DETECTABILITY OF ANOMALIES FROM A FEW NOISY TOMOGRAPHIC PROJECTIONS

Lionel Fillatre, Florent Retraint and Igor Nikiforov

ISTIT, FRE CNRS 2732, Université de Technologie de Troyes, BP 2060, 10010 Troyes, Cedex, France
E-mail : {fillatrl,retraint,nikiforov}@utt.fr

Abstract: The problem of detecting an anomaly from a limited number of noisy tomographic projections is addressed from the statistical point of view. An unknown two (or three)-dimensional scene is composed of a background, considered as a nuisance parameter, with a possibly hidden anomaly. A parametric approach is proposed to reduce the lack of *a priori* information and an optimal test is designed. To decide between two hypotheses (absence or presence of the anomaly), the statistical test eliminates the background, which can hide the anomaly. New results on anomaly detectability are proposed and discussed in this paper. It is shown that a size-limited anomaly is better detectable with several projections.

Keywords: computer tomography, parametric model, statistical tests, fault detection, incomplete data, invariance, signal detection.

1. INTRODUCTION

Computerized tomography (CT) is a technique for examining the internal structure of an object, which is very useful in quantitative non-destructive testing, object recognition and (biomedical) system monitoring, among others. In certain applications like baggage X-ray scanning or welding defects detection, it is desirable to detect an anomaly/target possibly hidden into an unknown deterministic background from a very limited number of noisy tomographic projections. A key issue of such a detection problem is to state the significance of the observed deviation due to the anomaly/target with respect to the unknown background considered as a nuisance parameter. An optimal statistical invariant test is proposed to detect an anomaly/target. Because the unknown background can hide the anomaly, a very special attention is paid to establish the problem of anomaly detectability. It appears that several projections can increase the detectability of a size-limited anomaly. The paper is organized as follows. The anomaly detection problem is stated in Section 2. Next, an optimal decision rule is designed for detecting the anomaly in Section 3. Finally, the detectability of anomalies is studied in Section 4.

2. PROBLEM STATEMENT

A more detailed description of the problem can be found in (Fillatre and Nikiforov, 2003).

2.1 Description of the imaging system

Let us assume the imaged object (or the original scene) $s(x, y)$ is a two-dimensional (2D) real function in the x-y coordinate system (the extension to the three-dimensional case is straightforward), where $s(x, y)$ has a compact support $\mathcal{D} \subset \mathbb{R}^2$. This situation is depicted in Fig. 1. In CT, the function

s represents the X-ray attenuation coefficient of the studied object, i.e. its physical property to absorb an X-ray flux. It is assumed that the object s is a real-valued square-integrable function : $s \in \mathcal{L}_2(\mathcal{D})$. The line integral of s along the ray $L(t,\omega)$: $x\cos\omega + y\sin\omega = t$, with $-1 \le t \le 1$ and $0 \le \omega \le \pi$, is the 2D function, denoted as $P_s(t,\omega)$, called the Radon transform of s (Natterer, 1986), given by :

$$P_s(t,\omega) = \int_{a(t,\omega)}^{b(t,\omega)} s(l)dl,$$

where $s(l) = s(t\cos\omega - l\sin\omega, t\sin\omega + l\cos\omega)$ and scalars $a(t,\omega)$ and $b(t,\omega)$ are defined by both the geometry of the original scene and the geometry of the acquisition system (see Fig. 1). In CT, projections are sampled along a linear numerical detector with n points of sampling whose abscissas belong to the set $\tau = \{t_1, t_2, \dots, t_n\}$. This leads to the discrete parallel-beam line integral projection, $\mathbf{R}_{\omega,\tau}$: $\mathcal{L}_2(\mathcal{D}) \mapsto \mathbb{R}^n$, of the scene s taken at the view angle ω for the set of sampling points τ :

$$\mathbf{R}_{\omega,\tau}(s) = (P_s(t_1,\omega), \dots, P_s(t_n,\omega))^T. \quad (1)$$

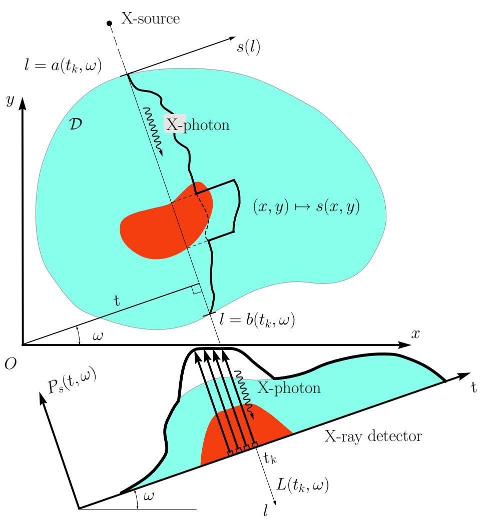


Fig. 1. Geometry of a 2D object imaged with a tomographic system.

2.2 Parametric background model

The background corresponds to the original scene without any anomaly. Sometimes such a background is assumed to be a known function or a zero-mean Gaussian random field (Frakt *et al.*, 1998). But, in general, such an assumption is not verified and the background must be considered as an unknown deterministic function (see for example (Kazantsev *et al.*, 2002)), which will be called nuisance parameter in the rest of the paper. It is assumed that $h \in \mathcal{L}_2(\mathcal{D})$. To model the background, a parametric point of view is taken up. In particular, a linearly parameterized approximation of the background h is used in the paper :

$$h(x,y) \approx h_\mu(x,y) = \sum_{k=1}^{m} \mu_k h_k(x,y), \;\; \forall (x,y) \in \mathcal{D},$$

where $\{h_1, \dots, h_m\}$ is a known family of basic functions in $\mathcal{L}_2(\mathcal{D})$ and $\mu = (\mu_1, \dots, \mu_m)^T$ is a real-valued vector of unknown parameters. A discussion about advantages and disadvantages to use such a parametrization can be found in (Fillatre and Nikiforov, 2003). Certainly, the choice of basic functions $\{h_1, \dots, h_m\}$ is crucially important and it is assumed that the error of approximation is negligible. For example, to model a background with low frequencies variations, it is possible to use a polynomial background defined by :

$$h_\mu(x,y) = \sum_{k=(i,j)\in T_m} \mu_k \, x^i y^j, \quad (2)$$

where $T_m = \{(i,j) \in \mathbb{N} \times \mathbb{N} \,|\, 0 \le i \le m, 0 \le j \le m - i\}$ and m is the maximum order of the polynomial function.

2.3 Anomaly model

The goal is to detect any significant deviations from the background. Let us define the function $(x,y) \mapsto f_{g,\mu}(x,y) \in \mathcal{L}_2(\mathcal{D})$ representing a local variation of the attenuation coefficient in the original scene due to the presence of the anomaly :

$$f_{g,\mu}(x,y) = \begin{cases} g(x,y) - h_\mu(x,y) & \text{if } (x,y) \in d \\ 0 & \text{if } (x,y) \in \mathcal{D} \setminus d \end{cases},$$

where $(x,y) \mapsto g(x,y)$, $(x,y) \in d$, is the anomaly attenuation coefficient and $d \subsetneq \mathcal{D}$. Practically, the quantity $f_{g,\mu}$ represents a *contrast* between the background h_μ and the true anomaly g.

2.4 Measurement model

Two situations are possible : $\mathcal{H}_0 = \{$there is no anomaly$\}$ and $\mathcal{H}_1 = \{$there is an anomaly$\}$. It follows that :

$$s(x,y) = \begin{cases} h_\mu(x,y) \text{ under } \mathcal{H}_0 \\ f_{g,\mu}(x,y) + h_\mu(x,y) \text{ under } \mathcal{H}_1 \end{cases}. \quad (3)$$

Putting together equations (1) and (3), we obtain the following measurement model for a particular view angle ω :

$$\mathbf{Y}_\omega = \begin{cases} \mathbf{H}_\omega \mu + \xi \text{ under } \mathcal{H}_0 \\ \theta_\omega + \mathbf{H}_\omega \mu + \xi \text{ under } \mathcal{H}_1 \end{cases}, \quad (4)$$

where $\theta_\omega = \mathbf{R}_{\omega,\tau}(f_{g,\mu})$, $\mathbf{H}_\omega = (\mathbf{H}_\omega^1, \dots, \mathbf{H}_\omega^m)$ with $\mathbf{H}_\omega^k = \mathbf{R}_{\omega,\tau}(h_k)$ for $k = 1, \dots, m$ and $\xi \sim \mathcal{N}(0_{n,1}, \sigma^2 I_n)$ corresponds to errors of measurement (I_n is the unit matrix of order n and

$0_{n,m}$ is the $n \times m$ matrix whose all elements are zero). It is assumed that $m < n$ and $\sigma > 0$ is known. If P projections are available, the vectors $\mathbf{Y}_{\omega_1,\ldots,\omega_P}$, $\theta_{\omega_1,\ldots,\omega_P}$ and the matrix $\mathbf{H}_{\omega_1,\ldots,\omega_P}$ can be easily designed from "elementary" components $\mathbf{Y}_{\omega_i}$, θ_{ω_i} and $\mathbf{H}_{\omega_i}$ by concatenation. In the rest of the paper, the subscript ω will be omitted to simplify the notations except if it is necessary to avoid misinterpretations.

3. DETECTION OF AN ANOMALY

3.1 Problem of detection

To detect a possible anomaly, it is natural to consider the following hypotheses testing problem between the null hypothesis :

$$\mathcal{H}_0 = \{\mathbf{Y} \sim \mathcal{N}(\mathbf{H}\mu, \sigma^2 I_n), \mu \in \mathbb{R}^m\} \qquad (5)$$

and the alternative one :

$$\mathcal{H}_1 = \{\mathbf{Y} \sim \mathcal{N}(\theta + \mathbf{H}\mu, \sigma^2 I_n), \theta \notin R(\mathbf{H}), \mu \in \mathbb{R}^m\} \qquad (6)$$

where $R(\mathbf{H})$ denotes the linear space spanned by the columns of $\mathbf{H}$, while considering μ as a *nuisance parameter*. Because $R(\mathbf{H})$ is the nuisance parameter space, it is natural to assume that $\theta \notin R(\mathbf{H})$ under $\mathcal{H}_1$: the anomaly must not coincide with the background.

3.2 Testing between two composite hypotheses

The quality of a statistical test $\delta : \mathbb{R}^n \mapsto \{\mathcal{H}_0, \mathcal{H}_1\}$ is defined with the probability of false alarm and the power of the test (Lehman, 1986). Let us consider the class of tests $\mathcal{K}_\alpha = \{\delta : \sup_{\theta \in R(\mathbf{H})} \Pr_\theta(\delta = \mathcal{H}_1) \leq \alpha\}$ where the probability $\Pr_\theta$ stands for the vector of observations $\mathbf{Y}$ being generated by the distribution $\mathcal{N}(\theta, \sigma^2 I_n)$. The power function $\beta_\delta(\theta)$ is defined with the probability of detection : $\beta_\delta(\theta) = \Pr_\theta(\delta = \mathcal{H}_1)$. The hypotheses testing problem (5)-(6) presents two main difficulties : (i) the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ are composite (a hypothesis is composite when it does not define a unique distribution of probability for the vector of observations) and (ii) there is an unknown nuisance parameter μ. There is no general way to design a test between two composite hypotheses (especially with nuisance parameters) (Lehman, 1986). An important particular case of the composite hypotheses testing problem was proposed by Wald. His idea is to impose an additional constraint on the class of considered tests, namely, a *constant power function* $\beta_\delta(\theta)$ over a family of surfaces defined on the parameter space $\mathbb{R}^n$ (see details in (Wald, 1943)) The reasons which justify such an approach in the case of anomaly detection are discussed in (Fillatre and Nikiforov, 2003).

Definition 1. (UBCP test (Wald, 1943)). A test $\delta^* \in \mathcal{K}_\alpha$ is said to have uniformly best constant power (UBCP) on the family of surfaces $\mathcal{S} = \{S_c\, ; c \geq 0\}$ defined over $\mathbb{R}^n$, if the following conditions are fulfilled :

(1) for any pair of points θ' and θ'' which lie on the same surface $S_c \in \mathcal{S}$, $\beta_{\delta^*}(\theta') = \beta_{\delta^*}(\theta'')$;
(2) for another test $\delta \in \mathcal{K}_\alpha$, which satisfies the previous condition, we have $\beta_{\delta^*}(\theta) \geq \beta_\delta(\theta)$ for all $\theta \in S_c$ and all surface S_c.

It is worth to note that the choice of the family of surfaces is naturally imposed by the statistical nature of the studied problem.

3.3 Optimal test

Consider now the hypotheses testing problem (5)-(6), given the measurement model (4). Let be the decision function $\Lambda(\mathbf{Y}) = \frac{1}{\sigma^2}\mathbf{Y}^T P_{\mathbf{H}}^{\perp}\mathbf{Y}$ and $\mathcal{S} = \{S_c : \frac{1}{\sigma^2}\theta^T P_{\mathbf{H}}^{\perp}\theta = c^2,\ c \geq 0\}$ where $P_{\mathbf{H}}^{\perp} = I_n - \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-}\mathbf{H}^T$ is the projection matrix on the left null space of the matrix $\mathbf{H}$ and $\mathbf{A}^-$ is a generalized inverse of $\mathbf{A}$.

Theorem 1. (UBCP test). Let $\alpha \in]0,1[$ and δ^* be defined by :

$$\delta^*(\mathbf{Y}) = \begin{cases} \mathcal{H}_0 & \text{if } \Lambda(\mathbf{Y}) < \lambda_\alpha \\ \mathcal{H}_1 & \text{else} \end{cases}, \qquad (7)$$

where the threshold λ_α is chosen to satisfy the false alarm bound α : $\Pr_{0_{n,1}}(\Lambda(\mathbf{Y}) \geq \lambda_\alpha) = \alpha$. Then, the test δ^* is UBCP in the class $\mathcal{K}_\alpha$ over the family of surfaces $\mathcal{S}$.

The statistics Λ is distributed according to the χ^2 law with $n - q$ degrees of freedom where $q = \text{rank}(\mathbf{H})$. This law is central under $\mathcal{H}_0$ and noncentral under $\mathcal{H}_1$ with a parameter of noncentrality equal to $c^2 = \frac{1}{\sigma^2}\theta^T P_{\mathbf{H}}^{\perp}\theta$. An illustration of the theorem 1 is provided by the Fig. 2. Anomalies are assumed to belong to the space $\mathbb{R}^3$ and the nuisance parameter is scalar ($\text{rank}(\mathbf{H}) = 1$). It is straightforward to show that each surface S_c is a cylinder whose the axis of revolution coincides with the straight line oriented by $\mathbf{H}$ and passing through the origin O. The radius of the cylinder S_c is σc. The *parity space* is the plane orthogonal to $\mathbf{H}$. It can be interpreted like an "informative space" : if the anomaly moves in a direction parallel to this space, the power of detection changes. At the opposite, the axis of revolution of the surface can be interpreted like a "slippery space" : the anomaly can "slide" parallel to this space without changing its probability to be detected. In fact, the test δ^* is invariant with respect to the group of translations $\mathbf{Y} \mapsto \mathbf{Y} + \mathbf{H}\mu$, $\mu \in \mathbb{R}^m$.

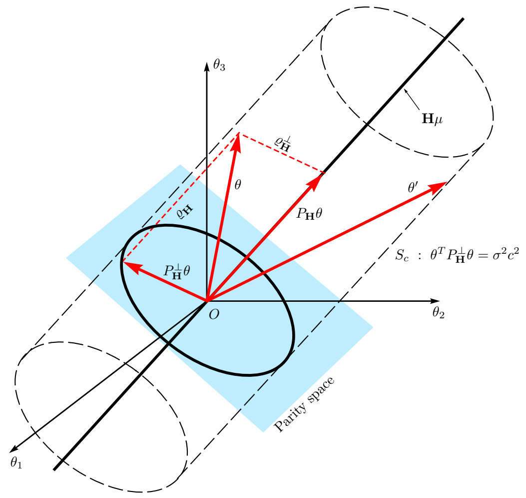


Fig. 2. An illustration of the parameter space in the 3D case with rank $(\mathbf{H}) = 1$.

4. DETECTABILITY OF SIZE-LIMITED ANOMALIES

The test δ^* can detect an anomaly by completely eliminating the unknown background. Obviously, the rejection of the background can damage the anomaly in the projection and the question which must be asked concerns the detectability of anomalies, that is to say the ability of the decision function to keep a trace of the anomaly after the rejection of the background. In this section, some definitions and results about detectability are given.

4.1 Motivation

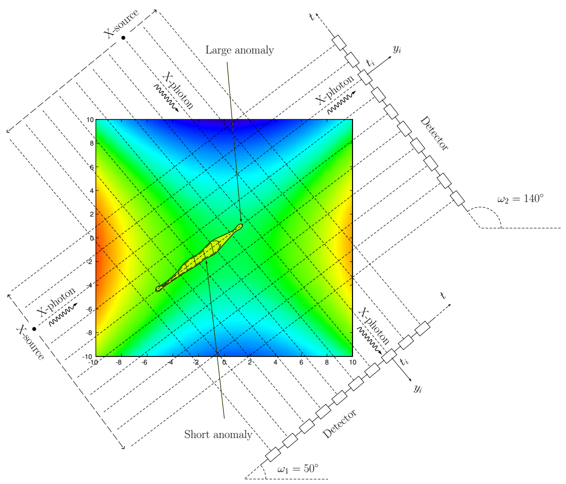


Fig. 3. An original scene composed of a polynomial background with a large anomaly (full line) or with a short anomaly (dotted line).

To motivate the analysis of the detectability of anomalies, let us consider the detection of an anomaly into a polynomial background :

$$h_\mu(x,y) = 1 - 2\,x - \,y + 0.1\,x\,y - x^2 + y^2,$$

with -10 cm $\leq x, y \leq 10$ cm. The linear detector is composed of $n = 10$ sensors regularly spaced (see Fig. 3). It is assumed that measures are "ideal" (no noise). Two different anomalies are considered, a large one (Fig. 3, anomaly with a full line) and a short one (Fig. 3, hatched anomaly with a dotted line), whose the projections, taken at the view angle $\omega_1 = 50°$, differ only for the 6-th sensor (Fig. 4.(a)). On its support of definition, the short anomaly perfectly coincides with the large one and they have the same constant attenuation coefficient. At the naked eye, it appears that the two projections are quite similar. The Fig. 4.(b) presents the vector $Z = P_{\mathbf{H}}^{\perp}\mathbf{Y}$ after the elimination of the background. Clearly, the large anomaly is completely eliminated by the rejection process whereas the shorter is still visible. This elementary example shows that anomalies can be seriously damaged by the decision process and it is necessary to identify anomalies which may be "erased".

4.2 Detectability criterion

First, definitions and results are proposed when only one projection is available. Next, results are extended to several projections.

4.2.1. One projection case : It is natural to define the detectability of an anomaly as follows (Basseville and Nikiforov, 1993).

Definition 2. (detectability). An anomaly $\theta \neq 0 \in \mathbb{R}^n$ is detectable if and only if :

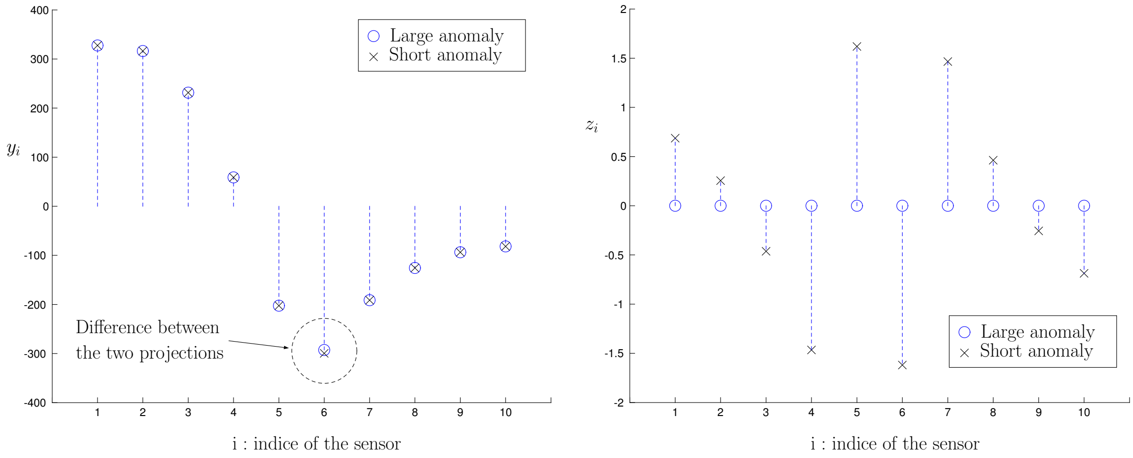
$$\inf_{\nu\in\mathbb{R}^m,\ \mu\in\mathbb{R}^m} \varrho(\mathbf{H}\nu,\ \theta + \mathbf{H}\mu) > 0,$$

where $\varrho(\theta, \theta')$ is the Kullback-Leibler Information (KLI) between two distributions of probability parameterized by vectors θ and θ'.

In other words, a fixed anomaly θ is detectable with respect to the background $\mathbf{H}$ if it is impossible to nullify the KLI between $\mathcal{H}_0$ and $\mathcal{H}_1$ by a simple variation of the nuisance parameter μ. Considering the test δ^* (7), its is straightforward to prove that this definition is equivalent to the following one.

Corollary 1. An anomaly $\theta \neq 0 \in \mathbb{R}^n$ is detectable if and only if $P_{\mathbf{H}}^{\perp}\theta \neq 0$.

Hence, if the anomaly belongs to the null space of the matrix $P_{\mathbf{H}}^{\perp}$, it is completely eliminated by the rejection of the background : the test δ^* (7) becomes to be "blind" to the anomaly. In the general case $\theta \in \mathbb{R}^n$, it always exists anomalies that are not detectable. A natural restriction



(a) (b)

Fig. 4. Difference of detectability between two anomalies : (a) projections $\mathbf{Y}$ at 50° of the two possible scenes presented on Fig. 3 and (b) vectors $\mathbf{Z} = P_{\mathbf{H}}^{\perp}\mathbf{Y}$ calculated from (a) for the two possible anomalies.

on the anomaly is to bound the number of its elements which are not zero : this is a *size-limited anomaly.* Let us note $\theta = (\theta_1, \ldots, \theta_n)^T$, $\text{supp}\,(\theta) = \{i \mid 1 \leq i \leq n \mid \theta_i \neq 0\}$ the set of indices of θ whose associated elements are not zero and $\text{card}\,(\theta) = \text{card}\,(\text{supp}\,(\theta))$ the number of elements of the vector θ which are different from zero. For example, if $\theta = (0\;1\;4\;7\;0)^T$, then $\text{supp}\,(\theta) = \{2, 3, 4\}$ and $\text{card}\,(\theta) = 3$.

Definition 3. (k-detectability). Anomalies of size k are detectable, or k-detectable, if and only if all anomalies θ such that $\text{card}\,(\theta) \leq k$ are detectable.

Let $J_n^k = \{I \,|\, I \subset \{1, \ldots, n\}, \text{card}\,(I) = k\}$ the set of all possible subsets of positive integers between 1 and n which have exactly k elements and $\mathbf{H}_I$ the sub-matrix of $\mathbf{H}$ obtained from $\mathbf{H}$ by eliminating the rows $i_1, \ldots, i_k$ such that $I = \{i_1, \ldots, i_k\}$. A brief study on the linearly dependency of the rows of $\mathbf{H}$ leads to the following lemma :

Lemma 1. Anomalies are k-detectable if and only if $\text{rank}\,(\mathbf{H}_I) = \text{rank}\,(\mathbf{H})$ for all $I \in J_n^k$.

A direct deduction from lemma 1 shows that anomalies are never k-detectable for $n - q < k \leq n$. Hence, the maximum size of detectable anomalies is $n - q$ but this maximum size is not always reached : it depends on the redundance which exists between the rows of $\mathbf{H}$.

4.2.2. Several projections case : It is assumed now that $\mathbf{H} = \mathbf{H}_{\omega_1,\ldots,\omega_P}$ and all matrices $\mathbf{H}_{\omega_i}$ have the same rank q. Let us note k_i the maximum value for which anomalies are k_i-detectable for the view angle ω_i. The lemma 1 can be generalized as follows.

Lemma 2. The maximum size $k_{1,\ldots,P}$ of detectable anomalies $\theta_{\omega_1,\ldots,\omega_P}$ verifies the inequality :

$$k_{1,\ldots,P} \;\geq\; P\,(1 + \min_{1 \leq i \leq n} k_i) - 1.$$

To illustrate the lemma 2, projections of the scene presented on Fig. 3 are taken at two different view angles $\omega_1 = 50°$ and $\omega_2 = 140°$. The two projections are then concatenated to obtain a whole projection with $n = 20$ measures. Such a whole projection is taken for the scene with the large anomaly and the short one (see Fig 5.(a)). The two projections are identical except for the 6-th and 16-th sensors. The background is eliminated by the rejection process and results are visible on Fig. 5.(b). The large anomaly which was undetectable with only one projection is now detectable.

4.3 Application : detection of anomalies into a polynomial background

Let us consider the general polynomial background (2). Due to the discrete Radon transform (1), the content of $\mathbf{H}$ is directly related to the shape of the support $\mathcal{D}$ of the original scene via the boundaries $a(t, \omega)$ and $b(t, \omega)$. Hence, to study the structure of $\mathbf{H}$, it is necessary to precise the role of its support of definition $\mathcal{D}$. Two kinds of boundaries for $\mathcal{D}$ are considered : an unspecified boundary and a rectangular boundary. For the unspecified boundary, the shape of $\mathcal{D}$ may be very complicated (Fig. 1) and no general results about detectability can be proposed for the moment.

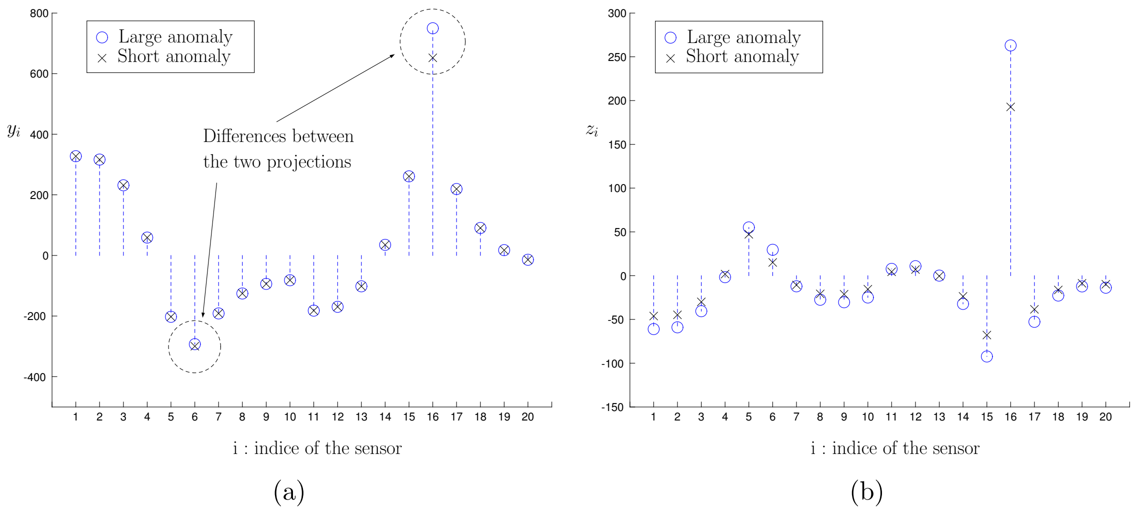


Fig. 5. Difference of detectability between two anomalies for two view angles $\omega_1 = 50°$ and $\omega_2 = 140°$: (a) projections $\mathbf{Y} = \mathbf{Y}_{\omega_1,\omega_2}$ of the two possible scenes presented on Fig. 3 and (b) vectors $\mathbf{Z} = P_{\mathbf{H}}^{\perp}\mathbf{Y}$ calculated from (a) for the two possible anomalies.

At the opposite, a rectangular boundary corresponds to a domain such that $a(t,\omega) = a(\omega)$ and $b(t,\omega) = b(\omega)$. Such a boundary can be found in medical applications like chest X-ray examination and breast imaging.

To study the detectability of anomalies of size k for a rectangular background, the lemma 1 needs to compute the rank of $C_n^k = \frac{n!}{k!(n-k)!}$ matrices, which is computable infeasible when k and n are large enough. However, it can be shown that $\mathbf{H} = \mathbf{V}(\tau)\,\mathbf{U}$ where $\mathbf{H}$ is a $n \times ((m+1)(m+2)/2)$ matrix with rank $(\mathbf{H}) = m+1$, $\mathbf{U}$ is an upper triangular matrix of size $(m+1) \times ((m+1)(m+2)/2)$ with the constant rank $m+1$ and $\mathbf{V}(\tau) = (\mathbf{v}_1, \ldots, \mathbf{v}_n)^T$ is a Vandermonde matrix of size $n \times (m+1)$ such that $\mathbf{v}_i = (1, t_i, t_i^2, \ldots, t_i^m)^T$. Due to this factorization of the matrix $\mathbf{H}$ and elementary algebra, the following lemma follows :

Lemma 3. Assume the family τ of sampling points verify $t_i \neq t_j$ for all $(t_i, t_j) \in \tau \times \tau$. For a polynomial background with a rectangular support, anomalies are $(n-m+1)$-detectable where m is the maximum degree of polynomial functions.

Hence, when the background is a polynomial function, the maximum size of detectability is reached. This reasoning can be apply to other kinds of geometrical configurations of the acquisition system with minor changes.

5. CONCLUSION

The problem of anomaly detection from a few noisy tomographic projections has been considered as a composite hypotheses testing problem. The unknown background is assumed to be linearly parameterized and is considered as a nuisance parameter. An UBCP invariant statistical test has been proposed. New results on the anomaly detectability constitutes the main contribution of the paper. It appears that an increasing number of projections improves the detectability of a size-limited anomaly. The detection and detectability of an anomaly into a non-linear background will be studied in future works.

REFERENCES

Basseville, Michèle and Igor V. Nikiforov (1993). *Detection of abrupt changes : theory and application*. Prentice Hall.

Fillatre, Lionel and Igor Nikiforov (2003). A statistical detection of an anomaly/target from noisy tomographic projections. In: *ISPA'03, Rome, Italy*. pp. 399–404.

Frakt, A.B., W.C. Karl and A.S. Willsky (1998). A multiscale hypothesis testing approach to anomaly detection and localization from noisy tomographic data. *IEEE Trans. Image Processing* **7**(6), 825–837.

Kazantsev, I.G., I. Lemahieu, G.I. Salov and R. Denys (2002). Statistical detection of defects in radiographic images in nondestructive testing. *Signal Processing* **82**(5), 791–801.

Lehman, E.L. (1986). *Testing Statistical Hypotheses, Second Edition*. Chapman & Hall.

Natterer, F. (1986). *The Mathematics Of Computerized Tomography*. John Wiley & Sons.

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Trans. Amer. Math. Soc.* **54**, 426–482.