Establishment and validation of the Channelized Hotelling Model Observer for image assessment in industrial radiography

Sebastian Eckel\textsuperscript{a,}\textsuperscript{*}, Peter Huthwaite \textsuperscript{a}, Michael Lowe \textsuperscript{a}, Andreas Schummb, Pierre Guérin \textsuperscript{b}

\textsuperscript{a} Imperial College London, Department of Mechanical Engineering, London, United Kingdom
\textsuperscript{b} EDF S.A., R&D Department, Paris, France

\textbf{A R T I C L E   I N F O}

Keywords: Radiography Image processing Qualification Frequency domain

\textbf{A B S T R A C T}

A new method for industrial radiography is presented to assess image quality objectively. The assessment is performed by a modelled observer developed to interpret radiographic images in order to rate the detectability of structural defects. For the purpose of qualifying radiographic NDE procedures, computational tools simulate the image, but should additionally automatically assess the associated image quality instead of relying on human interpretation. The Channelized Hotelling Model Observer (CHO) approach, originally developed for medical imaging, is here developed for industrial NDE applications to measure objectively the defect's detectability. A validation study based on a comparison of the model's efficiency of observing circular and elongated flaws shows that the CHO outperforms other detectability models used by industry. Furthermore, the model's reliability was verified by comparing it to psychophysical data.

1. Introduction

X-ray imaging methods are used by industry, in a similar manner to the medical field, to investigate the inner structure of a specimen without causing harm by cutting. In the field of Non-Destructive Evaluation (NDE), conventional film based or digital X-ray inspection is applied for testing critical technical components across a number of industry sectors to identify structural defects like cracks. Hence, reliable optical image interpretation is essential.

Establishing the most appropriate radiographic setup for an inspection problem at hand can be very costly with regard to mockups and labour. Therefore, increasing the use of computational tools to simulate radiographic inspections is advisable. These tools help to find the test configuration which leads to the best possible image result. Moreover, reliable models allow the qualification of test procedures to be achieved more efficiently.

Radiography is based on the phenomenon of electromagnetic rays passing through the investigated object and attenuating as a function of the object's material properties. The remaining photonic energy having totally passed through the object induces a chemical reaction in the film exposed on the backside of the object. This reaction causes localised blackening of the film, producing an image linked to the actual physical conditions of the object.

The interpretation of the image is difficult because it contains not just a specific visual defect signal but also signals from the surrounding geometry, while scattering, noise etc. also influence the image. Therefore, the image needs interpretation by a human expert to identify damage signatures in the image. The human based evaluation step is crucial, because here it will be decided if there is a defect present or not. To improve the image quality, computer tools should be able to predict the defect's visibility for image assessment reasons, which can then be used to optimise the setup. This prediction completes the frame of a holistic simulation, besides modelling the test setup and simulating the image, and is also a prerequisite for Probability of Detection (POD) studies. In particular, the qualified expert's defect detection ability investigating an actual film based image should be anticipated by a computer model. So far, visibility criteria applied to perform that prediction task are available for industry [1–4], but these are only applicable with restrictions or under strong assumptions.

The objective of this work is to establish a new technique to measure the detectability more generally to predict better the defect's visibility. The presented method is fundamentally different to the methods currently used in NDE. This new method will improve the process of qualifying test procedures and allows the best radiographic test setup

\textsuperscript{*} Corresponding author. Imperial College London, South Kensington Campus, Department of Mechanical Engineering, Exhibition Road, London, SW7 2AZ, United Kingdom.

E-mail address: s.eckel16@imperial.ac.uk (S. Eckel).

https://doi.org/10.1016/j.ndteint.2018.04.001

Received 26 October 2017; Received in revised form 23 March 2018; Accepted 4 April 2018

Available online 5 April 2018

0963-8695/© 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
with regard to visibility to be found.

This article is divided into the following sections: First, a review of existing visibility criteria that are already used in NDE is given. Then a transfer of a Model Observer (MO) approach from the medical field to industrial NDE is provided, including its validation for industrial applications.

2. State of the art

The following paragraphs describe existing models to predict the human detectability of visual signals in industrial radiographic images.

2.1. Rose Model

The stochastic phenomenon of detecting photon quanta follows the Poisson distribution. The characteristic of the Poisson distribution is that its expected value and variance are equal. The accuracy of measuring the detected number of photons \( N \) depends on its standard deviation \( \sigma \): the lower \( \sigma \), the more likely is \( N \) close to the expected value. Or put another way, the lower \( \sigma \) is, the easier it is to distinguish smaller changes of \( N \), referred to as \( \Delta N \), the smallest distinguishable change. This leads to the proportionality

\[
\Delta N \propto \sigma. \tag{1}
\]

Making use of the denoted distribution characteristic and the definition of the standard deviation

\[\sigma = \sqrt{\text{Var}(N) = \sqrt{N}} \tag{2}\]

it can be stated that

\[\Delta N \propto \sqrt{N} \tag{3}\]

or by introducing a proportionality factor \( k \) and rearranging

\[
\frac{\Delta N}{\sqrt{N}} = k \tag{4}
\]

which is called the Rose Model representing a signal-to-noise ratio. Visual experiments showed that a good approximated value of \( k \) should be around 5 to assure visibility [5–7].

2.2. EDF criterion

The EDF visibility criterion generates a visibility map of the whole image. This is done by measuring the average contrast in an elliptical area (area: 1.6 mm\(^2\); eccentricity: 0.89) around every pixel of the image, while the ellipse is rotated stepwise. The maximum of the average contrast values in any position of the ellipse is taken as the visibility value associated to the pixel in the centre of the ellipse. The global visibility measure of the whole image is given by the maximum measure of all pixels. The visibility value is afterwards divided by the noise to get a contrast-to-noise ratio [1,2,8].

2.3. CEA criterion

The CEA criterion is based on the Rose Model but generalises it beyond circular features by also considering elongated shapes. It rests on the industrial standard [9] giving a reference that elongated shapes of a specific width should be distinguished by the same visibility as circular shapes of a specific diameter [5].

2.4. INSA criterion

The INSA criterion’s major advantage compared to the classical Rose Model is that it takes the gradient of optical density that surrounds an image of a defect into account. That gradient can degrade the visibility wherefore it is introduced as a penalty factor to decrease the gross result of Rose’s model [4].

2.5. CSF criterion

The CSF criterion is basing on the contrast sensitivity of the human eye [2]. The Contrast Sensitivity Function (CSF) describes how the human eye is sensible to a visual signal in dependence of its contrast and spatial frequency [10,11]. The image is 2D Fourier transformed and an averaged 1D spatial frequency spectrum is calculated. That generalised frequency spectrum is afterwards weighted by the CSF, then summed, normalised and divided by the noise level, in order to get a visibility measure [2].

3. Model Observers

Medical radiographic imaging has pressure to reduce radiation exposure to the patient, but also ensure that lesions are still detectable. Therefore models have been developed in the medical field to quantify the image quality in terms of signal detectability. This lends the previous work conveniently to our subject of NDE, where we have similar objectives.

The task of image quality definition has to be performed objectively. Image quality is defined as how well desired information can be extracted from an image [12], generally by an observer. Model Observers (MO) are mathematical constructs to imitate the human process of visual recognition. They are developed in comparison with psychological studies in order to imitate human perception. Image quality measurements can be done by quantifying the detection performance of a mathematical observer model treating the image under investigation. As a result, the image quality is measured objectively and the human performance can be derived from the MO’s performance [12].

The formation of an exemplary MO will be described in the following paragraphs.

3.1. Theory and definitions

The process of imaging an object can be described by

\[\mathcal{W}(f) + n = g\]  \tag{5}

where \( f \) denotes the object being imaged, \( \mathcal{W} \) the imaging function representing the imaging system, \( n \) the noise generated during the measurement and \( g \) the image [13]. For better convenience, the image, normally stored as a matrix of pixel values, shall be transformed into a vector by lexicographical indexing: the rows of the matrix are transposed and successively concatenated to form a vector containing all the information of the original image representation.

All MOs compute a scalar variable, called test statistic \( \lambda \). That test statistic is compared to a decision threshold \( \lambda_t \) to hypothesise whether the image belongs to a “signal-present” or “signal-absent” class [12,14,15]. Fig. 1 describes schematically the image production and decision process.

The key question is how the observer computes the test statistic. The work presented here is restricted to linear observers, which compute the
test statistic by creating a weighted total of the image’s intensities at all pixel points [14]:

$$\lambda = \sum_{m} w_{m} g_{m} = w \cdot g$$  \hspace{1cm} (6)$$

where $M$ denotes the total pixel number, $m$ is an index variable indicating a specific cell value of the vectors and $w$ is the weighting vector called “observer template” [14].

In a simplistic case the MO just uses an expected approximation of the signal as the observer template and takes the scalar product with the image under investigation [16]. The template is obtained by observer training, considering a set of images containing similar signals and another set of images without the signal. The difference between the means of the image sets is the template to which the observer was trained. Furthermore, the scalar product indicates how likely it is that the investigated image contains the signal: it is more likely if the resulting value becomes higher and vice versa. That simplistic MO can be summarised as

$$\lambda = (\overline{g} - \overline{g}_{0})^{T} g$$  \hspace{1cm} (7)$$

where $\overline{\text{\textit{g}}}$ denotes the mean value of a set (population) of images $\overline{\text{\textit{g}}}_{1}$ containing the signal and a set of images $\overline{\text{\textit{g}}}_{10}$ without the signal [16].

Furthermore, two statistical properties are of interest [14]. The first is the already mentioned image mean $\overline{\text{\textit{g}}}$ and the second is the covariance matrix $K$ defined as

$$K = \langle (g - \overline{\text{\textit{g}}}) (g - \overline{\text{\textit{g}}})^{T} \rangle$$ \hspace{1cm} (8)$$

where $(\cdot)^{T}$ denotes the transpose and the angle brackets denote the ensemble average. The covariance can be illustrated as a measure of intensity change between two pixels of the same image regarding all images in a population of images: it represents a measure of the correlation between any two pixels across a population of images. This information can be known a priori if available, else it needs to be estimated [14].

For illustration, the covariance between two pixels $x_{i}$ and $x_{j}$ denoted as $\text{Cov}(x_{i}, x_{j})$ is defined as

$$\text{Cov}(x_{i}, x_{j}) = \text{E}(x_{i} - \text{E}(x_{i})) \cdot (x_{j} - \text{E}(x_{j})) \cdot \frac{\sum_{k=1}^{N} (x_{i,k} - \overline{x}_{i}) (x_{j,k} - \overline{x}_{j})}{N}$$ \hspace{1cm} (9)$$

where $\text{E}(\cdot)$ describes the expectation value and $N$ is the number of samples in the population. The full covariance information of an image represented by an $M$ dimensional vector can be stored in an $M^{2}$ dimensional matrix.

After Bochud et al. [17], the statistical properties are completely known in the case of well-defined computer generated background. So the covariance is calculable instead of being estimated using a population of images. In the circumstances of the work presented here, the statistical properties of the simulated images are not known a priori: a Monte-Carlo based simulation process combined with a noise model representing the film graininess is used here, and is sufficiently sophisticated that a conclusion about the image’s covariance matrix needs to be derived from a population of images. The significant challenge is that an invertible estimation of the covariance matrix needs a population at least as big as the dimension of the covariance matrix to be estimated [18]. In the case of high-resolution images represented by vectors having the dimension of the number of pixels (typically millions) in the image, it would mean that a huge population of high-resolution images is necessary; even with modern computer capabilities this is not feasible.

3.2. Hotelling Observer

The Hotelling Observer is based on the Hotelling Trace Criterion [20]. That criterion measures the separability between two classes, e.g. signal-present/absent. If it is assumed in the binary problem case that the covariance matrices are similar for the signal present and absent case and that the image covariance matrix is regular, then the Hotelling Observer’s test statistic is given as [11,19].

$$\lambda_{\text{Hot}}(g) = s^{T} K^{-1} g$$ \hspace{1cm} (10)$$

with the signal $s = \overline{g}_{1} - \overline{g}_{0}$.

This kind of observer is called a Prewhitenning Matched Filter (PMF) because the application of $K^{-1}$ on $g$ leads to a decorrelation in $g$, and the product between $s$ and $g$ is called a “matched filtering” operation [12]. The difficulty here is to estimate the covariance matrix as mentioned before: the number of images in the population has to increase the pixel number thereby the covariance becomes invertible. A convergence study with a small $10 \times 10$ pixel image even showed that the detectability measure only converges when the population considerably exceeds the number of pixels. This is illustrated by Fig. 2: as the population exceeds the number of pixels, in that case 100, the MO’s output start to converge. The values for a population number less than 100 could only be calculated using the pseudo-inverse of $K$. To avoid that issue, it is possible to use a non-pre-whitening observer that does not need $K^{-1}$, but this would lead to a worse performance [12].

3.3. Channelized Hotelling Observer

A possibility to reduce the problem’s dimension is to advance to the Channelized Hotelling Observer (CHO) incorporating frequency or orientation selective channels [21]. The CHO can be built up in a more anthropomorphic way as frequency or orientation selectivity are a known characteristic of the human visual system [14]. The image is pre-processed by filtering through channels that are adjusted to specific special frequencies or orientations [15]. It is also possible to link the CHO with the human eye’s contrast-frequency sensitivity by application of channels tuned to the CSF. Thus, the frequency selective channels of number $N_{c}$ are applied to the image in the spatial-frequency domain either in a consecutive, non-overlapping [21–23] or in an overlapping manner [23]. The channels reduce the high dimensional image vector to a set of channel-responses by integrating the image information over the frequencies defined by the channel filter [14]. That integration is carried out by a scalar product between a filter vector defining the channel in the frequency domain, usually a band-pass filter, and the Fourier transformed image $\mathbf{g}$ [14]:

$$u = T^{T} \mathbf{g}$$  \hspace{1cm} (11)$$

![Fig. 2. Convergence study of the detectability measure in relation to the population size for a $10 \times 10$ pixel image including an interpolation.](image-url)
where $T$ is an $M \times N$ dimensional matrix representing the bank of channel filters in its columns and $u$ is the $N$ dimensional channel response. The problem's dimension will be reduced hugely if $N < M$, as would usually be the case [13,18].

If the Hotelling Observer is operating on the channelised image $u$, the test statistic of the CHO is given by [22]

$$
\lambda_{CHO}(g) = T^H K_{CHO} u
$$

where $K_{CHO}$ is the covariance matrix of the channel response and $i$ is the Fourier transformed signal $x$.

Various filter functions are used for the channel definition, such as constant functions (called constant-Q filters) [21] or Difference-Of-Gaussians profiles [23] that are defined to be rotationally symmetric in the spatial frequency domain.

The next paragraph describes the definition of the mentioned constant-Q filters.

### 3.3.1. Constant-Q filters

After Myers and Barrett [21], the constant-Q filters are defined as rectangular, nonoverlapping channels in the frequency domain. The matrix $T$ filters all spatial frequencies $\rho$ of the image under investigation. If $L$ describes the number of used filters and $m = 1, 2, \ldots, L$ then $T$ stacks the different channels that are defined as

$$
T_m(\rho) = \begin{cases} 
0 & \text{for } \rho \leq Q^{-1} \rho_c \\
1 & \text{for } Q^{-1} \rho_c < \rho < Q^* \rho_c \\
0 & \text{for } \rho \geq Q^* \rho_c
\end{cases}
$$

where the filter-width ratio $Q$ and the cut-off frequency $\rho_c$ are parameters defining the borders of the frequencies passing the filter $m$ [21].

Fig. 3 shows exemplarily the definition of the constant-Q channels in the frequency domain.

### 4. Results

To validate the CHO in NDE, it is necessary to show the ability of the CHO to predict the defect's visibility. The validation was performed using different kinds of references: an industrial standard describing the visibility in X-ray NDE and two psychophysical datasets. In this work, radiographic simulation and setup modelling are performed by the simulation package MODERATO [24].

#### 4.1. Validation by referring to standardised IQI visibility and comparison with existing visibility models

The first phase of validation is based on a comparison of the CHO result to an established industrial standard and comparison with other visibility models. Simulated images of different kinds of Image Quality Indicators (IQI) have been considered. An IQI is usually taped to the object and both are X-rayed together. An IQI set consists of indicators having different diameters and geometries such as wire or steps and holes. By assessing which indicator of a specific diameter is at the limit of visibility, the quality of the radiograph in terms of visibility can be defined. Both wire and hole IQIs of different diameters have been X-rayed by simulation. The basis for validation is given by an industrial standard [9] describing a relationship between the visibility of a hole IQI and a wire IQI. Specifically it provides a basis which allows one to state that a hole of a specific diameter provides the same visibility as a wire of another specific diameter.

The different hole and wire IQI diameters considered in the image simulation are summarised by Table 1. Any pair of hole and wire geometry listed in the same row of Table 1 should be characterised by the same visibility after [9]. It is worth mentioning that wires of different length have been used in the present work, namely 10, 25 and 50 mm, as the standard [9] does not specify the wire length.

To test the model, firstly, noise is added to the noiseless high-resolution images (6000 × 1200 pixels for the wire and 1200 × 1200 pixels for the hole) with a 16 bit grey scale. Noise creation is undertaken using the same procedure as described in Ref. [25] using Ref. [26], allowing to build up a population of images and to enable the covariance matrix to be estimated. Normally distributed noise is added to any pixel of the image with a standard deviation $\sigma_{\text{noise}}$ of

$$
\sigma_{\text{noise}} = \sigma_{\text{film}} \sqrt{\frac{\pi \cdot 10^3 \cdot D}{4 \cdot A}} \sqrt{\sum}
$$

where $A$ is the pixel size, $D$ is the optical density of the specific pixel and $\sigma_{\text{film}}$ is an experimental value obtained by measuring the noise of real films.

### Table 1

| Configuration number | Hole IQI diameter in mm | Wire IQI diameter in mm |
|----------------------|-------------------------|-------------------------|
| 1                    | 0.25                    | 0.1                     |
| 2                    | 0.32                    | 0.125                   |
| 3                    | 0.4                     | 0.16                    |
| 4                    | 0.5                     | 0.2                     |
| 5                    | 0.63                    | 0.25                    |
| 6                    | 0.8                     | 0.32                    |
| 7                    | 1.0                     | 0.4                     |
| 8                    | 1.25                    | 0.5                     |

Fig. 3. Example of four ($m = 1, 2, 3, 4$) constant-Q filters displayed in the spatial frequency domain with channel parameters set to $\rho_c = 0.015$ and $Q = 2$ [23].

Fig. 4. CHO result using constant-Q filters comparing the visibility of the hole IQI with the wire IQI of 50 mm length.
4.1. CHO using constant-Q filters

A routine has been written to calculate the test statistic by a CHO with constant-Q filters. Input parameters for these filters are the numbers of applied channels \( L \), the filter-width ratio \( Q \) and the cut-off frequency \( \rho_c \). It seemed reasonable to set \( Q = 2.0 \) as in Ref. [23]. The cut-off frequency was chosen to be \( \rho_c = 0.03 \) to also incorporate small spatial frequencies. Five channels have been used \((L = 5)\), which is a reasonable size comparable to Refs. [23] and [27]. An additional channel has been added that takes into account very low frequencies from \( \rho_c = 0 \) to \( \rho_c = 0.03 \). This supplementary channel is necessary because the highest content of information in the spatial frequency domain is located close to the centre \((\rho_c = 0)\) for all investigated images. The population size was set to \( N = 600 \) so that the results are reasonably converged and smooth. The process was run four times and the results averaged to improve the accuracy. The pixel size is set to \( A = 10\times10\) μm² and \( \sigma_{\text{film}} = 0.017 \) and the results are presented in Fig. 4.

The MO would be totally consistent with Ref. [9] if the two curves for wire and hole IQI visibility were matching for all investigated configurations. It is notable that the CHO result diverges for higher diameters, but up to configuration number 4 (hole and wire diameter of 0.8 and 0.32 mm respectively) the CHO results match well with the validation basis given by Ref. [9].

4.1.2. Comparison of the CHO model with existing visibility models

It is interesting to perform a comparison of the new CHO visibility model with those already used by industry. The visibility results of these models, namely the EDF, CEA, INSA and CSF criterion are shown in Fig. 5 after application to the simulated radiographic images containing the hole IQI and wire IQI. Any of the four subplots shows how the visibility measure given by the indicated visibility model changes with increasing IQI dimensions. In the perfect case of total accordance of the visibility model with the validation basis [9], the resulting curves for the hole respectively wire IQI would match completely. The same parameters as for the application of the CHO model have been used, namely a pixel size of \( A = 10\times10\) μm² and the same film characteristic \( \sigma_{\text{film}} = 0.017 \). Over all the visibility models, the best results were achieved by investigating the wire IQI of 50 mm length, but the wire length only significantly influences the CSF criterion. All existing criteria result in an increasing divergence for configurations with larger IQI diameters compared to Ref. [9] as is also the case for the implemented CHO. But it can be stated that the divergence is clearly weaker in the case of the CHO when compared to the existing models. The divergence can be visualised by the percentage divergence between the visibility results of the wire and hole IQI in comparison to Ref. [9]. The integrated percentage divergence between all investigated indicators is traced in Fig. 6 for the case of the 25 mm wire IQI as well in Fig. 7 for the case of the 50 mm wire IQI. In these figures, a lower value means that the visibility of both IQIs matches in average better for all investigated configurations after Ref. [9]. It can be clearly stated that the CHO with constant-Q filters reflects better the comparability between wire and hole IQI than it is the case for all other existing models. Thus the CHO model delivers a better visibility measure than the existing models currently considered in industry. However, it should be highlighted that this validation assumes the correctness of Ref. [9].

4.2. Validation by dataset using elongated shaped visual signals

This validation is based on the psychophysical dataset described in Ref. [4]. A radiograph was produced consisting of a high number of visual defect signals. These signals were gained by performing an X-ray of 192 objects taped to a quadratic steel plate of 30 cm edge length and 30 mm thickness. The objects are of different longitudinal shape (diameter of 0.05 to 3.2 mm and length from 5 to 20 mm) and materials (copper, iron, lead and tungsten). The objects’ attenuation was between 0.7% and 7% of the steel plate’s background attenuation and the objects’ position was randomly chosen on the source side with an orientation in respect to the steel plate of 0°, 45° and 90°.

The radiograph was evaluated by three trained experts (level 2 according to EN 473) ranking the visibility of every object. They were allowed to rank an object in an integer range of 0 to 4, respectively as invisible, very low, low, rather high or highly visible and their rankings were run four times and the results averaged to improve the accuracy. The pixel size is set to \( A = 10\times10\) μm² and \( \sigma_{\text{film}} = 0.017 \). Over all the investigated configurations after Ref. [9], the resulting curves for the hole IQI with the wire IQI of 50 mm length.

![Fig. 5. EDF, CEA, INSA and CSF criterion results comparing the visibility of the hole IQI with the wire IQI of 50 mm length.](image1)

![Fig. 6. Integrated percentage divergence between visibility models and Ref. [9] of all configurations (wire IQI of 25 mm).](image2)

![Fig. 7. Integrated percentage divergence between visibility models and Ref. [9] of all configurations (wire IQI of 50 mm).](image3)
were averaged for every object. Radiographic images of all objects were separately simulated with a pixel size of $A = 10-10 \, \mu m^2$ using the same setup. The film characteristic was set to $\sigma_{film} = 0.017$. Afterwards, the simulated images were processed by the MO. The CHO parameters were unchanged to the description above. Only the population size was decreased to $N = 300$ because of the big dataset. A drawback of this decreased population size is the high standard deviation of 21.6% when averaged over 4 calculations. The results are presented in Fig. 8, plotting the CHO's results against the human assessments. Each point represents one object's visibility. The scattering of the measurements is caused by the human based nature of the experiment: visibility is a subjective measure leading to scattered results when averaged over a limited number of human assessors.

A good correlation between the human and the simulated visibility assessment by the CHO is obvious. Additionally, an envelope around the result is remarkable: an important result showing that the MO is able to clearly separate between very low and highly visible. Another very good result is that the MO did not lead to false negative results, which would be located in the lower-right part of the graph. False negative results mean that the MO was not able to warn of an actually well-visible and so important defect. However, that area of the graph is empty, which is very promising.

### 4.3. Validation by dataset using circular shaped visual signals

Furthermore, circular shaped visual signals were investigated and the MO's results were compared to a human assessment of simulated images displayed on a computer screen. In that case, the three assessors were untrained but familiar with industrial radiography. While it would be ideal to utilise trained experts for this and match the data used before (taken from previous work [4]), unfortunately such a panel was unavailable for this study. We would still expect a good correlation between the human perception has been applied to the industrial environment and successfully validated. The results presented above prove that the CHO model is more appropriate than the existing models to predict human defect visibility in industrial applications. Thus the outcome of this research will hugely improve simulations carried out in NDE to get an objective measure of image quality in regard of human performance. In addition, the gained information, such as the visibility comparison of circular and elongated flaws, will be very useful to further improve existing industrial standards as Refs. [28,29].

A drawback is that the CHO results occur with a very high standard deviation (from 5% up to 28%) that will only decrease if the population size $N$ is increased further. This can be explained by the significant noise, whereas less noisy images lead to more stable results.

### 5. Conclusion

It is the very first time known that the CHO as an established model of the human perception has been applied to the industrial environment and successfully validated. The results presented above prove that the CHO model is more appropriate than the existing models to predict human defect visibility in industrial applications. Thus the outcome of this research will hugely improve simulations carried out in NDE to get an objective measure of image quality in regard of human performance. In addition, the gained information, such as the visibility comparison of circular and elongated flaws, will be very useful to further improve existing industrial standards as Refs. [28,29].

A drawback is that the CHO results occur with a very high standard deviation (from 5% up to 28%) that will only decrease if the population size $N$ is increased further. This can be explained by the significant noise, whereas less noisy images lead to more stable results.

### Funding

Sebastian Eckel is funded by the UK Engineering and Physical Sciences Research Council (EPSRC) ICASE number 15220134, with contributions from EDF S.A. R&D. Peter Huthwaite is funded by an EPSRC Early Career Fellowship EP/M020207/1.

### Acknowledgment

The authors thank Mr Sebastian Heinlein (Imperial College London) for helpful discussions to improve the model implementation.

### References

[1] Besnard E. Etablissement de critères de détectabilité de défauts en radiographie. EDF R&D Internal Report; 2003.

[2] Jonchière R, Guérin P. Implémentation d’indicateurs de visibilité améliorés dans le code MODERATO. EDF R&D Internal Report; 2016.

[3] Tisseur D, Vienne C, Guérin P, Parmentier AP, Kaftandjian V, Duvachelle P, et al. A modified detectability criterion for conventional radiography simulation. In: 19th world conf. Non-Destructive Test, Munich; 2016.

[4] Kaftandjian V, Duvachelle P, Tisseur D, Vienne C, Guérin P, Schuman A. Comparison of human visibility on film radiography and various detectability parameters computed on simulated radiographs – a statistical study. In: 19th World Conf. Non-Destructive Test, Munich; 2016.

[5] KTH Royal Institute of Technology. Photon counting, SNR and the Rose model Theoretical background n.d.:1-11. http://medim.kth.se/h/2009/Photoncounting-Rose_model.pdf. [Accessed 4 October 2016].
[6] Rose A. The sensitivity performance of the human eye on an absolute scale. J Opt Soc Am 1948;38:196–208. https://doi.org/10.1364/JOSA.38.000196.

[7] Rose A. Television pickup tubes and the problem of vision. Adv Electron Electron Phys 1948:131–66.

[8] Guérin P, Schumm A. Validation du code de simulation de contrôle radiographique MODERATO Version 4.0. EDF R&D Internal Report; 2014.

[9] Design and construction rules for mechanical components of PWR nuclear islands RCC-M, AFCEN code. 2007.

[10] Ginsburg AP. Proposed new vision standards for the 1980’s and beyond: contrast sensitivity. Report by Air Force Aerospace Medical Research Lab; 1981.

[11] Campbell BYFW, Robson JG. Application of fourier analysis to the visibility of gratings. J Physiol 1968;197:551–66.

[12] Barrett HH, Yao JIE, Rollandt JP, Myers J. Ideal observer models of visual signal detection. Handb Med Imaging 2000;1:559–92.

[13] Myers KJ, Barrett HH, Smith WE, Myers J. Hotelling trace criterion and its correlation with human-observer performance. J Opt Soc Am A 1987;4:945–53. https://doi.org/10.1364/josaa.4.000945.

[14] Myers KJ, Barrett HH. Addition of a channel mechanism to the ideal-observer model. J Opt Soc Am A 1987;4:2447–57. https://doi.org/10.1364/JOSAA.4.002447.

[15] Abbey CK, Bochud F. Human- and model-observer performance in ramp-spectrum noise: effects of regularization and object variability. J Opt Soc Am A 2001;18: 473–86.

[16] De las Heras Gala H. Imaging in nuclear medicine. In: Giussani A, Hoeschen C, editors. Imaging Nucl. Med., vol. 9783642314; 2013. p. 167–86. https://doi.org/10.1007/978-3-642-31415-5.

[17] Bochud O, Abbey CK, Eckstein MP. Visual signal detection in structured backgrounds. III. Calculation of figures of merit for model observers in statistically nonstationary backgrounds. J Opt Soc Am A 2000;17:193–205.

[18] Ullman G. Quantifying image quality in diagnostic radiology using simulation of the imaging system and model observers. Doctoral Thesis Linköping University; 2008.

[19] Myers KJ. Ideal observer models of visual signal detection. Handb Med Imaging 2000;1:559–92.

[20] Firte BD, Barrett HH, Smith WE, Myers KJ. Hotelling trace criterion and its correlation with human-observer performance. J Opt Soc Am A 1987;4:945–53. https://doi.org/10.1364/josaa.4.000945.

[21] Myers KJ, Barrett HH. Addition of a channel mechanism to the ideal-observer model. J Opt Soc Am A 1987;4:2447–57. https://doi.org/10.1364/JOSAA.4.002447.

[22] Yao J, Barrett HH. Predicting human performance by a Channelized Hotelling Observer Model. Vol. 1768. San Diego: SPIE; 1992. p. 161–8.

[23] Abbey CK, Barrett HH. Human- and model-observer performance in ramp-spectrum noise: effects of regularization and object variability. J Opt Soc Am A 2001;18: 473–86.

[24] Bonin A, Lavayssière B, Chalmond B. Moderato: a monte-Carlo radiographic simulation. In: AIP Conf. Proc. Rev. Prog. Quant. Nondestruct. Evalvol. 509; 2000. https://doi.org/10.1063/1.1306111. Ames, IA, USA.

[25] Guérin P. Validation de la prise en compte de la granularité des films dans MODERATO. Paris, France: EDF R&D Internal Report, unpublished; 2015.

[26] European Committee for Standardization. EN 584-1:2006 F - Film pour radiographie industrielle Partie 1: Classification des systèmes films pour radiographie industrielle. AFNOR; 2006.

[27] ASTM International. E747 – 04 (reapproved 2010) standard practice for design, manufacture, and material grouping classification of hole-type image quality indicators (IQI ) used for radiology. Annu Book ASTM Stand 2010:1–7. https://doi.org/10.1520/E0747-04R10.

[28] CEN. BS EN ISO 19232-3:2013 Non-destructive testing — image quality of radiographs Part 3: image quality classes. BSI Stand Publ; 2013.