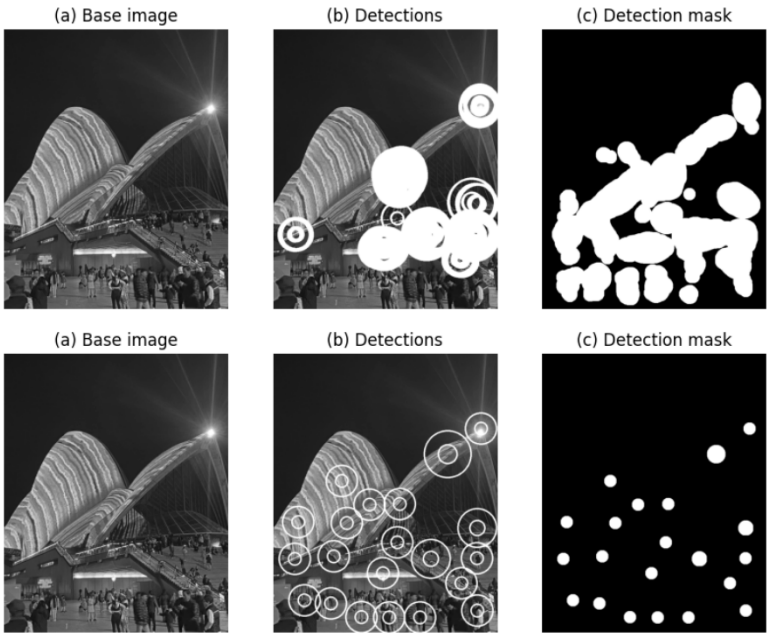
MATHÉMATIQUES
VISION
APPRENTISSAGE

Master MVA Detection Theory

Project

Detection of convergence areas in Digital Breast Tomosynthesis using a contrario modeling

Student :
Ugo Insalaco



Code available at: https://github.com/Ugo-Insalaco/MVA-Detection

March 23, 2024

1 Introduction

Tomography is a visualization technique for 3D volumes. By sequentially showing images of the same object at different depth, we humans are able to mentally reconstruct a representation of the latent object. In medical imaging, such images are acquired using X rays for a fixed angle and applied for instance for studying the presence of breast cancer (and is called Digital Breast Tomosynthesis).

Cancer lesions may take different shapes but one of them is visually characterized on tomography images by white masses with converging lines towards this mass (fig.1). The authors from [3, 2] introduce an *a contrario* modelization to detect such recognizable shapes. For reference, breast cancer touched around 2.3 million women in 2022 and caused 670.000 deaths in the world according to World Health Organization, and accounts for about 30% of all new female cancers.

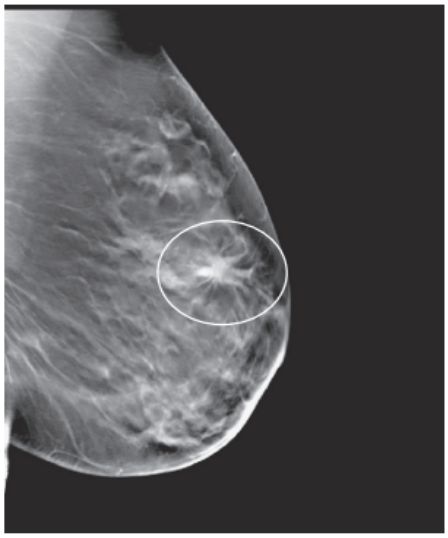
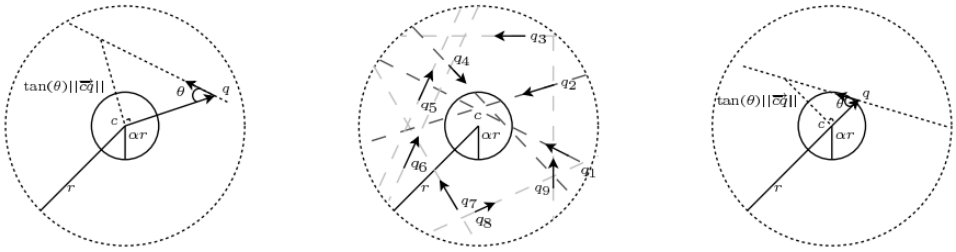


Figure 1: Example of a Digital Breast Tomosynthesis

1.1 *A contrario* model

The a contrario approach provides tools for detecting specific patterns in images like alignments of points, or crossing lines. According to the Gestalt theory, the elemental patterns (called gestalts) are directly linked to our perception of the image and strongly depend on their environment in the image, with phenomenons such as masking or conflicts. In our case the targeted gestalt corresponds the convergence of lines towards a mass. This thus separates into two parts: the mass and the converging lines, each line being itself composed of several pixels oriented in the same direction.

The authors build upon this concept for the modelization. For all possible position c of a lesion on the image, they first restrict the search of convergence inside a circle or radius r. In the middle of this circle should be lying the mass, which they separate with another circle of radius $\alpha * r$ ($\alpha \in [0; 1]$). Finally in the crown surrounded by these circles, they count the number of pixels q "directed towards" the mass (see fig.2). Using a threshold on the number of well-oriented pixels, they are able to detect the presence of a lesion: this is the *a contrario* approach.



(a) (b) (c)

Figure 2: Modelization setup for the *a contrario* approach

1.2 Constructing the NFA

The *a contrario* model in each elemental crown is defined such that each pixel q has a random direction characterized by an angle θ with the radius passing by q, sampled uniformly in $[-\pi/2; \pi/2]$. In practice, the direction will be computed as orthogonal to the gradient direction at the pixel location. This model is consistent with our intuition, as in a random image with no special structure, the gradients would also be distributed uniformly in any direction.

Using this simple hypothesis, we can define the random variable K_{crq} that is equal to 1 if the direction at q is crossing the inner circle αr and 0 otherwise. With simple computations, we have:

$$\begin{aligned}
\mathbb{P}(K_{cqr} = 1) &= \mathbb{P}\left(\|\overrightarrow{cq}\| |\tan\theta| \leq \alpha r\right) \\
&= \mathbb{P}\left(\frac{-\alpha r}{\|\overrightarrow{cq}\|} \leq \tan(\theta) \leq \frac{\alpha r}{\|\overrightarrow{cq}\|}\right) \\
&= \int_{-\arctan\left(\frac{\alpha r}{\|\overrightarrow{cq}\|}\right)}^{\arctan\left(\frac{\alpha r}{\|\overrightarrow{cq}\|}\right)} \frac{1}{\pi} d\theta \\
&= \frac{2}{\pi} \arctan\left(\frac{\alpha r}{\|\overrightarrow{cq}\|}\right)
\end{aligned}$$

which is the result found by the authors. Furthermore, for each convergence area, they define the unit tests based on

$$Z_{cr} = \sum_{q \in \Omega, \alpha r \leq \|\overrightarrow{cq}\| \leq r} K_{crq} \tag{1}$$

which sums over all pixels in image Ω lying inside the detection crown. The NFA is then

$$NFA((c, r), z_{cr}) = M\mathbb{P}(Z_{cr} \geq z_{cr}) \tag{2}$$

where M is the total number of tests, for instance in an image of size $H \times W$ and considering R different radius, $M = HWR$. In practice, we fix the meaningfulness threshold ϵ and compute $\lambda_r = \min\{\lambda | M\mathbb{P}(Z_{cr} \geq \lambda) \leq \epsilon\}$ (which does not depend on the center position) and consider that a sample z_{cr} is detected if $z_{cr} \geq \lambda$. Intuitively, λ_r is the maximum number of pixels q authorized in the detection areas of radius r in the *a contrario* model

for which we consider that we do not have a detection, up to ϵ detection over the entire image.

This definition is directly linked to the general definition of an NFA:

$$\forall \epsilon > 0, \mathbb{E}(\#|i, f(i, X_i) \leq \epsilon|) \leq \epsilon \tag{3}$$

with the tests i iterating over $[H] \times [W] \times [R]$ and $f((c, r), z_{cr}) = M\mathbb{P}(Z_{cr} \geq z_{cr})$

2 Implementation

In this section, I explain my approach in the implementation of the author's algorithm for detecting convergence areas, and the application of the masking principle for detection reduction.

2.1 *A contrario* probabilities

The first step is to compute the probability of having exactly k well-oriented points in a detection area. Formally, we seek for $p_k = \mathbb{P}(Z_{cr} = k)$. As Z_{cr} is defined as the sum of Bernouilli variables with different parameters, we do not have a closed form expression for p_k. However, as explained by the authors, we can use the generating function and the independance of each K_{crq} with the expression:

$$G_{cr}(x) = \prod_{q \in \Omega, \alpha r \leq \|\overrightarrow{cq}\| \leq r} \mathbb{P}(K_{crq} = 0) + \mathbb{P}(K_{crq} = 1)x = \sum_{k=1}^{+\infty} p_k x^k \tag{4}$$

Leading to each probability p_k. The equation is still complicated to apply as it is the product of a very large number of polynomials, involving very low numbers (the p_k) and very large ones if we compute it naively with the combination coefficients. The authors didn't provide their way at solving it, hence I used simple algorithm directly derived from this identity of the polynomials product:

$$P \times \sum_{i=0}^{n} a_i X_i = \sum_{i=1}^{n} a_i X^i P \tag{5}$$

In order to make the computations even faster, I took advantage of the invariance of $\mathbb{P}(K_{crq})$ with $\|\overrightarrow{cq}\|$ to not compute the product over all pixels in the crown, but instead binning the values of $\|\overrightarrow{cq}\|$ together in $B_r = B$ bins evenly spaced between αr and r.

$$G_{cr}(x) = \prod_{i=1}^{B} (\mathbb{P}(K_{crq} = 1) + \mathbb{P}(K_{crq} = 1))^{n_{cb_i}} \tag{6}$$

$$\text{With: } n_{cb} = \#\{q \in \Omega, b - 1 < \|\overrightarrow{cq}\| \leq b\} \tag{7}$$

Finally, in order to compute each power of the polynomials of degree 1, I implemented an algorithm based on

$$\left(\sum_{i=0}^{n} c_i X^i\right)(a_{n+1} + b_{n+1}X) = c_0 a_{n+1} + \sum_{i=1}^{n} (c_i a_{n+1} + c_{i-1} b_{n+1}) X^i + c_n b_{n+1} X^{n+1} \tag{8}$$

2.2 Binning effect

The binning in the first step reduces the number of polynomials to multiply drastically, at the cost of some approximation on the K_{crq} probabilities. However based on the results fig.3, It didn't seem to have a large effect on the final value of λ. In fact for $B >> (r - \alpha r)$ there should not be any approximation as the values of $\|\overrightarrow{cq}\|$ are already quantified.

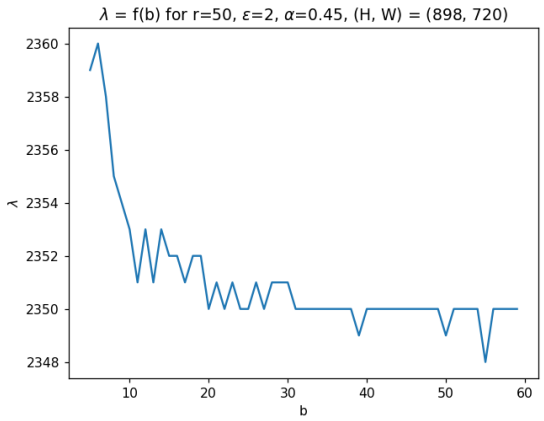


Figure 3: Evolution of λ as a function of the binning parameter B. The values of λ are constant for B large enough

2.3 Computing the detection

For my experiments on the implementation I used the image on fig.4. This image is relevant for the problem as it shows a strong converging pattern at the top of the Sydney Opera due to the light. The goal is to be able to automatically detect this gestalt. We can also observe the targeted pattern at the different intersections of the Opera spherical parts, and with the projected colored lines. I tried during my experiments to use DBT images, but, as I am not an expert in this field, I was not sure of which patterns we were willing to extract, and the converging patterns were also not as visible as the ones from my test image.

In order to compute the detection on the image, I implemented the authors version of a gradient computation. Since we made the hypothesis that the different angles θ were independent, the computation of the gradient had to reflect that idea. This is why the authors proposed a gradient that is based on 4×4 squares that do not overlap (fig.5). From this gradient, we define in each pixel an associated direction as the orthogonal line to the gradient.

In each possible center and radius of the image, using the direction of each pixel in the associated crown area, I computed the Z_{cr} value of the region and compared it with the λ values. To be perfectly sure that I computed the right K_{crq} values, I show on fig.6 and example of these values. Visually it seams correct with our expectations as the principle lines converging towards the center are denser in red dots that the rest of the area.

I fixed $\epsilon = 2$ in order to have False Alarms and to be sure that my implementation was valid. I manually selected a radius $r \in \{50; 75\}$ and $\alpha = 0.4$ by visually adjusting the crown marker on the image to have the white center area surrounded perfectly by the inner circle and a r not too large. On the first row of fig.7, I show the top 5000 most

Figure 4: Image on which to detect the convergence areas. The tip of the Sydney Opera has a strong converging pattern created by the light artefacts

significant detection on the image. Among them, We are pleased to observe our targeted point at the top of the opera.

2.4 Detection reduction with the masking principle

In the article, the authors simply aggregated the results of the *a contrario* method in a global mask, which lead to large patches of detection, like the third column of fig.7. Using the masking principle, I implemented a reduction of the number of detection on the image.

Definition 1: Given two ϵ-meaningful detection tests A, B, we say that A is masked by B if by removing the detection features of B from A, A is not anymore ϵ-meaningful.

In our specific case, we can compute the exclusion of a crown $A = (c, r)$ from a crown B and and compute the $Z_{A\setminus B} = \sum_{q \in A\setminus B} K_{crq}$ value using the remaining points. If the detection area is still ϵ-meaningful, then it is not masked. The masking principle was applied in a simple algorithm to only keep the maximum ϵ-meaningful detection areas (alg.1). This algorithm yielded the detection on the second row of fig.7. Visually we observe that the number of detection is well reduced, only keeping one detection per specific place (and keeping our beloved detection at the top of the opera). We can also see that most detection were made with the smaller radius $r = 50$, showing that the choice of $r = 75$ was not significant in the scale of our image, or that the smaller r detects too small structures.

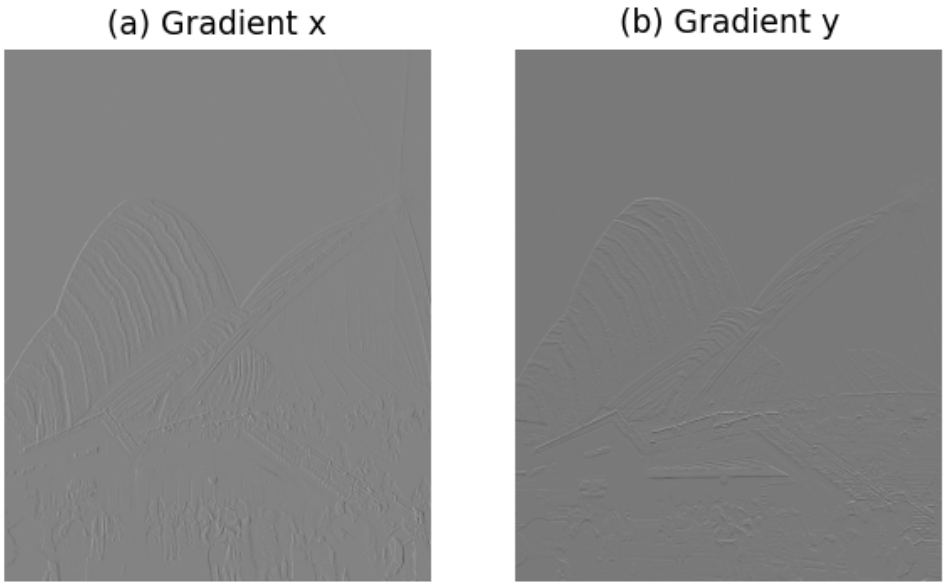


Figure 5: Gradients on my test image. The contrast is not perfect but it still well reflects the expected result from a gradient. For instance focus on the strong line in the bottom half of the gradient y map that disappears in the gradient x map

Algorithm 1 Masking principle algorithm

Require: Z_{cr} values of detected (c, r) tests
  Sort Z_{cr} by Descending order
  Keep $\leftarrow$ empty list
  for (c, r) detected tests **do**
    for do$(c', r') \in$ Keep
      if (c, r) is **Not** masked by (c', r') **then**
        Add (c, r) to Keep
      end if
    end for
  end for
  return Keep

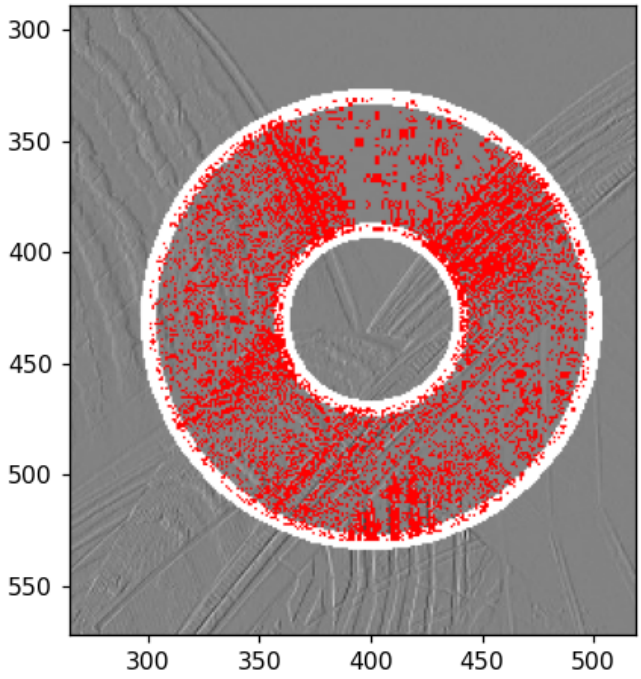


Figure 6: Example of computation of K_{cqr}. The red pixels are the ones well oriented towards the convergence area.

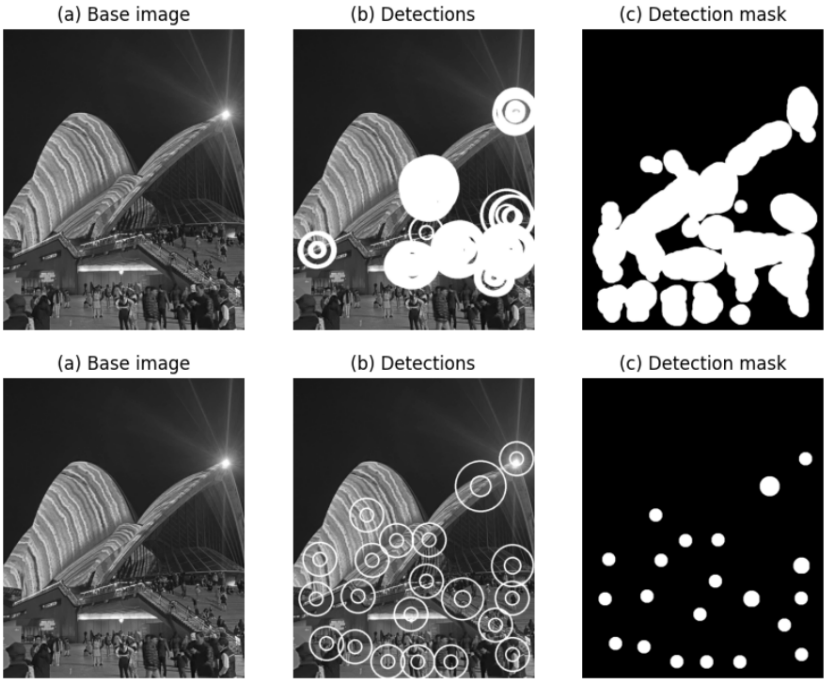


Figure 7: Detection using *a contrario* approach on the Sydney opera (top row) and filtered detection using the masking principle (bottom row)

3 More examples

To better experiment on the detection technique and the implementation, I show two extreme examples of images for this detection.

3.1 Random noise

Random noise images allow us to compare the defined maximum expected Number of False Alarms ϵ with our empirical results. Fixing a random 200×200 image with $r_{\min} = 10$ $r_{\max} = 20$, interpolated with 5 circles, $\alpha = 0.4$, and $\epsilon = 50$, I obtained the results in fig.8. We observe here that we indeed obtain a few dozen of detection on the random image, due to the selected $\epsilon = 50$.

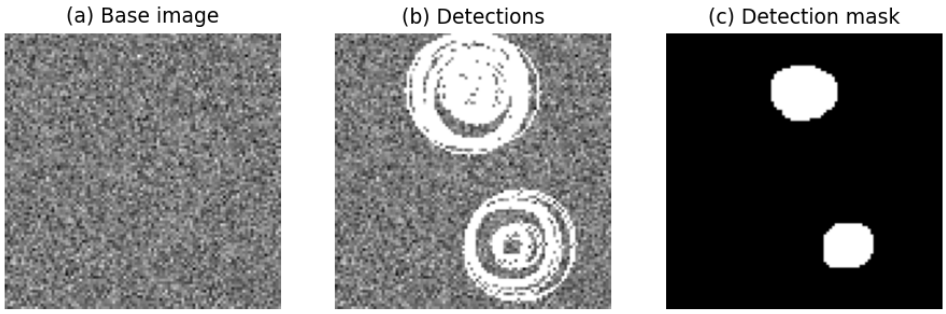


Figure 8: Detection on the noise image

3.2 Pupil

The epitome of a converging pattern could be seen in the human eye. In fig.9, I show the computed K_{crq} values for the image, which clearly shows strait lines emerging from the center. By reversing the previously implemented method, I obtained a meaningfulness for the selected detection area of 10^{-7}.

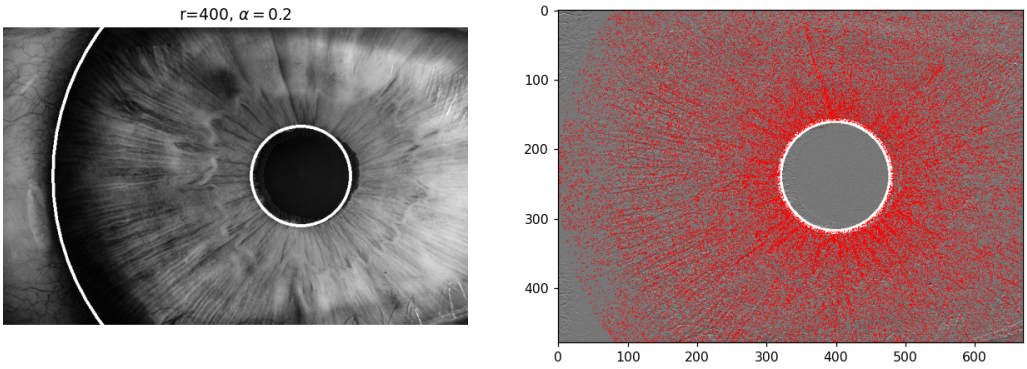


Figure 9: Eye example (left). K_{crq} values on a human eye (right)

4 Conclusion

This project allowed me to implement a fully working detection pipeline of convergence patterns in an image using the *a contrario* methodology. However it still has several limits and possible improvements.

Faster Polynomials: My implementation of the polynomial multiplication and power was still naive and based on a very simple formula. At the moment, computation of λ_r values for large r can take up to several minutes, which could be improved using smarter algorithms, for instance using the Fast Fourier Transform [1] or using Karatsuba algorithm

Binning effect: Binning the radius values allowed to speed up computations of the Z_{cr} values, however it introduces approximations. An interesting improvement would be to mathematically model this effect and choose an optimal B parameter. Furthermore, in my computation of the probabilities p_k, I obtained very smooth figures which may be analytically tractable and would obviate the detour through polynomials computations.

Edge effects: The λ_r values, in the method proposed by the authors does not depend on the center position c. This is an approximation for edge areas, that have less pixels, hence less possibility to reach the threshold λ_r for meaningfulness. In practice this would lead to less detection on the border of the image and in the case of Breast tomosynthesis, at the most shallow parts of the breast. This might not actually be a problem if cancerous cells were more likely to be deeper in the tissues, or using views from several angles.

Gestalt masking: In my experiment on the Sydney opera, I observed several detection at unexpected places, for instance surrounding pedestrians or false intersections of objects on the image due to 2D projection. In the article, this is partially solved by forcing a more uniform distribution on the converging angles θ using a threshold on their entropy. However this could be also linked to the detection of masked structures in the Gestalt theory. In our perception of shapes, we only interpret the strongest structures and forget about all possible intersections of objects in an image. In the case of this *a contrario* approach, it seems like we also managed to detect "weak" structures that were not perceptible at first sight. Once again, this still might not be an issue for DBT images

as all structures should be detected and cancerous lesions exactly correspond to the detection of converging areas, but in definitive, it raises the question of the application of this method to real images.

References

[1] Robert T. Moenck. "Practical fast polynomial multiplication". In: *Symposium on Symbolic and Algebraic Manipulation*. 1976. URL: https://api.semanticscholar.org/CorpusID:18500147.

[2] Giovanni Palma, Isabelle Bloch, and Serge L. Muller. "Detection of masses and architectural distortions in digital breast tomosynthesis images using fuzzy and a contrario approaches". In: *Pattern Recognit.* 47 (2014), pp. 2467–2480. URL: https://api.semanticscholar.org/CorpusID:11165769.

[3] Giovanni Palma et al. "Detection of convergence areas in Digital Breast Tomosynthesis using a contrario modeling". In: *Proceedings of SPIE - The International Society for Optical Engineering* 7260 (Feb. 2009). DOI: 10.1117/12.807547.