A-contrario detectability of spots in textured backgrounds

Bénédicte Grosjean

GE Healthcare
Mammography department
Benedicte.Grosjean@ge.com

Lionel Moisan

Université Paris Descartes
MAP5, CNRS UMR 8145
Lionel.Moisan@math-info.univ-paris5.fr

Abstract

Using the a-contrario framework recently introduced in the modeling of human visual perception, we build a statistical model to predict the detectability of a spot on a textured background. Contrary to classical formalisms (ideal observer and its extensions), which assume a known probability distribution for the signal to be detected, the a-contrario observer we build only relies on gestalt-driven measurements and on an approximate representation of the background texture. It extends the scope of previous a-contrario detectors by using a non-i.i.d. naive model and a notion of local context. The models we propose are first validated theoretically in the case of powerlaw textures, which are, in particular, classical models for mammograms. Then, going to more general microtextures (colored noise processes), we compute the relationship between the size of a spot and the minimum contrast required to reach a given detectability threshold according to the a-contrario observer. Three main types of microtextures pop out from this characterization, and in particular low-frequency textures for which curiously enough, the contrast being given, the most salient spots are the smallest ones. Last, we illustrate the interest of the a-contrario observer for two real applications: the detectability of opacities in mammograms and the perception of stains on pieces of clothing.

keywords: a-contrario, detection, texture, colored noise process, powerlaw, fractional Brownian motion, mammography.

1   Introduction

In this paper, we address the following issue: how to detect a "spot" on a grey-level texture image? By spot, we mean some local suspicious deviation of the intensity values, either a local overshoot (light spot) or a local undershoot (dark spot). We shall not assume a particular shape or pattern for the spot, and we shall consider that the spot does not mask the texture, but rather makes it look locally lighter or darker. As well, we shall not assume that the spot has a well-defined boundary (edges). Two examples of real images with such spots are shown on Figure 1. The initial motivation of this work was to measure the detectability of a mass in a X-ray mammographic image (see Figure 1 (left) or Figure 9(a) in the "Application examples" Section), but the problem is more general and we shall present another application, the detection of defects (or stains) in manufactured clothes (see Figure 14, left).
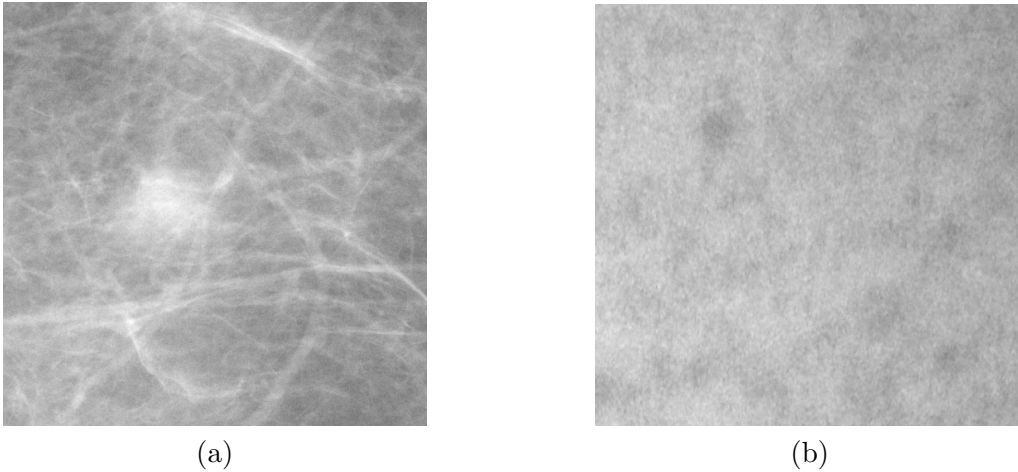
(a)　　　　　　　　　　　　　　　　　　　(b)

Figure 1: Spot examples in microtexture images: an abnormal sign in a real mammogram (approximately in the center of image (a)) and a dark stain on a wallpaper (in the upper-left corner of image (b)).

This spot detection issue is addressed in the context of human vision, that is, we aim at understanding the ability of the human eye to detect such spots. Thus, we shall be more concerned by proposing a detection model inspired from human vision than to search for the "best" detection model. This is consistent with the two applications we mention in this paper. For mammographic images, human observers are still considered as the best experts, even if Computer-Aided Diagnosis (CAD) is frequently used as a second reader. Hence, it is important to understand the human ability to detect masses, in order to improve and optimize the way mammographic images are displayed to the radiologist. As concerns the detection of defects in manufactured clothes, once again the important criterion in most applications (quality control for example) is to relate the detection to visual perception (in a way, a stain that nobody can see is not a stain). This is the reason why we focus on *detectability* rather than on detection, even if it is clear that these two notions involve some common issues.

It is difficult to define precisely what a texture is. Intuitively, it refers to an image where a "pattern" (in the most general sense) is repeated in a stochastic or deterministic way. If the randomness occurs down to the smallest scales (a few pixels), then the image is generally refered as a microtexture (e.g., a cloudy sky). If, on the contrary, the predictibility of the image is high below a certain scale, we rather have a macrotexture (e.g., a brick wall). Hence, the notion of microtexture is deeply related to scale: a grass field may be considered as a macrotexture or as a microtexture, depending at which scale (resolution) it is observed. Mathematically speaking, we shall restrict ourselves to textures modeled by colored noise processes. This class of textures is wide, and yields a good approximation of most microtextures without structured "edge" content (that is, contours). We shall pay special attention to the subclass of *powerlaw textures* (that is, white noise textures filtered with some powerlaw kernel), since they happen to model quite well numbers of natural textures, while their simple description enables deeper theoretical results. An important example is the power spectrum of breast tissue in mammographic images, that is known to decrease like the power -3 of the spatial frequency (see [5, 16]).

The detectability of spots in microtextures reveals a surprising property. In a white noise texture image, a spot with given contrast (that is, let us say, a given difference between the spot and the background average intensities) is, as one could expect, all the more visible than

it is large. However, in some textured images, this logical order can be reversed, that is small spots can be more visible than large ones (the contrast being fixed). This strange phenomenon, illustrated on Figure 2, occurs in particular in the case of powerlaw textures. It has been noticed by Kotre [20] and Burgess [6, 7] in the case of mammographic images, and explained for fixed template matching in $1/f^3$ noise textures. In this work, we shall provide another explanation of this in the case of powerlaw textures by using a general *a-contrario* model not involving the precise knowledge of the spot pattern.

Classical approaches concerned with decision-making in the presence of options are described by the *signal detection theory* (SDT) [27]. In binary hypothesis testing, the task is to decide whether a target is present or not in a given realization [18]: we decide H_0 — the null hypothesis — if we find the target is absent in the considered realization, and H_1 — the alternative hypothesis — otherwise. To make the decision, two main strategies can be followed according to the criterion to be optimized. First, Neyman-Pearson Lemma tells how to maximize the test power for a given probability of a false alarm. Another approach, based on Bayesian statistical decision theory, consists in making a trade-off between the probability of a false alarm and the probability of a miss. This is achieved by minimizing the Bayes risk, which assigns a cost to each possible decision. For a particular case of this cost, this strategy leads to minimizing the probability of error. The approach is then called Maximum A Posteriori (MAP) approach. It can be shown that each of these criteria give the same optimum test statistic. For a given image
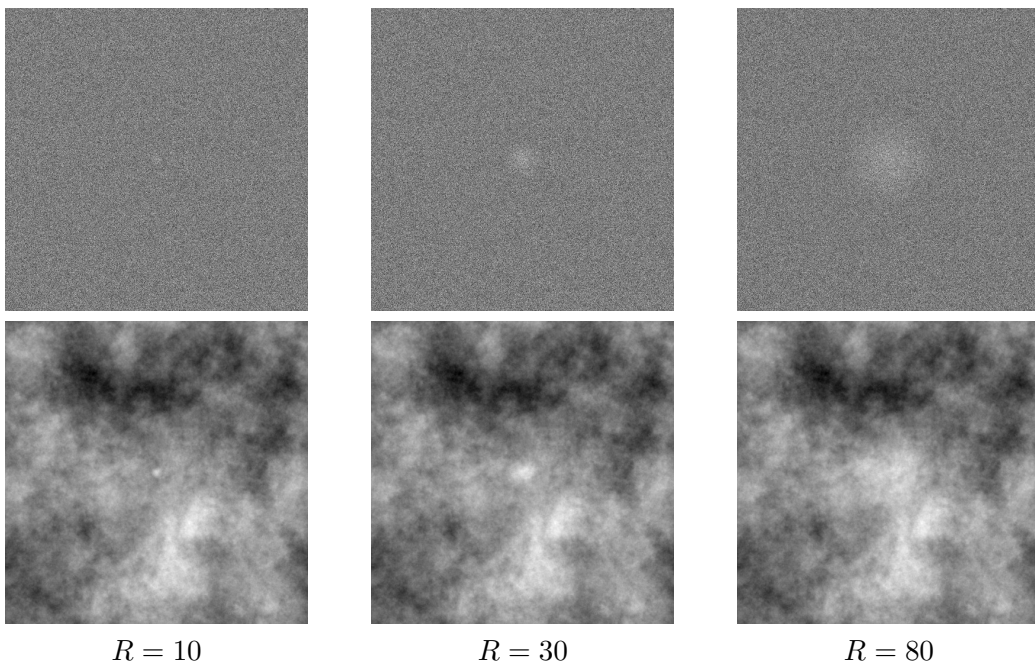

<div align="center">

$R = 10$ $R = 30$ $R = 80$

</div>

Figure 2: Examples of simulated spots with various sizes ($R = 10$, 30 and 80) but similar contrast, in a white noise texture (top row) and in a colored noise texture (bottom row). In the white noise texture, the saliency of the spot increases with its size. On the contrary, in the colored noise background, the unexpected reverse phenomenon occurs: the larger the spot, the less visible it is.

sample g, the test is equal to the likelihood ratio

$$L(g) = \frac{p(g|H_1)}{p(g|H_0)}. \tag{1}$$

Both conditional probabilities for the two hypotheses H_0 and H_1 need to be known in order the test statistic to be computed. Then the decision to detect the signal or not is made by comparing this ratio to a threshold, determined according to the chosen strategy.

A mathematical observer, called the *ideal observer*, has been derived from this theory. It uses the optimal strategy defined by the Bayesian approach. As required by this strategy, the ideal observer makes use of all information available in the image, as well as any prior information about the signal and the underlying processes involved in the image formation. Since it provides optimal decision on a given task, it is an upper bound of the detection performance achievable by any observer. The likelihood ratio is often highly non linear and difficult to determine, so linear discriminants have been developed. Their associated test statistic is the scalar product of the image with a given template. A particular case, the *prewhitening matched filter* (PW), uses the template $t = K^{-1}p$, where K is the covariance matrix of the image background and p is the pattern to detect [2]. In correlated noise, the PW observer derives a template that amounts to decorrelate the noise before the matched filtering process, so that detection issues can be reformulated in the case of a white noise background. In the particular SKE/BKE case (signal known exactly, background known exactly), when the image statistics are Gaussian, the ideal observer is the PW observer.

The template matching formalism has also been extended to further observer models in order to better match with human readers. Among them, the *Eye-filtered Non-PreWithening* observer (NPWE) uses an eye filter E (that takes into account the property of the human visual system), yielding a template $t' = Ep$. Other observer models, like the *Channelized Hotelling Observer* (CHO), first transform the image into channels, obtained by preprocessing the image with a set of spatial-frequency-selective filters inherited from the human visual system. A review of these mathematical observers can be found in [13].

All these observers share a common theoretical basis: the decision between two probabilistic models, H_0 and H_1. In this paper, since we would like to measure the detectability of spots in general, and not of a particular spot pattern, we cannot assume that we have a full knowledge of data statistics (H_1), as required by all these models. In fact, building probabilistic models of the structures to be detected is known to be extremely difficult in Computer Vision, and even apparently simple objects like spots raise severe issues. Indeed, building a complex model for spots would necessarily involve several parameters that would be difficult to estimate. On the contrary, if we try to build a simple probabilistic model for spots, then all results will be valid for this simple model, which will not necessarily tell us much about real spots. Let us give a simple example: for the ideal observer, two patterns p_1 and p_2 with same L^2 norm yield the same detection performances when superimposed to a Gaussian white noise. However, one can easily build two such patterns with different detectability levels for human perception: take for example a disc pattern for p_1, and a very blurry pattern with a complicated shape for p_2, while maintaining $\|p_1\| = \|p_2\|$. In this case, the gestalt processes involved in human vision [17] would clearly favor the perception of p_1.

In this study, we propose to use another statistical framework called *a-contrario theory*, introduced for the detection of alignments in images in [8], and then extended to several other

perceptual grouping laws (see [11]). Contrary to the observers described above, the a-contrario observer does not require a model H_0 for the background and a model H_1 for the structure to detect, but an approximate background model (still written H_0) and one or several measurements inspired from human perception. These ingredients are much more easy to deal with for the detection tasks usually involved in human vision. Moreover, some a-contrario models have shown promising similarities with human vision on simple detection tasks (see [10]). In this work, we shall reformulate and extend a little bit the a-contrario detection framework to distinguish principal and context measurements.

The plan of this paper is the following. We first develop in Section 2 the texture models (white or colored noise) we will use as background models. In section 3, we present an introduction to a-contrario detection, and then build three different spot detectors: one based on a single intensity measurement, one based on a contrast measurement, and the last one based on two measurements (a principal measurement and a context measurement). Then we study in Section 4 the detectability measures they provide for different kind of textures: white noise textures, powerlaw textures, pattern-noise textures, and real-world textures. In Section 5, we consider the two applications mentioned at the beginning of this paper: the detection/detectability of opacities in mammographic images and of stains or defects on clothes.

2 Texture models

There are many ways to define texture models in image analysis, and, as we mentioned before, the word "texture" itself refers more to the idea of the stochastic or deterministic repetition of a pattern than to a precise type of images. In this paper, we consider specific models of microtextures obtained as realizations of a colored noise.

2.1 Discrete white and colored noise

A discrete image is a function $u : \Omega \to \mathbb{R}$, where Ω is a finite subset of $\mathbb{Z}^2$, typically the cartesian product of two discrete intervals. The set Ω is the image domain, and for any $(x, y) \in \Omega$, $u(x, y)$ represents the intensity (grey level[1] of the pixel (x, y)). The size of Ω (number of image pixels) will be denoted by $|\Omega|$.

A discrete white noise image is simply a realization of a discrete 2-dimensional stochastic process defined by $|\Omega|$ independent and identically distributed random variables $N(x, y)$, where $(x, y) \in \Omega$. The distribution of the random variables $N(x, y)$ is generally chosen Gaussian, but other distributions are also possible. In the normalized Gaussian case, it is easy to show that the process N is characterized by the following property.

Proposition 1 *If N is a normalized discrete Gaussian white noise process (that is, all $N(x, y)$ are normal and i.i.d.), then for any n and any discrete functions $f_1, f_2, \ldots f_n \in \mathbb{R}^\Omega$, the random vector $(\langle N, f_i \rangle)_{1 \leq i \leq n}$ is a Gaussian random vector with zero mean and covariance matrix $((\langle f_i, f_j \rangle))_{1 \leq i,j \leq n}$.*

---

[1]Strictly speaking, the grey levels of a digital image are restricted to a finite set, but it will be more convenient in the following to work in the continuous domain, assuming real-valued images. This approximation is justified as long as the quantization step is small enough compared to the dynamics of image intensities.

Here we used the usual discrete inner product $\langle f, g \rangle = \sum_{(x,y) \in \Omega} f(x,y) g(x,y)$, still well defined when f or g is a stochastic process.

A discrete colored noise image is then obtained by convolving a white noise image N with a discrete kernel $k : \Omega \to \mathbb{R}$, that is

$$\forall x, y \in \Omega, \quad U(x,y) = k * N(x,y) = \sum_{(p,q) \in \Omega} k(p,q) N(x \ominus p, y \ominus q),$$

where $\ominus$ means the subtraction modulo Ω (that is, we assume that N is Ω-periodic). Sample examples of a white gaussian noise and colored noise images appear on several figures in this paper: Figure 2 (with a spot superimposed), 6 (b and d), 8 (middle column).

2.2 Continuous noise models

The notion of discrete Gaussian white noise can be transposed to a continuous domain, using the Wiener integral. In the following, we shall use as a definition the continuous analog of Proposition 1 given in the discrete case. This will be useful to establish theoretical results involving a (continuous) scale parameter in the image domain.

Definition 1 (white noise) *A normalized 2-dimensional Gaussian white noise is a stochastic process $N = \dot{B}$ (formal "derivative" of the Brownian process B) such that for any n and any functions $f_1, f_2, \ldots f_n \in L^2(\mathbb{R}^2)$, the random vector $(\langle N, f_i \rangle)_{1 \leq i \leq n}$ is a Gaussian vector with zero mean and covariance matrix $((\int f_i f_j))_{1 \leq i,j \leq n}$, where*

$$\langle N, f \rangle = \int f(x,y) \, \dot{B}(x,y) \, dx dy \tag{2}$$

is the 2-dimensional Wiener integral.

In particular, if $f, g \in L^2(\mathbb{R}^2)$, then $\langle N, f \rangle \sim \mathcal{N}\left(0, \int f^2\right)$ and the two random variables $\langle N, f \rangle$ and $\langle N, g \rangle$ are independent if and only if $\int fg = 0$. It is important to note that $\dot{B}$ is not a stochastic process in the usual sense, and in particular that a realization of $\dot{B}$ is not a L^2 function that could give a classical sense to the Wiener integral (2). However, in the following we will only use N through random variables like $\langle N, f \rangle$ (with $f \in L^2(\mathbb{R}^2)$), completely characterized by classical integrals.

The white noise process is not a good model for textures in general, since it assumes no spatial correlation, even for short distances. A more interesting class of models, relevant for microtextures, is given by colored noise processes.

Definition 2 (colored noise) *Let k be a tempered distribution over $\mathbb{R}^2$ such that its Fourier Transform $\hat{k}$ is bounded. A colored noise with kernel k is a stochastic process N_k such that for any n and any functions $f_1, f_2, \ldots f_n \in L^2(\mathbb{R}^2)$, the random vector $(\langle N_k, f_i \rangle)_{1 \leq i \leq n}$ is a Gaussian vector with zero mean and covariance matrix*

$$\left(\left(\frac{1}{4\pi^2} \int |\hat{k}|^2 \cdot \hat{f}_i \cdot \hat{f}_j^* \right) \right)_{1 \leq i,j \leq n},$$

where $\hat{f}_j^*$ stands for the complex conjugate of the Fourier transform of f_j.

6

Formally, N_k is obtained by the convolution a Gaussian white noise process N with k. In the case of a single function $f \in L^2(\mathbb{R}^2)$, Definition 2 implies that $\langle N_k, f \rangle$ is a Gaussian random variable with zero mean and variance $\|k * f\|^2$. Notice also that if k is the Dirac distribution, then N_k is a Gaussian white noise.

Among colored noise processes, powerlaw textures play a special role since they achieve a scale invariance often found in natural textures (see [1, 24]). As we mentioned before, mammographic background textures present interesting similarities with $1/f^3$ powerlaw textures[2] (case $\beta = 3$ below). A powerlaw texture can be defined rigorously from the fractional Brownian sheet, but it can also be viewed formally as a colored noise associated to a kernel k whose Fourier Transform is

$$\hat{k}(\xi) = \frac{1}{|\xi|^{\beta/2}}, \quad \xi \in \mathbb{R}^2 \setminus \{0\} \tag{3}$$

for some $\beta \in \mathbb{R}$ (the β parameter arises naturally, without the $1/2$ factor, in the power spectrum of k, that is $|\hat{k}|^2$). Since the function $\hat{k}$ is unbounded around $\xi = 0$, we shall consider a powerlaw texture as the limit when $\varepsilon \to 0$ of colored noises associated to the truncated kernels (k_ε) defined below, which converge pointwise to k. We will show that we can take the limit in the considered models.

Definition 3 (powerlaw truncated kernel) *Let $\beta \geq 0$ and $\varepsilon > 0$. We define k^ε as the unique tempered distribution whose Fourier Transform is defined by*

$$\forall \xi \in \mathbb{R}^2, \quad \hat{k}^\varepsilon(\xi) = \begin{cases} 0 & if \quad |\xi| \leq \varepsilon, \\ |\xi|^{-\beta/2} & if \quad |\xi| > \varepsilon. \end{cases} \tag{4}$$

3 A-contrario detection

3.1 Introduction

The a-contrario[3] detection theory has been primarily developed by Desolneux, Moisan and Morel in [8, 9], then developped by several other researchers (see, e.g. [12, 26, 28, 29]). A complete overview can be found in [11]. The principle of this method is inspired from the human perception of Gestalt Laws, that is, geometric grouping laws governing low-level human vision, e.g. alignment, convexity, color constancy, closeness, parallelism, ... (see [17]). This grouping process can be described by Helmholtz principle: we see a structure in a group of objects when their configuration, according to one or several Gestalt Laws, is unlikely to happen by chance in a random situation. The mathematical setting developed in a-contrario models was first intended for Gestalt grouping, but it rapidly appeared that it could be used for more general detection tasks (see [21] for an example in stereo matching).

An a-contrario model basically requires two ingredients: a *naive model*, and a measurement made on structures to be potentially detected. The naive model is a probabilistic model that

---

[2]Variations between about 2 and 4 in the exponent values between breasts images and inside one breast image seem to be due to variations in breast composition.

[3]The term a-contrario, which comes from Latin and means "by contradiction", refers to the fact that structures are detected as contradictions of a naive model. We shall write a-contrario instead of *a contrario* for better readability.

describes typical configurations where no structure is present. If, for example, we see some dots on a sheet of paper, we naturally assume a priori that these dots have been drawn uniformly and independently on the sheet area. This corresponds to a Bernoulli model (or a Poisson model if the number of dots is not fixed) for the distribution of the dots. Then, our visual system will see a structure (e.g. an alignment) if the observed configuration is, from the point of view of this structure, very unlikely in the naive model, hence very surprising (see Figure 3).

To weigh this amount of surprise, we need a *measurement*, let us say a number $x \in \mathbb{R}$. If the structure is more pregnant when x takes high values, then we may weigh the amount of surprise by the probability $\mathbb{P}(X \geq x)$, where x is the observed measurement and X is the random variable corresponding to the distribution of x in the naive model. Of course in general there will be several measurements $(x_i)_{i \in I}$, one for each group to be potentially detected (the index i may simultaneously encode the position, the size, the type, etc. of the structure to be detected, provided that all these parameters have been quantized on a finite set). In the classical a-contrario framework, the amount of surprise is measured by a number of false alarms, defined by

$$NFA(i, x_i) = |I| \cdot \mathbb{P}(X_i \geq x_i),$$

and the i-th group is said to be *ε-meaningful* if it satisfies $NFA(i, x_i) \leq \varepsilon$ (ε being a positive real number). This setting ensures a fundamental property of a-contrario detection: the average number of ε-meaningful groups detected by chance in the naive model is less than ε. In other words, ε controls the expected number of false alarms.
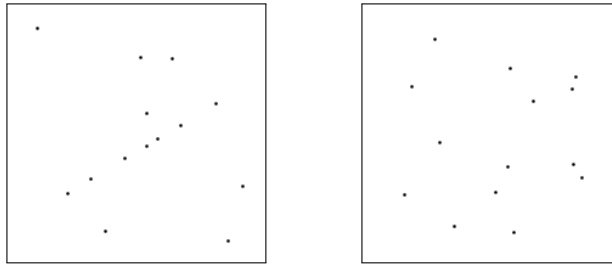

Figure 3: Why can't we help seeing an alignment of dots on the left image? According to Helmholtz principle, we a priori assume that the dots should have been drawn uniformly and independently as in the right image, and we see a structure (here a group of aligned dots) because such an alignment is very unlikely to happen by chance. Alignments of four dots can be *found* in the right image but they do not *pop out*, because they are likely to happen by chance in such an image. The formalization of this principle is realized in a-contrario detection models.

3.2   A general formalization of a-contrario detection

In this section, we propose a general definition of a number of false alarms (NFA), which includes previous constructions made in [11].

Definition 4 (Number of false alarms) *Let* $(X_i)_{1 \leq i \leq N}$ *be a set of random variables. A function* $F(i, x)$ *is a NFA (number of false alarms) for the random variables* (X_i) *if*

$$\forall \varepsilon > 0, \quad \mathbb{E}\left(\left|\{i, \ F(i, X_i) \leq \varepsilon\}\right|\right) \leq \varepsilon. \tag{5}$$

8

Let us comment this definition. Suppose that the joint distribution of the X_i's is given by some null hypothesis H_0. If the function F satisfies (5), then the family of tests $F(i, X_i) \leq \varepsilon$ ensures that the average number of false alarms (that is, the expected number of successful tests under H_0) is less than ε. Thus, a *number of false alarms* allows us to control globally the number of false detections associated to a family of random variables (type I error).

Proposition 2 *Let $(X_i)_{1 \leq i \leq N}$ be a set of random variables and $(n_i)_{1 \leq i \leq N}$ a set of positive real numbers. Then the function*

$$NFA(i, x_i) = n_i \cdot \mathbb{P}(X_i \geq x_i)$$

is a NFA as soon as

$$\sum_{i=1}^{N} \frac{1}{n_i} \leq 1 \tag{6}$$

and in particular if $n_i = N$ for all i.

Proof: The proof essentially relies on the sub-uniform distribution of p-values, precised in Lemma 1 (see appendix). Let $F_i(t) = \mathbb{P}(X_i \geq t)$, we have $NFA(i, x_i) = n_i \cdot F_i(x_i)$. Hence for any $\varepsilon > 0$,

$$
\begin{aligned}
\mathbb{E}\left(\left|\{i,\ NFA(i, X_i) \leq \varepsilon\}\right|\right) &= \mathbb{E}\left(\sum_{i=1}^{N} \mathbb{1}_{NFA(i,X_i)\leq\varepsilon}\right) = \sum_{i=1}^{N} \mathbb{E}\left(\mathbb{1}_{NFA(i,X_i)\leq\varepsilon}\right) \\
&= \sum_{i=1}^{N} \mathbb{P}(NFA(i, X_i) \leq \varepsilon) = \sum_{i=1}^{N} \mathbb{P}\left(F_i(X_i) \leq \frac{\varepsilon}{n_i}\right).
\end{aligned}
$$

We conclude thanks to Lemma 1, since

$$\sum_{i=1}^{N} \mathbb{P}\left(F_i(X_i) \leq \frac{\varepsilon}{n_i}\right) \leq \sum_{i=1}^{N} \frac{\varepsilon}{n_i} \leq \varepsilon. \qquad \square$$

Let us interpret Proposition 2. It basically says that the multiple test $NFA(i, X_i) \leq \varepsilon$, $1 \leq i \leq N$ is controlled by ε, in the sense that under H_0 (naive model) there is no more than ε false detections in average. It can be related to Bonferroni strategy for multiple tests in the following way. Since each test $NFA(i, X_i) \leq \varepsilon$ has a confidence level $\alpha_i = \varepsilon/n_i$ (as a consequence of Lemma 1), the probability of having at least one false alarm is

$$\alpha = \mathbb{P}(\exists i,\ NFA(i, X_i) \leq \varepsilon) \leq \sum_{i=1}^{N} \mathbb{P}(NFA(i, X_i) \leq \varepsilon) \leq \sum_{i=1}^{N} \alpha_i \leq \varepsilon.$$

Hence the $1/n_i$ coefficients can be viewed as a way to distribute the overall risk ε on the multiple tests. This is a reformulation of Bonferroni strategy, which says that an overall confidence level of α is obtained for a group of N tests as soon as each test is realized with a confidence level α/N. The definition of NFA is slightly different because it focus the control of the risk on the average number of false alarms (ε) instead of the probability of at least one false alarm (α). Contrary to Bonferroni strategy which does not make sense any more when $\varepsilon \geq 1$, testing a NFA with $\varepsilon \geq 1$ is not absurd at all. If you detect 1000 events with $\varepsilon = 10$ for example, then you know that

about 10 are false detections but that the remaining events (990 or so) are meaningful. Thus, the a-contrario multiple test strategy simply consists in replacing the notion of p-value (whose interpretation depends on the number of tests) by the notion of NFA, whose value yields an absolute decision criterion that does not depend any more on the number of tests. It would be interesting to study whether the simple multiple test strategy used in the NFA framework could be refined by using more advanced criteria like the area under the LROC curve (see [19, 25]) or the false detection rate (FDR, see [3]), but as far we know it is still an open question.

3.3 Naive model (H_0)

As we said above, a-contrario detection models are based on the contradiction of a naive model. In other terms, a-contrario detections reveal structures of a certain kind (depending on the chosen measurements) that should not occur in the naive model. Hence this naive model is not necessarily supposed to be a very realistic model of the observed images, but rather a simple probabilistic model of images in which we will guarantee that no detection will happen. This is the reason why white noises are often chosen as naive models, since it is generally acknowledged that no special structure should be detected in white noise. In the following, we go a little further and assume that the naive model (written H_0 in reference to the classical statistical setting) is a random image U defined by

$$(H_0): \qquad U = \mu + \sigma N_k, \tag{7}$$

where μ and σ are given real parameters (known or unknown), and N_k a colored noise model as described in Definition 2. We shall first give general detection thresholds for an arbitrary noise kernel k, and then consider more explicitly special kernels corresponding to white noise and powerlaw textures.

3.4 Measurements definition

Now we have defined the kind of naive model we are going to consider, we need to precise the measurements that will be made on the image to detect structures. The idea of a-contrario models is to specify, instead of a template of the "average" pattern to be found, one (or several) intuitive measurements supposedly responsible for the visual saliency of the structure looked for. Indeed, we do not want to specify a template because in general, the structures to be looked for are only vaguely known; here, it would be complicated to define the distribution of shape of what we mean by a "spot". Since we want to focus on spot detection, that is the detection of local suspicious deviation of intensity values, it seems reasonable to use as a measurement the local average intensity in a round-shaped domain, say a disc for example. More generally, we shall consider linear local measurements defined by

$$m(x, y) = \int_{\mathbb{R}^2} u(x - p, y - q)s(p, q) \, dpdq = (u * s)(x, y), \tag{8}$$

where s is a given measurement kernel. The average intensity in a disc is obtained when s is the indicator function of a disc, but the above formula allows more general measurements, that could take into account some psychovisual constraints for example. We shall only impose the following normalization, symmetry and regularity requirements.

Definition 5 (measurement kernel) *A measurement kernel is a function $s \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ such that*

(i) *s is even: $s(x,y) = s(-x,-y)$ almost everywhere;*

(ii) *the Fourier transform of s (noted $\hat{s}$) is a C^2 function;*

(iii) *s is normalized by $\int_{\mathbb{R}^2} s = 1$ (or equivalently by $\hat{s}(0,0) = 1$).*

In practice, s will be positive and bounded with a compact support, but these assumptions are not required in the computations to come. Notice the fundamental difference between the measurement kernel s (connected to the vision process, and independent of the naive model) and the classical notion of template (leading to different optimal statistics depending on the naive model).

3.5 Canonical formulation

Consider the naive model given by (7) and a measurement kernel s. Then the measurement at point $(0,0)$ is a random variable $M(0,0)$ defined by

$$M(0,0) = (U * s)(0,0) = ((\mu + \sigma N_k) * s)(0,0) = \mu + \sigma \langle N_k, s \rangle,$$

because s is normalized and even. Since the model distribution is stationary and the measurements are translation-invariant, the detection thresholds will be the same for each point (x,y), so that we will consider the typical point $(0,0)$ and shorten the notation $M(0,0)$ into

$$M = \mu + \sigma \langle N_k, s \rangle. \tag{9}$$

This can be put in a more general form. Let us define $\varphi = s * k$. This is a L^2 function, and its Fourier transform is $\hat{s} \cdot \hat{k}$, with $\hat{s} \in L^2$ and $\hat{k}$ bounded. Now from Definition 2, we know that there exists a white noise N such that $\langle N_k, s \rangle$ and $\langle N, \varphi \rangle$ follow exactly the same distribution for any measurement kernel s, that is $\langle N, \varphi \rangle \sim N(0, \|\varphi\|^2)$, with

$$\|\varphi\|^2 = \int_{\mathbb{R}^2} \varphi^2 = \frac{1}{4\pi^2} \int_{\mathbb{R}^2} |\hat{\varphi}|^2 = \frac{1}{4\pi^2} \int_{\mathbb{R}^2} |\hat{s}|^2 \cdot |\hat{k}|^2.$$

Hence, we can write M under the canonical form

$$M = \mu + \sigma \langle N, \varphi \rangle, \quad \text{with} \quad \varphi = s * k, \tag{10}$$

and remember that this random variable follows a Gaussian distribution

$$M \sim \mathcal{N}(\mu, \sigma^2 \|\varphi\|^2). \tag{11}$$

3.6 Model 1: no context

Let us begin with the simplest possible detection model. We want to perform some statistical tests on given locations of the image plane to detect "spots". Let us write $\mathcal{T}$ this set of tested

locations: in general, $\mathcal{T}$ will be the whole image grid, or an appropriate subgrid. At each point $(x, y) \in \mathcal{T}$, we make one measurement $m(x, y)$, defined from a measurement kernel s using (8). As we mentioned before, a typical choice for $m(x, y)$ is the average intensity in a disc centered at (x, y). Then we detect a bright spot if this measure $m(x, y)$ is above a certain threshold. As usual in a-contrario models, this threshold will be set in order to control the expected number of false alarms (NFA).

Proposition 3 (Number of false alarms for the model without context) *Let s be a measurement kernel. A number of false alarms associated to the measurements $m(x, y) = (u * s)(x, y)$ for the naive model (7) is given by*

$$NFA_1((x, y), m(x, y)) = |\mathcal{T}| \cdot \Phi_c\left(\frac{m(x, y) - \mu}{\sigma \|\varphi\|}\right).$$

where $\varphi = s * k$, $|\mathcal{T}|$ is the size of $\mathcal{T}$ (that is the number of tested locations) and

$$\Phi_c(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-t^2/2}\, dt = \mathbb{P}\left(X \geq x | X \sim \mathcal{N}(0, 1)\right) \tag{12}$$

is the tail of the Normal distribution.

Proof: From (11), we know that the random variable $\frac{M(x, y) - \mu}{\sigma \|\varphi\|}$ is normal, so that NFA_1 is a NFA thanks to Proposition 2. □

Let us change a little bit the notations. Since the test is the same for any point $(x, y) \in \mathcal{T}$, we can drop the reference to (x, y) in the definition of NFA_1. Conversely, we add a reference to σ and μ to recall that they are necessary parameters. Hence we shall write

$$NFA_1(\sigma, \mu, m) = |\mathcal{T}| \cdot \Phi_c\left(\frac{m - \mu}{\sigma \|\varphi\|}\right). \tag{13}$$

As usual, we shall say that the point $(x, y) \in \mathcal{T}$ is $\varepsilon-$meaningful if $NFA_1(\sigma, \mu, m(x, y)) \leq \varepsilon$. This corresponds to a point where a bright spot is detected in the sense that no more than ε such detections should occur in average in a realization of the naive model. Dark spots could be detected similarly (by changing formally u into $-u$) using the test $NFA_1(\sigma, -\mu, -m(x, y)) \leq \varepsilon$.

This observer does not have a null response at the null frequency, so it is not invariant with respect to grey-level translations, and more generally it is very sensitive to low-frequency image variations. Such models without context are nonetheless efficient in images containing white noise (or high frequency colored noise) background, provided that the image average (μ) is known.

3.7 Model 2: contrast related to the context

To bypass the sensitivity of the previous observer to low frequencies, we introduce another model, which takes into account a local context in its detection task. Since textured backgrounds can have large and slow variations (that is, strong low-frequency components), the local context allows us to have a reasonable local model of the background. It thus incorporates the prior assumption that the spots are in some way isolated from each other. This context area corresponds to the local neighborhood of the potentially detectable target, as represented in Figure 4.
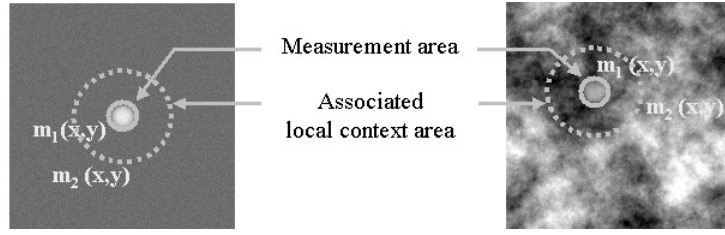
Figure 4: The observer with conditional context performs two measurements at each test location (x, y), typically the average intensity $m_1(x, y)$ in a domain centered at (x, y) (here a disc), and a context measure $m_2(x, y)$ in a larger domain (here the average intensity in a ring around the first disc). The two associated domains may overlap or not.

In this model, two measurements are made at each potential spot location $(x, y) \in \mathcal{T}$. Typically, one (principal) measurement (m_1) will be the average intensity in a disc centered at (x, y), and the second (context) measurement (m_2) will be the average intensity in a ring surrounding the previous disc. In a general setting, we consider two measurement kernels s_1 and s_2, that define through (8) two measurements m_1 and m_2. The corresponding random variables M_1 and M_2 are defined accordingly with (9). Now we can detect suspiciously high values of $M_1 - M_2$, that is a high contrast between the principal and the context measure. This yields the following a-contrario detection model.

Proposition 4 (Number of false alarms for the model with contrast to the context)
Let s_1 and s_2 be two measurement kernels. A number of false alarms associated to the measurements $m_i(x, y) = (u * s_i)(x, y)$ $(i = 1, 2, (x, y) \in \mathcal{T})$ for the naive model (7) is given by

$$NFA_2(\sigma, m_1, m_2) = NFA_1(\sigma, m_1 - m_2) = |\mathcal{T}| \cdot \Phi_c \left(\frac{m_1 - m_2}{\sigma \|\varphi_1 - \varphi_2\|} \right), \qquad (14)$$

provided that $\varphi_1 = s_1 * k$ and $\varphi_2 = s_2 * k$ differ (in the L^2 sense).

Proof: Since property (iii) of Definition (5) is not used in the proof of Proposition 3, the proof remains the same with $m = m_1 - m_2$, $s = s_1 - s_2$ and $\varphi = \varphi_1 - \varphi_2$. $\square$

This measure of contrast $m_1 - m_2$ has two advantages: first, it allows us to cancel the low-frequency components of the single measure m_1, and second it yields detection thresholds independent of μ, which is valuable when the precise value of μ is not known. In a sense, since the term $m - \mu$ in (13) is now replaced with $m_1 - m_2$, one could say that m_2 simply yields an estimate of μ. In fact, it does more than that in complex (i.e. non-white-noise) textures since it provides a *local* estimate of the texture average, which can be significantly different from μ in low-frequency textures.

3.7.1 Remark on the choice of the context

The choice of the measurement function associated with the context is wide, since the only requirement for s_2 is to satisfy the assumptions of Definition 5. We here highlight a particular property of the context measure: the detection levels remain exactly the same when the context area includes the area used for the principal measure (or, more generally, when the principal measure itself is used as a part of the context measure).

Proposition 5 *Let $\alpha \in \mathbb{R} \setminus \{-1\}$. Then for the model with contrast to the context (Model 2), the measurements (s_1, s_2) and (s_1, s_2') with*

$$s_2' = \frac{1}{1 + \alpha}(s_2 + \alpha s_1)$$

yield the same detection levels (same NFA_2 functions).

Proof: We simply need to notice that since $s_1 - s_2' = \frac{1}{1 + \alpha}(s_1 - s_2)$, we have

$$\|\varphi_1 - \varphi_2'\| := \|k * (s_1 - s_2')\| = \frac{1}{1 + \alpha}\|\varphi_1 - \varphi_2\| \qquad \text{and} \qquad m_1 - m_2' = \frac{1}{1 + \alpha}(m_1 - m_2),$$

so that

$$\frac{m_1 - m_2'}{\sigma\|\varphi_1 - \varphi_2'\|} = \frac{m_1 - m_2}{\sigma\|\varphi_1 - \varphi_2\|}$$

and the conclusion follows from (14). $\square$

In particular, if s_1 defines an average over some compact set D_1, that is $s_1 = \mathbb{1}_{D_1}/|D_1|$, then it is equivalent to define the context by the average over either some compact set D_2, or $D_2 \cup D_1$, or $D_2 \setminus D_1$. In particular, if D_1 is a disc, then it is equivalent to define the context by an average on a larger disc or on the ring delimited by the two discs.

3.8   Model 3: conditional context

Since we are now performing two measurements m_1 and m_2 at each point $(x, y) \in \mathcal{T}$, a natural question arises: is the contrast measure $m_1 - m_2$ the only way to combine the principal (m_1) and the context (m_2) information? Knowing the local context measure $m_2(x, y)$, a more subtle way to detect suspicious high values of $m_1(x, y)$ is to consider the probability of a high enough value of $M_1(x, y)$ *conditionally* to the observed value of M_2. This yields the following a-contrario detection model.

Proposition 6 (Number of false alarms for the model with conditional context) *Let s_1 and s_2 be two measurement kernels such that $\varphi_1 = s_1 * k$ and $\varphi_2 = s_2 * k$ are linearly independent. A number of false alarms associated to the measurements $m_i(x, y) = (u \, * s_i)(x, y)$ $(i = 1, 2,$ $(x, y) \in \mathcal{T})$ for the naive model (7) is given by*

$$NFA_3(\sigma, \mu, m_1, m_2) = |\mathcal{T}| \cdot \mathbb{P}(\, M_1 \geq m_1 \,|\, M_2 = m_2). \tag{15}$$

Proof: We need to prove that the NFA property

$$\forall \varepsilon > 0, \quad E_\varepsilon := \mathbb{E}\left(\left|\{(x, y) \in \mathcal{T}, \ NFA_2(\sigma, \mu, M_1(x, y), M_2(x, y)) \leq \varepsilon\}\right|\right) \leq \varepsilon \tag{16}$$

still holds with this kind of conditional measurement. Applying Lemma 2 (see appendix) to

$$F(m_1, m_2) = \mathbb{P}(\, M_1 \geq m_1 \,|\, M_2 = m_2),$$

we get

$$\mathbb{P}\left(F(M_1, M_2) \leq \frac{\varepsilon}{|\mathcal{T}|}\right) \leq \frac{\varepsilon}{|\mathcal{T}|}$$

(note that we shall prove in the next Proposition that since φ_1 and φ_2 are linearly independent, F satisfies the hypotheses of Lemma 2). Hence, for any $\varepsilon > 0$ we have as in Proposition 2

$$E_\varepsilon = \sum_{(x,y)\in\mathcal{T}} \mathbb{P}\left(NFA_2(\sigma, \mu, M_1(x,y), M_2(x,y)) \leq \varepsilon \right)$$

$$\text{hence} \quad E_\varepsilon = \sum_{(x,y)\in\mathcal{T}} \mathbb{P}\left(F(M_1(x,y), M_2(x,y)) \leq \frac{\varepsilon}{|\mathcal{T}|} \right) \leq \sum_{(x,y)\in\mathcal{T}} \frac{\varepsilon}{|\mathcal{T}|} \leq \varepsilon. \qquad \square$$

We now give an explicit form of $NFA_3(\sigma, \mu, m_1, m_2)$.

Proposition 7 *With the hypotheses made in Proposition 6, one has*

$$NFA_3(\sigma, \mu, m_1, m_2) = |\mathcal{T}| \cdot \Phi_c\left(\frac{(m_1 - \mu) - \gamma(m_2 - \mu)}{\sigma \, \|\varphi_1 - \gamma\varphi_2\|} \right), \quad \text{with} \quad \gamma = \frac{\langle \varphi_1, \varphi_2 \rangle}{\|\varphi_2\|^2}. \qquad (17)$$

Proof: Since φ_1 and φ_2 are linearly independent, $\|\varphi_2\| \neq 0$ so that Lemma 3 applies (see appendix) and we have

$$\begin{aligned} \left(M_1 \mid M_2 = m_2 \right) &= \left(\mu + \sigma\langle N, \varphi_1 \rangle \mid \langle N, \varphi_2 \rangle = \frac{m_2 - \mu}{\sigma} \right) \\ &\sim \mathcal{N}\left(\mu + \sigma\gamma\frac{m_2 - \mu}{\sigma} \,,\, \sigma^2\|\varphi_1 - \gamma\varphi_2\|^2 \right) \end{aligned}$$

which implies (17). $\square$

Note that if φ_1 and φ_2 are not linearly independent, either $\varphi_2 = 0$ and we are back to Model 1 (no context), or $\varphi_2 \neq 0$ and M_1 is completely determined by M_2, so that there can be no detection (no surprise arises from the observation of m_1).

It is important to notice the difference between the measurement "M_1 conditionally to M_2" and the joint measurement (M_1, M_2). In both models, the detection domain goes from $\{m_1 \geq \lambda\}$ (Model 1) to a more interesting geometry $\{F(m_1, m_2) \geq \lambda'\}$, but in the conditional framework, the role of the m_2 measurement is simply to adjust the threshold on m_1 to improve the detection performance, yielding a probability of false detection independent of m_2. On the contrary, for the "joint measure" model a spot could be detected only because the context is "rare", since the measures m_1 and m_2 play a symmetric role. This is why we chose a conditional measurement rather than a joint measurement for Model 3.

3.9 Comparison of the two context-related models

As we shall see now, Model 3 is the only NFA model that can be built from a principal measure m_1 and a context measure m_2 without introducing a biais with respect to the context value.

When $\gamma = 1$, Model 3 boils down to Model 2. This happens in particular in white noise textures when s_1 and s_2 are defined as the average over some compact domains D_1 and D_2 such that $D_1 \subset D_2$, since in that case

$$\gamma = \frac{|D_2|^2}{|D_1| \cdot |D_2|} \cdot \frac{\langle \mathbb{1}_{D_1}, \mathbb{1}_{D_2} \rangle}{\|\mathbb{1}_{D_2}\|^2} = \frac{|D_2|}{|D_1|} \cdot \frac{|D_1|}{|D_2|} = 1. \qquad (18)$$

As we shall see later, we still have $\gamma = 1$ for low-frequency powerlaw textures, but in the continuous setting only.

In general, the property $\gamma = 1$ is not satisfied and the measure $m_1 - m_2$ of Model 2 becomes $m_1 - \gamma m_2$. In that case, Model 3 is equivalent to Model 1 with $s = (s_1 - \gamma s_2)/(1 - \gamma)$, so that this coefficient γ can be viewed as the optimal way to combine the two measurements m_1 and m_2 in a single measurement m. This value of γ yields a fair distribution of the false alarms with respect to the context: the ratio of ε-meaningful spots expected under H_0 among locations sharing a particular value m_2 of the context is independent of m_2. As a consequence, under H_0 the distribution of the context values of ε-meaningful spots is exactly the distribution of context values in general (in the present case, a Gaussian distribution). If another detection model is built with a different value γ' of γ (Model 1 if $\gamma' \neq 1$, Model 2 if $\gamma' = 1$), then some context values will be favored (in the sense that the corresponding detection thresholds will decrease), but the price to pay to maintain the average NFA property will be a more strict NFA control for other context values, that is a raise of the other detection thresholds. Since

$$(m_1 - \mu) - \gamma'(m_2 - \mu) = (m_1 - \mu) - \gamma(m_2 - \mu) \quad - \quad (\gamma' - \gamma)(m_2 - \mu),$$

if $\gamma' > \gamma$ (e.g.), the detection will be more difficult for high values of the context $(m_2 - \mu > 0)$ and less difficult for small values $(m_2 - \mu < 0)$.

Hence, the value of γ given in (17) is the only combination of m_1 and m_2 that ensures an unbiased detection with respect to the context value. Since the context measurement is not supposed to be affected by the object to be detected, Model 3 is the most adequate of the two context-sensitive models we proposed. However, Model 2 remains interesting because it does not require the knowledge of μ, which can be very valuable in some situations. Furthermore, the values of γ, numerically measured in the experiments we made, were always close to 1 for low-frequency powerlaw textures, so that it is likely that the detection bias introduced by Model 2 when $\gamma \neq 1$ remains very small in these cases. Thus, in a way Model 2 can be viewed as a practical way to "force" $\gamma = 1$ (and hence eliminate μ) in Model 3.

In the following we study both Model 2 and 3 from a theoretical point of view, but in the applications we shall avoid the dependency on the (unknown) value of μ by systematically considering Model 2.

4 A contrario models: special cases

In this section, we explore further the behavior of the three previous a-contrario detection models, in particular in the case of white noise and powerlaw textures. Since we would like to understand the link between the size of a spot and its detectability (and in particular the contrast/size reversal shown in introduction), we introduce a scaling factor R (typically the radius of the disc defining the principal measurement kernel s_1) and change a little bit the notations as follows. The functions s_i $(i = 1, 2)$ become

$$s_{i,R}(x,y) = \frac{1}{R^2} s_i \left(\frac{x}{R}, \frac{y}{R} \right), \tag{19}$$

which can be rewritten in Fourier domain $\widehat{s_{i,R}}(\xi) = \widehat{s_i}(R\xi)$ for all $\xi \in \mathbb{R}^2$. These new functions $s_{i,R}$ still are measurement kernels, since they satisfy the three conditions of Definition 5. The

measures m_i (and the associated random variables M_i) are defined as before by

$$m_i = u * s_{i,R} \qquad \text{and} \qquad M_i = \mu + \sigma \langle N_k, s_{i,R} \rangle, \tag{20}$$

and so implicitly depend on R.

4.1 White noise textures

4.1.1 Model without context

Let us consider the case when the principal measurement is an averaging of u on a disc with radius R, that is

$$s_{1,R} = \frac{1}{\pi R^2} \mathbb{1}_{B(0,R)} \quad \text{with} \quad \mathbb{1}_{B(0,R)}(x,y) = \begin{cases} 1 & \text{if} \quad x^2 + y^2 \leq R^2 \\ 0 & \text{else.} \end{cases}$$

Since $\|\varphi_{1,R}\|^2 = \|s_{1,R}\|^2 = \frac{1}{\pi R^2}$, from (13) we get (with $s = s_{1,R}$)

$$NFA_1(\sigma, \mu, m, R) = |\Omega| \cdot \Phi_c \left(\sqrt{\pi} R \frac{m - \mu}{\sigma} \right). \tag{21}$$

This model yields a simple relation between the typical size of the principal measurement (R) and the contrast $c = m_1 - \mu$ required to ensure a given NFA, namely

$$\log \frac{c}{\sigma} = -\log R + constant. \tag{22}$$

In other words, to keep the same control on false detections, one should keep the contrast threshold c (or more precisely, the contrast-to-noise ratio c/σ) inversely proportional to R. This is coherent with the fact that in a white noise texture, a spot with given contrast is more visible when it is larger (notice that strictly speaking, Equation 22 links the contrast to the *detector* size R, and not to the unknwon real spot size). This -1 slope of the log-log contrast-detail curve has been noticed for a long time as a characteristic performance of human vision in white noise (see [23]). The ideal observer theory also yields the same slope (see [7] for example).

4.1.2 Model with contrast to the context

We now consider as a context measurement an averaging in a ring with outer radius αR ($\alpha > 1$) and inner radius R (that is, a ring with thickness $(\alpha - 1)R$ surrounding the principal measurement), defined by

$$s_{2,R} = \frac{1}{(\alpha^2 - 1)\pi R^2} \left(\mathbb{1}_{B(0,\alpha R)} - \mathbb{1}_{B(0,R)} \right).$$

Then, since

$$\|\varphi_{1,R} - \varphi_{2,R}\|^2 = \frac{1}{\pi R^2} + \frac{1}{(\alpha^2 - 1)\pi R^2} = \frac{1}{\pi R^2} \cdot \frac{1}{1 - \frac{1}{\alpha^2}},$$

we obtain

$$NFA_2(\sigma, \mu, m_1, m_2, R) = |\Omega| \cdot \Phi_c \left(\sqrt{\pi} \sqrt{1 - \frac{1}{\alpha^2}} R \frac{m_1 - m_2}{\sigma} \right). \tag{23}$$

As we noticed at the end of Section 3.7.1, we would obtain the same detection model by considering a context measurement defined by the averaging over the whole disc $B(0, \alpha R)$, that is

$$s'_{2,R} = \frac{1}{\alpha^2 \pi R^2} \mathbb{1}_{B(0,\alpha R)}.$$

The contrast-detail relation (22) is still satisfied for this model, but now with the relative contrast $c = m_1 - m_2$. In fact, since the texture is a white noise, M_1 and M_2 are independent and the role of the measurement m_2 is to give an estimate of μ. This explains why the detection becomes more efficient (NFA_2 smaller) when α grows: the estimate is simply better because the average is done on a larger support. When $\alpha \to \infty$, the estimate is asymptotically exact and (23) tends to (21). Of course, in practice it is generally not interesting to take a very large support for m_2, since such a global estimate of μ may be polluted by other spots or low-frequency variations.

4.1.3 Model with conditional context

Now what do we obtain with a conditional context? If the context measurement is an averaging in a ring (measurement $s_{2,R}$), then since M_1 and M_2 are independent ($\langle \varphi_{1,R}, \varphi_{2,R} \rangle = 0$), the context measure is useless and $NFA_3(\sigma, \mu, m_1, m_2, R) = NFA_1(\sigma, \mu, m_1, R)$. The situation is different if the context measurement is defined by an averaging on the whole disc $B(0, \alpha R)$ (kernel $s'_{2,R}$). Since s_1 and s_2 produce averages on two domains D_1 and D_2 with $D_1 \subset D_2$, one has $\gamma = 1$ (see Equation 18), so that Models 2 and 3 are equivalent. Thus, we can see that in the particular case of a white noise texture, the optimal model associated to a context boils down to a contrast measurement (Model 2) or to a single measurement (Model 1).

4.2 Powerlaw textures

Now we come to the case of powerlaw textures, defined by the limit when $\varepsilon \to 0$ of the colored noise N_{k^ε}, where k^ε has been defined in (4) and β is a fixed real number satisfying $0 \leq \beta < 6$ (the upper bound is for convergence purposes). Note that there is no link between the ε used here for the approximation k^ε of the ideal powerlaw kernel and the ε used elsewhere as a threshold to NFA for ε-meaningful spots (we maintain the usual notations since they are hardly ambiguous).

4.2.1 Model without context

As we said before, the detection model without context is adapted only to high-frequency textures, since it does not have a null response at the zero frequency. In the following theorem we can see that for powerlaw textures, the limit between low-frequency and high-frequency textures is $\beta = 2$: for $\beta < 2$, the texture energy is unbounded at infinity ($\int |\xi|^{-\beta} d\xi$ diverges at infinity) but bounded on a neighborhood of $\xi = 0$, so that a limit model without context can exist. For $\beta \geq 2$, the texture energy is unbounded on a neighborhood of $\xi = 0$, and the limit model collapses.

Theorem 1 *Let s_1 be a measurement kernel and $s_{1,R}$ its rescaled version by (19). For any $\varepsilon > 0$, we write $NFA_1(\sigma, \mu, k^\varepsilon, m_1, R)$ the number of false alarms proposed in Section 3.6,*

associated to the measurement kernel $s_{1,R}$ and to the naive model defined by (7) with μ, σ and $k = k^\varepsilon$. Then one has

$$NFA_1(\sigma, \mu, k^\varepsilon, m_1, R) \quad \underset{\varepsilon \to 0}{\longrightarrow} \quad \begin{cases} \overline{NFA}(\sigma, m_1 - \mu, s_1, R) & if \quad 0 \le \beta < 2, \\ \frac{1}{2}|\mathcal{T}| & if \quad 2 \le \beta < 6, \end{cases} \qquad (24)$$

with

$$\overline{NFA}(\sigma, m, s, R) = |\mathcal{T}| \cdot \Phi_c \left(\frac{1}{R^{\frac{\beta-2}{2}}} \frac{m}{\frac{\sigma}{2\pi}\sqrt{\int_{\mathbb{R}^2} |\hat{s}(\xi)|^2 \, |\xi|^{-\beta} \, d\xi}} \right). \qquad (25)$$

Proof: From Proposition 3 we know that

$$NFA_1(\sigma, \mu, k^\varepsilon, m_1, R) = |\mathcal{T}| \cdot \Phi_c \left(\frac{m_1}{\sigma \|\varphi_{1,R}^\varepsilon\|} \right), \quad \text{with}$$

$$4\pi^2 \|\varphi_{1,R}^\varepsilon\|^2 = \int_{\mathbb{R}^2} |\widehat{k^\varepsilon}|^2 |\widehat{s_{1,R}}|^2 = \int_{|\xi| \ge \varepsilon} |\widehat{s_1}(R\xi)|^2 \, |\xi|^{-\beta} \, d\xi = R^{\beta-2} \int_{|\zeta| \ge R\varepsilon} |\widehat{s_1}(\zeta)|^2 \, |\zeta|^{-\beta} d\zeta. \qquad (26)$$

Since $\widehat{s_1}(0) = 1$, the term under the last integral is equivalent to $|\zeta|^{-\beta}$ when $\zeta \to 0$, so that if $\beta \ge 2$, the integral over $\mathbb{R}^2$ diverges to $+\infty$ and $NFA_1(\sigma, \mu, k^\varepsilon, m_1, R)$ tends to $|\mathcal{T}| \cdot \Phi_c(0) = \frac{|\mathcal{T}|}{2}$ as $\varepsilon \to 0$. If $\beta < 2$, the integral over $\mathbb{R}^2$ is finite and by Lebegue's monotone convergence Theorem we get the desired limit $\overline{NFA}(\sigma, m_1, s_1, R)$ since Φ_c is continuous. $\square$

4.2.2 Models with context

Contrary to the model without context (NFA_1), the two models involving a context measure (NFA_2 and NFA_3) admit a non-trivial limit when $\varepsilon \to 0$ even for $\beta \ge 2$, as shown below.

Theorem 2 *Let s_1, s_2 be two measurement kernels and $s_{1,R}, s_{2,R}$ their rescaled versions by (19). For any $\varepsilon > 0$, we write $NFA_i(\sigma, \mu, k^\varepsilon, m_1, m_2, R)$ ($i = 2, 3$) the number of false alarms proposed respectively in Section 3.7 and 3.8, associated to the measurement kernels $s_{1,R}, s_{2,R}$ and to a naive model defined by (7) with μ, σ and $k = k^\varepsilon$. Then one has*

$$\forall i \in \{2, 3\}, \quad NFA_i(\sigma, \mu, k^\varepsilon, m_1, m_2, R) \quad \underset{\varepsilon \to 0}{\longrightarrow} \quad \overline{NFA}(\sigma, m_1 - \gamma m_2, s_1 - \gamma s_2, R), \qquad (27)$$

$$with \quad \gamma = \begin{cases} 1 & if \quad i = 2, \\ 1 & if \quad i = 3 \ and \ 2 \le \beta < 6, \\ \bar{\gamma} & if \quad i = 3 \ and \ 0 < \beta < 2, \end{cases} \quad and \quad \bar{\gamma} = \frac{\displaystyle\int_{\mathbb{R}^2} \widehat{s_1}(\xi)\widehat{s_2}(\xi) \, |\xi|^{-\beta} \, d\xi}{\displaystyle\int_{\mathbb{R}^2} |\widehat{s_2}(\xi)|^2 \, |\xi|^{-\beta} \, d\xi}. \qquad (28)$$

Let us first check that the announced limit $\overline{NFA}(\sigma, m_1 - \gamma m_2, s_1 - \gamma s_2, R)$ makes sense, since it involves the integral

$$I = \int_{\mathbb{R}^2} |\widehat{s_1}(\xi) - \gamma \widehat{s_2}(\xi)|^2 \, |\xi|^{-\beta} \, d\xi$$

19

that must be finite. If $\gamma \neq 1$, necessarily $\beta < 2$ and the same analysis as before shows that I is finite ($|\widehat{s_1}(\xi) - \gamma \widehat{s_2}(\xi)|^2$ is bounded around 0, and $|\xi|^{-\beta}$ is integrable around 0). If $\gamma = 1$, then since s_1 and s_2 are real and even, so are $\widehat{s_1}$ and $\widehat{s_2}$ and a Taylor expansion of order 2 gives

$$|\widehat{s_1}(\xi) - \widehat{s_2}(\xi)|^2 \, |\xi|^{-\beta} = \underset{\xi \to 0}{O}\left(|\xi|^{4-\beta}\right), \tag{29}$$

which is integrable around $\xi = 0$ as soon as $\beta < 6$.

4.2.3 Proof of Theorem 2

Let us begin with the case $i = 2$, which is quite easy. From Proposition 4 we know that

$$NFA_2(\sigma, \mu, k^\varepsilon, m_1, m_2, R) = |\mathcal{T}| \cdot \Phi_c\left(\frac{m_1 - m_2}{\sigma\|\varphi^\varepsilon_{1,R} - \varphi^\varepsilon_{2,R}\|}\right),$$

and as in (26),

$$4\pi^2\|\varphi^\varepsilon_{1,R} - \varphi^\varepsilon_{2,R}\|^2 = R^{\beta-2}\int_{|\zeta| \geq R\varepsilon} |\widehat{s_1}(\zeta) - \widehat{s_2}(\zeta)|^2 \, |\zeta|^{-\beta} \, d\zeta, \tag{30}$$

and since the integral over $\mathbb{R}^2$ (written I above) is finite, once again by Lebesgue's monotone convergence Theorem we get that

$$\|\varphi^\varepsilon_{1,R} - \varphi^\varepsilon_{2,R}\| \underset{\varepsilon \to 0}{\longrightarrow} \frac{1}{2\pi}R^{\frac{\beta-2}{2}}\sqrt{I}, \tag{31}$$

which leads to the expected limit for NFA_2.

Now let us come to the case $i = 3$. From Proposition 7 we have

$$NFA_3(\sigma, \mu, k^\varepsilon, m_1, m_2, R) = |\mathcal{T}| \cdot \Phi_c\left(\frac{m_1 - \gamma^\varepsilon m_2}{\sigma\|\varphi^\varepsilon_{1,R} - \gamma^\varepsilon\varphi^\varepsilon_{2,R}\|}\right) \quad \text{with} \quad \gamma^\varepsilon = \frac{\left\langle\varphi^\varepsilon_{1,R}, \varphi^\varepsilon_{2,R}\right\rangle}{\|\varphi^\varepsilon_{2,R}\|^2}. \tag{32}$$

Since γ^ε has a finite limit (written γ) when $\varepsilon \to 0$ (see Lemma 4 in appendix), $\varepsilon \mapsto |\gamma^\varepsilon|$ is bounded (say by $C_1 > 0$) on some nonempty interval $(0, \varepsilon_0)$. Moreover, the numerator of the argument of Φ_c in (32) converges to $m_1 - \gamma m_2$ when $\varepsilon \to 0$, and since

$$4\pi^2\|\varphi^\varepsilon_{1,R} - \gamma^\varepsilon\varphi^\varepsilon_{2,R}\|^2 = R^{\beta-2}\int_{|\zeta| \geq R\varepsilon} |\widehat{s_1}(\zeta) - \gamma^\varepsilon\widehat{s_2}(\zeta)|^2 \, |\zeta|^{-\beta} \, d\zeta$$

as in (26), the convergence of the denominator will be shown as soon as we establish that

$$\int_{|\xi| \geq R\varepsilon} |\widehat{s_1}(\xi) - \gamma^\varepsilon\widehat{s_2}(\xi)|^2 \, |\xi|^{-\beta} \, d\xi \underset{\varepsilon \to 0}{\longrightarrow} \int_{\mathbb{R}^2} |\widehat{s_1}(\xi) - \gamma\widehat{s_2}(\xi)|^2 \, |\xi|^{-\beta} \, d\xi. \tag{33}$$

We use Lebesgue's dominated convergence Theorem to prove this convergence. The identity $(a+b)^2 \leq 2(a^2 + b^2)$ implies that

$$|\widehat{s_1} - \gamma^\varepsilon\widehat{s_2}|^2|\widehat{k^\varepsilon}|^2 \leq 2(g_1 + g_2), \tag{34}$$

20

where $g_1 = |\widehat{s_1} - \widehat{s_2}|^2 |\widehat{k^\varepsilon}|^2$ belongs to $L^1(\mathbb{R}^2)$ as we already shown (Equation 31), and

$$g_2 = (1 - \gamma^\varepsilon)^2 |\widehat{s_2}|^2 |\widehat{k^\varepsilon}|^2.$$

Since $\widehat{s_2}(0) = 1$, we have $|\widehat{s_2}(\xi)| \leq 2$ on some non-empty neighborhood $|\xi| \leq \alpha_1$ of 0 (note that we can impose $\alpha_1 < R\varepsilon_0$), so that

$$\int_{\mathbb{R}^2} g_2 = \int_{|\xi| \geq \alpha_1} g_2(\xi)\, d\xi + \int_{|\xi| < \alpha_1} g_2(\xi)\, d\xi \leq C_2 + 4 \sup_{0 < \varepsilon < \alpha_1} \left((\gamma^\varepsilon - 1)^2 \int_{R\varepsilon < |\xi| < \alpha_1} |\xi|^{-\beta}\, d\xi \right)$$

where $C_2 = (1 + C_1)^2 \cdot \|\widehat{s_2}\|^2 \cdot \alpha_1^{-\beta}$ is a finite constant independent of ε. Now we consider the different possibilities for the asymptotics of γ^ε (see Lemma 4 in appendix), and show that $\int g_2$ can be bounded independently of ε (for $\varepsilon > 0$ small enough) in each case:

- case $0 \leq \beta < 2$: since $|\xi|^{-\beta}$ is integrable around $\xi = 0$,

$$\int g_2 \leq C_2 + 4(1 + C_1)^2 \int_{|\xi| \leq \alpha_1} |\xi|^{-\beta}\, d\xi \; < \; \infty.$$

- case $\beta = 2$: we have $\int_{R\varepsilon < |\xi| < \alpha_1} |\xi|^{-\beta}\, d\xi = \underset{\varepsilon \to 0}{O}(\log \varepsilon)$, so that

$$\int g_2 \leq C_2 + 4 \sup_{0 < \varepsilon < \alpha_1} \left(\underset{\varepsilon \to 0}{O}\left(\frac{1}{(\log \varepsilon)^2} \right) \cdot \underset{\varepsilon \to 0}{O}(\log \varepsilon) \right) \leq C_2 + \underset{\varepsilon \to 0}{o}(1).$$

- case $2 < \beta < 6$: we have $\int_{R\varepsilon < |\xi| < \alpha_1} |\xi|^{-\beta}\, d\xi = \underset{\varepsilon \to 0}{O}\left(\varepsilon^{2-\beta} \right)$ and once again $\int g_2 \leq C_2 + \underset{\varepsilon \to 0}{o}(1)$ since $\varepsilon^{2\beta-4} \cdot \varepsilon^{2-\beta} \to 0$ (case $2 < \beta < 4$), $(\log \varepsilon)^2 \varepsilon^4 \varepsilon^{-2} \to 0$ (case $\beta = 4$), and $\varepsilon^4 \varepsilon^{2-\beta} \to 0$ (case $4 < \beta < 6$).

To sum up, we proved that $g_2 \in L^1(\mathbb{R}^2)$, so that the left term of (34) is dominated by a L^1 function independent of ε (for ε small enough) and (33) follows. This allows us to take the limit in (32), and Theorem 2 is proven. $\square$

4.2.4 Contrast versus size

The two previous theorems show that the three previous NFA models still make sense for power-law textures, but not for all values of β. Model 1 remains valid only when $\beta < 2$ (high-frequency textures), and Model 3 boils down to Model 2 when $2 \leq \beta < 6$ (low-frequency textures). In all cases, the limit NFA model can be expressed as a function of $\overline{NFA}$, which combines the contrast c ($c = m_1 - \mu$ or $c = m_1 - m_2$, depending on the considered model) and the typical size R so as to yield the same given NFA when

$$\log \frac{c}{\sigma} = \frac{\beta - 2}{2} \log R + constant. \tag{35}$$

This generalization of (22) shows that in order to keep the same control on false detections, the relation between contrast and size can be reversed depending on the texture: in a high-frequency texture ($0 \leq \beta < 2$, hence including white noise), a spot with given contrast is more

visible when it is larger, whereas it is more visible when it is smaller in low-frequency textures ($2 < \beta < 6$, hence including mammographic image models), because the slope $(\beta - 2)/2$ then becomes positive. This explains the contrast reversal experiment shown on Figure 2.

In [7], Burgess obtained the same $(\beta - 2)/2$ slope for the prewhitening observer on powerlaw textures, as well as a good fit with the human observer in real mammography images (two-alternative forced choice task). Regarding the a-contrario observer (model with contrast to the context), one of the authors of the present paper recently showed a similar fit with the human observer on simulated powerlaw textures ($\beta = 0$ and $\beta = 3$) for a search detection task [15].

4.3 Relative number of false alarms

4.3.1 Number of tests

In our analysis of the relation between c and R (Equation 35), we implicitly assumed that $|\mathcal{T}|$ was independent of R. From a perceptual point of view, keeping the same number of tests for each detector size may not be the optimal strategy, since redundant tests artificially increase $|\mathcal{T}|$ without improving much the detection power (given the centers of two discs, the relative area of their intersection increases with their radius R, and so does the redundancy of the associated tests). Another strategy, more relevant in terms of visual system, consists in sampling the tests on a grid whose step is proportional to the detector measurement size. This way, large structures will be searched for and located with a worse precision than small ones. To take this into account in the NFA definition, we shall sample the tests on a grid with step $\delta(R) = 0.3 \cdot R$ in order to maintain a fixed overlap ratio (around 80 % for discs) between adjacent principal measurements. This constant 0.3 is arbitrary, but the order of magnitude seems relevant, and its precise value has not a strong influence in practice. With this convention, we can define a *relative NFA* by

$$NFA_r(\sigma, c, R) = \frac{|\Omega|}{\delta(R)^2} \cdot \Phi_c \left(\frac{c}{\sigma \|k * (s_{1,R} - s_{2,R})\|} \right). \tag{36}$$

For powerlaw textures, we shall take the limit of the right term for truncated kernels tending to the true powerlaw kernel (3). With this notion of relative NFA_r (that we shall use in all the following), the detectability of large structures is slightly increased and (35) becomes

$$\log \frac{c}{\sigma} = \frac{\beta - 2}{2} \log R + \log \Phi_c^{-1} \left(constant \cdot R^2 \right). \tag{37}$$

Compared to (35), this last relation yields (non-constant) slopes slightly smaller than $(\beta - 2)/2$, which is in agreement with the slopes measured for human observers (see [7]).

To measure the detectability of a spot, we may assume that R is fixed. In a detection issue, we need to test several values of R (say $R_1, R_2, \ldots R_k$, in geometric progression) while keeping the overall NFA property. This can be done with $NFA'(\sigma, c, R_i) = n(R_i, \Omega) \cdot NFA_r(\sigma, c, R_i)$, which is a NFA provided that

$$\sum_{i=1}^{k} \frac{|\Omega|}{\delta(R_i)^2 \cdot n(R_i, \Omega)} \leq 1$$

(see Proposition 2). If we assume that each value of R deserve the same overall control, then we may take $n(R_i, \Omega) = k \cdot |\Omega|/\delta(R_i)^2$, but if we want to give the same weight to each possible

detection, then we shall have $n(R_i, \Omega) = |\Omega| \cdot \sum_{i=1}^{k} \delta(R_i)^{-2}$. Both choices ensure the NFA property, but correspond to different strategy and prior assumptions on the size of the spots to be found (related to R, the measurement size).

4.3.2   Contrast-detail curves

The relation between contrast and size is usually represented by a curve c function of R, called CD-curve (for "contrast-detail"). We transpose this representation below for the a-contrario observer.

Definition 6 (contrast-detail curve) *We call CD-curve with level ε the set of pairs (c, R) for which $NFA_r(1, c, R) = \varepsilon$ for a contrast $c = m_1 - m_2$, a principal measurement $s_{1,R}$ defined by an averaging over a centered disc with radius R, and a context measurement $s_{2,R}$ defined by an averaging over a centered ring with inner radius R and outer radius $R\sqrt{2}$.*

The equation of CD-curves is easily deduced from (36): for a given naive model defined by a texture kernel k, the CD-curve with level ε is given by

$$c(R) = \|k * (s_{1,R} - s_{2,R})\| \cdot \Phi_c^{-1}\left(\frac{\varepsilon}{\delta(R)^2}\right). \tag{38}$$

In practice, we shall choose $\varepsilon = 1$ to compute the CD-curve of a given real image. For a theoretical CD-curve (with no special reference image), we shall take arbitrarily $|\Omega| = 10^6$ to get a reasonable order of magnitude on a typical 1000×1000 image. Contrast-detail curves for powerlaw textures are shown on Figure 5.

4.4   Pattern-noise textures

The CD-curves for powerlaw textures are monotone, yielding positive or negative dependency between size and detectability. Some textures may present a more complex detectability profile, with local extrema in the CD-curve. Such a situation arises when the texture kernel k is a simple generating pattern image with a restricted support (we shall call this "pattern-noise"). The corresponding texture generation process amounts to superimpose weighted copies of the generating pattern image at each point of the texture image, with Gaussian i.i.d weights. Two examples of pattern-noise are given on Figure 6. Since these textures are build by mixing copies of a single shape, one may expect the typical size of the shape to play a particular role in the associated CD-curve. This is indeed the case, as shown on Figure 7: the CD-curve presents a bell-shape, with a maximal required contrast (that is, minimum detectability) when the measurement size (R) approximately matches the shape size (say S). Since the texture naturally presents structures with size S, the a-contrario observer requires more contrast to detect a spot when $R \simeq S$ than when $R \ll S$ or $R \gg S$.

4.5   Real-world nonparametric textures

As a generalization of pattern-noise, a microtexture can also be created from a natural image u by using the image itself as a texture kernel. This amounts to keep the modulus of the
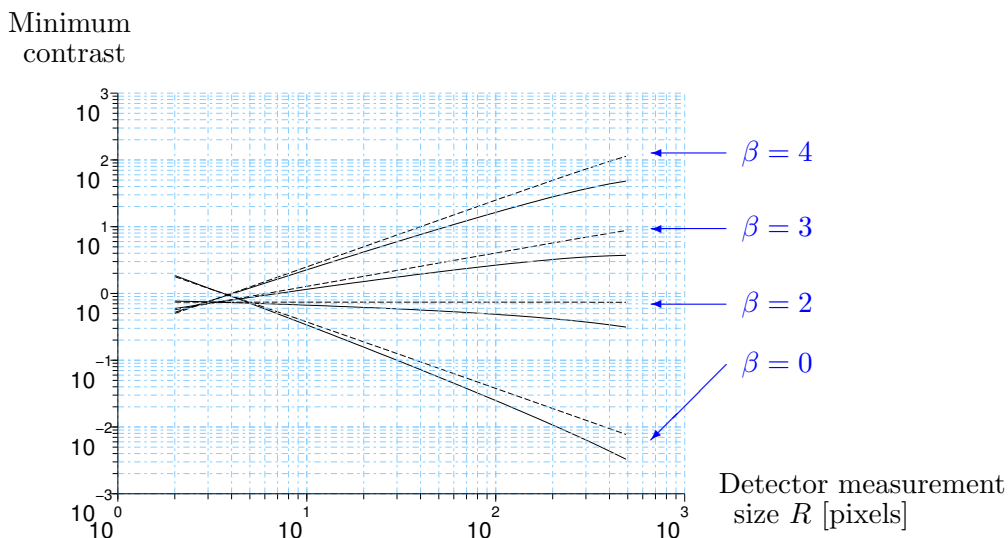
Figure 5: Theoretical contrast-detail (CD) curves (relation between the minimum contrast required for a detection and the size of the associated measurement) obtained by the relative a-contrario observer (NFA_r) in the particular case of powerlaw textures, with $\varepsilon = 1$ and $|\Omega| = 10^6$. Solid lines: CD-curves, defined from NFA_r with a variable number of tests ($\delta(R) = 0.3 \cdot R$, yielding a constant overlapping ratio around 80% for adjacent detectors). Dashed lines: the same curves, but for a fixed number of tests ($\delta(R) = 1$, less realistic from a perceptual point of view). The contrast-reversal noticed in introduction (Figure 2) appears here in the different slopes of the CD-curves: positive slope for $\beta < 2$ (detectability increases with size) and negative slopes for $\beta > 2$ (detectability decreases with size).
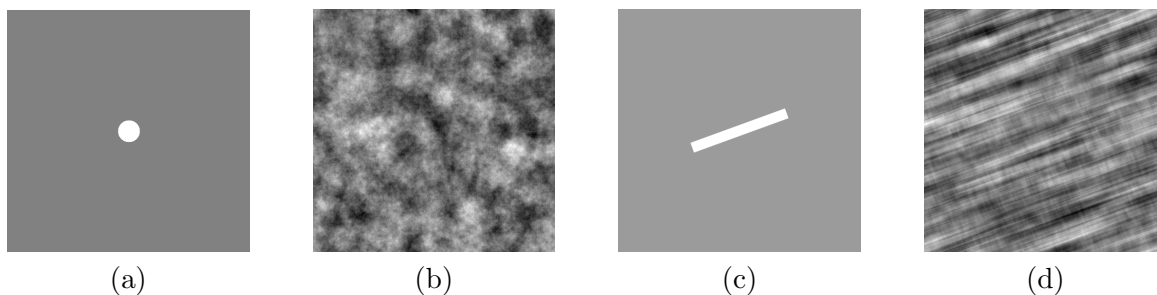

Figure 6: Pattern images (a,c) and samples of their corresponding pattern-noise textures (b,d).
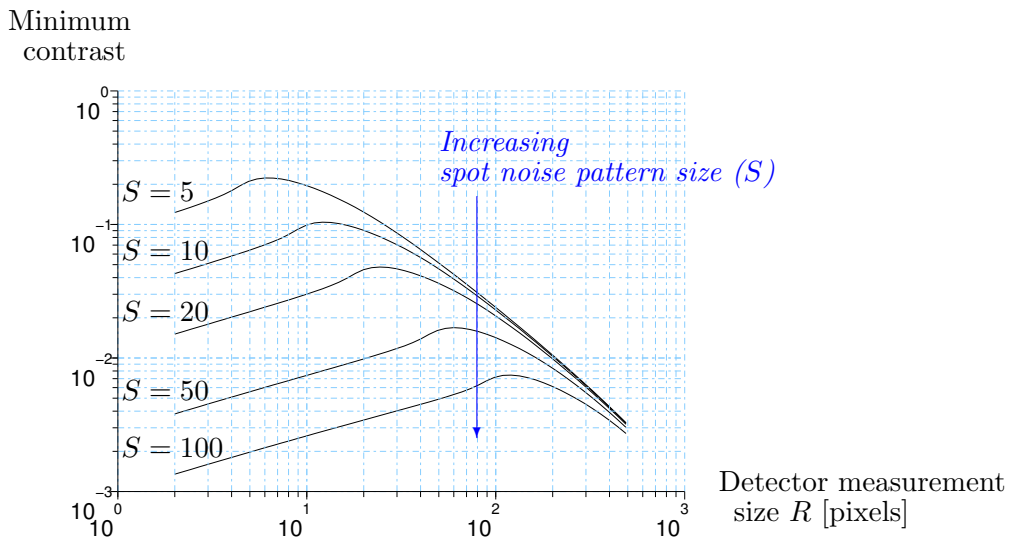
Figure 7: Contrast-detail (CD) curves with level $\varepsilon = 1$ for different pattern-noise textures. Generating pattern images consist in discs with fixed (unit) average grey level and various radius S (in pixels). The CD-curves have a "bell" shape, the maximum required contrast being reached when the measurement size R has the same order of magnitude than the radius of the pattern.

Fourier transform of u (since $\hat{k} = \hat{u}$ and $\hat{k} = |\hat{u}|$ define the same colored noise), or equivalently the autocorrelation of u by Wiener-Khintchine Theorem. Such a nonparametric texture model is fairly efficient for a texture without contour or alignment structure, and may deliver quite realistic sample images in that case. On Figure 8, we generated such texture samples for four different images: two Brodatz textures (see [4]), a picture of the see and a region of interest extracted from a mammogram. Their associated contrast-detail curves, also represented on this figure, can be classified in three types. First, we can notice that the Brodatz textures are strongly characterized by two different scales, respectively equal to about 3 and 10 pixels. For such textures, the contrast-detail curves are bell-shaped curves, like the ones already observed for the pattern-noise textures, meaning that the detection performance is degraded when the potential spot size is close to a particular critical size. For the sea picture, the contrast-detail curve appears to be monotone with a negative slope: the minimum measurement contrast decreases as the measurement kernel size increases. Indeed, for a given measured contrast, since this image is dominated by high frequencies, detection is all the more easy than the spot is large. On the contrary, the mammography image being dominated by lower frequencies, the detection performance globally falls down as the measurement kernel size increases, which means that for a given measured contrast, the larger the spot, the worse the detection. In these two last examples, we retrieve the same behavior as previously described for powerlaw textures, for which the contrast-detail curves were linear in a log-log scale.

As we can see, contrast-detail curves provide an intuitive way to understand the content of a texture in terms of relative contrast (grey levels) of structures with different sizes (pixels). The information in the average radial power spectrum would be essentially the same, but the CD-curve representation is much easier to interpret for pattern detection purposes.
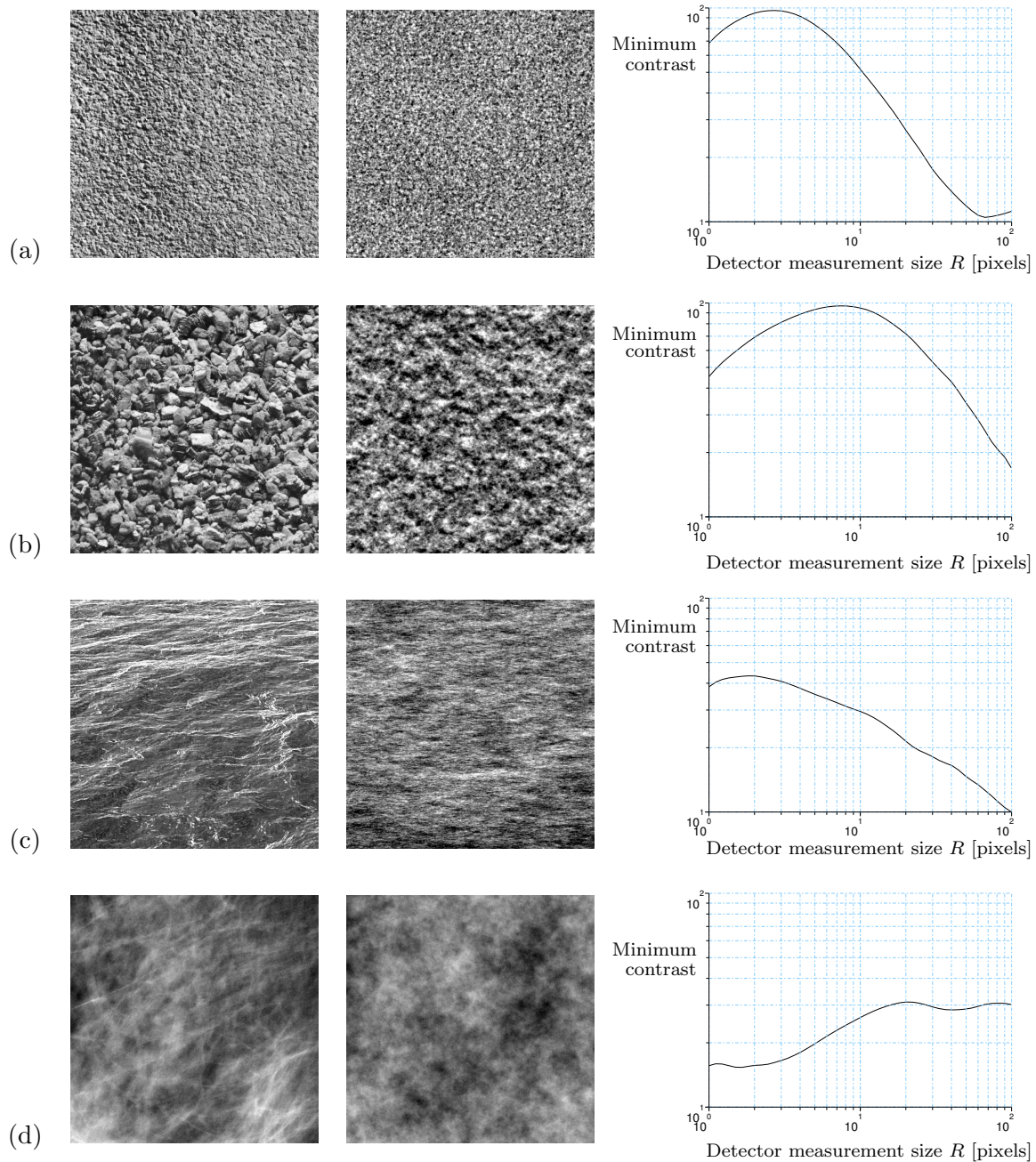
Figure 8: Left column: Real images of two Brodatz texture images (a,b), an image of the sea (c), and a 5 cm^2 region of interest extracted from a mammogram (d). Middle column: Samples of the associated nonparametric texture models obtained from their full power spectrum (see Section 4.5). Right column: CD-curves obtained by the a-contrario observer. For the textures with a strongly characteristic scale (a,b), the CD-curve reaches a maximum when the measurement size is similar to the texture characteristic scale. The sea texture (c) being dominated by higher frequencies, its CD-curve is monotone with a negative slope. On the contrary, the predominance of low frequencies in the mammography texture leads to a rather increasing CD-curve with a globally positive slope.

5   Application examples

5.1   Discrete framework

The a-contrario detection models presented in this study have been developed in a continuous framework, but all models and results of Section 3 remain valid in the discrete case, that is when the integrals defining convolutions and inner products are replaced with discrete sums. In particular, the γ constant involved in Proposition 7 still satisfies $\gamma = 1$ in the case of a Gaussian white noise texture when s_1 and s_2 are discrete averages on domains D_1 and D_2 such that $D_1 \subset D_2$ (Equation 18). However, the limit theorem 1 obtained for powerlaw textures (Section 4.2) is specific to the continuous setting (because scale invariance cannot be formulated on a discrete grid) and consequently discrete low-frequency powerlaw textures, as ordinary cases of general textures, will in general lead to $\gamma \neq 1$ (but close to 1, though). Therefore, the model with conditional context does not necessarily lead to a contrast measurement $m_1 - m_2$ for these textures. If, for a specific application, a rigorous contrast is required in the NFA expression, the detection model with contrast related to the context (model 2) could rather be used.

5.2   Lesion detectability in mammography images

In mammography images, two important kinds of lesions (masses and microcacifications) typically appear as spot-like structures. Hence, measuring the detectability of such kind of structures plays a key role in the optimization of the images devoted to human experts. The a-contrario detection models proposed in this paper naturally offer such a detectability measure. In this section, we propose to illustrate these models in two steps. First, we apply the NFA_2 criterion to detect spots in a real mammogram. As we said before, the goal of this study is not to build an automated detector of breast lesions as computer aided detection algorithms (CAD) do, but we may still want to check that the structures with a small NFA_2 correspond to the ones naturally catched by the human eye. Second, we synthesize several test images by putting artificial spots in real mammograms, and compare the detectability predicted by the NFA_2 criterion with our perception of these images.

5.2.1   Detection of real spots in mammography images

A 1500×2000 mammography image containing a low contrasted opacity with an approximately round shape is presented on Figure 9(a). The associated nonparametric breast texture model is defined as a colored noise matching the power spectrum of a region of interest (ROI) of the mammography image, located inside the breast area (see Section 4.5). As explained in the introduction, power spectra of mammography images approximately follow a powerlaw distribution. Thus, we could also have modeled the texture by a powerlaw process, whose parameters would be similarly estimated on a ROI inside the breast area. Note also that the model intends to represent the non-structured texture of actual mammograms, but not the breast curvilinear structures. This kind of highly structured details involve phase information that cannot be taken into account in a colored noise model (modeling these structures and understanding their influence on the detectability of spots is beyond the scope of this paper and would require a separate study). As well, the stationarity of the inside breast texture is a reasonable, though inexact, hypothesis.

27

(a) Original image   (b) $\varepsilon = 1$ (c) $\varepsilon = 10$
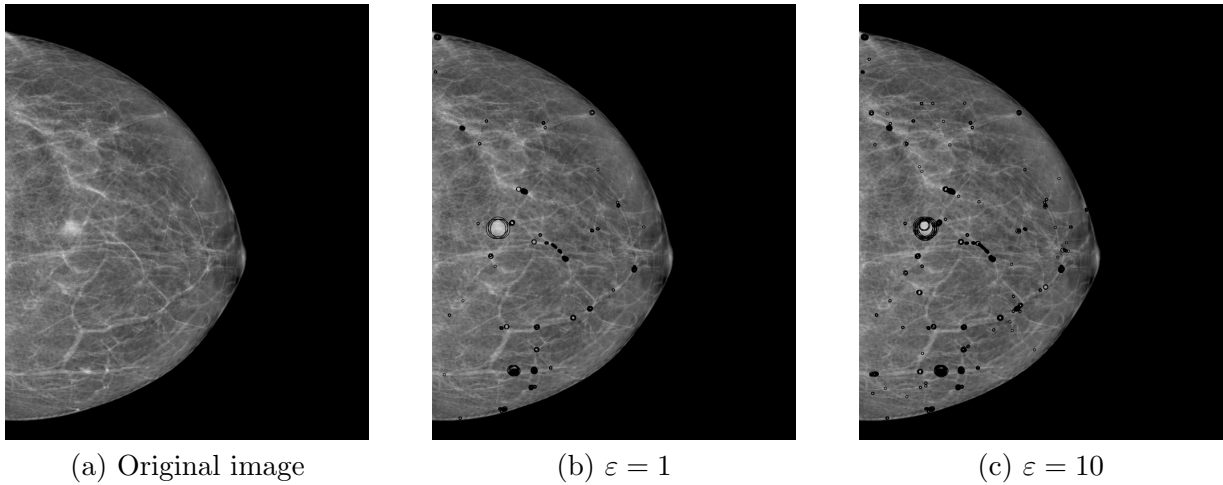
Figure 9: A-contrario detection of spots in a mammography image, for two values of the threshold ε applied to the detection metric NFA_2. The large opacity is well detected (its NFA_2 is equal to 0.15). Some clinically wrong detections also occur (small spots), mainly because the curvilinear breast structures are not taken into account by the texture model (these are false alarms clinically speaking, but are not with respect to the naive model used for the breast texture). The ε-dependency of the detection results is low: only a small amount of detections are added when the threshold is increased from 1 to 10.

The a-contrario detection model with contrast related to the context was used to detect spot-like lesions in the mammography image, with measurement kernel functions defined as normalized indicator function of discs. Five measurements sizes (radius R equal to 5, 10, 25, 50 and 100 pixels) and two context factors (equal to $\sqrt{2}$ and 2) were used, and the tests were sampled over a grid with step $\delta(R) = 0.3 \cdot R$ in order to ensure at least a 80% area overlapping of neighboring supports. At each tested location, we derived the NFA_2 associated to each combination of the object and context measurement sizes. Then, we selected the NFA_2 corresponding to the more significant detection (minimal NFA_2). The visible opacity is well detected, with a NFA_2 equal to 0.15. Smaller findings located along the breast curvilinear structures are found with smaller NFA_2 values. Indeed, since the breast curvilinear structures cannot be modeled by a colored noise (this kind of highly structured details involve phase information), it was expected that they will contradict the naive texture model hypothesis. Therefore, their detection is consistent with the detection model that intends to detect locally contrasted spots.

5.2.2 Detectability of simulated spots in real breast textures

We extracted from the mammogram represented in Figure 9(a) a 512×512 ROI corresponding to an area of homogeneous breast texture (this ROI is used on Figure 11). The contrast-detail curve obtained by the a-contrario observer for the detection of spots in this image is shown on Figure 10. The curve is found having (globally) a positive slope, meaning that larger spots require higher contrast $m_1 - m_2$ to be detected.

To illustrate this perception law, we added three simulated spots to the considered breast texture image with contrast and size values represented by points A, B and C of Figure 10. As shown on Figure 11, the two smallest spots are immediately noticed by the human eye, but
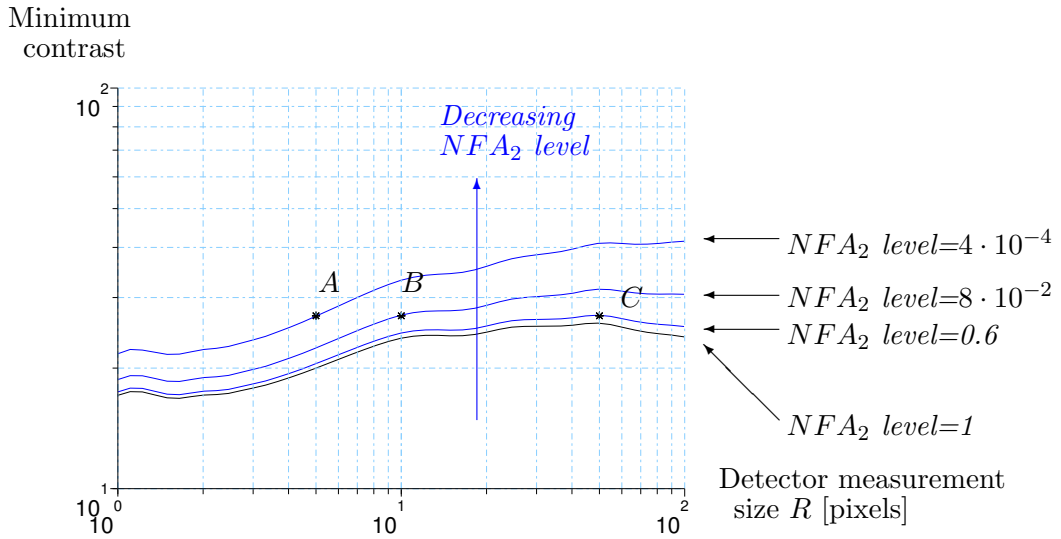
Figure 10: Contrast-detail curves with NFA_2 levels $4 \cdot 10^{-4}$, $8 \cdot 10^{-2}$, 0.6 and 1 obtained for a 512×512 ROI of the breast texture issued from the mammogram shown in Figure 9(a). The three points A, B and C correspond to spots with radius equal respectively to 5, 10 and 50 pixels, leading to the same contrast $m_1 - m_2$. At NFA_2 level 1 for instance, the lowest curve indicates that the spots associated to contrast-detail points A and B will be easily detected, whereas the one associated to C will be hard to detect (the higher the point with respect to the curve, the higher the detectability).

the larger one is really ambiguous. Notice that that these perception experiments should be realized on the electronic version of this paper (with the images displayed on a screen with sufficient resolution); on printed versions of this paper, the smallest spots may be difficult to see, especially on the non-zoomed images (top row). Notice also that the largest spot (right column) can be identified by comparing with the other images, but this comparison is not the visual task intended here (the task is absolute spot detection, not image comparison).

A similar experiment has been done in another breast texture image, extracted from the real mammogram of Figure 12(a). This texture is characterized by a power spectrum more dominated by lower frequencies than the previous one. Again, the contrast-detail curve is found having a positive slope (see Figure 13) but increasing more rapidly than the one found in Figure 10. Examples of images containing spots associated to the contrast-detail points A and B are shown on Figure 12. As predicted by our model, the smaller spot is highly visible in the texture image. On the contrary, the larger one is not easy to distinguish from the underlying breast structure.

These two experiments alone are not sufficient to prove that the proposed NFA_2 model yields a good prediction of the human ability to detect spots on mammographic backgrounds, but they are consistent with the idea that human CD-curves have a positive slope on such backgrounds. This property has been demonstrated by Burgess in [7] by evaluating human CD-curves from several psychovisual experiments. On powerlaw texture backgrounds, a good fit between the human observer and the a-contrario observer (NFA_2 model) has been observed as well [15].

(a) R=5 (NFA_2=1·10^{-5})   (b) R=10 (NFA_2=1.9·10^{-3})   (c) R=50 (NFA_2=0.8)

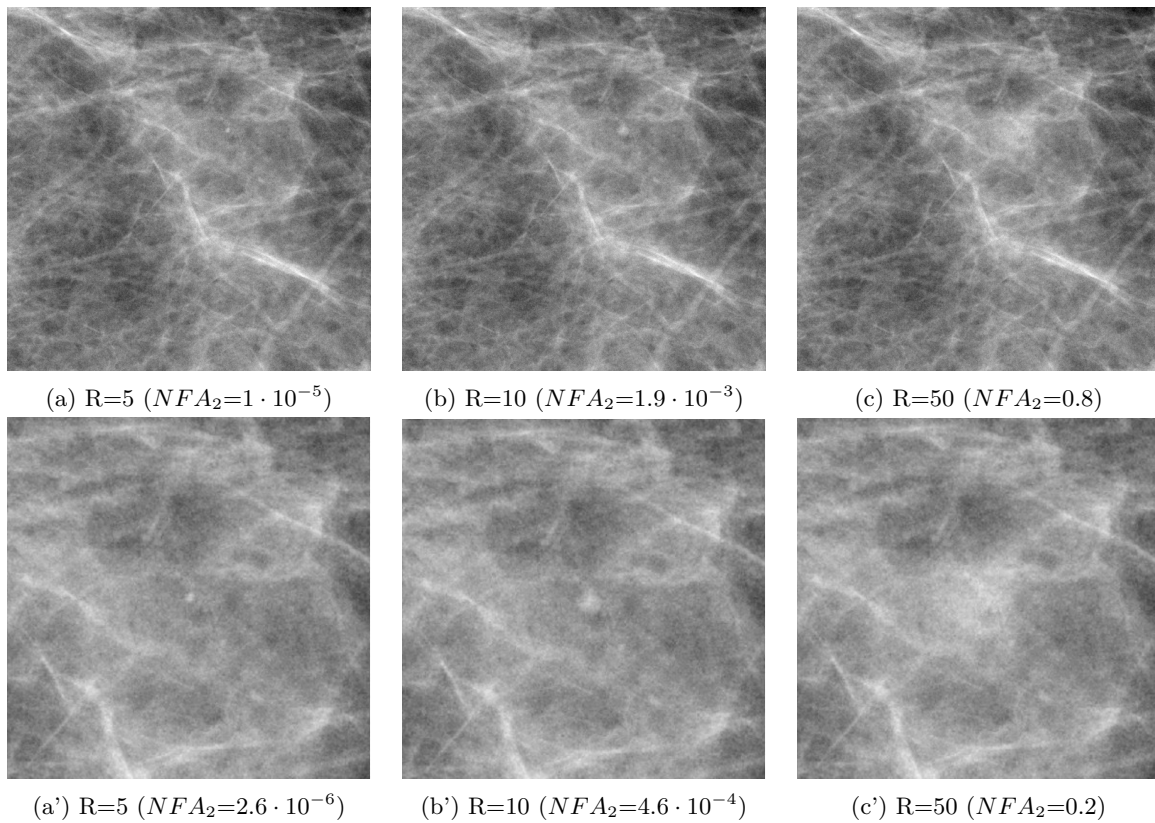(a') R=5 (NFA_2=2.6·10^{-6})   (b') R=10 (NFA_2=4.6·10^{-4})   (c') R=50 (NFA_2=0.2)

Figure 11: Examples of simulated spots added to a real breast texture. Bottom images are extracted from top images (hence, the NFA_2, like the number of tests, decreases). The considered spots are characterized by various radius sizes (R equal to 5, 10 and 50 pixels) but give rise to the same contrast $m_1 - m_2$ in the image. Whereas the two smaller spots can be detected, the larger one is less visible due to the presence of the surrounding texture. It should be noted that when looking at these images, the observer tends to detect spots by comparing the different images. To properly estimate the absolute visibility of the spot, the observer should look at a single image at a time, without comparing the observed image to the others.

[This area has been left blank for page layout purposes]

(a) Original mammogram

(b) R=5 (NFA_2= $6.8 \cdot 10^{-38}$)

(b') R=5 (NFA_2= $1.7 \cdot 10^{-38}$)
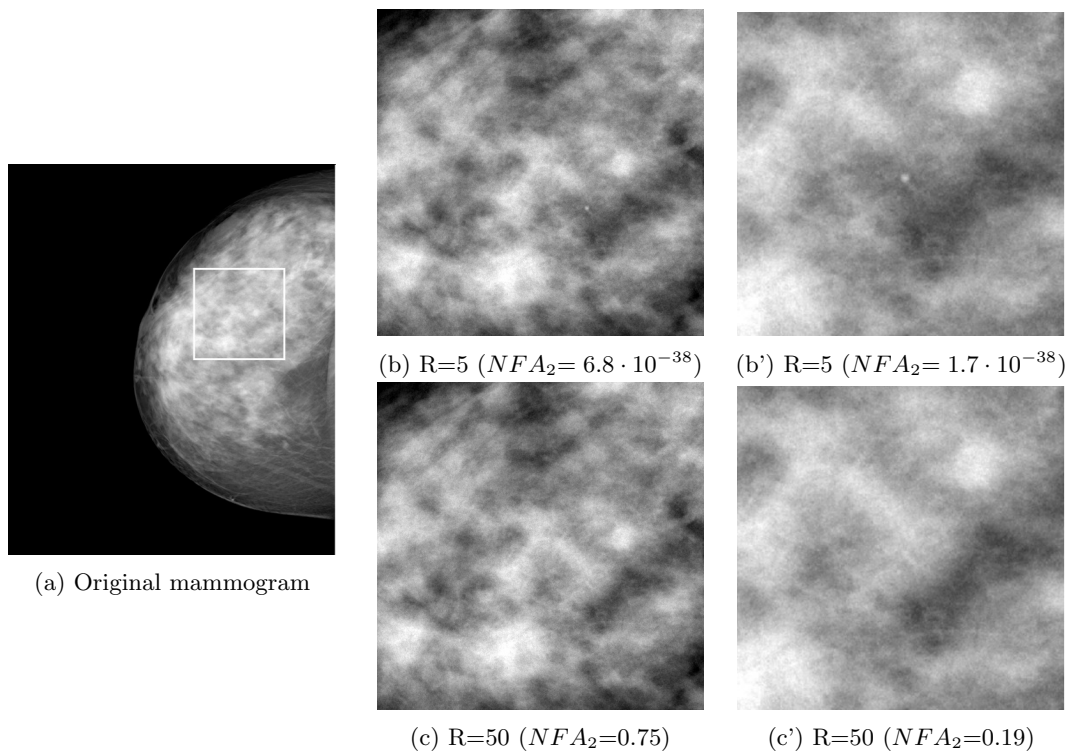
(c) R=50 (NFA_2=0.75)

(c') R=50 (NFA_2=0.19)

Figure 12: Simulated spots are added to a breast texture image extracted from (a). The spots have the same contrast, and a radius R equal to 5 (b,b') or 50 pixels (c,c'). Images represented in (b',c') are extracted from (b,c) respectively. Whereas the smaller spot is clearly visible, the larger one may not be distinguished from the background texture.
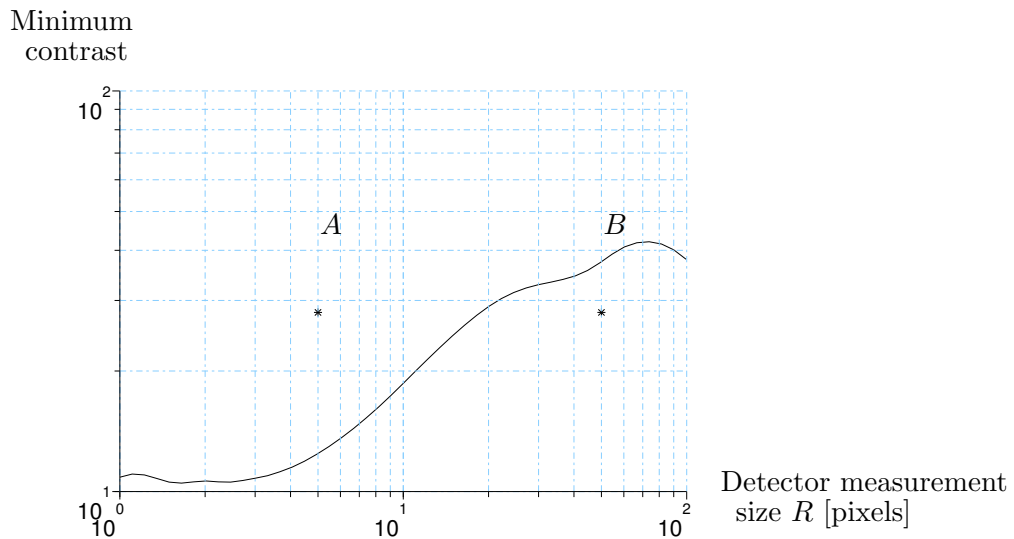

Figure 13: CD-curve with NFA_2 level 1 corresponding to the ROI highlighted in image 12(a). Points A and B correspond to the spots added on image 12 (b,b',c,c').

31

5.3   Stain detectability in textile images

We now present a second example where the additive spot hypothesis makes sense, the detection of stains (or manufacturing defects) in a piece of clothing. As for the mammography image, we realize two experiments on a wool texture image: the detection of actual stains and the measure of the detectability of an artificial spot superimposed to the original texture. In both cases, we compare the NFA_2 criterion with human perception.

5.3.1   Detection in textile images

We consider a 2048×2048 pixels image (shown in Figure 14(a)) extracted from a picture of wool textile. Figure 14(b) represents an image sample of its associated colored noise process. Some differences can be observed between the initial image and the image sample of the texture model, due to the lack of phase information in the model. The image power spectrum, used to generate the texture model, does not capture the global grid structure information. Furthermore, the power spectrum approach implies a contrast symmetry $u \mapsto -u$ in the texture image sample, which is not verified in the initial image (both white and black tiny spots can be distinguished in the sample image, whereas the real image contains mostly black ones). However, as we shall see later, the colored noise model seems to capture enough texture information to build a spot detection criterion that mimics human vision.
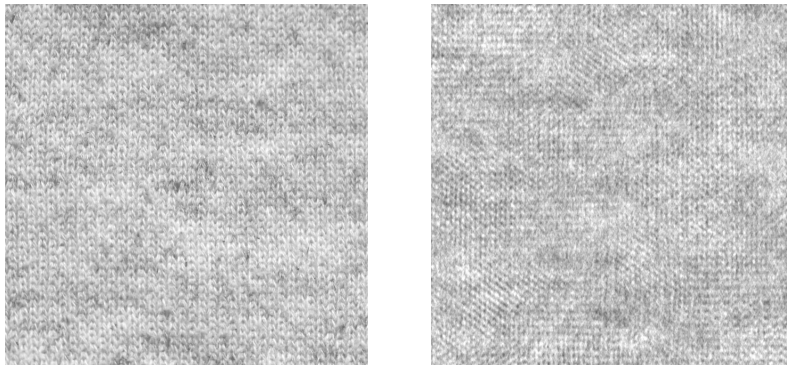

Figure 14: Left: image extracted from a picture of wool textile. Right: a sample of its associated texture model. Some aspects of the initial image are not represented in the texture model, such as the deterministic grid effects, since the image power spectrum does not catch this information.

Images have been scored by the same a-contrario observer as in the mammography example, but with larger potential stains sizes (see Figure 15). We tested both black stains and white ones, but found no detection for white stains even for a NFA_2 threshold equal to 10 (that is why this case in not illustrated in Figure 15). This seems visually consistent with the fact that we do not notice any particular white stain in this image. On the contrary, concerning black stains, the more meaningful detection occurs unsurprisingly at the image center. As the NFA_2 threshold is increased, new detected stains appear. They do not pop out as evidently as the stain located near the image center, but are undoubtfully perceived when our eye focuses on image subwindows. This focusing property is well modeled by a-contrario models, as noticed in [9] in the case of edge detection. Here, when considering a given potential stain in a sub-region

32

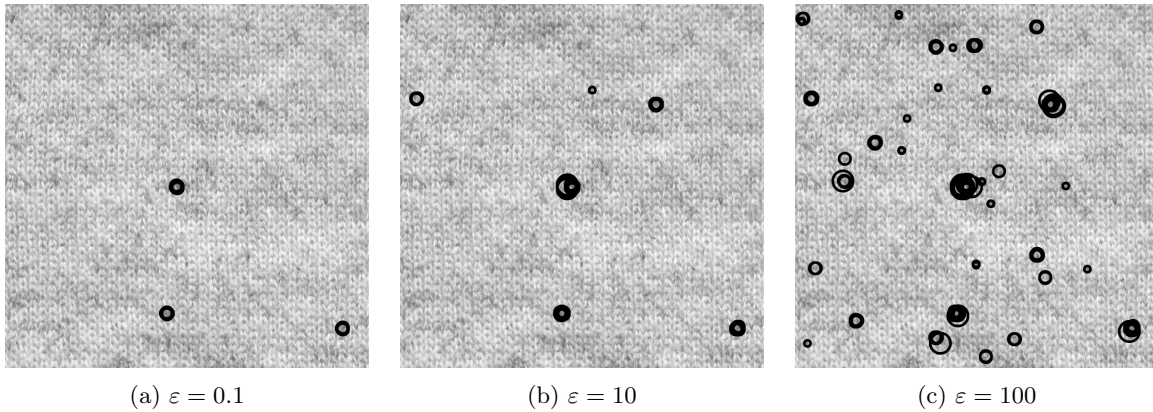(a) $\varepsilon = 0.1$ (b) $\varepsilon = 10$ (c) $\varepsilon = 100$

Figure 15: A-contrario detection of spots in the textile image (Figure 14, left), with 5 measurements sizes (radius R equal to 10, 25, 50, 100 and 200 pixels), 2 context factors (equal to $\sqrt{2}$ and 2), and a test sampling on a grid with step $\delta(R) = 0.3 \cdot R$ (like in Figure 9). A black stain, actually visually suspicious, is detected at the image center (a). When the NFA_2 threshold is increased, more stains are highlighted (b,c). These stains become visible if the eye focuses on a particular ROI of the image, and they would be detected as well by the focused a-contrario model.

of the full image, the NFA_2 derived by the a-contrario observer decreases because the number of tests is decreased. A potential stain with a NFA_2 equal to 100 in the full image would have a NFA_2 equal to 1 in a 0.1×0.1 ROI of the initial image. Thus, the a-contrario observer models the fact that human observers can focus their attention on a restricted ROI.

5.3.2 Detectability of simulated spots in textile images

The detectability of spots in textile textures is studied here in a 512×512 ROI of the whole piece of clothing. As shown in Figure 16, the minimal contrast needed to make a detection reaches a maximum for a measurement size equal to about 15 pixels. This critical size corresponds to the size of a typical texton[4] in the initial image: as structures with this scale naturally appear more often in the image, the contrast needed to detect them is increased. Furthermore, the bell shape of the detection curve indicates that the more the searched spot size departs from the texton size, the easier the detection. We retrieve here a detectability behavior similar to the detection of spots in the images of pattern noise associated to disc patterns.

Simulated spots with contrast and size values corresponding to points A, B, C and D of Figure 16 have been added to the wool texture image, as shown in Figure 17. As predicted by the contrast-detail curve, the spot with a 15-pixels radius is the less detectable one. Indeed, it corresponds to the typical "texton" size of the texture image and thus can be hardly distinguished from random spots due to normal texture. Conversely, the 100-pixels radius spot is the one which pops out the most evidently.

---

[4]We call texton the smallest basic pattern that is repeated (with variations) to produce the texture, when it exists.

Minimum
contrast

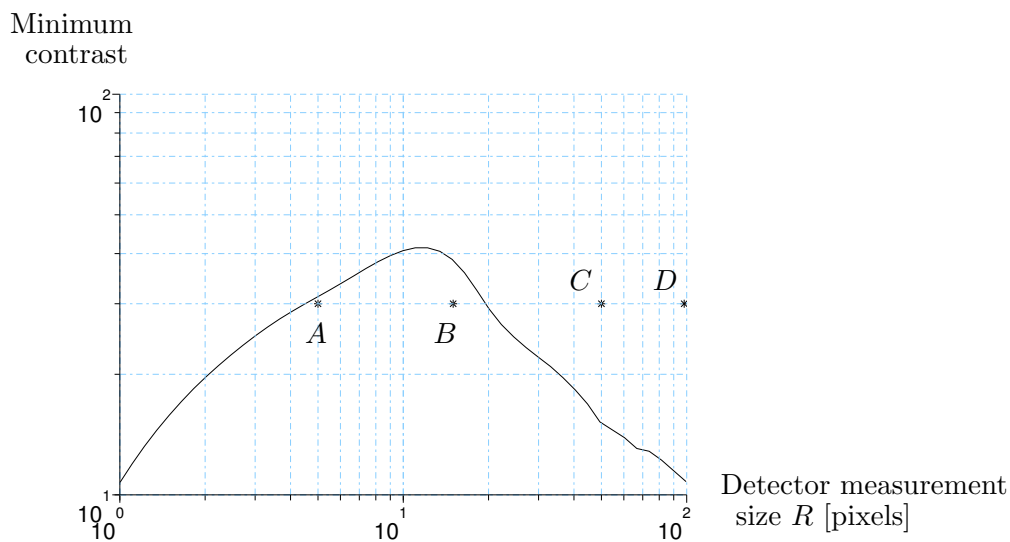
Detector measurement
size R [pixels]

Figure 16: CD-curve with NFA_2 level 1 obtained for the wool texture image. The minimum contrast takes its highest value at a measurement scale of about 15 pixels, corresponding to the typical "texton" size. Points A, B, C, and D correspond to the points added on Figure 17.

(a) R=5 (NFA_2=0.98)   (b) R=15 (NFA_2=145)   (c) R=50 (NFA_2=8.3·10^{-3})

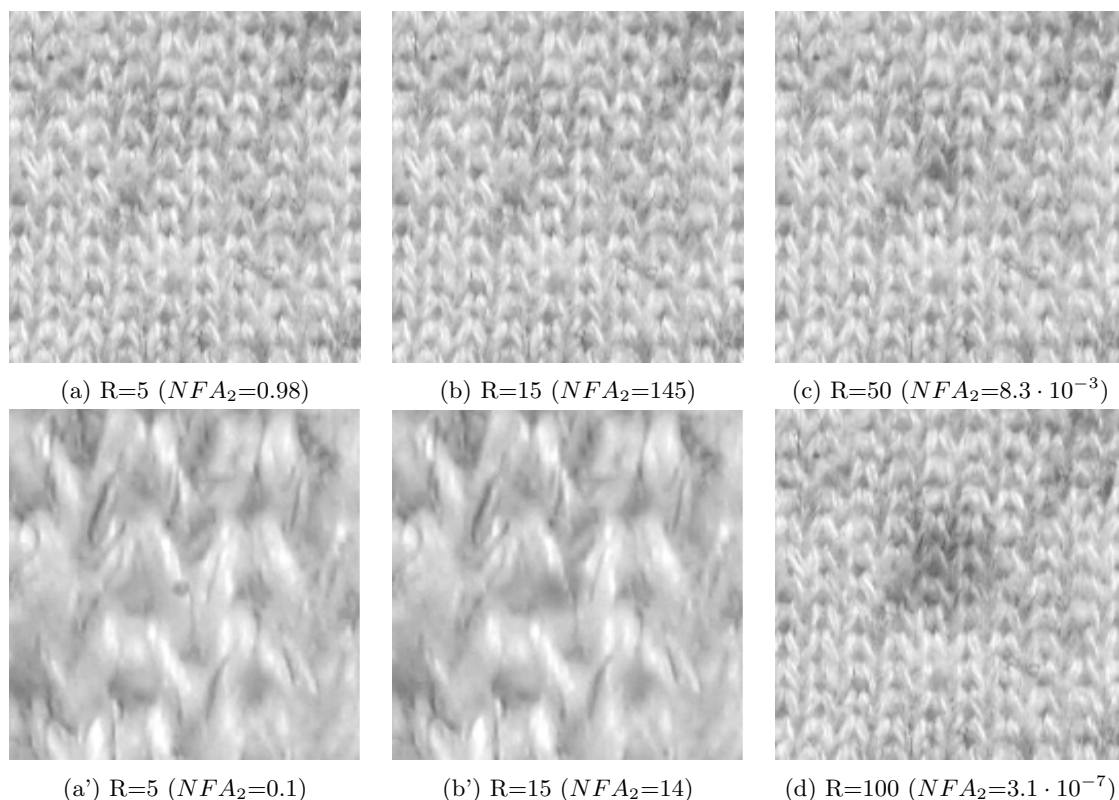(a') R=5 (NFA_2=0.1)   (b') R=15 (NFA_2=14)   (d) R=100 (NFA_2=3.1·10^{-7})

Figure 17: Examples of spots added to a wool textile image. The spots have the same contrast, but a different radius R. Images in (a',b') are zoomed from (a,b) respectively. The less visible spot has a radius of 15 pixels (image (b)), which corresponds to the scale of the image textons.

6 Conclusion

In classical statistical approaches of signal detection, both signal and background are described by a probabilistic model, and inference is made by deciding which model explains better the observed data. In the a-contrario framework, the approach is quite different. First, the structure to detect is not described by a probabilistic model but through one or several measurement functions generally inspired from Gestalt Theory. It relies on the idea that since the variability of potential structures is tremendously large, a description in terms of Gestalt grouping laws is more appropriate to mimic human vision than a complicated weighting of potentially observable structures. Second, the background model plays a different role in the two frameworks, even if it is described by a probability measure (generally written H_0) in both cases. In signal detection theory, the background model directly competes with a signal model (H_1), which requires a careful distinction between background and signal when the probabilistic models are established. In the a-contrario framework, the choice of the "naive model" does not change fundamentally the way structures are detected (the measurements above), but essentially drives two detection issues: the way the measurements and the detection parameters are structurally combined to form a single NFA measure, and the link between the detection threshold and the control of false alarms. When the naive model is a simplistic description of observed backgrounds, the corresponding NFA will generally yield an unprecise control of real false alarms, but may still remain a usable detection criterion. This property is illustrated in the case of powerlaw textures by the fact that $\gamma = 1$ independently of β (when $2 \leq \beta < 6$) for the model with conditional context (limit of NFA_2).

In this paper, we proposed a new way to model the human ability to detect "spots" in textured images. This led us to generalize in some ways the a-contrario framework previously developed in the literature: we gave a precise definition of a Number of False Alarms (NFA), and provided an easy way to build such a NFA, leaving the possibility of using arbitrary weights on the different tests (Proposition 2). We then developed three detection models dealing with the unusual case of a non i.i.d naive model, and, for the last one, introduced the notion of conditional context. This notion of context was proven to be essential for low-frequency textures, since the model without context collapses for powerlaw textures with $\beta \geq 2$ (Theorem 1). The same critical exponent ($\beta = 2$), establishing a boundary between high-frequency powerlaw textures ($\beta < 2$) and low-frequency ones ($\beta \geq 2$), was found when exploring the other context-dependent models. This corresponds exactly to the reversal of the contrast-size dependency for the human perception of spots, noticed on Figure 2. More generally, for an arbitrary microtexture the second a-contrario model we proposed (contrast to the context) allowed us to define a contrast-detail curve (CD-curve) that links the size of a spot (or, to be more accurate, the detector size) and its minimal perception contrast. From this point of view, real-world textures can be roughly classified in three types (see Figure 8): high-frequency textures corresponding to a decreasing CD-curve (e.g. white noise), low-frequency corresponding to an increasing CD-curve (e.g. mammographic textures), and pattern noise corresponding to a bell-shaped CD-curve (e.g. field of stones). More complicated CD-curves are also possible, though rarely encountered.

Concerning experiments, we illustrated the a-contrario models of spot detectability by providing a few examples of real or artificial spots on a mammographic image and on a wool texture image. On these images, there seems to be a good agreement between our relative ability to detect spots and the contrast-detail curves computed with the a-contrario model, even though the background images are quite more complex than the colored noise processes used as naive

35

models.

We believe that the a-contrario observer we proposed could provide a potential alternative to classical signal detection approaches to model the human perception of spots. It could be used in order to optimize the images devoted to human experts, as it is the case for mammography images for example. More generally, the ability of such models to mimic human vision in a framework compatible with gestalt principles could be useful to measure the detectability of other kinds of structures in textured images, in particular geometrical grouping laws like alignments or edges, already considered in white noise textures [8, 9]. A generalization of the present work to non-stationnary textures could also reveal interesting connexions with Constant False Alarm Rate techniques (CFAR, see [14, 22] for example) that have been popular for a long time in radar imaging.

Acknowledgments

7 Appendix

7.1 Lemma 1

Lemma 1 (distribution of p-values) *Let X be a real-valued random variable. Define $F(t) = \mathbb{P}(X \geq t)$. Then*

$$\forall s \in [0,1], \quad \mathbb{P}(F(X) \leq s) \leq s. \tag{39}$$

Proof: First note that F is left-continuous and nonincreasing. Consider a real number y, and define

$$z = \inf\{x, \ F(x) = F(y)\} \in [-\infty, y].$$

First assume that $z \neq -\infty$. Two possibilities arise.

1) Either $F(z) = F(y)$, then $F(x) \leq F(y) \iff x \geq z$ and

$$\mathbb{P}(F(X) \leq F(y)) = \mathbb{P}(X \geq z) = F(z) = F(y).$$

2) Or $F(z) > F(y)$, and $F(x) \leq F(y) \iff x > z$, so that

$$\mathbb{P}(F(X) \leq F(y)) = \mathbb{P}(X > z) = \lim_{\varepsilon \to 0^+} \mathbb{P}(X \geq z + \varepsilon) = \lim_{\varepsilon \to 0^+} F(z + \varepsilon) \leq F(y).$$

Now if $z = -\infty$, necessarily $F(y) = 1$ and as in case 1) we have $\mathbb{P}(F(X) \leq F(y)) = F(y)$. Hence, for real $y \in \mathbb{R}$ we have $\mathbb{P}(F(X) \leq F(y)) \leq F(y)$. Now if $s \in [0,1]$, we have, since F is nonincreasing,

$$\mathbb{P}(F(X) \leq s) = \sup_{y, \ F(y) \leq s} \mathbb{P}(F(X) \leq F(y)) = \sup_{y, \ F(y) \leq s} F(y) \leq s,$$

which ends the proof. Note that if X admits a probability density function, then $\mathbb{P}(F(X) \leq s) = s$ for all s. $\square$

7.2 Lemma 2

Lemma 2 (distribution of conditional p-values) *Let Y, Z be two real-valued random variables. Assume that there exists some set $U \subset \mathbb{R}$ with $\mathbb{P}(Z \in U) = 1$ for which the function $F(t, u) = \mathbb{P}(Y \geq t | Z = u)$ is well-defined on $\mathbb{R} \times U$. Then*

$$\forall s \in [0, 1], \quad \mathbb{P}(F(Y, Z) \leq s) \leq s. \tag{40}$$

Proof: For any $u \in U$, the random variable $X = Y | (Z = u)$ is well-defined by its survival function $\mathbb{P}(X \geq t) = F(t, u)$. Hence, from Lemma 1 we get

$$\forall \in [0, 1], \quad \mathbb{P}\left(F(X, u) \leq s\right) \leq s.$$

This implies that

$$G(u) := \mathbb{P}(F(Y, Z) \leq s | Z = u) = \mathbb{P}(F(X, u) \leq s) \leq s,$$

and as a consequence

$$\mathbb{P}(F(Y, Z) \leq s) = \mathbb{E}_Z G(Z) \leq s,$$

which is the announced result.

7.3 Lemma 3

Lemma 3 *Let $\varphi_1, \varphi_2 \in L^2(\mathbb{R}^2)$ such that $\|\varphi_2\| \neq 0$, and N a normalized two-dimensional Gaussian white noise. Then we have the following conditional distribution:*

$$\forall t \in \mathbb{R}, \quad \langle N, \varphi_1 \rangle \mid \left(\langle N, \varphi_2 \rangle = t\right) \sim \mathcal{N}\left(\gamma t, \|\varphi_1 - \gamma \varphi_2\|^2\right), \quad \text{with} \quad \gamma = \frac{\langle \varphi_1, \varphi_2 \rangle}{\|\varphi_2\|^2}.$$

Proof: Conditionally to $\langle N, \varphi_2 \rangle = t$, we have

$$\langle N, \varphi_1 \rangle = \langle N, \varphi_1 \rangle - \gamma \langle N, \varphi_2 \rangle + \gamma \langle N, \varphi_2 \rangle = \langle N, \varphi_1 - \gamma \varphi_2 \rangle + \gamma \langle N, \varphi_2 \rangle.$$

Now since

$$\langle \varphi_1 - \gamma \varphi_2, \varphi_2 \rangle = \gamma \|\varphi_2\|^2 - \gamma \|\varphi_2\|^2 = 0,$$

the random variables $\langle N, \varphi_1 - \gamma \varphi_2 \rangle$ and $\langle N, \varphi_2 \rangle$ are independent and we have

$$\left(\langle N, \varphi_1 \rangle \mid \langle N, \varphi_2 \rangle = t\right) = \langle N, \varphi_1 - \gamma \varphi_2 \rangle + \gamma t \sim \mathcal{N}\left(\gamma t, \|\varphi_1 - \gamma \varphi_2\|^2\right).$$

7.4   Lemma 4

Lemma 4 *With the notations of Theorem 2, one has*

$$
\gamma^\varepsilon := \frac{\left\langle \varphi_{1,R}^\varepsilon, \varphi_{2,R}^\varepsilon \right\rangle}{\|\varphi_{2,R}^\varepsilon\|^2} =
\begin{cases}
\bar{\gamma} + \underset{\varepsilon\to 0}{o}(1) & \text{if } 0 \le \beta < 2, \\[2mm]
1 + \underset{\varepsilon\to 0}{O}\left(\dfrac{1}{\log \varepsilon}\right) & \text{if } \beta = 2, \\[2mm]
1 + \underset{\varepsilon\to 0}{O}\left(\varepsilon^{\beta-2}\right) & \text{if } 2 < \beta < 4, \\[2mm]
1 + \underset{\varepsilon\to 0}{O}\left(\varepsilon^2 \cdot \log \varepsilon\right) & \text{if } \beta = 4, \\[2mm]
1 + \underset{\varepsilon\to 0}{O}\left(\varepsilon^2\right) & \text{if } 4 < \beta.
\end{cases}
$$

Proof: As in (26), the scaling properties of $\widehat{s_1}$ and $\widehat{s_2}$ with respect to R lead to

$$
\gamma^\varepsilon = \frac{\left\langle \varphi_{1,R}^\varepsilon, \varphi_{2,R}^\varepsilon \right\rangle}{\|\varphi_{2,R}^\varepsilon\|^2} = \frac{A_\varepsilon}{B_\varepsilon}
$$

with $\quad A_\varepsilon = \displaystyle\int_{|\xi|\ge R\varepsilon} \widehat{s_1}(\xi)\cdot\widehat{s_2}(\xi)\,|\xi|^{-\beta}\,d\xi \quad$ and $\quad B_\varepsilon = \displaystyle\int_{|\xi|\ge R\varepsilon} |\widehat{s_2}(\xi)|^2\,|\xi|^{-\beta}\,d\xi$

(note that in all the following we suppose that ε is small enough to ensure that $B_\varepsilon \neq 0$).

• If $0 \le \beta < 2$, both integrals A_0 and B_0 are finite, so that by Lebesgue's monotone convergence Theorem we have the existence of a limit (say $\bar{\gamma}$) to γ^ε when $\varepsilon \to 0$, which ends the proof for this case.

• Now we assume that $\beta \ge 2$. Since $\widehat{s_2}(0) = 1$, we have $|\widehat{s_2}(\xi)|^2\,|\xi|^{-\beta} \underset{\xi\to 0}{\sim} |\xi|^{-\beta}$, and since the integral of $|\xi|^{-\beta}$ diverges around 0,

$$
B_\varepsilon \underset{\varepsilon\to 0}{\sim} 2\pi \int_{R\varepsilon}^1 t^{-\beta+1}\,dt \underset{\varepsilon\to 0}{\sim}
\begin{cases}
\frac{2\pi R^{2-\beta}}{\beta-2}\varepsilon^{2-\beta} & \text{if } \beta > 2, \\[1mm]
2\pi \log \varepsilon & \text{if } \beta = 2.
\end{cases}
\tag{41}
$$

Now since $\widehat{s_2}$ is real,

$$
\gamma^\varepsilon = 1 + \frac{A_\varepsilon'}{B_\varepsilon} \quad \text{with} \quad A_\varepsilon' = \int_{|\xi|\ge R\varepsilon}\left(\widehat{s_1}(\xi) - \widehat{s_2}(\xi)\right)\cdot\widehat{s_2}(\xi)\,|\xi|^{-\beta}\,d\xi,
\tag{42}
$$

and since $\widehat{s_1}$ and $\widehat{s_2}$ are even functions taking the same value (1) in $\xi = 0$, a Taylor expansion yields as in (29)

$$
\widehat{s_1}(\xi) - \widehat{s_2}(\xi) = \underset{\xi\to 0}{O}\left(|\xi|^2\right), \quad \text{so that}
$$

$$
A_\varepsilon' = \underset{\varepsilon\to 0}{O}\left(\int_{R\varepsilon\le|\xi|\le 1}|\xi|^{2-\beta}\,d\xi\right) =
\begin{cases}
\underset{\varepsilon\to 0}{O}(1) & \text{if } \beta < 4 \\[1mm]
\underset{\varepsilon\to 0}{O}(\log\varepsilon) & \text{if } \beta = 4 \\[1mm]
\underset{\varepsilon\to 0}{O}\left(\varepsilon^{4-\beta}\right) & \text{if } 4 < \beta
\end{cases}
\tag{43}
$$

The remaining of the proof is obvious considering the combination of possibilities for B_ε and A_ε' (equations 41 and 43). $\square$

References

[1] L. Alvarez, Y. Gousseau, and J.-M. Morel. The size of objects in natural and artificial images. *Advances in Imaging and Electron Physics*, 111:167–242, 1999.

[2] H.H. Barrett, J. Yao, J.P. Rolland, and K.J. Myers. Model observers for assessment of image quality. *Proc. Natl. Acad. Sci. USA*, 90:9758–9765, 1993.

[3] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B*, 57:289–300, 1995.

[4] P. Brodatz. *Textures: A Photographic Album for Artists and Designers*. Dover, New York, 1966.

[5] A.E. Burgess. Mammographic structures: Data preparation and spatial statistics analysis. In *Medical Imaging 1998, Image Processing, San Diego, CA, edited by K. Hanson (Proceedings of the Society of Photo-optical Instrumentation Engineers, Bellingham, WA, 1999)*, 3661:642–653, 1998.

[6] A.E. Burgess, F.L. Jacobson, and P.F. Judy. Human observer detection experiments with mammograms and power-law noise. *Medical Physics*, 28(4):419–437, 2001.

[7] A.E. Burgess, F.L. Jacobson, and P.F. Judy. Lesion detection in digital mammograms. In *Proceedings of SPIE, Medical Imaging*, 4320:555–560, 2001.

[8] A. Desolneux, L. Moisan, and J.-M. Morel. Meaningful alignements. *International Journal of Computer Vision*, 40(1):7–23, 2000.

[9] A. Desolneux, L. Moisan, and J.-M. Morel. Edge detection by Helmholtz principle. *Journal of Mathematical Imaging and Vision*, 14:271–284, 2001.

[10] A. Desolneux, L. Moisan, and J.-M. Morel. Computational Gestalts and Perception Thresholds. *Journal of Physiology* 97(2-3):311-324, Paris, special issue on neurogeometry and visual perception, 2003.

[11] A. Desolneux, L. Moisan, and J.-M. Morel. *From Gestalt Theory to Image Analysis - A Probabilistic Approach*. Springer-Verlag, collection "Interdisciplinary Applied Mathematics", vol. 34, 2008.

[12] F. Dibos, S. Pelletier, G. Koepfler Real-time segmentation of moving objects in a video sequence by a contrario detection. Proceedings of *Int. Conf. Image Processing* 1:1065-1068, 2005.

[13] M.P. Eckstein, C.K. Abbey, and F.O. Bochud. *A practical guide to model observers for visual detection in synthetic and natural noisy images*, chapter 10, pages 593–628. Handbook of Medical Imaging Physics and Psychophysics, J. Beutel, H.L. Kundel, R.L. Van Metter, ed., SPIE Press, Bellingham, WA, 2000.

[14] P.P. Gandhi and S.A. Kassam. Analysis of CFAR Processors in Nonhomogeneous Background. *IEEE Transactions on Aerospace and Electronic Systems* 24(4):427-445, 1988.

[15] B. GROSJEAN, S. MULLER and H. SOUCHAY. Lesion detection using an a-contrario detector in simulated digital mammograms. In *Proceedings of SPIE, Medical Imaging* 6146:216–227, 2006.

[16] J.J. HEINE and R.P. VELTHUIZEN. Spectral analysis of full field digital mammography data. *Medical Physics*, 29(5):647–661, 2002.

[17] G. KANIZSA. *Grammatica del Vedere, Traduction: La grammaire du voir, Diderot Editeur, Arts et Sciences, 1996.* Il Mulino, Bologna, 1980.

[18] M. KAY. *Fundamentals of statistical signal processing, Volume II, Detection theory.* Prentice Hall, 1998.

[19] P. KHURD and G. GINDI. Decision Strategies That Maximize the Area Under the LROC Curve. *IEEE Transactions on Medical Imaging* 24(12):1626–1636, 2005.

[20] C.J. KOTRE. The effect of background structure on the detection of low contrast objects in mammography. *The British Journal of Radiology*, 71:1162–1167, 1998.

[21] L. MOISAN and B. STIVAL. A probabilistic criterion to detect rigid point matches between two images and estimate the fundamental matrix. *International Journal of Computer Vision*, 57(3):201–218, 2004.

[22] F.C. ROBEY, D.R. FUHRMANN, E.J. KELLY and R. NITZBERG. A CFAR adaptive matched filter detector. *IEEE Transactions on Aerospace and Electronic Systems*, 28(1):208–216, 1992.

[23] A. ROSE. Sensitivity performance of the human eye on an absolute scale. *Journal of optometry society of* America, 38:196–208, 1948.

[24] D. RUDERMAN. Origins of scaling in natural images. *Vision Research*, pages 814–817, 1996.

[25] R.G. SWENSSON. Using localization data from image interpretation to improve estimates of performance accuracy. *Med. Decision Making* 20:170-185, 2000.

[26] S. TILIE, L. LABORELLI and I. BLOCH. A contrario False Alarms Removal for Improving Blotch Detection in Digitized Films Restoration Proceedings of *6th EURASIP Conference focused on Speech and Image Processing, Multimedia Communications and Services - IWSSIP*, pages 410-413, 2007.

[27] H.L. VAN TREES. *Detection, estimation and modulation theory, Vol.1.* Ed. John Wiley and Sons, 1968.

[28] T. VEIT, F. CAO and P. BOUTEMY. An a contrario Decision Framework for Region-Based Motion Detection. *International Journal of Computer Vision* 68(2):163-178, 2006.

[29] F. XUE, Q. LIU and J. FROMENT. An a Contrario Approach for Parameters Estimation of a Motion-Blurred Image. In *Lecture Notes in Computer Science* 4679:267-279, 2007.