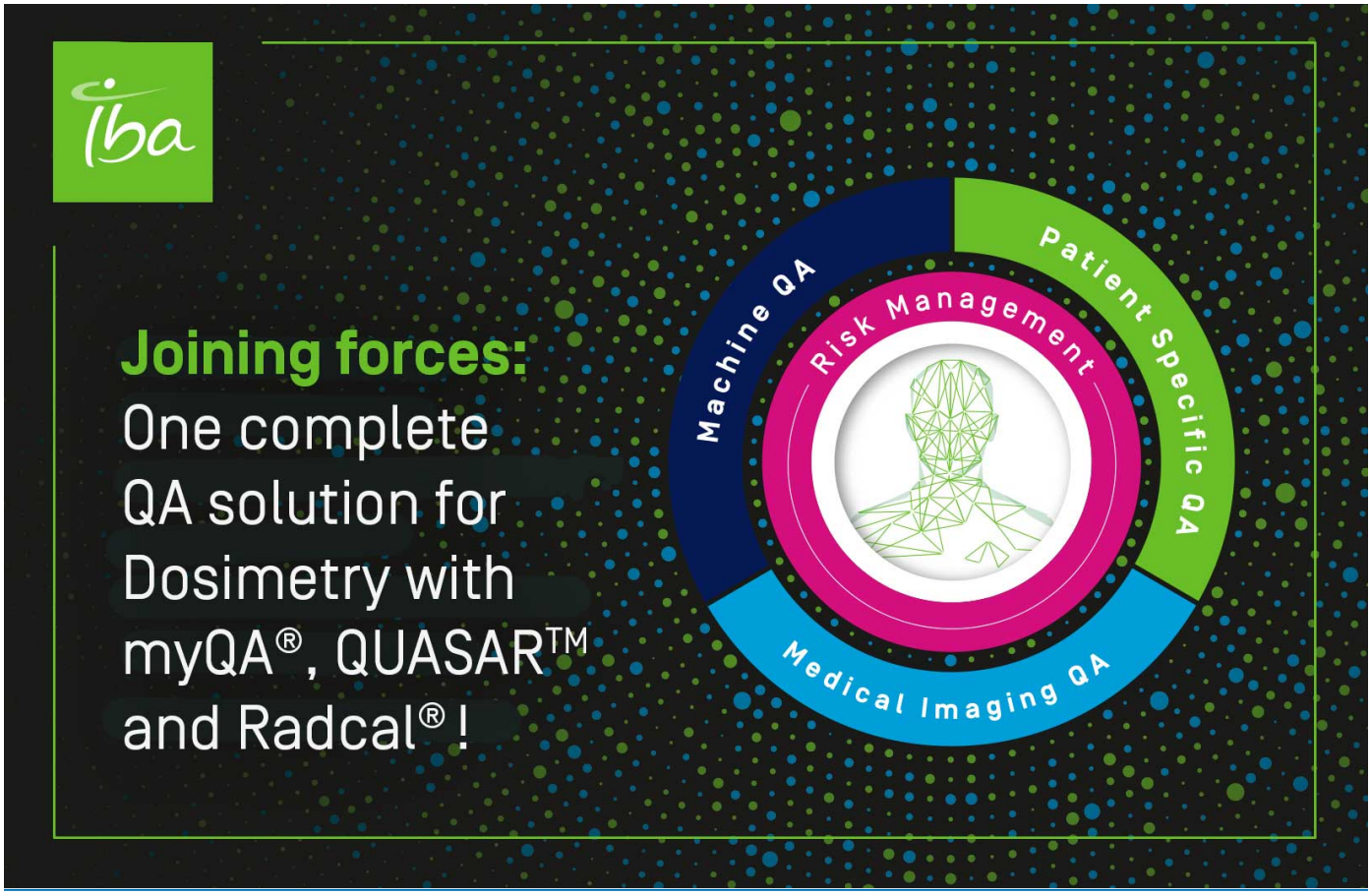
Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

PAPER • OPEN ACCESS

A nonparametric measure of contrast in x-ray images

To cite this article: M Anton *et al* 2024 *Phys. Med. Biol.* **69** 155013



IOP Publishing Phys. Med. Biol. 69 (2024) 155013 https://doi.org/10.1088/1361-6560/ad6119

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

PAPER

A nonparametric measure of contrast in x-ray images

M Anton[1,*], M Reginatto[1], S Schopphoven[2], C Abou Jaoude[2], U Mäder[3], M Fiebich[3], F Mauter[1,4], I Sechopoulos[4,5] and R van Engen[5]

[1] Physikalisch-Technische Bundesanstalt, Braunschweig and Berlin, Germany
[2] Reference Centre for Mammography Screening Southwest Germany, Giessen, Germany
[3] Institute of Medical Physics and Radiation Protection, University of Applied Sciences, Giessen, Germany
[4] Radboud University Medical Center, Nijmegen, The Netherlands
[5] LRCB Dutch Expert Centre for Screening, Nijmegen, The Netherlands
[*] Author to whom any correspondence should be addressed.

E-mail: mathias.anton@ptb.de

Keywords: quality assurance, contrast, grey value distance, nonparametric statistics, x-ray diagnostics

OPEN ACCESS

RECEIVED
22 March 2024

REVISED
18 June 2024

ACCEPTED FOR PUBLICATION
9 July 2024

PUBLISHED
19 July 2024

Abstract

Objective. We propose a nonparametric figure of merit, the contrast equivalent distance CED, to measure contrast directly from clinical images. *Approach.* A relative brightness distance δ is calculated by making use of the order statistic of the pixel values. By multiplying δ with the grey value range R, the mean brightness distance MBD is obtained. From the MBD, the CED and the distance-to-noise ratio DNR can be derived. The latter is the ratio of the MBD and a previously suggested nonparametric measure τ for the noise. Since the order statistic is independent of the spatial arrangement of the pixel values, the measures can be obtained directly from clinical images. We apply the new measures to mammography images of an anthropomorphic phantom and of a phantom with a step wedge as well as to CT images of a head phantom. *Main results.* For low-noise images of a step wedge, the MBD is equivalent to the conventional grey value distance. While this measure permits the evaluation of clinical images, it is sensitive to noise. Therefore, noise has to be quantified at the same time. When the ratio σ/τ of the noise standard deviation σ to τ is available, validity limits for the CED as a measure of contrast can be established. The new figures of merit can be calculated for entire images as well as on regions of interest (ROI) with an edge length not smaller than 32 px. *Significance.* The new figures of merit are suited to quantify the quality of clinical images without relying on the assumption of a linear, shift-invariant system. They can be used for any kind of greyscale image, provided the ratio σ/τ can be estimated. This will hopefully help to achieve the optimisation of image quality vs dose required by radioprotection laws.

1. Introduction

For medical diagnostics involving ionising radiation, a higher radiation dose usually leads to lower noise and hence to a better image. However, this has to be balanced with the health risk for the patient, which also increases with increasing dose. Therefore, radioprotection laws require an optimisation of image quality vs radiation dose according to the principle that 'doses should [...] be kept as low as reasonably achievable, taking into account economic and societal factors.' (ICRP 2007). While computed tomography (CT) provides the largest contribution to the collective effective dose from all x-ray examinations (UNSCEAR 2022), mammography screenings deserve special attention due to the fact that a mainly healthy population is exposed.

Image quality is enhanced by modern image reconstruction in CT and modern image processing in mammography. At the same time, unfortunately, techniques like iterative reconstruction, artificial intelligence-based (AI) approaches (e.g. for CT) or contrast enhancement by local histogram equalisation (e.g. for mammography) render established image quality assessment methods inadequate: For example, the 3D image reconstruction and image processing is non-linear. AI-based image processing and -reconstruction as well as local contrast enhancement are not only non-linear but can be dependent on local structures at a

specific location in the image. Therefore, the requirement in traditional image quality evaluations that the system is linear and shift-invariant is not met. In the case of mammography, an additional complication is that technical phantoms, which are often used for quality assessment, do not lead to the same image processing as a real breast or an anthropomorphic phantom would.

Model observers have been widely researched and are used to overcome these problems (Myers and Barrett 1987, Barrett and Myers 2003, ICRU 2006, Monnin *et al* 2011, Vaishnav *et al* 2014, Balta *et al* 2019, Samei *et al* 2019, Anton *et al* 2021). However, this approach is hampered by a huge workload. In addition, to use model observers with modern image reconstruction and image processing schemes, anthropomorphic phantoms would be required that are often not readily available. To compound this problem, currently the relation between a model observer and human observers has to be established case by case, making the implementation of model observers in quality assurance protocols and guidelines very difficult.

Therefore, it appears highly desirable to derive a set of parameters suitable for the quantification of image quality directly from the clinical image. For natural images, there is an abundance of publications dealing with various approaches for no-reference image quality assessment[6] (see, for example, Ferzli and Karam 2009, Shen *et al* 2009, Zhu and Milanfar 2010, Moorthy and Bovik 2011, Mittal *et al* 2012, Vu and Chandler 2012, Liu *et al* 2014, Ma *et al* 2017, Panetta *et al* 2018). A review of comparable methods for medical images can be found in a publication by Chow and Paramesran (2016). For x-ray images, however, not many attempts to derive figures of merit (FOM) for the image quality directly from clinical images have been published yet (Christianson *et al* 2015, Sanders *et al* 2016, Abadi *et al* 2017, Smith *et al* 2017).

What is required (at a minimum) is a set of figures of merit that quantify noise, noise texture, contrast, and spatial resolution or sharpness of the image under investigation. As a first step, the authors have published a suggestion for a new, nonparametric measure τ for noise that is derivable from mammography images (Anton *et al* 2023) and that is based on work by Bielecka *et al* (2020) and by Obuchowicz *et al* (2020).

In the present work, we suggest the contrast equivalent distance CED as a measure of contrast which can—in principle—be derived from any digital greyscale image, not only from x-ray images. Different from suggestions by other authors (Peli 1997, Rizzi *et al* 2004, Simone *et al* 2012, Lecca *et al* 2021), we do not attempt to produce a measure of perceived contrast but restrict the investigations to an objective measure that can, at a later stage, be related to perceived contrast.

Just like the measure of noise, τ, the new measure of contrast is based on nonparametric statistics, or, more precisely, on the order statistics of the pixel values. In section 2, we define the relative brightness distance δ and the mean brightness distance MBD as auxiliary quantities. From those, the contrast equivalent distance CED and the distance-to-noise ratio DNR can be derived. The latter is obtained with the help of τ and is also applicable to any greyscale image. We then investigate limitations of the approach by using simulated images of a step function with different noise levels and different noise textures (section 3). Finally, we describe how CED and DNR can be determined locally on regions of interest (ROI) of mammographic images to quantify the contrast enhancement achieved by the manufacturer's image processing (section 4). The DNR is required to establish the regions of validity of the measure MBD and hence of CED.

In appendix A, we show how the new measures relate to established methods, using mammograms involving a grey value step wedge according to the German standard DIN 6868-14 (NAR 2020) as an example. An application of the proposed methods to CT images is demonstrated in appendix B. Finally, appendix C shows how the relative mutual information obtained from co-occurrence matrices (Haralick *et al* 1973, Nowosad and Stepinski 2019) might help to interpret the new Figures of merit.

The proposed measures CED and DNR will be part of a larger, yet incomplete set of parameters for the quality assessment of x-ray diagnostic images. These parameters will have to be combined in a suitable way and then be related to the quality assessment by experts. Since the parameters presented up to now can be calculated quickly, they could be produced immediately after acquisition. In this way, in the future, reference values specific to certain diagnostic tasks could be established, similar to the diagnostic reference values for the CT dose (compare Ria *et al* 2019). This could then be the base for the required image quality vs dose optimisation.

2. The new quantities: CED and DNR

Usually, contrast is defined as a brightness or luminance difference divided by a mean value. One of the basic quantities is the grey value distance[7], defined as the difference of the mean grey values $m_1 - m_2$ of two

[6] No-reference image quality assessment (NR-IQA) methods aim to provide a measure for the image quality without referring to a reference. The opposite would be FR-IQA, full reference image quality assessment: here, an 'ideal' image is available to which the image under investigation is compared. The latter is done, e.g. when image compression methods are evaluated.

[7] See, for example, the German standard DIN 6868-14 (NAR 2020) for the quality assurance of mammography systems.

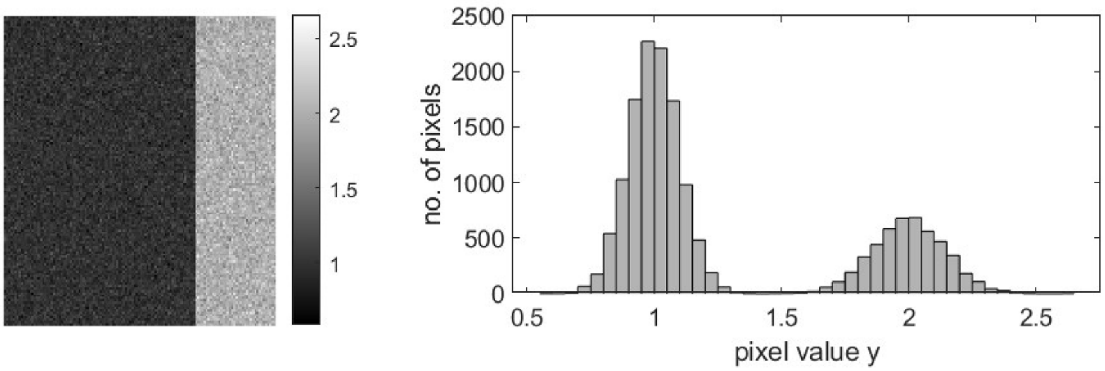


Figure 1. Square image ($\nu = 128$) with an asymmetric step from $y = 1$ to $y = 2$, with white noise. Left: greyscale image; Right: histogram.

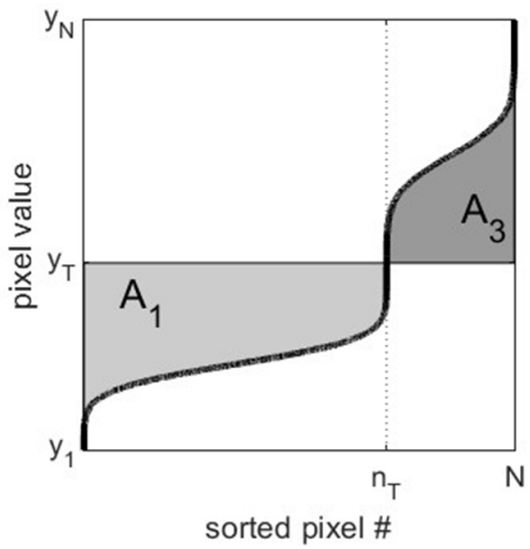


Figure 2. Illustration of the definition of δ: the pixel values $y_i, i = 1 \ldots N$ are sorted in ascending order and then plotted. The threshold y_T is reached at pixel number n_T. In this example, $y_1 = 0.59$ and $y_N = 2.7$, $N = 16384$, $y_T = 1.5$ at $n_T = 11521$, resulting in a relative brightness distance $\delta = 0.48$.

distinct, preselected and homogeneous spatial regions 1 and 2 of an image ($m_1 > m_2$). The Michelson contrast C_M, e.g. is defined in terms of the grey value distance as $C_M = (m_1 - m_2)/(m_1 + m_2)$.

Since we aim at a measure that can be obtained from patient images, we need to define similar quantities which make no use of predefined regions of 'high' or 'low' grey values. However, it is desirable that—at least in limiting cases—the new measures for the grey value distance and the contrast will yield the same values as the established ones, for example in the case of a high signal-difference-to-noise ratio (SDNR).

We will first introduce the relative brightness distance δ and the mean brightness distance MBD as auxiliary quantities. The definitions will be illustrated by a few instructive examples. From the MBD, the contrast equivalent distance CED and the distance-to-noise ratio DNR are derived.

2.1. Definition of the auxiliary quantities δ and MBD

The definition of the relative brightness distance δ is easily understood by using a simple illustrative examples Assume a square image of edge length ν, totalling $N = \nu^2$ pixels with pixel values $y_i,\ i = 1 \ldots N$. An example with a step function from $y = 1$ to $y = 2$, $\nu = 128$ with additional white noise is shown in figure 1, along with its histogram. In figure 2, the corresponding pixel values are sorted in ascending order and then plotted (in

nonparametric statistics, this is called the 'order statistics' of the pixel values, see, e.g. Hoaglin *et al* 1983). The graph is subdivided into four quadrants by the threshold value y_T and its abscissa n_T.

A_1 and A_3 correspond to the shaded regions indicated in figure 2. Their sum is the area between the order statistic and the horizontal line defined by the threshold value y_T. With the range $R = y_N - y_1 = \max(y) - \min(y)$, the total area of the graph is given by $A_{tot} = N \cdot R$. An integral quantity δ can then be defined as[8]

$$\delta = \frac{A_1 + A_3}{\frac{1}{2}A_{tot}}, \tag{1}$$

where A_1 and A_3 are given by

$$A_1 = \sum_{i=1}^{n_T} (y_T - y_i) \quad \text{and} \quad A_3 = \sum_{i=n_T+1}^{N} (y_i - y_T).$$

We obtain

$$\delta = \frac{\sum_{i=n_T+1}^{N} y_i - \sum_{i=1}^{n_T} y_i + (2 \cdot n_T - N) \cdot y_T}{\frac{N}{2} \cdot R}. \tag{2}$$

The resulting value of δ depends on the choice of the threshold y_T, i.e. $\delta = \delta(y_T)$.

If we use the abbreviations μ_H and μ_L for the mean values of y_i above and below the threshold y_T, respectively,

$$\mu_H = \frac{1}{N - n_T} \cdot \sum_{i=n_T+1}^{N} y_i \quad \text{and} \quad \mu_L = \frac{1}{n_T} \cdot \sum_{i=1}^{n_T} y_i, \tag{3}$$

we obtain

$$\delta(y_T) = \frac{2}{R} \cdot \left[\left(1 - \frac{n_T}{N}\right) \cdot \mu_H - \frac{n_T}{N} \cdot \mu_L - \left(1 - \frac{2 \cdot n_T}{N}\right) \cdot y_T \right]. \tag{4}$$

The expression becomes particularly simple in case N is even and $n_T = N/2$: the right hand side of equation (4) then equals $(\mu_H - \mu_L)/R$.

δ quantifies the distribution of the grey values within the range R. Multiplication of δ with the range R yields the mean brightness distance MBD:

$$\text{MBD} = \delta \cdot R \quad \text{where} \quad R = \max(y) - \min(y). \tag{5}$$

In the case where is little overlap between the grey value distributions above and below the threshold, respectively, the resulting value of the MBD converges to the conventional grey value distance (see appendix A).

We use Otsu's method (Otsu 1979) to determine the threshold value y_T from the grey value histograms. The method chooses a threshold such that the variance between the two classes defined in this way ('above' and 'below' the threshold) is maximum whereas the intra-class variances are minimal[9].

Some care has to be taken concerning the abscissa n_T: When none of the y_i-values coincides with the threshold y_T, n_T is defined as

$$n_T = \max(\{i | y_i < y_T\}).$$

If the number of y_i values coinciding with y_T is uneven, i.e. if the cardinality of the set $\{i | y_i = y_T\}$ is uneven

$$\text{card}(\{i | y_i = y_T\}) = 2 \cdot m + 1, \quad m \in \mathbb{N}_0,$$

n_T is chosen as the $(m+1)$th element of $\{i | y_i = y_T\}$. For an even number of y_i values coinciding with y_T, i.e. for

[8] The normalisation to $\frac{1}{2}A_{tot}$ was chosen in order to achieve that the value of the mean brightness distance MBD $= \delta \cdot R$ for a grey value step will coincide with the conventional value, see appendix A.

[9] In case more than two well-distinguishable grey levels are expected (i.e. three or more distinct peaks in the histogram), other methods may be used. One possibility is to use Otsu's method for multiple thresholds, another one is the use of a peakfinder algorithm and use the midpoint between peaks as a threshold. See, e.g. appendices A and B.

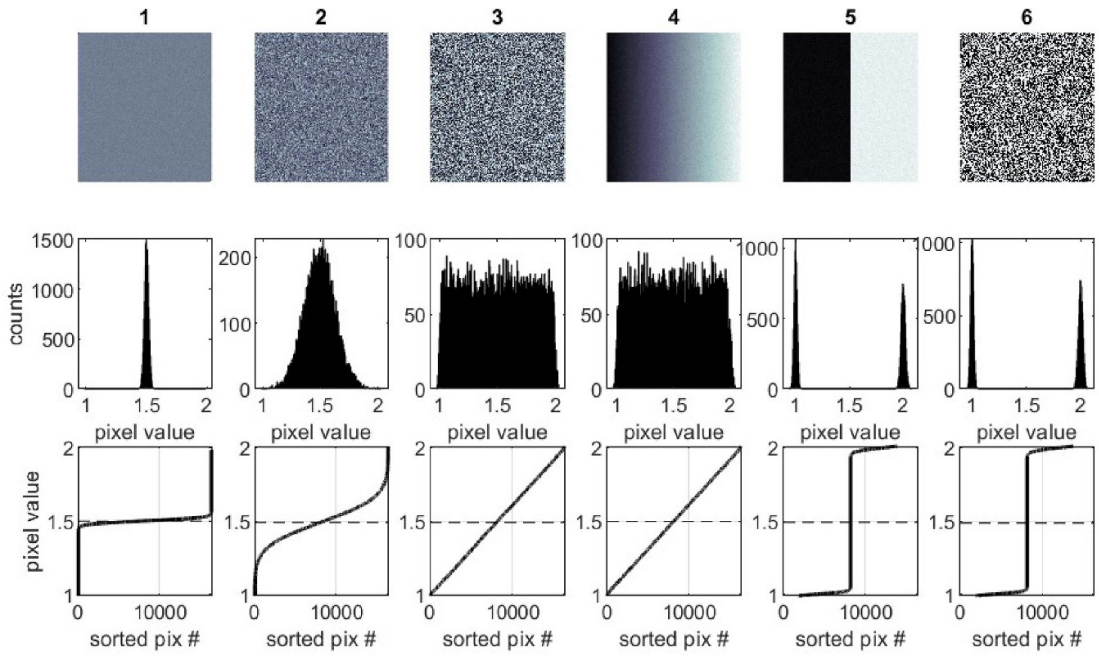


Figure 3. Images with an edge length of $\nu = 128$ px (top row), the corresponding histograms (middle row), i.e. counts as a function of pixel value, and the plots of the sorted pixel values (order statistics, bottom row). In the bottom row, the dashed line represents the threshold. Panel 1: constant grey value with low noise and just two outliers; Panel 2: white noise; Panel 3: uniformly distributed noise; Panel 4: a grey value ramp; Panel 5: a step; Panel 6: the same pixel values as in Panel 5, but distributed randomly. Note that $R \approx 1$ in all cases. The corresponding parameters δ and MBD are listed in table 1.

Table 1. Parameters describing the contrast in example images 1-6 (columns) of figure 3. The rows list the relative brightness distance δ and the mean brightness distance, MBD$= \delta \cdot R$.

	1	2	3	4	5	6
δ	0.028	0.206	0.477	0.466	0.885	0.890
MBD	0.028	0.215	0.501	0.501	1.000	1.000

$$\mathrm{card}\left(\{i|y_i = y_\mathrm{T}\}\right) = 2 \cdot m, \quad m \in \mathbb{N},$$

n_T is chosen as the mth element of $\{i|y_i = y_\mathrm{T}\}$.

2.2. Illustrative examples

By definition, δ can—in principle—assume values between zero and two. However, for well-behaved cases (grey values distributed smoothly), δ varies between zero and unity.

Some simple examples are displayed in figure 3: the top row shows images with ν =128 px, the second row shows the corresponding histograms (counts as a function of pixel value) and the third row shows the order statistics, where the chosen threshold is indicated by the dashed line. The columns are labelled by the numbers 1… 6. In panels 1, 4, 5 and 6 a small amount of white noise has been added to make the example more realistic. Panels 2 and 3 contain nothing but noise. Resulting parameters δ and MBD are listed in table 1.

The comparisons between panels 3 and 4 and between panels 5 and 6 deserve special attention: although the histograms and hence the δ-values are very similar, the appearances of the images are significantly different. This indicates that an additional parameter that quantifies the amount of structure in the image would be helpful to judge the meaningfulness of the MBD: The 'structuredness', e.g. can be measured by the relative mutual information (Nowosad and Stepinski 2019) as is outlined in appendix C.

2.3. The quantities CED and DNR

With the help of the MBD and the mean values μ_L and μ_H from equation (3), the contrast equivalent distance, CED, can be defined as

$$\mathrm{CED} = \frac{\mathrm{MBD}}{\mu_L + \mu_H}. \tag{6}$$

The crucial step is the calculation of the MBD including μ_L and μ_H, as explained in section 2.1. By normalising MBD to the sum of the mean grey values, it is guaranteed that the CED thus obtained can be used to compare the contrast in images, independent of their overall grey value range. Under low-noise conditions, the CED converges to the Michelson contrast C_{M}, as is demonstrated in section 3.

In a previous publication (Anton *et al* 2023), we have introduced a parameter τ to quantify noise. τ is proportional to the standard deviation σ of the noise and can be obtained directly from diagnostic images. By using this quantity and the MBD as defined above, we can define a distance-to-noise ratio DNR as

$$\mathrm{DNR} = \frac{\delta \cdot R}{\tau} = \frac{\mathrm{MBD}}{\tau}. \tag{7}$$

DNR can be viewed as a generalisation of the signal-difference-to noise ratio (SDNR): a higher DNR usually means a better image (see section 3).

We shall use the new parameters in several examples in the following sections and in the appendix: We will first investigate the limitations of the approach in the presence of noise, the influence of noise covariance and also the influence of the size of the region of interest (ROI) chosen to determine the parameters with the help of simulated data (section 3). In section 4, the new figures of merit are applied to a mammogram of an anthropomorphic phantom. A comparison of the MBD to the conventional grey value distance is described in appendix A using a mammogram of a quality assurance phantom. A CT example is investigated in appendix B.

2.4. Implementation

The calculation of the quantities defined above has been implemented in Matlab® (The MathWorks, Massachusetts, USA). We plan to make the software publicly available via the Matlab® Central File exchange.

3. Contrast of a step function—the influence of noise, noise covariance and ROI size

In the case of separate, well-defined regions in image space with different grey values, their means and standard deviations can be easily calculated which yields the conventional grey value distance $m_1 - m_2$ and the corresponding signal-difference-to noise ratio SDNR. When calculating δ, however, the classification is not achieved by associating grey values to spatial regions but by thresholding in grey level space. It turns out that the MBD and hence CED will be corrupted as a measure of contrast when noise becomes dominant[10]. Therefore, the limits of validity have to be determined by quantifying the noise at the same time. Since we are aiming at a measure which can also be applied locally, the size of the ROI chosen for the evaluation is considered as well.

We evaluated δ, MBD, CED, and DNR for a series of simulated images exhibiting a symmetric grey value step, adding uncorrelated (white) as well as correlated noise. To model the covariance, we assumed an exponential decay with the px-px distance as in a previous publication (please consult section 3.3 in Anton *et al* 2023), where such a dependence has been shown to describe the noise in mammograms rather well. The half-width $x_{1/2}$ of the exponential was varied between 0 and 5 px. Square ROIs with an edge length of 32 px and 128 px were compared. The standard deviation σ of the noise was assumed to be proportional to the square root of the grey values, i.e. $\sigma_i = f \cdot \sqrt{m_i}$ $(i = 1, 2)$ for the two regions, where f is a proportionality factor that was varied between ≈ 0.05 and 0.28. For $m_1 = 6$, $m_2 = 5$, this results in a range of values of the (nominal) signal-difference to noise ratio SDNR between 1.5 and 8, where

$$\mathrm{SDNR} = \frac{m_1 - m_2}{\sqrt{0.5 \cdot \sigma_1^2 + 0.5 \cdot \sigma_2^2}}.$$

Samples of the simulated images are shown in figure 4. The corresponding Michelson contrast (without considering noise) is $C_{\mathrm{M}} = 1/11 \approx 0.091$. The quantities δ, R, MBD, CED and DNR were calculated according to equations (2) and (5)–(7). The threshold y_{T} was obtained using Otsu's method (Otsu 1979). The noise parameter τ was calculated as described by Anton *et al* (2023) for eight nearest neighbours.

Some representative results are shown in figure 5. In the top row, MBD $= \delta \cdot R$ is shown; in the middle row, the CED is displayed; in the bottom row, the resulting DNR is seen, all as a function of the simulated SDNR. The three columns correspond to different values of the covariance half-width $x_{1/2}$. The left column represents the results for uncorrelated (white) noise, i.e. for $x_{1/2} = 0$. The middle column exhibits the results for $x_{1/2} = 0.3$ px, which is a representative value for noise in mammography images. The right column shows the data for the case of $x_{1/2} = 3$ px. Filled and open triangles represent the results for an ROI edge length of

[10] In the limit of pure normally distributed noise with a standard deviation σ, it can be shown that MBD$= \sqrt{\pi/8} \cdot \sigma$.

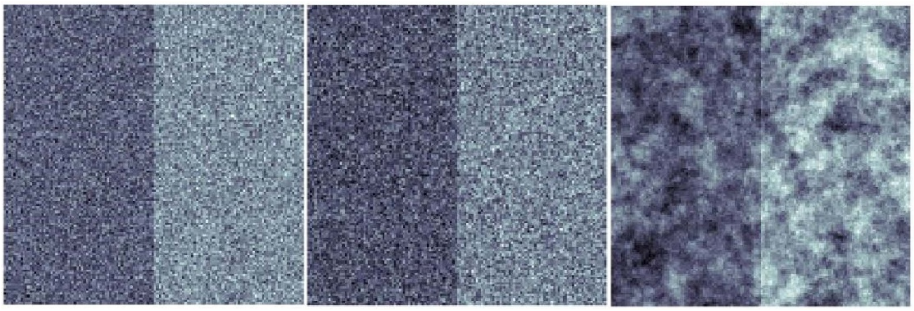


Figure 4. Simulated sample images with $\nu = 128$ px, SDNR = 1.5, but different values of the covariance length $x_{1/2}$ and hence different noise textures. Left: $x_{1/2} = 0$ px, Middle: $x_{1/2} = 0.3$ px, Right: $x_{1/2} = 3$ px.

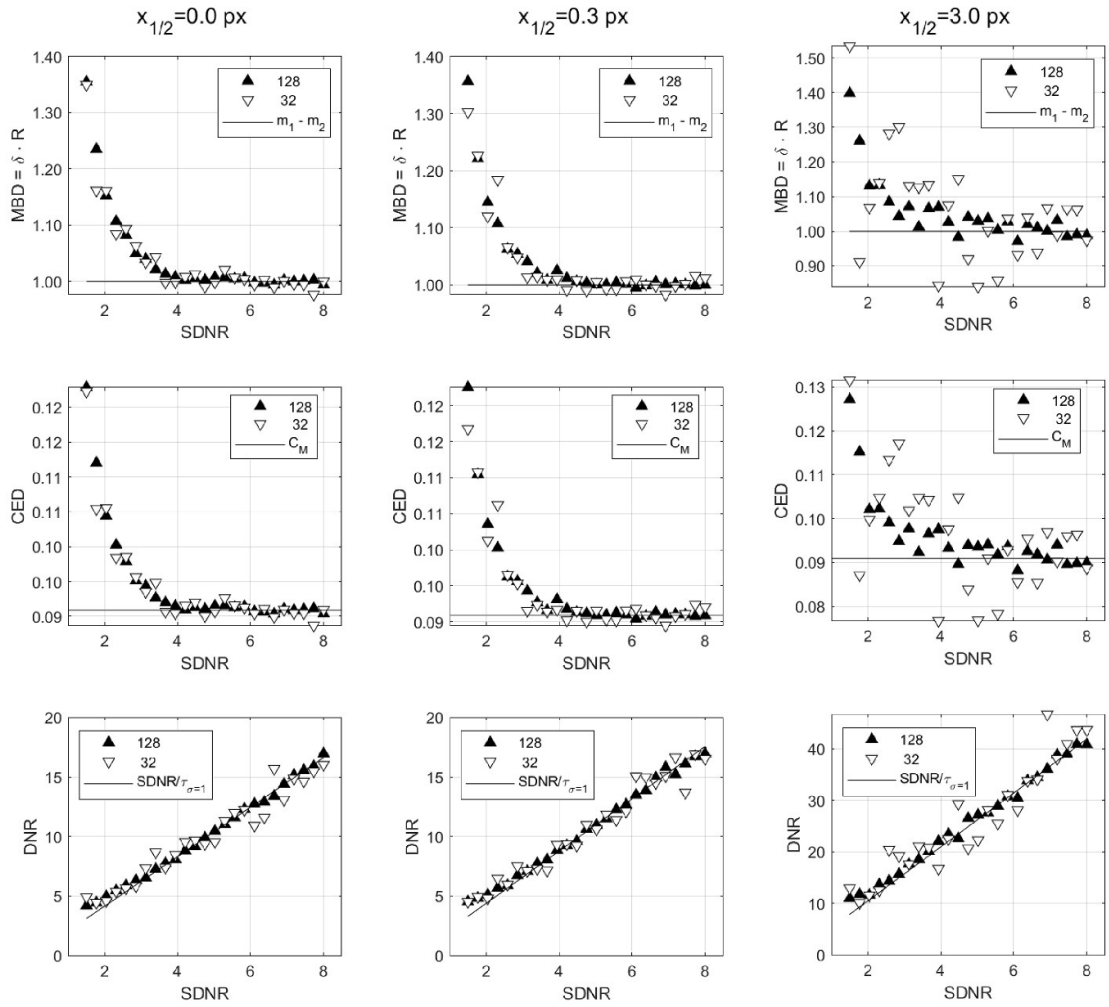


Figure 5. Mean brightness distance MBD (top), contrast equivalent distance CED (middle) and distance-to-noise ratio DNR (bottom) as a function of the simulated SDNR. Left column: uncorrelated (white) noise, $x_{1/2} = 0$; middle column: correlated noise with an exponential decay, covariance length $x_{1/2} = 0.3$ px (typical for Mammography); right column: highly correlated noise, $x_{1/2} = 3$ px. For further explanations please consult the text.

128 px and 32 px, respectively[11]. The horizontal line in the top row shows the simulated (true) grey value distance $m_1 - m_2 = 1$. In the middle row, the line represents the Michelson contrast C_M. The continuous line in the bottom row is the SDNR divided by the $\tau-$value for pure noise with $\sigma = 1$.

[11] As was demonstrated in the previous publication (Anton *et al* 2023), there is a size limit for the ROIs: ideally, the edge length of the patch should not be smaller than $\nu = 32$ px. A limit is reached at values of $\nu = 16$ px due to the increasing statistical uncertainty of τ.

Table 2. Validity limits for DNR. The parameter $x_{1/2}$ is the half-width of the exponential decay of the covariance (Anton *et al* 2023), σ/τ is the ratio of the noise standard deviation σ to the noise measure τ, and the following columns show the values of DNR which correspond to an equivalent SDNR = 2 and = 4 according to relation equation (8), respectively.

	$x_{1/2}$		$\mathrm{DNR_{lim}}$	
	in px	σ/τ	SDNR = 2	SDNR = 4
for processing	0.3	2.20	5.2	9.3
for presentation	0.5	2.45	5.8	10.3

The first and most important observation is that the MBD overestimates the true grey value distance when SDNR is lower than 4. For SDNR = 3, MBD overestimates the true grey value distance by about 5%, for SDNR = 2 the relative difference increases to 10% and increases further for lower SDNR values. The CED overestimates the contrast C_M accordingly, as is seen in the middle row. It has to be stressed that this is not an 'error' in the classical sense but a deviation from the conventional grey value distance caused by the definition of MBD$= \delta \cdot R$. Only when noise is low, i.e. for SDNR> 4, DNR and SDNR are proportional to each other.

This systematic effect appears to depend neither on the ROI size nor on the covariance length $x_{1/2}$ when the image 'content', in this case a step function, is the same for both ROI sizes. Both a lower ROI size and a higher covariance length only lead to an increased scatter in the data.

Another systematic effect is the known influence of $x_{1/2}$ on the proportionality factor between σ and τ (see figures 6 and 7 in Anton *et al* 2023) which is responsible for the change of the ratio of DNR to SDNR with varying $x_{1/2}$. This can be seen in the right part of figure 5.

From similar simulated data, a simple proportionality between σ/τ and the DNR values corresponding to SDNR = 2 and SDNR = 4 was found:

$$\mathrm{DNR_{lim}} = a \cdot \frac{\sigma}{\tau}, \tag{8}$$

where $a = 2.36 \pm 0.01$ for SDNR = 2 and $a = 4.22 \pm 0.02$ for SDNR = 4, respectively. The uncertainties are standard uncertainties determined from repeated simulations (type A, according to the *Guide to the expression of uncertainty in measurements, JCGM100 2008)*,. The scaling relation equation (8) is valid for noise with an exponential decay of the covariance (mammography) as well as for noise with a Gaussian decay of the covariance, which is suitable to model noise in CT (see, for example, appendix A of Anton *et al* 2020). In the next section, we shall demonstrate how to make use of these findings.

4. Application to mammograms of an anthropomorphic phantom

To investigate the applicability of the new measures to realistic images, images of the 3D-printed anthropomorphic breast phantom described by Schopphoven *et al* (2019), acquired on a digital mammography system (Mammomat Inspiration, Siemens, Forchheim, Germany) at 27 kV (tungsten target, 50 μm rhodium filtration) and increasing exposure time-current products from 10 mAs to 200 mAs were used, as in our previous publication (Anton *et al* 2023). The mammograms of the phantom adequately resemble those of a real breast of 32 mm compressed thickness (Schopphoven *et al* 2019). Imaging the 3D-printed phantom triggers a realistic image processing.

From our previous investigation using the same dataset (Anton *et al* 2023), it is known that the covariance length $x_{1/2}$ of the noise is ≈ 0.3 px for the *for processing* data and ≈ 0.5 px for the *for presentation* data, which is comparable to the simulated data in the middle column of figure 5. The corresponding ratios σ/τ and the resulting limits for DNR according to equation (8) are summarised in table 2.

As an example, we chose the mammogram with an exposure time-current product of 71 mAs which is close to 68.1 mAs, the value that would have been selected by the automatic exposure control (AEC)[12]. Images of the inverted for processing mammogram and the for presentation image are shown in figure 6. The mammogram was subdivided into ROIs with an edge length of 32 px. For each ROI, the image quality parameters MBD, CED, τ and DNR were calculated. The time for the calculations was only a few seconds.

The quantities CED and DNR are well suited for a comparison: by definition, they do not depend on the overall grey value range, which is different for the *for processing* and *for presentation* images.

[12] The AEC value is the one that would have been used on a patient. It is determined by a mammography device-specific, complex automatic program. Usually, a mechanical breast thickness measurement determines the voltage and the filtering for the x-ray source. After a pre-image with a small fraction of the envisaged exposure, the appropriate exposure time-current product is selected such that the required image quality is achieved with minimal dose. The criteria used are not publicly available.

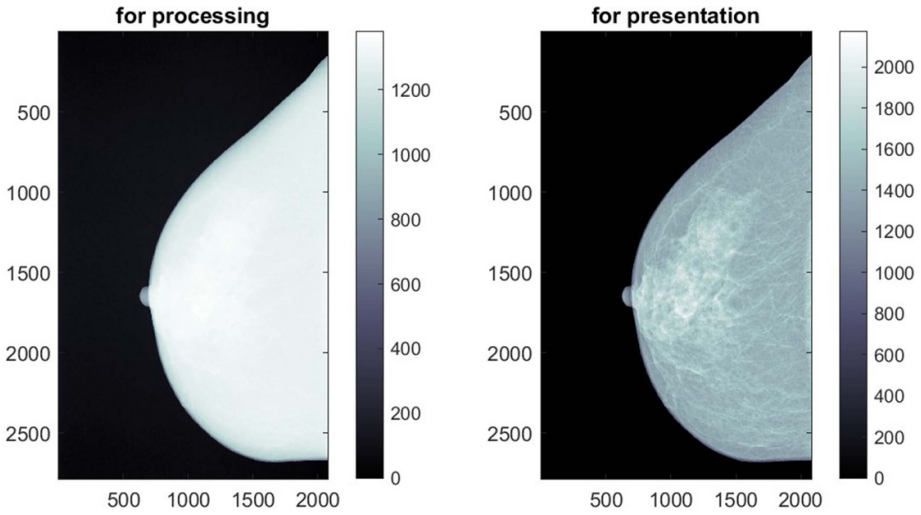


Figure 6. Mammograms of the anthropomorphic phantom for an exposure-time current product of 71 mAs. Left: inverted *for processing* image, Right: *for presentation*.

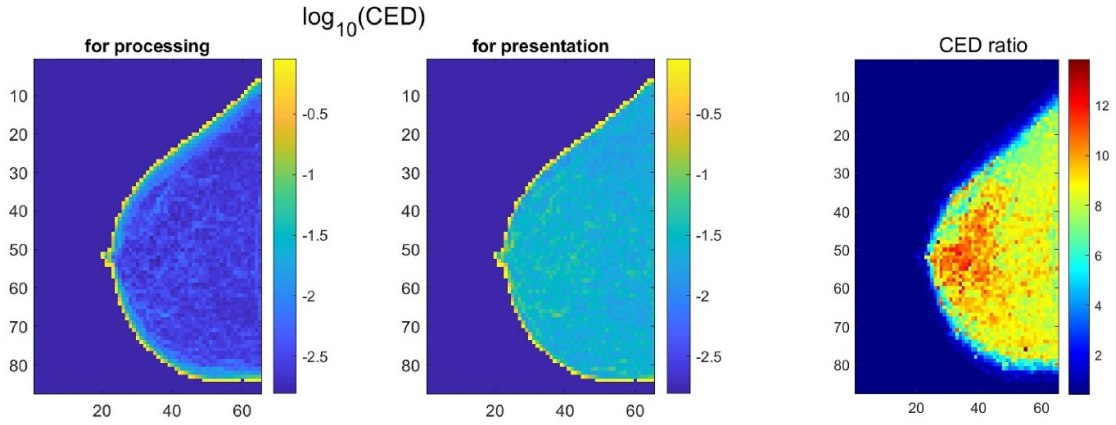


Figure 7. Left and middle panels show $\log_{10}$(CED). The *for processing* results are shown on the left, the *for presentation* ones in the middle. On the right, the ratio of CED (*for presentation*) to CED (*for processing*) is shown.

Colourmaps of the CED are shown in figure 7. To facilitate a visual comparison of the results, the logarithm of CED is shown. Already by direct comparison between the left and middle panel it is clearly seen that the image processing has increased the contrast significantly. On the right of figure 7, the ratio of CED (*for presentation*) to CED (*for processing*) is displayed, which highlights the strong increase of contrast in the breast region with a high fraction of glandular tissue. The median increase over the whole breast is of about a factor of 8. Near the skin-air boundary, on the other hand, a small or no improvement is achieved by the processing. However, the contrast is highest there, anyway.

A similar comparison for the DNR is shown in figure 8. On the left, the *for processing* results are shown, in the middle, the *for presentation* data are displayed, and on the right, the ratio of DNR (*for presentation*) to DNR (*for processing*) is shown. As was the case for the CED data, the logarithm is displayed to make a direct visual comparison possible (see left and middle panel). The comparison is much easier when the ratio is considered, which can be seen on the right side: by the processing, the DNR is increased by about 32% (median of the ratio *for presentation*/*for processing*).

Comparing CED and DNR, it is observed that the contrast measured by CED increases more strongly and with a greater local variability than the distance-to-noise ratio DNR.

The question may be raised whether the presented data are reliable. In figure 9, the ROIs which fulfil the criterion DNR $\geqslant$ $\mathrm{DNR_{lim}}$ (SDNR $= 4$) are marked in white, others are black. On the left, the data are shown for the *for processing* image, on the right for the *for presentation* image. In case of the *for processing* image, 38.5% of the ROIs inside the breast fulfil the criterion, in case of the *for presentation* image, the fraction is

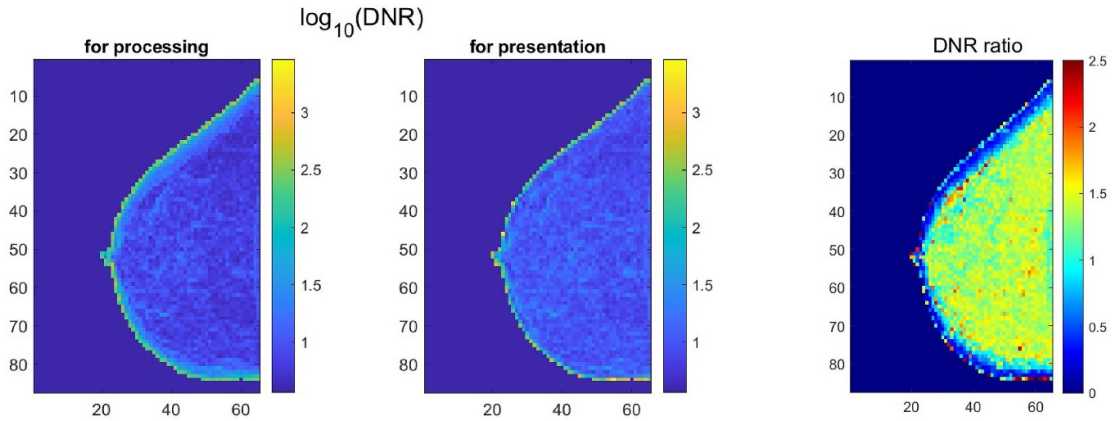


Figure 8. The $\log_{10}$(DNR) for the *for processing* image is shown on the left, $\log_{10}$(DNR) for the *for presentation* image in the middle. On the right, the ratio of DNR (*for presentation*) to DNR (*for processing*) is shown.

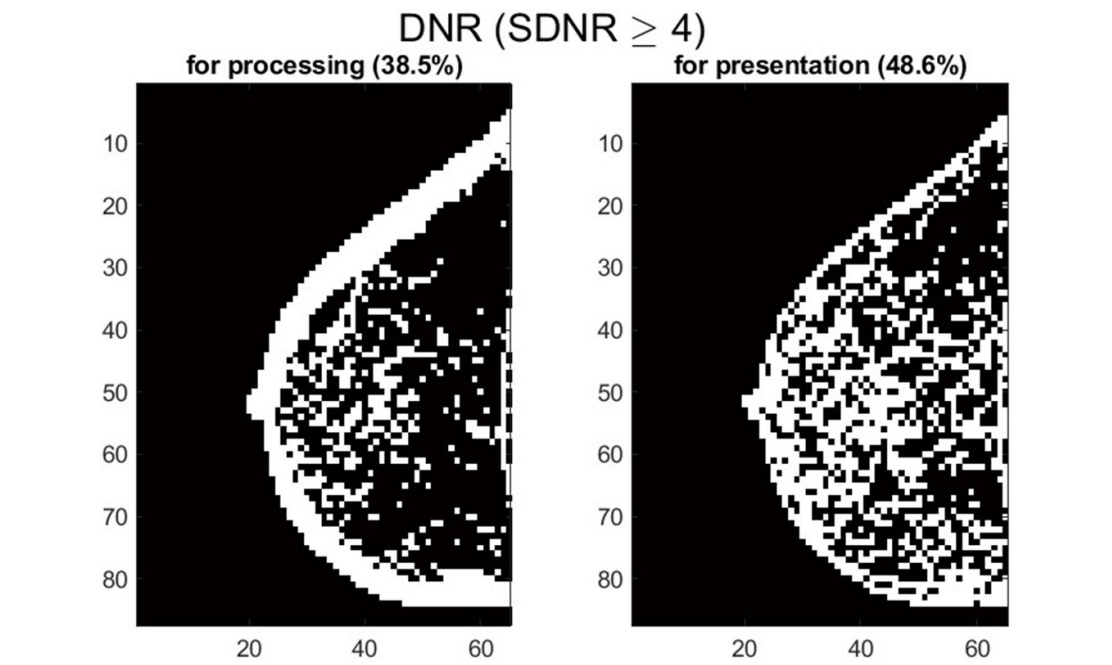


Figure 9. Regions (ROIs) where DNR exceeds $\mathrm{DNR_{lim}}$(SDNR = 4) are shown in white (compare table 2). The *for processing* results are shown on the left, the *for presentation* ones on the right. The values in the title indicate the percentage of ROIs inside the breast which fulfil the criterion DNR$\geqslant \mathrm{DNR_{lim}}$(SDNR = 4). For DNR$\geqslant \mathrm{DNR_{lim}}$(SDNR = 2), the breast regions would be almost completely white.

48.6%. The corresponding maps for DNR $\geqslant \mathrm{DNR_{lim}}$ (SDNR = 2) are not shown, since the breast regions would almost completely be shown in white: the fraction fulfilling DNR $\geqslant \mathrm{DNR_{lim}}$ (SDNR = 2) is 92.8% (*for processing*) and 99% (*for presentation*), respectively. This means that a quantitative evaluation of the image quality parameters CED and DNR is indeed reasonable, since almost everywhere inside the breast region systematic deviations are smaller than 10% (compare section 3). The maps shown in figure 9 also indicate which regions deserved special attention if the mammogram was of a real patient: regions with a lower SDNR (the black regions) would have to be inspected more carefully.

Histograms obtained for the parameters shown in figures 7 and 8 generally bear few similarities with normal distributions. Therefore, calculating means and standard deviations would make no sense. Instead, a convenient way to summarise the information contained in these maps is the use of box-and-whisker charts or the corresponding five-number summary statistics (see, for example, Hoaglin *et al* 1983). As an example, box-and-whisker charts for the ratios CED (*for presentation*) to CED (*for processing*) and DNR (*for presentation*) to DNR (*for processing*) are displayed in figure 10. Information on CED and DNR is

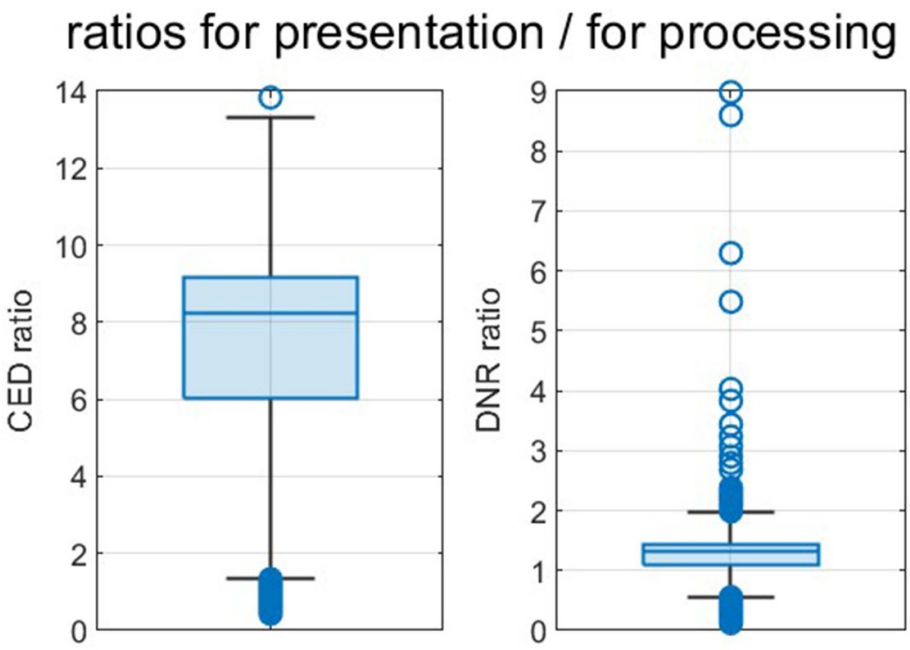


Figure 10. Left: box-and-whisker chart for the ratio CED (*for presentation*) to CED (*for processing*) (compare figure 7, right), the median value is 8.25. Right: box-and-whisker chart for the ratio DNR (*for presentation*) to DNR (*for processing*). The median of the ratio is 1.32. Note that only ROIs with DNR$\geqslant$ DNR$_{\mathrm{lim}}$(SDNR = 2) are taken into account.

Table 3. Summary statistics Q_1 (first quartile), Median (second quartile) and Q_3 (third quartile) for the examples shown in figure 6. $Q_3 - Q_1$ is the so-called interquartile range and represents a 50% coverage interval for the distribution of the parameters.

	for processing			*for presentation*		
Parameter	Q_1	Median	Q3	Q_1	Median	Q3
CED in %	0.248	0.309	0.482	2.012	2.463	3.184
DNR	6.561	8.287	12.847	8.387	10.257	13.435

summarised in table 3: The first quartile Q_1, the second quartile (the median) and the third quartile Q_3 are given, which correspond to the limits of the box and the horizontal line within the box of the corresponding box-and-whisker-chart, for both the *for processing* and the *for presentation* image.

Summarising, one can state a median increase of the CED by about a factor of 8 due to the image processing and a median increase of the DNR by about 30%.

Obviously, the procedure described above does not only work on entire images but also on parts of it: larger regions of interest can be specified (and subdivided in smaller ROIs) to analyse local quality parameters.

5. Discussion

Based on the mean brightness distance MBD$= \delta \cdot R$, the contrast measure CED can be determined.

For two distinct regions with different grey values and low noise, the MBD is equivalent to the conventional grey value distance that is defined as the difference of the mean grey values of these regions, as is demonstrated in appendix A.

However, when noise is more prominent, the direct correspondence of the MBD to the conventional grey value distance is lost. In the limiting case when only noise is present, MBD is proportional to the noise standard deviation σ.

In order to determine whether MBD and hence CED is a valid measure of contrast, the distance-to-noise ratio DNR $=$ MBD $/\tau$, a generalisation of the SDNR, has to be determined. τ is a nonparametric measure of noise that can also be determined directly from patient images (Anton *et al* 2023).

A relation between the SDNR of a simple grey value step and the DNR can be established when the ratio of the standard deviation σ of the noise to τ is known. The ratio σ/τ is ≈ 2 for white noise but increases for

noise with a slower decay of the covariance with the px-px distance (see Anton *et al* 2023). Using simulated data with noise of different textures, i.e. with different noise covariances, we have shown in section 3 that DNR-values that correspond to specific SDNR values can be obtained by multiplying σ/τ with a simple proportionality factor (see equation (8)). With the help of these limits, validity regions for the MBD and hence for CED can be established.

When DNR exceeds a value $\mathrm{DNR_{lim}}$ that is equivalent to an SDNR of 4, no systematic deviations between the MBD and the true grey value distance are expected. If the DNR is higher or equal than the value equivalent to SDNR = 2, but smaller than the value equivalent to SDNR = 4, systematic deviations will not exceed 10%. Therefore, we consider the CED a valid measure of contrast when DNR is higher or equal to $\mathrm{DNR_{lim}}(\mathrm{SDNR} = 2)$.

It has to be stressed that the deviation for higher noise levels is not an error in the classical sense, but can be viewed as the price to pay for the possibility to measure contrast without imaging a special SDNR phantom.

To determine the validity limits of the new measures, it is essential to measure τ, but also σ. For simulated data, the ratio σ/τ is readily available. In some cases, like in the mammography example of section 4, the ratio of σ/τ may be available from separate measurements, in this case of a homogeneous PMMA phantom of equivalent thickness. In other cases, such as the CT example shown in appendix B, homogeneous regions in the image may be present so that σ/τ can be evaluated directly from the image under investigation. In general, however, a universal tool which allows one to estimate the σ/τ ratio under any circumstances is needed and will be the subject of future work.

The described image quality parameters can be applied to an entire image or to regions of it. In the mammography example of section 4, the whole mammogram was subdivided into ROIs with an edge length of 32 px, whereas in the CT example in appendix B, the images were investigated as a whole. When the image quality parameters are evaluated on smaller ROIs, graphical presentations of their local variation are available, as we have shown. To summarise the information thus obtained, box-and-whisker charts or their underlying five-number summary statistics (see, for example, Hoaglin *et al* 1983) are convenient, since the resulting distributions of the parameters are seldom normal distributions. The limiting value $\mathrm{DNR_{lim}}(\mathrm{SDNR} = 2)$ has to be calculated to be able to exclude ROIs with invalid Δ-values from the summary statistics. The actual values of the summary statistics may, however, also depend on the ROI size chosen. Since it is necessary to calculate τ along with the MBD for the ROIs, it is recommended to use ROIs with an edge length of 32 px or more.

With our mammography example, we have shown that the increase of the contrast measured by CED and the increase of the distance-to-noise ratio DNR due to the image processing can be quantified. The local character of the enhancement is apparent: the contrast enhancement is strongest in the region with more glandular tissue, which appears reasonable. In addition, by looking at the local $\mathrm{DNR_{lim}}$-maps, regions with a lower signal-difference-to noise ratio can be identified, as we have shown as well. The CT example of appendix B served to demonstrate that the concept also works when more than two distinct grey value levels are distinguished, e.g. by using Otsu's method for multiple thresholds. Handling and keeping track of multiple thresholds when the image is subdivided into smaller ROIs is the subject of ongoing work.

6. Summary and outlook

Using the order statistics of an image (or of ROIs of an image), the relative brightness distance δ can be derived from any kind of image. Multiplication of δ with the grey value range R yields the mean brightness distance MBD. For conventional cases with a defined grey value step and low noise, values obtained for MBD coincide with values of the grey value distance obtained from established methods. Not using the information on the location of 'high' and 'low' grey value regions, as is the case when the MBD is determined from clinical images, requires the knowledge of the noise level as well as knowledge about the noise texture. If both informations are available, it is possible to determine whether the contrast equivalent distance CED, a measure derived from the MBD, is a valid measure of contrast. The measures can be determined locally on small ROIs of an image down to a size of 32×32 px. As a summary statistic for the quantification of the corresponding parameters for the whole image, a five-number summary like the one used for box-and-whisker plots is suggested. The proposed procedure is demonstrated for mammograms of an anthropomorphic phantom and for CT images of a head phantom.

From the results, it is obvious that the envisaged set of non-parametric figures of merit for the quantification of image quality will not consist of a set of independent ones: MBD and hence CED cannot be interpreted without knowing τ and a measure for the σ/τ ratio which is related to the noise texture. Possible

additional parameters like, e.g. the relative mutual information, may turn out to be helpful as well. As a next step, a nonparametric, local measure of the noise texture is required. This is the subject of ongoing work. Once a set of figures of merit is considered complete, their relation to the assessment by experts will have to be established.

Data availability statement

The data that support the findings of this study are available upon request from the authors.

Appendix A. The MBD for a step wedge of a quality assurance phantom

The left side of figure A1 shows a mammography image (for processing) of a quality assurance phantom according to the German standard DIN 6868-14 (NAR 2020). The image was acquired on a *Hologic Selenia Dimensions* device with a Tungsten anode, 28 kV, rhodium filtering, and an exposure time-current product of ≈100 mAs. The right side of figure A1 shows the ROI containing only the step wedge that was used for the evaluation of the grey value distance between successive grey levels. The size of the ROI is approximately 600×2400 px.

For the standard evaluation of the grey value distance, separate, square (sub-) ROIs with an edge length of 200 px were centred in each grey level area and the mean grey value m_n for each level n was calculated. The same was done independently at the German reference centre for mammography south-west (RZSW) using the software that is regularly used for quality assurance measurements.

For the evaluation using δ, a peakfinding algorithm (peakfinder, Yoder 2023), retrieved from the Matlab® Central file exchange was applied to the histogram and the mid-points between the peaks were used as thresholds, see figure A2. Equation (2) was then applied to each consecutive pair of grey levels. One of the pairs is highlighted in figure A2 by the shaded area. As a result, fourteen MBD $= \delta \cdot R$-values were obtained.

The results are listed in table A1: the leftmost column gives the indices of the grey levels, where 'offset' corresponds to the offset area (the black rectangular region on top of the right part of figure A1), the other numbers follow counterclockwise, from highest to lowest level. The white square on the top left is level 0, the bottom squares correspond to levels 7 (left) and 8 (right). Mean grey values m_n and corresponding distances $m_{n-1} - m_n$ are listed in the next two columns for the RZSW evaluation. The grey value distance listed for level 13 is the difference to the offset. The following two columns list m_n and $m_{n-1} - m_n$ for the centred ROI evaluation at PTB using Matlab®, the column marked 'rel. diff.' gives the relative difference between the ROI evaluations at PTB and RZSW in per cent. The following three columns display the results from the evaluation of the MBD $= \delta \cdot R$. The column with header MBD contains the mean brightness distance, the next one gives the relative difference to the RZSW $(m_{n-1} - m_n)$-evaluation in per cent and the last column yields the reproducibility. The latter was estimated from ten repeated evaluations and is quantified by the coefficient of variation CV, i.e. the relative standard deviation. Differences between the repeated evaluations are due to the manual positioning of the ROI.

Only for the lowest grey levels (11–13) does the relative difference between RZSW and PTB ROI evaluation exceed one per cent; for levels 0-9 the relative difference is smaller than 0.5%. The difference between $m_{n-1} - m_n$ (RZSW) and the MBD is slightly higher. It is largest for the two highest levels (0 and 1). For all other levels, the relative difference exceeds 1% just once. The comparatively large relative differences of 3%–4% for levels 0 and 1 are not explained by a lack of reproducibility, as can be deduced from the rightmost column of the table. Since none of the SDNR values in this example is less than 5, the simulation results in section 3, figure 5 suggest that the systematic deviations between $m_{n-1} - m_n$ measured according to the German standard and the MBD are also not caused by noise.

Instead, they result from the fact that for the centred ROI evaluations (RZSW and PTB), boundary regions between different steps are excluded, whereas for the $\delta \cdot R$-evaluation, the whole of the larger ROI containing the entire step wedge is used, which comprises also edge regions which exhibit in some regions slightly different grey values.

However, this means that for the case when there are well separated grey levels, the MBD$= \delta \cdot R$ evaluation yields results equivalent to established methods.

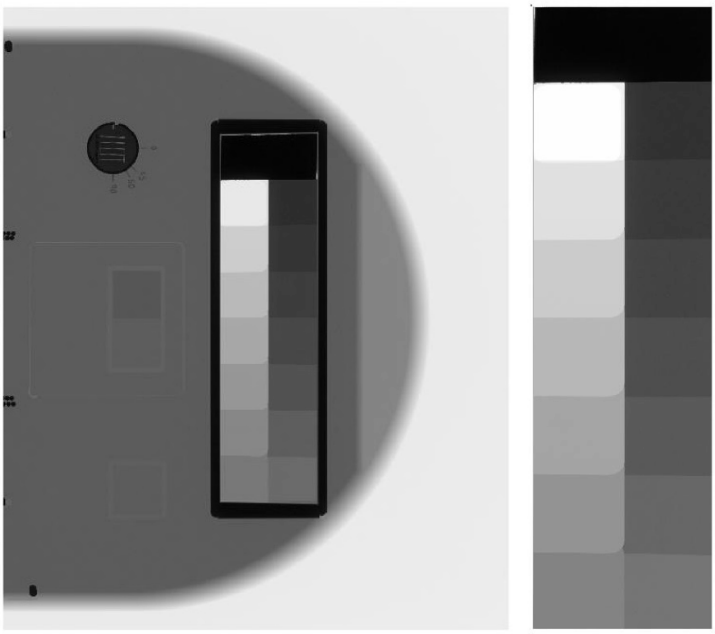


Figure A1. Left: mammography image (for processing) of a quality assurance phantom according to DIN 6868-14; Right: the ROI containing only the step wedge for the evaluation with δ. For the graphic representation, the logarithm of the pixel values is shown.

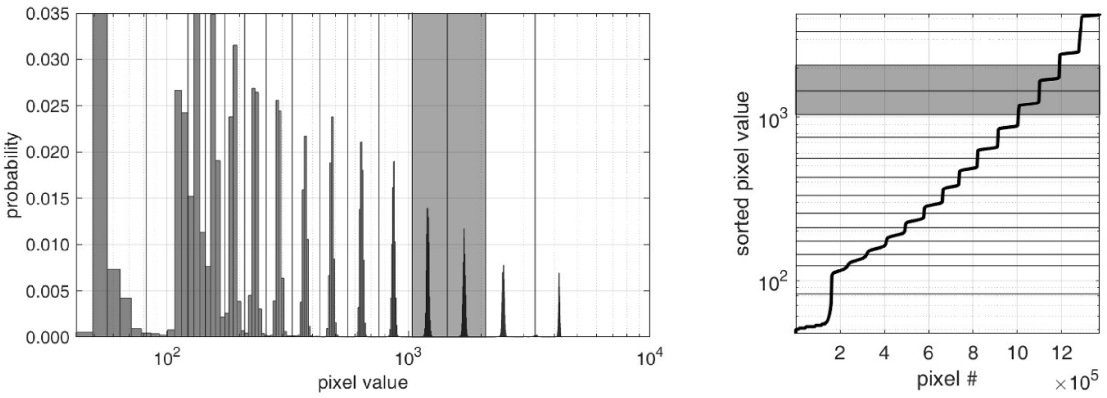


Figure A2. Left: Histogram of the pixel values present in the ROI of figure A1, right. Please note the logarithmic x-axis, which is used for better visibility of the separate peaks. Vertical lines indicate the thresholds determined by a peakfinder algorithm. Right: the pixel values of the same ROI in ascending order. The same couple of grey levels is highlighted by the shaded area in both parts of the figure. Horizontal lines represent the thresholds. Please note the logarithmic y-axis. For further explanations please consult the text.

Table A1. Grey value distances for a step wedge with 14 grey levels + offset. The leftmost column indicates grey level numbers, where '0' corresponds to the highest level (white, top left in figure A1) and '13' to the lowest one (top right). 'offset' is the black rectangular area on top. The columns titled ROI/RZSW contain the results of an evaluation by the RZSW with the software used for routine quality assurance measurements. The columns ROI/PTB give the results of an evaluation of the centred ROIs using Matlab®. m is the mean grey value of a level, $m_{n-1} - m_n$ is the grey value distance. The columns with header $\delta \cdot R$ list the results from the δ-calculation using the midpoints between the peaks of the histogram as thresholds. Columns marked 'rel. diff.' list the relative difference to the RZSW evaluation in per cent. The rightmost column yields the reproducibility of the $\delta \cdot R$-Results as a coefficient of variation in per cent derived from ten repeated evaluations.

Evaluation	ROI/RZSW		ROI/PTB			$\delta \cdot R$		
Level n	m	$m_{n-1} - m_n$	m	$m_{n-1} - m_n$	rel. diff. in %	MBD	rel. diff. in %	CV_{10} in %
Offset	52.13	—	52.00	—	—	—	—	—
0	4204.71	1731.61	4203.32	1731.65	0.0	1661.31	−4.1	0.01
1	2473.10	769.97	2471.67	769.09	−0.1	790.89	2.7	0.1
2	1703.13	500.98	1702.58	500.39	−0.1	499.33	−0.3	0.2
3	1202.15	334.24	1202.19	333.92	−0.1	334.01	−0.1	0.1
4	867.91	228.87	868.27	229.04	0.1	228.33	−0.2	0.1
5	639.04	158.47	639.23	158.61	0.1	158.57	0.1	0.1
6	480.57	107.66	480.62	108.15	0.5	108.27	0.6	0.1
7	372.91	82.31	372.47	82.09	−0.3	81.44	−1.0	0.3
8	290.61	59.56	290.38	59.31	−0.4	59.43	−0.2	0.3
9	231.04	43.27	231.07	43.09	−0.4	43.07	−0.5	0.2
10	187.77	31.81	187.98	31.59	−0.7	32.07	0.8	0.7
11	155.96	23.96	156.39	23.62	−1.4	23.93	−0.1	0.8
12	132.00	18.14	132.77	17.86	−1.5	18.11	−0.1	1.7
13	113.86	61.73	114.91	62.91	1.9	61.68	−0.1	1.7

Table B1. Parameters of the four CT images: The first column lists the label (compare figures B1–B3), the second one the designation of the reconstruction kernel, the third column lists the ratio σ/τ and the following columns list the DNR limits for an equivalent SDNR = 2 and 4, respectively, according to equation (8).

			DNR_{lim}	
Label	Kernel	σ/τ	SDNR = 2	SDNR = 4
A	H20f	14.41	34.01	60.82
B	H40f	9.23	21.78	38.95
C	H50f	5.39	12.72	22.75
D	H60f	4.32	10.19	18.22

Appendix B. Applying CED and DNR to CT images

We have applied the methods outlined in the previous sections to CT images of a head phantom. Data was acquired on a Siemens Somatom CT (already in 2013) and reconstructed using filtered back projection with different convolution kernels. A protocol for the investigation of a sinusitis had been used, with a voltage of 100 kV, an exposure time-current product of 100 mAs and different convolution kernels, ranging from a soft, smoothing kernel (H20f) to a rather 'hard' kernel (H60f, see table B1).

The segmentation was performed using Otsu's method for multiple thresholds (Matlab's® built-in function `multithresh`) using the histograms shown in figure B1.[13]

The resulting segmented images are displayed in figure B2. The segmentation is less successful for the image with the hard kernel (image D, bottom right), especially level 2 and 3 are hard to separate. The image with the softest kernel (A, top left) exhibits the clearest segmentation of the different materials. The harder kernel, however, appears to produce a sharper picture (D, bottom right).

For the estimation of σ/τ, a region of interest (ROI) was marked in the homogeneous, w-shaped region that represents a part of the brain, see the lower middle part of the images in figure B2. The results for σ/τ as well as the resulting DNR_{lim} values corresponding to SDNR = 2 and SDNR = 4, respectively, are listed in table B1. Results for the contrast equivalent distance CED and for the distance-to-noise ratio DNR are shown in figure B3.

Comparison of figure B3, right, with table B1 shows that in all cases the DNR exceeds the limiting value with an equivalent SDNR of 2. Only the DNR value associated to transition 2-3 for the hardest reconstruction kernel (D, H60f) is smaller than the DNR limit for SDNR = 4. Therefore, CED and DNR can be considered

[13] To obtain HU numbers instead of the given pixel values, a value of 1024 has to be subtracted. Dealing with negative numbers would cause undue difficulties when calculating the image quality parameters.

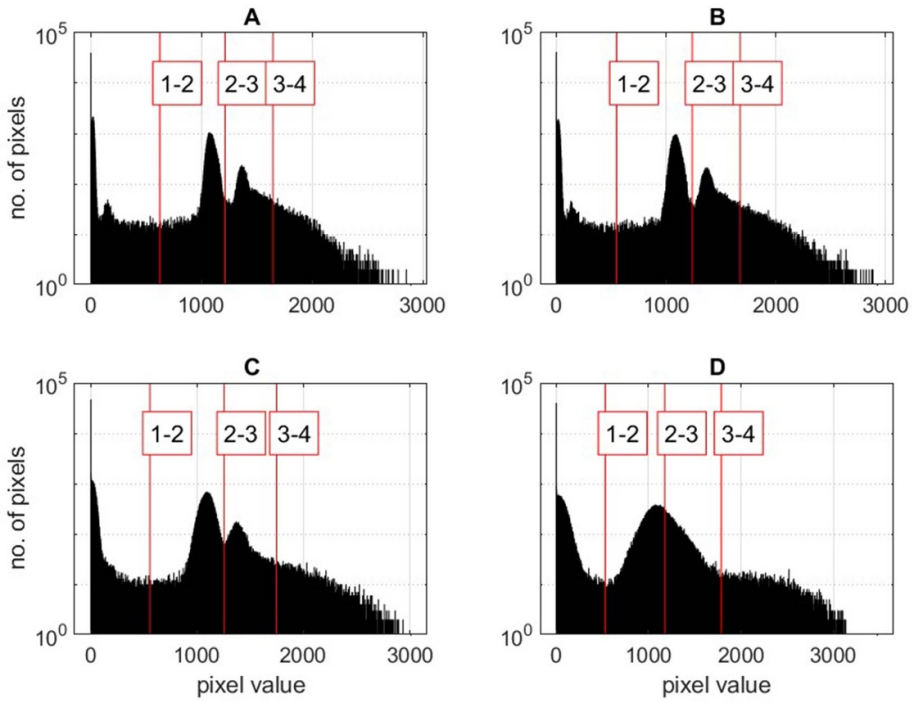


Figure B1. Histograms of four CT images of a head phantom. The raw data are identical (see text), only different convolution kernels were used ranging from a strongly smoothing kernel (A, H20f) to a rather 'hard' kernel (D, H60f), see table B1. Here, the logarithm of the counts (number of pixels) in each bin is plotted vs the pixel value (grey value). Vertical lines indicate the threshold values determined by Otsu's method for multiple thresholds. Labels a-b indicate the threshold between level a and level b.

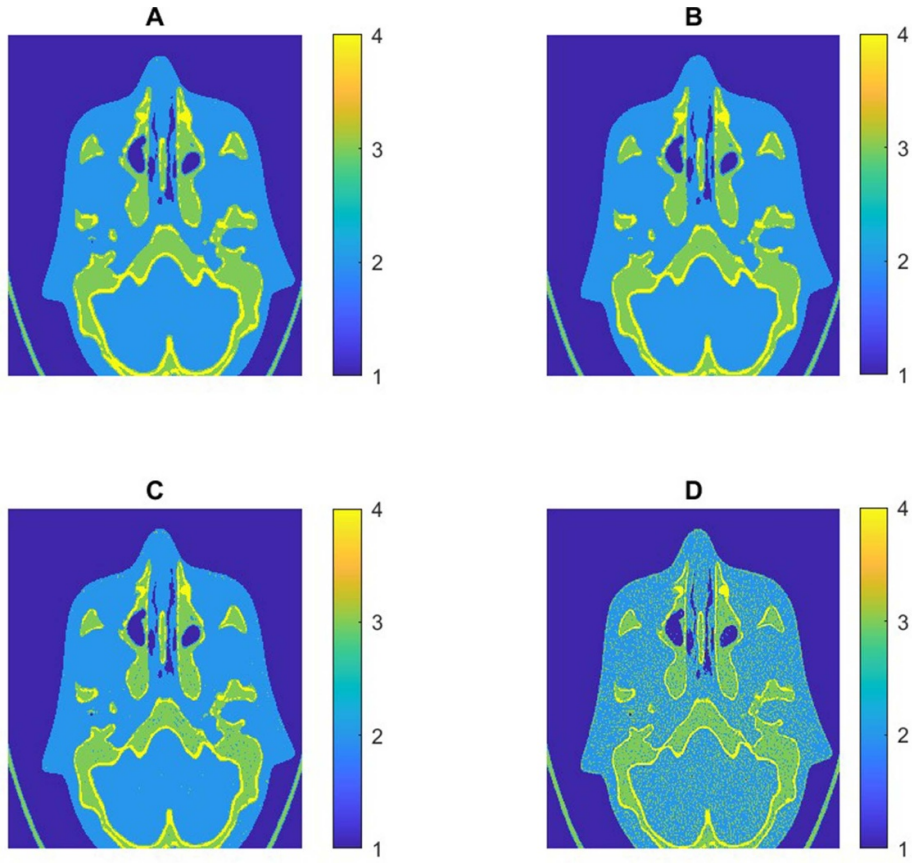


Figure B2. Images segmented using the thresholds given in figure B1. The labels in the colourbar are the same as the ones indicated in figures B1 and B3. Parameters are given in table B1.

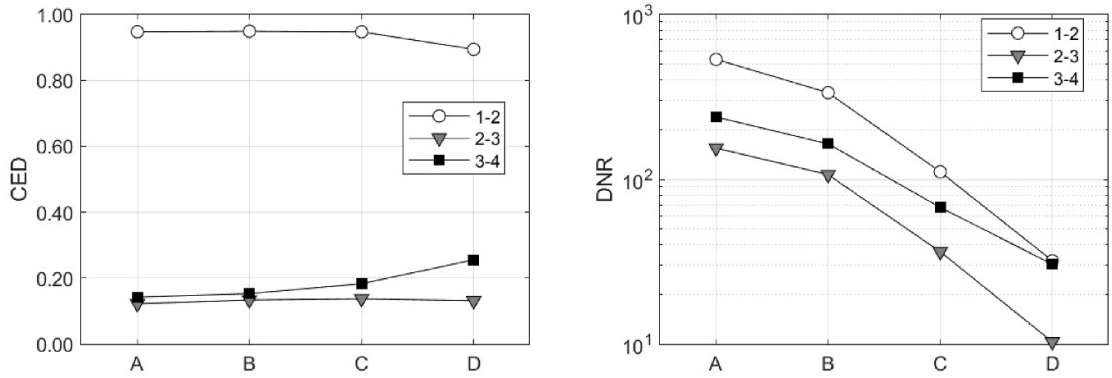


Figure B3. Image quality parameters for different convolution kernels (compare table B1). The legend refers to the threshold indicated in figure B1. Left: the contrast equivalent distance CED. Right: distance-to-noise ratio DNR. Note the logarithmic scale for the DNR.

valid measures of contrast and contrast-to-noise ratio. Note that the reduction of DNR with harder kernels is so strong that a logarithmic scale was chosen for the graphical presentation (see figure B3, right).

As is obvious, the contrast between air and head is highest (transition 1-2), the smallest difference occurs between levels 2-3. According to the results shown on the left of figure B3, the contrast between the two types of bone (green and yellow in figure B2, transition 3-4) is increased by using the harder convolution kernel, whereas other contrasts appear less pronounced than with softer kernels. This corresponds to the somewhat 'sharper', albeit noisier impression of image D.

Using this example, it is demonstrated that the proposed measures can also be applied to entire CT images. The software for dealing with more than one threshold when the image is subdivided into smaller ROIs is work in preparation.

Appendix C. Mutual information as a measure of structure

The relative grey value distance δ alone may not be sufficient to characterise the perceived contrast in an image (or in a part of it), which is easily seen considering an image with the same number of just black and white pixels, distributed randomly: while the measure δ would be unity (where the threshold can be anywhere between the two extremes), indicating a maximum relative grey value distance, the image itself would not be perceived as high-contrast, due to the lack of order.

In 2019, Nowosad and Stepinski (2019) published their results on the quantification and classification of landscape patterns, based on the co-occurrence matrix introduced by Haralick *et al* (1973). In order to use the concept, the pixel values of the images have to be arranged in classes. In case of grey-level images as in our case, this simply means a segmentation into a (pre-selected) number M of grey levels c_i. The co-occurrence matrices contain the relative frequencies for a central pixel having grey value c_i that its neighbours in vertical, horizontal and two diagonal directions have a grey value c_j. Thereby, an $M \times M$ co-occurrence matrix can be set up (for every one of the four directions, and one summarizing all directions). From this two-dimensional histogram, different entropy values can be calculated and finally a relative mutual information U which can take on values between 0 and 1.

We have used the definitions of the co-occurrence matrices as given by Haralick *et al* (1973) (see also Shahbahrami *et al* 2012) and the information-theoretical approach including the definition of the relative mutual information U from Nowosad and Stepinski (2019), therefore these equations will not be reproduced here.

For the calculation of δ, some threshold has to be chosen. Therefore, it appears natural to choose the same threshold for a binarisation of the image (i.e. $M = 2$) and calculate the mutual information from the binarised image. As an example, the corresponding binarised versions of the images from figure 3 are displayed in figure C1, along with the resulting values of the relative mutual information U. For the constant grey value image with low noise and just a few outliers (panel 1 on the left), the relative mutual information is zero. The same value results for the two images that contain only noise (panel 2 and panel 3). Although the δ-values for panel 3 (uniformly distributed noise) and panel 4 (a grey value ramp) are the same, the relative mutual information U is much higher for the ramp. It is almost equal to the mutual information of the step (panel 5), but due to the less sharp boundary of the ramp at the threshold value, which is caused by the noise, U is smaller for the ramp. When comparing panels 5 and 6 that both contain as many 'black' as 'white' pixels

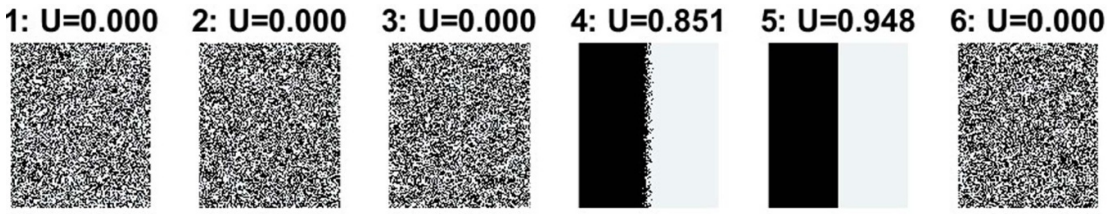


Figure C1. Binarised versions of the images from figure 3 using the same threshold as the one used for the calculation of δ.

(apart from the low noise), it is observed that the relative mutual information for the step is much higher, the mutual information for the randomised image is zero, although the FOM for the contrast will be the same.

This example is added to illustrate the fact, that, in order to interpret the figures of merit for the contrast like MBD and CED, it will be helpful to supply parameters for noise (τ) and noise texture (to be developed) as well and also a parameter such as U that is capable of measuring the structure within the image or the ROI under investigation. How this set of parameters is to be combined for image quality assessment is the subject of ongoing work.

ORCID iDs

M Anton https://orcid.org/0009-0009-7143-1725
M Reginatto https://orcid.org/0000-0002-5811-6036
U Mäder https://orcid.org/0000-0002-7330-4916
F Mauter https://orcid.org/0009-0008-9560-9957

References

Abadi E, Sanders J and Samei E 2017 Patient-specific quantification of image quality: an automated technique for measuring the distribution of organ Hounsfield units in clinical chest CT images *Med. Phys.* **44** 4736–46

Anton M, Mäder U, Schopphoven S and Reginatto M 2023 A nonparametric measure of noise in x-ray diagnostic images – mammography *Phys. Med. Biol.* **68** 045003

Anton M, Reginatto M, Elster C, Mäder U, Schopphoven S, Sechopoulos I and van Engen R 2021 The regression detectability index RDI for mammography images of breast phantoms with calcification-like objects and anatomical background *Phys. Med. Biol.* **66** 225015

Anton M, Veldkamp W, Hernandez-Giron I and Elster C 2020 RDI—a regression detectability index for quality assurance in x-ray imaging *Phys. Med. Biol.* **65** 085017

Balta C, Bouwman R W, Sechopoulos I, Broeders M J M, Karssemeijer N, van Engen R E and Veldkamp W J H 2019 Can a channelized Hotelling observer assess image quality in acquired mammographic images of an anthropomorphic breast phantom including image processing? *Med. Phys.* **46** 714–25

Barrett H H and Myers K J 2003 *Foundations of Image Science* (Wiley-VCH)

Bielecka M, Bielecki A, Obuchowicz R and Piórkowski A 2020 Universal measure for medical image quality evaluation based on gradient approach *Computational Science—ICCS 2020*, ed V V Krzhizhanovskaya, G Závodszky, M H Lees, J J Dongarra, P M A Sloot, S Brissos and J Teixeira (Springer International Publishing) pp 406–17

Chow L S and Paramesran R 2016 Review of medical image quality assessment *Biomed. Signal Process. Control* **27** 145–54

Christianson O, Winslow J, Frush D P and Samei E 2015 Automated technique to measure noise in clinical CT examinations *Am. J. Roentgenol.* **205** W93–W99

Ferzli R and Karam L J 2009 A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB) *IEEE Trans. Image Process.* **18** 717–28

Haralick R M, Shanmugam K and Dinstein I H 1973 Textural features for image classification *IEEE Trans. Syst. Man Cybern.* **6** 610–21

Hoaglin D C, Mosteller F and Tukey J W 1983 *Understanding Robust and Exploratory Data Analysis* (Wiley)

ICRP 2007 The 2007 recommendations of the International Commission on Radiological Protection: ICRP publication 103 *Ann. ICRP* **37**

ICRU 2006 ICRU Report No. 54: Medical imaging—the assessment of image quality *J. ICRU* **6** N

JCGM100 2008 Evaluation of measurement data—guide to the expression of uncertainty in measurement. GUM 1995 with minor corrections *Technical Report* (BIPM, Working Group 1 of the Joint Committee for Guides in Metrology (JCGM/WG 1))

Lecca M, Rizzi A and Serapioni R P 2021 An image contrast measure based on Retinex principles *IEEE Trans. Image Process.* **30** 3543–54

Liu L, Liu B, Huang H and Bovik A C 2014 No-reference image quality assessment based on spatial and spectral entropies *Signal Process. Image Commun.* **29** 856–63

Ma K, Liu W, Zhang K, Duanmu Z, Wang Z and Zuo W 2017 End-to-end blind image quality assessment using deep neural networks *IEEE Trans. Image Process.* **27** 1202–13

Mittal A, Moorthy A K and Bovik A C 2012 No-reference image quality assessment in the spatial domain *IEEE Trans. Image Process.* **21** 4695–708

Monnin P, Marshall N W, Bosmans H, Bochud F O and Verdun F R 2011 Image quality assessment in digital mammography: part II. NPWE as a validated alternative for contrast detail analysis *Phys. Med. Biol.* **56** 4221

Moorthy A K and Bovik A C 2011 Blind image quality assessment: from natural scene statistics to perceptual quality *IEEE Trans. Image Process.* **20** 3350–64
Myers K J and Barrett H H 1987 Addition of a channel mechanism to the ideal-observer model *J. Opt. Soc. Am.* A **4** 2447–57
NAR 2020 DIN 6868-14: Image quality assurance in x-ray departments—Part 14: constancy testing of x-ray installations for digital mammography *Technical Report* (DIN Normenausschuss Radiologie (NAR))
Nowosad J and Stepinski T F 2019 Information theory as a consistent framework for quantification and classification of landscape patterns *Landscape Ecol.* **34** 2091–101
Obuchowicz R, Oszust M, Bielecka M, Bielecki A and Piórkowski A 2020 Magnetic resonance image quality assessment by using non-maximum suppression and entropy analysis *Entropy* **22** 220
Otsu N 1979 A threshold selection method from gray-level histograms *IEEE Trans. Syst. Man Cybern.* **9** 62–66
Panetta K, Samani A and Agaian S 2018 A robust no-reference, no-parameter, transform domain image quality metric for evaluating the quality of color images *IEEE Access* **6** 10979–85
Peli E 1997 In search of a contrast metric: matching the perceived contrast of Gabor patches at different phases and bandwidths *Vis. Res.* **37** 3217–24
Ria F, Davis J T, Solomon J B, Wilson J M, Smith T B, Frush D P and Samei E 2019 Expanding the concept of diagnostic reference levels to noise and dose reference levels in CT *AJR. Am. J. Roentgenol.* **214** 1–6
Rizzi A, Algeri T, Medeghini G and Marini D 2004 A proposal for contrast measure in digital images *Conf. on Colour in Graphics, Imaging and Vision* vol 2004 (Society for Imaging Science and Technology) pp 187–92
Samei E *et al* 2019 Performance evaluation of computed tomography systems: summary of AAPM Task Group 233 *Med. Phys.* **46** e735–56
Sanders J, Hurwitz L and Samei E 2016 Patient-specific quantification of image quality: an automated method for measuring spatial resolution in clinical CT images *Med. Phys.* **43** 5330–8
Schopphoven S, Cavael P, Bock K, Fiebich M and Mäder U 2019 Breast phantoms for 2D digital mammography with realistic anatomical structures and attenuation characteristics based on clinical images using 3D printing *Phys. Med. Biol.* **64** 215005
Shahbahrami A, Pham T A and Bertels K 2012 Parallel implementation of Gray Level Co-occurrence Matrices and Haralick texture features on cell architecture *J. Supercomput.* **59** 1455–77
Shen J, Li Q and Erlebacher G 2009 Curvelet based no-reference objective image quality assessment *2009 Picture Coding Symp.* (IEEE) pp 1–4
Simone G, Pedersen M and Hardeberg J Y 2012 Measuring perceptual contrast in digital images *J. Vis. Commun. Image Represent.* **23** 491–506
Smith T B, Solomon J B and Samei E 2017 Estimating detectability index *in vivo*: development and validation of an automated methodology *J. Med. Imaging* **5** 031403
UNSCEAR 2022 UNSCEAR 2020/2021 Report: Sources, effects and risks of ionizing radiation: Annex a: evaluation of medical exposure to ionizing radiation *Technical Report* (United Nations Scientific Committee on the Effects of Atomic Radiation)
Vaishnav J Y, Jung W C, Popescu L M, Zeng R and Myers K J 2014 Objective assessment of image quality and dose reduction in CT iterative reconstruction *Med. Phys.* **41** 071904
Vu P V and Chandler D M 2012 A fast wavelet-based algorithm for global and local image sharpness estimation *IEEE Signal Process. Lett.* **19** 423–6
Yoder N 2023 peakfinder(x0, sel, thresh, extrema, includeEndpoints, interpolate), MATLAB Central file exchange (available at: www.mathworks.com/matlabcentral/fileexchange/25500-peakfinder-x0-sel-thresh-extrema-includeendpoints-interpolat) (Accessed 22 September 2023)
Zhu X and Milanfar P 2010 Automatic parameter selection for denoising algorithms using a no-reference measure of image content *IEEE Trans. Image Process.* **19** 3116–32