davies, A. G., Schelkens, P., Taeymans, Y., and Philips, W., "Selecting stimuli parameters for video quality assessment studies based on perceptual similarity distances," in [*Proc. SPIE, Image Processing: Algorithms and Systems XIII*], **9399**, 93990F–1–93990F–10 (Feb. 2015).

[41] Dodge, J. T., Brown, B. G., Bolson, E. L., and Dodge, H. T., "Lumen diameter of normal human coronary arteries. influence of age, sex, anatomic variation, and left ventricular hypertrophy or dilation.," *Circulation* **86**, 232–246 (July 1992).

[42] Platiša, L., Van Brantegem, L., Kumcu, A., Ducatelle, R., and Philips, W., "Influence of study design on digital pathology image quality evaluation: the need to define a clinical task," *Journal of Medical Imaging* **4**(2), 021108–1–021108–12 (2017).