An anisotropic a contrario framework for the detection of convergences in images

Agnès Desolneux · Fanny Doré

Abstract The *a contrario* framework for the detection of convergences in an image consists in counting, for each tested point, the number of elementary linear structures that converge to it (up to a given precision), and when this number is high enough, the point is declared to be a meaningful point of convergence. This is so far analogous to a Hough transform, and the main contribution of the *a contrario* framework is to provide a statistical definition of what "high enough" means: it means large enough to ensure that in an image where all elementary structures are distributed according to a background noise model, there is, in expectation, less than 1 detection. Our aim in this paper is to discuss, from a methodological viewpoint, the choice and the influence of the background noise model. This model is generally taken as the uniform independent distribution on elementary linear structures, and here we discuss the case of images that have a natural anisotropic distribution of structures. Our motivating example is the one of mammograms in which we would like to detect stellate patterns (that appear as local convergences of spicules), and in which the linear structures are naturally oriented towards the nipple. In this paper, we show how to tackle the two problems of: (a) defining and estimating an anisotropic "normal" distribution from an image, and of (b) computing the probability that a random structure, following an anisotropic distribution, converges to any given convex region. We illustrate the whole approach with several examples.

Keywords *a contrario* methodology · points of convergence in images · stellate lesions in mammograms · stochastic geometry · estimation of a parametric model

A. Desolneux
CMLA (CNRS, ENS Cachan, Université Paris-Saclay)
61 avenue du Président Wilson,
94235 Cachan Cedex - France
E-mail: agnes.desolneux@cmla.ens-cachan.fr

F. Doré
MAP5 (Université Paris Descartes)
45 rue des Saints-Pères,
75270 Paris Cedex 06 - France
E-mail: fanny.dore@parisdescartes.fr

1 Introduction

There are many types of convergences of linear structures in images. One can mainly distinguish two classes: the global convergences (that are generally related to vanishing points) and the local convergences. This second class of convergences contains many different situ-

ations. For instance they may correspond to corners or junctions in the image, and in the framework of medical images they correspond to some type of geometric patterns that can, for example, characterize lesions. This is in particular the case of stellate lesions in mammograms. This last type of convergences will be of special interest for us, and it is the situation that led us to develop the method proposed in this paper.

Among all possible methods to detect convergences in an image, we will be interested here in the so-called *a contrario* methods. The general idea underlying this methodology is the following principle: some geometric structures that occur in an image are meaningful when they have a low probability of happening by chance. Under this form, this general principle is sometimes called Helmholtz principle (as in the paper of S.C. Zhu [20]) or also the non-accidentalness principle (as in the paper of Witkin and Tenenbaum [18] or in the book of D. Lowe [12]). In [5], a general methodology has been proposed to translate this principle into an efficient tool, that is called the *a contrario* methodology. It works like this: first define a noise model (also called naive model, background model or *a contrario* model), and then, given a geometric pattern observed in an image, compute the expected number of such patterns occurring in the noise model. When this expected number is less than ε (a positive small real number, less or equal to 1 in general), then the observed geometric pattern is declared ε-meaningful. One of the aim of the paper will be to discuss the choice of the noise model. Indeed, in many situations, the noise model is taken to be the independent uniform distribution on elementary objects (where geometric patterns correspond to a certain arrangement of these elementary objects). Now, this noise model is assumed to correspond to a "normal" situation in which no meaningful patterns are to be found. It has therefore to be chosen carefully.

There is a huge litterature on the issue of detecting global convergences such as vanish-

ing points in man-made environments images. In the *a contrario* framework, this issue has already been addressed by Almansa et al. in [1]. The authors consider the line segments of an image and define vanishing points as regions of the plane that are intersected by a large number of the supporting lines of the segments. Their *a contrario* noise model is that the support lines are independent and uniformly distributed. Using results of stochastic geometry, they build a partition of the plane such that the probability that a random line meets a region is constant for every region. Finally, thanks to the definition of a number of false alarms for each region, they are able to detect meaningful vanishing points.

As we already mentioned it, the local convergences in an image can be of very different types: corners, junctions, stellate patterns, etc. Many of these situations have already been tackled in the *a contrario* framework. For instance, Cao in [3] detects corners in a image, and in [19], Xia, Delon and Gousseau set a robust *a contrario* framework for the detection of junctions in images. These authors consider the pixels as elementary objects and compute their orientation in order to count the number of pixels with similar orientation in the neigborhood of each point. The *a contrario* noise model is that the norm of the locally normalized gradient follows a Rayleigh distribution, that its phase is uniformly distributed and that both variables are independent. They define a notion of strength of each potential branch of a junction, and of strength of the junction as the minimal strength of its branches. Finally, meaningful junctions are the ones that have a strength larger than a threshold, computed thanks to the *a contrario* methodology. This threshold ensures that there are less than ε-meaningful events in a random image following the noise model.

In medical images and more precisely in mammograms, stellate lesions and architectural distortions are also characterized by a local convergence pattern. There are a lot of works on the detection of such lesions. These works

often rely on the extraction of distinguishing features (see for instance [21], [2]) and then perform the aggregation of these features in a decision process (for instance a decision binary tree or a Bayesian classifier). One of the features of stellate lesions and of architectural distortions is the local organization of elementary oriented structures. In [11], Kegelmeyer introduced the ALOE feature that measures the flatness of the local orientation histogram: small values of the ALOE feature indicate that the histogram of the local orientation is flat, which happens in the case of stellate lesions, whereas high values of the ALOE are the consequence of a mode in the local orientation histogram which can describe parallel fibers in this region. Other statistics have been defined to characterize stellate patterns. For instance in [10], Karssemeijer et. al compute the orientation of each pixel using Gaussian filters and then count, at each location, the number of pixels in a ring-shaped neighborhood that are oriented towards the center. To differentiate stellate patterns and parallel fibers they define a second feature, caracterizing the fact that well-oriented pixels are all around the center. Then, having these two features, they build a classifier using a dataset of example feature vectors taken from digitized mammograms, all showing a stellate lesion labeled by an expert radiologist. Recently Palma et al. in [15] and [16] (see also more details in [14]) proposed an *a contrario* framework for the detection of stellate patterns in digital breast tomosynthesis. The stellate patterns are modeled as concentric circles (annulus) such that there is a large number of pixels in this annulus that are oriented towards its center. The *a contrario* noise model is that the pixels are independent and that their orientations are uniform. And the *a contrario* methodology gives the threshold for an annulus to be a meaningful stellate pattern.

Our main concern in this paper will be the choice of the *a contrario* noise model. In many applications, it is usually chosen as the i.i.d. (independent identically distributed) uniform distribution on elementary structures or ob-

jects. This choice is made for the sake of simplification of theoretical computations and it is also often a rough approximation of the distribution of elementary objects in a "normal" image. However in the case of mammograms, the uniform model on the orientation of the linear structures doesn't fit well their normal distribution. Indeed, the linear structures naturally converge towards the nipple, so that the "normal" distribution is not isotropic. We will therefore propose anisotropic models describing a principal normal convergence and derive the *a contrario* method for the detection of convergences in this new framework. As another example of departure from the i.i.d. uniform choice, let us also mention the work of Grosjean et al. in [8] and [9] on the detectability of bright masses in mammograms in an *a contrario* setting. In their work the noise model is chosen as a power law Gaussian texture, that is not an i.i.d. uniform distribution and that is known, when the power law exponent is $\beta \simeq 3$, to model well the texture of mammograms. They are then able to explain a visual perceptual phenomenon that relates the size and brightness of visible spots to the power law exponent of the texture.

Finally, let us emphasize that our contribution in this paper is mainly a methodological one. The experimental results we show on the detection of stellate lesions in mammograms are not as accurate as the state-of-the-art ones. The reason for this is that we consider here only one feature of these lesions and detecting them accurately requires the aggregation of more than one feature.

The paper is organized as follows. The first part of the paper deals with the detection of global convergences in the image plane, and the second part with local ones. More precisely, in Section 2 we describe the *a contrario* framework for the detection of regions of the plane that are met by a large number of lines. We first start with the i.i.d. uniform model on lines and give useful results of stochastic geometry in that framework, mainly about the probability for a random line to meet a given con-

vex region of the plane. Then we introduce a
Gaussian distribution on lines that takes into
account a natural convergence towards a given
point M. We then show how the results of
stochastic geometry are extended to that case.
We also give some asymptotic results on the
behavior of the Gaussian distribution on lines.
In Section 3, we give an algorithm for the esti-
mation of a uniform-Gaussian mixture model
from a set of lines. And in Section 4 we show
several examples of detection of global conver-
gences illustrating the influence of the choice
of the *a contrario* noise model. In the second
part of the paper (Section 5), we will be in-
terested in local convergences and here again
we will show how to estimate an anisotropic
model from a set of line segments in an image,
and we will illustrate the methodology on sev-
eral examples. Finally we end the paper with
a conclusion and discussion in Section 6.

2 The a contrario framework for the detection of global convergences

2.1 Definitions

Let us consider an image defined on a domain
$\Omega \subset \mathbb{R}^2$. Let N be the number of lines in
the image. These lines can be, for instance,
the supporting lines of the segments detected
in the image by the Line Segment Detector
(LSD) algorithm of Grompone et al. [6], also
available online on IPOL[1]. Each line D is rep-
resented by its polar coordinates denoted by
(ρ, φ) with $(\rho, \varphi) \in \mathbb{R} \times [0, \pi)$ (see Figure 1).
The lines are the elementary objects we study.
Due to measurement errors and unprecise con-
vergences, we won't be able to detect "points
of convergence" and we shall rather search for
"regions of convergence", that are regions of the
plane such that "a lot of" lines intersect them.

For a region V of the plane, the number
of lines that meet V is computed and denoted
by $k(V)$. We will be interested in regions such
that $k(V)$ is "large". To define what "large"
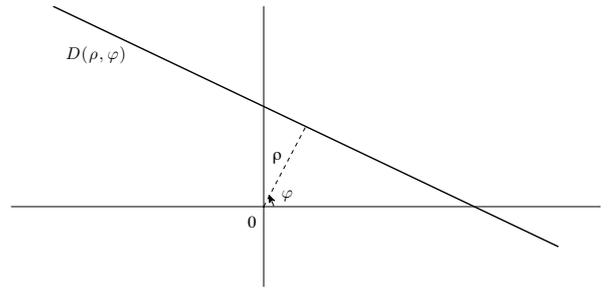
Fig. 1 The image domain Ω is centered at 0, that
is also the origin of the polar parametrization of
straight lines. The distance ρ is the signed distance
from 0 to the line $D = D(\rho, \varphi)$. Notice that lines
with different ρ but with same angle φ are parallel.

means, we first need to compute the proba-
bility of what we observe under an *a contrario*
noise model.

The noise model is the following: we assume
that the N lines in the image domain Ω are
independent and identically distributed with a
measure μ_f of the form

$$d\mu_f = f(\rho, \varphi)d\rho d\varphi$$

where f is a non-negative function.

If f is chosen to be the constant 1 then the
measure $d\mu_1 = d\rho d\varphi$ is the uniform measure
on lines (also called the Poincaré's measure)
and it is the only measure, up to a positive
multiplicative constant, on lines that is invari-
ant under translations and rotations (see the
proof of this in [17]).

We can now introduce the following defini-
tion of meaningful regions of convergence un-
der an *a contrario* noise model.

Definition 1 Let $V \subset \mathbb{R}^2$ be a convex set.
The probability that a random line of the im-
age domain Ω intersects the convex set V un-
der the measure μ_f is defined by

$$p_f(V) := \frac{\mu_f(D \cap V \neq \emptyset \text{ and } D \cap \Omega \neq \emptyset)}{\mu_f(D \cap \Omega \neq \emptyset)}. \quad (1)$$

The number of false alarms of the region V
under the measure μ_f is then defined by

$$\text{NFA}_f(V) = N_T \cdot \mathcal{B}(N, k(V), p_f(V)), \quad (2)$$

where N_T is the number of tests (number of tested regions), $k(V)$ is the number of lines, among the N lines of the image, that intersect the region V and

$$\mathcal{B}(l,k,p) := \sum_{j=k}^{l} \binom{l}{j} p^j (1-p)^{l-j} \qquad (3)$$

is the tail of the binomial distribution of parameters l and p.

Finally, let $0 < \varepsilon \leq 1$. When $\mathrm{NFA}_f(V) \leq \varepsilon$, the region V is said ε-meaningful under the measure μ_f.

This definition of the number of false alarms (NFA) has for consequence the following proposition, that provides a real meaning to the parameter ε: it is an upper-bound of the number of ε-meaningful events that can occur in an image where the lines follow the *a contrario* noise model.

Proposition 1 *Let M_ε be the random variable that counts the number of ε-meaningful regions in an image where the N lines are independent and identically distributed with the law μ_f. Then the expectation of M_ε satisfies*

$$\mathbb{E}(M_\varepsilon) = \mathbb{E}\left(\sum_V \mathbb{1}_{\{V \ is \ \varepsilon\text{-}meaningful\}} \right) \leq \varepsilon. \quad (4)$$

Proof When the N lines are randomly distributed, then the integers $k(V)$ become random variables and we have:

$$\mathbb{E}(M_\varepsilon) = \sum_V \mathbb{P}(\mathrm{NFA}_f(V) \leq \varepsilon)$$

$$= \sum_V \mathbb{P}\left(\mathcal{B}(N, k(V), p_f(V)) \leq \frac{\varepsilon}{N_T} \right)$$

$$\leq \sum_V \frac{\varepsilon}{N_T} = \varepsilon,$$

where the inequality comes from the well-known result in statistics that p-values are uniformly distributed under the null hypothesis. $\qquad \square$

As shown by Formula (2) the number of false alarms is the product of the number of

tests that is usually very high, with the binomial tail that can be very small in comparison, since the parameter $p_f(V)$ is often small. This can produce computational issues that can fortunately be avoided by using the large deviations approximation available when $k(V) \geq N p_f(V)$ (see [7] for instance):

$$\log \mathrm{NFA}_f(V) \approx \log N_T - k(V) \log \frac{k(V)/N}{p_f(V)}$$

$$- (N - k(V)) \log \frac{1 - k(V)/N}{1 - p_f(V)}. \quad (5)$$

2.2 Results of stochastic geometry under the uniform measure on lines

In order to calculate the probability that a random line meets a region V one needs first to calculate the measure of the set of lines, that meet the image domain Ω and then the measure of the set of lines that meet both convex sets V and Ω as in Formula (1). We will first recall in this section some well-known results of stochastic geometry under the uniform measure on lines defined by

$$d\mu_1 = d\rho \, d\varphi.$$

In the next section we will see how these results change when we use a Gaussian measure on lines instead of the uniform one. The main result we will need is the following theorem, the proof of which can be found in the first chapter of [17].

Theorem 1 (see [17]) *Let K_1 and K_2 be two bounded closed convex sets of the plane $\mathbb{R}^2$. Then the μ_1-measure of the set of lines D meeting K_1 is given by*

$$\mu_1(D \cap K_1 \neq \emptyset) = \mathrm{Per}\, K_1, \qquad (6)$$

where Per *denotes the perimeter (i.e. the length of the boundary) of a set.*
And the μ_1-measure of the set of lines meeting

both K_1 and K_2 is given by

$$\mu_1(D \cap K_1 \neq \emptyset \text{ and } D \cap K_2 \neq \emptyset) =$$
$$\begin{cases} L_i - L_e & \text{if } K_1 \cap K_2 = \emptyset, \\ \operatorname{Per} K_2 & \text{if } K_2 \subset K_1, \\ \operatorname{Per} K_1 + \operatorname{Per} K_2 - L_e & \text{otherwise,} \end{cases}$$
$$(7)$$

where L_i and L_e are the interior and exterior perimeters of K_1 and K_2, respectively defined as the length of the boundary of the interior cover of K_1 and K_2 and the length of the boundary of the convex hull of $K_1 \cup K_2$ as shown on Figure 2.
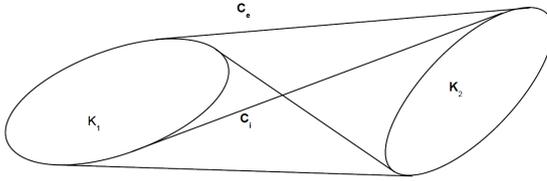
Fig. 2 Exterior and interior perimeters of two convex sets K_1 and K_2 in the case where $K_1 \cap K_2 = \emptyset$. The convex set C_e is the convex hull of K_1 and K_2 so that the exterior perimeter is $L_e = \operatorname{Per} C_e$. And C_i is the interior cover of K_1 and K_2, so that the interior perimeter is $L_i = \operatorname{Per} C_i$.

From both Equations (6) and (7) one can deduce the probability, under μ_1, for a random line D to meet the convex set K_2 conditionally to the event that it meets K_1, as it is simply given by :

$$\mathbb{P}_{\mu_1}(D \cap K_2 \neq \emptyset \,|\, D \cap K_1 \neq \emptyset) =$$
$$\frac{\mu_1(D \cap K_1 \neq \emptyset \text{ and } D \cap K_2 \neq \emptyset)}{\operatorname{Per} K_1}.$$

In the following, to have simple formulas, we will often assume that the image domain Ω is the disk $B(0, R_I)$ of radius R_I and centered in 0. This is just a simplifying hypothesis, and all results can also be stated using a rectangular domain. Formulas would just be less simple.

An important consequence of the results of Theorem 1 is that they allow us to construct the set of test regions. We will use here the construction proposed in [1], and that we briefly recall now. The family of test regions V cover the whole image plane $\mathbb{R}^2$ and there are two kinds of regions: the interior regions that are inside the image domain Ω assumed to be $\Omega = B(0, R_I)$ and the exterior regions that are outside Ω. A value r for the precision of the convergences is first fixed. The interior regions are then chosen to be disks centered at pixels $\mathbf{x} \in \Omega$ and with radius r. Not all pixels need to be taken, it is enough to take them at distance $r/2$ for instance. The probability that a random line under the uniform distribution meets an interior region $V = B(\mathbf{x}, r)$ conditionally to the fact that it already meets the image domain Ω is given by Theorem 1 and it is equal to

$$p_1(V) = \frac{\operatorname{Per} B(\mathbf{x}, r)}{\operatorname{Per} B(0, R_I)} = \frac{r}{R_I},$$

and it is independent of $\mathbf{x}$. Then, exterior regions are chosen to be portions of circular sectors with angle $\Theta = 2r/R_I$ between two distances d_i and d_{i+1}. The sequence of distances d_i is defined in such a way that the probability for a random line to meet an exterior region conditionally to the fact that it already meets the image domain Ω is also equal to r/R_I. Therefore all regions, interior or exterior, are of equal probability. The details of the construction of regions and precise formulas for d_i can be found in [1]. In particular, the sequence d_i if finite, and there is an integer i for which the angular sector region is between the two distances d_i and the infinity. This "last distance" d_i will be denoted d_∞ in the following. See Figure 3, for an example of exterior regions.

The procedure of detection is applied with several scales r. Thus we have to adjust the number of tests N_T in (2) in order to guarantee that we have less than ε-meaningful events in total. This implies that the number of tests is

in fact the number of all tested regions, that is

$$N_T = \sum_r \left(\#\{V_e^r\}_e + \#\{V_i^r\}_i \right),$$

where $\{V_e^r\}_e$ (resp. $\{V_i^r\}_i$) denotes the set of exterior (resp. interior) regions at scale r.
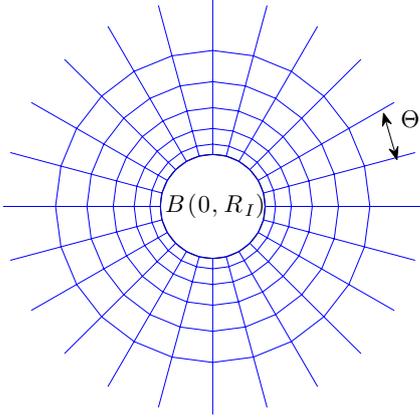
Fig. 3 The image domain is assumed to be the disk $B(0, R_I)$. The exterior regions are portions of circular sectors with angle Θ and between two distances d_i and d_{i+1}. They form a partition of the outside planar domain $\mathbb{R}^2 \setminus \Omega$.

2.3 A Gaussian law on lines

In this section, we want to go beyond the uniform measure on lines and use an *a contrario* noise model that will be able to take into account a main "normal" convergence of the lines in the image. To do this, we introduce the following Gaussian measure μ_g on lines given by

$$d\mu_g = \frac{1}{\pi\sqrt{2\pi}\sigma} e^{-(\rho - x_M cos\varphi + y_M \sin \varphi)^2/2\sigma^2} d\rho d\varphi,$$
$$(8)$$

where (x_M, y_M) are the cartesian coordinates of a point M in the image plane $\mathbb{R}^2$ and $\sigma > 0$ is a parameter. This measure μ_g models the fact that the lines have a tendency to converge towards the point M (see Figure 4). The parameter σ controls the "precision" of the convergence (as in the usual 1D Gaussian distribution where it controls the "width of the peak").

The density function of the Gaussian measure μ_g defined by (8) is given for every $(\rho, \varphi) \in \mathbb{R} \times [0, \pi)$ by

$$g(\rho, \varphi) = \frac{1}{\pi\sqrt{2\pi}\sigma} e^{-(\rho - x_M cos\varphi + y_M \sin \varphi)^2/2\sigma^2}$$
$$= \frac{1}{\pi\sqrt{2\pi}\sigma} e^{-(\rho - r_M \cos(\varphi - \theta_M))^2/2\sigma^2},$$

where (r_M, θ_M) are the polar coordinates of M. The constant $\frac{1}{\pi\sqrt{2\pi}\sigma}$ is the normalization constant making the integral of g on $\mathbb{R} \times [0, \pi)$ equal to 1.

Proposition 2 *The variable $\rho - r_M \cos(\varphi - \theta_M)$ is the signed distance from the line $D(\rho, \varphi)$ to the point M. Under μ_g this variable follows the 1D Gaussian law with mean 0 and variance σ^2, and the variable φ follows the uniform law on the interval $[0, \pi)$.*

Proof With the change of variables $\tilde{\rho} = \rho - r_M \cos(\varphi - \theta_M)$, the density probability of $(\tilde{\rho}, \varphi)$ is given by:

$$\tilde{g}(\tilde{\rho}, \varphi) = g(\tilde{\rho} + r_M \cos(\varphi - \theta_M), \varphi)$$
$$= \frac{1}{\pi\sqrt{2\pi}\sigma} e^{-\tilde{\rho}^2/2\sigma^2}.$$

By integration on φ we deduce that the marginal density probability of $\tilde{\rho}$ is the 1D Gaussian law with mean 0 and variance σ^2. And the integration on ρ shows that the marginal distribution on φ is the uniform law on $[0, \pi)$. □

On Figure 4 we show some sets of lines sampled from the uniform distribution and from a Gaussian distribution given by Eq. (8).

2.4 Measure of the set of lines meeting convex sets under the Gaussian law

We show in this section how Equations (6) and (7) change when, instead of the uniform law on lines, we use a Gaussian law on lines defined by Equation (8). We first recall the definition
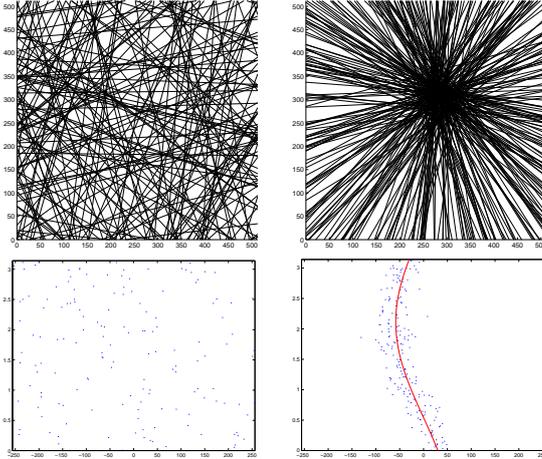
Fig. 4 The image domain Ω is here a square of side length 512. First line: the left picture represents a set of 200 lines under the uniform measure conditioned to meet the image domain Ω; the right picture represents a set of 200 lines under the Gaussian law conditioned to intersect Ω with $\sigma = 20$ and $(x_M, y_M) = (286, 306)$. Second line: the same samples represented as points (ρ, φ) in the rectangular domain $[-256, 256] \times [0, \pi)$. Left picture: the dots are uniformly distributed. Right picture: the dots are concentrated around the curve $(r_M \cos(\theta_M - \varphi), \varphi)$ plotted in red. The width of the concentrated region is of the order of the standard deviation σ.

of the support function of a convex set since we will need it extensively in the following.

Definition 2 (Support function of a convex set) Let K be a closed bounded convex set. The support function of K is defined for all $\varphi \in [0, 2\pi)$ by

$$s_K(\varphi) = \sup_{\mathbf{x} \in K} \langle \mathbf{x}, \mathbf{e}_\varphi \rangle, \tag{9}$$

where $\mathbf{e}_\varphi$ is the unit vector having an angle φ with the horizontal axis, and $\langle \cdot, \cdot \rangle$ is the usual Euclidean scalar product in $\mathbb{R}^2$.

The perimeter of a convex set K can be computed from its support function s_K, and more precisely we have the following result (see [17] for a proof):

$$\text{Per } K = \int_0^{2\pi} s_K(\varphi) d\varphi. \tag{10}$$

We now present some results of stochastic geometry under the Gaussian distribution (Eq. (8)) on lines. Such a Gaussian distribution is described by two parameters: a point M (its cartesian coordinates are denoted by (x_M, y_M) and the polar ones by (r_M, θ_M)) and a width (or standard deviation) parameter σ.

Proposition 3 *Let K be a closed bounded convex set and let s_K denote its support function. The measure under the law μ_g of the set of lines meeting K is given by*

$$\mu_g(D \cap K \neq \emptyset) =$$

$$\frac{1}{\pi} \int_{\varphi=0}^{\pi} \left[\Phi \left(\frac{s_K(\varphi) - r_M \cos(\theta_M - \varphi)}{\sigma} \right) \right.$$

$$\left. - \Phi \left(\frac{-s_K(\varphi + \pi) - r_M \cos(\theta_M - \varphi)}{\sigma} \right) \right] d\varphi, \tag{11}$$

where Φ is the cumulative distribution function of the standard normal distribution, i.e.

$$\forall t \in \mathbb{R}, \quad \Phi(t) := \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \, du.$$

Proof By definition of the support function of K, a line of polar coordinates (ρ, φ) meets the set K if and only if $-s_K(\varphi + \pi) \leq \rho \leq s_K(\varphi)$. As a consequence we have:

$$\mu_g(D \cap K \neq \emptyset) = \int_{\varphi=0}^{\pi} \int_{-s_K(\varphi+\pi)}^{s_K(\varphi)} g(\rho, \varphi) d\rho d\varphi.$$

Using the change of variable $\rho \mapsto \frac{\rho - r_M \cos(\varphi - \theta_M)}{\sigma}$, we get

$$\mu_g(D \cap K \neq \emptyset) =$$

$$\int_{\varphi=0}^{\pi} \frac{1}{\pi} \int_{(-s_K(\varphi+\pi)-r_M \cos(\varphi-\theta_M))/\sigma}^{(s_K(\varphi)-r_M \cos(\varphi-\theta_M))/\sigma} \frac{1}{\sqrt{2\pi}} e^{-\rho^2/2} d\rho d\varphi,$$

which is the announced result. □
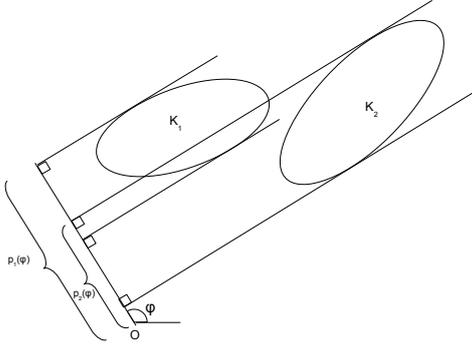
Fig. 5 Illustration of the condition for a line with an angle parameter φ to meet both convex sets K_1 and K_2 of respective support function $s_{K_1} = p_1$ and $s_{K_2} = p_2$.

Proposition 4 *Let K_1 and K_2 be two closed bounded convex sets with respective support function s_{K_1} and s_{K_2}. The measure, under the Gaussian law μ_g, of the set of lines meeting both K_1 and K_2 can be written as*

$$\mu_g(D \cap K_1 \neq \emptyset \text{ and } D \cap K_2 \neq \emptyset) =$$

$$\frac{1}{\pi} \int\limits_{\varphi=0}^{\pi} \mathbb{1}_{\{\alpha(\varphi) \leq \beta(\varphi)\}} \Big[\Phi\left(\frac{\beta(\varphi) - r_M \cos(\theta_M - \varphi)}{\sigma} \right)$$

$$- \Phi\left(\frac{\alpha(\varphi) - r_M \cos(\theta_M - \varphi)}{\sigma} \right) \Big] d\varphi,$$

where the two functions α and β are defined by:

$$\alpha(\varphi) = \max(-s_{K_1}(\varphi + \pi), -s_{K_2}(\varphi + \pi))$$
$$\text{and} \quad \beta(\varphi) = \min(s_{K_1}(\varphi), s_{K_2}(\varphi)).$$

Proof As illustrated on Figure 5, the lines with an angle φ that meet both convex sets K_1 and K_2 are those satisfying the following conditions :

$$- s_{K_1}(\varphi + \pi) \leq \rho \leq s_{K_1}(\varphi)$$
$$\text{and} \quad - s_{K_2}(\varphi + \pi) \leq \rho \leq s_{K_2}(\varphi).$$

These conditions are compatible if and only if $\alpha(\varphi) \leq \beta(\varphi)$. And when they are compatible

they sum up into the single condition

$$\alpha(\varphi) \leq \rho \leq \beta(\varphi). \tag{12}$$

The final result is then obtained by integrating over φ:

$$\mu_g(D \cap K_1 \neq \emptyset \text{ and } D \cap K_2 \neq \emptyset) =$$

$$\int\limits_{\varphi=0}^{\pi} \mathbb{1}_{\{\alpha(\varphi) \leq \beta(\varphi)\}} \int\limits_{\alpha(\varphi)}^{\beta(\varphi)} g(\rho, \varphi) \, d\rho \, d\varphi,$$

and using again, as in the previous proposition, the change of variable $\rho \mapsto \frac{\rho - r_M \cos(\varphi - \theta_M)}{\sigma}$. $\qquad \square$

As a direct consequence of Proposition 3, we can compute the measure, under μ_g, of the set of lines that meet any disk $B(Q, r)$ of center Q and radius r. Indeed, we first notice that the support function of $B(Q, r)$ is given by $s_B(\varphi) = r_Q \cos(\theta_Q - \varphi) + r$, where (r_Q, θ_Q) are the polar coordinates of Q. Then a simple computation (using trigonometric formulas and the fact that we integrate a periodic function on a period) shows that we get

$$\mu_g(D \cap B(Q, r) \neq \emptyset) = \frac{1}{\pi} \int\limits_0^{\pi} \Big[\Phi\left(\frac{r - d(M, Q) \cos \varphi}{\sigma} \right)$$

$$- \Phi\left(\frac{-r - d(M, Q) \cos \varphi}{\sigma} \right) \Big] d\varphi, \quad (13)$$

where $d(M, Q)$ is the usual Euclidean distance between the point of convergence M and the center Q of the disk.

A particular case of the above results occurs when we take for the disk the image domain $\Omega = B(0, R_I)$, and it simplifies into:

$$\mu_g(D \cap \Omega \neq \emptyset) = \frac{1}{\pi} \int\limits_0^{\pi} \Big[\Phi\left(\frac{R_I - r_M \cos \varphi}{\sigma} \right)$$

$$- \Phi\left(\frac{-R_I - r_M \cos \varphi}{\sigma} \right) \Big] d\varphi. \quad (14)$$

Thanks to Formula (13) we can compute the probability under the Gaussian measure that a random line meets an interior region V

that, we recall, is chosen to be a disk of radius r and center $\mathbf{x} \in \Omega$. For the exterior regions, they are of two types: bounded or not. The bounded exterior regions are trapezoids and to apply Proposition 4, we need to compute $\alpha(\varphi)$ and $\beta(\varphi)$. They are simply related to the projections of the four vertices of the trapezoid on the line of angle φ passing through the origin 0 of the image domain $\Omega = B(0, R_I)$. More precisely, if the region is between two distances d_1 and d_2 from 0 and two angles θ_1 and θ_2, then we have

$$\alpha(\varphi) = \max(-R_I, \min_{i,j=1,2}\{d_i \cos(\theta_j - \varphi)\})$$

$$\text{and} \quad \beta(\varphi) = \min(R_I, \max_{i,j=1,2}\{d_i \cos(\theta_j - \varphi)\}).$$

For an unbounded exterior region starting at distance d_∞, we can still use the formula of Proposition 4. We simply need to slightly modify the definitions of $\alpha(\varphi)$ and $\beta(\varphi)$ such that

$$\alpha(\varphi) = \max(-R_I, \min_{j=1,2}\{d_\infty \cos(\theta_j - \varphi)\})$$

$$\text{and} \quad \beta(\varphi) = \min(R_I, \max_{j=1,2}\{d_\infty \cos(\theta_j - \varphi)\}).$$

2.5 Asymptotic behavior of the Gaussian law

As we already mentioned it, the Gaussian distribution on lines is given by two parameters: the point M of convergence and the "precision" σ of the convergence. In this section, we are interested in the influence of these parameters and more precisely in the asymptotic behavior of the measure of lines meeting a fixed convex set K in one the three following situations:
1. The parameter σ goes to infinity. In this case we show that the Gaussian measure behaves (in some sense that is precisely defined below) like the uniform measure on lines.
2. The parameter σ goes to 0. Then the Gaussian measure degenerates and we show that it converges (here again, in some sense that is defined below) to the Dirac measure where all lines pass through the point M.
3. The point M goes to infinity, in the sense that θ_M is fixed while r_M goes to infinity. In

this case, the Gaussian measure degenerates into the Dirac measure where all lines are parallel and uniform in that direction.

All these results match the intuition, and they are precisely stated in the following theorem. Since the proofs are rather technical, they are postponed to the Appendix.

Theorem 2 (Asymptotic behavior of the Gaussian law) *Let g be the density function $g(\rho, \varphi) = \frac{1}{\pi\sqrt{2\pi}\sigma} e^{-(\rho - r_M \cos(\theta_M - \varphi))^2/2\sigma^2}$ on lines. We then have the three following results.*

1. *For any bounded closed convex set K of the plane, as σ goes to $+\infty$, we have*

$$\mu_g(D \cap K \neq \emptyset) = \frac{1}{\pi\sqrt{2\pi}\sigma}\int_0^{2\pi} s_K(\varphi)d\varphi + O\left(\frac{1}{\sigma^2}\right)$$

$$= \frac{1}{\pi\sqrt{2\pi}\sigma}\operatorname{Per}K + O\left(\frac{1}{\sigma^2}\right).$$

As a direct consequence, this implies that

$$\frac{\mu_g(D \cap K \neq \emptyset)}{\mu_g(D \cap \Omega \neq \emptyset)} \xrightarrow[\sigma \to +\infty]{} \frac{\operatorname{Per}K}{2\pi R_I}$$

$$= \frac{\mu_1(D \cap K \neq \emptyset)}{\mu_1(D \cap \Omega \neq \emptyset)},$$

where $d\mu_1 = d\rho d\varphi$ is the uniform measure on lines.
2. *Given any bounded closed and smooth (i.e. with a smooth support function) convex set K of the plane, as σ goes to 0, we have*

$$\mu_g(D \cap K \neq \emptyset) \xrightarrow[\sigma \to 0^+]{} \mu_h(D \cap K \neq \emptyset)$$

$$= \frac{1}{\pi}\int_0^\pi \mathbb{1}_{\{-s_K(\varphi+\pi)\leq r_M \cos(\theta_M - \varphi)\leq s_K(\varphi)\}} d\varphi,$$

where here we denote $d\mu_h = h(\rho, \varphi)d\varphi$ with $h(\rho, \varphi) = \frac{1}{\pi}\delta_{\rho = r_M \cos(\theta_M - \varphi)}$.
3. *For every bounded closed convex set K of the plane, as r_M goes to $+\infty$, we have*

$$\mu_g(D \cap K \neq \emptyset) =$$

$$\frac{1}{\pi r_M}(s_K(\theta_M + \tfrac{\pi}{2}) + s_K(\theta_M - \tfrac{\pi}{2})) + o\left(\frac{1}{r_M}\right).$$

And as a direct consequence,

$$\frac{\mu_g(D \cap K \neq \emptyset)}{\mu_g(D \cap \Omega \neq \emptyset)} \xrightarrow[r_M \to +\infty]{} \frac{\mu_z(D \cap K \neq \emptyset)}{\mu_z(D \cap \Omega \neq \emptyset)}$$

$$= \frac{s_K(\theta_M + \frac{\pi}{2}) + s_K(\theta_M - \frac{\pi}{2})}{2R_I},$$

where $d\mu_z = z(\rho, \varphi)d\rho$ with $z(\rho, \varphi) = \delta_{\varphi = \theta_M \pm \frac{\pi}{2}}$.

3 Estimation of a Gaussian mixture model from lines in an image

The detection results obtained by an *a contrario* method are directly related to the choice of the "naive" background noise model. This naive noise model should represent the "normal" model in which no detection is expected. For instance in mammograms, linear structures converge roughly to the nipple, this is a "normal" convergence and therefore it should not be detected as a potential lesion. The solution we propose in this article is to take into account this principal convergence in the *a contrario* model. The model we choose is a parametric mixture model, with a principal convergence term and a uniform term, conditioned to meet the image domain Ω, and that we will estimate directly from the set of lines.

More precisely, we assume that the distribution on lines we want to estimate is of the form:

$$f_{p,M,\sigma}(\rho, \varphi) = \mathbb{1}_{\{D_{(\rho,\varphi)} \cap \Omega \neq \emptyset\}} \times$$

$$\left[\frac{1-p}{\text{Per}\,\Omega} + \frac{pe^{-(\rho - x_M \cos \varphi + y_M \sin \varphi)^2 / 2\sigma^2}}{\pi\sqrt{2\pi}\sigma Z_{M,\sigma}(\Omega)} \right]$$

$$(15)$$

The above formula contains two terms, the first one is the uniform measure, the second one is the Gaussian model introduced in Section 2.3. Both density functions are conditioned on the set of lines that meet the image domain Ω. The term Per Ω stands for the perimeter of the image domain. When the image domain is the disk of radius R_I and center 0, then we have Per $\Omega = 2\pi R_I$, and the condition $D_{(\rho,\varphi)} \cap \Omega \neq \emptyset$ is equivalent to $|\rho| \leq R_I$.

In the second term, we use again the Gaussian measure μ_g on (ρ, φ) given by

$$d\mu_g(\rho, \varphi) = g_{M,\sigma}(\rho, \varphi)d\rho d\varphi$$

$$= \frac{1}{\pi\sqrt{2\pi}\sigma} e^{-(\rho - x_M \cos \varphi + y_M \sin \varphi)^2 / 2\sigma^2} d\rho d\varphi,$$

where (x_M, y_M) are the coordinates of a point M towards which a proportion p of the lines converges (up to a "precision" σ). The constant $Z_{M,\sigma}(\Omega)$ in Equation (15) is given by $Z_{M,\sigma}(\Omega) = \mu_g(D_{(\rho,\varphi)} \cap \Omega \neq \emptyset)$. It is the Gaussian measure of the set of lines that meet the image domain Ω. Notice that this constant $Z_{M,\sigma}(\Omega)$ has already been explicitly calculated in Section 2.3 with Formula (14).

To estimate the parameters of a density of the form (15) from a set of lines, we choose to do it in two steps:

1. Estimation of the point of convergence M. This point is chosen to be the center of the most meaningful region of convergence (*i.e.* with the smallest NFA) under the uniform *a contrario* noise model $d\mu_1 = d\rho d\varphi$ and where multiple scales are used (as described in Section 2.2).
2. Once the point M is estimated, we proceed to the simultaneous estimation of both parameters p and σ by maximizing the log-likelihood on the variable

$$\tilde{\rho} = \rho - x_M \cos \varphi + y_M \sin \varphi,$$

which is the signed distance of the line $D_{(\rho,\varphi)}$ to the point M.

When a line $D_{(\rho,\varphi)}$ follows the distribution $f_{p,M,\sigma}$ defined by Equation (15), then the marginal law followed by the variable $\tilde{\rho}$ can easily be calculated. We denote it by $h_{p,\sigma}$ (the point M is now assumed to be fixed) and, in a way similar to Proposition 2, it is given by:

$$h_{p,\sigma}(\tilde{\rho}) =$$

$$\left[\frac{1-p}{2\pi R_I} + \frac{p}{\pi\sqrt{2\pi}\sigma Z_{M,\sigma}(\Omega)} e^{-\tilde{\rho}^2 / 2\sigma^2} \right] \times J(\tilde{\rho})$$

where J is a function independent of p and σ. More precisely, J is given by

$$
J(\tilde{\rho}) = \int\limits_0^\pi \mathbb{1}_{\{D_{(\rho,\varphi)} \cap \Omega \neq \emptyset\}} \, d\varphi
$$

$$
= \int\limits_0^\pi \mathbb{1}_{\{|\tilde{\rho} + x_M \cos\varphi - y_M \sin\varphi| \leq R_I\}} \, d\varphi.
$$

Then, given N lines in the images, the log-likelihood of (p, σ) given the observed variables $\tilde{\rho}_1, ..., \tilde{\rho}_N$ is

$$
\mathcal{LL}(p, \sigma | \tilde{\rho}_1, ..., \tilde{\rho}_N) := \sum_{i=1}^N \log h_{p,\sigma}(\tilde{\rho}_i). \qquad (16)
$$

We consider it as a 2-variable function and we estimate its maximum directly (by an exhaustive search on discretized values of p and σ), which gives us the "most likely" values of p and σ.

Once a Gaussian-uniform mixture model of the form (15) has been estimated, we run the *a contrario* detection of convergence with the estimated model as the noise model. In this anisotropic noise model, one needs to calculate the probability $p_{f_{p,M,\sigma}}(V)$ that a line D meets a region V under the model $f_{p,M,\sigma}$. Formula (1) shows that it uses both measures $\mu_{f_{p,M,\sigma}}(D \cap V \neq \emptyset$ and $D \cap \Omega \neq \emptyset)$ and $\mu_{f_{p,M,\sigma}}(D \cap \Omega \neq \emptyset)$. This last measure is equal to 1 since $f_{p,M,\sigma}$ is a density function on the lines that meet Ω. Whereas the first measure is given by :

$$
\mu_{f_{p,\sigma}}(D \cap V \neq \emptyset \text{ and } D \cap \Omega \neq \emptyset) =
$$

$$
\frac{(1-p)}{2\pi R_I} \mu_1(D \cap V \neq \emptyset \text{ and } D \cap \Omega \neq \emptyset)
$$

$$
+ \frac{p}{Z_{M,\sigma(\Omega)}} \mu_{g_{M,\sigma}}(D \cap V \neq \emptyset \text{ and } D \cap \Omega \neq \emptyset).
$$

The two terms above are then simply computed by respectively the formulas of Theorem 1 and of Proposition 4.

Examples of the whole procedure (detection under the uniform noise model, estimation of a mixture model and *a contrario* detection under this estimated anisotropic model) are given in the next section.

4 Examples

We present here examples with three different types of images: a synthetic image, a "natural" image and a mammogram. For each example are displayed:

1. the original image with the set of line segments detected (or simulated),
2. the set of supporting lines,
3. $-\log \mathrm{NFA}_1$, where NFA_1 is the number of false alarms against the uniform model for each interior region and, when relevant, for each exterior region,
4. the estimated point M of principal convergence,
5. the log-likelihood function $\mathcal{LL}(p, \sigma)$,
6. $-\log \mathrm{NFA}_{f_{\hat{p},M,\hat{\sigma}}}$, where $\mathrm{NFA}_{f_{\hat{p},M,\hat{\sigma}}}$ is the number of false alarms of the regions against the estimated mixture model.

Figure 6 is a synthetic example where there are two convergences but one is "stronger" than the other one. This main convergence is the most meaningful one under the uniform model. But then, once this principal convergence is integrated into the *a contrario* mixture model by the proposed procedure, the secondary convergence appears as the most meaningful event whereas the principal one is no longer a meaningful event.

On the flower image of Figure 7 there is a clear principal convergence at the center of the flower. It is indeed detected as the most meaningful region under the uniform distribution on lines. Then, against the estimated mixture distribution, there are no meaningful region at all. One can therefore say that, in this sense, the estimated model is a better description of the distribution of lines in this image than the uniform model.

On Figure 8, we illustrate the result of the whole procedure on a mammogram. On this mammogram, there is a very meaningful convergence around the nipple, which is normal since the normal linear structures visible in mammograms (fibrous tissue patterns and blood vessels) are roughly oriented towards this point in a normal breast tissue. Once this principal

(a)

(b)

(c)

(d)
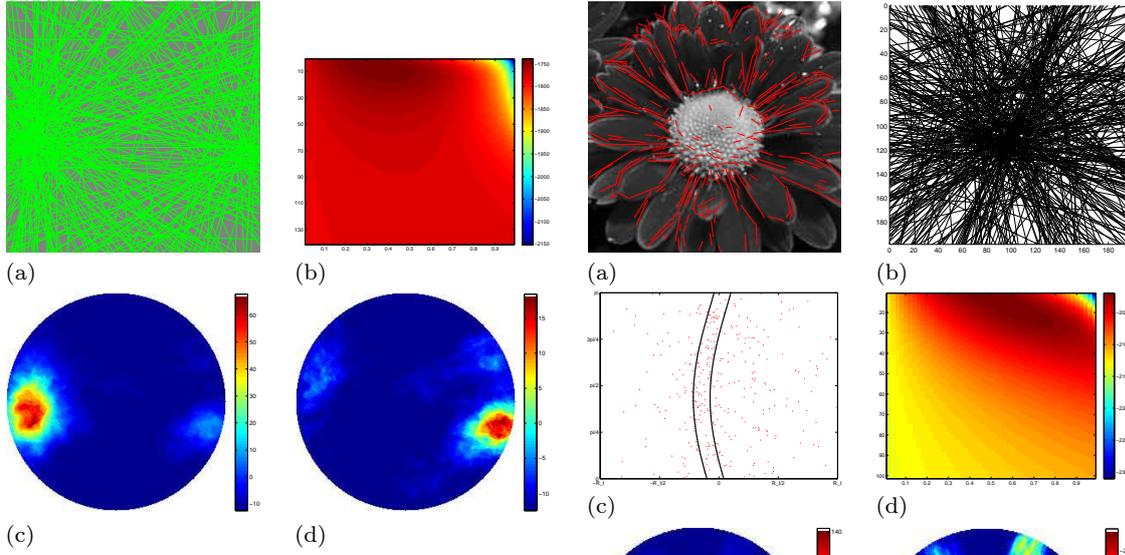
Fig. 6 (a) We simulated $N = 300$ straight lines with a parametric mixture model with two Gaussian terms and one uniform term: the main convergence is centered in $(x_M, y_M) = (-100, 14)$ with a standard variation $\sigma = 20$ and a weight $p = 0.4$, and the secondary convergence has center $(x_{M'}, y_{M'}) = (114, 28)$, $p' = 0.2$ and $\sigma' = 10$. The uniform term has thus weight $1 - p - p' = 0.4$. (b) Image of the log-likelihood as a function of p and σ. It is maximal for $(\hat{p}, \hat{\sigma}) = (0.37, 17)$. (c) Image of $-\log \text{NFA}_1$ showing that the main convergence is detected as the most meaningful region whereas the secondary convergence is not detected under the uniform model. (d) Image of $-\log \text{NFA}_{f_{\hat{p}, M, \hat{\sigma}}}$, showing that, under the estimated distribution $f_{\hat{p}, M, \hat{\sigma}}$, the secondary convergence is the most meaningful region of convergence whereas the principal convergence is no longer meaningful.

(a)

(b)

(c)

(d)

(e)

(f)

Fig. 7 (a) The original flower images, with $N = 332$ line segments detected (by the LSD Algorithm). The radius of the image is $R_I = 99$. (b) The set of support lines. (c) The same set of lines seen as points (ρ_n, φ_n) in $[-R_I, R_I] \times [0, \pi)$. (d) Image of the log-likelihood. It is maximal for $(\hat{p}, \hat{\sigma}) = (0.53, 13)$. (e) Image of $-\log \text{NFA}_1$. Against the uniform model the most meaningful point is $(x_M, y_M) = (-3, 14)$. (f) Image of $-\log \text{NFA}_{f_{\hat{p}, M, \hat{\sigma}}}$. There is no meaningful region against the parametric mixture model.

convergence is integrated in the mixture model it is no longer detected as significant. However another convergence appears: it is highlighted when the *a contrario* model changes (see the log NFA scales on the images). This could be a suspicious region. However, lesions are local events and detecting them using the lines in the image may not be adapted. We therefore have to use the line segments themselves (and not their support lines since these lines "forget" where the line segments were). This will be the aim of the next section.

5 Detection of local convergences

In this section, we will be interested in the detection of local convergences (as opposed to global ones in the previous section). Indeed, in many applications (for instance in mammograms), the convergences we are interested in are local ones, in the sense that only the way the elementary linear structures are organized in a neighborhood of a given point matters, even if these elementary linear structures may have a global "natural" orientation. Therefore,

(a)

(b)

(c1)

(d1)

(c2)

(d2)
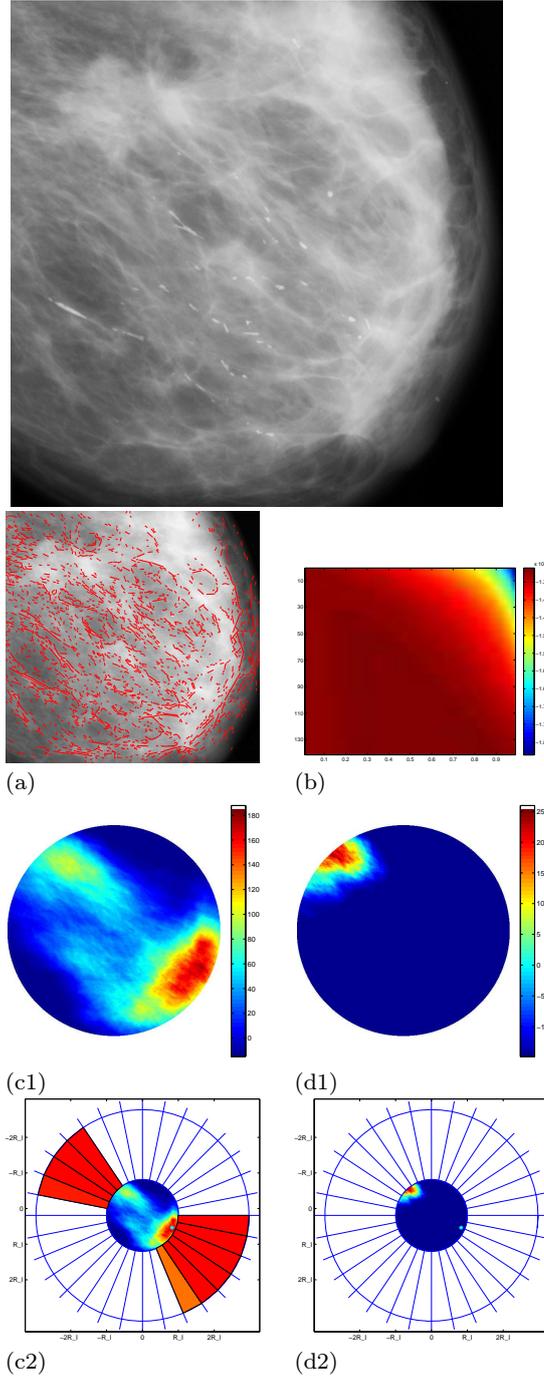
Fig. 8 (a) The image radius is $R_I = 274$. The number of line segments is $N = 1781$. These are the line segments detected by the LSD Algorithm of [6]. (b) Image of the log-likelihood. The estimated parameter values are $(\hat{p}, \hat{\sigma}) = (0.62, 145)$. (c1) and (c2): Images of $-\log \text{NFA}_1$ on respectively interior and exterior regions. The most meaningful point is inside the image domain, it is the point $(x_M, y_M) = (226, 94)$. (d1) and (d2): Images of $-\log \text{NFA}_{f_{\hat{p}, M, \hat{\sigma}}}$ on respectively interior and exterior regions. This second run of the *a contrario* method gives a new convergence inside the image, whereas no exterior region is meaningful.

as we did it in the previous section for lines, we will introduce here an anisotropic *a contrario* model on line segments that is a mixture of a uniform distribution on lines segments and of a second anisotropic "Gaussian" term that models the natural convergence of some of the segments to a particular point in the image domain.

5.1 An anisotropic law on line segments

The elementary objects we deal with in this section are line segments. A line segment S is given by the three coordinates (x_S, y_S, θ) where $\mathbf{x}_S = (x_S, y_S)$ are the cartesian coordinates of the center of the line segment S and where $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2})$ is the angle between the horizontal axis and the line segment. Note that this description of line segments doesn't take their length into account. The reason for this is that here, in the detection of convergence, we are only interested in their position and orientation, and we don't need their length (they are all assumed to be "small segments").

The uniform model on the line segments is given by the measure

$$d\mu_1 = dx_S dy_S d\theta. \tag{17}$$

It corresponds to a uniform distribution for the center $\mathbf{x}_S$ of the segment, a uniform distribution also for its orientation θ, and both being independent. A natural question is then to give the link between the distributions on line segments and the distributions on their supporting lines. In particular, we may ask: does a uniform distribution on line segments imply a uniform distribution on their support lines ? To explore this, we first derive the relationship between the coordinates (x_S, y_S, θ) of the line segments and those (ρ, φ) of its support line. The support line (ρ, φ) of the segment is parallel to the vector $\mathbf{v}_\varphi = (\sin \varphi, \cos \varphi)$ with $\varphi = \theta + \frac{\pi}{2}$. An origin point on the support line is taken as being the orthogonal projection of the origin 0 of the domain Ω on the line. The parameter that gives the segment position on

the support line is denoted by t and therefore the segment can be described by the new system of coordinates (ρ, φ, t). See Figure 9 for an illustration of this. The change of coordinates can be written by

$$\begin{cases} x_S & = \rho \cos \varphi + t \sin \varphi \\ y_S & = -\rho \sin \varphi + t \cos \varphi \\ \theta & = \varphi - \frac{\pi}{2}. \end{cases} \qquad (18)$$
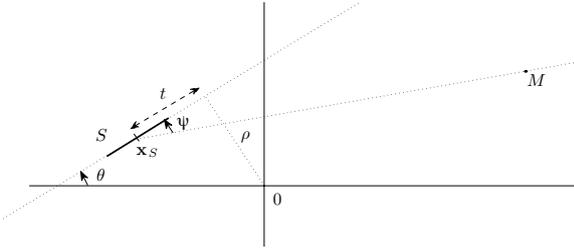
Fig. 9 Parametrization of the segments with $(\mathbf{x}_S, \theta)$ or (ρ, φ, t), and angle ψ between the segment and a global convergence point M.

The Jacobian of this change of coordinates is equal to 1. It implies that every density of probability $\tilde{f}$ on the variables (x_S, y_S, θ) leads to a density f on (ρ, φ, t) related to $\tilde{f}$ by

$$f(\rho, \varphi, t) = \tilde{f}(x_S, y_S, \theta)$$
$$= \tilde{f}(\rho \cos \varphi + t \sin \varphi, -\rho \sin \varphi + t \cos \varphi, \varphi - \frac{\pi}{2}).$$

In particular, if $\tilde{f}$ is the uniform law on (x_S, y_S, θ) conditioned to be in the image domain, that is conditioned to $x_S^2 + y_S^2 \leq R_I^2$, then the law on (ρ, φ, t) is given by

$$f(\rho, \varphi, t) = \mathbb{1}_{\{\rho^2 + t^2 \leq R_I^2\}} \mathbb{1}_{\{\varphi \in [0, \pi)\}}.$$

By integration over t we find that the marginal density of probability on support lines (ρ, φ) is the function

$$(\rho, \varphi) \mapsto 2\sqrt{R_I^2 - \rho^2} \mathbb{1}_{\{\rho^2 \leq R_I^2\}} \mathbb{1}_{\{\varphi \in [0, \pi)\}}, \quad (19)$$

which is not the uniform law on lines conditioned to meet the image domain. Indeed, as we already saw it in Section 2.2, the uniform

law on lines was given by the density $(\rho, \varphi) \mapsto \mathbb{1}_{\{\rho^2 \leq R_I^2\}} \mathbb{1}_{\{\varphi \in [0, \pi)\}}$. The law on lines induced by uniformly distributed segments may look paradoxical since if we count, in an informal way, the "number" of points at a distance between ρ and $\rho + d\rho$ from 0, then we find that it is proportional to $2\pi \rho d\rho$. This increases with ρ, whereas the density given by Equation (19) behaves the opposite way: it decreases with ρ. The explanation is that line segments are not only points, we have to add an orientation. Therefore the "number" of segments such that their supporting line is at a distance between ρ and $\rho + d\rho$ from 0 is proportional to the number of points lying in a strip intersecting the disk, which area is roughly $2\sqrt{R_I^2 - \rho^2} d\rho$. An illustration of the phenomenon that uniformly distributed segments don't imply uniformly distributed lines is given on Figure 10. On this figure, we show how uniformly distributed segments create meaningful convergences under the uniform noise model on lines. Such convergences would not exist if the support lines were uniformly distributed.
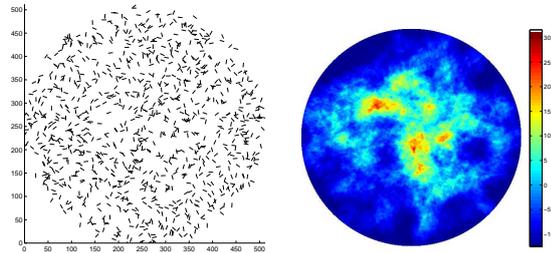
Fig. 10 Left: 1000 uniformly distributed line segments on the image domain. Right: Image of $-\log \mathrm{NFA}_1$ under the uniform distribution on line. Every point such that $-\log \mathrm{NFA}_1 \geq 0$ is the center of a 1-meaningful region. There are many meaningful events, mainly close to the image center, corresponding to small values of ρ, which are more frequent than they would be if the lines were uniformly distributed on $(-R_I, R_I) \times [0, \pi)$.

Now, in order to include more information in the *a contrario* model (in particular the existence of a natural convergence, as in the case of mammograms), we have to define some anisotropic

model on the line segments in a way similar
to the Gaussian law on lines in Section 2.3.
The model considered here is again a para-
metric mixture of the uniform law on line seg-
ments and of another term, which is still called
"Gaussian", modelling the convergence to a point
M. In this second term we will need the anal-
ogous, for angular values, of the Gaussian dis-
tribution on $\mathbb{R}$. It is called the von Mises dis-
tribution or sometimes the circular normal dis-
tribution. It is defined by:

Definition 3 (von Mises distribution on
$[-\frac{\pi}{2}, \frac{\pi}{2})$ **)** The von Mises distribution on $[-\frac{\pi}{2}, \frac{\pi}{2})$
is given by its density of probability on $[-\frac{\pi}{2}, \frac{\pi}{2})$,
which is :

$$\forall t \in [-\tfrac{\pi}{2}, \tfrac{\pi}{2}), \ \ g_\kappa(t) = \frac{e^{\kappa \cos(2t)}}{2\pi I_0(\kappa)}, \tag{20}$$

where κ is a parameter inversely proportional
to the "concentration" or "width", and I_0 is the
modified Bessel function of order 0. This law is
symmetric and it is the uniform law on $[-\frac{\pi}{2}, \frac{\pi}{2})$
when $\kappa = 0$. See also Figure 11 to visualize the
shape of the von Mises distribution for differ-
ent values of κ.
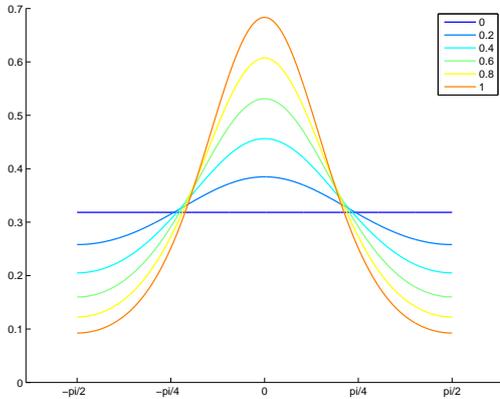
Fig. 11 The von Mises distribution for different
values of κ.

Given a point M in the image domain, as
in the case of lines, we want to model here
the fact that a proportion p of the line seg-
ments converges towards M. To take this into

account, we first consider the angle ψ between
the line segment (x_S, y_S, θ) and the line de-
fined by the two points $\mathbf{x}_S = (x_S, y_S)$ and M.
See Figure 9. Then we make the assumption
that the density of ψ is given by the mixture
distribution defined for all $u \in [-\frac{\pi}{2}, \frac{\pi}{2})$ by

$$h_{p,\kappa}(u) = (1-p)\frac{1}{\pi} + p\frac{e^{\kappa \cos(2u)}}{\pi I_0(\kappa)}. \tag{21}$$

The second term is the one describing the con-
vergence towards M, and it makes small values
of ψ more likely. The first term is simply the
uniform law on $[-\frac{\pi}{2}, \frac{\pi}{2})$. The angles ψ and θ
are related by the equation

$$\theta = \psi + \arctan\left(\frac{y_M - y_S}{x_S - x_M}\right) \pmod{\pi}, \tag{22}$$

so that the law of θ knowing (x_S, y_S) is easily
derived from the law of ψ. More precisely, the
density of probability of θ knowing (x_S, y_S) is
the function, defined on $[-\frac{\pi}{2}, \frac{\pi}{2})$, by

$$t \mapsto h_{p,\kappa}\left(t - \arctan\frac{y_M - y_S}{x_S - x_M}\right), \tag{23}$$

where the function $h_{p,\kappa}$ is considered as being
π-periodic. Notice that when the parameter p
is null, then ψ follows the uniform law.

Finally, the whole *a contrario* distribution
on line segments is: the line segments are in-
dependent, the position (x_S, y_S) is uniformly
distributed on the image domain Ω, and the
law of θ knowing (x_S, y_S) is given by the mix-
ture distribution of Equations (21) and (23).

Estimation. As in Section 3 we address the is-
sue of the estimation of such models given a
real image and its line segments $\{(x_S^{(n)}, y_S^{(n)}, \theta_n)\}$,
with $n = 1, \ldots, N$. We first choose to estimate
the point M of global convergence the same
way as in Section 3, by minimizing the number
of false alarms against the uniform model on
the support lines, and setting M as the cen-
ter of the most meaningful region among all
scales r. Then the parameters p and κ are es-
timated by maximizing the log-likelihood on

(p, κ) given the data $\psi_1, \ldots, \psi_N$ in a way similar to the estimation of p and σ in Section 3. The log-likelihood on (p, κ) given the data $\psi_1, \ldots, \psi_N$ is:

$$\mathcal{LL}(p, \kappa | \psi_1, \ldots, \psi_N) =$$
$$\sum_{n=1}^{N} \log \left((1-p)\frac{1}{\pi} + p\frac{e^{\kappa \cos(2\psi_n)}}{\pi I_0(\kappa)} \right),$$

where

$$\forall n, \quad \psi_n = \theta_n - \arctan \left(\frac{y_M - y_S^{(n)}}{x_S^{(n)} - x_M} \right)$$

Then log-likelihood maximum is found by an exhaustive search after the discretization of the set of parameters (p, κ).

5.2 The *a contrario* framework for local convergences
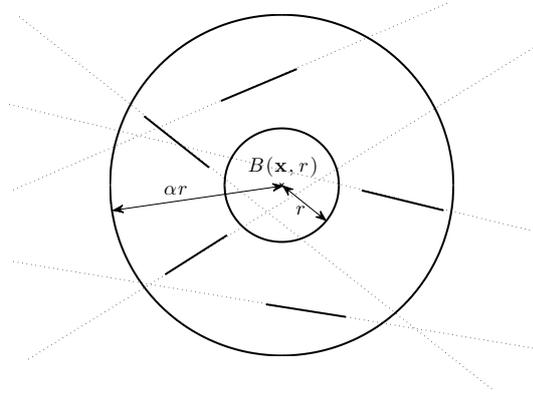
Fig. 12 Geometric description of a local convergence: we count the number of line segments in the annulus and among them, the ones such that their support line meets the central disk $B(\mathbf{x}, r)$.

The local convergences are searched under the form of annuli denoted by $\mathcal{A}(\mathbf{x}, r, \alpha)$ that are rings delimited by two concentric circles of center $\mathbf{x}$ and respective radius r and αr, where $\alpha > 1$ (and denoted respectively by $\mathcal{C}(\mathbf{x}, r)$ and

$\mathcal{C}(\mathbf{x}, \alpha r)$). For each potential convergence (annulus) we consider the variables $N(\mathbf{x}, r, \alpha)$ and $K(\mathbf{x}, r, \alpha)$ which are respectively the number of line segments between the two circles $\mathcal{C}(\mathbf{x}, r)$ and $\mathcal{C}(\mathbf{x}, \alpha r)$ and the number of line segments between the two circles whose support lines meet the central disk $B(\mathbf{x}, r)$. See also Figure 12 for an illustration of this.

Local convergences in an image are likely to have large values for both variables $K(\mathbf{x}, r, \alpha)$ and $N(\mathbf{x}, r, \alpha)$. The *a contrario* method provides the threshold on these variables in order for a potential convergence to be meaningful.

Definition 4 Let N segments $(x_S^{(i)}, y_S^{(i)}, \theta_i)$, $1 \leq i \leq N$, be observed in an image domain Ω, assumed to be a disk of radius R_I. The number of false alarms of an annulus $\mathcal{A}(\mathbf{x}, r, \alpha)$ under a probability density f on segments is defined by

$$\mathrm{NFA}_f(\mathbf{x}, r, \alpha) =$$
$$N_a \cdot \mathcal{B}(N(\mathbf{x}, r, \alpha), K(\mathbf{x}, r, \alpha), p_f(\mathbf{x}, r, \alpha)),$$
$$(24)$$

where N_a is the number of tested annulus, $\mathcal{B}$ denotes the tail of the binomial distribution and p_f is the probability of convergence towards the center of the annulus. It is given by

$$p_f(\mathbf{x}, r, \alpha) = \frac{\int_{\Omega \times [0,\pi]} \mathbb{1}_{E_{x,y}} \mathbb{1}_{E_\theta} f(x_S, y_S, \theta) \, dx_S dy_S d\theta}{\int_{\Omega \times [0,\pi]} \mathbb{1}_{E_{x,y}} f(x_S, y_S, \theta) \, dx_S dy_S d\theta},$$

where $E_{x,y}$ and E_θ are respectively the events

$$E_{x,y} = \{ r \leq \|(x_S, y_S) - \mathbf{x}\| \leq \alpha r \} \quad \text{and}$$

$$E_\theta = \{ |\theta - \arctan \frac{y_S - y_\mathbf{x}}{x_S - x_\mathbf{x}}| \leq \arcsin \frac{r}{\|(x_S, y_S) - \mathbf{x}\|} \}.$$

Let $\varepsilon > 0$ be a "small number" (less than 1), then an annulus $\mathcal{A}(\mathbf{x}, r, \alpha)$ such that $\mathrm{NFA}_f(\mathbf{x}, r, \alpha) < \varepsilon$ is said to be ε-meaningful.

Proposition 5 *The above definition indeed defines a number of false alarms, in the sense that: in an image where N line segments are following the a contrario model with density f, then the mean number of ε-meaningful events is less than ε.*

Proof This proposition can be proved in way analogous to Proposition 1. Notice however that here there are two random variables: $N(\mathbf{x}, r, \alpha)$ and $K(\mathbf{x}, r, \alpha)$, and not only just $k(V)$ as in Proposition 1. Notice also that the definition of NFA used here is analogous to the one of [8] where a "context" information is used. $\square$

We will mainly use the above number of false alarms NFA_f with two different probability densities f. We will first consider the case of the uniform distribution on positions and orientations for segments, which is equivalent to say that we take

$$f_1(x_S, y_S, \theta) = \frac{1}{\pi^2 R_I^2} \mathbb{1}_{\{x_S^2 + y_S^2 \le R_I^2\}}.$$

Then the probability of convergence has a simple expression. Indeed using polar coordinates and an integration by parts, we get

$$p_1(\mathbf{x}, r, \alpha) =$$
$$\frac{2}{\pi(\alpha^2 - 1)} \left[\alpha^2 \arcsin\left(\frac{1}{\alpha}\right) + \sqrt{\alpha^2 - 1} - \frac{\pi}{2} \right]. \tag{25}$$

The second case we will consider is the one of a mixture density of the form

$$f(x_S, y_S, \theta) =$$
$$\frac{1}{\pi R_I^2} h_{p,\kappa}\left(\theta - \arctan\frac{y_M - y_S}{x_S - x_M} \right) \mathbb{1}_{\{x_S^2 + y_S^2 \le R_I^2\}},$$

where $h_{p,\kappa}$ is the uniform/von Mises mixture distribution defined by Equation (21). Here the point M and the parameters p and κ are estimated from the line segments of an image, as described in the previous section.

5.3 Examples

As in the case of global convergences, we will here illustrate the proposed method on three examples: a synthetic set of line segments, a natural image and a mammogram.

On Figure 13, we illustrate the influence of the *a contrario* background noise model on a
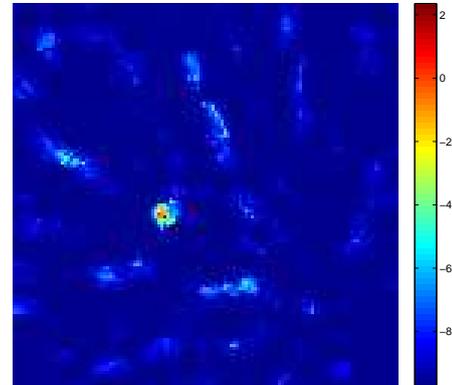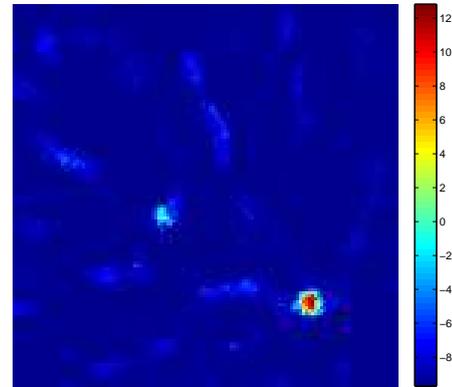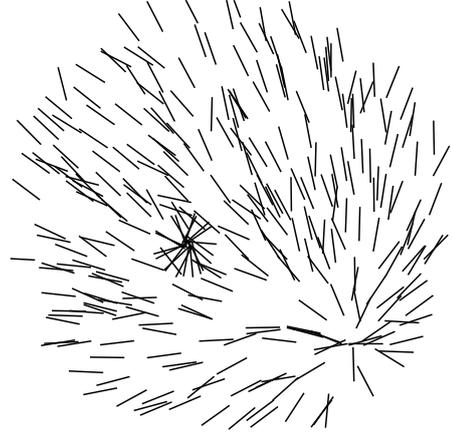
Fig. 13 Synthetic example: this set of line segments contains a main convergence and a secondary one (top image). The middle image represents $-\log(\mathrm{NFA}_1)$, where NFA_1 is the number of false alarms of local convergences under the uniform distribution on line segments: only the main convergence is meaningful. The bottom image represents $-\log(\mathrm{NFA}_f)$, where NFA_f is the number of false alarms of local convergences under the von Mises distribution on line segments: only the secondary convergence is meaningful.

synthetic set of line segments. Under the uniform distribution on line segments, the main point of convergence is the only meaningful local convergence (middle image). Whereas under the von Mises distribution, this main convergence is not meaningful anymore and the secondary convergence becomes the only meaningful local convergence. In this example, the main point M of global convergence and the parameters of the von Mises distribution are known (and we don't estimate them).

On Figure 14, we show the results of the proposed approach on a natural image. This image contains many geometric structures, and a point M of global convergence is found below the center of the image (represented as a red star). The histogram of the values of ψ (angle between the support line of the segment and the line defined by the center of the segment and M) has a narrow mode in 0. The red curve on the histogram represents the estimated mixture distribution. On the second line of the figure, we show on the left the image of $-\log(\text{NFA}_1)$, where NFA_1 is the number of false alarms under the uniform law on line segments. The point M of global convergence is also here the most meaningful local convergence. On the right, we show the image of $-\log(\text{NFA}_f)$, where NFA_f is the number of false alarms under the estimated mixture distribution on line segments. Notice how the point M is now "erased" (not meaningful anymore) under the mixture distribution.

On Figure 15, we perform exactly the same experiments: estimation of the point M, shown as a red star; estimation of a mixture uniform/von Mises distribution for the angles ψ and computation of the number of false alarms for local convergences under both: the uniform distribution on line segments and the estimated one. This image is a mammogram and it is not as geometric as the natural image of Figure 14. The point M that is found as the most meaningful global convergence is close to the nipple but they do not coincide. This point M is also the most meaningful point of local convergence under the uniform distribution on line
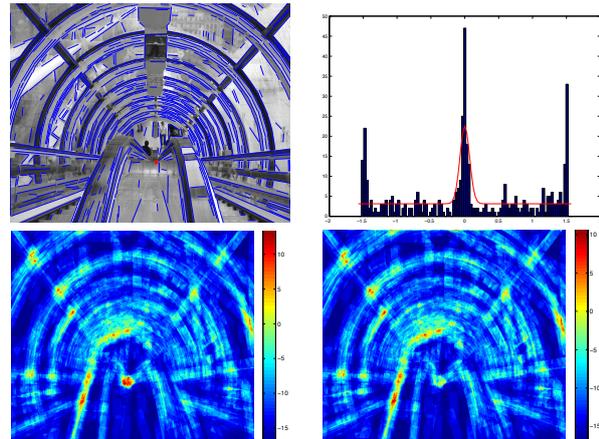
Fig. 14 First line: original image (Beaubourg center), its line segments, and the estimated point M of global convergence (represented as a red star). On the right: histogram of angles ψ and the estimated mixture distribution plotted in red. Second line: on the left, image of $-\log(\text{NFA}_1)$ under the uniform law on line segments, and on the right, image of $-\log(\text{NFA}_f)$ under the estimated mixture distribution on line segments.

segments, and it could be the center of a suspicious region. Then under the estimated mixture distribution, another point becomes the most meaningful one. This example illustrates the hierarchical organization of local convergences: we start by detecting the most meaningful one, then we integrate it to the model and this allows us to enhance the secondary convergences.

5.4 From Theory to Practice

As pointed out by one of the reviewers of the paper, looking at Figure 15, one may wonder "whether the method works at all". Indeed, on that example, the point with lowest NFA_1 is not on the nipple, and other detections are not improved when using NFA_f instead of NFA_1. For the point with lowest NFA_1, we don't use any spatial information (like being on the boundary of the breast, for instance), and therefore, in some cases, this point is not on the nipple. This point only corresponds to the point that
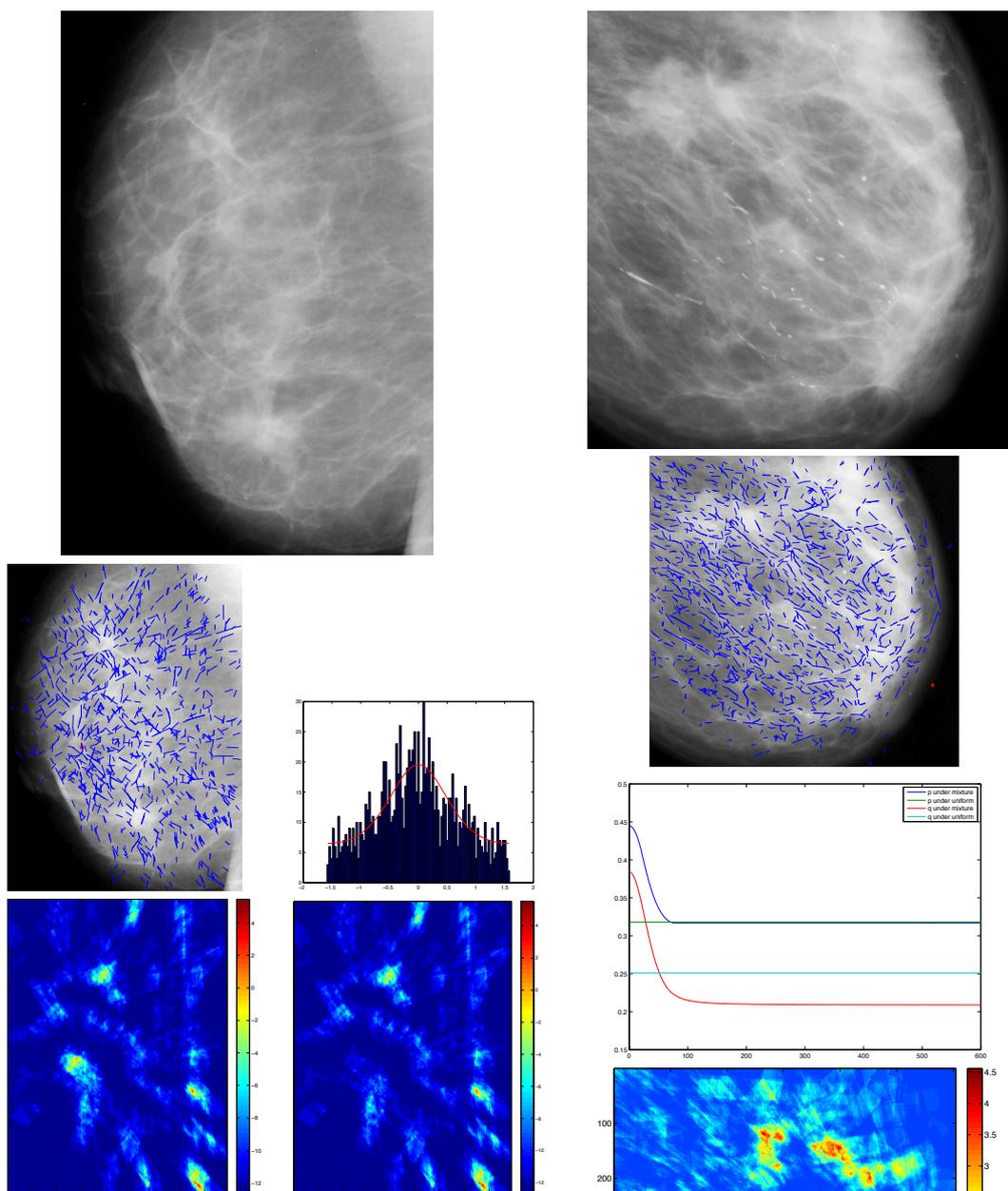
Fig. 15 First line: original image (a mammogram), its line segments, and the estimated point M of global convergence (represented as a red star). On the right: histogram of ψ and the estimated mixture distribution plotted in red. Second line: on the left, image of $-\log(\mathrm{NFA}_1)$ under the uniform law on line segments, and on the right, image of $-\log(\mathrm{NFA}_f)$ under the estimated mixture distribution on line segments.
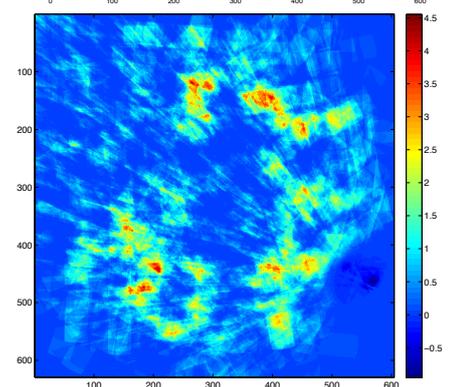
Fig. 16 Original mammogram, its line segments, and a point M of global convergence (red star, manually chosen). Plot of the probabilities p_1, p_f, q_1 and q_f as a function of the distance between $\mathbf{x}$ and M. Image of $-\log \mathrm{NFA}_f^{\perp} + \log \mathrm{NFA}_1^{\perp}$, showing the enhancement of the detections when using the anisotropic *a contrario* model.

would be in the highest position in a kind of hierarchical description of the local convergences of the image. Now, for the other detections, the probability of convergence $p_f(\mathbf{x}, r, \alpha)$ only depends on the distance of $\mathbf{x}$ from M, it decreases and becomes very close to the uniform probability of convergence $p_1(\mathbf{x}, r, \alpha)$ as the distance increases. This explains why NFA_1 and NFA_f are almost the same for detections that are far away the point M.

Therefore, one can ask if our approach can be used in practice. We have presented here a theoretical framework, and the validation of this approach for the detection of stellate lesions in mammograms remains to be done. Ideally, we would like to prepare a companion paper, intended to a medical imaging audience, showing the interest of our approach, using for instance ROC curves on a dataset of annotated mammograms.

Now, we can give a hint on the features of stellate patterns that could be used, and then need to be validated by a statistical learning approach on a dataset, as it is done for instance by Karssmeijer et. al in [10]. For an annulus $\mathcal{A}(\mathbf{x}, r, \alpha)$, we have already considered two features, namely its NFA_1 and its NFA_f. Now, we could consider a second type of features, counting now the number of linear structures that have their center at a distance larger than r from the line $(M\mathbf{x})$ and still converging to $\mathbf{x}$. This is, in spirit, similar to the second feature defined by Karssmeijer et. al in [10], where they consider the directions where the oriented pixels come from. Thus, a second number of false alarms of an annulus $\mathcal{A}(\mathbf{x}, r, \alpha)$ under a probability density f on segments can be defined by

$$\mathrm{NFA}_f^{\perp}(\mathbf{x}, r, \alpha) =$$
$$N_a \cdot \mathcal{B}(N^{\perp}(\mathbf{x}, r, \alpha), K^{\perp}(\mathbf{x}, r, \alpha), q_f(\mathbf{x}, r, \alpha)),$$
$$(26)$$

where, as before, N_a is the number of tested annulus, $\mathcal{B}$ denotes the tail of the binomial distribution and q_f is now the probability of convergence towards the center of the annulus, for

a point that is not in the "main direction". It is given by

$$q_f(\mathbf{x}, r, \alpha) = \frac{\int_{\Omega \times [0,\pi]} \mathbb{1}_{E_{x,y}^{\perp}} \mathbb{1}_{E_\theta} f(x_S, y_S, \theta) \, dx_S dy_S d\theta}{\int_{\Omega \times [0,\pi]} \mathbb{1}_{E_{x,y}} f(x_S, y_S, \theta) \, dx_S dy_S d\theta},$$

where $E_{x,y}^{\perp}$ and E_θ are respectively the events

$$E_{x,y}^{\perp} = \{r \leq \|(x_S, y_S) - \mathbf{x}\| \leq \alpha r \text{ and } d((x_S, y_S), (M\mathbf{x})) \geq r\}$$

and

$$E_\theta = \{|\theta - \arctan \frac{y_S - y_\mathbf{x}}{x_S - x_\mathbf{x}}| \leq \arcsin \frac{r}{\|(x_S, y_S) - \mathbf{x}\|}\}.$$

On Figure 16, we have plotted the four different probabilities of convergence p_1, p_f, q_1 and q_f as a function of the distance between $\mathbf{x}$ and M. As we already pointed it out, p_1 and p_f are almost equal when the distance is large. Now, for q_1 and q_f, the situation is very different, since q_f (red curve) is much lower q_1 for large distances. As a consequence, we have a enhancement of the detections when using $\mathrm{NFA}_f^{\perp}$ compared to $\mathrm{NFA}_1^{\perp}$. This is illustrated by the last image of Figure 16, that represents $-\log \mathrm{NFA}_f^{\perp} + \log \mathrm{NFA}_1^{\perp}$.

6 Conclusion and discussion

In this paper, we have proposed and studied changes in the *a contrario* framework for the detection of points of convergences of linear structures in images. These convergences are of two types : global (like vanishing points or the normal convergence of linear structures in mammograms), or local (like stellate patterns in mammograms). For the global convergences we considered the supporting lines of previously detected line segments and for local convergences the line segments themselves were considered. The *a contrario* noise model, which is often chosen as the uniform independent distribution on linear structures, was changed into an anisotropic model to take into account a principal normal convergence of the linear structures. We proposed a parametric mixture model with a uniform term and a "Gaussian" term

modeling the principal convergence. The estimation of such a model on a sample of linear structures was addressed. The estimation of the point of global convergence was solved by detecting the most meaningful point of convergence against the uniform model whereas the remaining parameters were estimated by maximizing the likelihood. We tested the *a contrario* method on different images and in most of them, the principal convergence detected against the uniform model was no longer detected against the estimated parametric model. And sometimes new convergences were highlighted against the anisotropic model.

This procedure was devoted to improve the detection of stellate lesions in mammograms, by taking into account the normal convergence of the linear structures towards the nipple. Indeed against the uniform model this convergence is often detected although it is a normal convergence and shouldn't be detected. By using the simple uniform model the meaningful events are biased: the naive model should describe the normal distribution of linear structures in mammograms, and against this model only abnormal patterns should be detected.

There are several future directions of research for this work. A first point is that the hypothesis of independence of the linear structures could be modified. Indeed, the line segments detected in images are often organised in long chains (in natural images as well as in mammograms). Even though the independence hypothesis is convenient, it could be suppressed by adding some correlation between neighbouring line segments, in a way similar to what is done by Myaskouvskey, Gousseau and Lindenbaum in [13] for the detection of alignments. Under such a hypothesis, the detection performances can indeed be improved, but the computational complexity is much higher.

A second point is that the mixture model was obtained by adding a "convergence term" to the model. The convergence term was centered on the most meaningful point against the uniform noise model. Such a procedure could be carried on and we could add a second convergence term, and then a third one, etc. This would certainly make sense in particular in man-made environment images, where there are often two or three vanishing points. This step-by-step enrichment of the noise model would lead to a full (and not only up to the second order as presented here) hierarchical description of the convergences in the image. A third point, that is closely related to the previous one, is that for the estimation of the *a contrario* model we have restrained ourselves to a family of parametric Gaussian/uniform mixture models (notice also that the theoretical analysis of the proposed estimator could be studied). However one may ask what happens if there is no restriction on the form of the estimated model. The only condition should be that against the new model the principal convergence (against the uniform model) is no longer meaningful. Now, among all distributions that satisfy this condition which one is at the same time "the most random" (for instance in the sense of maximizing the entropy)? And also, is it possible to define a model against which there would be no detections? All these questions were recently rised, in a similar framework, in [4].

Finally, as we already emphazised it, we have proposed here in this paper a theoretical and methodological framework. It remains to turn it into an effective detection algorithm of stellate patterns in mammograms and to perform extensive tests on datasets of mammograms. But such a goal requires a more accurate definition of stellate lesions (besides the stellate organization of spicules, there is also often a dense mass in the center). As shown in the paper of Palma et. al [16], the efficient detection of potential lesions has to be performed using the aggregation of several partial detectors. We started in Section 5.4 explaining how to define features from NFA numbers, that could be used in a statistical learning methodology to validate the approach on a dataset of annotated mammograms. Ideally, this would be the aim of a companion paper, following this one.

Acknowledgements

A Appendix: Proof of Theorem 2

We give, in this appendix, the proof of Theorem 2. We first recall the result of Proposition 3: let K be a bounded closed convex set of the plane with support function s_K, then

$$\mu_g(D \cap K \neq \emptyset) = \frac{1}{\pi} \int_{\varphi=0}^{\pi} \left[\Phi \left(\frac{s_K(\varphi) - r_M \cos(\theta_M - \varphi)}{\sigma} \right) \right.$$
$$\left. - \Phi \left(\frac{-s_K(\varphi + \pi) - r_M \cos(\theta_M - \varphi)}{\sigma} \right) d\varphi \right],$$

where Φ is the cumulative distribution function of the standard normal distribution, that is

$$\forall t \in \mathbb{R}, \quad \Phi(t) := \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \, du.$$

Proof of case 1: σ goes to $+\infty$

The proof uses the development of the function Φ in the neighborhood of 0. For every $x \in \mathbb{R}$ we have

$$\Phi(x) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} x + O(x^2).$$

Therefore, taking first $x = \frac{s_K(\varphi) - r_M \cos(\theta_M - \varphi)}{\sigma}$, then $x = \frac{-s_K(\varphi + \pi) - r_M \cos(\theta_M - \varphi)}{\sigma}$ and computing the difference, it comes:

$$\Phi \left(\frac{s_K(\varphi) - r_M \cos(\theta_M - \varphi)}{\sigma} \right)$$
$$- \Phi \left(\frac{-s_K(\varphi + \pi) - r_M \cos(\theta_M - \varphi)}{\sigma} \right)$$
$$= \frac{1}{\sqrt{2\pi}} \frac{s_K(\varphi) + s_K(\varphi + \pi)}{\sigma} + O \left(\frac{1}{\sigma^2} \right),$$

where the term $O \left(\frac{1}{\sigma^2} \right)$ can be made independent of φ because the support function s_K is bounded. Finally, the measure of the set of lines meeting the convex set K under μ_g can be written:

$$\mu_g(D \cap K \neq \emptyset) = \frac{1}{\pi\sqrt{2\pi}\sigma} \int_0^{2\pi} s_K(\varphi) d\varphi + O \left(\frac{1}{\sigma^2} \right)$$

$$= \frac{1}{\pi\sqrt{2\pi}\sigma} \operatorname{Per} K + O \left(\frac{1}{\sigma^2} \right),$$

which is the announced result. □

Proof of case 2: σ goes to 0

We will need the following elementary result. Let $g_\sigma(x) := \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}$ denote the 1D Gaussian density of variance σ^2. Then if w is a piecewise continuous bounded function of the real variable, we have

$$\lim_{\sigma \to 0+} \int_{\mathbb{R}} w(x) g_\sigma(x) \, dx = \frac{w(0+) + w(0-)}{2}, \qquad (27)$$

where $w(0+) := \lim_{x \to 0, x > 0} w(x)$, resp. $w(0-) := \lim_{x \to 0, x < 0} w(x)$.

The proof of this result is simple. We first make the change of variable $y = x/\sigma$, such that

$$\int_{\mathbb{R}} w(x) g_\sigma(x) \, dx = \int_{\mathbb{R}} w(\sigma y) g_1(y) \, dy$$

$$= \int_{-\infty}^{0} w(\sigma y) g_1(y) \, dy + \int_{0}^{+\infty} w(\sigma y) g_1(y) \, dy.$$

And the result follows thanks to the dominated convergence theorem and to the fact that the integral of g_1 on $(0, +\infty)$ or on $(-\infty, 0)$ is equal to $\frac{1}{2}$.

Now, the measure of the set of lines D that meet the convex K with support function s_K can be written:

$$\mu_{g_\sigma}(D \cap K \neq \emptyset) = \frac{1}{\pi} \int_0^{\pi} \int_{\mathbb{R}} \mathbb{1}_{I_M(\varphi)}(\rho) g_\sigma(\rho) d\rho d\varphi,$$

where for $\varphi \in [0, \pi)$, we denote $I_M(\varphi)$ the interval

$$I_M(\varphi) :=$$
$$[-s_K(\varphi + \pi) - r_M \cos(\theta_M - \varphi), s_K(\varphi) - r_M \cos(\theta_M - \varphi)].$$

Thanks to the result (27) it comes that for all $\varphi \in [0, \pi)$

$$\lim_{\sigma \to 0+} \int_{\mathbb{R}} \mathbb{1}_{I_M(\varphi)}(\rho) g_\sigma(\rho) d\rho :=$$

$$\lim_{\sigma \to 0+} H_\sigma(\varphi) := H(\varphi),$$

where the function H is given by:

$$H(\varphi) = \begin{cases} 0 & \text{if } 0 \notin I_M(\varphi), \\ \frac{1}{2} & \text{if } 0 \in \partial I_M(\varphi), \\ 1 & \text{if } 0 \in I_M(\varphi) \setminus \partial I_M(\varphi). \end{cases}$$

Moreover since the function H_σ converges pointwise to H and since $\forall \varphi \in [0, \pi), \forall \sigma > 0, |H_\sigma(\varphi)| \leq 1$ it comes, thanks to the dominated convergence theorem, that

$$\lim_{\sigma \to 0+} \int_0^{\pi} H_\sigma(\varphi) d\varphi = \int_0^{\pi} H(\varphi) d\varphi.$$

Therefore, we finally get:

$$\lim_{\sigma \to 0^+} \mu_{g_\sigma}(D \cap K \neq \emptyset) = \frac{1}{\pi} \int_0^\pi H(\varphi)\, d\varphi$$

$$= \frac{1}{\pi} \int_0^\pi \mathbb{1}_{\{-s_K(\varphi+\pi) \leq r_M \cos(\theta_M - \varphi) \leq s_K(\varphi)\}}\, d\varphi.$$

where the last equality comes from the fact that, for a smooth bounded convex set K, the Lebesgue measure of the set of $\varphi \in [0, \pi)$ such that $-s_K(\varphi + \pi) = r_M \cos(\theta_M - \varphi)$ or $s_K(\varphi) = r_M \cos(\theta_M - \varphi)\}$ is 0. □

Proof of case 3: r_M goes to $+\infty$

Let K be a smooth bounded closed convex set with support function s_K. We then denote

$$F_K(r_M) := \mu_g(D \cap K \neq \emptyset) =$$

$$\frac{1}{\pi} \int_0^\pi \int_{-s_K(\varphi+\pi)}^{s_K(\varphi)} \frac{1}{\sigma\sqrt{2\pi}} e^{-(\rho - r_M \cos(\theta_M - \varphi))^2/2\sigma^2}\, d\rho d\varphi.$$

To find the equivalent of $F_K(r_M)$ as r_M goes to infinity, we first make the change of variable $\psi = \varphi - \theta_M + \frac{\pi}{2}$. Thanks to the π-periodicity of the function that is integrated in $F_K(r_M)$, we get

$$F_K(r_M) =$$

$$\frac{1}{\pi} \int_{-\pi/2}^{\pi/2} \int_{-s_K(\psi+\theta_M+\frac{\pi}{2})}^{s_K(\psi+\theta_M-\frac{\pi}{2})} \frac{1}{\sigma\sqrt{2\pi}} e^{-(\rho - r_M \sin\psi)^2/2\sigma^2}\, d\rho d\psi.$$

In a way similar to the Laplace's method, when r_M goes to infinity, the above integral will "concentrate" around the value $\psi = 0$. More precisely, let $\delta \in (0, \frac{\pi}{2})$, then define

$$I_\delta :=$$

$$\frac{1}{\pi} \int_{-\delta}^{\delta} \int_{-s_K(\psi+\theta_M+\frac{\pi}{2})}^{s_K(\psi+\theta_M-\frac{\pi}{2})} \frac{1}{\sigma\sqrt{2\pi}} e^{-(\rho - r_M \sin\psi)^2/2\sigma^2}\, d\rho d\psi$$

and

$$J_\delta :=$$

$$\frac{1}{\pi} \int_{|\psi|>\delta} \int_{-s_K(\psi+\theta_M+\frac{\pi}{2})}^{s_K(\psi+\theta_M-\frac{\pi}{2})} \frac{1}{\sigma\sqrt{2\pi}} e^{-(\rho - r_M \sin\psi)^2/2\sigma^2}\, d\rho d\psi.$$

In I_δ, we make the change of variable $y = \sin\psi$. We then have $dy = \sqrt{1 - y^2}d\psi$, and since $y \in [-\sin\delta, \sin\delta]$ we can bound I_δ by

$$I_\delta' \leq I_\delta \leq \frac{1}{\cos\delta} I_\delta'$$

with

$$I_\delta' :=$$

$$\frac{1}{\pi} \int_{-\sin\delta}^{\sin\delta} \int_{-s_K(\arcsin y + \theta_M + \frac{\pi}{2})}^{s_K(\arcsin y + \theta_M - \frac{\pi}{2})} \frac{1}{\sigma\sqrt{2\pi}} e^{-(\rho - r_M y)^2/2\sigma^2}\, d\rho dy.$$

Then making a new change of variable $x = r_M y$ in I_δ', we get

$$r_M I_\delta' =$$

$$\frac{1}{\pi} \int_{-r_M\sin\delta}^{r_M\sin\delta} \int_{-s_K(\arcsin(x/r_M) + \theta_M + \frac{\pi}{2})}^{s_K(\arcsin(x/r_M) + \theta_M - \frac{\pi}{2})} \frac{e^{-(\rho - x)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}\, d\rho dx.$$

Thanks to the dominated convergence theorem (we recall that the support function s_K is bounded), as r_M goes to infinity we have that

$$r_M I_\delta' \xrightarrow[r_M \to \infty]{} \frac{1}{\pi} \int_{-\infty}^{\infty} \int_{-s_K(\theta_M+\frac{\pi}{2})}^{s_K(\theta_M-\frac{\pi}{2})} \frac{1}{\sigma\sqrt{2\pi}} e^{-(\rho-x)^2/2\sigma^2}\, d\rho dx$$

$$= \frac{1}{\pi}(s_K(\theta_M + \frac{\pi}{2}) + s_K(\theta_M - \frac{\pi}{2})).$$

Now we control the integral J_δ. We develop the term $(\rho - r_M \sin\psi)^2$, and using again the fact that s_K is bounded, we get

$$0 \leq J_\delta \leq e^{-(r_M\sin\delta)^2/2\sigma^2} e^{(\|s_K\|_\infty r_M)/\sigma^2},$$

which shows that $r_M J_\delta$ goes to 0 as r_M goes to infinity.

Finally, putting all this together, we have: for any $\varepsilon > 0$, fix δ in such a way that $1/\cos\delta \leq (1+\varepsilon)$. Then for this fixed value of δ, there exist $C > 0$ such that for all $r_M > C$ we have

$$(1 - \varepsilon)\frac{1}{\pi}(s_K(\theta_M + \frac{\pi}{2}) + s_K(\theta_M - \frac{\pi}{2})) \leq r_M I_\delta$$

$$\leq (1+\varepsilon)^2 \frac{1}{\pi}(s_K(\theta_M + \frac{\pi}{2}) + s_K(\theta_M - \frac{\pi}{2})),$$

and

$$0 \leq r_M J_\delta \leq \varepsilon.$$

Summing I_δ and J_δ, we have the announced result that is

$$r_M F_K(r_M) \xrightarrow[r_M \to \infty]{} \frac{1}{\pi}(s_K(\theta_M + \frac{\pi}{2}) + s_K(\theta_M - \frac{\pi}{2})).$$

□

References

1. Almansa, A., Desolneux, A., Vamech, S.: Vanishing point detection without any a priori information. IEEE Transactions on Pattern Analysis and Machine Intelligence **25**(4), 502–507 (2003)
2. Bornefalk, J.: Use of quadrature filters for detection of stellate lesions in mammograms. In: Scandinavian Conference on Image Analysis (SCIA), pp. 649–658 (2005)
3. Cao, F.: Application of the Gestalt principles to the detection of good continuations and corners in image level lines. Computing and Visualization in Science **7**(1), 3–13 (2004)
4. Desolneux, A.: When the a contrario approach becomes generative. International Journal of Computer Vision (2015). To appear
5. Desolneux, A., Moisan, L., Morel, J.M.: From Gestalt theory to image analysis: a probabilistic approach, *Interdisciplinary Applied Mathematics*, vol. 34. Springer (2008)
6. Grompone von Gioi, R., Jakubowicz, J., Morel, J.M., Randall, G.: LSD: A fast line segment detector with a false detection control. IEEE Transactions on Pattern Analysis and Machine Intelligence **32**(4), 722–732 (2010)
7. Grimmett, G.R., Stirzaker, D.R.: Probability and random processes, third edn. Oxford University Press, New York (2001)
8. Grosjean, B., Moisan, L.: A-contrario detectability of spots in textured backgrounds. Journal of Mathematical Imaging and Vision **33**(3), 313–337 (2009)
9. Grosjean, B., Muller, S., Souchay, H.: Lesion detection using an a-contrario detector in simulated digital mammograms. In: Proc. of SPIE Vol, vol. 6146 (2006)
10. Karssemeijer, N., Te Brake, G.M.: Detection of stellate distortions in mammograms. IEEE Transactions on Medical Imaging **15**(5), 611–619 (1996)
11. Kegelmeyer, W., Pruneda, J., Bourland, P., Hillis, A., Riggs, M., Nipper, M.: Computer-aided mammographic screening for spiculated lesions. Radiology **191**, 331–337 (1994)
12. Lowe, D.: Perceptual Organization and Visual Recognition. Kluwer Academic Publishers, Amsterdam (1985)
13. Myaskouvskey, A., Gousseau, Y., Lindenbaum, M.: Beyond independence: An extension of the a contrario decision procedure. International Journal of Computer Vision **101**(1), 22–44 (2013)
14. Palma, G.: Détection automatique des opacités en tomosynthèse numérique du sein. Ph.D. thesis, Telecom Paris Tech (2010)
15. Palma, G., Bloch, I., Muller, S.: Spiculated lesions and architectural distortions detection in digital breast tomosynthesis datasets. In: International Workshop on Digital Mammography (IWDM), Girona, Spain (2010)
16. Palma, G., Bloch, I., Muller, S.: Detection of masses and architectural distortions in digital breast tomosynthesis images using fuzzy and a contrario approaches. Pattern Recognition **47**(7), 2467–2480 (2014)
17. Santalo, L.A.: Integral Geometry and Geometric Probability, second edn. Cambridge University Press (2004)
18. Witkin, A., Tenenbaum, J.: On the role of structure in vision. In: Human and Machine Vision, pp. 481–543. A. Rosenfeld ed., Academic Press, New York (1983)
19. Xia, G., Delon, J., Gousseau, Y.: Accurate junction detection and characterization in natural images. International Journal of Computer Vision **106**(1), 31–56 (2014)
20. Zhu, S.: Embedding Gestalt Laws in Markov Random Fields. IEEE Transactions on Pattern Analysis and Machine Intelligence **21**(11), 1170–1187 (1999)
21. Zwiggelaar, R., Parr, T.C., Schumm, J., Hutt, I.W., Taylor, C.J., Astley, S.M., Boggis, C.R.M.: Model-based detection of spiculated lesions in mammograms. Medical Image Analysis **3**(1), 39–62 (1999)