IEEE COPYRIGHT AND CONSENT FORM

To ensure uniformity of treatment among all contributors, other forms may not be substituted for this form, nor may any wording of the form be changed. This form is intended for original material submitted to the IEEE and must accompany any such material in order to be published by the IEEE. Please read the form carefully and keep a copy for your files.

Author Online Use

6. Personal Servers. Authors and/or their employers shall have the right to post the accepted version of IEEE-copyrighted articles on their own personal servers or the servers of their institutions or employers without permission from IEEE, provided that the posted version includes a prominently displayed IEEE copyright notice and, when published, a full citation to the original IEEE publication, including a link to the article abstract in IEEE Xplore. Authors shall not post the final, published versions of their papers.

7. Classroom or Internal Training Use. An author is expressly permitted to post any portion of the accepted version of his/her own IEEE-copyrighted articles on the authors personal web site or the servers of the authors institution or company in connection with the authors teaching, training, or work responsibilities, provided that the appropriate copyright, credit, and reuse notices appear prominently with the posted material. Examples of permitted uses are lecture materials, course packs, e-reserves, conference presentations, or in-house training courses.
A NEW TOMOGRAPHY MODEL FOR ALMOST OPTIMAL DETECTION OF ANOMALIES

Rémi Cogranne, Florent Retraint

ICD - LM2S - Université de Technologie de Troyes (UTT) - UMR STMR CNRS,
12 rue Marie Curie - CS 42060 - 10004 Troyes cedex - France.
E-mail : \{remi.cogranne, florent.retraint\}@utt.fr

ABSTRACT

In this paper a new methodology for detecting anomaly from few tomography projections is presented. This methodology exploits a statistical model adapted to the content of radiographs together with hypothesis testing theory. The main contributions are the following. First, using a generic model of the tomography acquisition pipeline, the whole non-destructive testing process is entirely automated. Second, by using testing theory the statistical properties of the proposed test are analytically established. This particularly permits the guaranteeing of a prescribed false-alarm probability and allows us to show that the proposed test is almost optimal. Experimental results show the sharpness of the established results and the relevance of the methodology.

Index Terms— Hypothesis testing theory, Parametric model, Background subtraction, Optimal detection, Non-destructive testing.

1. INTRODUCTION

During the past decades X-ray or gamma tomography inspection have been widely used in non-destructive testing processes. Usually, the detection process based on tomographic projections rely on manual interpretation of acquired images by human experts. Hence, there is a wide need for automatic detection methods. However, depending on the inspected object internal geometry the detection may be difficult due to the non-anomalous background.

The prior methods for detection of internal defects using tomography can be divided into three categories [1, 2]: 1) The category of generic methods that do not require any priori knowledge includes image processing enhancement tools (field flattening, contrast enhancement, etc.) [3], pattern recognition [4] and supervised machine learning [5, 6]. 2) Some methods require a ground truth or a reference image used as a model [1]. A sufficient difference between this ground truth and the inspected tomography projections lead to classify the inspected object as defective [7]. 3) Finally, most of the Computerized Tomography (CT) methods rely on the reconstruction of the inspected object from its tomographic projections. This reconstruction is usually not necessary to only perform a detection and two main approaches have been proposed to this end: Bayesian and non-Bayesian approaches. The methodology proposed in the present paper belongs to the non-Bayesian CT approach.

In fact, generic approaches are sensitive to the noise as well as object and anomaly geometry. Similarly, the use of a ground truth is only possible when all the objects are exactly identical and the whole inspection process is calibrated precisely. On the contrary, Bayesian methods are very efficient when apriori knowledges on both the inspected object and the potential anomaly are available which is not always possible. In such a situation, another approach is to model the expected non-anomalous background of a tomography to perform a so-called background subtraction, see [6, 8] for instance.

In this paper, an original model of tomography projections is proposed based on a study of the whole tomography acquisition process [9]. This model allows the subtraction of background which is necessary to perform a statistical test to detect anomalies. The main contribution are the following: 1) By modelling the whole acquisition process, the proposed model can be applied to a wide range of inspected object without prior knowledge. 2) Thanks to an original linearisation of the proposed model, the whole detection process is computationally efficient. 3) The statistical properties of the proposed test are established; this allows the guaranteeing of a given false-alarm probability and also proves the almost optimality of the test, i.e. the loss of performance compared to the optimal statistical test is bounded. The relevance of theoretical findings are emphasized through numerical experimentations.

The paper is organized as follows. Section 2 states the problem of anomaly detection and presents the theoretical optimal test. Section 3 presents the proposed model of tomography used in Section 4 to design a statistical test with established performance. Section 5 presents numerical results and Section 6 concludes the paper.

Research partially funded by Troyes University of Technology (UTT) strategic program COLUMBO.
2. STATISTICAL DETECTION OF ANOMALIES AND OPTIMAL TEST

Let $Z^{(k)} = \{z^{(k)}_{m,n}\}, (m, n) \in \{1, \ldots, M\} \times \{1, \ldots, N\}, k \in \{1, \ldots, K\}$ be a set of $K$ tomographic projections represented as images of $M \times N$ pixels. For the sake of clarity the projection index is omitted in this section because they are processed individually. In this paper, tomography projection are split into $J$ non-overlapping sub-blocks of $J$ pixels. Without loss of generality, in this paper sub-blocks are vectors denoted $z_i = (z_{i,1}, z_{i,2}, \ldots, z_{i,J})^T$. When the inspected object is free from anomaly, each vector $z_i$ can be decomposed as [10, 11, 12]:

$$z_i = \mu_i + \xi_i \Rightarrow z_i \sim N(\mu_i, \sigma_i^2I_J) \quad (1)$$

where $\mu_i = (\mu_{1,i}, \ldots, \mu_{J,i})^T$ is the deterministic expectation of vector $z_i$, $\sigma_i^2$ is the $i$-th vector variance, assumed constant over the $J$ pixels, and $I_J$ is the identity matrix of size $J \times J$. Note that in practice, model (1) can be obtained for any covariance matrix $\Sigma$ by applying the change of variable $g(z_i) = \Sigma^{-1/2}z_i$, where $\Sigma^{-1} = \Sigma^{-1/2}\Sigma^{-1/2}$.

In the present paper the expectation of each vector is modelled as a linear combination of basis functions as follows:

$$\mu_i = Hx_i. \quad (2)$$

Here, $H$ is known full-column rank matrix of size $J \times p$ and $x_i \in \mathbb{R}^p$ is a set of $p < J$ nuisance parameters defining $\mu_i$. Similarly, when an anomaly is present in the inspected object, the value of each $z_i$ can be decomposed as:

$$z_i \sim N(\mu_i + a_i, \sigma_i^2I_J) \quad (3)$$

where $a_i = (a_{1,i}, \ldots, a_{J,i})^T$ is the vector representing the anomaly on $i$-th vector. It follows from Equations (1)-(2) that for any covariance matrix $\Sigma$ by applying the change of variable $g(z_i) = \Sigma^{-1/2}z_i$, where $\Sigma^{-1} = \Sigma^{-1/2}\Sigma^{-1/2}$.

In order to design a statistical test with known performance, it is impossible to model vectors $a_i$.

For testing problem such as (4), Wald [14] proposed the Uniformly Best Cost Constant (UBPC) as a criterion of optimality because the power of a test depends on both the non-anomalous background and the anomaly. To present the optimal UBPC test, let us define $R(H) \subseteq \mathbb{R}^p$ the column space spanned by $H$ and $R(H)^\perp \subseteq \mathbb{R}^{J-p}$ its orthogonal complement, sometimes referred to as the “parity space” [15]. The projection onto the parity space is obtained by $\tilde{z}_i = Wz_i$ where $W$ spanned the null space of $H$ and besides verifies, the following properties: $WH = 0$, $WW^T = I_{J-p}$. Hence, by projecting the tomographic observation $z_i$ onto the parity subspace $R(H)^\perp$ yields: $\tilde{z}_i = Wz_i = W\xi_i + Wa_i$, which shows how algebraic subtraction of background or nuisance parameters is easily possible using the parity subspace.

It is shown in [15] that to solve testing problems such as (4) the test defined by:

$$\delta^*(Z) = \begin{cases} H_0 & \text{if } \Lambda^*(Z) = \sum_{i=1}^{I} \sigma_i^{-2}\|Wz_i\|_2^2 < \tau_{\alpha_0}^* \\ H_1 & \text{if } \Lambda^*(Z) = \sum_{i=1}^{I} \sigma_i^{-2}\|Wz_i\|_2^2 \geq \tau_{\alpha_0}^* \end{cases} \quad (7)$$

is UBPC over the class $\mathcal{K}_{\alpha_0}$ provided that the threshold $\tau_{\alpha_0}^*$ is the solution of equation $\mathbb{P}_{H_0}[\Lambda^*(Z) \geq \tau_{\alpha_0}^*] = \alpha_0$. Denoting $g^2 = \sum_{i=1}^{I} \sigma_i^{-2}\|Wz_i\|_2^2$, and the number of “degree of freedom” $I = I(J-p)$, the properties of Gaussian random variables allows the establishing of UBPC test properties in the following Theorem 1.

**Theorem 1.** For any $\alpha_0 \in ]0; 1[$ the decision threshold:

$$\tau_{\alpha_0}^* = F_{\chi^2_Y}^{-1}\left(1 - \alpha_0, 0\right) \quad (8)$$

guarantees that $\mathbb{P}_{H_0}[\Lambda^*(Z) \geq \tau_{\alpha_0}^*] = \alpha_0$ so that the UBPC test $\delta^*$ (7) is in the class $\mathcal{K}_{\alpha_0}$. Here, $F_{\chi^2_Y}$, and $F_{\chi^2_Y}^{-1}$, represents the non-central chi-squared cumulative distribution function and its inverse.

Choosing the threshold $\tau_{\alpha_0}^*$ as defined in (8), the power function associated with the UBPC test $\delta^*$ (7) is given by:

$$\beta^*(g, \alpha_0) = 1 - F_{\chi^2_Y}\left(\tau_{\alpha_0}^*, g^2\right) \quad (9)$$

**Proof.** of Theorem 1 is omitted due to lack of space.

Theorem 1 emphasizes that, provided that the model $H$ (2) is correctly chosen, the decision threshold only depends on $\alpha_0$ and $Y$. This allows the guaranteeing of a give
false-alarm probability whatever the inspected object might be. Moreover, Theorem 1 also provides an upper bound on the detection power one can expect from any test that aims at detecting an anomaly.

3. STATISTICAL MODEL OF CT IMAGES

In practice, the hardest difficulty is to design a model $H$ for the non-anomalous background; to this end the whole acquisition process is modelled. An object is characterized by its absorption function $a(x, y, z, \nu)$ where $\nu$ is the photons energy. Without loss of generality, the sensor lies at plane $z = 0$ and the photons source is at $z = z_M$. It follows from Beer-Lambert law [3], that the “ideal” number of photons $\omega(x, y)$ passing through the inspected object and reaching the sensor at location $(x, y)$ is given by [9]:

$$\omega(x, y) = N_0 \int_{R^2} \phi(\nu) \exp \left( -\int_0^{z_M} a(l(z), z, \nu) dz \right) d\nu.$$  

where $N_0$ represents the mean number of photons emitted by the source, $\phi(\nu)$ represents the source energy spectrum, and $l : \mathbb{R} \to \mathbb{R}^2$ represents the path of photons from the source to the sensor.

In practice, the mean number of incident photons $\mu(x, y)$ widely differs from $\omega(x, y)$ due to physical blurring phenomena, such as scattering, diffraction, etc. Taking into account these deterministic phenomena, the mean number of incident photons is given by [9]:

$$\mu(x, y) = \iint_{R^2} \omega(x, y) h_{(x,y)}(u-x, v-y) dxdy$$  

where (10) represents a two dimensional convolution with kernel $h_{(x,y)}$ at location $(x, y)$. In the present paper, the convolution kernel is assumed to be locally modelled by a Gaussian blur:

$$h_{(x,y)} = \frac{1}{2\pi \sigma_{\text{psf}}} \exp \left( -\frac{x^2 + y^2}{2\sigma_{\text{psf}}^2} \right)$$

where $\sigma_{\text{psf}} = \sigma_{\text{psf}}(x, y) > 0$ is the local PSF parameter which represents non the non-stationary nature of the scattering process over the whole image plane (due to aberrations, materials properties, etc...).

The model of the inspected object absorption function used in this paper is inspired from [16]. Roughly speaking it is based on the idea that the inspected object is composed of different materials separated by abrupt discontinuities. From the physical properties of light, one can expect that the following properties hold. 1) The absorption function varies smoothly over areas defined by each material. 2) The absorption function $a$ is discontinuous across the boundary of these domains. It follows from these properties that the “theoretical” intensity reaching $i$-th vector is a piecewise continuous function and hence, can be written: $\omega_i(x) = \omega_i^{(c)}(x) + \omega_i^{(s)}(x)$ where $\omega_i^{(c)}$ is continuous and $\omega_i^{(s)}$ is the singular part. The function $\omega_i^{(s)}$ is piecewise constant and hence can be written:

$$\omega_i^{(s)}(x) = \sum_{d=1}^{r_i} u_{i,d} 1_{R^+}(x - t_{i,d})$$

with $r_i$ the number of discontinuities in $i$-th vector $z_i$ and $1_{R^+}$ the indicator function of the set $\mathbb{R}^+$. The parameters $u_{i,d}$ and $t_{i,d}$ characterize, respectively, the intensity and the location of the $d$-th discontinuity in the $i$-th vector $z_i$. From the linearity of Equation (10), it follows that the measured (blurred or scattered) mean number of photons reaching the sensor over the $i$-th vector $z_i$ is given by:

$$\mu_i(x) = \mu_i^{(c)}(x) + \mu_i^{(s)}(x)$$

where $\mu_i^{(c)}$ and $\mu_i^{(s)}$ respectively correspond to the convolved continuous and singular part of “ideal” unaltered flow $\omega_i$.

Denoting $\Phi(u) = \int_{-\infty}^{u} (2\pi)^{-1/2} \exp (-x^2/2) dx$ the Gaussian integral function, the function $\mu_i^{(s)}$ can be written by [17, 18]:

$$\mu_i^{(s)}(x) = \sum_{d=1}^{r_i} u_{i,d} \Phi \left( \frac{x - t_{i,d}}{\varsigma_{i,d}} \right)$$

with $\varsigma_{i,d}$ the local blurring parameter at $d$-th discontinuity location. On the contrary, the function $\mu_i^{(c)}$ results from the convolution between $\omega_i^{(c)}$ and $h_{(x,y)}$ and, hence, is assumed to be very smooth or low frequency. Consequently, it can be accurately modelled by an algebraic polynomial of degree $p - 1$ [17, 19, 20]. Using the piecewise polynomial model of the continuous part $\mu_i^{(c)}$ and the Equation (14) to model $\mu_i^{(s)}$, it follows from Equation (13) that:

$$\mu_i(x) = \sum_{q=0}^{p-1} s_{i,q} x^q + \sum_{d=0}^{r_i} u_{i,d} \Phi(x; \eta_{i,d})$$

where $s_{i,q}$ are the polynomial coefficients representing the continuous part, $\eta_{i,d} = (t_{i,d}, \varsigma_{i,d})$ represents $d$-th discontinuity parameter and $\Phi(x; \eta_{i,d}) = \Phi(x - t_{i,d}/\varsigma_{i,d})$.

After integration over each sensor cell, it follows from (15) that the expectation of pixels on the $i$-th vectore, denoted, $\mu_i$ is given by:

$$\mu_i = Hs_i + F(\eta_i)u_i = G(\eta_i)v_i.$$  

Here, $H$ is the matrix of size $L \times p$ whose element $(i,j)$ is $x_{i,j}^{p-1}$ and $F(\eta_i)$ is the matrix of size $L \times r_i$ whose element $(i,d)$ is $\Phi(x_i; \eta_{i,d}) = \Phi(x_i - t_{i,d}/\varsigma_{i,d})$. It is obvious that $G(\eta_i) = (H|F(\eta_i))$ and $v_i = (s_i, u_i)^T$. 

4. ALMOST OPTIMAL DETECTION OF ANOMALY

Using the proposed locally adapted model (16) to estimate the expectation $\mu_i$ from the observations $x_i$ is difficult because the discontinuities are non-linear with respect to parameters $\eta_{i,d} = (t_{i,d}, \varsigma_{i,d})$. To tackle this difficulty, it is proposed in this paper to follow the general methodology proposed in [21]; this approach is based on a Taylor series expansion of the non-linearity which allows the writing of:

$$\mu_i = Hx_i + u_i F(\hat{\eta}_i) + u_i \hat{\Phi}(\hat{\eta}_i) (\eta_i - \hat{\eta}_i) + o(\|\eta_i - \hat{\eta}_i\|_1)$$

where $\hat{\Phi}(\hat{\eta}_i)$ is the $N \times 2r_i$ Jacobian matrix of $F(\hat{\eta}_i)$ and $\hat{\eta}_i$ an estimation of discontinuity parameter $\eta_i$. This yields the linear model:

$$\mu_i = \hat{G}(\hat{\eta}_i) v_i + o(\|\eta_i - \hat{\eta}_i\|_1),$$

with $\hat{G}(\hat{\eta}_i) = \left( H \mid F(\hat{\eta}_i) \mid \hat{\Phi}(\hat{\eta}_i) \right)$ and $v_i = \left( x_i \quad u_i (\eta_i - \hat{\eta}_i) \right)$

Using the linearisation approach (17) allows the design of the proposed linearised GLRT:

$$\hat{\delta}(Z) = \begin{cases} H_0 & \text{if } \hat{\Lambda}(Z) = \sum_{i=1}^{I} \hat{\sigma}_i^{-2} \| W_{\hat{G}(\hat{\eta}_i)} Z_i \|_2^2 < \hat{\tau}_{\alpha_0} \\ H_1 & \text{if } \hat{\Lambda}(Z) = \sum_{i=1}^{I} \hat{\sigma}_i^{-2} \| W_{\hat{G}(\hat{\eta}_i)} Z_i \|_2^2 \geq \hat{\tau}_{\alpha_0} \end{cases}$$

with $\hat{\sigma}_i^2$ the estimation of variance obtained from [12, 22] in this paper.

To establish the properties of the GLR test $\hat{\delta}$, Equation (18), the main difficulty is to analyse the error term $o(\|\eta_i - \hat{\eta}_i\|_1)$, in Equation (17), which is due to linearisation around estimation $\eta_i \neq \hat{\eta}_i$. For the sake of clarity, it is assumed that one discontinuity (at most) is present in each vector $x_i$; the extension to the case of multiple discontinuities is straightforward at the cost of complicated notations. Besides, it is assumed that the error on non-linear parameter estimation is bounded by a constant $\|\eta_i - \hat{\eta}_i\|_1 \leq \zeta$. The literature proposes methods which give such estimates [23, 24]; in this paper the one proposed in [23] is used. Assuming that the estimation error $\|\eta_i - \hat{\eta}_i\|_1 \leq \zeta$ is bounded by $\zeta$, permits the writing of:

$$\sum_{i=1}^{I} \frac{1}{\hat{\sigma}_i^2} \| W_{\hat{G}(\hat{\eta}_i)} Z_i \|_2^2 \leq \sum_{i=1}^{I} \frac{u_i^2}{4\pi^2} \zeta^2 = b_{\text{max}}$$

Denoting, $\hat{\theta}^2 = \sum_{i=1}^{I} \hat{\sigma}_i^{-2} \| W_{\hat{G}(\hat{\eta}_i)} Z_i \|_2^2$ the estimated “Anomaly-to-Noise Ratio” (ANR), Eq. (19) and a short algebra allow the establishing of proposed GLR $\hat{\delta}$ (18) properties in Theorem 1.

**Theorem 2.** Assuming that the model (16) holds and that $\hat{\eta}_i$ is an unbiased estimator of $\eta_i$, satisfying $\|\eta_i - \hat{\eta}_i\|_1 \leq \zeta$, then for any $\alpha_0 \in [0; 1]$ the decision threshold:

$$\hat{\tau}_{\alpha_0} = F_{\chi^2_k}^{-1} \left( 1 - \alpha_0; b_{\text{max}} \right)$$

guarantees that $P_{\hat{\delta}} \left[ \hat{\Lambda}(Z) \geq \hat{\tau}_{\alpha_0} \right] \leq \alpha_0$ so that the GLR test $\hat{\delta}$ (18) is in the class $K_{\alpha_0}$.

Choosing the threshold $\hat{\tau}_{\alpha_0}$ as defined in (20), the power function associated with the GLR test $\hat{\delta}$ (18) is bounded by:

$$1 - F_{\chi^2_k} \left( \hat{\tau}_{\alpha_0}; \hat{\theta}^2 + b_{\text{max}} \right) \leq \hat{\beta}(\hat{\tau}_{\alpha_0}^2; \alpha_0) \leq 1 - F_{\chi^2_k} \left( \hat{\tau}_{\alpha_0}; \hat{\theta}^2 \right).$$

**Proof.** of Theorem 2 is omitted due to the lack of space. $\Box$

Theorem 2 shows that the loss of power of the proposed GLR test $\hat{\delta}$, compared to the optimal UBPC test $\delta^*$, is due to the two following reasons. First, the matrix $G(\eta_i)$ has less columns than the one which is used in practice $\hat{G}(\hat{\eta}_i)$, the smaller number of “degree of freedom” degrades the quality of the estimate. Second, the linearisation (17) is not perfect. There is an error of approximation due to estimation error $\|\eta_i - \hat{\eta}_i\|_1 \leq \zeta$ which cause a bias bounded by $b_{\text{max}}$ (19) in the calculation of GLR $\hat{\Lambda}$. Fortunately, provided that the number of discontinuities $r$ and error on estimation of discontinuity parameters $\zeta$ are sufficiently small, the GLR test $\hat{\delta}$ performs almost as well as the optimal UBPC test $\delta^*$. 

---

**Comment:** For a clearer understanding, the Linearisation Sections (17) and (18) should be highlighted.
5. NUMERICAL RESULTS

To verify numerically the relevance and the sharpness of the theoretical results, it is proposed to perform a Monte-Carlo simulation based on acquired data. The vectors used for this experimentation are shown by the blue line in Figure 3a. This line, modelled by two vectors of 75 pixels described by a polynomial of order \( p - 1 = 3 \), is analysed 50 000 times, after addition of a Gaussian noise. Two vectors, with and without anomaly, are used to obtain results under both hypotheses \( H_0 \) and \( H_1 \), see also data presented in Figure 2.

Figure 1 presents a comparison between empirical and theoretical results through a ROC curve; that is, the detection power as a function of false-alarm probability \( \alpha_0 \). The results presented in Figure 1 show the relevance of the proposed methodology and the sharpness of the theoretically established results.

To complete the previous numerical results, Figure 2 shows few data used for the previous Monte-Carlo simulation together with the results obtained from background subtraction. Figure 2 shows that the proposed methodology allows the rejection of nuisance parameter while preserving the presence of the anomaly.

The inspection of a nuclear fuel rod in presented through a tomography projections shown in Figure 3. Figure 3a shows an acquired tomography projection which contains almost invisible anomalies (slight voids). The residuals obtained from proposed background subtraction are shown in Figure 3b.

In order to show potential application of proposed methodology to a wide range of anomalies detection problem, Figure 4 presents results obtained from a medical image. Figure 4a shows the original longitudinal radiographic image of forearm focusing on the radius area which exhibits a slight fracture. Figure 4b shows the result obtained from proposed background subtraction. Note that each line is made of 248 pixels which are divided into four vectors of 62 pixels whose smooth part \( \mu_i(c) \) are modelled by a polynomial of degree \( p - 1 = 3 \). For comparison, Figure 4c shows the results obtained from wavelet background subtraction [24] which obviously do not separated anomaly from background making detection harder.

6. CONCLUSION

This paper proposes a new methodology for anomaly detection from a few noisy tomography projections. The main contribution are the following. First, the physical properties of tomographic system is studied in order to obtain a locally adapted model of the acquired projections. Then, this locally adapted model is used to subtract non-anomalous background accurately and efficiently. This finally allows the designing of a statistical test whose properties are analytically established in order to guarantee a prescribed false-alarm probability. Besides, the test is also shown to be almost optimal. Numerical results show the relevance of the proposed approach and the sharpness of theoretically established results.

7. REFERENCES


